

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

Leonardo Assis Assis da Silva

University of Brasilia Institute of Biological Sciences: Universidade de Brasilia Instituto de Ciencias Biologicas

Marcos Fernando Basso

IMAmt: Instituto Mato-Grossense do Algodao

Bergmann Morais Ribeiro (✉ bergmann@unb.br)

University of Brasília <https://orcid.org/0000-0002-1061-196X>

Research Article

Keywords:

Posted Date: September 19th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1975972/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Archives of Virology on January 4th, 2023.

See the published version at <https://doi.org/10.1007/s00705-022-05672-8>.

Abstract

Cotton boll weevil (CBW; *Anthonomus grandis*; Coleoptera: Curculionidae), is considered the major cotton insect-pest causing considerable losses in yield and fiber quality. The increase in boll weevil population due to the increasingly inefficient chemical control measures are 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. Herein, we used a metagenomic approach based on RNA deep sequencing to investigate the presence of viruses and coding viral RNA in native apparently healthy adult CBW insects collected in cotton crops at Mato Grosso state, Brazil. Using the Illumina HiSeq 2000 paired-end platform, 138,798 virus-related reads were obtained and a consensus sequence of a putative new virus with 10,632 nucleotides in length. The 5'- and 3'-untranslated sequences (UTR) were obtained by Rapid Amplification of cDNA ends (RACE), followed by Nanopore sequencing. The complete sequence showed a 5'-UTR sequence (1,158 nucleotides), a 3'-UTR (561 nucleotides), and a single ORF of 8,913 nucleotides encoding a large polyprotein was identified. Sequence analysis of the putative polyprotein showed several regions with high sequence homology to structural and non-structural proteins of viruses from the *Iflaviridae* family. Pairwise alignments from the polyprotein amino acid sequence showed higher identity (32.13%) with a partial polyprotein sequence of a putative iflavirus (QKN89051.1) found in samples of wild zoo birds in China. Phylogenetic analyses with full polyprotein sequences from different iflaviruses indicate that this new picorna-like virus is more closely related to iflaviruses found in lepidopteran insects. Thus, this new picorna-like virus was tentatively named *Anthonomus grandis* Iflavirus 1 (AgIV-1). This is to our knowledge the first complete virus sequence found in CBW that could be a beginning for further studies about its infectivity, transmission, and possible symptoms or acute disease. The AgIV-1 might be used to develop biological or molecular tools such as a viral vector to carry RNA interfering molecules for CBW control.

Full Text

Running

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 exhibits the highest incidence in cotton crops during the transition period ranging from flowering to fructification [5, 6]. CBW adults feed on and lay eggs within the cotton reproductive structures, causing often flower bud abortion [7, 8]. Since CBW are endophytic, its larvae can cause damage to flower buds when they are not aborted, impacting fiber quality [6, 9]. Its high reproductive capacity, plasticity and genetic variability associated with the occurrence of crop residues or stumps in cotton crops helped to increase the incidence, density and geographic distribution of CBW worldwide [6, 10–12]. Until now, no conventional or transgenic cotton cultivar with satisfactory resistance to CBW are available to cotton producers.

Consequently, numerous insecticide applications have been used annually for its management [13]. Unfortunately, the frequent occurrence of CBW populations with reduced susceptibility to insecticides and failure in chemical control has already been reported in Brazilian cotton crops [14, 15]. Meanwhile, the identification of new viruses infecting CBW may provide information that might result in the development of molecular or biological tools for their effective control.

Herein, we used an RNA-sequencing approach through next-generation sequencing (NGS) to investigate the presence of viruses and coding viral RNA in native, apparently healthy, adult CBW insects collected in September 2020 on 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 a mortar and pestle containing SM buffer (100 mM NaCl, 8 mM MgSO₄, 50 mM Tris-Cl, pH 7.5). The homogenate was filtered once through cheesecloth and centrifuged three times at 4,000 × *g* for 10 min for supernatant clarification. Then, clarified supernatant was used for RNA extracted using QIAamp Viral RNA Mini Kit (QIAGEN, Hilden, Germany) according to the manufacturer's instructions. Total RNA sample was processed for rRNA removal using the Ribo-Zero rRNA removal kit (Illumina, San Diego, CA, USA) and cDNA library was constructed using the 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 *de novo* assembled using MEGAHIT software [16]. The resulting contigs closely related to viruses were retrieved using BLASTx against an *in house* viral RefSeq database. 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]. As well, genome annotation was also performed using Geneious 11.1.5 software, whereas open reading frames (ORFs) were confirmed using BLASTx search against the NCBI non-redundant protein database (08/2021).

