

Isolation and genomic characterization of *Culex theileri* flaviviruses in field-collected mosquitoes from Turkey

**Koray Ergünay^{a,b}, Nadine Litzba^b, Annika Brinkmann^b, Filiz Günay^c,
Sırrı Kar^d, Kerem Öter^e, Serra Örsten^a, Yasemen Sarıkaya^c,
Bülent Alten^c, Andreas Nitsche^b, Yvonne-Marie Linton^{f-h}**

^a*Hacettepe University; Faculty of Medicine, Department of Medical Microbiology, Virology Unit, Ankara, TURKEY*

^b*Robert Koch Institute; Centre for Biological Threats and Special Pathogens 1 (ZBS 1), Berlin, GERMANY*

^c*Hacettepe University; Faculty of Sciences, Department of Biology, Division of Ecology, Ankara, TURKEY*

^d*Namık Kemal University, Faculty of Arts and Sciences, Department of Biology, Tekirdağ, TURKEY*

^e*Istanbul University; Faculty of Veterinary Medicine, Department of Parasitology, Istanbul, TURKEY*

^f*Walter Reed Biosystematics Unit, Museum Support Center MRC-534, Smithsonian Institution, Suitland, Maryland, USA*

^g*Department of Entomology, Walter Reed Army Institute of Research, Silver Spring, Maryland, USA*

^h*Department of Entomology, National Museum of Natural History, Smithsonian Institution, Washington, DC, USA*

Author for Correspondence:

Koray Ergünay MD, PhD
Hacettepe University Faculty of Medicine
Department of Medical Microbiology, Virology Unit
Morphology Building 3rd Floor.
06100 Sıhhiye Ankara TURKEY
Ph: 90-312-305 1560 -1465; 90-312-6802447
ekoray@hacettepe.edu.tr; korayergunay@gmail.com

Keywords:

Flavivirus, mosquito, culex, Turkey, genome

ABSTRACT

Vector surveillance for the arthropod-borne infections has resulted in the isolation of a growing number of novel viruses, including several flavivirus strains that exclusively replicate in insects. This report describes the isolation and genomic characterization of four insect-specific flaviviruses from mosquitoes, previously collected from various locations in Turkey. C6/36 *Aedes albopictus* and Vero cell lines were inoculated with mosquito pools. On C6/36 cells, mild cytopathic effects, characterized as rounding and detachment, were observed in four pools that comprised female *Culex theileri* mosquitoes. Complete (3 isolates, 10697 nucleotides) or near-complete (1 isolate, 10452 nucleotides) genomic characterization was performed in these culture supernatants via next generation sequencing. All strains demonstrated high genetic similarities, with over 99% identity match on nucleotide and amino acid alignments, revealing them to be different isolates of the same virus. Sequence comparisons identified the closest relative to be the *Culex theileri* flavivirus (CTFV) strains, originally characterized in Portugal. Phylogenetic analyses demonstrated that the isolates remained distinct as a cluster but formed a monophyletic group with CTFV strains, and shared a common ancestor with Quang Binh or related *Culex* flaviviruses. The organization of the viral genome was consistent with the universal flavivirus structure and stem-loops; conserved motifs and imperfect tandem repeats were identified in the non-coding ends of the viral genomes. A potential ribosomal shifting site, resulting in the translation of an additional reading frame, was detected. The deduced viral polyprotein comprised 3357 amino acids and was highly-conserved. Amino acid variations, presumably associated with adaptive environmental pressures, were identified. These isolates comprise the first fully characterized insect-specific flaviviruses in Turkey. Their impact on West Nile virus circulation, which is also endemic in the study region, remains to be explored.

1. Introduction

Viruses tentatively assigned to the genus *Flavivirus* that are mainly detected and exclusively replicating in insects are frequently named as mosquito-only or insect-specific flaviviruses (ISF) (Calzolari et al., 2016). Flaviviruses infect several vertebrate species, usually via arthropod-borne transmission, sometimes independent of vectors (Simmonds et al., 2011). However, ISFs are distinct for the inability to be propagated in vertebrate cell lines and lack of vertebrate hosts (Calzolari et al., 2016). ISFs have been recognized as a major phylogenetic group along with mosquito-borne, tick-borne and no-known-vector flaviviruses (Moureau et al., 2015) and have been suggested to represent an ancestral lineage of flaviviruses (Cook and Holmes, 2006; Cook et al., 2012). Following the isolation of the prototype ISF, cell-fusing agent virus (CFAV) from a mosquito cell line (Stollar and Thomas, 1975), several genetically-related but distinct viruses have been identified, especially during the last decade (Simmonds et al., 2011; Calzolari et al., 2016). Currently, two groups of ISFs have been recognized where the first group comprises phylogenetically-distinct isolates, including the prototype strain. Viruses in the second group are not monophyletic and are phylogenetically related to the mosquito/vertebrate flaviviruses (Blitvich and Firth, 2015).

