Arch Virol (2015) 160:3143–3147
DOI 10.1007/s00705-015-2605-x

ANNOTATED SEQUENCE RECORD

Sequence and genome organization of papaya meleira virus infecting papaya in Brazil

Emanuel F. M. Abreu[1] · Cleidiane B. Daltro[1] · Elsa O. P. L. Nogueira[2] · Eduardo C. Andrade[1] · Francisco J. L. Aragão[2]

Received: 6 June 2015 / Accepted: 7 September 2015 / Published online: 14 September 2015

Abstract Papaya sticky disease (‘*meleira*’) was first observed in Brazil at the beginning of the 1980s. The disease is characterized by intense latex exudation from the fruit surface that becomes dark as it oxidizes, which makes it difficult to sell. The causal agent, which has been called papaya meleira virus (PMeV), has been identified as an isometric virus particle, approximately 50 nm in diameter, with a double-stranded RNA genome. Here, we report the first complete sequence and organization of the 8.7-kb viral dsRNA genome. Two ORFs coding for a putative coat protein and RNA-dependent RNA polymerase (RdRp) were predicted. *In silico* analysis revealed that the translated ORF2 contains the conserved domains characteristic of an RdRp protein (pfam02123:RdRP 4), which is a family that includes RdRps from members of the genera *Luteovirus*, *Totivirus* and *Rotavirus*. Evolutionary analysis with amino acid sequences with the RdRps from members of the family *Totiviridae* and some dsRNA viruses showed that PMeV RdRp did not root itself in any genus.

Papaya sticky disease (named ‘*meleira*’ in Portuguese) was first observed in Brazil at the beginning of the 1980s [1] and, more recently, in Mexico [2]. It is now found in the most important Brazilian papaya-production areas, and it has become one of the main limiting factors for the papaya industry. The disease is characterized by intense latex exudation from the fruit surface, which becomes dark as it oxidizes [1]. In some cases, the pulp may be affected, assuming a spongy aspect and having changed its flavor. Necrotic symptoms can appear at the edges of new leaves, resulting in latex exudation [1, 3]. As a consequence, the fruit becomes stained and sticky, which makes it difficult to sell.

The causal agent of papaya sticky disease has been identified as an isometric virus particle, approximately 50 nm in diameter. The virus has been called papaya meleira virus (=papaya sticky disease virus) and has a double-stranded RNA (dsRNA) genome, which was initially determined by its migration in acrylamide and agarose gels to be approximately 10 kb (6×10^6 Da) [1] and 12 kb [4]. Since the disease has a long asymptomatic incubation period, diagnostic methods based on the detection of viral dsRNA have been developed [5–8]. Diversity analysis based on RT-PCR amplification of a 560-bp fragment from 31 Brazilian isolates have suggested similarities between PMeV and dsRNA viruses of the family *Totiviridae* [9].

Despite being considered one of the main viral diseases of papaya in Brazil, with potential to be spread to other countries, knowledge about the sequence and genomic organization of PMeV is still limited. Here, we report for the first time the complete sequence and organization of a dsRNA associated with papaya meleira virus infection.

Viral dsRNA was isolated from infected fruits (collected in Baraúnas, RN, Brazil) as described previously [9]. Sequencing was carried out at Macrogen Inc. (Korea) using the 454 GS FLX system. From 1,947 reads (424,287 bases; 166 singletons) a contig with 8,924 bases was identified

Electronic supplementary material The online version of this article (doi:10.1007/s00705-015-2605-x) contains supplementary material, which is available to authorized users.

