AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Transcriptome Sequencing of Immature Ants Reveals the Complete Genome Sequence of a New Isolate of *Solenopsis invicta Virus 2* from the Mississippi Delta

Margaret L. Allen[a]

[a]Biological Control of Pests Research Unit, National Biological Control Laboratory, U.S. Department of Agriculture-Agricultural Research Service, Stoneville, Mississippi, USA

ABSTRACT *Solenopsis invicta virus 2* (SINV-2) is an RNA virus that infects red imported fire ants. I report the genome sequence of SINV-2MSD, an isolate infecting ants collected from Mississippi. The obtained genome is 11,303 nucleotides, including six open reading frames encoding four structural proteins, a helicase, and an RNA-dependent RNA polymerase.

Invasive red imported fire ants (RIFA), *Solenopsis invicta* (Buren), plague much of the world and have proven to be challenging and expensive to control (1–3). Ant viruses have been sought for potential use as biological control agents against invasive ants, including RIFA (4, 5). *Solenopsis invicta virus 2* (SINV-2) is a positive-sense single-stranded picorna-like RNA virus. A proposal to establish a new family of polycistronic picorna-like RNA viruses, *Polycipiviridae*, has been submitted to the International Committee on Taxonomy of Viruses (ICTV) (6). The species exemplar of the virus was isolated from RIFA collected in Florida (7). Transcriptome sequencing of the juvenile ants revealed the genome sequence of a unique SINV-2 isolate, designated SINV-2MSD (where "MSD" is Mississippi Delta [Washington County, MS]).

Six transcriptomes of RIFA at juvenile stage were prepared with a primary aim to identify RNA interference targets for future invasive ant pest control research, as described previously (8). Briefly, total RNA was extracted from three independent ant colonies. Extractions were made from either 20 larvae or 20 pupae using USB PrepEase (USB Corporation, Cleveland, OH, USA), according to the manufacturer's instructions. Samples were qualified on a NanoDrop spectrophotometer and Bioanalyzer, and rRNA was depleted using RiboZero Gold (human/mouse/rat). Libraries were prepared according to the Illumina TruSeq stranded total RNA sample preparation guide. Sequencing was performed on the Illumina HiSeq 2000 platform in the 100-bp paired-end (PE) configuration, according to the manufacturer's instructions. Cutadapt (9) was used to remove the reads that contained adapter contamination and low-quality or unresolved bases. Sequence quality was verified using FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). To accomplish our primary differential expression goals (8), raw reads were mapped to the genome of *S. invicta* using TopHat v2.1.0 and assembled using Cufflinks v2.2.1 (transcriptome preparation by LC Sciences, Houston, TX). Then, to further mine the transcriptomes, with an aim to reveal the microbial ecology associated with juvenile ants, individual *de novo* assemblies of the sequences were completed with the LaserGene 14 software (SeqMan NGen and SeqMan Pro software; DNAStar, Inc., Madison, WI), using RefSeq Invertebrate (transcript annotation package downloaded from the DNAStar website) genes as guides. The unassembled reads from each of these six assemblies were then resubmitted for assembly, again using the SeqMan NGen software and using microbial (viral, bacterial, fungal, and protozoan) data files for reference guidance. Each assembly contained long assembled contigs matching the

Citation Allen ML. 2019. Transcriptome sequencing of immature ants reveals the complete genome sequence of a new isolate of *Solenopsis invicta virus 2* from the Mississippi Delta. Microbiol Resour Announc 8:e01115-18. https://doi.org/10.1128/MRA.01115-18.

Editor Jelle Matthijnssens, KU Leuven

Address correspondence to meg.allen@ars.usda.gov.

Received 20 August 2018
Accepted 12 April 2019
Published 2 May 2019

TABLE 1 Assembly statistics for individual sample assemblies using reference-based workflow and the SINV-2 genome with accession number NC_009544 as the reference

	Data by Bam file[a]					
Statistic	**A01L4toSinV2-0.bam**	**B02L4toSinvV2-0.bam**	**C03L4toSinV2again-0.bam**	**X01wptoSinvV2-0.bam**	**Y02wptoSinvV2-0.bam**	**Z03wptoSinvV2-0.bam**
Total no. of reads in unassembled file	104,172,889	99,193,824	84,562,843	103,025,940	111,623,447	75,041,763
Contig length, with gaps (bases)	11,320	11,322	12,235	11,319	11,320	11,322
Contig length, without gaps (bases)	11,319	11,319	11,319	11,319	11,319	11,319
Avg length/sequence (bases)	93	94	94	93	92	96
Total sequence length (bases)	518,588	716,308	439,686,475	363,846	484,675	1,461,106
Top strand (no. of sequences)	2,316	3,583	2,797,837	1,419	1,921	7,707
Bottom strand (no. of sequences)	3,253	4,001	1,866,336	2,486	3,335	7,357
Total no. of sequences	5,569	7,584	4,664,173	3,905	5,256	15,064
Median coverage (×)	46.19	68.89	151,514.17	29.37	40.76	221.74
Median observed pair distance (bp)	151	155	157	151	150	171

