Complete Genome Sequence of Chikungunya Virus Isolated in the Philippines

Kent D. Kawashima,* Lady-Anne C. Suarez, Hannah Karen M. Labayo, Veni R. Liles, Noel C. Salvoza, David C. Klinzing,* Maria Luisa G. Daroy, Ronald R. Matias, Filipinas F. Natividad

Research and Biotechnology, St. Luke’s Medical Center, Quezon City, Philippines

* Present address: Kent D. Kawashima, Department of Genetics, School of Life Science, the Graduate University for Advanced Studies (SOKENDAI), Mishima, Japan; David C. Klinzing, AITbiotech Pte. Ltd., Singapore.

Chikungunya virus is an alphavirus of the *Togaviridae* family, which causes a febrile illness with arthralgia in humans. We report here on the complete genome sequence of chikungunya virus strain CHIKV-13-112A isolated from a patient in the Philippines who was suspected to have dengue virus. Phylogenetic analysis revealed that the strain is of the Asian genotype.

Chikungunya virus (CHIKV) is an alphavirus that causes chikungunya fever. It is transmitted by the bite of an infected *Aedes* mosquito. CHIKV occurs throughout the Philippines, with outbreaks reported in many provinces in 2012 to 2013 (see [http://www.promedmail.org/direct.php?id=1494945](http://www.promedmail.org/direct.php?id=1494945) and [http://www.promedmail.org/direct.php?id=1980717](http://www.promedmail.org/direct.php?id=1980717)).

Here, we report the complete coding sequence of the Asian genotype strain CHIKV-13-112A isolated from a Filipino patient initially suspected to have dengue. CHIKV was initially detected by reverse transcription-PCR (RT-PCR) targeting the E1 gene (1) and propagated in C6/36 *Aedes albopictus* cells for 7 days. CHIKV RNA isolated from the culture supernatant was used to construct a tagged random cDNA library by a modified single-primer amplification (SISPA) method (2).

First-strand tagged cDNA was produced using SuperScript III reverse transcriptase (Invitrogen) and the second-strand cDNA using the Klenow fragment (Vivantis). Tagged double-stranded (ds)-cDNA was amplified using iProof high-fidelity (HF) Taq polymerase (Bio-Rad) and a 20-nucleotide (nt) primer, FR20RV, which is complementary to the palindromic tag on ds-cDNA. Amplified ds-cDNA was used for library preparation for next-generation sequencing on MiSeq (Illumina). Paired-end sequencing was carried out using MiSeq version 2 reagents. The reads were assembled *de novo* using SGA (version 0.10.13) (3) and mapped to the CHIKV complete genome sequence most similar to the assembled contigs (accession no. KF318729) using the Burrows-Wheeler aligner (BWA) (version 0.7.8) (4). The consensus sequence of the mapped reads was determined using SAMtools (version 0.19-44428cd) (5).

The CHIKV-13-112A genome obtained is a positive single-stranded RNA (ssRNA) that is 11,933 nt long. Its coding sequence consists of two open reading frames (ORFs) flanked by 5′ and 3′ untranslated regions (UTR) whose termini were not verified. The first 7,413-nt ORF encodes a 2,471-amino acid (aa) nonstructural polyprotein containing nonstructural (ns)P1, 2, 3, and 4 proteins, respectively. There is a read-through site at positions 5586 to 5588 within the ORF that ignores an internal stop codon. From positions 7438 to 7502, there is a 65-nt UTR separating the two ORFs. The second 3,747-nt ORF encodes a 1,249-aa structural polyprotein containing the C, E2, E3, and E1 proteins, respectively.

The CHIKV-13-112A and other CHIKV complete genome sequences in GenBank, together with those of O’nyong-nyong virus (accession no. NC_001512), were aligned using MUSCLE (version 3.8.31) (6) and manually adjusted in UGENE (version 1.11.5) (7). Phylogenetic tree reconstruction based on the complete coding sequences was performed through RAXML (version 7.2.8) (8) using the GTR+gamma model of rate heterogeneity. The results revealed that strain CHIKV-13-112A belongs to the Asian genotype. A BLASTn search revealed 99% similarity to an isolate from Zhejiang, China (accession no. KF318729), while a phylogenetic analysis of complete coding sequences showed close relation to another Chinese strain (accession no. KC488650) in addition to the Zhejiang isolate, and to an Indonesian strain (accession no. FJ807897), all of which form a subclade in the Asian genotype. Likewise, two other CHIKV Philippine strains isolated in 1985 were classified as belonging to the Asian genotype (9). This confirms our previous analyses based on the nsP1 and E1 genes of 12 CHIKV strains isolated in 2012 from Manila, Philippines, which showed them to be of the Asian genotype (K. Kawashima, L. Suarez, H. Labayo, N. Salvoza, M. Daroy, R. Matias, and F. Natividad, unpublished data).

**Nucleotide sequence accession numbers.** The complete genome of CHIKV-13-112A has been deposited in DDBJ with the accession no. AB860301. The version described in this paper is the third version, AB860301.3.

**ACKNOWLEDGMENTS**

This work was supported by an equipment grant from Tahoe Research Initiative of the Illumina MiSeq next-generation sequencer and a funding...
grant under the Dengue Research Program (project code 07-024) from the Research and Biotechnology, St. Luke’s Medical Center, Philippines.

REFERENCES

1. Hasebe F, Parquet MC, Pandey BD, Mathenge EG, Morita K, Balasubramaniam V, Saat Z, Yusop A, Sinniah M, Natkunam S, Igarashi A. 2002. Combined detection and genotyping of chikungunya virus by a specific reverse transcription-polymerase chain reaction. J. Med. Virol. 67:370–374. http://dx.doi.org/10.1002/jmv.10085.

2. Djikeng A, Halpin R, Kuzmickas R, Depasse J, Feldhuyum J, Sengamalay N, Afonso C, Zhang X, Anderson NG, Spiro DJ. 2008. Viral genome sequencing by random priming methods. BMC Genomics 9:5. http://dx.doi.org/10.1186/1471-2164-9-5.

3. Simpson JT, Durbin R. 2010. Efficient construction of an assembly string graph using the FM-index. Bioinformatics 26:i367–i373. http://dx.doi.org/10.1093/bioinformatics/btp217.

4. Li H, Durbin R. 2010. Fast and accurate long-read alignment with Burrows-Wheeler Transform. Bioinformatics 26:589–595. http://dx.doi.org/10.1093/bioinformatics/btp698.

5. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map (SAM) format and SAMtools. Bioinformatics 25:2078–2079. http://dx.doi.org/10.1093/bioinformatics/btp352.

6. Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32:1792–1797. http://dx.doi.org/10.1093/nar/gkh340.

7. Okonechnikov K, Golosova O, Fursov M, UGENE team. 2012. Unipro UGENE: a unified bioinformatics toolkit. Bioinformatics 28:1166–1167. http://dx.doi.org/10.1093/bioinformatics/bts091.

8. Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22:2688–2690. http://dx.doi.org/10.1093/bioinformatics/btl446.

9. Volk SM, Chen R, Tsetsarkin KA, Paige Adams A, Garcia TJ, Sall AA, Nasar F, Schuh AJ, Holmes EC, Higgs S, Maharaj PD, Brault AC, Weaver SC. 2010. Genome-scale phylogenetic analyses of chikungunya virus reveal independent emergences of recent epidemics and various evolutionary rates. J. Virol. 84:6497–6504. http://dx.doi.org/10.1128/JVI.01603-09.