Full-Genome Sequencing of a Virus from a 33-Year-Old Sample Demonstrates that *Arracacha Mottle Virus* Is Synonymous with *Arracacha Virus Y*

Ian P. Adams,[a] Neil Boonham,[b] Roger A. C. Jones[c,d]

[a]Fera Science Ltd., Sand Hutton, York, United Kingdom
[b]School of Natural and Environmental Sciences, Newcastle University, Newcastle upon Tyne, United Kingdom
[c]Institute of Agriculture, Faculty of Science, University of Western Australia, Crawley, Western Australia, Australia
[d]Department of Primary Industries and Regional Development, South Perth, Western Australia, Australia

ABSTRACT We describe here the first genome sequence of *Arracacha virus Y* (ArVY) derived from an arracacha (*Arracacia xanthorrhiza*) sample originally collected in 1976 in Peru and compare it with other potyvirus genome sequences. It had a 79% nucleotide identity with a 2013 Brazilian *Arracacha mottle virus* (AMoV) sequence, suggesting that AMoV is ArVY.

When the samples taken in a 1975 survey of subsistence plantings containing crop mixtures at Umari, Huanuco Department, in the Peruvian Andes were analyzed, two viruses with isometric particles, *Arracacha virus A* (genus *Nepovirus*) and *Arracacha virus B* (AVB; genus *Cheravirus*) were isolated from arracacha (Peruvian parsnip; *Arracacia xanthorrhiza*) plants showing pronounced yellow mosaic symptoms in young leaves (1, 2). After a second visit to the same location in 1976, in addition to AVA and AVB being found again in some samples, a third virus with 750-nm flexuous filamentous particles was detected in samples from symptomless arracacha plants. As its particles were typical of the family *Potyviridae*, which is named after *Potato virus Y*, the third virus was named *Arracacha virus Y* (ArVY). A fourth virus with 650-nm filamentous particles typical of the genus *Carlavirus* was also present. It was first named *Arracacha latent virus* (3) but was later identified as another *Carlavirus*, *Potato virus S* (4). In 1981, ArVY was recovered from leaves of arracacha plants that were freeze-dried in 1979 in the United Kingdom. The infected arracacha plants had been propagated serially from tubers since the original tubers from Umari were sent to the United Kingdom for electron microscopy in 1976. The recovered ArVY was mechanically transmissible to several plant species in the families *Aizoaceae*, *Chenopodiaceae*, and *Solanaceae* and was transmitted nonpersistently by the aphid species *Aphis gossypii* and *Myzus persicae* (3). ArVY-infected leaf material was freeze-dried in glass vials in 1984 and kept thereafter in what is now called the Fera Science Ltd. plant virus collection in York, England. In 2009, a potyvirus was found infecting symptomatic arracacha in Brazil and called *Arracacha mottle virus* (AMoV), and a complete genomic sequence of it was obtained in 2013 (5, 6).

In this paper, the methods used were the same as those described previously (7–9). Briefly, an RNeasy kit (Qiagen, UK), was used to extract total RNA from the freeze-dried ArVY-infected leaf material in 2017. Following the manufacturer's instructions, a ScriptSeq complete plant leaf kit (Illumina, USA) was employed to construct an indexed plant ribosome subtracted sequencing library. Utilizing a 600 cycle V3 kit, this indexed library was sequenced along with others on a MiSeq instrument (Illumina). Sickle (https://github.com/najoshi/sickle) was used in paired-end mode to 3′ trim the result-

Address correspondence to Roger A. C. Jones, roger.jones@uwa.edu.au.

ing 593,960 paired reads to a quality score of 20. The reads were assembled with Trinity v2, with the maximum memory allocation set to 99 gigabytes of RAM, and the process allocated 64 central processing units (10). We used BLAST+ (11) to compare the contigs obtained with the GenBank nonredundant (nr) and nucleotide databases. Next, the extract reads function in MEGAN (12) was used to extract reads of viral origin. One 9,647-nucleotide (nt)-long contig which had been assembled from 439,090 of the paired reads (minimum depth, 10 reads; average depth, 17,918 reads) resembled the genome of a potyvirus. The reads were *de novo* assembled, and the read count was from postassembly mapping back. The ArVY genome had a 79% nucleotide identity to the AMoV genome (GenBank accession number DQ925486). Given that ArVY had a nucleotide identity of >75% with AMoV, both belong to the same potyvirus species (13). As the name ArVY was used first for this virus, it should take precedence over the later name, AMoV.

