Complete genome analysis of a novel chuvirus from a southern green stink bug (Nezara viridula)

Peng-Peng Ren1 · Zhuang-Xin Ye1 · Sai-Nan Wang1 · Jun-Min Li1 · Jian-Ping Chen1 · Chuan-Xi Zhang1 · Jia-Bao Lu1

Received: 14 April 2022 / Accepted: 24 June 2022 / Published online: 23 August 2022
© The Author(s), under exclusive licence to Springer-Verlag GmbH Austria, part of Springer Nature 2022

Abstract
A novel chuvirus from a southern green stink bug (Nezara viridula) was identified by RNA sequencing in this study and was tentatively named “Ningbo southern green stink bug chuvirus 1” (NBSGSBV-1). The complete genome sequence of NBSGSBV-1 consists of 11,375 nucleotides, and the genome was found to be circular by ‘around-the-genome’ reverse transcription polymerase chain reaction (RT-PCR) and Sanger sequencing. Three open reading frames (ORFs) were predicted in the NBSGSBV-1 genome, encoding a large polymerase protein (L protein), a glycoprotein (G protein), and a nucleocapsid protein (N protein). A phylogenetic tree was constructed based on all of the currently available RNA-dependent RNA polymerase amino acid sequences of viruses of the family Chuviridae, and NBSGSBV-1 was found to cluster together with Sanya chuvirus 2 and Hubei odonate virus 11, indicating that NBSGSBV-1 might belong to the genus Odonatavirus. Five conserved sites were identified in the L proteins of NBSGSBV-1 and other chuviruses. The abundance and characteristics of the NBSGSBV-1-derived small interfering RNAs suggested that NBSGSBV-1 actively replicates in the host insect. To the best of our knowledge, this is the first report of a chuvirus identified in a member of the insect family Pentatomidae. The discovery and characterization of NBSGSBV-1 will help us to understand the diversity of chuviruses in insects.

With the advancement of next-generation sequencing (NGS) technology and metagenomic analysis in recent years, a growing number of novel viruses have been discovered and identified [1, 2]. Many have been discovered in arthropods such as insects, arachnids, and chilopods [2, 3]. Arthropods are therefore assumed to be a major reservoir of viral genetic variety and are likely to play an important role in viral evolution [2, 4]. As the number of novel viruses increases dramatically, new families of viruses are gradually divided. For instance, the family Chuviridae, which belongs to the order Jingchuvirales, class Monjiviricetes, was created in 2015 [2]. Most of the viruses in the family Chuviridae were discovered in the ancient Chinese region called Chu, which referred to the middle and lower reaches of the Yangtze River at that time [2, 5]. The phylogenetic diversity of chuviruses is between those of segmented and non-segmented viruses, and their genome structures are also diverse, including unsegmented, bi-segmented, and circular forms [2, 4, 5]. Typically, the circular structure of the chuvirus genome is distinct from the pseudo-circular structure of those of other negative-sense RNA viruses such as members of the families Peribunyaviridae and Orthomyxoviridae. The genomes of the linear chuviruses encode a glycoprotein (G), a nucleoprotein (N), and a polymerase (L), while the gene order in circular chuviruses such as Bole tick virus 3 (GenBank No. NC_028259.1) is L-(G)-N (displayed in a linear form). It has been reported that the G gene has probably been lost during the long-term evolution of some chuviruses [2, 4, 6].

The southern green stink bug (Nezara viridula), belonging to family Pentatomidae, order Hemiptera, is a highly polyphagous cosmopolitan pest. N. viridula feeds on a variety of important economic crops and is widely distributed in the Americas, Asia, Australia, and Europe [7, 8]. The damage caused by N. viridula is mainly due to its piercing-sucking mouthparts resulting in plant damage, reduced seed germination and survival, and the spread of plant pathogens [9, 10]. Previous studies have revealed the presence of several viruses in N. viridula. In 1992, two pathogenic viruses,
Nezara viridula virus 1 and Nezara viridula virus 2, were isolated and identified in *N. viridula* [11]. A honeybee virus called Israeli acute paralysis virus (IAPV) has recently been discovered in *N. viridula*, indicating that some viruses can spread among species [12]. In this study, a novel chuvirus present in *N. viridula* was identified by RNA sequencing (RNA-seq). This is the first chuvirus identified in a member of the family Pentatomidae.

