Archives of Virology (2018) 163:3471–3475
https://doi.org/10.1007/s00705-018-3980-x

ANNOTATED SEQUENCE RECORD

Analysis of the complete genomic sequence of a novel virus, areca palm necrotic spindle-spot virus, reveals the existence of a new genus in the family *Potyviridae*

Ke Yang[1] · Minyuan Ran[1] · Zengping Li[1] · Meijiao Hu[2] · Li Zheng[1] · Wenbo Liu[1] · Pengfei Jin[1] · Weiguo Miao[1] · Peng Zhou[3] · Wentao Shen[3] · Hongguang Cui[1]

Received: 20 May 2018 / Accepted: 26 July 2018 / Published online: 22 August 2018

Abstract

A novel virus, tentatively named "areca palm necrotic spindle-spot virus" (ANSSV), was identified in *Areca catechu L.* in Hainan, China, and its complete genomic sequence was determined. Its positive-sense single-stranded RNA genome is comprised of 9,437 nucleotides (nt), excluding the poly (A) tail, and contains one large open reading frame encoding a polyprotein of 3,019 amino acids (aa). A Blastp search showed that the polyprotein of ANSSV shared a maximum of 31%-32% aa sequence identity (with 86%-95% coverage) with all seven known macluraviruses. Nucleotide sequence comparison of the ORF of ANSSV to those of macluraviruses revealed identities ranging from 41.0% to 44.6%, which is less than the inter-genus identity values for the family *Potyviridae*. Phylogenetic analysis based on either the aa or nt sequence of the polyprotein did not cluster ANSSV into any established or unassigned genus of the family *Potyviridae*. Therefore, we suggest that ANSSV is the first member of a previously unrecognized genus of the family *Potyviridae*.

The family *Potyviridae* is the largest family of plant-infecting RNA viruses, members of which include many agriculturally and economically important viral pathogens that cause significant losses in a wide range of crops [1, 2]. Currently, the family is divided into 10 definitive genera (*Potyvirus*, *Bevemovirus*, *Brambyvirus*, *Bymovirus*, *Ipomovirus*, *Macluravirus*, *Poacevirus*, *Roymovirus*, *Rymovirus* and *Tritimovirus*), which are distinguished by host range, genomic features, and phylogeny, as well as an unassigned genus [1, 3, 4]. Most viruses within the family have monopartite, positive-sense single-stranded genomes (8.2 to 11.3 kb) and flexuous filamentous particles that are 680-900 nm in length and 11-20 nm in width [1]. However, members of the genus *Bymovirus* have bipartite genomes with particles of two modal lengths of 250-300 and 500-600 nm [1]. For most members of the family, the RNA genome contains a long open reading frame (ORF) and another relatively short ORF transcribed by RNA polymerase slippage in the P3 coding sequence [5–7], and the resulting two polyproteins are proteolytically processed into 11 mature viral proteins, namely P1, HC-Pro, P3, P3N-PIPO, 6K1, CI, 6K2, VPg, NIa-Pro, NIb and CP. Notably, P1, which is a serine proteinase, is absent in all known viruses in the genus *Macluravirus*.

Areca palm (*Areca catechu L.*) is the second largest commercial crop in Hainan, China. Unfortunately, in recent years it has been severely damaged by a variety of pathogens and insects, although viral pathogens have rarely been reported, except for areca palm velarivirus 1 (a member of the genus *Velarivirus* of the family *Closteroviridae*), which has been identified in yellowing leaves of areca palm in Hainan, China [8]. In October 2017, a suspected virus-infected areca palm showing chlorosis, necrosis and spindle-spot symptoms (Fig. 1A) was observed in Baoting, Hainan, and crude

Handling Editor: Ioannis E. Tzanetakis.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00705-018-3980-x) contains supplementary material, which is available to authorized users.