ISFs share many similar properties with widely-known pathogenic arthropod-borne flaviviruses such as Dengue virus (DENV), West Nile virus (WNV) and tick-borne encephalitis virus (TBEV). They are enveloped, single-stranded RNA viruses, with a single open reading frame (ORF) that encodes the viral polyprotein, which is then cleaved by viral and host proteases to form three structural (C, preM and E), and seven nonstructural (NS1, NS2A, NS2B, NS3, NS4A, NS4B and NS5) proteins (Simmonds et al., 2011). ISFs seem to have a global distribution and have been

detected in field-collected mosquitoes from Europe, East Asia, Africa, the American Continent and Australia (Blitvich and Firth, 2015; Calzolari et al., 2016).

The ISFs possess potentially distinct and unique features among flaviviruses which are likely to have an impact on the epidemiology of their related vector-borne pathogens. Their widespread distribution without obvious reservoirs or amplification hosts requires exploration of potential mechanisms, such as vertical transmission, contributing to their dispersion (Blitvich and Firth, 2015). Some ISFs have been shown to produce DNA forms of their genomic RNA, and ISF-related sequences can be detected in the genomes of particular mosquitoes, a trait not previously documented for other flaviviruses (Crochu et al., 2004; Cook et al., 2009). Although ISFs have been mainly reported to infect various species of mosquitoes, viral RNA has also been detected in sandflies (family Psychodidae) and other hematophagous dipterans taxonomically distant from mosquitoes (family Culicidae), suggesting a broader spectrum of species susceptible to infection (Moureau et al., 2010; Sanchez-Seco et al., 2010). With no documentation of replication in vertebrates, either from cell lines or animal inoculations, ISFs are considered to pose no threat to human and animal health (Calzolari et al., 2016). However, their impact on live mosquitoes and their potential to harbor pathogenic flaviviruses and other microbial pathogens require deeper investigation and might have significant implications for infection control (Vasilakis and Tesh, 2015; Newman et al., 2016).

Located in the transboundary region of the temperate climate zone, connecting Asia, Europe and Africa, Turkey provides suitable habitats for activity of several mosquito species, some of which are well-established arbovirus vectors (Ramsdale et al., 2001). The widespread distribution of WNV with human and equine infections have been reported in various regions of Anatolia (Ergunay et al., 2011;

Ergunay et al., 2014). This study was performed to isolate and characterize mosquito-borne flaviviruses previously detected during a field surveillance campaign in Anatolia.

2. Materials and Methods

2.1. Ethics Statement

Previously processed pools of field-collected mosquitoes were employed for virus isolation and characterization; thus, local or regional ethics committee approval was not required.

2.2 Specimens and Virus Cultivation

Mosquito pools previously observed as reactive in flavivirus screening were used for virus cultivation and genomic characterization. The specimens were initially collected during a field campaign via CDC Miniature Light traps (John W. Hock Company, Gainesville, FL, USA) during June-October 2014 and 2015 (Ergunay et al., 2016). Following morphological identification to species level, they were pooled according to collection site, species and sex, processed and screened for flaviviruses employing a nested PCR assay described previously (Vazquez et al., 2012). Six mosquito pools that comprise *Culex theileri* (n:5) and *Culex pipiens* (n:1) species, in which PCR product sequencing revealed Mediterranean *Culex* and *Culex theileri* flavivirus-related sequences were included in the study. The pools were subjected to DNA barcoding by PCR amplification and sequencing of the cytochrome c oxidase I gene (Folmer et al., 1994).

