

Nucleotide sequence and genome organization of Dweet mottle virus and its relationship to members of the family *Betaflexiviridae*

Subhas Hajeri · Chandrika Ramadugu ·
Manjunath Keremane · Georgios Vidalakis ·
Richard Lee

Received: 10 February 2010 / Accepted: 12 July 2010 / Published online: 20 July 2010
© The Author(s) 2010. This article is published with open access at Springerlink.com

Abstract The nucleotide sequence of Dweet mottle virus (DMV) was determined and compared to sequences of members of the families *Alphaflexiviridae* and *Betaflexiviridae*. The DMV genome has 8,747 nucleotides (nt) excluding the 3' poly-(A) tail. DMV genomic RNA contains three putative open reading frames (ORFs) and untranslated regions of 73 nt at the 5' and 541 nt at 3' termini. ORF1 potentially encoding a 227.48-kDa polyprotein, which has methyltransferase, oxygenase, endopeptidase, helicase, and RNA-dependent RNA polymerase (RdRP) domains. ORF2 encodes a movement protein of 40.25 kDa, while ORF3 encodes a coat protein of 40.69 kDa. Protein database searches showed 98–99% matches of DMV ORFs with citrus leaf blotch virus (CLBV) sequences. Phylogenetic analysis based on the RdRP core domain revealed that DMV is closely related to

CLBV as a member of the genus *Citrivirus*. DMV did not satisfy the molecular criteria for demarcation of an independent species within the genus *Citrivirus*, family *Betaflexiviridae*, and hence, DMV can be considered a CLBV isolate.

Introduction

Dweet mottle virus (DMV) was reported from Riverside, California, in 1968 during reindexing of a 'Cleopatra' mandarin variety introduced from Florida in the Citrus Variety Improvement Program, the forerunner of the Citrus Clonal Protection Program (CCPP) [11]. DMV produced leaf-mottling symptoms only in 'Dweet' tangor that was similar to but distinct from the symptoms of psorosis and concave gum [18]. A partial sequence analysis showed that DMV has very high sequence homology (over 96%) with citrus leaf blotch virus (CLBV) [26]. CLBV was first reported in Spain in an introduction of 'Nagami' kumquat from Corsica [16]. Both DMV and CLBV induce mottling in 'Dweet' tangor and stem pitting in 'Etrog' citron [6]. However, only CLBV causes vein clearing in 'Pineapple' sweet orange, and bud union crease on trifoliate and trifoliate hybrids rootstocks and has been reported to be seed transmitted [6–8, 16, 26]. Experiments with CLBV infectious clones have suggested that the bud union crease and vein clearing symptoms may be caused by a different agent associated with 'Nagami' kumquat sources [25]. Since DMV and CLBV have distinct biological similarities and differences, a direct comparison of the full DMV genome at the genetic level was undertaken. The full genomic sequence and genome organization of DMV are presented here and compared to the members of families *Alphaflexiviridae* and *Betaflexiviridae*.

Nucleotide sequence data reported is available in the NCBI GenBank databases under the accession number 'FJ009367'.

Electronic supplementary material The online version of this article (doi:10.1007/s00705-010-0758-1) contains supplementary material, which is available to authorized users.

S. Hajeri · G. Vidalakis (✉)
Department of Plant Pathology and Microbiology,
University of California, 900 Univ. Ave, Riverside,
CA 92521, USA
e-mail: vidalg@ucr.edu

C. Ramadugu
Department of Botany and Plant Sciences,
University of California, Riverside, CA 92521, USA

M. Keremane · R. Lee
USDA-ARS National Clonal Germplasm Repository for Citrus
and Dates, 1060 Martin Luther King Blvd, Riverside,
CA 92507, USA

Provenance of virus material and sequence analysis

DMV isolate-932, according to the CCPP citrus disease bank records, was used for the present study. Total RNA was extracted using an RNeasy kit (Qiagen, CA, USA) according to manufacturer's instructions from DMV-932-infected plants maintained at the USDA-ARS National Clonal Germplasm Repository for Citrus and Dates (NCGRCD), Riverside, CA. Primers were designed using sequence information of CLBV from the online database (NCBI accession no. NC_003877). Primers were designed in such a way that overlapping clones can be generated along the length of the genome (Supplementary Table 1). Reverse transcription-polymerase chain reaction (RT-PCR) was performed with RNA from the DMV-infected plants as a template (Supplementary Fig. 1). RT-PCR products with a 3'-A tail (*Taq* DNA polymerase activity) were cloned into pGEM-T Easy Vector (Promega Corp., USA) using a TA-cloning strategy followed by transformation of chemically competent *Escherichia coli* DH5- α cells [20].

