Archives of Virology
https://doi.org/10.1007/s00705-020-04592-9

ANNOTATED SEQUENCE RECORD

Identification and complete genomic sequence of a novel sadwavirus discovered in pineapple (*Ananas comosus*)

Adriana Larrea-Sarmiento[1] · Alejandro Olmedo-Velarde[1] · James C. Green[1] · Maher Al Rwahnih[2] · Xupeng Wang[1] · Yun-He Li[3] · Weihuai Wu[4] · Jingxin Zhang[5] · Tracie K. Matsumoto[6] · Jon Y. Suzuki[6] · Marisa M. Wall[6] · Wayne Borth[1] · Michael J. Melzer[1] · John S. Hu[1]

Received: 19 December 2019 / Accepted: 15 February 2020

Abstract

The complete genomic sequence of a putative novel member of the family *Secoviridae* was determined by high-throughput sequencing of a pineapple accession obtained from the National Plant Germplasm Repository in Hilo, Hawaii. The predicted genome of the putative virus was composed of two RNA molecules of 6,128 and 4,161 nucleotides in length, excluding the poly-A tails. Each genome segment contained one large open reading frame (ORF) that shares homology and phylogenetic identity with members of the family *Secoviridae*. The presence of this new virus in pineapple was confirmed using RT-PCR and Sanger sequencing from six samples collected in Oahu, Hawaii. The name "pineapple secovirus A" (PSVA) is proposed for this putative new sadwavirus.

Members of the family *Secoviridae* are non-enveloped viruses containing linear positive-sense single-stranded RNA [(+)-ssRNA] genomes that are monopartite or bipartite. Each genome segment encodes a large polyprotein, is covalently linked to a viral protein (VPg) at its 5'- end, and has a poly-A tract at its 3'- end. Secovirids form isometric viral particles and are classified within eight genera: *Nepovirus*, *Comovirus*, *Fabavirus*, *Sadwavirus*, *Cheravirus*, *Torradovirus*, *Sequivirus*, and *Waikavirus* [1]. Recently, some unassigned secoviruses were classified within the genus *Sadwavirus.* It has been proposed that this genus should be further divided into three subgenera: "*Stramovirus*", "*Satsumavirus*", and "*Cholivirus*" [2]. An additional ninth genus, "*Stralarivirus*", was recently proposed based on sequencing of several isolates of strawberry latent ringspot virus and phylogenetic analysis [3]. For bipartite secovirids, replication-associated proteins are encoded in the RNA1-coded polyprotein, while the RNA2-encoded polyprotein may contain a movement protein and up to three coat proteins (CP). Some members may encode additional proteins, including a glutamic protease and proteins of unknown function. An additional open reading frame coding for a protein with unknown function is found in RNA2 of torradoviruses [1, 2]. The host range of members of the family *Secoviridae* varies among the genera, and these viruses are experimentally sap-transmissible. Natural transmission of most viruses is mediated by insects and nematodes, although sequiviruses require a helper virus, and others do not have a known biological vector [1].

Originating in South America, pineapple (*Ananas comosus*) is a widely distributed and consumed perennial monocot, ranking as the eleventh-most-cultivated fruit worldwide

Communicated by Jesús Navas-Castillo.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00705-020-04592-9) contains supplementary material, which is available to authorized users.

✉ John S. Hu
johnhu@hawaii.edu

[1] Department of Plant and Environmental Protection Sciences, University of Hawaii, Honolulu, HI, USA

[2] Department of Plant Pathology, University of California, Davis, CA, USA

[3] South Subtropical Crop Research Institute, Chinese Academy of Tropical Agricultural Science, Zhanjiang, China

[4] Environment and Plant Protection Institute, Chinese Academy of Tropical Agricultural Sciences, Haikou, China

[5] Key Laboratory of New Technique for Plant Protection in Guangdong, Institute of Plant Protection, Guangdong Academy of Agricultural Sciences, Guangzhou, China

[6] United States Department of Agriculture, Agricultural Research Service, Pacific Basin Agricultural Research Center, Hilo, HI, USA

Published online: 29 March 2020

[4]. Pineapple was introduced to Hawaii in 1813 and became an economically significant crop [5]. Viral infections of pineapple by various members of the genus *Ampelovirus* have been associated with mealybug wilt of pineapple (MWP), which affects different cultivars of pineapple globally. Although pineapple mealybug wilt-associated virus 1 (PMWaV-1), PMWaV-2, and PMWaV-3 have been associated with MWP, the etiology of the disease is not well understood [6, 7]. In Hawaii, symptom expression of MWP is associated with the presence of both PMWaV-2 and concurrent feeding by insects of the pineapple mealybug complex [7, 8]. However, this association differs from that occurring in Australia, where MWP symptoms may be expressed in plants not infected with PMWaV-2 but infected by other PMWaVs [9]. Additional factors, including the presence of different strains of PMWaV-2, interactions between members of PMWaV species and badnaviruses that have also been found infecting pineapple, and other uncharacterized viruses, may also be involved in the etiology of MWP [7].