✉ Francisco J. L. Aragão
francisco.aragao@embrapa.br

[1] Embrapa Mandioca e Fruticultura Rua, Embrapa s/n, CP 007, Cruz das Almas, BA 44380-000, Brazil

[2] Embrapa Recursos Genéticos e Biotecnologia, PqEB W5 Norte, Brasília, DF 70770-900, Brazil

(99.41 % Q40). Based on this contig, primers were designed to re-amplify fragments throughout the primary sequence (Fig. 1, Supplementary Table 1). In addition, to confirm the 5′ and 3′ ends of the viral genome, specific primers were designed from the internal sequenced fragment and amplified by using the 5′ and 3′-RACER System (Invitrogen, Carlsbad, CA, USA), using reverse specific nested primers (5′RACER 1 and 3′RACER 2, respectively, Fig. 1, Supplementary Table 1) in combination with the GeneRacer Nested Primer (Invitrogen). PCR products were cloned into pGEM-T Easy Vector (Promega) and sequenced at Macrogen Inc. (Korea).

A sequence of 8,768 bp corresponding to the genome of PMeV was assembled and entered into the NCBI GenBank database with the accession number KT013296. The PMeV dsRNA was estimated by electrophoresis to be 10-12 kb in length [1, 4]. However, dsDNA was used as the reference in the determination of dsRNA size. The RNA double helix differs from the DNA double helix by a lower rise per base pair, resulting in a higher linear charge density. This causes the apparent mass of an RNA molecule to be higher than that of a DNA molecule with the same number of base pairs in agarose gels with a concentration lower than 1.2 % [10]. This may explain the difference in size of the PMeV dsRNA observed in the electrophoresis analysis and the genome sequence. Comparison of the 8.7-kb viral dsRNA sequence with a 629-bp fragment (within the *pol* gene) that was amplified from purified isometric viral particles [6] revealed 76 % identity (E-value = $6e^{-128}$). Comparison of the deduced amino acid sequence revealed 81% identity (E-value = $5e^{-124}$). Nevertheless, considering that this *pol* region was diverse, ranging from 86 % to 99 % in a comparison of 31 PMeV Brazilian isolates [9] with the fragment published by Araujo et al. [6], the sequence

