



A novel picorna-like virus identified in the cotton boll weevil *Anthonomus grandis* (Coleoptera: Curculionidae)

Leonardo Assis da Silva¹ · Marcos Fernando Basso² · Bergmann Morais Ribeiro¹

Received: 18 August 2022 / Accepted: 27 September 2022 / Published online: 4 January 2023
© The Author(s), under exclusive licence to Springer-Verlag GmbH Austria, part of Springer Nature 2022

Abstract

The cotton boll weevil (CBW; *Anthonomus grandis*; Coleoptera: Curculionidae) is considered the major insect pest of cotton, causing considerable losses in yield and fiber quality. An increase in the boll weevil population due to increasingly inefficient chemical control measures is of great concern among cotton producers. The absence of conventional or transgenic cultivars with minimal resistance to CBW has stimulated the search for new molecular and biological tools for efficient control of this insect pest. In this study, we used a metagenomic approach based on RNA deep sequencing to investigate the presence of viruses and coding viral RNA in apparently healthy native adult CBW insects collected from cotton crops in Mato Grosso state, Brazil. Using an Illumina HiSeq 2000 paired-end platform, 138,798 virus-related reads were obtained, and a consensus sequence of a putative new virus, 10,632 nucleotides in length, was assembled. The sequences of the 5' and 3' untranslated regions (UTRs) were determined by rapid amplification of cDNA ends (RACE), followed by Nanopore sequencing. The complete genome sequence included a 5'-UTR (1,158 nucleotides), a 3'-UTR (561 nucleotides), and a single ORF of 8,913 nucleotides encoding a large polyprotein. Sequence analysis of the putative polyprotein showed several regions with high sequence similarity to structural and non-structural proteins of viruses of the family *Iflaviridae*. Pairwise alignments of polyprotein amino acid sequences showed the highest sequence identity (32.13%) to a partial polyprotein sequence of a putative iflavirus (QKN89051.1) found in samples from wild zoo birds in China. Phylogenetic analysis based on full polyprotein sequences of different iflaviruses indicated that this new picorna-like virus is most closely related to iflaviruses found in lepidopteran insects, and it was therefore tentatively named "Anthonomus grandis iflavirus 1" (AgIV-1). This is, to our knowledge, the first complete viral genome sequence found in CBW, and it could provide a basis for further studies about the infectivity and transmission of this virus and its possible association with symptoms or acute disease. AgIV-1 could potentially be used to develop biological or molecular tools, such as a viral vector to carry interfering RNA molecules for CBW control.

Handling Editor: Simona Abba'

Accession number: The partial genome sequence was deposited in the GenBank database under accession number OK413669.

✉ Bergmann Morais Ribeiro
bergmann@unb.br

¹ Laboratory of Baculovirus, Cell Biology Department, University of Brasília (UnB), Campus Universitário Darcy Ribeiro, Brasília, DF CEP 70910-900, Brazil

² BIOMOL/BIOTEC Laboratory, Mato Grosso Cotton Institute (IMAmt), Rondonópolis, MT 78740-970, Brazil

Introduction

Cotton (*Gossypium hirsutum*) is one of the major socioeconomically important crops cultivated worldwide, serving as the main source of fiber for the textile industry [1–3]. Cotton crops are constantly challenged by several insect pest species [4]. The cotton boll weevil (CBW), *Anthonomus grandis* (Coleoptera: Curculionidae), is considered the major insect pest in South and North America, and cotton crops exhibit the highest incidence of infestation during the transition period from flowering to fructification [5, 6]. CBW adults feed on and lay eggs within the cotton reproductive structures, often causing flower bud abortion [7, 8]. Since CBWs are endophytic, their larvae can cause damage to flower buds when they are not aborted, impacting fiber quality [6, 9]. Their high reproductive

capacity, plasticity, and genetic variability and their occurrence in crop residues or stumps has helped to increase the incidence, density, and geographic distribution of CBWs worldwide [6, 10–12]. There is currently no conventional or transgenic cotton cultivar with satisfactory resistance to CBW available to cotton producers. Consequently, insecticides are applied many times each year for its management [13]. Unfortunately, the frequent occurrence of CBW populations with reduced susceptibility to insecticides and failure of chemical control in cotton crops has already been reported in Brazil [14, 15]. Meanwhile, the identification of new viruses that infect CBW may provide information leading to the development of molecular or biological tools for their effective control.