In August 2019, a single adult stink bug was collected from a rice field in Hangzhou, Zhejiang, China, and total RNA was later extracted using TRIzol Reagent (Invitrogen, MA, USA) according to the manufacturer's instructions. After confirming the quality of the extracted RNA using a NanoDrop spectrophotometer (Thermo Scientific, MA, USA), paired-end (150 bp) sequencing of the RNA library was performed using an Illumina HiSeq 4000 sequencer (Novogene, Tianjin, China). After trimming adapter sequences [13], a total of 45,044,688 clean reads were obtained. The assembled contigs were first compared with Barcode of Life Data Systems (https://www.boldsystems.org/), and one contig representing the potential cytochrome oxidase subunit 1 (COI) sequence of the stink bug was extracted and subjected to a search on the National Center for Biotechnology Information (NCBI) website (https://www.ncbi.nlm.nih.gov/) and found to be 99.70% identical to the deposited COI sequence of *N. viridula*, confirming that the stink bug used in this study was indeed *N. viridula*. This contig was confirmed by Sanger sequencing and submitted to the GenBank database (accession number ON171205) (Supplementary File S1). The assembled contigs were then used to search a local virus database downloaded from the NCBI viral reference database (https://www.ncbi.nlm.nih.gov/genome/viruses). The results showed that one contig corresponded to the nearly complete genome sequence (about 11,500 nt) of a virus resembling previously reported chuviruses. To avoid false positives and to identify homologous viruses, the putative viral sequence was compared to sequences in the NCBI nucleotide (NT) and non-redundant (NR) protein databases (Supplementary Table S1). Both Bowtie2 and Samtools were used to map the adaptor- and quality-trimmed reads of the transcriptome back to the chuvirus contig, and a high level of coverage (~14,000X) was observed. To verify this contig, it was divided into six regions, and specific primers were designed accordingly (Supplementary Table S2) and used to perform RT-PCR and Sanger sequencing for each region. The results showed that the sequences of all six regions were perfectly matched to the original contig obtained from the transcriptome. Moreover, ‘around-the-genome RT-PCR’ was performed from the 3’ end to the 5’ end, and Sanger sequencing clearly showed that this chu-like viral contig is in a circular form (Supplementary Fig. S1).

The full genome sequence, consisting of the chu-like virus of 11,375 nucleotides, was submitted to the NCBI GenBank database (accession number ON191814) (Supplementary File S2), and this virus was tentatively named "Ningbo southern green stink bug chuivirus 1" (NBSGSBV-1). Using the NCBI Open Reading Frame Finder (ORF Finder, https://www.ncbi.nlm.nih.gov/orffinder/), three non-overlapping ORFs (ORF1, nt 48-6578; ORF2, nt 6619-8628; ORF3, nt 8698-10179) were found (Fig. 1A and B). ORFs 1-3 were predicted to encode a 261.12-kDa large polymerase protein (L protein), a 80.28-kDa glycoprotein (G protein), and a 59.16-kDa nucleoprotein (N protein), respectively, which is consistent with the typical structure of a previously reported circular chuivirus (L-(G)-N) [2]. To identify chuviruses related to NBSGSBV-1, a BLASTp search of the NCBI reference viral sequence database was performed, and the results indicated that the NBSGSBV-1 L protein and G protein sequences shared the highest sequence similarity to the NBSGSBV-1.

---

**Fig. 1** Characterization of Ningbo southern green stink bug chuivirus 1 (NBSGSBV-1). (A) Genome structure and transcript coverage of NBSGSBV-1. (B) Schematic diagram of the circular genome structure of NBSGSBV-1. RNA-dependent RNA polymerase, glycoprotein, and nucleoprotein genes are shaded blue, orange, and gray, respectively. L, G, and N represent the large polymerase protein, glycoprotein, and nucleoprotein, respectively.
A novel chuvirus from a southern green stink bug

similarity to a sequence identified as Chuviridae sp. (GenBank no. QXV86379.1), with 41.58% and 48.91%, amino acid sequence identity, respectively. The N protein sequence shared the highest sequence similarity with Changping tick virus 2 (GenBank No. YP_009177706.1) with 27.69% amino acid sequence identity (Supplementary Table S1). Using InterProScan (https://www.ebi.ac.uk/interpro), three conserved domains in the L protein were identified, including a Mononeg_mRNAcap domain, a Mononeg_RNA_pol_cat domain, and a Mononega_L_MeTrfase domain, whereas no conserved elements were found in the G or N protein sequences (Fig. 1A). To determine the abundance and coverage of the NBSGSBV-1-derived sequence reads, the RNA-seq reads were realigned to the confirmed full genome sequence of NBSGSBV-1. Notably, viral reads appeared to be accumulated in the 3′ region of the genome, especially in ORF3 (N) (Fig. 1A).

