



# A Metagenomic Approach Identified a Novel Phasi Charoen-Like Virus Coinfecting a Chikungunya Virus-Infected *Aedes aegypti* Mosquito in Brazil

Marielton dos Passos Cunha,<sup>a</sup> Rafaella Sayuri Ioshino,<sup>b</sup> André Luis Costa-da-Silva,<sup>b</sup> Vivian Petersen,<sup>b</sup> Margareth Lara Capurro,<sup>b</sup> Paolo Marinho de Andrade Zanotto<sup>a</sup>

<sup>a</sup>Department of Microbiology, Institute of Biomedical Sciences, University of Sao Paulo, Sao Paulo, Brazil

<sup>b</sup>Department of Parasitology, Institute of Biomedical Sciences, University of Sao Paulo, Sao Paulo, Brazil

**ABSTRACT** Insect-specific viruses do not replicate in vertebrates. Here, we report the genome sequence of a novel strain of a Phasi Charoen-like virus (PCLV) that was isolated from a wild *Aedes aegypti* mosquito collected in Aracajú, Sergipe State, Brazil. The coding-complete genome of the PCLV is described in this report.

The *Aedes aegypti* mosquito is the main vector of several arboviruses. Although this species can be infected by viruses that circulate among invertebrate and vertebrate hosts, finding an infected wild mosquito is frequently challenging. Studies of the mosquito microbiome using metagenomics led to the discovery of a growing number of insect-specific viruses (ISVs) (1, 2). ISVs naturally infect and replicate in insects or laboratory insect cells but do not replicate in vertebrates or vertebrate cells (3).

Here, we describe the coding-complete genome of a member of the *Bunyaviridae* family that was obtained from an *A. aegypti* mosquito, which was naturally infected by chikungunya virus (CHIKV), as part of a surveillance study of arboviruses in urban mosquitoes (4). To identify whether there was any novel virus in urban *A. aegypti* mosquitoes, we collected mosquitoes in Sergipe State, Brazil, as described previously (4). The RNA was extracted from the abdomen of an *A. aegypti* female mosquito, which was initially macerated and subsequently subjected to extraction using the QIAamp viral RNA minikit (Qiagen, Valencia, CA, USA), purified (treated with DNase I), and concentrated using the RNA Clean and Concentrator-5 kit (Zymo Research, Irvine, CA, USA). The RNA was subjected to next-generation genomic sequencing. Briefly, paired-end RNA libraries were constructed using the TruSeq stranded total RNA HT sample preparation kit, and sequencing was performed using the Illumina NextSeq platform.

Taxonomic analysis of the reads revealed the presence of a large number of reads classified as Phasi Charoen-like virus (PCLV) (79.15%), followed by the presence of CHIKV (1.99%), as previously reported for the same PCLV-positive mosquito (4). All of the other reads were classified as belonging to other viruses and cellular organisms, such as *Archaea* and *Bacteria* (Fig. 1A). Taxonomic analysis of contigs assembled using *de novo* assembly led to the identification of three genome segments of a PCLV (segment L: 6,792 bp; GC content, 38.03%; segment M: 3,907 bp; GC content, 38.52%; segment S: 1,420 bp; GC content, 37.82%), with complete sizes (including coding and noncoding regions) similar to those of the RefSeq PCLV genome and those reported for this virus in the literature (segment L, GenBank accession number [NC\\_038262](https://www.ncbi.nlm.nih.gov/nuccore/NC_038262); segment M, [NC\\_038261](https://www.ncbi.nlm.nih.gov/nuccore/NC_038261); segment S, [NC\\_038263](https://www.ncbi.nlm.nih.gov/nuccore/NC_038263)). Genomic coverage was recovered for all sites using a map-to-reference strategy (Fig. 1B to D). Phylogenies indicate that the segments grouped with other PCLV sequences from around the world (1, 2, 5). Interestingly, the viral sequences announced here presented evolutionary relationships closer

**Citation** Cunha MDP, Ioshino RS, Costa-da-Silva AL, Petersen V, Capurro ML, Zanotto PMDA. 2020. A metagenomic approach identified a novel Phasi Charoen-like virus coinfecting a chikungunya virus-infected *Aedes aegypti* mosquito in Brazil. *Microbiol Resour Announc* 9:e01572-19. <https://doi.org/10.1128/MRA.01572-19>.

