The Complete Nucleotide Sequence and Biotype Variability of Papaya leaf distortion mosaic virus

Tetsuo Maoka and Tatsuji Hataya

First author: Plant Virology Laboratory, National Agricultural Research Center for Hokkaido Region, Sapporo, Japan; and second author: Laboratory of Pathogen-Plant Interactions, Plant Breeding Science, Graduate School of Agriculture, Hokkaido University, Sapporo, Japan.

Accepted for publication 2 September 2004.

ABSTRACT

Maoka, T., and Hataya, T. 2005. The complete nucleotide sequence and biotype variability of Papaya leaf distortion mosaic virus. *Phytopathology* 95:128-135.

The complete nucleotide sequence of the genome of Papaya leaf distortion mosaic virus (PLDMV) was determined. The viral RNA genome of strain LDM (leaf distortion mosaic) comprised 10,153 nucleotides, excluding the poly(A) tail, and contained one long open reading frame encoding a polyprotein of 3,269 amino acids (molecular weight 373,347). The polyprotein contained nine putative proteolytic cleavage sites and some motifs conserved in other potyviral polyproteins with 44 to 50% identities, indicating that PLDMV is a distinct species in the genus *Potyvirus*. Like the W biotype of *Papaya ringspot virus* (PRSV), the non-papaya-infecting biotype of PLDMV (PLDMV-C) was found in

plants of the family Cucurbitaceae. The coat protein (CP) sequence of PLDMV-C in naturally infected-*Trichosanthes bracteata* was compared with those of three strains of the P biotype (PLDMV-P), LDM and two additional strains M (mosaic) and YM (yellow mosaic), which are biologically different from each other. The CP sequences of three strains of PLDMV-P share high identities of 95 to 97%, while they share lower identities of 88 to 89% with that of PLDMV-C. Significant changes in hydrophobicity and a deletion of two amino acids at the N-terminal region of the CP of PLDMV-C were observed. The finding of two biotypes of PLDMV implies the possibility that the papaya-infecting biotype evolved from the cucurbitaceae-infecting potyvirus, as has been previously suggested for PRSV. In addition, a similar evolutionary event acquiring infectivity to papaya may arise frequently in viruses in the family Cucurbitaceae.

Papaya (*Carica papaya*) is one of the most important fruit trees in tropical and subtropical areas. *Papaya ringspot virus* biotype P (PRSV-P) is regarded as the most destructive factor in papaya production in many countries (26). In Japan, however, Papaya leaf distortion mosaic virus (PLDMV) has been reported as the most destructive virus (19,23). PLDMV was first reported on Okinawa Island in 1954. Based on a ringspot symptom observed on fruit and the presence of flexuous filamentous particles, the virus was initially considered to be PRSV (19,37). However, because no serological relationship was found between the virus and PRSV, it was later named PLDMV (19).

Both viruses belong to the genus *Potyvirus*. They are similar to each other in several properties such as their symptoms on papaya plants, their host range, transmission by aphids, and physical properties (23). The PLDMV RNA genome has been partially characterized at the 3'-terminal region including the coat protein (CP) coding region (22), whereas PRSV RNA has been completely characterized (36). The CP sequence of PLDMV shares only 55 to 59% amino acid identity with that of PRSV, supporting the finding that they are serologically unrelated (22). To understand the molecular relationship between PLDMV and PRSV in detail, we have fully sequenced the PLDMV genome. We present here molecular evidence that PLDMV is a new species in the genus *Potyvirus*.

Between 1990 and 1995, 335 leaf samples were collected from PLDMV-infected papaya plants in Japan. Three strains that produced different symptoms in papaya, leaf distortion mosaic (LDM), mosaic (M), and yellow mosaic (YM), were found in 71, 24, and 4% of samples, respectively (*unpublished data*). In this study, the biological, serological, and molecular properties of each strain were determined.

PRSV is classified into two biotypes; they cannot be distinguished in serological tests, but they differ in their ability to infect papaya. PRSV-W isolated from the family Cucurbitaceae cannot infect papaya, unlike PRSV-P (26). Likewise, a virus that reacted with a PLDMV antiserum was found for the first time in a cucurbitaceous weed (PLDMV-C(T)). The C biotype of PLDMV (PLDMV-C) is unable to infect papaya and indistinguishable from PLDMV-P in serological tests. The CP sequence of PLDMV-C was compared with those of the three strains of papaya-infecting biotype P. In this study, we discuss the variability of PLDMV and the evolutionary relationship of biotypes P and C.

MATERIALS AND METHODS

Virus sources. The three strains of P biotype (LDM, M, YM) and the *Trichosanthes* (T) strain of the C biotype were used. PLDMV-P(LDM) isolate J56P was obtained from a distorted leaf of a papaya plant in Ishigaki Island, Okinawa Prefecture of Japan in 1990. This isolate was called isolate P56 in a previous report (22). PLDMV-P(M) isolate J69P was obtained from a mosaic leaf of a papaya plant cultivated on Okinawa Island in 1991. PLDMV-P(YM) isolate J179P was obtained from a papaya leaf, which had only the yellow mosaic symptom without the distortion symptom, on Ishigaki Island in 1992. The three strains of PLDMV-P were propagated in *Cucumis metuliferus* line Acc.2459. PLDMV-C(T) isolate J199C was collected from a mosaic leaf of *Trichosanthes*

Corresponding author: T. Maoka; E-mail address: maokat@affrc.go.jp

The nucleotide sequences reported in this paper have been submitted to the DDBJ/EMBL/GenBank nucleotide sequence databases under accession nos. AB088221, AB092814, AB092815, and AB092816.

DOI: 10.1094/PHYTO-95-0128

© 2005 The American Phytopathological Society

bracteata grown on Ishigaki Island in 1995 and then propagated in cucumber (*Cucumis sativus*) cv. Sagami-hanshiro. The viruses were purified as described previously (22) with the exception of PLDMV-P(M). Because of its low quantity in the propagation host, only a few particles of PLDMV-P(M) were purified. After purification, the preparations were used immediately for immunization. The remainder of the preparations were stored at -80°C until use for cDNA cloning.

