Genome Sequence of Pineapple Secovirus B, a Second Sadwavirus Reported Infecting *Ananas Comosus*

Adriana Larrea-Sarmiento
University of Hawai'i at Manoa https://orcid.org/0000-0003-3355-1522
Andrew D.W. Geering
Queensland Alliance for Agriculture and Food Innovation
Alejandro Olmedo-Velarde
University of Hawai'i at Mānoa: University of Hawai'i at Manoa
Xupeng Wang
University of Hawai'i at Mānoa: University of Hawai'i at Manoa
Wayne Borth
University of Hawai'i at Mānoa: University of Hawai'i at Manoa
Tracie K Matsumoto
USDA-ARS Daniel K Inouye US Pacific Basin Agricultural Research Center
Jon Y Suzuki
USDA-ARS Daniel K Inouye US Pacific Basin Agricultural Research Center
Marisa M. Wall
USDA-ARS Daniel K Inouye US Pacific Basin Agricultural Research Center
Michael J Melzer
University of Hawai'i at Mānoa: University of Hawai'i at Manoa
Richard Moyle
The University of Queensland School of Agriculture and Food Sciences
Murray Sharman
DAF: Queensland Department of Agriculture and Fisheries
John Hu (johnhu@hawaii.edu)
University of Hawai'i at Mānoa: University of Hawai'i at Manoa https://orcid.org/0000-0003-3003-4140
John E. Thomas
Queensland Alliance for Agriculture and Food Innovation https://orcid.org/0000-0002-8941-0959

Research Article

Keywords:

Posted Date: March 14th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1416846/v1

Abstract

The complete genome sequence of a member of a novel species of *Sadwavirus* (family *Secoviridae*) infecting pineapple on the island of Oahu, Hawaii was determined by high throughput sequencing. The genome comprised two RNA molecules of 5,956 nt and 3,808 nt in length, excluding the poly-A tails at the 3' end and each encoding a single large polyprotein. The RNA-1 polyprotein contained five conserved domains, all associated with replication, while the RNA-2 polyprotein is cleaved into the movement protein and coat protein. Analysis of the Pro-Pol region revealed less than 75% amino acid identity with pineapple secovirus A, chocolate lily virus A, and Dioscorea mosaic-associated virus, all members and candidate members of the proposed subgenus *Cholivirus*. Two specific primer sets derived from HTS data were used in RT-PCR assays, confirming the presence of this new virus infecting pineapple. The name "pineapple secovirus B" is proposed for this putative new virus.

Full Text

Plant viruses in the family *Secoviridae* (order *Picornavirales*) mainly infect dicotyledonous plant species and are naturally transmitted by nematodes or arthropods. The capsids of these viruses are icosahedral, 25- to 30-nm in diameter, and composed of 60 coat protein (CP) subunits [9, 10]. The secovirus genome can be monopartite or bipartite. Bipartite genomes are divided between two RNA segments (RNA-1 and -2), each modified by addition of a 3'-terminal poly(A) tail, and a covalently bound VPg protein at the 5'-end. RNA-1 ranges in length from 6 to 8 kb and encodes the proteins necessary for cytoplasmatic replication. In contrast, RNA-2 is smaller at 2 to 4 kb and encodes the movement protein (MP) and up to three coat proteins (CP) [7]. The Pro-Pol sequence on RNA-1 is used to infer phylogenetic relationships within the family and with other members of the order [7]. Currently, members of the *Secoviridae* family are classified into eight genera: *Comovirus*, *Fabavirus*, *Nepovirus*, *Cheravirus*, *Sadwavirus*, *Torradovirus*, *Sequivirus*, and *Waikavirus* [7, 10]. A ninth genus, "*Stralarivirus"*, and the division of the genus *Sadwavirus* into the three subgenera, "Stramovirus", "Satsumavirus", and "Cholivirus" was proposed in the revision taxonomy of the family *Secoviridae* [7].

Pineapple secovirus A (PSV-A), an unassigned member within the family *Secoviridae*, was first detected from an *Ananas comosus* germplasm accession (HANA 187) from the Pacific Basin Agricultural Research center (PBARC) in Hilo, Hawaii [5]. PSV-A survey was carried out in 2019 on the island of Oahu and revealed the presence of the virus in six out of twelve plants presenting mealybug wilt of pineapple (MWP) symptoms (reddening and wilting of the leaves) but not in any the 13 asymptomatic plants [5]. Various combinations of ampelovirus from the pineapple mealybug wilt-associated virus (PMWaV) complex of species were also detected in the 12 symptomatic plants (Larrea-Sarmiento et al, unpublished results).