To investigate the taxonomic status of NBSGSBV-1, all of the available RNA-dependent RNA polymerase (RdRp) sequences of members of the family Chuviridae were retrieved from NCBI to generate a phylogenetic tree, and Atrato chu-like virus 5 (GenBank no. QHA33675.1) of the family Aliusviridae was used as an outgroup. The substitution model was first evaluated using ModelTest-NG, and a maximum-likelihood tree was then constructed using RAxMLNG (version 0.9.0) with 1000 bootstrap replicates [14–16]. As shown in Fig. 2A, NBSGSBV-1 clustered with Sanya chuvirus 2 (GenBank no. UHK03098.1) with a bootstrap value of 100. Another virus in the same branch with NBSGSBV-1 was Hubei odonate virus 11 (GenBank no. YP 009336946.1), which belongs to the genus Odonatavirus and has the same genome organization as NBSGSBV-1. According to the International Committee on Taxonomy of Viruses (ICTV, https://talk.ictvonline.org/), the genus Odonatavirus currently has three members. We suggest that NBSGSBV-1 may be added as a new member in this genus. Although NBSGSBV-1 has a circular genome, the genomes of Sanya chuvirus 2 and Hubei odonate virus 11 are both linear, which may imply unique evolutionary characteristics of chuvirus genomes.

MEME (https://meme-suite.org/meme/tools/meme) was used with default parameters to identify conserved motifs in the L genes of NBSGSBV-1, and five closely related chuviruses that clustered in the same branch of the tree, and five conserved motifs were identified (Supplementary Fig. S2). The conserved regions identified in these chuviruses might...
provide information for the classification of novel chuviruses in the future. In summary, the above results indicate that the newly discovered NBSGSBV-1 may belong to the genus *Odonatavirus* in the family *Chuviridae*.

To better understand small interfering RNA (siRNA)-based host antiviral immunity in response to NBSGSBV-1 infection, small RNAs (sRNAs) from *N. viridula* were sequenced and comprehensively characterized. First, a library of sRNA from *N. viridula* was prepared using an Illumina TruSeq sRNA Sample Preparation Kit (Illumina, USA) and sequenced using an Illumina HiSeq 2500 platform. Next, the sRNAs were processed to obtain clean reads, and sRNAs with a length of 18-30 nt were extracted. Finally, Bowtie was employed to map the processed sRNA reads back to the whole viral genome sequence of NBSGSBV-1 (allowing zero mismatches) for the identification of vsiRNAs. The obtained vsiRNAs were then analyzed further using custom Perl scripts and Linux Bash scripts [17]. A total of 12,907 vsiRNA reads (including 3,187 unique ones) mapped perfectly to the assembled genome sequence of the NBSGSBV-1. The length of vsiRNAs was mainly concentrated at 22 nt, accounting for 31.9% and 23.7% of the total and unique vsiRNAs, respectively. Moreover, they were almost equally derived from the sense and antisense strands of the viral genome (Fig. 2B). The vsiRNAs displayed a strong A/U preference in their 5’-terminal nucleotides (Fig. 2B) and were uniformly distributed throughout the viral genome, as shown in Fig. 2B and C. The typical characteristics of vsiRNAs suggested that the host antiviral RNAi pathway was actively involved in the response to NBSGSBV-1 infection.

In conclusion, a novel circular chuvirus, NBSGSBV-1, was identified in *N. viridula*. This is the first report of a chuvirus in a member of the insect family Pentatomidae. The discovery and identification of NBSGSBV-1 will contribute to a better understanding of the diversity of chuviruses in insects.

**Supplementary Information** The online version contains supplementary material available at [https://doi.org/10.1007/s00705-022-05560-1](https://doi.org/10.1007/s00705-022-05560-1).

**Acknowledgements** This work was supported by the National Natural Science Foundation of China (U20A2036) and Ningbo Science and Technology Innovation 2025 Major Project (2019B10004).