Host range studies and aphid transmission. Twenty species of herbaceous plants belonging to seven families were grown in a greenhouse at 25 to 30°C and used for studies of host range of each PLDMV strain. The plants were mechanically inoculated with a crude sap containing each PLDMV strain as described previously (23). Cucurbitaceous plants were inoculated at the cotyledon stage, leguminous plants were inoculated at the primary leaf stage, and all other plants were inoculated at the three- to four-leaf stage. At least two plants were inoculated with each strain of PLDMV. Symptoms were recorded visually, and asymptomatic test plants were examined by double-antibody sandwich enzyme-linked immunosorbent assay (ELISA) (10) for the presence of virus 1 month after inoculation.

The aphid transmissibility of each PLDMV strain was tested as described previously (23). Five viruliferous *Myzus persicae* were placed on each of five papaya or *Cucumis metuliferus* Acc.2459 plants. The plants were assessed for infection by symptom appearance 1 month after inoculation.

Serological tests. The antiserum against PLDMV-P(LDM) was made against the purified virus from *Cucumis metuliferus* as described previously (24). The antisera against PLDMV-P(YM) and PLDMV-C(T) were made as follows. Fresh purified virus was mixed 1:1 with Freund's complete adjuvant and injected into the leg muscles of a rabbit two times at 2- (for YM strain) or 5- (for T strain) week intervals. An intravenous injection was given 1 (YM) or 3 (T) weeks after the final intramuscular injection. Blood was collected 10 (YM) or 12 (T) days after the booster injection. Serological relationships among the four strains were tested by western blotting and sodium dodecyl sulfate (SDS)-immunodiffusion as described previously (24).

Cloning and sequencing. Viral RNA was extracted from purified particles of PLDMV-P(LDM) using SepaGene (Sanko Junyaku Co., Ltd., Tokyo) according to the manufacturer's instructions. The 3'-terminal genomic sequence consisting of 1,404 nucleotides was determined from the cDNA clone PL50 and reported previously (22). The oligonucleotide primer Nib-1 (Table 1) was synthesized based on the 5'-terminal sequence of the PL50

clone (Fig. 1). The cDNA was synthesized using the Nib-1 primer and Marathon cDNA Amplification Kit (Clontech Laboratories, Palo Alto, CA) according to the manufacturer's instructions, and cDNA clone Nib99 was obtained. Likewise, cDNA clones Nia41 and CI64 were obtained using the oligonucleotide primers Nia-1 and CI-1 (Table 1, Fig. 1), respectively. The oligonucleotide primer 6K1-2 (Table 1) was synthesized based on the 5'-terminal sequence of cDNA clone CI64 and used for the cDNA synthesis as described previously (Fig. 1). The cDNA synthesis was repeated, and the resultant cDNA clone (6K146) was partially sequenced to design the reverse primer HC-1 (Table 1). The cDNA clone HC23 was obtained using the HC-1 primer and was partially sequenced to design the forward primer P1-2 (Table 1). The large cDNA clone P16K111 was obtained by reverse transcription-polymerase chain reaction (RT-PCR) using a pair of primers P1-2 and 6K1-2 (Fig. 1). The cDNA synthesis was repeated as described, and the cDNA clone P140 was obtained using the oligonucleotide primer P1-4 (Table 1). A fragment containing the 5' end of the genome was obtained by 5' rapid amplification of cDNA ends (5' RACE). Oligonucleotide adapter primer A (9) (Table 1) was ligated to the 3' ends of cDNA, which was synthesized with the primer P1-4, using T4 RNA ligase as described (14). The genome-specific product was obtained by nested PCR using a pair of outer primers B (9) and P1-3 and a pair of inner primers C (9) and P1-7 (Table 1). The amplified cDNA fragments were purified on Size Sep 400 Spun Columns (Amersham Life Science, Cleveland, OH), cloned in pT7Blue T-Vector (Novagen, Madison, WI), and sequenced. Cycle sequencing in both directions was performed using an automated DNA sequencer (Li-Cor Model 4000L; Li-Cor, Lincoln, NE) with Thermo Sequenase fluorescence-labeled primer cycle sequencing kit (Amersham Life Science) according to the manufacturer's instructions.

To amplify CP-coding regions of the PLDMV-P(YM) and PLDMV-C(T) genomes, a primer set, PL3 and PL4 (Table 1), was designed. PL6 (Table 1) was used instead of PL4 for PLDMV-P(M). The RT-PCR was carried out as follows: 1 cycle at 42°C for 15 min and at 99°C for 5 min followed by 30 cycles at 94°C for 1 min, at 50°C (PL4) or 65°C (PL6) for 2 min, and at 72°C for 3 min, and finally 1 cycle at 72°C for 7 min. Amplified cDNA fragments were cloned as described previously. The RT-PCR clones were sequenced using a DNA sequencer (Amersham Life Science, Model ALF Express) with an AutoRead Sequencing Kit (Amersham Life Science) according to the manufacturer's instructions.