The NGS resulted in 56,210,608 total reads of which 138,798 were considered as virus-related sequences. The *de novo* assembling of these viral reads generated a consensus sequence with 10,632 nucleotides in length (Genbank accession number OK413669, Supplementary Material S1). The genome coverage was of 1,440X. A single ORF of 8,913 nucleotides encoding a large polyprotein was predicted, additional to a 5'-UTR sequence (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 the MinION platform with the Rapid RBK110.96 kit, following the manufacturer's instructions (Oxford Nanopore Technologies) and the sequences analyzed using the Geneious 11.1.5 software. We identified by sequences alignment functional ORFs flanked by putative proteolytic sites such as structural proteins, and the non-structural proteins (Fig. 1). From this data, virus genome organization clearly resembled other members of the family *Iflaviridae* [18, 19].

From the amino acid sequence alignment using the MAFFT method [20], this putative amino acid polyprotein showed 32.13% identity with a sequence of a putative iflavirus (QKN89051.1) found in

samples of wild zoo birds in China. According to the International Committee on Taxonomy of Viruses (ICTV) [19], iflaviruses presents two demarcation criteria for new species recognition into genus, such as: (1) natural host range and (2) sequence identity at the amino acid level of the capsid proteins under 90% [21, 22]. Therefore, we concluded that this new picorna-like virus describe here accomplishes all ICTV requirements to be recognized as new specie belonging to the *Iflavirus* genus. So, we have tentatively named this putative new virus as Anthonomus grandis Iflavirus 1 (AgIV-1).

Phylogenetic analysis with the whole genome sequence of this putative novel virus was compared to other iflaviruses (Table S1). Sequence alignments were carried out using the MAFFT method [20] with whole polyprotein-coding genes with highest identity (Fig. 2). A maximum-likelihood tree was inferred using the Fast-tree method implemented into Geneious 11.1.5 software [17] while the branch support was estimated by a Shimodaira-Hasegawa-like test. According to Silva et al. [2015], iflaviruses do not form a single clade according to the insect-infected order, suggesting that they did not follow the same evolutionary path as their insect hosts at order level [23, 24]. Our results corroborate this observation, where we can see that AgIV-1 is ancestral within a clade with three iflaviruses found in lepidopteran hosts (Fig. 2).

Given this, the identification of new viruses in this insect-pest increases our knowledge of their diversity, evolution, and provides new information to the development of biotechnological tools for its control. Until now few iflaviruses were associated with symptoms or acute disease [25, 26], except the deformed wing virus of honey bees [27]. Future studies with this new picorna-like virus identified in CBW will be focused on virus prevalence, infectivity, and possible use of this virus as a viral vector to carry RNA interfering strategies for CBW biological control.

Declarations

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

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

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).

References

1. Dados. <https://www.abrapa.com.br/Paginas/Dados.aspx>. Accessed 28 Sep 2021
2. Cotton: World Markets and Trade | USDA Foreign Agricultural Service. <https://www.fas.usda.gov/data/cotton-world-markets-and-trade>. Accessed 28 Sep 2021

3. Conab - Boletim da Safra de Grãos. <https://www.conab.gov.br/info-agro/safras/graos/boletim-da-safra-de-graos>. Accessed 28 Sep 2021
4. Trapero C, Wilson IW, Stiller WN, Wilson LJ (2016) Enhancing Integrated Pest Management in GM Cotton Systems Using Host Plant Resistance. *Front Plant Sci* 0:500. <https://doi.org/10.3389/FPLS.2016.00500>
5. 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* 2013 323 32:630–643. <https://doi.org/10.1007/S11105-013-0674-0>
6. 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>
7. 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>
8. 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>
9. 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>
10. 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
11. Pimenta M, Mata RA, Venzon M, et al (2016) Survival and preference of cotton boll weevil adults for alternative food sources. *Brazilian J Biol* 76:387–395. <https://doi.org/10.1590/1519-6984.16214>
12. Stadler T, Buteler M (2007) Migration and dispersal of *Anthonomus grandis* (Coleoptera: Curculionidae) in South America. *Rev la Soc Entomológica Argentina* 66:205–217
13. Ribeiro TP, Basso MF, Carvalho MH de, 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>
14. 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* 41:. <https://doi.org/10.4025/ACTASCIAGRON.V41I1.42714>
15. 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>
16. 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>

17. 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>
18. Van Oers MM. (2010) Genomics and biology of I flaviruses. *Insect Virol*
19. Valles SM, Chen Y, Firth AE, et al (2017) ICTV virus taxonomy profile: I flaviridae. *J Gen Virol*. <https://doi.org/10.1099/jgv.0.000757>
20. 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>
21. Chen Y, Nakashima N, Christian P, et al (2012) Virus Chapter: Dicistrovidae. *Taxon Viruses* 840–845
22. Kuhn JH, Jahrling PB (2010) Clarification and guidance on the proper usage of virus and virus species names. *Arch Virol* 155:445–453
23. 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>
24. 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>
25. Ghosh RC, Ball B V., 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>
26. Lanzi G, Miranda JR de, 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>
27. 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>

Figures



Figure 1

Genome organization of the putative new picorna-like virus, tentatively named here as *Anthonomus grandis I flavivirus 1* (AgIV-1) identified in cotton boll weevil (*Anthonomus grandis*). The viral genome consists of 10,632 nucleotides with a 5'-UTR and 3'-UTR polyadenylated tail, encoding a large polyprotein ORF of 8,913 nucleotides with structural proteins such as VP1 to VP4 located at the N-terminal region and

the non-structural proteins such as helicase, protease, and RNA-dependent RNA polymerase (RdRp) at the C-terminal region. The N-terminal of the polyprotein contains a typical L-peptide.

