Genome Sequence of a Novel Iflavirus from mRNA Sequencing of the Pupa of Bombyx mori Inoculated with Cordyceps militaris

doi: 10.1128/genomeA.01039-15

genomeA
Journals.ASM.org

Genome Sequence of a Novel Iflavirus from mRNA Sequencing of the Pupa of *Bombyx mori* Inoculated with *Cordyceps militaris*

Tomohiro Suzuki,[a,b] Yoshino Takeshima,[c] Toshiyuki Mikamoto,[d] Jun-David Saeki,[d] Tatsuya Kato,[b,c,e] Enoch Y. Park,[b,c,e] Hirokazu Kawagishi,[b,c,e] Hideo Dohra[b]

Center for Bioscience Research and Education, Utsunomiya University, Utsunomiya, Tochigi, Japan[a]; Research Institute of Green Science and Technology, Shizuoka University, Suruga-ku, Shizuoka, Japan[b]; Department of Applied Biological Chemistry, Faculty of Agriculture, Shizuoka University, Suruga-ku, Shizuoka, Japan[c]; Nichihara Research & Development Laboratories, Inc., Tsuwano, Kanoashi, Shimane, Japan[d]; Graduate School of Science and Technology, Shizuoka University, Suruga-ku, Shizuoka, Japan[e]

We discovered a novel iflavirus from the transcriptome of the *Bombyx mori* pupa inoculated with the insect-pathogenic fungus *Cordyceps militaris*. The assembled iflavirus genome has 10,119 nucleotides, with a 3′-polyadenylated tail, and it encodes a polyprotein composed of 3,004 amino acids.

Received 2 August 2015 **Accepted** 7 August 2015 **Published** 17 September 2015

Citation Suzuki T, Takeshima Y, Mikamoto T, Saeki J-D, Kato T, Park EY, Kawagishi H, Dohra H. 2015. Genome sequence of a novel iflavirus from mRNA sequencing of the pupa of *Bombyx mori* inoculated with *Cordyceps militaris*. Genome Announc 3(5):e01039-15. doi:10.1128/genomeA.01039-15.

Address correspondence to Hirokazu Kawagishi, kawagishi.hirokazu@shizuoka.ac.jp, or Hideo Dohra, dora.hideo@shizuoka.ac.jp.

The domestic silkworm *Bombyx mori* is well studied and has been of interest due to its excellent characteristic as a textile fiber; recently, it has also been used as a host for the expression of eukaryotic proteins. In this study, the pupa of *B. mori* was used as a host of the insect-pathogenic fungus *Cordyceps militaris*. The RNA sequencing (RNA-seq) analysis of *B. mori* inoculated with *C. militaris* detected an iflavirus genome sequence. Iflaviruses are positive-sense single-stranded RNA viruses that infect insect hosts, such as moths and butterflies (*Lepidoptera*), honey bees and ants (*Hymenoptera*), and brown planthoppers and aphids (*Hemiptera*) (1–3). Several iflaviruses are known to have pathogenicity to their hosts, often leading to diarrhea, developmental malformations, and death of the host (2).

We sequenced the mRNA of *B. mori* (Kinsyu-Showa strain) (Nichihara Research & Development Laboratories, Inc.) inoculated with *C. militaris* (NBRC 100741). Total RNA was extracted from three male samples of *B. mori* using the TRIzol reagent (Life Technologies) and further purified using the RNeasy plant mini-kit (Qiagen). Strand-specific RNA sequencing libraries were prepared using the SureSelect strand-specific RNA library prep kit (Agilent Technologies) and sequenced using an Illumina MiSeq sequencer. In total, 11.7 million 76-bp paired-end reads were generated, of which 238,946 reads (2.05%) were derived from the iflavirus genome. The raw reads were cleaned up with cutadapt (4) by trimming adapter sequences and low-quality ends (quality score, <30), discarding reads <50 bp, and with the FASTX-Toolkit (5) by trimming the last 76 bases, resulting in 236,539 paired-reads totaling approximately 35.3 Mb. The cleaned reads were *de novo* assembled using Trinity (6). The assembled single contig has 10,119 nucleotides, with a G+C content of 38.7%, terminates in a 3′-polyadenylated tail, and shows an average 3,486× coverage of the total length of the iflavirus genome. As a result of the annotation by Prokka (7), the iflavirus genome was found to encode a polyprotein composed of 3,004 amino acids. The BLASTp search to the NCBInr protein database for the polyprotein showed 73% amino acid sequence identity with that of the gypsy moth *Lymantria dispar* iflavirus 1 (3), but it did not show such high similarity (23% identical) with that of the same host *B. mori* infectious flacherie virus (1). These results and the phylogenetic analysis (data not shown) of the amino acid sequences of the polyproteins suggest that the iflavirus is a novel species of the genus *Iflavirus*. The iflavirus was scarcely detected in hot-air-drying pupae, suggesting that the iflavirus genome was degraded at a high temperature.

Here, we report the genome sequence of a novel iflavirus detected from the pupa of *B. mori* inoculated with *C. militaris*. However, little is known about interactions between the iflavirus and *C. militaris* in a host pupa. It might be a good target for future studies to elucidate the effects of the iflavirus on developmental stages of *C. militaris*, such as infection to a pupa, reproduction, and fruiting body formation.

Nucleotide sequence accession number. The genome sequence has been deposited in DDBJ under the accession no. LC068762. The version described in this paper is the first version.

ACKNOWLEDGMENT

This work was supported by the Functional Genomics Section, Research Institute of Green Science and Technology, Shizuoka University.

REFERENCES

1. **Isawa H, Asano S, Sahara K, Iizuka T, Bando H.** 1998. Analysis of genetic information of an insect picorna-like virus, infectious flacherie virus of silkworm: evidence for evolutionary relationships among insect, mammalian and plant picorna(-like) viruses. Arch Virol **143:**127–143. http://dx.doi.org/10.1007/s007050050273.
2. **Oers MM.** 2010. Genomics and biology of iflaviruses, p 231–250. *In* Asgari S, Johnson KN (ed), Insect virology. Caister Academic Press, Norfolk, United Kingdom.
3. **Carrillo-Tripp J, Krueger EN, Harrison RL, Toth AL, Miller WA, Bonning BC.** 2014. *Lymantria dispar* iflavirus 1 (LdIV1), a new model to study iflaviral persistence in lepidopterans. J Gen Virol **95:**2285–2296.

4. **Martin M.** 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J **17:**10–12. http://dx.doi.org/10.14806/ej.17.1.200.
5. **Pearson WR, Wood T, Zhang Z, Miller W.** 1997. Comparison of DNA sequences with protein sequences. Genomics **46:**24–36. http://dx.doi.org/10.1006/geno.1997.4995.
6. **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. http://dx.doi.org/10.1038/nbt.1883.
7. **Seemann T.** 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics **30:**2068–2069. http://dx.doi.org/10.1093/bioinformatics/btu153.