To examinate for the presence of further undiscovered viruses infecting *Ananas comosus*, high-throughput sequencing (HTS) was done on the same field plants detailed in the study by Larrea-Sarmiento et al. (2020). Total RNAs were extracted from the basal portions of individual pineapple leaf

samples using Spectrum™ total RNA kit (Sigma Aldrich) following the manufacturer's instructions. Total RNA extracted from the twelve MWP symptomatic field samples and thirteen healthy-looking plants were pooled into two composite RNA samples followed by a ribodepletion step to remove the ribosomal RNA (rRNA). cDNA library synthesis was followed by HTS using Illumina® NovaSeq6000 to obtain paired-end reads (2 × 100 bp) at the Genomics High-Throughput Sequencing Facility at the University of California, Irvine.

Data obtained from the ~40 million raw reads per composite ribosomal RNA-depleted total RNA were curated and assembled following the methods of Green et al. [2]. The resulting contigs were annotated by doing BLASTX searches of the NCBI virus sequence database. Annotated contigs revealed shared sequence identity with the previously characterized PMWaVs and secoviruses. Two contigs retrieved from the symptomatic composite sample had significant matches to PSV-A and other sadwaviruses but were sufficiently divergent to suggest they represented the two RNA components of a new species. Similar to PSV-A and the majority of secovirids, the potential new virus has a bipartite genome encoding two positive-sense RNA molecules.

To complete the genome sequence of the virus, 5' and 3' rapid amplification of cDNA ends (RACE) was performed. Both 5' and 3' ends were obtained using the Takara SMARTer RACE 5'/3' kit according to the manufacturer's instructions followed by PCR with a universal anchored primer and sequence specific primers (Table S1). Amplicons were cloned, and five to seven clones were Sanger sequenced. The complete genome comprised two RNA molecules; RNA-1 is 5,956 nt long (GenBank accession code OM777135) and RNA-2 is 3,808 nt long (GenBank accession code OM777136), each coding for large polyproteins referred as P1 and P2, respectively. The name pineapple secovirus B (PSV-B) is proposed for this putative new virus infecting pineapple.

The PSV-B P1 of 1875 aa is predicted to code for proteins involved in replication: protease cofactor (Pro-C), helicase (Hel), genome-linked viral protein (VPg), protease (Pro), and RNA-dependent RNA-polymerase (Pol). Likewise, PSV-B P2 of 1143 aa is predicted to code for the movement protein (MP) and one large coat protein (CP) (Fig. 1). Like other secovirids, both PSV-B RNAs are expected to possess a genome-linked viral protein termed "VPg" bound at the 5' end and a poly(A) tail at the 3' end, respectively [1, 7]. Analysis of conserved domains using the NCBI Conserved Domain Search Tool (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) revealed conserved motifs for the RdRp and Helicase for the P1; whereas a conserved MP motif was predicted for the P2. Parallel analysis using HMMER (https://www.ebi.ac.uk/Tools/hmmer/) and Pfam (http://pfam.xfam.org/) further predicted the presence of two CP domains corresponding to nepovirus and picornavirus CP domains for P2, although other members within the subgenus *Cholivirus* (Genus *Sadwavirus*) are predicted to encode only one large CP [4-6, 11]. Analysis of the predicted cleavage sites for P1 located four Q/S, and five E/G dipeptides were also found [3]. The cleavage sites recognized by the RNA-1-encoded 3C-like protease (3CL-Pro) likely cleave P1 at five sites, defining six domains; while 3CL-Pro likely cleaves P2 at one site, defining two domains (Fig. 1) [1, 6, 7].