Primer extension analysis. To determine the exact 5' end of the viral genome of PLDMV-P(LDM), primer extension analysis

TABLE 1. Synthetic oligonucleotide primers

Primer name ^a	Primer sequence (5'-3')	Genome location
Primers for complete sequencing		
Nib-1(-)	CGCAAATCGCTTCTAATCTGTGTCTCTGGT	8,780-8,808
Nia-1(-)	CAACCCATAAGTTGTGGTGACCGA	6,851-6,876
CI-1(-)	CCCTTATTGATTACATCTCGTGCAGACC	4,864-4,891
6K1-2(-)	CAGTCAAGATGTCTTGAATGTCGTCCAAG	4,151-4,179
HC-1(-)	CAACATCGATTAAGCTTGGCCATTGACC	2,703-2,730
P1-2(+)	GCCTGAGCGCATACACAAAAACCTCC	304-329
P1-4(-)	TCTAAGGCCACAGTGGAAACAACCTGC	370-395
API	CCATCCTAATACGACTCACTATAGGGC	Adaptor
Primers for 5' rapid amplification of cDNA ends		
P1-3(-)	ACAGTGGAACAACCTGCCAGATGCAG	361-386
P1-7(-)	TCCAGGAACAAGACGAACAC	196-215
A	GTAGGAATTCGGGTGTAGGGAGGTCGACATTGCC	Adaptor
B	GGCAATGTCGACCTCCCTACAAC	Adaptor
C	CTCCCTACAACCCGAATTCCTAC	Adaptor
Primers for coat protein sequencing		
PL3(+) ^b	gttaagcttTGGGTTCTKGARCAAGCWCCA	8,871-8,891
PL4(-) ^b	attggatccGAGGGCTACCTCACTG	10,045-10,061
PL6(-) ^b	gcggatccTATGCTTAGGTAATA	9,946-9,961

^a Signs (+ and -) after the primer indicate whether the primer sequence is homologous or complementary to the genome, respectively.

^b Lowercase letters have no relationship to the genome sequence.

was performed as described by Hataya et al. (16) using P1-7 (Table 1) primer complementary to the viral RNA. Reference sequencing ladders were produced by sequencing a 5' RACE clone using a primer set anchor primer C (9) and P1-7 (Fig. 1).

Sequence analysis. Nucleotide and amino acid sequence data were analyzed using the DNASIS software (version 7.0 for Windows) from Hitachi Software Engineering Co. (Tokyo, Japan). For the multiple alignment, the program CLUSTAL X version 1.81 (32) was used with the default parameters. Relationship dendrograms were calculated with 1,000 bootstrap replicates using the neighbor-joining option in CLUSTAL X and drawn using the TreeView program (25). Table 2 lists the viruses discussed in this report and their sequence accession numbers in the DDBJ/EMBL/GenBank databases. When the CP amino acid sequences were analyzed, *Barley yellow mosaic virus* in the genus *Bymovirus* was used as the outgroup.

RESULTS

Host range, symptoms, and aphid transmission. PLDMV-P systemically infected 5 of the 20 plant species inoculated (Table 3). However, the infectivity or symptoms of the three strains differed from each other in four species with the exception of mosaic symptoms on *Cucumis metuliferus* Acc.2459. On *Carica papaya*, PLDMV-P(LDM) and -P(M) induced symptoms similar to those caused by the original virus source in the field; however, PLDMV-P(YM) induced distortion and mosaic symptoms, but not yellow mosaic. The yellow mosaic symptom appeared approximately 6 months after inoculation (data not shown). *Cucumis melo* var. *conomom* was used as a differential host for three strains; it develops a mosaic symptom after infection by PLDMV-P(LDM), a latent infection by PLDMV-P(M), and is not infected by PLDMV-P(YM). PLDMV-P(M) differed from the other strains by the chlorotic spots on systemically infected leaves of *Cucurbita pepo*. The mosaic symptom on *Cucumis sativus* cv. Sagami-hansiro induced by PLDMV-P(YM) was milder than that induced by the other strains.

The host range of P biotype was limited to caricaceous and cucurbitaceous plants. On the other hand, PLDMV-C(T) infected *Nicotiana benthamiana* in the family Solanaceae but not *Carica papaya* in the family Caricaceae. In addition, PLDMV-C(T) could be distinguished from P biotype strains by its infectivity to *Cucumis sativus* (House King L) in the family Cucurbitaceae and by its symptoms on *Cucumis metuliferus* Acc.2459 (Table 3).

All PLDMV strains were transmitted by *Myzus persicae* in a nonpersistent manner at a very high rate (78 to 100%).

Serological relationships. The serological relationships among the four strains of PLDMV were examined by western blotting and SDS-immunodiffusion tests (Fig. 2). Antisera against PLDMV-P(LDM), -P(YM), and -C(T) reacted well with all antigens in both serological tests. No spur formed between them in SDS-immunodiffusion tests (Fig. 2B). The reaction between PLDMV-P(M) and antisera appeared to be slightly weaker than the others in the SDS-immunodiffusion test. This weaker reaction might be caused by the lower quantity of PLDMV-P(M) than in other strains in infected leaves. The CP of PLDMV-C(T) migrated slower than that of other strains in SDS-polyacrylamide gels.

Complete nucleotide sequence of the PLDMV genome. The PLDMV-P(LDM) genome comprised 10,153 nucleotides, excluding the 3' poly(A) tail. The 5'-terminal region of the viral genome was determined from 18 independent 5' RACE clones. Eight clones contained five A residues at the viral cDNA end attached to the anchor primer C. Seven clones had four A residues, and the remaining three clones had six A residues at the viral cDNA end. To map the 5' end of the viral genome, a run-off cDNA product was synthesized with IRD800-labeled primer P1-7. The run-off product migrated as a single band. From these results, we concluded that the 5' sequence of PLDMV genome was 5'-AAA-AAUAUAAAAA...3'. The complete nucleotide sequence of the PLDMV genome was submitted to the DDBJ/EMBL/GenBank nucleotide sequence databases under accession no. AB088221. No identified potyviruses were identical to the nucleotide sequence of PLDMV-P(LDM) in a computed homology search (BLAST;

Fig. 1. Genetic map of **A**, the genomic RNA of Papaya leaf distortion mosaic virus (PLDMV)-P strain LDM and **B**, cDNA clones used for sequencing. **A**, Amino acids on either side of each cleavage site of polyprotein are below the site. Numbers above the polyprotein denote the first nucleotide coding for each protein. **B**, Primers used for amplifications are indicated by arrows with their names. The cDNA clones are shown with rectangles and their names. Clone PL50, shown with a black rectangle, was sequenced previously (22). Clones Nlb99, Nla41, CI64, 6K146, HC23, and P140 were made with the AP1 adaptor of Marathon cDNA Amplification Kit (Clontech Laboratories, Palo Alto, CA). Clones 6K146 and HC23 were partially sequenced to design the reverse primer HC-1 and forward primer P1-2. Adaptor primer C (9) was used for the 5' rapid amplification of cDNA ends.

provided online by the National Center for Biotechnology Information [NCBI]).