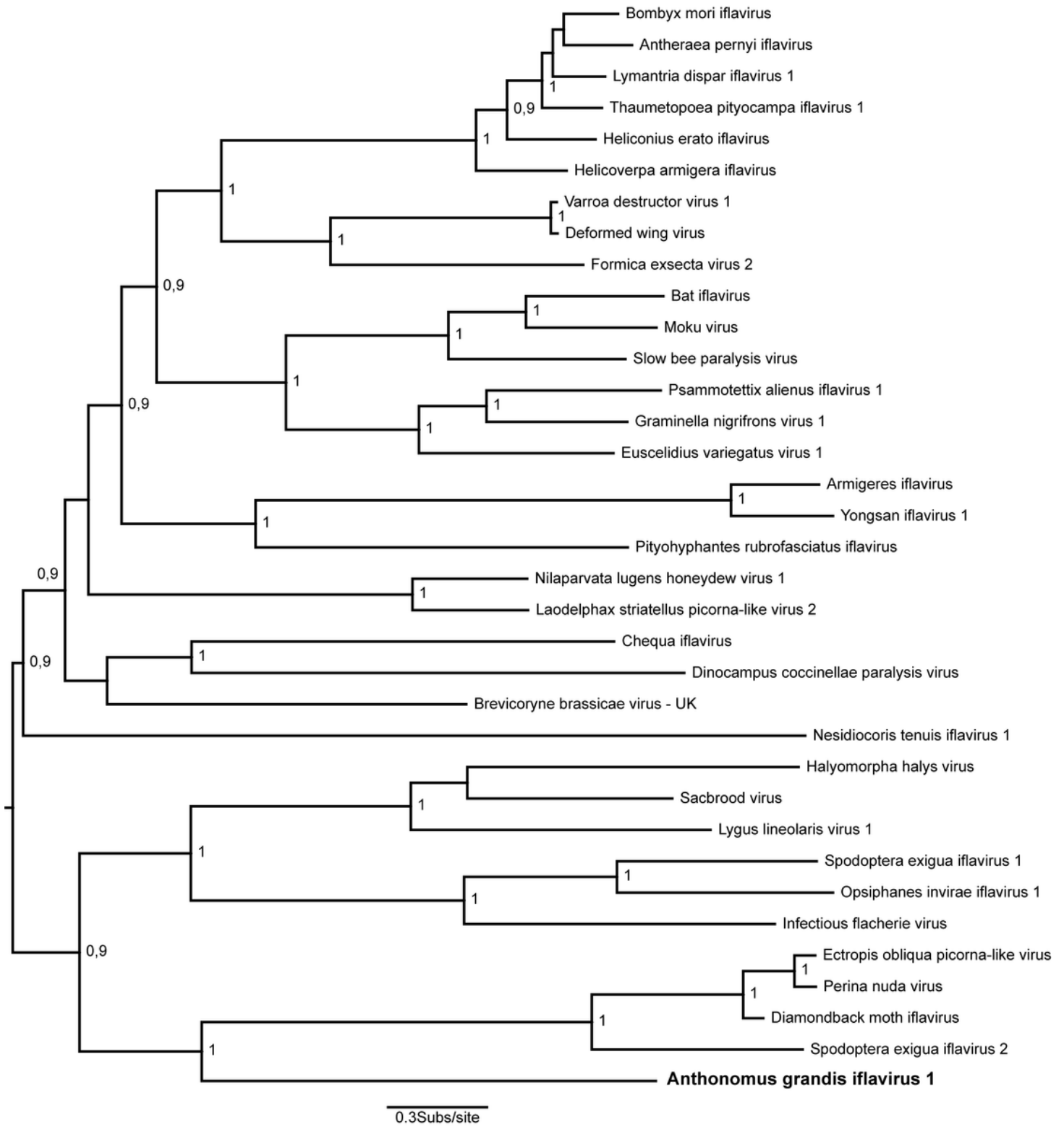


Figure 2

Phylogenetic analysis of the new picorna-like virus, tentatively named here as *Anthonomus grandis iflavirus 1* (AgIV-1), identified in cotton boll weevil (*Anthonomus grandis*), with respect to viruses

belonging to the family *Iflaviridae*. Midpoint rooted phylogenetic tree was constructed with the whole polyprotein, while phylogenetic tree was inferred by maximum-likelihood tree using the Fast-tree method implemented in Geneious 11.1.5 software. The branch support was estimated by a Shimodaira-Hasegawa-like test.

Supplementary Files

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

- [SupplementaryMaterialS1new.docx](#)
- [TableS1.docx](#)