The recently characterized PSV-A was found to be closely related to Dioscorea mosaic associated virus (DMaV) and to chocolate lily virus A (CLVA) [5]. In 2020, the proposed revision of the family *Secoviridae* classified DMaV and CLVA, previously denotated as unassigned secoviruses, as members of the *Cholivirus* subgenus within the *Sadwavirus* genus [7]. To study the taxonomic position PSV-B, and relatedness with PSV-A and other members of the family *Secoviridae*, phylogenetic analysis using the Maximum Likelihood method based on the aa sequence of the Pro-Pol region was carried out using LG+G as the best model of protein evolution. This analysis suggests that PSV-B is a new *Sadwavirus* member that is related to, but distinct from, the previously characterized *Cholivirus* member PSV-A (Fig. 2). It is placed on a branch distinct from PSV-A and basal to that containing DMaV and CLVA (Fig. 2). For members within the family *Secoviridae*, species demarcation criteria are <80% identity for the aa sequence of the Pro-Pol region, and <75% identity for the large and small CP together [10]. Identities of 45.1% and 53.5% were observed when comparing the Pro-Pol region of PSV-B to PSV-A and CLVA homolog regions, respectively, using pairwise comparisons. These results are consistent with the findings reported in Australia in 2002 where two isometric viruses were reported to infect pineapple and partial sequences showed similarities to strawberry mottle virus [8]. Combined results support the establishment of PSV-B as a member of a new species in the subgenus *Cholivirus* under the genus *Sadwavirus* in the family *Secoviridae* [7].

Specific primers were used for the detection of each of the two PSV-B RNAs from the 25 field samples collected in 2019. Primer pairs targeting the RdRp-coding region and CP-coding region for the detection of RNA1 and RNA2, respectively, were used in RT-PCR assays (Table S1). Total RNA was used for cDNA synthesis using the M-MLV reverse transcriptase from Promega. RT-PCR reactions were performed as follows: initial denaturation at 95ºC for 5 min, 35 cycles of 95ºC for 40 s, 57ºC for 30 s and 72ºC for 40 s, and the final extension at 72ºC for 10 minutes. Four out of the 12 MWP symptomatic samples revealed the expected band of 702 bp for RNA-1 and 380 bp for RNA-2, and their identity was confirmed by direct Sanger sequencing of the amplicons. The presence of PMWaVs and PSV-A was assessed as detailed previously (Green et al. 2019; Larrea-Sarmiento et al. 2020; Larrea-Sarmiento et al. 2021). Samples infected with PSV-B were also infected with PMWaV-2, PMWaV-3, PMWaV-6, and PSV-A. Nine out of the 13 healthy looking plant samples presented only single infections with PMWaV-3; while the remaining five did not test positive for any virus [5]. Further research is needed to determine the prevalence of PSV-B in other pineapple producing countries and to evaluate if PSV-B is involved in the etiology of MWP.

Declarations

Funding

The research was supported in part by grants from the United State 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.

Acknowledgements

The research was supported in part by grants from the United State 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).

Ethics declarations

Conflict of interest

The authors declare that they have no conflicts 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. Dullemans AM, Botermans M, de Kock MJD, de Krom CE, van der Lee TAJ, Roenhorst JW, Stulemeijer IJE, Verbeek M, Westenberg M, van der Vlugt RAA (2020) Creation of a new genus in the family Secoviridae substantiated by sequence variation of newly identified strawberry latent ringspot virus isolates. Archives of Virology 165:21-31
2. Green JC, Rwahnih MA, Olmedo-Velarde A, Melzer MJ, Hamim I, Borth WB, Brower TM, Wall M, Hu JS (2020) Further genomic characterization of pineapple mealybug wilt-associated viruses using high-throughput sequencing. Tropical Plant Pathology 45:64-72
3. Halgren A, Tzanetakis IE, Martin RR (2007) Identification, Characterization, and Detection of Black raspberry necrosis virus. Phytopathology 97:44-50
4. Hayashi EAI, Blawid R, de Melo FL, Andrade MS, Pio-Ribeiro G, de Andrade GP, Nagata T (2017) Complete genome sequence of a putative new secovirus infecting yam (Dioscorea) plants. Arch Virol 162:317-319
5. Larrea-Sarmiento A, Olmedo-Velarde A, Green JC, Al Rwahnih M, Wang X, Li YH, Wu W, Zhang J, Matsumoto TK, Suzuki JY, Wall MM, Borth W, Melzer MJ, Hu JS (2020) Identification and complete

genomic sequence of a novel sadwavirus discovered in pineapple (Ananas comosus). Arch Virol 165:1245-1248