Here, we used an RNA sequencing approach involving next-generation sequencing (NGS) to investigate the presence of viruses and coding viral RNA in apparently healthy native adult CBW insects collected in September 2020 in a cotton field situated in Serra da Petrovina (16°47'53''S and 54°07'53''W), Pedra Preta city, Mato Grosso state, Brazil. Pooled CBW insects (200 mg) were macerated in SM buffer (100 mM NaCl, 8 mM MgSO₄, and 50 mM Tris-Cl, pH 7.5), using a mortar and pestle. The homogenate was filtered once through cheesecloth and centrifuged three times at 4,000 × g for 10 min to clarify the supernatant, which was then used for RNA extraction using a QIAamp Viral RNA Mini Kit (QIAGEN, Hilden, Germany) according to the manufacturer's instructions. rRNA was removed from the total RNA sample using a Ribo-Zero rRNA Removal Kit (Illumina, San Diego, CA, USA), and a cDNA library was constructed using a TruSeq RNA Library Preparation Kit (Illumina, San Diego, CA, USA). The cDNA sample was sequenced at Macrogen (Gangnam-gu, Seoul, Republic of Korea) using an Illumina HiSeq 2000 paired-end platform. The raw reads were quality trimmed and assembled *de novo* using MEGAHIT software [16]. The resulting contigs that were closely related to viral sequences were retrieved from an in-house viral RefSeq database using BLASTx. To extend the assembled sequences as far as possible, generated/trimmed reads were mapped back to the respective viral genomes using Geneious 11.1.5 software [17], which was also used for genome annotation. Open reading frames (ORFs) were confirmed using a BLASTx search against the NCBI non-redundant protein database (08/2021).

The NGS resulted in 56,210,608 total reads, 138,798 of which were considered virus-related sequences. *De novo* assembly of these viral reads generated a consensus sequence that was 10,632 nucleotides in length (GenBank accession number OK413669, Supplementary Material S1). The genome coverage was 1,440X. A single ORF of 8,913 nucleotides encoding a large polyprotein was predicted, in addition to a 5'-UTR (1,158 nucleotides) and a 3'-UTR (561 nucleotides). The 5' and 3' ends of the viral genome were confirmed by rapid amplification of cDNA ends (RACE) using 5' and 3' RACE System for Rapid Amplification of cDNA Ends, version 2.0 (Thermo Fisher Scientific) according to the manufacturer's protocol (data not shown). The amplified 5' and 3' products were sequenced using a MinION platform, using a Rapid RBK110.96 kit, following the manufacturer's instructions (Oxford Nanopore Technologies), and the sequences were analyzed using Geneious 11.1.5 software. Functional regions of the polyprotein encoding structural and non-structural proteins flanked by putative proteolytic cleavage sites were identified by sequence alignment (Fig. 1). From these data, the genome organization of the novel virus clearly resembled that of other members of the family *Iflaviridae* [18, 19].

An amino acid sequence alignment made using the MAFFT method [20] showed that this putative polyprotein was 32.13% identical to the corresponding sequence of a putative iflavirus (QKN89051.1) found in samples from wild zoo birds in China. The International Committee on Taxonomy of Viruses (ICTV) [19] has established two demarcation criteria for creating new species in the genus *Iflavirus*: (1) natural host range and (2) amino acid sequence identity in the capsid protein under 90% [21, 22]. We therefore conclude that this new picorna-like virus should be recognized as a member of a new species in the genus *Iflavirus*. We have tentatively named this putative new virus "Anthonomus grandis iflavirus 1" (AgIV-1).