✉ Hongguang Cui
hongguang.cui@hainu.edu.cn

[1] Institute of Tropical Agriculture and Forestry, Hainan University, Haikou 570228, Hainan, People's Republic of China

[2] Environment and Plant Protection Institute, Chinese Academy of Tropical Agricultural Sciences, Haikou 571101, Hainan, People's Republic of China

[3] Institute of Tropical Bioscience and Biotechnology, Chinese Academy of Tropical Agricultural Sciences, Haikou 571101, Hainan, People's Republic of China

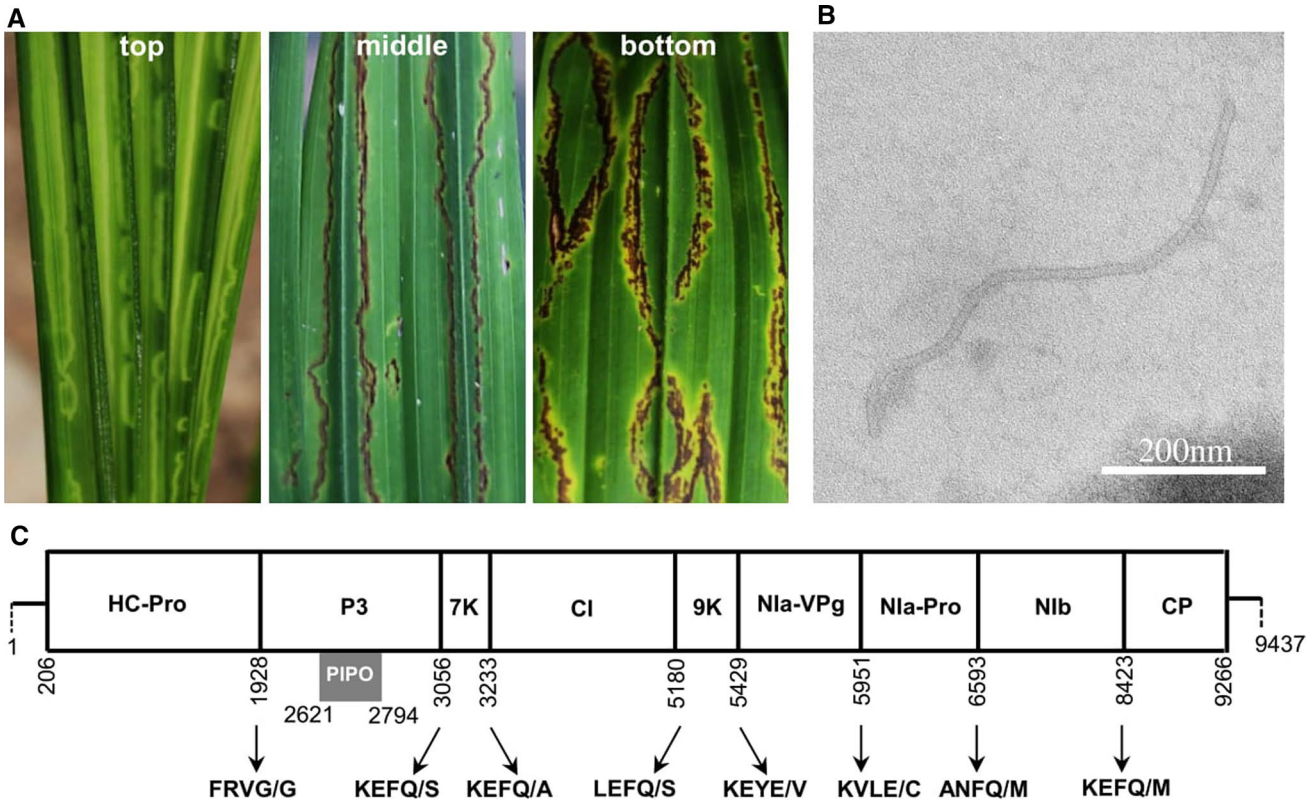

Fig. 1 Disease symptoms and full-length genomic organization diagram of areca palm necrotic spindle-spot virus (ANSSV). **a.** ANSSV-infected areca palm showing chlorosis on top leaves and necrosis and spindle-spot on medium and bottom leaves. **b.** Transmission electron micrographs of ANSSV virions from crude extracts of diseased areca palm leaves. **c.** Schematic representation of the genomic organization of ANSSV. Two short horizontal lines, 5' and 3' UTR, respectively; large box, the long ORF; smaller boxes, the mature proteins resulting from the proteolytic processing of the large polyprotein; arrows, putative conserved sites recognized by HC-Pro or NIa-Pro proteinases

extracts of diseased leaves were subjected to screening of viral pathogens by transmission electron microscopy. As expected, viral particles with flexuous filaments of 15 × 780 nm were observed (Fig. 1B). To characterize the viral agent, the diseased leaves were subjected to small-RNA analysis, essentially as described previously [9]. A total of 27,746,020 clean sRNA reads were obtained and assembled *de novo* into larger contigs. The resulting 693 contigs were subjected to a Blastx search against the non-redundant protein database of NCBI GenBank, and a total of 44 contigs with sizes ranging from 56 to 300 bp showed relatively high sequence similarity to the genomic sequences of viruses in the genus *Macluravirus* (or *Bymovirus*) of the family *Potyviridae*. It was not found that any contig or its deduced amino acid sequence was homologous to genomic or protein sequences of other plant-infecting viruses. Herein, the characterized viral agent was tentatively named areca palm necrotic spindle-spot virus (ANSSV).