**Declarations**

**Compliance with ethical standards** The authors declare no conflict of interest.

**References**

1. Edgar RC, Taylor J, Lin V, Altman T, Barbera P, Meleshko D, Lohr D, Novakovsky G, Buchfink B, Al-Shayeb B, Banfield JF, de la Pena M, Korobeynikov A, Chikhi R, Babaian A (2022) Petabase-scale sequence alignment catalyses viral discovery. Nature 602:142–147
2. Li C-X, Shi M, Tian J-H, Lin X-D, Kang Y-J, Chen L-J, Qin X-C, Xu J, Holmes EC, Zhang Y-Z (2015) Unprecedented genomic diversity of RNA viruses in arthropods reveals the ancestry of negative-sense RNA viruses. Elife 4:e05378. [https://doi.org/10.7554/eLife.05378](https://doi.org/10.7554/eLife.05378)
3. Niu J, Li XL, Wu YL, Sun QZ, Zhang W, Cao M, Wang JJ (2020) RNA virome screening in diverse but ecologically related citrus pests reveals potential virus-host interactions. J Invertebr Pathol 170:107329
4. Qin XC, Shi M, Tian JH, Lin XD, Gao DY, He JR, Wang JB, Li CX, Kang YJ, Yu B, Zhou DJ, Xu J, Plyusnin A, Holmes EC, Zhang YZ (2014) A tick-borne segmented RNA virus contains genome segments derived from unsegmented viral ancestors. Proc Natl Acad Sci U S A 111:6744–6749
5. Shi M, Lin XD, Tian JH, Chen LJ, Chen X, Li CX, Qin XC, Li J, Cao JP, Eden JS, Buchmann J, Wang W, Xu J, Holmes EC, Zhang YZ (2016) Redefining the invertebrate RNA virosphere. Nature 540:539–543
6. Cook S, Chung BY, Bass D, Moureau G, Tang S, McAlister E, Culverwell CL, Glucksman E, Wang H, Brown TD, Gould EA, Harbach RE, de Lamballerie X, Firth AE (2013) Novel virus discovery and genome reconstruction from field RNA samples reveals highly divergent viruses in dipteran hosts. PLoS ONE 8:e80720
7. Todd JJA (1989) Ecology and behavior of *Nezara viridula*. Anna Rev Entomol 34:273–292
8. Panizzzi A, Lucini T (2016) What happened to *Nezara viridula* (L.) in the Americas? Possible reasons to explain populations decline. Neotrop Entomol 45:619–628
9. Miller L, Rose H, McDonald F (1978) The effects of damage by the green vegetable bug, *Nezara viridula* (L.) on yield and quality of soybeans. Aust J Entomol 16:421–426
10. Leskey TC, Hamilton GC, Nielsen AL, Polk DF, Rodriguez-Saona C, Bergh JC, Herbert DA, Kuhar TP, Saona C, Bergh JC, Herbert DA, Kuhar TP, Gp D (2012) Pest status of the brown marmorated stink bug, *Halyomorpha halys* in the USA? Possible reasons to explain populations decline. Neotrop Entomol 45:619–628
11. Miller L, Rose H, McDonald F (1978) The effects of damage by the green vegetable bug, *Nezara viridula* (L.) on yield and quality of soybeans. Aust J Entomol 16:421–426
12. Leskey TC, Hamilton GC, Nielsen AL, Polk DF, Rodriguez-Saona C, Bergh JC, Herbert DA, Kuhar TP, Pfeiffer D, Gp D (2012) Pest status of the brown marmorated stink bug, *Halyomorpha halys* in the USA. Outlooks Pest Manag 23:218–226
13. Williamson C, Von Wechmar M (1992) Two novel viruses associated with severe disease symptoms of the green stinkbug *Nezara viridula*. J Gen Virol 73:2467–2471
14. Susevich ML, Reynaldi FJ, Marti GA, Echeverría MG (2019) Primer hallazgo del Virus de la Parálisis Aguda Israeli (IPAV) en *Nezara viridula* (Hemiptera: Pentatomidae). Rev Soc Entomol Argentina 78:1–5
15. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q (2011)
Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol 29:644–652

14. Nguyen L-T, Schmidt HA, Von Haeseler A, Minh BQ (2015) IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol 32:268–274

15. Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis A (2019) RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. Bioinformatics 35:4453–4455

16. Darriba D, Posada D, Kozlov AM, Stamatakis A, Morel B, Flouri T (2020) ModelTest-NG: a new and scalable tool for the selection of DNA and protein evolutionary models. Mol Biol Evol 37:291–294

17. Li J, Andika IB, Shen J, Lv Y, Ji Y, Sun L, Chen J (2013) Characterization of rice black-streaked dwarf virus-and rice stripe virus-derived siRNAs in singly and doubly infected insect vector Laodelphax striatellus. PLoS ONE 8:e66007

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

Springer Nature or its licensor 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.