![American Society for Microbiology | Microbiology Resource Announcements logo]

Genome Sequences of Novel Iflaviruses in the Gray Lawn Leafhopper, an Experimental Vector of Corn Stunt Spiroplasma

Sharon A. Andreason,[a] Zachary Lahey,[a] Christian Ayala-Ortiz,[b] Alvin M. Simmons[a]

aUnited States Department of Agriculture, Agricultural Research Service, U.S. Vegetable Laboratory, Charleston, South Carolina, USA

bDepartment of Entomology and Plant Pathology, Oklahoma State University, Stillwater, Oklahoma, USA

Sharon A. Andreason and Zachary Lahey contributed equally to this work. Author order was determined alphabetically.

ABSTRACT   Two novel iflaviruses were detected in the metatranscriptome of the gray lawn leafhopper, *Exitianus exitiosus* (Uhler). The assembled genome sequence of *Exitianus exitiosus virus 1* was 9,858 nucleotides (nt) long and encodes a 3,083-amino acid (aa) polyprotein. *Exitianus exitiosus virus 2* was 10,219 nt long and encodes a 2,947-aa polyprotein.

The gray lawn leafhopper, *Exitianus exitiosus* (Uhler) (Hemiptera: Cicadellidae), is a vector of maize chlorotic dwarf virus (MCDV) (1) and an experimental vector of *Spiroplasma kunkelii* (2). An earlier study on the transcriptional response of *E. exitiosus* to the acquisition of *S. kunkelii* revealed transcripts with homology to iflaviruses (*Iflaviridae*: *Iflavirus*) (3). Subsequent sequencing of *E. exitiosus* from the same colony confirmed the presence of two novel iflavirus genomes, *Exitianus exitiosus virus 1* (EeV1) and *Exitianus exitiosus virus 2* (EeV2).

Exitianus exitiosus was originally collected from Bermuda grass in Stillwater, OK, in September 2014 and reared on a mix of cereal grasses. Total RNA was extracted from adults using the E.Z.N.A. mollusk RNA kit (Omega Bio-Tek Inc., Norcross, GA). High-quality total RNA (RNA integrity number [RIN], >8) was used to prepare 150-bp paired-end TruSeq RNA libraries, which were sequenced on an Illumina HiSeq 4000 instrument by Novogene Co., Ltd. (Beijing, China). Fifty million paired-end reads were sequenced (SRA accession number SRR15166997). The *E. exitiosus* metatranscriptome was assembled *de novo* using Trinity v.2.13.2 (4), following Chen et al. (5). Open reading frames were determined using TransDecoder v.5.5.0 (https://github.com/TransDecoder/TransDecoder). A BLASTP search of the translated transcripts against the NCBI nonredundant (nr) protein sequences database identified transcripts with homology to a *Scaphoideus titanus iflavirus 1* polyprotein (GenBank accession number QIJ56901.1), which corresponded to EeV1, and an *Exitianus capicola iflavirus 2* polyprotein (GenBank accession number ULR75443.1), corresponding to EeV2.

The EeV1 draft genome was sequenced to 188× coverage, with 12,350 high-quality reads mapping to the viral sequence. EeV1 has a G+C content of 36.7% and comprises a 5′ untranslated region (UTR; 362 nucleotides [nt] assembled), a 247-nt 3′ UTR with a poly(A) tail, and a 9,249-nt coding sequence (CDS). The 5′ UTR forms several stem-loop structures characteristic of an internal ribosome entry site (IRES), a hallmark of the genus *Iflavirus*. The CDS contains a single open reading frame (ORF) that encodes a 3,083-amino acid (aa) polyprotein (351.2 kDa). A BLASTP search of the EeV1 polyprotein identified *Scaphoideus titanus iflavirus 1* as the best hit (sequence identity, 53.7%; query coverage, 99%). Annotation of the EeV1 polyprotein using InterProScan v.5.55-88.0 (6) recovered four conserved domains for viral capsid proteins at the N terminus (two rhv_like, one Rhv, and one CPRV_capsid) and three conserved domains at the C terminus (RNA_helicase, Peptidase_C3, and RdRp_1).