To clone and sequence the entire genome of ANSSV, eight fragments covering nearly the complete genome were amplified using RT-PCR with the corresponding primer sets designed based on the sequences of the contigs (Supplementary Table S1). The 5'- and 3'-terminal cDNA sequences were obtained using 5' and 3'RACE kits (Invitrogen). The conventional cloning strategy and sequence analysis were essentially the same as reported recently [9]. Eventually, a complete genomic sequence with a length of 9,437 nucleotides (nt) excluding the poly(A) tail was generated for ANSSV (Fig. 1C, and deposited in the GenBank database with accession number MH330686.

The complete genome of ANSSV contains a large ORF (9,060 nt), flanked by a 5' untranslated region (UTR) of 205 nt and a 3'-UTR of 172 nt and encoding a polyprotein of 3,019 aa. Sequence comparisons between ANSSV and all other macluraviruses revealed eight conserved cleavage sites (Supplementary Table S2), which were predicted to be recognized by HC-Pro or NIa-Pro to cleave the polyprotein into nine mature proteins, denoted HC-Pro, P3, 7K, CI, 9K, VPg, NIa, NIb and CP (Fig. 1C). The genome organization is similar to that of macluraviruses, as the ANSSV genome lacks the P1-encoding sequence (Fig. 1C). In addition, a putative small ORF (PIPO) embedded within the P3 cistron, resulting from RNA polymerase slippage at GA7 (nt 2621-2794) [5–7], was also predicted. Similar to other viruses