One aliquot (200-400µl) of supernatant from each positive pool was diluted in an equal volume of phosphate buffered saline, filtered through a 0.22 µm sterile

membrane filter (Merck Millipore, Darmstadt, Germany) and inoculated onto semi-confluent monolayers of African green monkey (Vero, ATCC-CCL81) and *Aedes albopictus* (C6/36, ATCC-CRL1660) cells in T25 flasks (Nunc, Roskilde, Denmark). After one hour at room temperature for enabling virus attachment to cells, 5 ml of Leibovitz L15 Medium (C6/36 cells) or Dulbecco's Modified Eagle Medium (Vero cells) were added, supplemented with 5% fetal bovine serum, L-glutamine, 100 U/ml penicillin, 100 µg/ml streptomycin, and 1 µg/ml fungizone. The cells were incubated at 28°C (C6/36 cells) or 37°C in 5% CO₂ (Vero cells) and observed daily for cytopathic effects (CPE). Blind passages were performed weekly up to the 3rd passage, when culture supernatants were tested for viral nucleic acids via the screening assay. Culture supernatants from positive inoculations were then collected, stored at -80°C and used as viral stocks.

2.3 Next generation sequencing and data analysis

Total RNA, purified from the cell culture supernatant via QIAamp viral RNA extraction kit (Qiagen, Valencia, CA, USA), was subjected to next generation sequencing (NGS) for the determination of viral genomes. For this purpose, libraries were prepared from each specimen using the NexteraXT DNA Library Preparation Kit (Illumina Inc., San Diego, CA, USA). Briefly, 5 ng of RNA was reverse transcribed into double-stranded cDNA via SuperScript IV Reverse Transcriptase (Thermo Fisher Scientific, Hennigsdorf, Germany) and NEBNext mRNA Second Strand Synthesis Module (New England Biolabs, Frankfurt am Main, Germany). The cDNA was cleaned up using Agencourt AMPure XP Reagent (Beckman Coulter Biosciences, Krefeld, Germany) and analyzed via the Agilent 2100 Bioanalyzer (Agilent Technologies, Waldbronn, Germany) for yield and size distribution.

Fragmentation, adaptor ligation and amplification steps were performed via Nextera XT DNA Library Prep kit (Illumina Inc), as suggested by the manufacturer. The sequencing runs were performed in one lane of the Illumina HiSeq 1500 (Illumina Inc.) in the high output mode.

The raw NGS data was de-multiplexed and extracted in fastq format. Paired-end reads of 2 x 250 basepairs (bp) were collected and reads of each library were aligned to the RefSeq viral nucleotide and protein genome database using MALT (MEGAN alignment tool, v0.3.8) (Huson et al., 2016) and DIAMOND (v0.7.1) (Buchfink et al., 2015) tools. Subsequent de novo assembly of the full genomes was carried out via Geneious software v9.1 (Biomatters Ltd, Auckland, New Zealand).

2.4 Phylogenetics, RNA structure and protein analyses

Complete and near-complete viral genomes and putative polyprotein sequences were analyzed using CLC Main Workbench v7.7.2 (CLCBio, Aarhus, Denmark), Bioedit v.7.0.9.0 (Hall, 1999) and MEGA v.6.06 (Tamura et al., 2013). Nucleotide and protein similarity searches in the public databases were assessed by the Basic Local Alignment Search Tool (Altschul et al., 1990), implemented via the National Center for Biotechnology Information website (www.ncbi.nlm.nih.gov/blast/), using BLASTn, BLASTn optimized for highly similar sequences (MEGABLAST) and BLASTx algorithms. Nucleotide and amino acid sequence alignments were generated via the CLUSTAL W program (Thompson et al., 1994) implemented in the Bioedit software. Pairwise sequence comparisons were carried out via MEGA and CLC softwares. The optimal model for the phylogenetic and molecular evolutionary analyses were determined using the Find best DNA/protein-substitution model tools built into MEGA software and verified using complementary

models. Maximum likelihood trees based on the nucleotide and amino acid sequences were constructed using the Tamura-Nei and Jones-Taylor-Thornton models, respectively. The reliability of the inferred trees was evaluated by bootstrap analysis of 1000 replicates.

The non-coding 5' and 3' ends of the viral genomes were analyzed using the MFOLD v2.3 (<http://mfold.rna.albany.edu/?q=mfold>) and RNAfold (<http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi>) servers for folding patterns, as well as visual inspection to identify conserved or repeated sequence motifs. Putative protease cleavage sites were predicted via SignalP 4.0 (<http://www.cbs.dtu.dk/services/SignalP/>) and PROSPER (protease specificity prediction server, <https://prosper.erc.monash.edu.au/home.html>) (Petersen et al., 2011; Song et al., 2012). The identified cleavage sites were compared to other flaviviruses. Conserved domain and motif searches were performed using the Web CD-search tool (<http://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi>) and MOTIF Search (<http://www.genome.jp/tools/motif/>) in the PFAM database (Bateman et al., 2002; Marchler-Bauer et al., 2015).