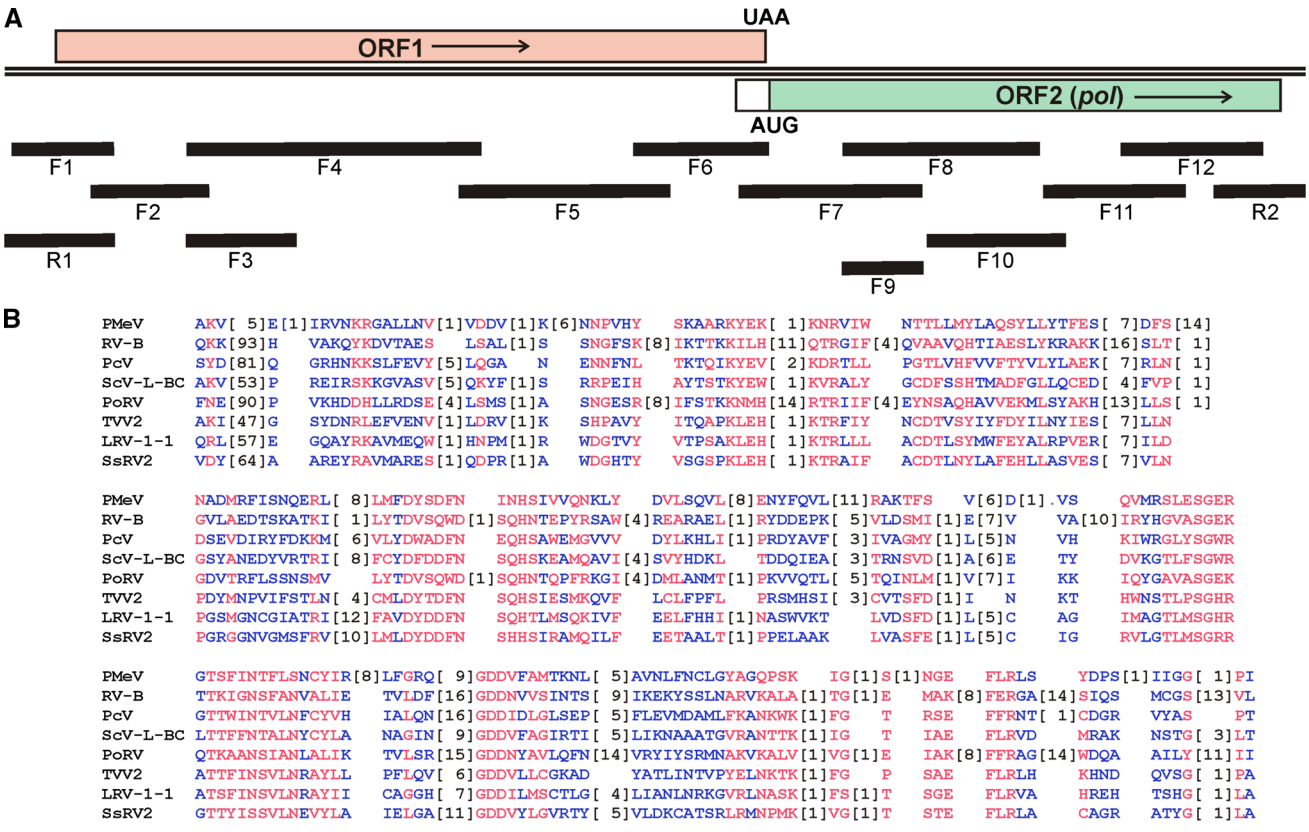
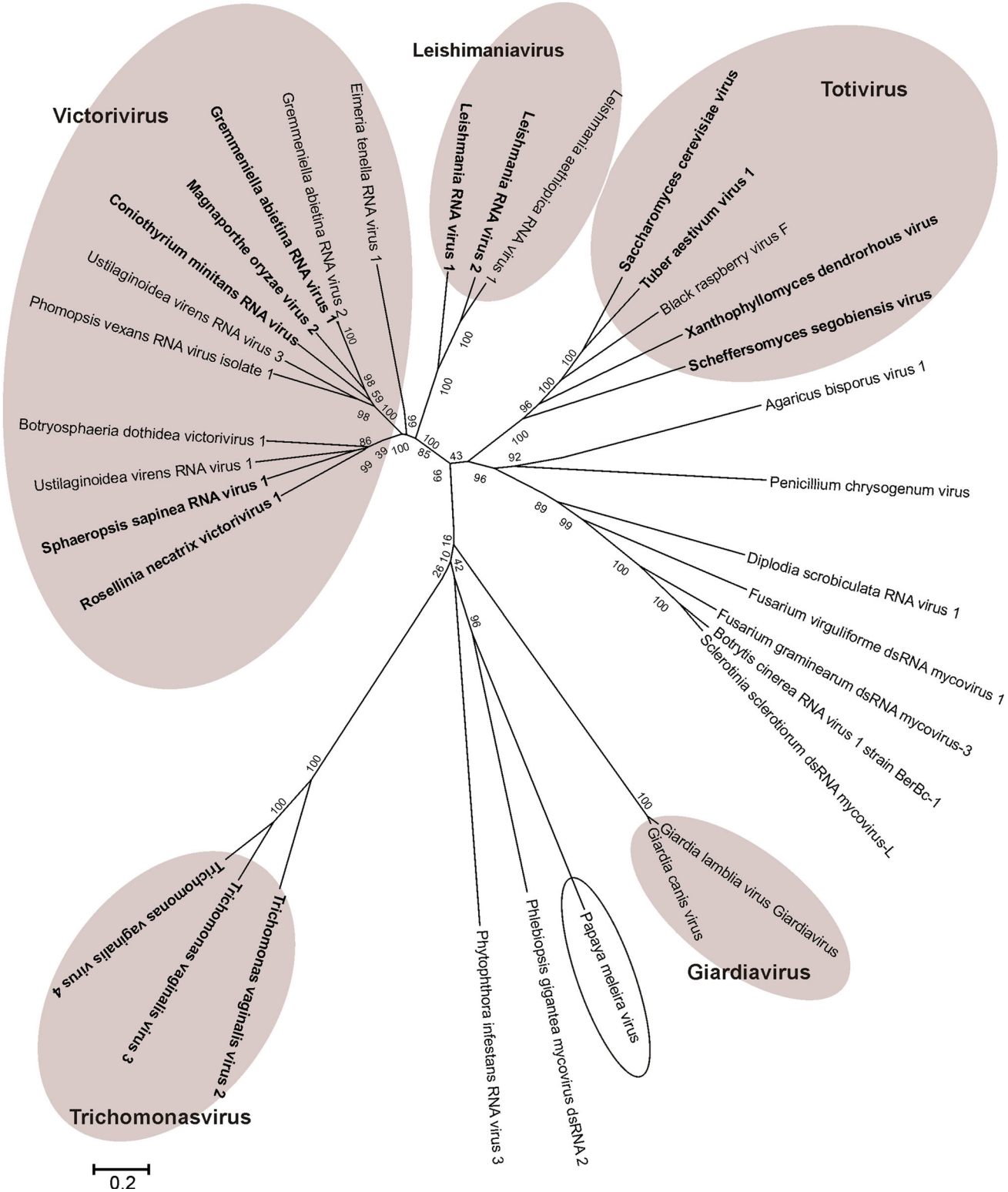