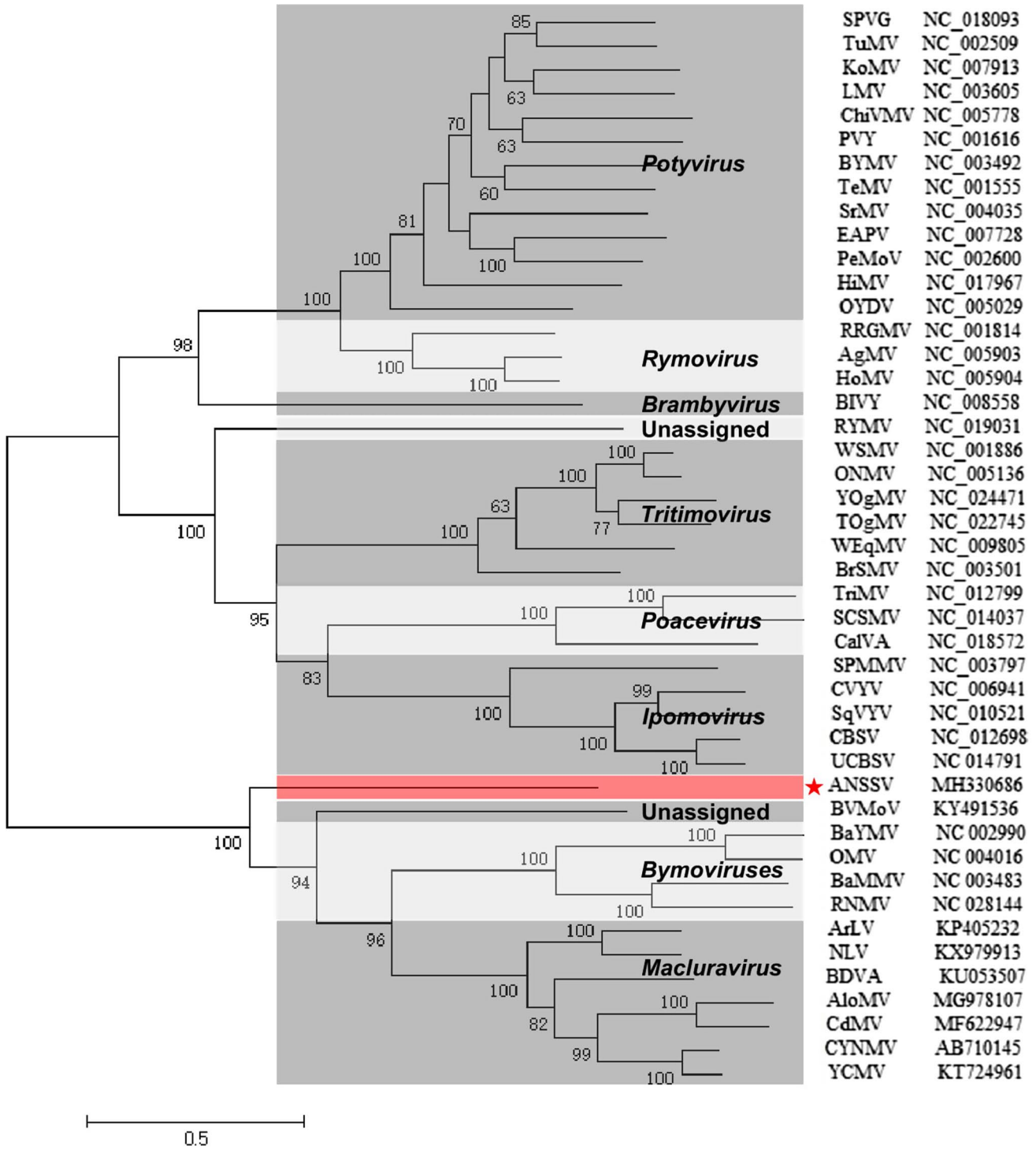

Fig. 2 Phylogenetic analysis of aa sequences of the polyprotein of ANSSV, all seven reported macluraviruses, and 37 representative members in the other definite and unassigned genera within the family *Potyviridae*. These sequences were completely aligned using ClustalX, and a phylogenetic tree was constructed with MEGA 7.0 and the maximum-likelihood method. The numbers at the nodes indicate the percentage of 1000 bootstrap replicates supporting this group. Values below 60% are not shown. The bar represents 0.5 substitutions per site. ANSSV is indicated by a red asterisk

within the family *Potyviridae*, the conserved motif ^{461}G-X-C-X30-L-X2-WP-X36-HF^{536}, the catalytic active site, was identified in HC-Pro, but the 'PTK' and 'KITC' motifs, which are involved in aphid transmission, and the 'FRNK' motif, which is involved in RNA silencing suppression, were absent. The conserved motifs ^{1093}GSGKS-X3-P^{1101} and

^{1186}DE-X-H^{1189}, which are important for helicase activity, were present in CI, as was the NIb motif ^{2752}GDD^{2754}, which is associated with replicase activity. A conserved motif, 'DAG', located near the amino-proximal region of the CP of most potyviruses and implicated in aphid transmission, was absent for ANSSV.

A Blastp search against the NCBI GenBank non-redundant protein database showed that the polyprotein of ANSSV shares a maximum of 31%-32% aa sequence identity (with 86%-95% coverage) with those of all seven known macluraviruses, followed by 32% (with 70% coverage) with bellflower veinal mottle virus (BVMoV, a member of an unassigned genus) and 26%-29% (with 73%-78% coverage) with all five known bymoviruses in the family *Potyviridae*. A nucleotide sequence comparison of the polyprotein-encoding region of ANSSV to orthologues of these viruses revealed identity values ranging from 38.2% to 44.6%, which is below the threshold for genus demarcation within the family *Potyviridae* [10]. Moreover, an aa sequence comparison of each mature protein was performed. Consistently, ANSSV was identified to have higher sequence similarity to macluraviruses than to BVMoV or bymoviruses for each putative mature protein except PIPO (Supplementary Table S3). CI and NIb were identified to be relatively conserved proteins, sharing 30.2%-38.5% and 38.4%-47.7% sequence identity, respectively. In contrast, extremely low sequence identities were found for PIPO and 9K (8.5%-23.1% and 11.4%-30%, respectively; Supplementary Table S3). As the CP sequence is an important taxonomic criterion for classification of members of the family *Potyviridae* [10], we further compared nt and aa sequences of the ANSSV CP with those of all seven known macluraviruses and 38 representative members of the other definite and unassigned genera of the family *Potyviridae*. The results showed that the CP of ANSSV had maximum sequence identity to macluraviruses (44.4%-47.0% and 30.8%-37.6% at the nt and aa level, respectively), followed by bymoviruses (38.1%-40.7% and 26.1%-28.0%, respectively) (Supplementary Table S4).