Phylogenetic analysis was performed based on whole genome sequences of the putative novel virus and those of other iflaviruses (Supplementary Table S1). Sequence alignments were made by the MAFFT method [20] using the whole polyprotein coding regions of related viruses (Fig. 2). A maximum-likelihood tree was constructed by the FastTree method implemented in Geneious 11.1.5 software [17], and branch support was estimated using a

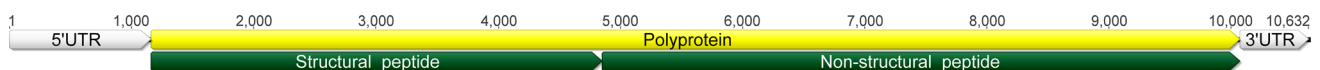


Fig. 1 Genome organization of the putative new picorna-like virus, tentatively named "Anthonomus grandis iflavirus 1" (AgIV-1), identified in cotton boll weevil (*Anthonomus grandis*). The viral genome consists of 10,632 nucleotides with a 5'-UTR and a 3'-UTR with a polyadenylated tail, containing a large polyprotein ORF of 8,913

nucleotides encoding the structural proteins VP1 to VP4 in the N-terminal region and the non-structural proteins helicase, protease, and RNA-dependent RNA polymerase (RdRp) in the C-terminal region. The N-terminus of the polyprotein contains a typical L-polypeptide.