Exitianus exitiosus virus 2 was sequenced to 1,484× coverage (101,134 reads) and has a G+C content of 35.1%. The 10,219-nt EeV2 draft genome comprises a 5′ UTR (1,205 nt
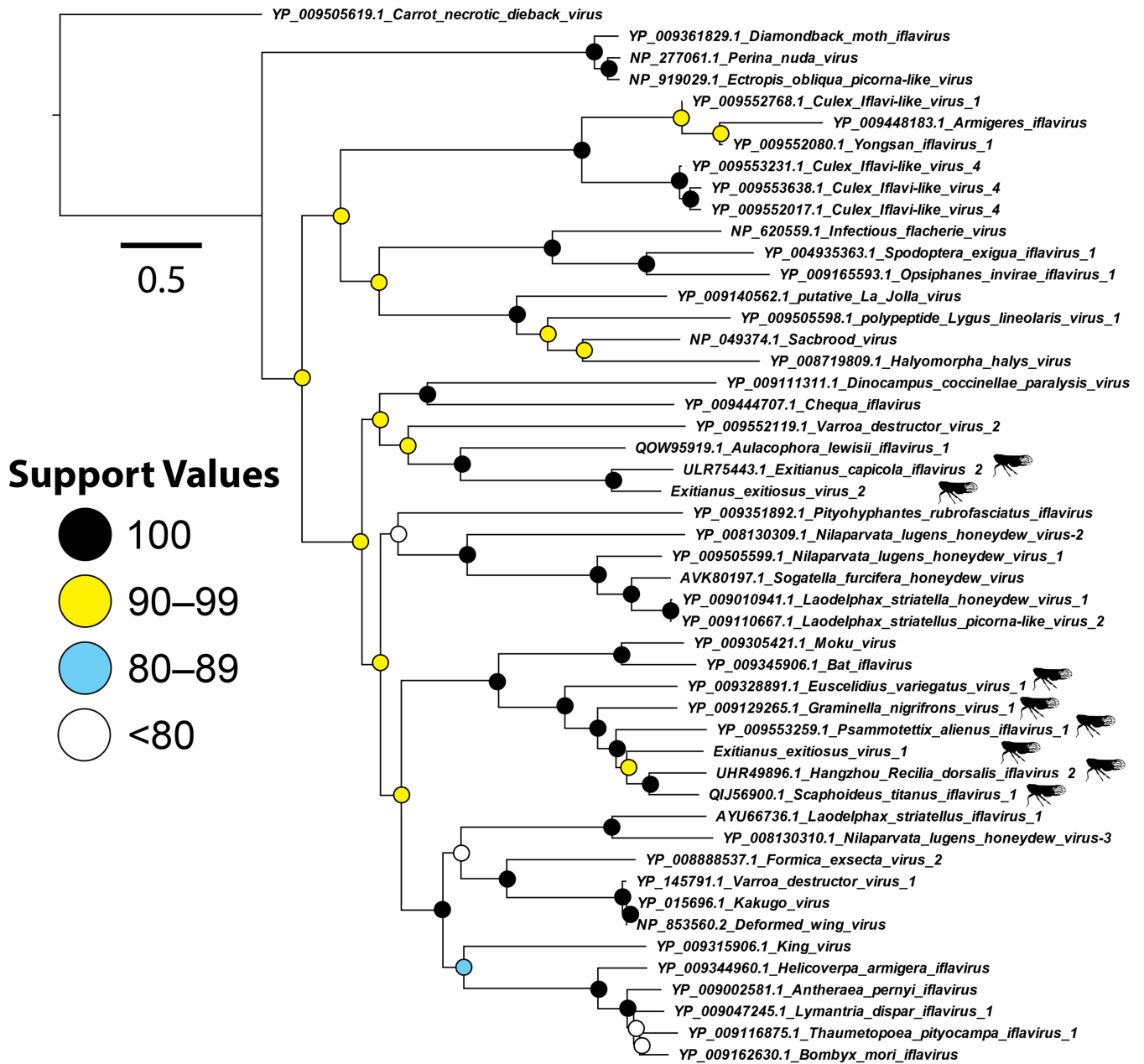
FIG 1 Maximum likelihood phylogenetic analysis of reference and putative iflavirus polyprotein sequences (amino acids) retrieved on 4 February 2022 from NCBI GenBank. The tree depicted is the best tree (lowest log-likelihood) from 10 independent runs. Sequences were aligned with MAFFT v.7.490 using the FFT-NS-i algorithm. The phylogenetic analysis was performed using IQ-TREE v.2.1.3, the Q.pfam+F+R6 substitution model, and *Carrot necrotic dieback virus* (*Secoviridae*) as the outgroup. Branch support was measured with 1,000 ultrafast bootstrap replicates and the –bnni flag to reduce the impact of potential model violations. Branch support is depicted as colored circles at the nodes. Viruses discovered in other deltocephaline leafhoppers are indicated by leafhopper icons next to the virus names.