A single, large open reading frame (ORF) that encodes 3,269 amino acids was identified (molecular weight of 373,347) that commenced at nucleotide 135 and ended with a UAG termination codon at nucleotide 9942 to 9944 followed by the 3'-untranslated region (3'UTR) of 209 nucleotides. The 5'UTR of the genome was 134 nucleotides long and comparable to those of other potyviruses (75 to 213 nucleotides). The overall identity of the PLDMV 5'UTR with those of other potyviruses ranged from 41 to 68%. The 5'UTR of PLDMV had high (68%) identity with that of *Turnip mosaic virus* (TuMV); 20 nucleotides of the 5'-terminal region were completely identical. A sequence like the 'potybox' motif (1) was found at nucleotide 15 to 26 (UCAACAAACUU), in which only two underlined bases are different from those of the proposed consensus (UCAACACAACAU) and a box 'a' motif is included (20). Another motif box 'b' was found at nucleotide 40 to 47 (UCAAUACA), in which the underlined U is different from that of the proposed consensus G and the underlined A is inserted in the consensus (UCAAGCA) (20).

Polyprotein sequence features. From analogy with the other potyviruses, the processing sites of the polyprotein were deduced (Fig. 1). Amino acid residues in the cleavage site between P1 and HC-Pro (helper component-proteinase) of PLDMV are tyrosine (Y) and serine (S), which are identical to those found in 15 of 35 other potyviruses listed in Table 2. The cleavage site between HC-Pro and P3 is between glycine (G) and glycine (G), conserved in

all potyviruses in Table 2 except *Potato virus A* (PVA) (S/A). The other seven sites cleaved by NIa-Pro proteinase of PLDMV have the consensus sequence E-X-V-X-(H/A/F)-(Q/E)/(S/G/A), resembling those of other potyviruses. The polyprotein of PLDMV had 44 to 50% identities with those of other potyviruses. The amino acid sequences of the putative 10 functional proteins of PLDMV were compared with those of other potyviral proteins. The lowest identity was observed between PLDMV and Scallion mosaic virus (ScaMV) in the P1 protein (22%). A slightly higher identity between PLDMV and TuMV was observed in the polyprotein (50%) and the 6K1 (68%) and 6K2 (58%) proteins.

The highly conserved amino acid sequence motifs previously reported for potyviruses (31) were present at analogous positions in the PLDMV proteins. In the P1 protein, the conserved motif concerned with proteinase activity, H₃₈₈-X₈-D₃₉₇-X₃₂-G₄₂₉-X-S₄₃₁-G₄₃₂ (34), was found. Another conserved sequence, F-I-V-R-G, involved with the autoproteolytic activity of the P1 (28), was found at amino acid position 452 to 456, although in PLDMV phenylalanine (F) was substituted with isoleucine (I). The same substitution was observed in the P1 protein of PVA (27). In the second protein, HC-Pro, we found a possible zinc finger metal-binding motif, C-X₈-C-X₁₃-C-X₄-C (28). However, the third cysteine is replaced by methionine in the PLDMV HC-Pro protein (that is, C₅₀₈-X₈-C₅₁₇-X₁₃-M₅₃₁-X₄-C₅₃₆). The same replacement from cysteine to methionine was observed in the HC-Pro of *Sweet potato feathery mottle virus* (SPFMV) (29). The motif C-C-C-V-T involving long-distance movement (11) was observed in the PLDMV HC-Pro protein in the form of A₇₇₃-C-C-V-T₇₇₇, which is similar to *Peanut mottle virus* (PeMoV, accession no. AF023848). The K-I-T-C (2) and P-T-K (17) amino acid motifs of HC-Pro in-

TABLE 2. Viruses in the genus *Potyvirus* and their sequence accession numbers in the DDBJ/EMBL/GenBank database^a

Virus species	Abbreviation	Accession no.
<i>Bean common mosaic necrosis virus</i>	BcMNV	U19287
<i>Bean common mosaic virus</i>	BcMV	U05771
<i>Bean yellow mosaic virus</i>	BYMV	U47033
<i>Chilli veinal mottle virus</i>	ChiVMV	AJ237843
<i>Clover yellow vein virus</i>	CIYVV	AB011819
<i>Cocksfoot streak virus</i>	CSV	AF499738
<i>Cowpea aphid-borne mosaic virus</i>	CABMV	AF348210
<i>Dasheen mosaic virus</i>	DsMV	AJ298033
<i>Japanese yam mosaic virus</i>	JYMV	AB016500
<i>Johnsongrass mosaic virus</i>	JGMV	Z26920
<i>Leek yellow stripe virus</i>	LYSV	AJ307057
<i>Lettuce mosaic virus</i>	LMV	X97704
<i>Maize dwarf mosaic virus</i>	MDMV	AJ001691
<i>Onion yellow dwarf virus</i>	OYDV	AJ510223
<i>Papaya ringspot virus</i>	PRSV-HA	S46722
	PRSV-Tw	X97251
<i>Pea seed-borne mosaic virus</i>	PSbMV	D10930
<i>Peanut mottle virus</i>	PeMoV	AF023848
<i>Pepper mottle virus</i>	PepMoV	M96425
<i>Peru tomato mosaic virus</i>	PTV	AJ437280
<i>Plum pox virus</i>	PPV	X16415
<i>Potato virus A</i>	PVA	Z21670
<i>Potato virus V</i>	PVV	AJ243766
<i>Potato virus Y</i>	PVY	X12456
<i>Scallion mosaic virus</i>	ScaMV	AJ316084
<i>Sorghum mosaic virus</i>	SrMV	AJ310197
<i>Soybean mosaic virus</i>	SMV	S42280
<i>Sugarcane mosaic virus</i>	SCMV	AJ297628
<i>Sweet potato feathery mottle virus</i>	SPFMV	D86371
<i>Tamarillo mosaic virus</i>	TamMV	AJ131403
<i>Tobacco etch virus</i>	TEV	M15239
<i>Tobacco vein mottling virus</i>	TVMV	X04083
<i>Turnip mosaic virus</i>	TuMV	D10927
<i>Wild potato mosaic virus</i>	WPMV	AJ437279
<i>Yam mosaic virus</i>	YMV	U42596
<i>Zucchini yellow mosaic virus</i>	ZYMV	L29569
<i>Barley yellow mosaic virus</i> ^b	BaYMV	D01091