3. Results

3.1 Virus Cultivation and NGS

Out of the 6 mosquito pool macerates used to inoculate both C6/36 and Vero cell monolayers, four of the former revealed very mild CPE, characterized by cell rounding and detachment 5-7 days post inoculation, during the 2nd and 3rd passages. The culture supernatants were screened by PCR, revealing positive amplification results. No CPE or amplification was observed in Vero cells inoculated with any of the macerates. Attempts at virus titration by plaque assays were

unsuccessful on C6/36 cells. Supernatants from C6/36 cells with PCR-positive results (n=4) were processed and submitted to NGS. All reactive pools originated from *Cx. theileri* pools with female individuals collected in Canakkale and Kirklareli provinces during 2014 and 2015 (**Table 1**).

The NGS data resulted in 29,737,056, 2,013,712, 1,766,838 and 4,213,596 reads (250 bp) from specimens CTFVtur-1, CTFVtur-2, CTFVtur-3 and CTFVtur-4, respectively. Reads were then trimmed for quality (Phred quality score >33 with >99.9% base call accuracy) and the obtained reads were nucleotide- and protein-aligned against the complete viral database from NCBI, using MALT and DIAMOND software tools. Total read numbers specified to flaviviruses were 48,424 (CTFVtur-1), 3,025 (CTFVtur-2), 69,541 (CTFVtur-3) and 56,123 (CTFVtur-4). Following de novo assembly of the sequences, they were mapped to the closest related strain, *Culex theileri* flavivirus (CTFV) RP-2011 (Parreira et al., 2012), via Geneious software. Thus, the isolates are tentatively called the *Culex theileri* flavivirus Turkey (CTFVtur).

3.2 Genomic characterization of CTFVtur isolates and phylogeny

The genome of 10697 nucleotides was assembled for CTFVtur-1 (GenBank accession numbers: KX652375), CTFVtur-3 (KX652378) and CTFVtur-4 (KX652377). The isolate CTFVtur-2 (KX652376) yielded a contig of 10452 nucleotides, missing 245 bases from the 5' end of the viral genome on alignment, presumably due to the relatively low number of flavivirus-mapped reads in the supernatant (see above). Since three complete viral genomes were available, no further attempt was made to fill the gap for this isolate.

The analysis of the genomes revealed a 10,073 nucleotides long polyprotein-coding region, flanked by non-coding 5' and 3' ends of 199 and 425 nucleotides, respectively. Alignment and pairwise comparisons of the coding region demonstrated similarity rates of 99.09–99.64% and 99.73–99.88% on nucleotide and putative amino acid levels, respectively, among isolates (**Table 2**). BLASTn (MEGABLAST) searches revealed highest identity matches to the sequences belonging to two CTFV isolates characterized in Portugal (Parreira et al., 2012). Similarity rates of 94.51–99.04% were noted in pairwise nucleotide comparisons with these viruses (**Table 2**). CTFVtur strains are 148 and 13 nucleotides longer than CTFVs at the 5' and 3' ends of the genome, respectively.

Phylogenetic relationships of the CTFVtur isolates with several mosquito and tick-borne flaviviruses as well as insect-specific strains and viruses with no known vectors were determined. Complete nucleotide and putative amino acid sequences as well as the NS5 amino acid sequences were analyzed by using the maximum likelihood method. In all analyses, CTFVtur isolates clearly form a monophyletic group with CTFV isolates, which shares a common ancestor with mosquito flaviviruses YDFV, LSFlaviV-A20-09 and Quang Binh viruses from *Culex* mosquitoes, with high bootstrap values (**Figures 1–3**).

3.3 Analysis of 5' and 3' non-coding regions

The 5' region was highly conserved (0.5% variation) with a single nucleotide variation (G–T, residue 149) in CTFVtur-1 in 199 nucleotides. The predicted secondary structure of the 5' non-coding region at 28°C demonstrated the typical stem structure with top and side loops, as observed in several insect-specific and mosquito-borne flaviviruses (Hoshino et al., 2007). An identical 5' folding pattern was

noted in all strains. Among the isolates, the 3' region was also conserved, with two nucleotide changes (A–T in 10,449, C–G in 10,608) in CTFVtur-1 strain, leading to a 0.47% variation. The conserved pentanucleotides (CACAG and/or CACCG) found in almost all mosquito-borne and insect-specific flavivirus 3' ends (Hoshino et al., 2007) were present in the isolates. An imperfect tandem repeat, incorporating 30–31 nucleotides comparable to those in *Culex* flavivirus, cell-fusing agent virus and Kamithi river virus, was identified (**Table 3**) (Gritsun et al., 2014). The predicted folding patterns of the 3' end revealed several distinct or duplicated short hairpins, a typical feature of the insect-specific flaviviruses (Villordo et al., 2016).

