AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Complete Genome Sequences of Dengue Virus Type 2 Strains from Kilifi, Kenya

Everlyn Kamau,[a] Charles N. Agoti,[a] Joyce M. Ngoi,[a] Zaydah R. de Laurent,[a] John Gitonga,[a] Matthew Cotten,[b] My V. T. Phan,[b] D. James Nokes,[a,c] Eric Delwart,[d] Eduard Sanders,[a,e] George M. Warimwe[a,e]

[a]KEMRI-Wellcome Trust Research Programme, Kilifi, Kenya
[b]Department of Viroscience, Erasmus MC, Rotterdam, The Netherlands
[c]School of Life Sciences, Zeeman Institute, University of Warwick, Coventry, United Kingdom
[d]Blood Systems Research Institute, University of California, San Francisco, California, USA
[e]Centre for Tropical Medicine and Global Health, University of Oxford, Oxford, United Kingdom

ABSTRACT Dengue infection remains poorly characterized in Africa and little is known regarding its associated viral genetic diversity. Here, we report dengue virus type 2 (DENV-2) sequence data from 10 clinical samples, including 5 complete genome sequences of the cosmopolitan genotype, obtained from febrile adults seeking outpatient care in coastal Kenya.

Dengue virus (DENV) is a mosquito-borne, enveloped flavivirus with an RNA genome of ~10.7 kb (1). DENV, further classified into four distinct serotypes (DENV-1 to DENV-4), is widespread globally, causing an estimated 390 million infections yearly (2, 3). A licensed vaccine, Dengvaxia, is available for use in individuals with proof of prior infections (4), while other candidate vaccines are in development (5). Although an estimated >60 million DENV infections occur in Africa annually (1), DENV epidemiology is poorly characterized, including the circulating viral genetic diversity (6–8). Studies reporting new DENV genomes have the potential to facilitate the development of DENV vaccines. Here, we report 5 complete and 5 partial DENV-2 genome sequences identified from plasma samples of patients participating in a cross-sectional study on the burden of DENV and chikungunya virus in coastal Kenya (9). A detailed description of the study design is provided elsewhere (9, 10). The participants, aged 18 to 35 years, were seeking care at health facilities in Mtwapa, Kilifi County between February 2014 and January 2015. DENV detection used the CDC DENV-1 to -4 real-time reverse transcriptase PCR (RT-PCR) assay kit (catalog number KK0128). All participants provided written informed consent, and the study protocol was approved by the Scientific and Ethics Review Unit (Kenya) and the University of Oxford Tropical Research Ethics Committee in United Kingdom.

Sequencing was performed as previously described (11, 12). Total nucleic acid was extracted from 10 DENV-positive samples using the TRIzol LS reagent (Invitrogen) and DNase treated (TURBO DNase; Invitrogen). cDNA was synthesized using SuperScript III reverse transcriptase (Invitrogen) and nonribosomal hexanucleotide primers with reduced rRNA targets. Second-strand synthesis was performed using a Klenow fragment (New England BioLabs). Standard Illumina libraries were prepared using the Nextera XT kit (Illumina), and paired-end sequencing (2 × 250 bp) was performed with the MiSeq reagent v2 kit (Illumina). Short reads were filtered for quality using quality assessment of short read (QUASR) v.7.03 (13) and *de novo* assembled using SPAdes v.3.3.11 (14), and N_{50} values were determined using QUAST v.3.2.0 (15). Mean genome coverage was calculated by mapping short reads onto individual assemblies with Bowtie2 v.2.3.4.3 (16), followed by using the SAMtools v.1.9 sort and index functions (17) on the aligned

Citation Kamau E, Agoti CN, Ngoi JM, de Laurent ZR, Gitonga J, Cotten M, Phan MVT, Nokes DJ, Delwart E, Sanders E, Warimwe GM. 2019. Complete genome sequences of dengue virus type 2 strains from Kilifi, Kenya. Microbiol Resour Announc 8:e01566-18. https://doi.org/10.1128/MRA.01566-18.

Editor Julie C. Dunning Hotopp, University of Maryland School of Medicine

Address correspondence to Everlyn Kamau, ekamau@kemri-wellcome.org.