In 2018, leaf tissues from several asymptomatic pineapple accessions were collected from the United States Department of Agriculture, Agricultural Research Service (USDA-ARS), National Plant Germplasm Repository (NPGR) at the Daniel K. Inouye U.S. Pacific Basin Agricultural Research Center (PBARC) in Hilo, Hawaii, USA. Total RNA was extracted from about 0.1 g of basal portions of pineapple leaf tissue using an RNeasy® Plant Mini Kit (QIAGEN) or a Spectrum® Plant Total RNA Kit (Sigma Aldrich) following the manufacturers' protocols. Pineapple accessions HANA 158, 160, and 187 were selected to further characterize the genomes of PMWaV-1, -2, -3, and -4 based on the presence of PMWaVs screened by reverse transcription (RT)-PCR [10]. Total RNA was subjected to ribosomal RNA depletion, and high-throughput sequencing (HTS) was performed on an Illumina® NexSeq 500 platform at the Foundation Plant Services, University of California at Davis.

Approximately 38M, 23M, and 36M Illumina single-end reads from HANA 158, 160, and 187, respectively, were adapter trimmed and subsequently assembled, and contigs were annotated as described by Green et al. [10]. In addition to the complete genome sequences of PMWaV-1 (13,071 nt), -2 (16,259 nt), -3 (13,298 nt), and -4 (12,932 nt), two contigs (4,161 and 6,128 nt) retrieved from HANA 187 were found, suggesting the presence of a putative new member of the family *Secoviridae* infecting pineapple (Table S1). A search for potential protein-encoding segments in the genome was conducted using ORF Finder, and translation products were verified using BLASTp. The sequences of the 5'- and 3'-termini of RNA1 and RNA2 of the new virus were determined using rapid amplification of cDNA ends (RACE) reactions. To determine the 5'-terminal sequence, either double-stranded (ds)RNA isolated by the method of Morris and Dodds [11] or total RNA was used as a template for polyadenylation. The 3'-terminal sequence was determined by sequencing the products of RT-PCR using an oligo-dT primer to target the poly-A tails of RNA1 and RNA2 (Table S2).

Sequence lengths obtained for the two genomic RNAs of the new virus, including the untranslated regions (UTR) and excluding poly-A tails, were 6,128 nt for RNA1 (MN809923) and 4,161 nt for RNA2 (MN809924). BLASTx analysis of the two contigs showed that RNA1 shared 35% identity with dioscorea mosaic-associated virus of the proposed subgenus "*Cholivirus*" and strawberry mottle virus of the proposed subgenus "*Stramovirus*", and RNA2 shared 26% identity with dioscorea mosaic associated virus and chocolate lily virus A of the proposed subgenus "*Cholivirus*". These related viruses are currently classified as sadwaviruses within the family *Secoviridae* [2]. Based on these findings, we propose the name "pineapple secovirus A" (PSVA) for this putative new sadwavirus. In addition, Australia in 2002, two isometric viruses infecting pineapple and resembling strawberry mottle virus were partially characterized [12], but no sequence information on them is publicly available.

RNA1 contains a large ORF encoding a 1,865-amino-acid (aa) polyprotein with a predicted molecular weight of 210 kDa and carries domains associated with replication. Cleavage of the polyprotein is predicted to occur at specific dipeptides to produce five putative proteins: a protease cofactor (Pro-C), a helicase (Hel), a viral-genome-linked protein (VPg), a protease (Pro), and an RNA-dependent RNA polymerase (RdRp) (Fig. 1). RNA2 contains one ORF encoding a 1,119-aa polyprotein with a predicted molecular weight of 124 kDa. Analysis of RNA2 using the NCBI Conserved Domain Search Tool (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) revealed no conserved domains. In contrast, RNA2-encoded polyproteins in other members of the family *Secoviridae* commonly contain CP and MP domains [1]. Conserved motifs in the putative MP and CP were also found by multiple alignment of RNA2-encoded polyprotein sequences of PVSA and other stramoviruses and choliviruses. The most probable cleavage sites in the RNA1- and RNA2-encoded polyproteins, based on alignments with other sadwaviruses, are shown in Fig. 1. Interestingly, the last 185 nt in the 3'-UTRs of both RNA1 and RNA2 shared 88% identity. This feature has been observed in other secoviruses as well [1, 13].