3.4 Analysis of Putative Coding Region

A putative viral polyprotein of 3,357 amino acids was present in all CTFVtur isolates. The polyprotein was highly conserved among strains, with 0.12–0.27% variation. Pairwise comparisons with the related CTFV strains demonstrated an identical polyprotein size and 98.65–99.31% overall similarity on the amino acid level (**Table 2**). Region-specific comparisons with CTFV isolate 153 further revealed 97.6–100% similarity rates on particular mature viral proteins (**Table 4**).

The CTFVtur polyprotein follows the universal flavivirus genomic organization and comprises structural proteins virion C-anchored C (C-AC), premembrane-membrane (PrM-M), envelope (E), and non-structural proteins (NS1, NS2a, NS2b, NS3, NS4a, NS4b and NS5) (Simmonds et al., 2011). The predicted cleavage sites were homologous to CTFV (Parreira et al., 2012) and seem to involve the viral NS3 serine protease in the processing of C-AC, PrM-M, NS2a–NS2b, NS2b–NS3, NS3–NS4a and NS4b–NS5.

The ISF-programmed ribosomal frameshifting site, which includes a slippery heptanucleotide motif followed by a 5–9 nucleotide spacer region and a pseudoknot or stem-loop structure (Blitvich and Firth, 2015), was present in all CTFVtur isolates. The heptanucleotide motif encompassed the nucleotides 3,498–3,504 and was located between NS1 and NS2a. The composition of the motif and the structure was nearly identical to that of CTFV, despite a single nucleotide variation leading to an identical spacer in *Culex* flaviviruses which was detected in the spacer region of CTFVtur-3.

Throughout the putative polyprotein, several amino acid variations were identified among CTFVtur strains and in comparison with CTFV isolates (**Table 5**). In particular, modifications detected in residues C-44, PrM-149, NS3-1661 and NS5-2658 were also present in CFTV isolates, indicating a probable role in adaptation regardless of the location. In contrast, variations in C-AnC (residues 39, 123), E (490, 604, 667) NS1 (853, 973), NS2a (1078, 1239, 1272), NS3 (1750, 1769), NS4b (2230, 2244, 2248, 2249, 2355, 2408) and NS5 (2709, 2836, 2888, 3327, 3355) were consistent between CTFVtur and CTFV strains and might represent adaptive changes to environmental/host factors.

Functional analyses of the polyprotein revealed conserved motifs in E, NS1, NS3 and NS5 regions. Central and dimerization domains of the flavivirus glycoprotein (PSSM-ID: 279241, PFAM-ID: PF00869) was identified in E protein (residues: 321-436). The conserved motif of 14 amino acids, present in several flaviviruses and ISFs and thought to be involved in viral endosomal fusion and cellular entry, was identified in all strains (**Table 6**) (Cook et al., 2012). In NS1, flavivirus non-structural protein 1 domain (PSSM-ID: 279316, PFAM-ID: PF00948; residues: 708-1067) was also detected. In NS3, peptidase S7-flavivirus NS3 serine

protease (PSSM-ID: 250249, PFAM-ID: PF00949; residues: 1471-1614), DEAD-like helicase, Flavivirus DEAD domain (PSSM-ID: 214692, PFAM-ID: PF07652; residues: 1621-1766) were identified. Furthermore, conserved ATP (residues: 1641-1645) and ion binding (residues: 1729-1732) sites (PSSM-ID: 238005) were discovered. Finally, NS5 included the Flavivirus RNA-dependent RNA polymerase (PSSM-ID: 250266, PFAM-ID: PF00972; residues: 2709-3353) and FtsJ-like methyltransferase motifs (PSSM-ID: 279986, PFAM-ID: PF001728; residues: 2510-2683).