^a The taxonomy of viruses and abbreviations of virus names are based on the ICTV 7th report (33).

^b The sequence of *Barley yellow mosaic virus* in the genus *Bymovirus* was used as an outgroup.

TABLE 3. Host range and symptoms^a of Papaya leaf distortion mosaic virus (PLDMV) strains

Host	Biotypes and strains			
	P			C
	LDM	M	YM	T
<i>Carica papaya</i> (Tainung No. 2)	DM ^b	MM ^b	DM ^b	–
<i>Beta vulgaris</i>	–	–	–	–
<i>Chenopodium amaranticolor</i>	–	–	–	–
<i>Chenopodium quinoa</i>	–	–	–	–
<i>Chenopodium murale</i>	–	–	–	–
<i>Cucumis metuliferus</i> (Acc. 2459) (PI292190)	M	M	M	DM
<i>Cucumis sativus</i> (House King L) (Sagami-hansiro)	–	–	–	M
<i>Cucumis melo</i> var. <i>conomom</i> (Kuromon)	M	+	–	M
<i>Cucurbita ficifolia</i>	–	–	–	–
<i>Cucurbita maxima</i> (Ebisu)	–	–	–	–
<i>Cucurbita pepo</i> (Diner)	M	CS	M	M
<i>Luffa acutangula</i>	–	–	–	–
<i>Gomphrena globosa</i>	–	–	–	–
<i>Nicotiana benthamiana</i>	–	–	–	M
<i>Nicotiana glutinosa</i>	–	–	–	–
<i>Nicotiana tabacum</i> (Bright yellow)	–	–	–	–
<i>Petunia hybrida</i>	–	–	–	–
<i>Phaseolus vulgaris</i>	–	–	–	–
<i>Raphanus sativus</i>	–	–	–	–
<i>Vigna sinensis</i>	–	–	–	–

^a CS, chlorotic spot; DM, distortion and mosaic; M, mosaic; MM, mild mosaic; –, no infection; and +, latent infection. Symptoms were recorded visually 1 month after inoculation.

^b PLDMV-P(LDM) and -P(M) induced symptoms similar to the original virus source in the field. However, PLDMV-P(YM) did not induce yellow mosaic. Yellow mosaic appeared about 6 months after inoculation (data not shown).

Fig. 2. Serological relationships among Papaya leaf distortion mosaic virus (PLDMV) strains in **A**, western blots and **B**, sodium dodecyl sulfate (SDS)-immunodiffusion. **A**, Western blots of extracts from *Cucumis metuliferus* leaves uninfected (H) and infected with the leaf distortion mosaic (LDM), mosaic (M), yellow mosaic (YM), or *Trichosanthes* (T) strain. The blots were probed with anti-PLDMV-P(LDM) (left), -P(YM) (center), and -C(T) (right) antibody. **B**, The center wells contain antiserum against the LDM (LDM), YM (YM), or T (T) strain. The peripheral wells contain SDS-treated extracts of *C. metuliferus* leaves uninfected (H) and infected with LDM strain (LDM), M strain (M), YM strain (YM), or T strain (T).

Fig. 3. Phylogenetic tree of the amino acid sequences of the polyprotein of all fully sequenced members of the genus *Potyvirus*. The values adjacent to the nodes indicate the bootstrap confidence values (for 1,000 replicates). Only values >600 are shown. The scale bar shows the number of substitutions per residue. Abbreviations are listed in Table 2.

involved in transmissibility by aphids were found in the PLDMV HC-Pro protein at amino acid positions 533 to 536 and 791 to 793, respectively. The motifs of a nucleotide phosphate-binding site involving helicase activity (15,21), G₁₄₂₃-X-X-G₁₄₂₆-X-G₁₄₂₈-K-S-T₁₄₃₁ and (hydrophobic residue) n-D₁₅₁₂-E₁₅₁₃, were found in the N-terminal regions of the PLDMV CI protein. In the region of NIa-Pro protein, the conserved amino acids involved in proteinase activity (H₂₂₅₈, D₂₂₉₃, and C₂₃₆₃) were found (13). The NIb protein of PLDMV contained the consensus sequence motif S-G-X₃-T-X₃-N-T-X₃₀-G-D-D that is conserved in positive-stranded viral RNA-dependent RNA polymerases (18) at amino acid position 2766 to 2809.

A phylogenetic analysis based on the polyprotein amino acid sequences of members of the genus *Potyvirus* demonstrated that PLDMV does not form a subgroup with other potyviruses (Fig. 3).

CP comparison among PLDMV strains. The molecular relationships among four strains of PLDMV were analyzed based on the CP amino acid sequences. The CP nucleotide sequences of the PLDMV-P(M), -P(YM), and -C(T) have been submitted to the DDBJ/EMBL/GenBank databases under accession nos. AB092814, AB092815, and AB092816, respectively. No potyvirus sequences identical to those of PLDMV-P(M), -P(YM), and -C(T) were found using a BLAST search. Furthermore, only low identities were observed between these sequences and the CP nucleotide sequences of other potyviruses, containing cucurbitaceous viruses, such as *Watermelon mosaic virus* (accession no. L22907) or *Moroccan watermelon mosaic virus* (accession no. AF305545).