Fig. 1 Genome organization of papaya meleira virus. (A) The dsRNA genomic plus strand includes three ORFs. ORF1 is a hypothetical protein, ORF2 encodes the coat protein (Gag) and ORF3 encodes a RNA-dependent RNA polymerase (RdRp). It is possible that there is an overlap between the ORF and a fusion Gal-Pol, generated by a translational -1 ribosomal frameshift produced by a 7-bp slippery site and an essential pseudoknot structure. Black bars represent sequences (F1 to F12) that were re-amplified using specific primers and RACER (R1 and R2). (B) Alignment of the putative RdRP of PMeV with the RdRP4 (pfam02123) domain of RdRPs. The numbers in brackets indicate the number of amino acid residues separating individual motifs. The red-to-blue scale represents the degree of conservation, with red designating highly conserved residues. RV-B, Rotavirus B (gi 548839, *Reoviridae*, *Rotavirus*); PcV, *Penicillium chrysogenum* virus (gi 81951345, *Chrysoviridae*, *Chrysovirus*); ScV-L-BC, *Saccharomyces cerevisiae* virus L-BC (gi 313104168, *Totiviridae*; *Totivirus*); PoRV, *porcine rotavirus* (strain YM) (gi 75568552, *Reoviridae*, *Rotavirus*); TVV2, *Trichomonas vaginalis* virus 2 (gi 81974669, *Totiviridae*, *Trichomonasvirus*); LRV-1-1, *Leishmania* RNA virus 1–1 (gi 81925554, *Totiviridae*, *Leishmaniavirus*); SsRV2, *Sphaeropsis sapinea* RNA virus 2 (gi 81982721, *Totiviridae*, *Victorivirus*)


Leishimaniavirus
Totivirus
Victorivirus
Eimeria tenella RNA virus 1
Gremmeniella abietina RNA virus 2
Gremmeniella abietina RNA virus 1
Magnaporthe oryzae virus 2
Coniothyrium minitans RNA virus
Ustilaginoidea virens RNA virus 3
Phomopsis vexans RNA virus isolate 1
Botryosphaeria dothidea victorivirus 1
Ustilaginoidea virens RNA virus 1
Sphaeropsis sapinea RNA virus 1
Rosellinia necatrix victorivirus 1
Leishmania RNA virus 1
Leishmania RNA virus 2
Leishmania aethiopica RNA virus 1
Saccharomyces cerevisiae virus
Tuber aestivum virus 1
Black raspberry virus F
Xanthophyllomyces dendrorhous virus
Scheffersomyces segobiensis virus
Agaricus bisporus virus 1
Penicillium chrysogenum virus
Diplodia scrobiculata RNA virus 1
Fusarium virguliforme dsRNA mycovirus 1
Fusarium graminearum dsRNA mycovirus-3
Botrytis cinerea RNA virus 1 strain BerBc-1
Sclerotinia sclerotiorum dsRNA mycovirus-L
Giardia lamblia virus Giardiavirus
Giardia canis virus
Giardiavirus
Papaya meleira virus
Phlebiopsis gigantea mycovirus dsRNA 2
Phytophthora infestans RNA virus 3
Trichomonas vaginalis virus 4
Trichomonas vaginalis virus 3
Trichomonas vaginalis virus 2
Trichomonasvirus
0.2

