Complete Coding Genome Sequence for a Novel Multicomponent Kindia Tick Virus

Detected from Ticks Collected in Guinea

Running title: Genome Sequence for Kindia Tick Virus from Guinea

Vladimir A. Ternovoi,a Elena V. Protopopova,a Alexander N. Shvalov,a Mikhail Yu. Kartashov,a
Roman B. Bayandin,a Tatyana V. Tregubchak,a Sergey A. Yakovlev,b Konstantin A. Nikiforov,b
Svetlana N. Konovalova,a Valery B. Loktev,a Alexander P. Agafonov,a Rinat A. Maksyutov,a
Anna Yu. Popovac

a State Research Center for Virology and Biotechnology “Vector”, Koltsovo, Novosibirsk Region, Russia
b Federal Government Health Institution Russian Research Anti-Plaque Institute “Microbe”, Saratov, Russia
c Russian Federal Service for Surveillance on Consumer Rights Protection and Human Wellbeing, Moscow, Russia

Corresponding contributor: Valery B Loktev, State Research Center of Virology and Biotechnology «Vector», Rospotrebnadzor, Koltsovo, Novosibirsk Region, 630559, Russia,

Phone number: +7 383 3634700 add 2453; Facsimile number: +7 383 3367409; email addresses: valeryloktev@gmail.com and loktev@vector.nsc.ru
ABSTRACT

Kindia tick virus (KITV) is a novel multicomponent virus first detected by direct sequencing of *Rhipicephalus geigyi* ticks in Guinea in 2017. Here, we present a complete coding genome sequence for all four segments of KITV/2017/1. This virus appears to be evolutionarily related to unclassified flaviviruses, such as Alongshan virus.

ANNOUNCEMENT

The multicomponent (segmented) tick flaviviruses evolutionarily related to the unsegmented viruses of the genus *Flavivirus* have been recognized since only 2014 (1-4). This group of unclassified segmented tick-borne flaviviruses now includes Alongshan virus, Jingmen tick virus (JMTV) and Mogiana tick virus (MGTV), which have been detected in Asia, Europe and South America, respectively. These diverse and globally distributed viruses are capable of infecting a wide range of hosts, such as ticks, animals and humans (1–8). However, only a few complete genome sequences have been reported for JMTV (Kosovo and China), Alongshan virus (China) and MGTV (Brasilia).

Here, we report the complete coding genome sequences for the first African isolate, named Kindia tick virus (KITV), from *Rhipicephalus geigyi* collected from domestic cattle (*Bos taurus*) in Kindia, Guinea, West Africa. Twenty-six pools of 5 ticks each were frozen in liquid nitrogen, crushed by plastic pestles, homogenized in phosphate-buffered saline and used for RNA isolation with TRizol reagent (Invitrogen Co., USA). Total RNA was quantified with a Qubit RNA Assay Kit (Invitrogen Co., USA) following the manufacturer’s instructions. RNA-seq libraries were constructed with an NEBNext Ultra RNA Library Prep Kit for Illumina (New England Biolabs). Sequencing was performed using a MiSeq Reagent Kit v3 for 600 cycles. Cutadapt (version 1.18) and SAMtools (version 0.1.18) were used to remove the Illumina adaptors and duplicate reads. After removing adapters, the read length was 108-118 bases, with the numbers of reads not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.
per pool ranging from 151,282 to 540,929. The contigs were assembled *de novo* using the MIRA assembler with default parameters (version 4.9.6). In five pools, we found Mogiana-like fragments using BLASTN. These fragments were aligned to the reference genome segments of Mogiana tick virus isolate MGTG/V4/11 (4). The average coverage of the four segments was 35, 18, 52, and 30, respectively. The complete sequences for KITV/2017/1 were verified by Sanger sequencing of Seg 2 (2 short fragments), Seg 3 (1 sf) and Seg 4 (2 sf), with gaps or low coverage, as previously described (9). Total RNA from tick pools positive for KITV/2017/1 was used for first-strand DNA synthesis using a Reverta-L Kit (Interlabservice, Russia). The KITV/2017/1 genetic material was amplified by PCR with specific primers designed based on draft NGS sequences (available on request), with subsequent PCR fragment isolation and sequencing. The sequences for two detected KITVs (KITV/2017/1 and KITV/2017/2) were also deposited in GenBank. The nucleotide identity between KITV/2017/1 and KITV/2017/2 was 99.7% (Seg 1), 99.4% (Seg 2), 98.1% (Seg 3), and 99.3% (Seg 4). KITV contains putative open reading frames (ORFs) congruent with JMTV and MGTG, namely, nonstructural protein 1 (Seg 1), VP1 (Seg 2), NSP2 (Seg 3), VP2 and VP3 (Seg 4). The sizes of the sequenced segments (ORFs) are 2968 (2743), 2805 (2262), 2667 (2427) and 2725 (2351) bases, with CG contents of 52.3%, 54.9%, 54.4% and 54.2% for each segment, respectively. The lengths of the 5' UTR and 3' UTR were different for each segment and were 97-156 and 121-387 bases, respectively. The 5' UTR conservative motif GCAAGTGCA typical for JMTV was found in four segments, and the 3' UTR conservative motifs GGCAAGTGC and CAAGTG were also found in Seg 2 and Seg 4 of KITV. The divergence between KITV and MGTV/JMTV based on nucleotide sequence was 7–28% (Seg 1), 7.7–20.6% (Seg 2), 2.3–29.3% (Seg 3), and 0.4–22% (Seg 4), and based on amino acid sequences, it was 3.1–21.8% (Seg 1), 3.4–20.3% (Seg 2), 1.3–20.8% (Seg 3), and 0.2–15% (Seg 4), as evaluated by BLAST. Given these similarities, we suggest that KITV is a new member of the multicomponent (segmented) tick-borne flavivirus group and possibly represents a new species together with MGTG.
Data availability. GenBank accession numbers for the viral sequences: MH678723 and MH678727 (Seg 1), MH678724 and MH678728 (Seg 2), MH678725 and MH678729 (Seg 3), and MH678726 and MH678730 (Seg 4); complete sequences MK673133 (Seg 1), MK673134 (Seg 2), MK673135 (Seg 3), and MK673136 (Seg 4); and Sequence Read Archive accessions SRX5930668 to SRX5930671 under project PRJNA545394. The annotations have also been deposited into GenBank for these sequences.

ACKNOWLEDGMENTS

The authors would like to gratefully acknowledge support from the Russian Federal Service for Surveillance on Consumer Rights Protection and Human Wellbeing, Project RN№2904, from 12/22/2017.

REFERENCES


encephalitis virus from human brain to different cell cultures induces multiple genomic