6. Mann KS, Walker M, Sanfaçon H (2017) Identification of Cleavage Sites Recognized by the 3C-Like Cysteine Protease within the Two Polyproteins of Strawberry Mottle Virus. Frontiers in Microbiology 8
7. Sanfaçon H, Dasgupta I, Fuchs M, Karasev AV, Petrzik K, Thompson JR, Tzanetakis I, van der Vlugt R, Wetzel T, Yoshikawa N (2020) Proposed revision of the family Secoviridae taxonomy to create three subgenera, “Satsumavirus”, “Stramovirus” and “Cholivirus”, in the genus Sadwavirus. Archives of Virology 165:527-533
8. Thompson H, Botella, J., Fairbairn, D., , Moyle, R., Thomas, J. E., Sharman, M., Aitken, E.A., Geering, A.D.W. (2005) Characterization and detection of isometric viruses in pineapple. In: Proceedings of the 15th Biennial Australasian Plant Pathology Conference, Australia. Australasian Plant Pathology Society: Geelong, Victoria, p 138
9. Thompson JR, Kamath N, Perry KL (2014) An evolutionary analysis of the Secoviridae family of viruses. PLoS One 9:e106305
10. Thompson JR, Dasgupta I, Fuchs M, Iwanami T, Karasev AV, Petrzik K, Sanfaçon H, Tzanetakis I, van der Vlugt R, Wetzel T, Yoshikawa N, Ictv Report C (2017) ICTV Virus Taxonomy Profile: Secoviridae. The Journal of general virology 98:529-531
11. Wylie SJ, Luo H, Li H, Jones MGK (2012) Multiple polyadenylated RNA viruses detected in pooled cultivated and wild plant samples. Archives of Virology 157:271-284