To examine the taxonomic classification of ANSSV, phylogenetic analysis was performed using the aa and nt sequences of the polyprotein of ANSSV and those of all seven reported macluraviruses (including our newly reported Alpinia oxyphylla mosaic virus with GenBank accession no. MG978107 [9]) and 37 representative members of the other definite and unassigned genera of the family *Potyviridae*. Phylogenetic analysis based on aa and nt sequences both consistently placed ANSSV in a separate clade that was most closely related to two definite genera *Macluravirus* and *Bymovirus* and an unassigned genus represented by BVMoV [4] (Fig. 2, Supplementary Figure S1).

In summary, the complete genomic sequence of a novel virus infecting areca palm was determined and annotated. We show that the polyprotein of ANSSV has maximum sequence identity to those of all reported macluraviruses within the family *Potyviridae* (31%-32% and 41.0%-44.6% at the aa and nt level, respectively). Furthermore, phylogenetic analysis did not cluster ANSSV in any of the established or unassigned genera of this family. Consequently, we suggest that ANSSV is the first member of a putative new genus in the family *Potyviridae*, which might be called "*Arepavirus*" (It has to be mentioned that another novel virus, namely areca palm necrotic ring-spot virus (ANRSV), newly identified from areca palm by our group, belongs to the new genus as well, unpublished data). However, it is not known whether the disease symptoms observed in areca palm were caused by ANSSV. Therefore, an infectivity test (either with an infectious cDNA clone or with purified viral particles) is needed to fulfill Koch's postulates.

Acknowledgements The authors thank Dr. Xiaofei Cheng (Northeast Agricultural University) for analysis of sRNA data.

Funding This study was funded by the Hainan Major Research Project for Science and Technology (Research and Demonstration of Key Technologies on Yellowing Disaster Control and Ecological Highly-efficient Cultivation of Areca Palm), Hainan Provincial Natural Science Foundation of China (Grant no. 318MS006) and the Scientific Research Foundation for Advanced Talents, Hainan University (Grant no. KYQD(ZR)1734).

Compliance with ethical standards

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

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

References

1. Wylie SJ, Adams M, Chalam C, Kreuze J, López-Moya JJ, Ohshima K, Praveen S, Rabenstein F, Stenger D, Wang A, Zerbini FM, ICTV Report Consortium ICTV (2017) Virus Taxonomy Profile: Potyviridae. J Gen Virol 98:352–354
2. Revers F, García JA (2015) Molecular biology of potyviruses. Adv Virus Res 92:101–199
3. Mollov D, Lockhart B, Zlesak D (2013) Complete nucleotide sequence of rose yellow mosaic virus, a novel member of the family Potyviridae. Arch Virol 158:1917–1923
4. Seo JK, Kwak HR, Kim MK, Kim JS, Choi HS (2017) The complete genome sequence of a novel virus, bellflower veinal mottle virus, suggests the existence of a new genus within the family Potyviridae. Arch Virol 162:2457–2461

5. Chung BY, Miller WA, Atkins JF, Firth AE (2008) An over-lapping essential gene in the Potyviridae. Proc Natl Acad Sci USA 105:5897–5902
6. Olspert A, Chung BY, Atkins JF, Carr JP, Firth AE (2015) Transcriptional slippage in the positive-sense RNA virus family Potyviridae. EMBO Rep 16:995–1004
7. Rodamilans B, Valli A, Mingot A, San León D, Baulcombe D, López-Moya JJ, García JA (2015) RNA polymerase slippage as a mechanism for the production of frameshift gene products in plant viruses of the Potyviridae family. J Virol 89:6965–6967
8. Yu H, Qi S, Chang Z, Rong Q, Akinyemi IA, Wu Q (2015) Complete genome sequence of a novel velarivirus infecting areca palm in China. Arch Virol 160:2367–2370
9. Hu W, Li Z, Wang X, Liu W, Huang C, Miao W, Cui H (2018) Complete genomic sequence of a novel macluravirus, alpinia oxyphylla mosaic virus (AloMV), identified in Alpinia oxyphylla. Arch Virol. https://doi.org/10.1007/s00705-018-3879-6
10. Adams MJ, Antoniw JF, Fauquet CM (2005) Molecular criteria for genus and species discrimination within the family Potyviridae. Arch Virol 150:459–479