Nucleotide sequence data were assembled using the ContigExpress tool (Vector NTI Advance 10-InforMax, USA). ORF Finder at NCBI was used to search for potential ORFs in the DMV genome. Conserved domains in the amino acid sequences were identified by CD-Search on NCBI [13, 14]. Sequence information was used in the Basic Local Alignment Search Tool (BLAST) program to identify regions of similarity between DMV sequences and the sequence database. Comparative analysis of amino acid and nucleotide sequences of the RNA-dependent RNA polymerase (RdRP) domain of the replicase polyprotein and the capsid protein (CP) of DMV was performed with selected members of the families *Alphaflexiviridae* and *Betaflexiviridae* using the ClustalX2 and GeneDoc programs [10, 12, 17].

Phylogenetic and molecular evolutionary analyses were performed using the conserved amino acid sequences of the RdRP domains from 6 members of the family *Alphaflexiviridae* and 17 members of the family *Betaflexiviridae* using the MEGA version 4.1 [21] and MrBayes [19] programs. Neighbor-Joining, Minimum Evolution, Maximum Parsimony, UPGMA (Unweighted Pair Group Method with Arithmetic mean) and Bayesian methods of phylogeny estimation were utilized [28]. Multiple alignments of amino acid sequences were obtained using ClustalX2 [12]. Bootstrap values were obtained by including 10,000 replicates in the MEGA4.1 program, and the MrBayes program was run for 2 million generations to estimate the posterior probabilities.

Total RNA from the leaves of healthy and DMV-infected 'Dweet' tangor plants (1 g) was extracted, and 20 μ g of total RNA was used for northern blot analysis using the method of Annamalai and Rao [2]. For the

detection of DMV genomic (g) and subgenomic (sg) RNAs, radiolabeled riboprobes complementary to the 5'-terminal region (1–870 nt) and 3'-terminal region (8,357–8,747 nt) were used.

Genome organization and phylogenetic relationship to members of the families *Alphaflexiviridae* and *Betaflexiviridae*

Overlapping cDNA fragments ranging from 1.0 to 1.87 kb were used to compile the nucleotide sequence of the DMV genome (Supplementary Fig. 1). The consensus nucleotide sequence was submitted to GenBank (accession FJ009367). The highest nucleotide sequence homology (97–98%) was with CLBV isolates from Spain and New Zealand. The genomic organization of DMV is presented in Fig. 1a. The complete genome of the virus consists of 8,747 nucleotides (nt), excluding the poly-A tail at the 3' end. The overall A + U and G + C content of DMV genomic RNA was 60 and 40%, respectively. ORF-1, which is at the 5'-proximal end, comprises 5,889 nt (nt position 74–5,962) with a putative translation product of 1,962 aa (227.48 kDa), a putative polyprotein with replicase activity. ORF-2 is 1,089 nt long and is located at position 5,962–7,050. The putative translation product of ORF-2 is a movement protein that is 362 aa long (40.25 kDa). The 3'-proximal ORF is located between residues 7,115 and 8,206 and encodes a putative coat protein of 363 aa (40.69 kDa). The 5' untranslated region (UTR) is about 74 nt long, while the 3' UTR is 541 nt long. An intergenic region of 64 nt is present between ORF-2 and ORF-3. A one-nucleotide overlap between ORF-1 and ORF-2 indicated a potential +1 frameshift translation strategy employed by the virus. The conserved core domains on the amino acid sequences of the replicase polyprotein (Supplementary Table 3), as identified by the CD-Search on NCBI, are presented within the ORF-1 in Fig. 1a. Northern blot hybridization of total RNA with probes of different gRNA regions revealed that DMV produces two 3'-coterminal and two 5'-coterminal subgenomic RNAs (Fig. 1b).