assembled), a 173-nt 3′ UTR with a poly(A) tail, and an 8,841-nt CDS. The single ORF encodes a 2,947-aa polyprotein (336.5 kDa). A BLASTP of the EeV2 polyprotein identified *Exitianus capicola iflavirus 2* as the best hit (sequence identity, 60.7%; query coverage, 84%). Annotation of the EeV2 polyprotein recovered all but one (Rhv) of the conserved domains identified in EeV1.

The homology of the two iflaviruses was assessed by constructing a phylogenetic tree (7–9) based on the polyprotein sequences of classified and unclassified Iflaviridae (Fig. 1). The effects of EeV1 and EeV2 on host fitness remain to be explored.

Data availability. The *E. exitiosus* transcriptome shotgun assembly project has been deposited at DDBJ/EMBL/GenBank under the accession number GJYH00000000.1. The version described in this paper is the first version, GJYH01000000. The *E. exitiosus* reads have

been deposited in the NCBI Sequence Read Archive under the accession number SRR15166997. The genome sequences of *Exitianus exitiosus virus 1* and *Exitianus exitiosus virus* 2 have been deposited at GenBank under the accession numbers ON402368 and ON402369, respectively.

ACKNOWLEDGMENTS

REFERENCES

1. Nault LR, Madden L. 1988. Phylogenetic relatedness of maize chlorotic dwarf virus leafhopper vectors. Phytopathology 78:1683–1687. https://doi.org/10.1094/Phyto-78-1683.
2. Nault LR. 1980. Maize bushy stunt and corn stunt: a comparison of disease symptoms, pathogen host ranges, and vectors. Phytopathology 70:659. https://doi.org/10.1094/Phyto-70-659.
3. Andreason SA. 2016. Discrimination and transcriptional response analysis of hemipteran phytopathogen vectors. ProQuest. PhD dissertation. Oklahoma State University, Stillwater, OK. https://www.proquest.com/docview/1841269071. Accessed 7 March 2022.
4. Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, Couger MB, Eccles D, Li B, Lieber M, MacManes MD, Ott M, Orvis J, Pochet N, Strozzi F, Weeks N, Westerman R, William T, Dewey CN, Henschel R, LeDuc RD, Friedman N, Regev A. 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat Protoc 8:1494–1512. https://doi.org/10.1038/nprot.2013.084.
5. Chen H, Lahey Z, Talamas EJ, Johnson NF. 2021. Identification and expression of chemosensory receptor genes in the egg parasitoid *Trissolcus basalis*.

Comp Biochem Physiol Part D Genomics Proteomics 37:100758. https://doi.org/10.1016/j.cbd.2020.100758.
6. Jones P, Binns D, Chang H-Y, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G, Pesseat S, Quinn AF, Sangrador-Vegas A, Scheremetjew M, Yong S-Y, Lopez R, Hunter S. 2014. InterProScan 5: genome-scale protein function classification. Bioinformatics 30:1236–1240. https://doi.org/10.1093/bioinformatics/btu031.
7. Hoang DT, Chernomor O, Von Haeseler A, Minh BQ, Vinh LS. 2018. UFBoot2: improving the ultrafast bootstrap approximation. Mol Biol Evol 35:518–522. https://doi.org/10.1093/molbev/msx281.
8. Kalyaanamoorthy S, Minh BQ, Wong TK, Von Haeseler A, Jermiin LS. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods 14:587–589. https://doi.org/10.1038/nmeth.4285.
9. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, Von Haeseler A, Lanfear R. 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol Biol Evol 37:1530–1534. https://doi.org/10.1093/molbev/msaa015.