A multiple amino acid sequence alignment of the CP region of PLDMV strains is shown in Figure 4. The CPs of the P biotype

strains are composed of 293 amino acids, and the deduced molecular weights of CP of PLDMV-P(LDM), -P(M), and -P(YM) are 33,277, 33,093, and 33,221, respectively. When the three strains of the P biotype were compared, amino acids different from the two other strains are observed mostly in PLDMV-P(M): five amino acids at the N-terminal half and five amino acids at the C-terminal region of the CP. PLDMV-P(YM) differs from the two other strains in three amino acids at the CP core region. PLDMV-P(LDM) differs from the two other strains in three amino acids: two amino acids at the N-terminal region and one amino acid at the C-terminal region of the CP. The three PLDMV-P strains share 95 to 98% identity. In contrast, the PLDMV-C(T) CP is composed of 291 amino acids with a molecular weight of 33,019. In PLDMV-C(T) CP, two amino acid (N-K) deletions were found in the three N-K repeat sequences at amino acids 28 to 33 of the PLDMV-P CP. The PLDMV-C(T) CP has lower identities (88 to 89%) than those of the three strains of P biotype. Significant changes in hydrophobicity (A to E, N to I, A to T, and S to P) were observed at the N-terminal region of the CP of PLDMV-C(T).

The phylogenetic analysis based on the CP amino acid sequences grouped PLDMV strains in one cluster, distinct from clusters of other potyviruses. In the PLDMV cluster, P biotype strains (LDM, YM, and M) are closely related to each other but slightly separated from strain T of biotype C (data not shown).

DISCUSSION

The size of the potyviral genome ranges from 10,820 (SPFMV) to 9,324 nucleotides (ScaMV) (5). The PLDMV genome of

Fig. 4. Multiple alignment of the amino acid sequences of the coat proteins of Papaya leaf distortion mosaic virus (PLDMV) strains. Gaps (–) have been introduced for maximum alignment. Amino acids that are not conserved among four viruses are written in white letters on a black background. Asterisks indicate the positions of amino acids that are different between PLDMV-C(T) and the consensus of three PLDMV-P strains.

10,153 nucleotides is the fourth largest potyvirus ever reported, after SPFMV (10,820 nucleotides), *Onion yellow dwarf virus* (10,538 nucleotides) (6), and PRSV (10,326 nucleotides). PLDMV has a genome organization similar to other species in the genus *Potyvirus*. The fact that amino acid sequence motifs were conserved in potyviral proteins suggests that the putative PLDMV proteins are functionally homologous to other potyviral proteins. In the sequence comparisons of PLDMV with 35 other potyviruses, no virus showed a high level of identity with PLDMV over the entire genomic region, and no recombination event occurred in the PLDMV genome; therefore, PLDMV is genetically unrelated to other potyviruses. From these results and the previous report (22), we conclude that PLDMV is a distinct species of the genus *Potyvirus*.

The host range studies indicate that PLDMV can be classified into two biotypes; the P biotype infects papaya, whereas the C biotype is unable to infect papaya. Analysis of molecular relationships based on the CP sequence supports this classification, although the two biotypes are indistinguishable serologically using polyclonal antisera. Two new strains, M and YM, were found in papaya during field investigations, although the majority of the strains in the field were LDM. In the P biotype, each strain differs in characteristics in addition to the symptoms it causes on papaya. For example, the CP amino acid sequence of M strain was slightly different from other strains. Ten substitutions of amino acids in the CP of PLDMV-P(M) were found. However, they were serologically related, except that the J179P isolate of the YM strain could not be detected with PLDMV-P(LDM) antiserum in ELISA (*unpublished data*). It is not known whether all isolates of YM strain do not react with PLDMV-P(LDM) antiserum in ELISA. Further studies are needed to characterize the YM strain.

The C biotype differed from the P biotype in its host range and CP amino acid sequence. Significant changes of hydrophobicity were found at the N-terminal region of the CP of PLDMV-C(T). Although the N-terminal region constitutes the major epitope in the CP of viruses in the genus *Potyvirus* (30), no significant difference was found in the serology between the C and P biotype strains. The deduced molecular mass of the CP of all strains was 33 kDa. However, the CP of PLDMV-C(T) migrated slower than other strains in polyacrylamide gel electrophoresis (PAGE) of western blotting (Fig. 2). Numerous amino acid substitutions in the N-terminal region of the CP of PLDMV-C(T) may explain the discrepancy between the calculated molecular weight and the mobility in PAGE of western blotting.

Interestingly, another papaya-infecting potyvirus, PRSV, also has a papaya-infecting biotype (P) and a non-papaya-infecting biotype (W). We have identified common features among these potyviruses. For example, neither PRSV-W nor PLDMV-C infect papaya. Most size differences of PRSV-CP resulted from differences in the number of K-E repeats, located in the N-terminal region of the CP (4). In the case of PLDMV, the size differences among strains also were located in the N-K repeats at the N-terminal region of the CP. PRSV-P was rarely detected in naturally infected cucurbitaceous plants (3). In our field investigations, PLDMV-P was not isolated from cucurbitaceous plants (*unpublished data*).

Bateson et al. (3) and Wang et al. (35) compared the CP amino acid sequences in various strains of PRSV-P and -W and reported that sequence variability did not correspond to biotype differences but to geographical regions. PRSV-P was proposed as being derived locally from PRSV-W (3). In addition, Bateson et al. (4) compared the CP amino acid sequences of 93 PRSV strains and reported that PRSV may have originated in Asia and that populations of the PRSV-P biotype have evolved on several occasions from PRSV-W populations. They also reported that the mutation from PRSV-W to PRSV-P involves changes to one or more single amino acids, and these changes arise frequently in PRSV populations of cucurbits as a result of random, natural mutation (4).