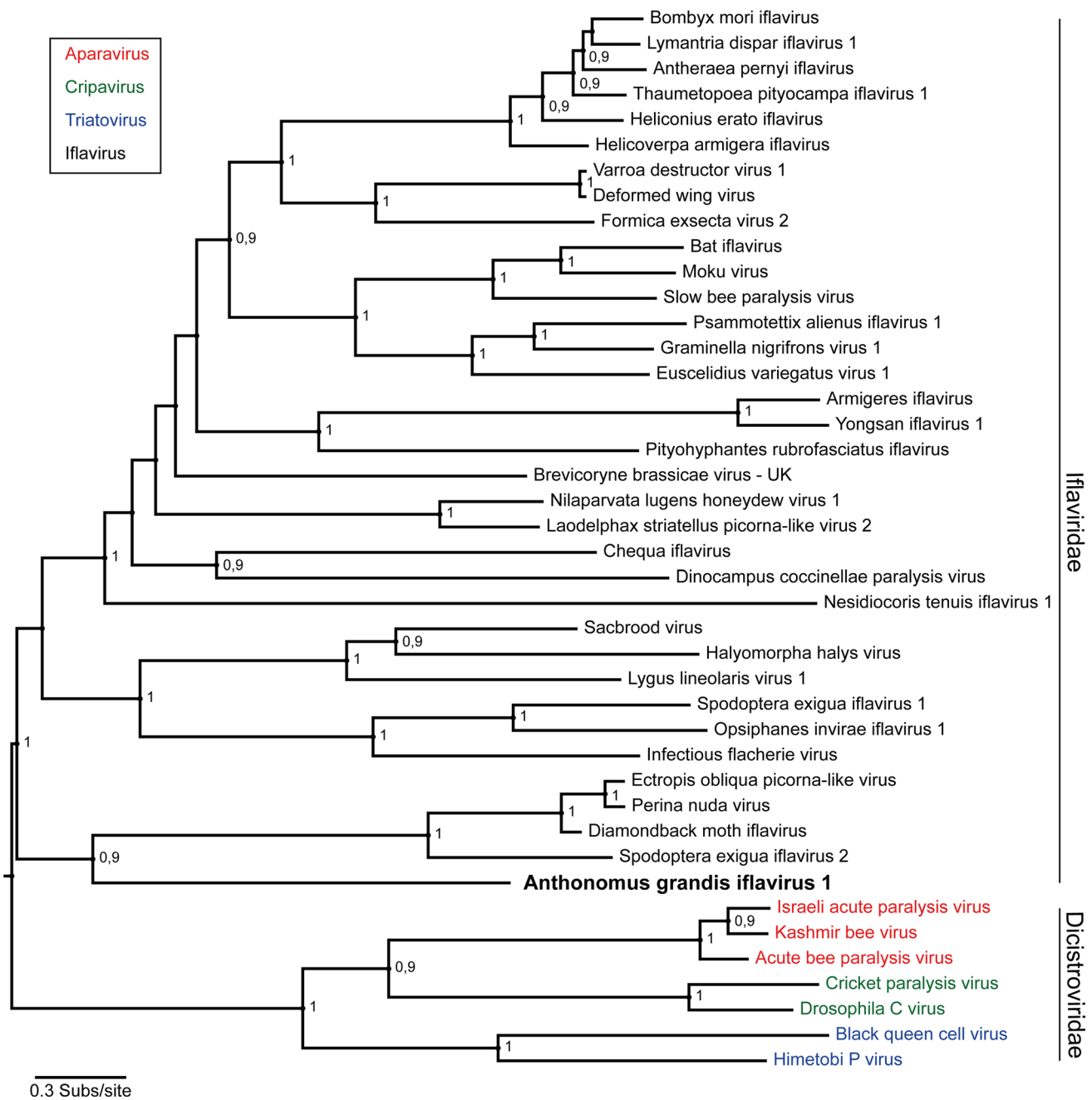


Fig. 2 Phylogenetic analysis of the new picorna-like virus, tentatively named "*Anthonomus grandis iflavirus 1*" (AgIV-1), identified in cotton boll weevil (*Anthonomus grandis*) and other viruses belonging to the families *Iflaviridae* and *Dicistroviridae*. The midpoint-rooted

maximum-likelihood phylogenetic tree was constructed based on whole polyprotein sequences, using the FastTree method implemented in Geneious 11.1.5 software. The branch support was estimated using a Shimodaira-Hasegawa-like test.

Shimodaira-Hasegawa-like test. According to Silva et al. [23], iflaviruses do not form separate clades according to the insects they infect, suggesting that they did not follow the same evolutionary path as their insect hosts at the order level [23, 24]. The results of this study corroborate this observation, and AgIV-1 can be seen to be ancestral within a clade with three iflaviruses found in lepidopteran hosts (Fig. 2).

The identification of new viruses infecting cotton boll weevil increases our knowledge of their diversity and evolution and provides new information that could lead to the development of biotechnological tools for its control. Few iflaviruses have been associated with symptoms or acute disease [25, 26], with deformed wing virus of honey bees being an exception [27]. Future studies with this new picorna-like

virus will be focused on its prevalence, infectivity, and possible use as a viral vector to carry interfering RNA for biological control strategies.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00705-022-05672-8>.

Funding This research was supported by grants provided by Fundação de Apoio à Pesquisa do Distrito Federal (FAP-DF, grant number 193.001532/2016) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, grant number 305756/2017-6).