Data availability. The ArVY sequence described here was deposited in GenBank under the accession number MH716807. Raw data were deposited in the SRA under BioSample number SAMN10081144 and SRA run number SRR7873470, which are part of SRA study number SRP162046.

ACKNOWLEDGMENTS

REFERENCES

1. Jones RAC, Kenten RH. 1978. Arracacha virus A, a newly recognised virus infecting arracacha (Arracacia xanthorrhiza; Umbelliferae) in the Peruvian Andes. Ann Appl Biol 90:85–91. https://doi.org/10.1111/j.1744-7348.1978.tb02613.x.

2. Kenten RH, Jones RAC. 1979. Arracacha virus B, a second isometric virus infecting arracacha (*Arracacha xanthorrhiza*; *Umbelliferae*) in the Peruvian Andes. Ann Appl Biol 93:31–36. https://doi.org/10.1111/j.1744-7348.1979.tb04722.x.

3. Brunt AA, Crabtree K, Gibbs AJ (ed). 1990. Viruses of tropical plants. Descriptions and lists from the VIDE database. CAB International, Wallingford, UK.

4. Santillan FW, Fribourg CE, Adams IP, Gibbs AJ, Boonham N, Kehoe MA, Maina S, Jones RAC. 2018. The biology and phylogenetics of *Potato virus S* isolates from the Andean region of South America. Plant Dis 102:869–885. https://doi.org/10.1094/PDIS-09-17-1414-RE.

5. Orilio AF, Dusi AN, Madeira NR, Inoue-Nagata AK. 2009. Characterization of a member of a new *Potyvirus* species infecting arracacha in Brazil. Arch Virol 154:181–185. https://doi.org/10.1007/s00705-008-0289-1.

6. Orílio AF, Lucinda N, Dusi AN, Nagata T, Inoue-Nagata AK. 2013. Complete genome sequence of arracacha mottle virus. Arch Virol 158:291–295. https://doi.org/10.1007/s00705-012-1473-x.

7. Adams PI, Boonham N, Jones RAC. 2017. First complete genome sequence of *Arracacha virus A* from an 38-year-old sample from Peru. Genome Announc 5:e00141-17. https://doi.org/10.1128/genomeA.00141-17.

8. Adams PI, Fox A, Boonham N, Jones RAC. 2017. Complete genomic sequence of the potyvirus, *Mashua virus Y*, obtained from a 33-year-old mashua (*Tropaeaolum tuberosum*) sample. Microbiol Resour Announc 7:e01064-18. https://doi.org/10.1128/MRA.01064-18.

9. Adams PI, Abad J, Fribourg CE, Boonham N, Jones RAC. 2017. Complete genome sequence of *Potato virus T* from Bolivia obtained from a 33-year-old sample. Microbiol Resour Announc 7:e01066-18. https://doi.org/10.1128/MRA.01066-18.

10. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol 29:644–652. https://doi.org/10.1038/nbt.1883.

11. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: architecture and applications. BMC Bioinformatics 10:421. https://doi.org/10.1186/1471-2105-10-421.

12. Huson DH, Beier S, Flade I, Górska A, El-Hadidi M, Mitra S, Ruscheweyh H-J, Tappu R. 2016. MEGAN community edition—interactive exploration and analysis of large-scale microbiome sequencing data. PLoS Comput Biol 12:e1004957. https://doi.org/10.1371/journal.pcbi.1004957.

13. Adams MJ, Antoniw JF, Fauquet CM. 2005. Molecular criteria for genus and species discrimination within the family *Potyviridae*. Arch Virol 150:459–479. https://doi.org/10.1007/s00705-004-0440-6.