The highest amino acid sequence similarity for the RdRP of DMV was with CLBV isolates (99–100%), followed by cherry mottle leaf virus (60%) and apple stem pitting virus (55%) (Fig. 2). The highest amino acid sequence similarity for CP of DMV was with CLBV isolates (98–99%), followed by apple stem pitting virus (6%), apple stem grooving virus (5%) and cherry mottle leaf virus (4%) (Supplementary Fig. 2). DMV and CLBV isolates were clustered in a single clade within the genus *Citrivirus*, in both neighbor-joining and maximum-parsimony topologies (100% bootstrap) with close ancestral relationships with members of the genus *Trichovirus*,

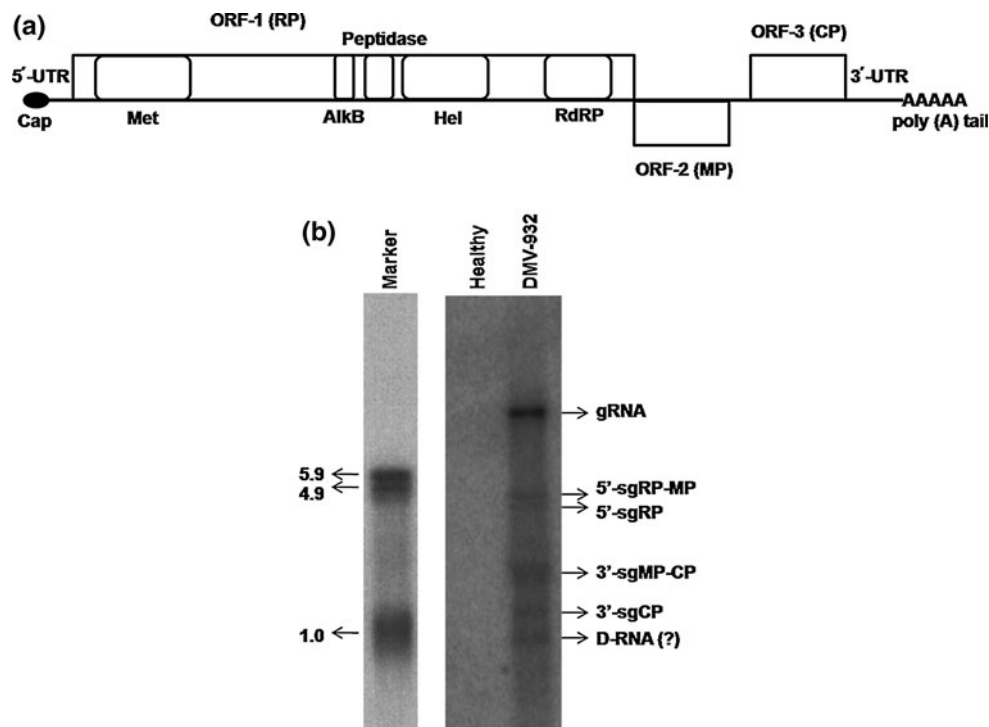


Fig. 1 **a** Schematic representation of the genomic organization of Dweet mottle virus (DMV). The open reading frames (*ORFs*) are shown as *boxes*. ORF-1 potentially encodes a 227-kDa replicase polyprotein, ORF-2 encodes a movement protein of 40 kDa, and ORF-3 encodes a coat protein of 41 kDa. The genome is 8,747 nt long, excluding 3'-poly(A) tail. The 5' untranslated region (*UTR*) is about 74 nt long, while the 3' UTR is 541 nt long. An intergenic region of 64 nt is present between ORF-2 and ORF-3. Conserved core domains of the replicase polyprotein (1,962 aa) of DMV are shown as *boxes* within ORF-1. Methyl transferase is involved in capping, AlkB (oxygenase protein) in integrity of RNA by oxidative demethylation,

peptidase in auto-proteolysis, helicase and RNA-dependent RNA polymerase (RdRP) in virus replication. **b** Northern blot hybridization analysis of total RNA from 1 g of healthy plant tissue (*left lane* of right membrane) and 1 g DMV infected tissue (*right lane* of right membrane), using radiolabeled riboprobes from the 5'-terminal and 3'-terminal regions of the DMV genome. *Arrowheads* indicate the positions of the genomic RNA (gRNA), 5'-coterminal subgenomic RNAs (5'-sgRP-MP and 5'-sgRP), 3'-coterminal subgenomic RNAs (3'-sgMP-CP and 3'-sgCP), and unidentified defective RNA (D-RNA) reacting with the probes. Sizes (kb) of marker are shown on the *left* membrane

family *Betaflexiviridae* (Supplementary Fig. 3). These phylogenetic relationships were also supported by the minimum-evolution, UPGMA, and Bayesian predicted topologies (data not shown).