Two sets of primers designed based on the HTS-derived sequences were used in tandem to detect the presence of PSVA in pineapple. PSVA-RNA1-F/R and PSVA-RNA2-F/R (Table S2) amplified fragments of 566 nt and 395 nt, located in the RdRp-coding region of RNA1 and in the putative CP-coding region of RNA2, respectively. We collected 13 MWP-symptomatic and 14 asymptomatic pineapple plants from Oahu. Total RNA was extracted as described above for use in an RT-PCR assay using cDNA and random

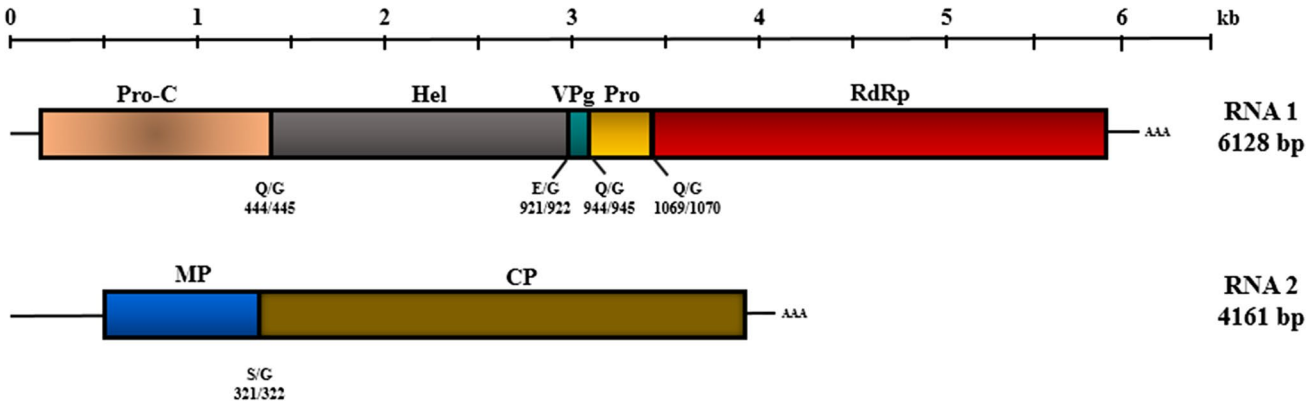

Fig. 1 The genomic organization of RNA1 and RNA2 of pineapple secovirus A. Vertical lines within the boxes indicate the putative cleavage sites with their respective amino acid sequences, based on conserved protease cleavage sites and compared with other sadwaviruses. The scale at the top is in kilobases (kb). Pro-C, protease factor; Hel, helicase; VPg, viral-linked protein; Pro, protease; RdRp, RNA-dependent RNA polymerase; MP, movement protein; CP, coat protein. "AAA" at the 3' end of each RNA segment represents the poly-A tail

primers to detect single or mixed infections with PSVA and/or PMWaV-2. RT-PCR reactions were used to detect PMWaV-2 in pineapple samples as reported previously by Sether et al. [13]. Six of the 13 symptomatic samples were infected with PSVA, while all 13 of the MWP-symptomatic plants were infected by PMWaV-2. PSVA and PMWaV-2 were not detected in the asymptomatic plants. Amplicons corresponding to fragments of RNA1 (MN833715) and RNA2 (MN833716) from PSVA isolate symptomatic-1 (S1) were sequenced directly. Sequence analysis using Geneious 2019.2.3 showed that PSVA-S1 shared 94.2% identity at the nucleotide level and 100% at the amino acid level for RNA1, and 89.6% identity at the nucleotide level and 95.4% at the amino acid level for RNA2 with PVSA from HANA187 derived from HTS. These preliminary results confirmed the presence of PSVA in Hawaiian pineapple fields.