Recently, Chen et al. (7) reported that a chimeric PRSV-W containing the CI-, 6K2-, and NIa-coding regions of PRSV-P acquired infectivity to papaya. They concluded that the host determinant gene of PRSV for papaya infection was located between the CI and NIa genes (7).

PLDMV was found in 1954 in Japan, although papaya arrived in Japan in the 18th century (12). PLDMV-C was also detected from three cucurbitaceous plants: *Cucumis melo*, *Cucurbita moschata*, and *Lagenaria siceraria* on Ishigaki Island (*unpublished data*). This indicates that PLDMV-C is widely spread in the family Cucurbitaceae. If the hypothesis for PRSV is applicable to PLDMV, PLDMV is primarily a pathogen of the family Cucurbitaceae and PLDMV-P could be expected to have been adapted to infect papaya from PLDMV-C by a mutation. The common phylogenetic divergence of papaya-infecting potyviruses into biotypes may signify that a similar evolutionary event occurred in viral genomes, so that the virus acquired infectivity on papaya and that the mutation for this adaptation to infect papaya may arise frequently in these viruses in cucurbits.

Although the complete nucleotide sequence of PLDMV-P is available, no sequence data for PLDMV-C is available, except for the CP gene. Comparative studies between PLDMV-P and PLDMV-C may identify the site for adaptation for papaya infection. We have sequenced the CP genes for PLDMV from only four isolates, but more sequence data for the CP gene are needed to estimate the origin of PLDMV.

Chen et al. (8) reported the occurrence of PLDMV in Taiwan. They also reported that the CP amino acid of the Taiwanese PLDMV had 95% identity with that of the Japanese PLDMV. All of the cucurbitaceous plants from which PLDMV was detected in Japan also grow in Taiwan. However, it is not known whether the Taiwanese PLDMV was introduced from Japan or if it existed locally but was missed in the wake of the damage by PRSV. Further comparative studies between PLDMV-P and -C are needed to determine the host determinant gene for papaya infection and the origin of this species.

ACKNOWLEDGMENTS

We thank F. Fukumoto, T. Usugi, and I. Uyeda for their useful suggestions; and the late M. Iwasaki, the late S. Kashiwazaki, and the late C. Noda for helpful discussions. A part of this work was supported by the funds provided by JIRCAS (Japan International Research Center for Agricultural Sciences).

LITERATURE CITED

- Atreya, C. D. 1992. Application of genome sequence information in potyvirus taxonomy: An overview. Pages 17-23 in: *Potyvirus Taxonomy*. O. W. Barnett, ed. Springer-Verlag, New York.
- Atreya, C. D., and Pirone, T. P. 1993. Mutational analysis of the helper component-proteinase gene of a potyvirus: Effects of amino acid substitutions, deletions, and gene replacement on virulence and aphid transmissibility. *Proc. Natl. Acad. Sci. USA* 90:11919-11923.
- Bateson, M. F., Henderson, J., Chaleeprom, W., Gibbs, A. J., and Dale, J. L. 1994. Papaya ringspot potyvirus: Isolate variability and the origin of PRSV type P (Australia). *J. Gen. Virol.* 75:3547-3553.
- Bateson, M. F., Lines, R. E., Revill, P., Chaleeprom, W., Ha, C. V., Gibbs, A. J., and Dale, J. L. 2002. On the evolution and molecular epidemiology of the potyvirus Papaya ringspot virus. *J. Gen. Virol.* 83:2575-2585.
- Che, J., Zheng, H. Y., Chen, J. P., and Adams, M. J. 2002. Characterisation of a potyvirus and a potexvirus from Chinese scallion. *Arch. Virol.* 147:683-693.
- Chen, J., Adams, M. J., Zheng, H. Y., and Chen, J. P. 2003. Sequence analysis demonstrates that Onion yellow dwarf virus isolates from China contain a P3 region much larger than other potyviruses. *Arch. Virol.* 148:1165-1173.
- Chen, K. C., Wang, C. H., Liu, F. L., Su, W. C., and Yeh, S. D. 2003. Analysis of host determinant genes of *Papaya ringspot virus* for papaya infection. Page 303 in: *Abstracts 8th International Congress of Plant Pathology*. ICPP Programme Committee, Christchurch, New Zealand.