Discussion

The complete nucleotide sequence of DMV was determined and compared to sequences of members of the families *Alphaflexiviridae* and *Betaflexiviridae*. Analyses of the entire nucleotide sequence and northern blot analysis revealed that the DMV genome is closely related to that of CLB, with very similar ORFs and protein products [24]. The new virus genus *Citivirus* has been considered for CLB because of the peculiar biological, structural and molecular characteristics compared to other members of *Betaflexiviridae* [4, 15, 23]. In addition, Vives et al. [27] compared 14 isolates of CLB from different geographical regions of the world and found low genetic diversity. The

historical background of the California DMV isolate in combination with the complete genome sequence and the phylogenetic analysis presented here strongly supports the inclusion of DMV in the genus *Citivirus*. According to the molecular criteria for species demarcation within a genus of the family *Betaflexiviridae*, amino acid sequences of polymerase and CP genes must differ by more than 10% [1]. Sequence homology and genome organization support the hypothesis that DMV is a CLB isolate and not a member of a distinct species.

An interesting revelation of the organization study of conserved core domains of the DMV replicase polyprotein was the AlkB protein. Numerous single-stranded plant RNA viruses encode AlkB domains, and remarkably, the majority of these belong to the family *Betaflexiviridae* [3, 15]. Bacterial and mammalian AlkB proteins include 2-oxoglutarate (2-OG) and Fe(II)-dependent oxygenase, which reverse methylation damage in RNA and DNA [5, 9]. AlkB maintains the integrity of the viral RNA genome by oxidative demethylation through repair of