[a] The file names designate the colony origin (01, 02, or 03) and the insect stage (L4, late-instar larva; wp, white pupa). The sample C03L4 assembly was repeated because of the much higher coverage (indicated by file name "...again..."), and the second assembly run matched the first.

genome of SINV-2. To obtain full-genome sequences, each of the six paired-end Illumina file sets was again assembled, mapping reads to the NCBI genome of SINV-2 (GenBank accession number NC_009544 [now NC_039236]). Each assembly yielded a complete viral genome encoding all six of the predicted protein-encoding regions and terminating in a 16-nucleotide polyadenylation sequence (Table 1). Multiple alignment using MegAlign (DNAStar, Inc.) verified that the sequences were identical. The consensus genome differed from the GenBank NC_009544 reference by lacking the first 11 nucleotides of the 5′ untranslated region, having two gaps, with 1 nucleotide at position 4452, just upstream of the fourth open reading frame (ORF), and one gap of 6 nucleotides at positions 11296 to 11301 just upstream of the polyadenylation tail. Also, isolate SINV-2MSD had 203 nucleotide differences and amino acid changes, as follows: 2 substitutions in ORF1, 1 substitution in ORF2, 6 substitutions in ORF3, 3 substitutions in ORF4, and 14 substitutions in ORF5. This finding demonstrates the prevalence of the recently identified SINV-2 throughout the expansion range of the host, RIFA, as well as the presence of regional mutations of this novel RNA virus.

For all software cited, default settings were used except as noted.

Data availability. The complete genome sequence of SINV-2MSD was deposited in GenBank under the accession number MG676340; the transcriptomes are deposited under BioProject number PRJNA393960, BioSample numbers SAMN07345557 and SAMN07345558, and Sequence Read Archive numbers SRR5868358, SRR5868359, SRR5868360, SRR5868361, SRR5868362, and SRR5868363.

ACKNOWLEDGMENTS

I thank Jian Chen, Mark Weaver, and an anonymous reviewer for providing comments on a previous draft of this announcement.

Trade, firm, or corporation names in this publication are for the information and convenience of the reader. Such use does not constitute an official endorsement or approval by the U.S. Department of Agriculture of the Agricultural Research Service (USDA-ARS).

REFERENCES

1. Ascunce MS, Yang C-C, Oakey J, Calcaterra L, Wu W-J, Shih C-J, Goudet J, Ross KG, Shoemaker D. 2011. Global invasion history of the fire ant *Solenopsis invicta*. Science 331:1066–1068. https://doi.org/10.1126/science.1198734.
2. Kenis M, Auger-Rozenberg M-A, Roques A, Timms L, Péré C, Cock MJW, Settele J, Augustin S, Lopez-Vaamonde C. 2009. Ecological effects of invasive alien insects. Biol Invasions 11:21–45. https://doi.org/10.1007/s10530-008-9318-y.
3. Wang H, Wang H, Tao Z, Ge Q. 2018. Potential range expansion of the red imported fire ant (*Solenopsis invicta*) in China under climate change. J Geogr Sci 28:1965–1974.
4. Hsu H-W, Chiu M-C, Shoemaker D, Yang C-C. 2018. Viral infections in fire ants lead to reduced foraging activity and dietary changes. Sci Rep 8:13498. https://doi.org/10.1038/s41598-018-31969-3.
5. Valles SM, Porter SD, Calcaterra LA. 2018. Prospecting for viral natural enemies of the fire ant *Solenopsis invicta* in Argentina. PLoS One 13:e0192377. https://doi.org/10.1371/journal.pone.0192377.
6. Olendraite I, Lukhovitskaya NI, Porter SD, Valles SM, Firth AE. 2017. *Polycipiviridae*: a proposed new family of polycistronic picorna-like RNA viruses. J Gen Virol 98:2368–2378. https://doi.org/10.1099/jgv.0.000902.
7. Valles SM, Strong CA, Hashimoto Y. 2007. A new positive-strand RNA virus with unique genome characteristics from the red imported fire ant, *Solenopsis invicta*. Virology 365:457–463. https://doi.org/10.1016/j.virol.2007.03.043.
8. Allen ML, Rhoades JH, Sparks ME, Grodowitz MJ. 2018. Differential gene expression in red imported fire ant (*Solenopsis invicta*) (Hymenoptera: Formicidae) larval and pupal stages. Insects 9:E185. https://doi.org/10.3390/insects9040185.
9. Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J 17:10–12. https://doi.org/10.14806/ej.17.1.200.