◀**Fig. 2** Phylogenetic relationship between papaya meleira virus and members of the family *Totiviridae*. The unrooted tree was generated by the neighbor-joining method using the amino acid sequences of RdRPs of members of the family *Totiviridae*. Numbers indicate the percentage of bootstrap support from 1,000 replicates. Genera are clustered and labeled with colored dots. Classified and ICTV-approved species are denoted in bold. GenBank accession numbers are as follows: *Aspergillus foetidus* slow virus, CCD33024.1; *Beauveria bassiana* victorivirus, YP_009032633.1; black raspberry virus F, YP_001497151.1; *Botryotinia fuckeliana* totivirus 1, YP_001109580.1; *Botryosphaeria dothidea* victorivirus 1, YP_009072433.1; *Botrytis cinerea* RNA virus 1 strain BerBc-1, NC 026139; *Coniothyrium minitans* RNA virus, YP_392467.1; *Diplodia scrobiculata* RNA virus 1, NC 013699; *Eimeria brunetti* RNA virus 1, NP_108651.1; *Eimeria tenella* RNA virus 1, YP_009115500.1; *Fusarium graminearum* dsRNA mycovirus 3, NC 013469; *Fusarium virguliforme* dsRNA mycovirus 1, JN671444; *Giardia canis* virus, ABB36743.1; *Giardia lamblia* virus, NP_620070.1; *Gremmeniella abietina* RNA virus L1, AAK11656.1; *Gremmeniella abietina* RNA virus 2, YP_044807.1; *Helminthosporium victoriae* virus, NP_619670.2; *Helicobasidium mompa* totivirus, NP_898833.1; *Leishmania aethiopica* RNA virus 1, YP_009030005.1; *Leishmania* RNA virus 1-1, NP_041191.1; *Leishmania* RNA virus 2-1, NP_043465.1; *Magnaporthe oryzae* virus 2, YP_122352.1; *Rosellinia necatrix* victorivirus 1 RNA (BAM36400.1); *Penicillium chrysogenum* virus segment 1, YP_392482.1; *Phomopsis vexans* RNA virus isolate 1, YP_009115492.1; *Phlebiopsis gigantea* mycovirus dsRNA 2, CAJ34334.2; *Phytophthora infestans* RNA virus 3, JN603241; *Saccharomyces cerevisiae* virus L-A, NP_620495.1; *Scheffersomyces segobiensis* virus, AGG68771.1; *Sclerotinia sclerotiorum* dsRNA mycovirus L, YP 006331064.1; *Sphaeropsis sapinea* RNA virus 1, NP_047558.1; *Tolypocladium cylindrosporum* virus 2, YP_004089630.1; *Trichomonas vaginalis* virus 2-1, AAF29445.1; *Trichomonas vaginalis* virus 3-1, NP_659390.1; *Trichomonas vaginalis* virus 4, AED99798.1; *Tuber aestivum* virus 1, ADQ54106.1; *Ustilaginoidea virens* RNA virus 3, YP_009004156.1; *Ustilaginoidea virens* RNA virus 1, AHH25153.1; *Xanthophyllomyces dendrorhous* virus, AFH09414.1

presented here might nevertheless represent the genome of the causal agent of sticky disease. This is supported by the observation that the 8.7-kb dsRNA was only found in plants with sticky disease symptoms.