Amino acid sequences of the conserved domains within the protease-polymerase (Pro-Pol) region were aligned using MEGA 7.0.1 [14] with those of orthologous members of the family *Secoviridae* to generate a phylogenetic tree (Fig. 2). A maximum-likelihood (ML) algorithm based on the Le-Gascuel model [15] was used with 1,000 bootstrap repetitions as branch support. PSVA clustered together with other sadwaviruses in the proposed subgenus "*Cholivirus*", namely, dioscorea mosaic associated virus and chocolate lily virus A. Pairwise alignment of the Pro-pol region of PVSA and dioscorea mosaic associated virus showed 77% identity. Considering the ICTV species demarcation criteria for secovirids (<80% identity in the Pro-Pol region or <75% identity in the CP) [16], PVSA should be considered a new sadwavirus subclassified within the proposed subgenus "*Cholivirus*". These results support the placement of PSVA as a putative member of the genus *Sadwavirus*, family *Secoviridae*. Further research is needed to identify the biological vector of PSVA, investigate whether PSVA is related to the two isometric viruses infecting pineapple in Australia, and determine if this putative new pineapple virus is involved in the etiology of MWP.

Funding This research was supported in part by grants from the United States Department of Agriculture National Institute of Food and Agriculture, Hatch HAW09025-H (1001478), and the United State Department of Agriculture -Agricultural Research Service (58-5320-4-012).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Human and animal rights This study did not include experiments with human or animal participants performed by any of the authors.

References

1. Thompson JR, Dasgupta I, Fuchs M et al (2017) ICTV virus taxonomy profile: *Secoviridae*. J Gen Virol 98(4):529–531
2. Sanfaçon H, Dasgupta I, Fuchs M et al (2020) Proposed revision of the family *Secoviridae* taxonomy to create three subgenera, "*Satsumavirus*", "*Stramovirus*" and "*Cholivirus*", in the genus *Sadwavirus*. Arch Virol 165:527–533
3. Dullemans AM, Botermans M, de Kock MJD et al (2020) Creation of a new genus in the family *Secoviridae* substantiated by sequence variation of newly identified strawberry latent ringspot virus isolates. Arch Virol 165(1):21–32
4. FAO (2019) FAOSTAT. Food and Agriculture Organization of the United Nations. Pineapple Crop Information. http://www.fao.org/land-water/databases-and-software/crop-information/pineapple/en/. Accessed 26 Mar 2020