**Editor** John J. Dennehy, Queens College

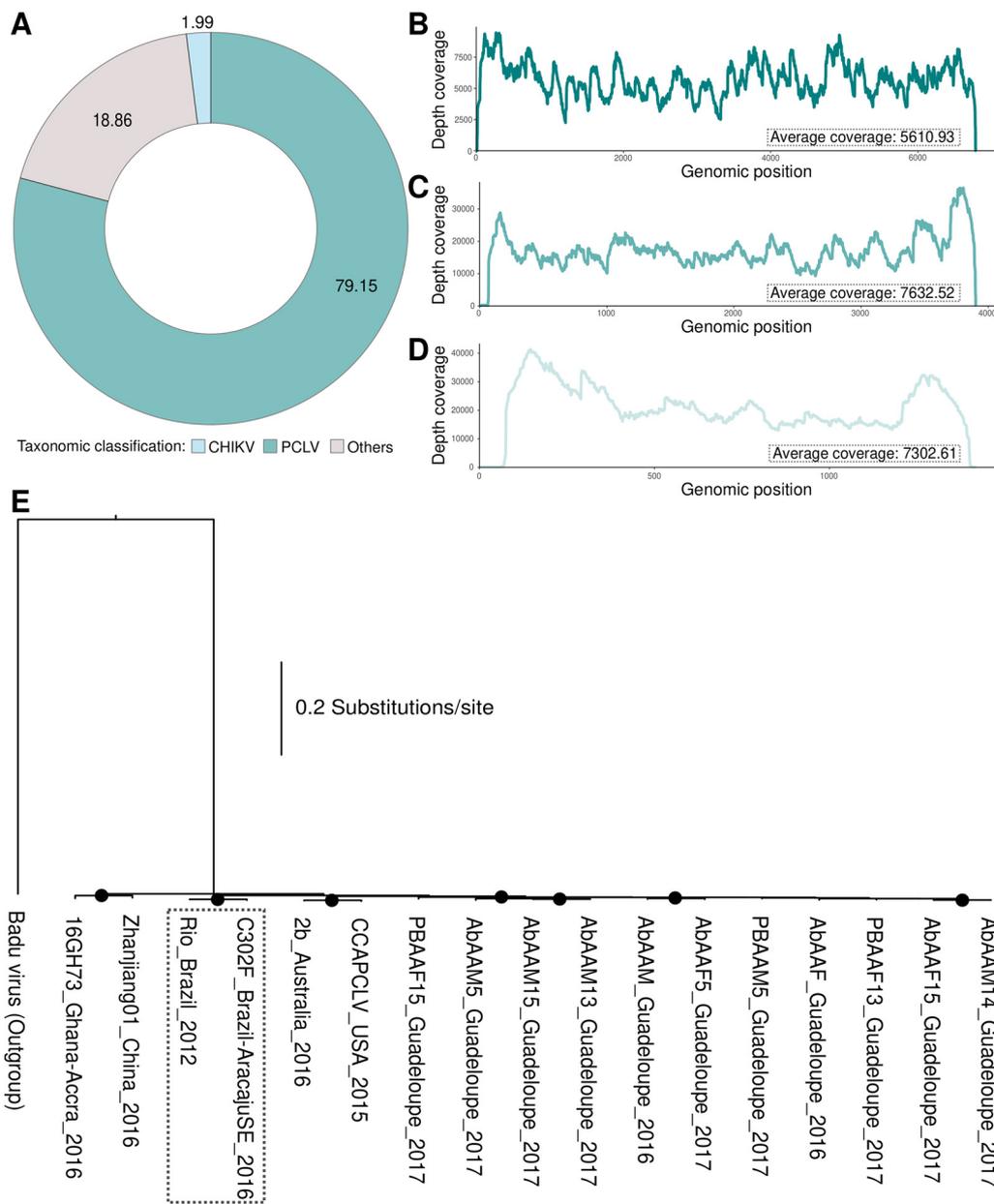
**Copyright** © 2020 Cunha et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Marielton dos Passos Cunha, [marieltondospassos@gmail.com](mailto:marieltondospassos@gmail.com), or Paolo Marinho de Andrade Zanotto, [pzanotto@usp.br](mailto:pzanotto@usp.br).

**Received** 17 January 2020

**Accepted** 6 July 2020

**Published** 30 July 2020



**FIG 1** Taxonomic characterization of a novel PCLV. (A) Taxonomic classification of all paired-end reads classified for the mosquito. Short unpaired reads and low-quality bases and reads were removed using Trimmomatic version 0.39 (leading, 20; trailing, 20; slidingwindow, 4;25; minlen, 36). The taxonomic classification of each paired-end read was performed using the Kaiju version 1.7.2 program, based on the reference sequence database provided by the program (completely assembled and annotated reference genomes of archaea, bacteria, and viruses from the NCBI RefSeq database); 79.15% of reads were PCLV, 1.99% were classified as CHIKV, and the rest were classified as reads compatible with other taxa. (B to D) Paired-end reads were assembled *de novo* with SPAdes version 3.13.1 (meta). The taxonomic classification of each assembled contig was determined using the Kaiju version 1.7.2 program, based on the reference sequence database provided by the program. In a second step, we use the three contigs classified as PCLV segments L, M, and S to map to the reference (PCLV contigs) using the program Bowtie 2 version 2.4.1 to determine the average coverage and the coverage for each nucleotide site along in the genome. Shown are the average coverage and the depth coverage for the three PCLV segments, segment L (B), segment M (C), and segment S (D). (E) Phylogenetic characterization of the PCLV based on amino acid sequences (aligned using MUSCLE version 3.8.1551 with default parameters) of all of the concatenated segments (concatenated using a supermatrix approach) of the virus, using a maximum likelihood approach implemented in IQ-TREE version 1.6.12, with the inclusion of other PCLV genomic sequences available in GenBank (up to May 2020, using PCLV sequences that contain all segments and that have the location and year of isolation). To root the tree, we used the reference genome of another phasivirus, the Badu virus (segment L, GenBank accession number [NC\\_038257](#); segment M, [NC\\_038258](#); segment S, [NC\\_038259](#)). The black dots indicate bootstrap values above 70. Briefly, the results indicate that the characterized sequences clustered with a virus previously isolated from a mosquito in Brazil (strain Rio) and that all of the Guadeloupe sequences were aggregated into a single group.

Downloaded from <http://mra.asm.org/> on September 21, 2020 by guest