Two ORFs were predicted. ORF1 encodes a putative structural protein with a predicted amino acid sequence that is 26 % identical to that of the protein encoded by ORF1 of *Phlebiopsis gigantea* mycovirus (dsRNA 2), which is provisionally considered a member of the family *Totiviridae*. In addition, ORF1 showed sequence similarity to a hypothetical protein of the unclassified viruses *Sclerotinia sclerotiorum* dsRNA mycovirus-L (22 %) and *Botrytis cinerea* RNA virus 1, and also to a putative structural/gag protein of the unclassified viruses *Fusarium virguliforme* dsRNA mycovirus 2 (22 %) and *Fusarium virguliforme* dsRNA mycovirus 1 (21 %). However, a Pfam database (EMBL-EBI) search revealed no similarities to protein domain families.

ORF2 was predicted to encode a protein containing the conserved domains characteristic of members of the RdRp protein family pfam02123:RdRP_4), which includes RNA-dependent RNA polymerases (RdRps) from members of the genera *Luteovirus*, *Totivirus* and *Rotavirus* (Fig. 1). BLASTn analysis revealed no significant similarity of the PMeV genome to those of other viruses. BLASTx analysis of ORF2 (RdRp) sequences revealed the highest similarity to the RdRp of *Diplodia scrobiculata* RNA virus 1 (33 %), grapevine-associated totivirus 2 (33 %), *Phlebiopsis gigantea* mycovirus dsRNA 2 (32 %), *Fusarium graminearum* dsRNA mycovirus-3 (31 %), *Botrytis cinerea* RNA virus 1 (31 %), *Sclerotinia sclerotiorum* dsRNA mycovirus-L (31 %), *Phytophthora infestans* RNA virus 3 (31 %), *Cytospora sacchari* RNA virus (31 %) and *Fusarium virguliforme* dsRNA mycovirus 1 (29 %). Surprisingly, no significant similarity was found to the partial sequence of a RdRp-coding gene from a virus named papaya meleira virus isolate Q25, which was found in Mexico (GenBank accession number KF781635) [8].

Another ORF of 252 bp was recognized at the 5′ end, coding for an 83-amino-acid hypothetical protein.

The length of the 5′-UTR of PMeV was determined in a reliable manner, using the RACER procedure. An analysis of the 3′-UTR region using MFOLD [11] suggested a secondary structure that could act as a *cis* signal for replication ($\Delta G = -3.80$ kcal/mol) (data not shown).

Although an ATG (from *pol*) was found positioned after the *gag* stop codon, the fact that *gag* is located in the frame +3 and *pol* is located in the frame +1 creates the possibility that both ORFs could be translated as a fusion protein if ribosomal -1 frameshifting were to occur within the region of overlap. Similar genomic arrangements have been described for several members of the family *Totiviridae* [12]. Ribosomal -1 frameshifting is often associated with a shifty (slippery) heptamer motif, XXX.YYY.Z (where X = A, C, G or U; Y = A or U; Z = is A, C or U), followed by a 4- to11-nt spacer and an RNA pseudoknot (or stem-loop) [13]. In fact, these structures are present within the *gag* gene (sequence AAAAAUC at position 5,100, followed by a stem-loop). However, identifying frameshifts is difficult because of their diverse nature, and current computational models could make false positive predictions. Consequently, reference protein sequences will be needed to confirm the PeMV translational mechanism.

Evolutionary analysis with amino acid sequences of the RdRps from members of the family *Totiviridae* was carried out in MEGA6 [14]. A phylogenetic tree was constructed using the neighbor-joining algorithm with a bootstrap test (1,000 replicates). Evolutionary distances were computed using the Poisson correction method. All positions containing gaps and missing data were eliminated. The rootless tree showed five main branches, corresponding to the genera *Giardiavirus*, *Trichomonasvirus*, *Totivirus*, *Victorivirus* and *Leishmaniavirus*. Although PMeV showed

similarities to mycoviruses such as the *Phlebiopsis gigantea* mycovirus, it could not be included in any of these clusters (Fig. 2). Further analysis will be required to determine whether PMeV is a novel member of the *Totiviridae* or a member of a novel family of plant viruses similar to mycoviruses.