Figures

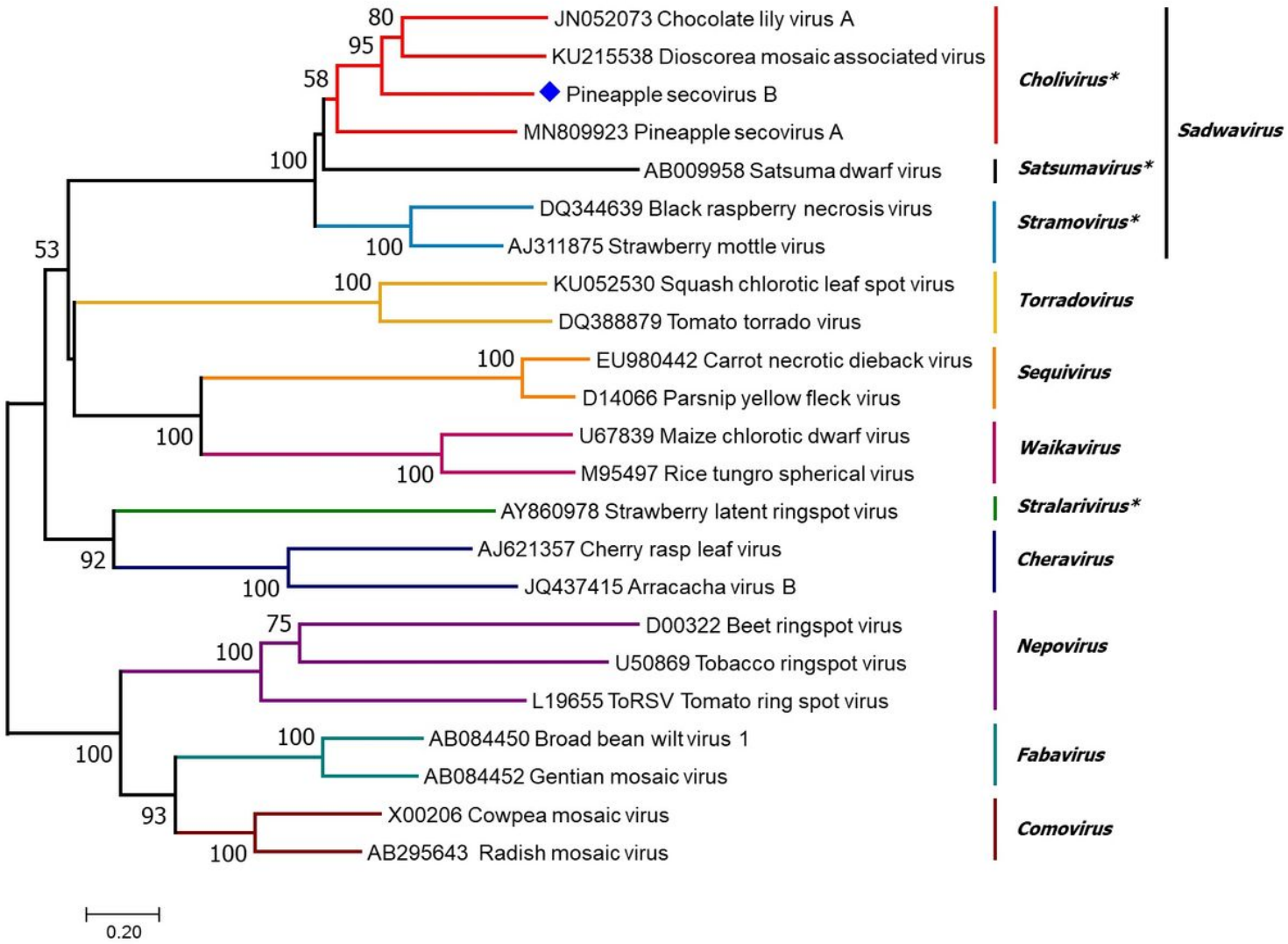

Figure 1

Phylogenetic relationships of the Pro-Pol region of pineapple secovirus B (PSV-B) with other members of the family *Secoviridae*. The Maximum Likelihood method with the LG+G matrix-based model were used with 1,000 bootstrap pseudo-replicates as percentage values for branch support. Protein predicted sequences were used and the respective GenBank accessions are provided with each virus name. Alignment was generated using Clustal and implemented in Mega v.7.0.1. Asterisks represent the three proposed subgenera within the genus *Sadwavirus* and the proposed genus "*Stralarivirus*", which have yet to the ratified by the ICTV. The blue diamond indicates PSV-B.

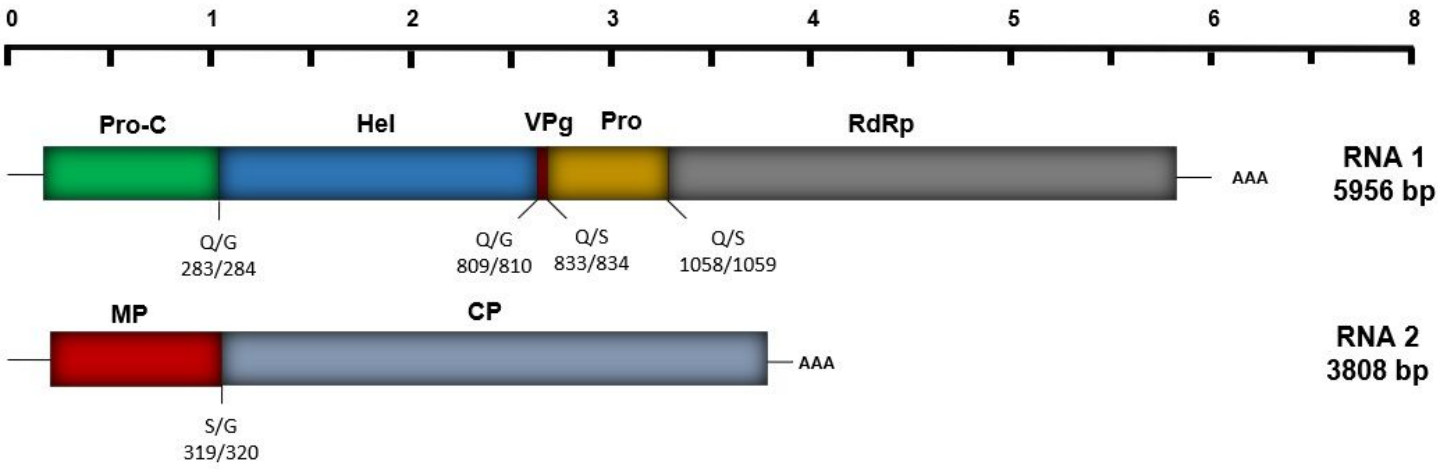

Figure 2

Genome organization of pineapple secovirus B (PSV-B): predicted cleavage sites with their corresponding dipeptide in P1 and P2 polyproteins are shown as vertical lines under the complete full-length of each RNA fragment of PSV-B. Numbers indicates the amino acid position. RNA-1 (5,956 bp) encodes a large polyprotein predicted to be cleaved into 5 proteins: protease cofactor (Pro-C), helicase (Hel), genome-linked viral protein (VPg), protease (Pro), and RNA-dependent RNA-polymerase (Pol). RNA-2 (3,808 bp) is predicted to code for the movement protein (MP) and one large coat protein (CP). Q, glutamine; G, glycine; S, serine. “AAA” at the 3’ end position of each RNA segment represents the poly-A tail.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- FASTAPSVB.docx
- TABLES1PSVB.docx