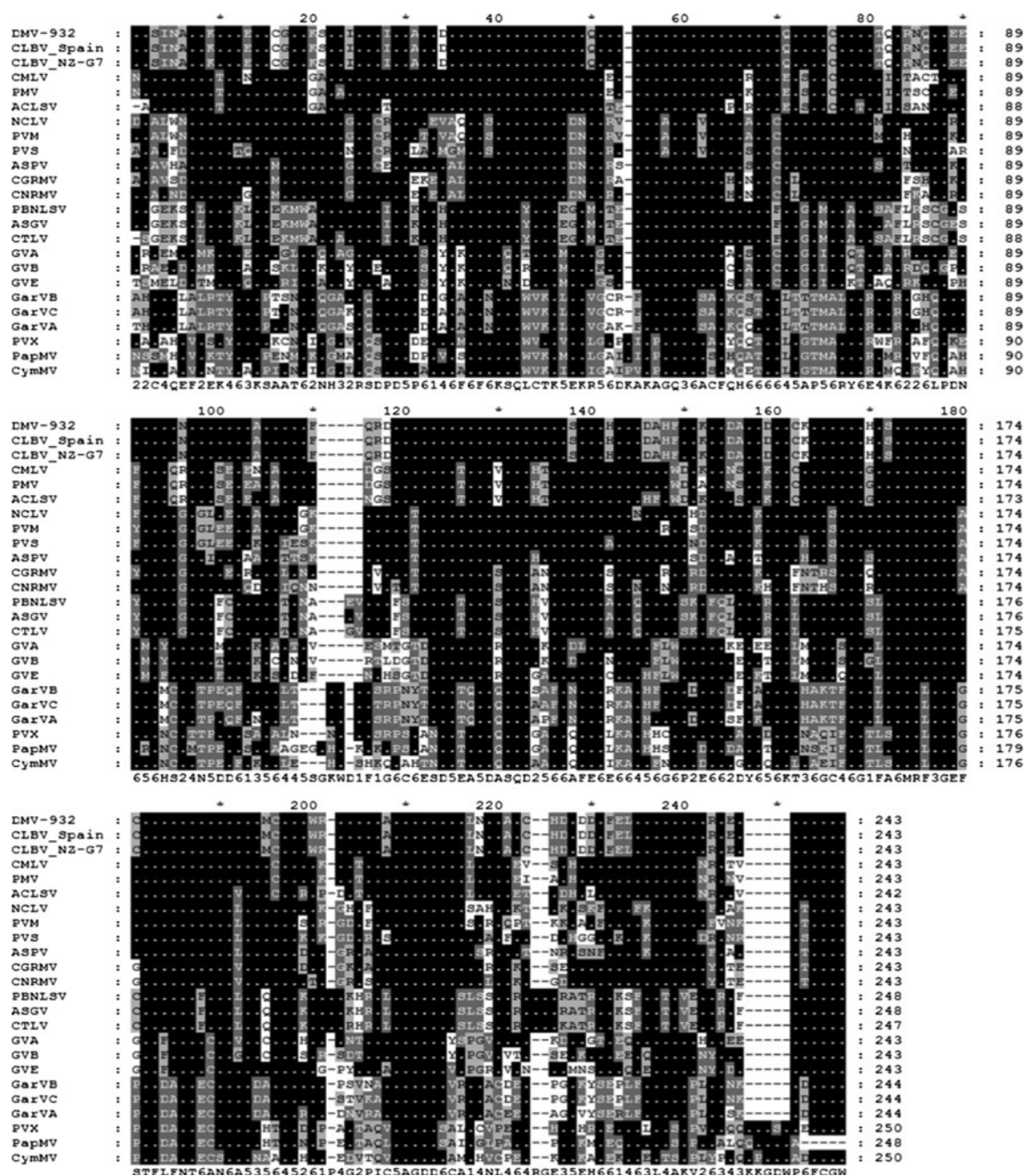


Fig. 2 Alignment of amino acid sequence of RdRP domains of selected members of the families *Alphaflexiviridae* and *Betaflexiviridae* using the GeneDoc program [15]. Black and gray shading represent sequences that are identical and conserved among all of the aligned sequences. The intensity of the shading represents the degree of conservation. Three viruses from selected genera of the families *Alphaflexiviridae* and *Betaflexiviridae* were used for analysis (viruses and GenBank accession numbers are as follows: Dweet mottle virus isolate 932 [DMV-932: FJ009367], citrus leaf blotch virus Spain [CLBV Spain: NC_003877], CLB_V New Zealand-G78 [CLBV_NZ-G78: EU857540], apple chlorotic leaf spot virus [ACLSV: NC_001409], cherry mottle leaf virus [CMLV: NC_002500], peach mosaic virus [PMV: NC_011552], cherry green ring mottle virus

[CGRMV: NC_001946], cherry necrotic rusty mottle virus [CNRMV: NC_002468], apple stem pitting virus [ASPV: NC_003462], potato virus S [PVS: NC_007289], narcissus common latent virus [NCLV: NC_008266], potato virus M [PVM: NC_001361], citrus tatter leaf virus [CTLV: FJ355920], pear black necrotic leaf spot virus [PBNLSV: AY596172], apple stem grooving virus [ASGV: NC_001749], grapevine virus E [GVE: NC_011106], grapevine virus A [GVA: NC_003604], grapevine virus B [GVB: NC_003602], garlic virus B [GarVB: EF596816], garlic virus C [GarVC: NC_003376], garlic virus A [GarVA: NC_003375], *Cymbidium mosaic virus* [CymMV: NC_001812], potato virus X [PVX: NC_011620], papaya mosaic virus [PapMV: NC_001748])

deleterious methylation damage [22]. Interestingly, most of the AlkB-containing viruses infect woody or perennial plants such as citrus, which is the only known natural host for both DMV and CLBV. The long-term survival of viruses within a single infected plant might be attributed to the functional advantages provided by the AlkB protein [15].