Currently, papaya sticky disease or *meleira* is one of the most important constraints on the papaya industry in Brazil. The sequence of papaya meleira virus will provide support for (1) a conclusive taxonomic classification, (2) studies of virus genetic diversity and evolution, (3) development of more-appropriate diagnostic methods, (4) generation of mutant mild strains that could be used in a cross-protection strategy, and (5) generation of strategies for the production of resistant varieties by both conventional methods and genetic engineering.

Acknowledgments This work was supported by Grant 475899/2013-0 from the CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico, Brazil). C. B. Daltro was supported by a fellowship from FAPESB (Fundação de Amparo à Pesquisa do Estado da Bahia, Brazil).

References

1. Kitajima E, Rodrigues C, Silveira J, Alves F, Ventura J, Aragão FJL, Oliveira C (1993) Associantion of isometric vírus-like particles, restricted to laticifers, with meleira ("sticky disease") of papaya (*Carica papaya*). Fitopatol Bras 8:118–122
2. Perez-Brito D, Tapia-Tussell R, Cortes-Velazquez A, Quijano-Ramayo A (2012) First report of Papaya meleira virus (PMeV) in Mexico. Afr J Biotechnol 11:13564–13570
3. Abreu PMV, Antunes TFS, Magaña-Álvarez A, Pérez-Brito D, Tapia-Tussell R, Ventura JA, Fernandes AAR, Fernandes PMB (2015) A current overview of the Papaya meleira virus, an unusual plant virus. Viruses 7:1853–1870
4. Maciel-Zambolim E, Kunieda-Alonso S, Matsuoka K, de Carvalho M, Zerbini F (2003) Purification and some properties of Papaya meleira virus, a novel virus infecting papayas in Brazil. Plant Pathol 52:389–394
5. Rodrigues SPGOP, Andrade JS, Ventura JA, Fernandes PMB (2005) Simplified molecular method for the diagnosis of Papaya meleira virus in papaya latex and tissues. Summa Phytopathol 31:281–283
6. Araújo MMM, Tavares ET, Silva FR, Marinho VLA, Júnior MTS (2007) Molecular detection of Papaya meleira virus in the latex of *Carica papaya* by RT-PCR. J. Virol. Methods. 146:305–310
7. Abreu P, Piccin JG, Rodrigues SP, Buss DS, Ventura JA, Fernandes P (2012) Molecular diagnosis of Papaya meleira virus (PMeV) from leaf samples of *Carica papaya* L. using conventional and real-time RT-PCR. J Virol Methods 180:11–17
8. Zamudio-Moreno E, Ramirez-Prado JH, Moreno-Valenzuela AO, Lopez-Ochoa LA (2015) Early diagnosis of a Mexican variant of Papaya meleira virus (PMeV-Mx) by RT-PCR. Genet Mol Res 14:1145–1154
9. Daltro CB, Abreu EFM, Aragão FJL, Andrade EC (2014) Genetic diversity studies of Papaya meleira virus. Trop Plant Pathol 39:104–108
10. Livshits MA, Amosova OA, Lyubchenko YL (1990) Flexibility difference between double-stranded RNA and DNA as revealed by gel electrophoresis. J Biomol Struct Dyn 7:1237–1249
11. Zuker M (2003) Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res 31(13):3406–3415
12. King AMQ, Adams MJ, Carstens EB, Lefkowitz EJ (2012) Virus Taxonomy: classification and nomenclature of viruses. International Committee on Taxonomy of Viruses, 9th edn. Elsevier Academic Press, Amsterdam
13. Moon S, Byun Y, Kim HJ, Jeong S, Han K (2004) Predicting genes expressed via -1 and +1 frameshifts. Nucleic Acids Research 31:4884–4892
14. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol 30:2725–2729