Received 15 November 2018
Accepted 2 January 2019
Published 24 January 2019

TABLE 1 Sequencing results and data assembly metrics for the partial and complete DENV-2 genomes from Kilifi, Kenya[a]

Strain	Total no. of reads[b]	No. of contigs[c]	N_{50} value	Genome length (bp)	G+C content (%)	Mean coverage (×)	GenBank accession no.	BioSample accession no.
DENV2/KLF/001/2014	196,326	637	1,161	10,093	45.9	1,023.3	MH456892	SAMN10606308
DENV2/KLF/004/2014	22,182	59	10,632	10,173	46.1	236.45	MH456893	SAMN10606311
DENV2/KLF/006/2014	2,505,710	5,989	613	8,372	45.7	3,452.32	MH456894	SAMN10606313
DENV2/KLF/005/2014	231,334	282	1,315	10,173	46.0	2,981.43	MH456895	SAMN10606312
DENV2/KLF/009/2014	4,963,410	22,485	791	4,272	45.5	3,324.8	MH456896	SAMN10606316
DENV2/KLF/002/2014	4,399,878	15,101	620	10,173	46.0	6,172.7	MH456897	SAMN10606309
DENV2/KLF/003/2014	195,090	208	825	10,173	46.1	3,589.76	MH456898	SAMN10606310
DENV2/KLF/008/2014	210,582	106	2,464	7,859	46.2	2,290.83	MH456899	SAMN10606315
DENV2/KLF/007/2014	75,768	359	791	2,407	45.2	451.82	MH456900	SAMN10606314

[a] QUASR parameters, "-d –q –l 125 –m 30"; SPAdes parameters, "–careful"; QUAST parameters, "minimum contig length: 500, ambiguity: one, threshold for extensive mis-assembly size: 1000"; Bowtie2: parameters, "-q –S –local."
[b] Short read length ranged from 35 to 250 bases.
[c] Contig refers to contiguous length of genomic sequence in which the order of bases is known to a high confidence level.

bam files, and then using the bedtools v.2.27.0 genomecov function (18) for generating the coverage statistics.

Sequencing and data assembly results and parameters are shown in Table 1. Full-length DENV genomes (>10 kb) were obtained from 5 samples while the remaining 5 samples had genomes between 2,407 and 8,372 nucleotides. Maximum likelihood phylogeny of envelope gene sequences, together with representative sequences of known DENV genotypes, classified the strains as the DENV-2 cosmopolitan genotype (19). Sequence annotation of the full-length genomes using Geneious R8.1.5 identified the expected polyprotein (10,176 nucleotides [nt], 3,392 amino acids) that yields 3 structural proteins, 7 nonstructural proteins, and a flanking 214-nt 3′ untranslated region (UTR) segment. No deletions, insertions, or premature stop codons were identified within the polyprotein-coding region. All five genomes showed a high level (>99.4%) of nucleotide similarity. Our results demonstrate an application for unbiased next-generation sequencing (NGS) without pathogen-specific enrichment. The new data provide a useful reference for the design of local diagnostics and for studies aimed at understanding DENV evolution and transmission in Kenya.

Data availability. The assembled sequences for full (five) and partial (four) genomes are available in the GenBank nucleotide database under accession numbers MH456892 to MH456900. The raw data are available in the NCBI SRA archive under BioProject accession number PRJNA510506.

ACKNOWLEDGMENTS

We acknowledge the staff at KEMRI-Wellcome Trust Research program who recruited patients and processed the samples.

This work was funded by a joint award from the Medical Research Council, Wellcome Trust, and Newton Fund as part of the EDCTP2 program supported by the European Union (grant number MC_PC_15092). Wellcome Trust grant number 102975 to D.J.N. also supported the work.

The manuscript was submitted for publication with permission from the director of the Kenya Medical Research Institute.

REFERENCES

1. Bhatt S, Gething PW, Brady OJ, Messina JP, Farlow AW, Moyes CL, Drake JM, Brownstein JS, Hoen AG, Sankoh O, Myers MF, George DB, Jaenisch T, Wint GR, Simmons CP, Scott TW, Farrar JJ, Hay SI. 2013. The global distribution and burden of dengue. Nature 496:504–507. https://doi.org/10.1038/nature12060.
2. Kraemer MU, Sinka ME, Duda KA, Mylne AQ, Shearer FM, Barker CM, Moore CG, Carvalho RG, Coelho GE, Van Bortel W, Hendrickx G, Schaffner F, Elyazar IR, Teng HJ, Brady OJ, Messina JP, Pigott DM, Scott TW, Smith DL, Wint GR, Golding N, Hay SI. 2015. The global distribution of the arbovirus vectors Aedes aegypti and Ae. albopictus. Elife 4:e08347. https://doi.org/10.7554/eLife.08347.
3. Leta S, Beyene TJ, De Clercq EM, Amenu K, Kraemer MUG, Revie CW. 2018. Global risk mapping for major diseases transmitted by Aedes aegypti and Aedes albopictus. Int J Infect Dis 67:25–35. https://doi.org/10.1016/j.ijid.2017.11.026.
4. World Health Organization. 2018. Meeting of the Strategic Advisory Group of Experts on immunization, April 2018—conclusions and recommendations. World Health Organization, Geneva, Switzerland.