Acknowledgments We thank Dr. ALN Rao and Venkatesh Sivandam (UC Riverside) for assisting in the northern blot analysis. We also acknowledge the technical assistance of Toan Khuong (UC Riverside) and the CCPP personnel.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Adams MJ, Antoniw JF, Bar-Joseph M, Brunt AA, Candresse T, Foster GD, Martelli GP, Milne RG, Fauquet CM (2004) The new plant virus family *Flexiviridae* and assessment of molecular criteria for species demarcation. *Arch Virol* 149:1045–1060
- Annamalai P, Rao AL (2005) Replication-independent expression of genome components and capsid protein of brome mosaic virus *in planta*: a functional role for viral replicase in RNA packaging. *Virol* 338:96–111
- Aravind L, Koonin EV (2001) The DNA-repair protein AlkB, EGL-9, and leprecan define new families of 2-oxoglutarate- and iron-dependent dioxygenases. *Genome Biol* 2: research0007.1–research0007.8
- Carstens EB (2010) Ratification vote on taxonomic proposals to the International Committee on Taxonomy of Viruses (2009). *Arch Virol* 155:133–146
- Falnes PO (2004) Repair of 3-methylthymine and 1-methylguanine lesions by bacterial and human AlkB proteins. *Nucleic Acids Res* 32:6260–6267
- Galipienso L, Navarro L, Ballester-Olmos JF, Pina J, Moreno P, Guerri J (2000) Host range and symptomatology of a graft-transmissible pathogen causing bud union crease of citrus on trifoliolate rootstocks. *Plant Pathol* 49:308–314
- Galipienso L, Vives MC, Moreno P, Milne RG, Navarro L, Guerri J (2001) Partial characterisation of *Citrus leaf blotch virus*, a new virus from Nagami kumquat. *Arch Virol* 146:357–368
- Guerri J, Pina JA, Vives MC, Navarro L, Moreno P (2004) Seed transmission of *Citrus leaf blotch virus*: implications in quarantine and certification programs. *Plant Dis* 88:906
- Koivisto P, Robins P, Lindahl T, Sedgwick B (2004) Demethylation of 3-methylthymine in DNA by bacterial and human DNA dioxygenases. *J Biol Chem* 279:40470–40474
- Koonin EV, Dolja VV (1993) Evolution and taxonomy of positive-strand RNA viruses: implications of comparative analysis of amino acid sequences. *Crit Rev Biochem Mol Biol* 28:375–430
- Krueger RR, Bash JA, Lee RF (2005) Dweet mottle virus and Citrus leaf blotch virus. *Topics Subtrop Newsl* 3:1
- Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* 23:2947–2948
- Marchler-Bauer A, Anderson JB, Chistaz F, Derbyshire MK, De Weese-Scott C, Fong JH, Geer LY, Geer RC, Gonzales NR, Gwadz M, He S, Hurwitz DI, Jackson JD, Ke Z, Lanczycki CJ, Liebert CA, Liu C, Lu F, Lu S, Marchler GH, Mullokandov M, Song JS, Tasneem A, Thanki N, Yamashita RA, Zhang D, Zhang N, Bryant SH (2009) CDD: specific functional annotation with the Conserved Domain Database. *Nucleic Acids Res* 37:205–210
- Marchler-Bauer A, Bryant SH (2004) CD-Search: protein domain annotations on the fly. *Nucleic Acids Res* 32:327–331
- Martelli GP, Adams MJ, Kreuze JF, Dolja VV (2007) Family *Flexiviridae*: a case study in virion and genome plasticity. *Annu Rev Phytopathol* 45:73–100
- Navarro L, Pina JA, Ballester JF, Moreno P, Cambra M (1984) A new graft transmissible disease found in ‘Nagami’ kumquat. In: *Proc 9th Conf Org Citrus Virologists*, pp 234–240
- Nicholas KB, Nicholas HB Jr (1997) GeneDoc: analysis and visualization of genetic variation. *EMBASE News* 4
- Roistacher CN, Blue RL (1968) A psorosis-like virus causing symptoms only on Dweet tangor. In: *Proc 4th Conf Int Org Citrus Virologists*, pp 13–18
- Ronquist F, Huelsenbeck JP (2003) MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572–1574
- Sambrook J, Russell DW (2001) *Molecular cloning. A laboratory manual*, 3rd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York
- Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol Biol Evol* 24:1596–1599
- van den Born E, Omelchenko MV, Bekkelund A, Leihne V, Koonin EV, Dolja VV, Falnes PO (2008) Viral AlkB proteins repair RNA damage by oxidative demethylation. *Nucleic Acids Res* 36:5451–5461
- Vives MC, Galipienso L, Navarro L, Moreno P, Guerri J (2001) The nucleotide sequence and genomic organization of citrus leaf blotch virus: candidate type species for a new virus genus. *Virol* 287:225–233
- Vives MC, Galipienso L, Navarro L, Moreno P, Guerri J (2002) Characterization of two kinds of subgenomic RNAs produced by Citrus leaf blotch virus. *Virol* 295:328–336
- Vives MC, Martin S, Ambros S, Renovell A, Navarro L, Pina JA, Moreno P, Guerri J (2008) Development of a full-genome cDNA clone of Citrus leaf blotch virus and infection of citrus plants. *Mol Plant Pathol* 9:787–797
- Vives MC, Pina JA, Juárez J, Navarro L, Moreno P, Guerri J (2004) Dweet mottle disease is probably caused by *Citrus leaf blotch virus*. Abstracts, 16th Conf Int Soc Citrus Virologists, 98
- Vives MC, Rubio L, Galipienso L, Navarro L, Moreno P, Guerri J (2002) Low genetic variation between isolates of Citrus leaf blotch virus from different host species and different geographical origins. *J Gen Virol* 83:2587–2591
- Yang Z, Rannala B (1997) Bayesian phylogenetic inference using DNA sequences: a Markov chain Monte carlo method. *Mol Biol Evol* 14:717–724