8. Chen, L. F., Bau, H. J., and Yeh, S. D. 2000. Identification of virus capable of breaking transgenic resistance of papaya conferred by the coat protein gene of *Papaya ringspot virus*. Pages 51-52 in: Abstracts of International Symposium on Tropical and Subtropical Fruits. Griffith University, Queensland.
9. Chen, Z. 1996. Simple modifications to increase specificity of the 5' RACE procedure. *Trends Gen.* 12:87-88.
10. Clark, M. F., and Adams, A. N. 1977. Characteristics of the microplate method of enzyme-linked immunosorbent assay for the detection of plant viruses. *J. Gen. Virol.* 34:475-483.
11. Cronin, S., Verchot, J., Haldeman-Cahill, R., Schaad, M. C., and Carrington, J. C. 1995. Long-distance movement factor: A transport function of the potyvirus helper component proteinase. *Plant Cell* 7:549-559.
12. Degi, K. 1998. Breeding of papaya and its possibility in Okinawa. (in Japanese) *Okinawa Agric.* 33:87-101.
13. Dougherty, W. G., Parks, T. D., Cary, S. M., Bazan, J. F., and Fletterick, R. J. 1989. Characterization of the catalytic residues of the tobacco etch virus 49-kDa proteinase. *Virology* 172:302-310.
14. Dumas, J. B., Edwards, M., Delort, J., and Mallet, J. 1991. Oligodeoxy-ribonucleotide ligation to single-stranded cDNAs: A new tool for cloning 5' ends of mRNAs and for constructing cDNA libraries by in vitro amplification. *Nucleic Acids Res.* 19:5227-5232.
15. Gorbalenya, A. E., Koonin, E. V., Donchenko, A. P., and Blinov, V. M. 1988. A novel superfamily of nucleoside triphosphate-binding motif containing proteins which are probably involved in duplex unwinding in DNA and RNA replication and recombination. *FEBS Lett.* 235:16-24.
16. Hataya, T., Uchino, K., Arimoto, R., Suda, N., Sano, T., Shikata, E., and Uyeda, I. 2000. Molecular characterization of *Hop latent virus* and phylogenetic relationships among viruses closely related to carlaviruses. *Arch. Virol.* 145:2503-2524.
17. Huet, H., Gal-On, A., Meir, E., Lecoq, H., and Raccah, B. 1994. Mutations in the helper component protease gene of zucchini yellow mosaic virus affect its ability to mediate aphid transmissibility. *J. Gen. Virol.* 75:1407-1414.
18. Kamer, G., and Agros, P. 1984. Primary structural comparison of RNA-dependent polymerase from plant, animal and bacterial viruses. *Nucleic Acids Res.* 12:7269-7282.
19. Kawano, S., and Yonaha, T. 1992. The occurrence of papaya leaf-distortion mosaic virus in Okinawa. *Tech. Bull. FFTC (Food and Fertilizer Technology Center for the Asian and Pacific Region, Taipei)* 132:13-23.
20. Lain, S., Riechmann, J. L., and Garcia, J. A. 1989. The complete nucleotide sequence of plum pox potyvirus RNA. *Virus Res.* 13:157-172.
21. Lain, S., Riechmann, J. L., and Garcia, J. A. 1990. RNA helicase: A novel activity associated with a protein encoded by a positive strand RNA virus. *Nucleic Acids Res.* 18:7003-7006.
22. Maoka, T., Kashiwazaki, S., Tsuda, S., Usugi, T., and Hibino, H. 1996. Nucleotide sequence of the capsid protein gene of papaya leaf-distortion mosaic potyvirus. *Arch. Virol.* 141:197-204.
23. Maoka, T., Kawano, S., and Usugi, T. 1995. Occurrence of the P strain of papaya ringspot virus in Japan. *Ann. Phytopathol. Soc. Jpn.* 61:34-37.
24. Maoka, T., Tsuda, S., Usugi, T., Noda, C., and Iwasaki, M. 2002. Changes in serological reactivity of papaya leaf distortion mosaic virus caused by papain in *Carica papaya* L. and its detection by using antipain or papain. *J. Gen. Plant Pathol.* 68:89-93.
25. Page, R. D. M. 1996. TREEVIEW: An application to display phylogenetic trees on personal computers. *Comput. Appl. Biosci.* 12:357-358.
26. Purcifull, D. E., Edwardson, J. R., Hiebert, E., and Gonsalves, D. 1984. Papaya ringspot virus. No. 292 in: *Descriptions of Plant Viruses*. Commonw. Mycol. Inst., Kew, England.
27. Puurand, U., Makinen, K., Paulin, L., and Saarma, M. 1994. The nucleotide sequence of potato virus A genomic RNA and its sequence similarities with other potyviruses. *J. Gen. Virol.* 75:457-461.
28. Robaglia, C., Durand-Tardif, M., Tronchet, M., Boudazin, G., Astier-Manificier, S., and Casse-Delbart, F. 1989. Nucleotide sequence of potato virus Y (N strain) genomic RNA. *J. Gen. Virol.* 70:935-947.
29. Sakai, J., Mori, M., Morishita, T., Tanaka, M., Hanada, K., Usugi, T., and Nishiguchi, M. 1997. Complete nucleotide sequence and genome organization of sweet potato feathery mottle virus (S strain) genomic RNA: The large coding region of the P1 gene. *Arch. Virol.* 142:1553-1562.
30. Shukla, D. D., Stride, P. M., Tracy, S. L., Gough, K. H., and Ward, C. W. 1988. The N and C termini of the coat proteins of potyviruses are surface-located and the N terminus contains the major virus-specific epitopes. *J. Gen. Virol.* 69:1497-1508.
31. Shukla, D. D., Ward, C. W., and Brunt, A. A. 1994. Genome structure, variation and function. Pages 74-112 in: *The Potyviridae*. CAB International, West Sussex.
32. Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F., and Higgins, D. G. 1997. The CLUSTAL X windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* 25:4876-4882.
33. Van Regenmortel, M. H. V., Fauquet, C. M., Bishop, D. H. L., Carstens, E. B., Estes, M. K., Lemon, S. M., Maniloff, J., Mayo, M. A., McGeoch, D. J., Pringle, C. R., and Wickner, R. B. (eds.) 2000. *Virus Taxonomy*. Seventh Report of the International Committee on Taxonomy of Viruses. Academic Press, San Diego.
34. Verchot, J., Koonin, E. V., and Carrington, J. C. 1991. The 35-kDa protein from the amino-terminus of the potyviral polyprotein functions as a third virus-encoded proteinase. *Virology* 185:527-535.
35. Wang, C. H., Bau, H. J., and Yeh, S. D. 1994. Comparison of the nuclear inclusion b protein and coat protein gene of five papaya ringspot virus strains distinct in geographic origin and pathogenicity. *Phytopathology* 84:1205-1210.
36. Yeh, S. D., Jan, F. J., Chiang, C. H., Doong, T. J., Chen, M. C., Chung, P. H., and Bau, H. J. 1992. Complete nucleotide sequence and genetic organization of papaya ringspot virus RNA. *J. Gen. Virol.* 73:2531-2541.
37. Yonaha, T. 1976. Viruses isolated from papaya in Okinawa (Japan). 1: Properties of papaya ringspot virus. *Bull. Coll. Agric. Univ. Ryukyus* 23:115-124.