5. Torresi J, Ebert G, Pellegrini M. 2017. Vaccines licensed and in clinical trials for the prevention of dengue. Hum Vaccin Immunother 13:1059–1072. https://doi.org/10.1080/21645515.2016.1261770.
6. Jaenisch T, Junghanss T, Wills B, Brady OJ, Eckerle I, Farlow A, Hay SI, McCall PJ, Messina JP, Ofula V, Sall AA, Sakuntabhai A, Velayudhan R, Wint GR, Zeller H, Margolis HS, Sankoh O, Dengue in Africa Study Group. 2014. Dengue expansion in Africa—not recognized or not happening? Emerg Infect Dis 20. https://doi.org/10.3201/eid2010.140487.
7. Ward T, Samuel M, Maoz D, Runge-Ranzinger S, Boyce R, Toledo J, Velayudhan R, Horstick O. 2017. Dengue data and surveillance in Tanzania: a systematic literature review. Trop Med Int Health 22:960–970. https://doi.org/10.1111/tmi.12903.
8. Amarasinghe A, Kuritsk JN, Letson GW, Margolis HS. 2011. Dengue virus infection in Africa. Emerg Infect Dis 17:1349–1354. https://doi.org/10.3201/eid1708.101515.
9. Ngoi CN, Price MA, Fields B, Bonventure J, Ochieng C, Mwashigadi G, Hassan AS, Thiong'o AN, Micheni M, Mugo P, Graham S, Sanders EJ. 2016. Dengue and chikungunya virus infections among young febrile adults evaluated for acute HIV-1 infection in coastal Kenya. PLoS One 11:e0167508. https://doi.org/10.1371/journal.pone.0167508.
10. Ngoi CN, Siqueira J, Li L, Deng X, Mugo P, Graham SM, Price MA, Sanders EJ, Delwart E. 2016. The plasma virome of febrile adult Kenyans shows frequent parvovirus B19 infections and a novel arbovirus (Kadipiro virus). J Gen Virol 97:3359–3367. https://doi.org/10.1099/jgv.0.000644.
11. Cotten M, Oude Munnink B, Canuti M, Deijs M, Watson SJ, Kellam P, van der Hoek L. 2014. Full genome virus detection in fecal samples using sensitive nucleic acid preparation, deep sequencing, and a novel iterative sequence classification algorithm. PLoS One 9:e93269. https://doi.org/10.1371/journal.pone.0093269.
12. Phan MVT, Anh PH, Cuong NV, Munnink BBO, van der Hoek L, My PT, Tri TN, Bryant JE, Baker S, Thwaites G, Woolhouse M, Kellam P, Rabaa MA, Cotten M. 2016. Unbiased whole-genome deep sequencing of human and porcine stool samples reveals circulation of multiple groups of rotaviruses and a putative zoonotic infection. Virus Evol 2:vew027. https://doi.org/10.1093/ve/vew027.
13. Watson SJ, Welkers MR, Depledge DP, Coulter E, Breuer JM, de Jong MD, Kellam P. 2013. Viral population analysis and minority-variant detection using short read next-generation sequencing. Philos Trans R Soc Lond B Biol Sci 368:20120205. https://doi.org/10.1098/rstb.2012.0205.
14. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19:455–477. https://doi.org/10.1089/cmb.2012.0021.
15. Mikheenko A, Saveliev V, Gurevich A. 2016. MetaQUAST: evaluation of metagenome assemblies. Bioinformatics 32:1088–1090. https://doi.org/10.1093/bioinformatics/btv697.
16. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357. https://doi.org/10.1038/nmeth.1923.
17. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25:2078–2079. https://doi.org/10.1093/bioinformatics/btp352.
18. Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26:841–842. https://doi.org/10.1093/bioinformatics/btq033.
19. Twiddy SS, Farrar JJ, Vinh Chau N, Wills B, Gould EA, Gritsun T, Lloyd G, Holmes EC. 2002. Phylogenetic relationships and differential selection pressures among genotypes of dengue-2 virus. Virology 298:63–72.