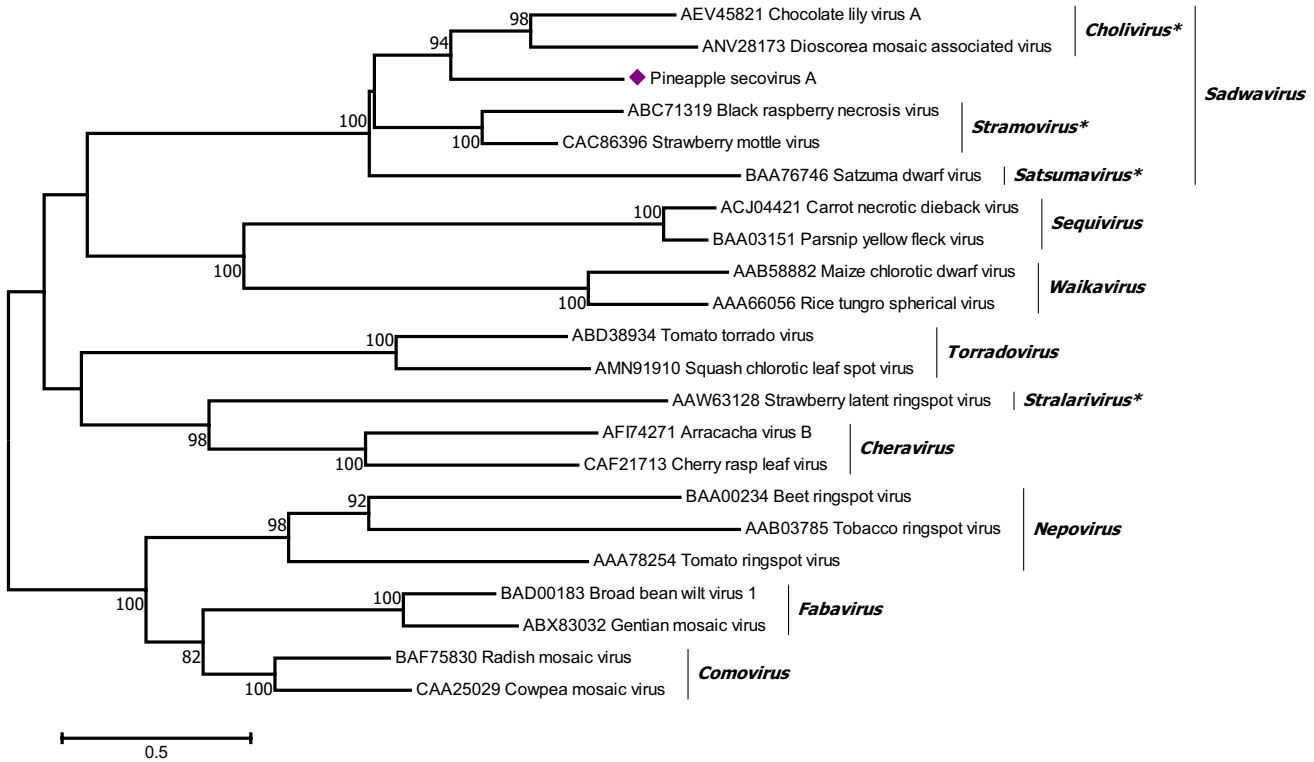

Fig. 2 Phylogenetic relationship of pineapple secovirus A (PSVA) with selected members of the family *Secoviridae* based on an alignment of their Pro-Pol amino acid sequences. The tree was constructed using the maximum-likelihood (ML) method based on the Le-Gascuel model, with 1,000 bootstrap pseudo-replicates as percentage values for branch support. GenBank sequence accession numbers are provided before each virus name. The scale represents the number of substitutions per unit branch length. Asterisks represent the three proposed subgenera ("*Stramovirus*", "*Satsumavirus*", and "*Cholivirus*") within the genus *Sadwavirus* and the proposed genus "*Stralarivirus",* which have yet to be ratified by the ICTV. The diamond indicates PSVA

5. Bartholomew DP, Hawkins RA, Lopez JA (2012) Hawaii pineapple: the rise and fall of an industry. American Society for Horticultural Science 47(10):1390–1398
6. Hernandez-Rodriguez L, Ramos-Gonzalez PL, Garcia-Garcia G et al (2014) Geographic distribution of mealybug wilt disease of pineapple and genetic diversity of viruses infecting pineapple in Cuba. Crop Prot 65:43–50
7. Dey K, Green J, Melzer MJ et al (2018) Mealybug wilt of pineapple and associated viruses. Horticulturae 4(4):52
8. Sether DM, Hu JS (2002) Closterovirus infection and mealybug exposure are both necessary factors for the development of mealybug wilt of pineapple disease. Phytopathology 92:928–935
9. Gambley CF, Geering ADW, Steele V et al (2008) Identification of viral and non-viral reverse transcribing elements in pineapple (*Ananas comosus*), including members of two new badnavirus species. Arch Virol 153(8):1599–1604
10. Green JC, Al Rwahnih M, Olmedo-Velarde A et al (2020) Further genomic characterization of pineapple mealybug wilt-associated viruses using high-throughput sequencing. Trop Plant Pathol 45:64–72
11. Morris TJ, Dodds JA (1979) Isolation and analysis of double-stranded RNA from virus-infected plant and fungal tissue. Phytopathology 69:854–858
12. Halgren A, Tzanetakis IE, Martin RR (2007) Identification, characterization, and detection of black raspberry necrosis virus. Phytopathology 97:44–50
13. Thompson HL, Botella J, Fairbaim D et al (2005) Characterization and detection of isometric viruses in pineapple. In: Proceedings of the 15th Biennial Australasian Plant Pathology Conference' (Australasian Plant Pathology Society: Geelong, Vic.), p 318
14. Sether DM, Melzer MJ, Busto J et al (2005) Diversity and mealybug transmissibility of ampeloviruses in pineapple. Plant Dis 89(5):450–456
15. Kumar S, Stecher G, Tamura K (2016) MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol 33(7):1870–1874
16. Le SQ, Gascuel O (2008) An Improved General Amino Acid Replacement Matrix. Mol Biol Evol 25(7):1307–1320
17. King AMQ, Adams MJ, Lefkowitz EJ (2012) Virus taxonomy: classification and nomenclature of viruses: ninth report of the international committee on taxonomy of viruses. Elsevier, New York