Declarations

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

References

- Dados. <https://www.abrapa.com.br/Paginas/Dados.aspx>. Accessed 28 Sep 2021
- Cotton: World Markets and Trade | USDA Foreign Agricultural Service. <https://www.fas.usda.gov/data/cotton-world-markets-and-trade>. Accessed 28 Sep 2021
- Conab - Boletim da Safra de Grãos. <https://www.conab.gov.br/info-agro/safras/graos/boletim-da-safra-de-graos>. Accessed 28 Sep 2021
- Trapero C, Wilson IW, Stiller WN, Wilson LJ (2016) Enhancing integrated pest management in GM cotton systems using host plant resistance. *Front Plant Sci*. <https://doi.org/10.3389/FPLS.2016.00500>
- Artico S, Lambret-Frotté J, Nardeli SM et al (2013) Isolation and characterization of three new promoters from *Gossypium hirsutum* that show high activity in reproductive tissues. *Plant Mol Biol Report* 323(32):630–643. <https://doi.org/10.1007/S11105-013-0674-0>
- Santos RC, Marcellino LH, Monnerat RG, Gander ES (2003) Mechanical damage in cotton buds caused by the boll weevil. *Pesqui Agropecuária Bras* 38:1351–1356. <https://doi.org/10.1590/S0100-204X2003001100015>
- Showler AT (2009) Relationships of different cotton square sizes to Boll Weevil (Coleoptera: Curculionidae) feeding and oviposition in field conditions. *J Econ Entomol* 98:1572–1579. <https://doi.org/10.1603/0022-0493-98.5.1572>
- Showler AT (2004) Influence of cotton fruit stages as food sources on Boll Weevil (Coleoptera: Curculionidae) fecundity and oviposition. *J Econ Entomol* 97:1330–1334. <https://doi.org/10.1093/JEE/97.4.1330>
- Ribeiro TP, Arraes FBM, Lourenço-Tessutti IT et al (2017) Transgenic cotton expressing Cry10Aa toxin confers high resistance to the cotton boll weevil. *Plant Biotechnol J* 15:997–1009. <https://doi.org/10.1111/pbi.12694>
- Martins W, Ayres C, Lucena W (2007) Genetic diversity of Brazilian natural populations of *Anthonomus grandis* Boheman (Coleoptera: Curculionidae), the major cotton pest in the New World. *Genet Mol Res* 6:23–32
- Pimenta M, Mata RA, Venzon M et al (2016) Survival and preference of cotton boll weevil adults for alternative food sources. *Braz J Biol* 76:387–395. <https://doi.org/10.1590/1519-6984.16214>
- Stadler T, Buteler M (2007) Migration and dispersal of *Anthonomus grandis* (Coleoptera: Curculionidae) in South America. *Rev Soc Entomol Argent* 66:205–217
- Ribeiro TP, Basso MF, de Carvalho MH et al (2019) Stability and tissue-specific Cry10Aa overexpression improves cotton resistance to the cotton boll weevil. *Biotechnol Res Innov* 3:27–41. <https://doi.org/10.1016/J.BIORI.2019.12.003>
- Oliveira-Marra SOD, Guedes RNC, Bastos CS et al (2019) Insecticide resistance and control failure likelihood among populations of the boll weevil (*Anthonomus grandis*) from Mato Grosso (Brazil). *Acta Sci Agron*. <https://doi.org/10.4025/ACTASCIAGRON.V41I11.42714>
- Rolim GG, Arruda LS, Torres JB et al (2019) Susceptibility of cotton Boll Weevil (Coleoptera: Curculionidae) to spinosyns. *J Econ Entomol* 112:1688–1694. <https://doi.org/10.1093/jee/toz066>
- Li D, Liu CM, Luo R et al (2015) MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* 31:1674–1676. <https://doi.org/10.1093/bioinformatics/btv033>
- Kearse M, Moir R, Wilson A et al (2012) Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28:1647–1649. <https://doi.org/10.1093/bioinformatics/bts199>
- Van Oers MM (2010) Genomics and biology of Iflaviruses. *Insect Virol* 20:20
- Valles SM, Chen Y, Firth AE et al (2017) ICTV virus taxonomy profile: iflavirus. *J Gen Virol*. <https://doi.org/10.1099/jgv.0.000757>
- Katoh K (2002) MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* 30:3059–3066. <https://doi.org/10.1093/nar/gkf436>
- Chen Y, Nakashima N, Christian P et al (2012) Virus chapter: Dicistroviridae. *Taxon Viruses* 20:840–845
- Kuhn JH, Jahrling PB (2010) Clarification and guidance on the proper usage of virus and virus species names. *Arch Virol* 155:445–453
- Silva LA, Ardisson-Araujo DMP, Tinoco RS et al (2015) Complete genome sequence and structural characterization of a novel iflavirus isolated from *Opsiphanes invirae* (Lepidoptera: Nymphalidae). *J Invertebr Pathol* 130:136–140. <https://doi.org/10.1016/j.jip.2015.08.001>
- dos Santos ER, Trentin LB, Ecker A et al (2019) An iflavirus found in stink bugs (Hemiptera: Pentatomidae) of four different species. *Virology* 534:72–79. <https://doi.org/10.1016/j.virol.2019.06.002>
- Ghosh RC, Ball BV, Willcocks MM, Carter MJ (1999) The nucleotide sequence of sacbrood virus of the honey bee: an insect picorna-like virus. *J Gen Virol* 80:1541–1549. <https://doi.org/10.1099/0022-1317-80-6-1541>
- Lanzi G, de Miranda JR, Boniotti MB et al (2006) Molecular and biological characterization of deformed wing virus of honeybees (*Apis mellifera* L.). *J Virol* 80:4998. <https://doi.org/10.1128/JVI.80.10.4998-5009.2006>
- de Miranda JR, Genersch E (2010) Deformed wing virus. *J Invertebr Pathol* 103:S48–S61. <https://doi.org/10.1016/j.jip.2009.06.012>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.