4. Discussion

During the past decade, a significant increase in newly-discovered ISFs was observed, in parallel with the biosurveillance efforts regarding arthropod-borne viruses and supported by powerful molecular tools such as NGS. They have resulted in the detection of several novel flaviviruses and the characterization of ISFs, nearly four decades after the isolation of the prototype virus (Stollar and Thomas, 1975; Calzolari et al., 2016). This report describes the genomic characterization of the CTFVs, isolated following a screening campaign in various regions in Turkey (Ergunay et al., 2016).

The study involved the inoculation of 6 mosquito pools, previously identified as positive for genetically related *Culex*-associated ISFs, onto cell lines of insect and vertebrate origin that are employed frequently for flavivirus isolation (Calzolari et al., 2016). Four pools demonstrated mild CPE solely on insect cells during 2nd to 3rd passages and were positive by an NS5-based PCR used for screening, suggesting the isolation of the infecting viruses. Complete (for 3 isolates) or near-complete (1 isolate) characterizations were accomplished from culture supernatants via NGS,

performed using standardized procedures and equipment. All strains demonstrated high genetic similarity with over 99% identity match on nucleotide and amino acid alignments, thus revealing them as different isolates of the same virus (**Table 2**). Sequence comparison identified the closest relative strain to be the two isolates (strains 153 and 178) of CTFV RP-2011 originally characterized in Portugal (Parreira et al., 2012), with 94–99% similarity rates (**Table 2**). A variation over 16% on the nucleotide level is required to establish any particular strain as a separate species in the *Flaviviridae* family (Kuno et al., 1998). Moreover, antigenic structure and associated cross-reactions, vector/host associations and environmental characteristics are also considered for official species demarcation and individual characterization of closely related viruses (Simmonds et al., 2011), some of which have not been fully explored for CTFVs. Phylogenetic analyses reveal ISFs to be distinct from other flaviviruses, and isolates originating from *Culex*, *Aedes*, *Mansonia* and *Coquilettidia* spp. mosquitoes form different clades and subclades (Blitvich and Firth, 2015; Calzolari et al., 2016). CTFV strains from Turkey and Portugal are placed with *Culex*-associated ISFs in various analyses, the most closely related strains being Quang Binh or similar viruses from China and Japan (**Figure 1-3**). Despite statistically supported separation of the strains originating from Portugal and Turkey based on complete genome phylogenies (**Figure 1 and 2**), current information indicates that these strains constitute local variants of CTFV. The acronym CTFVtur was used throughout this report for practical purposes to differentiate the strains detected in Turkey. The pairwise comparisons demonstrate that CTFVtur genomes are 161 nucleotides longer than CTFVs from Portugal, completing the non-coding ends of the genome, which had not been possible previously by using conventional methods (Parreira et al., 2012).

Although the most comprehensive characterization of the CTFV was performed from specimens collected in Portugal during 2009 and 2010, there is evidence for the circulation of several related and potentially identical strains (Blitvich and Firth, 2015). Partial NS5 sequences closely related to CTFV have been detected in Spain and Portugal from 2006 to 2010 with different acronyms (CxthFV) or names (Spanish *Culex* flavivirus, SCxFV) (Calzolari et al., 2012; Vazquez et al., 2012). Another short sequence of a related strain named Wang Thong virus from Thailand was deposited in the GenBank (Accession No. AY457040). The available sequences are 91–100% identical; therefore these strains are considered as the same flavivirus species (Blitvich and Firth, 2015). Virus isolation and subsequent antigenic characterization or at least more detailed genetic data are required to elucidate the properties and taxonomy of these strains.

The analysis of the nucleotide sequences of the CTFVtur strains revealed several genus- or species-specific structural properties. The stem-loop structures occurring in non-coding ends of the genome, presumably associated with various steps in the virus life cycle (Villordo et al., 2016), are detected in the strains by using software-based RNA folding predictions. Furthermore, conserved flavivirus pentanucleotide motifs and imperfect tandem repeats (**Table 3**) were identified in the 3' ends of the viral genome. The number and pattern of the repeats seem to differ in various ISFs, and a possible association with the host mosquito species (*Culex* and *Aedes*) has been suggested (Hoshino et al., 2007). All CTFVtur isolates were also observed to contain the potential ribosomal shifting site, resulting in the translation of a new reading frame common in ISFs (Blitvich and Firth, 2015). The expression pattern and the structure of the resulting *fifo* protein in CTFVs also require further investigation.

The deduced viral polyprotein is identical in size (3,357 amino acids) in the CTFVs from Turkey and Portugal and remain highly conserved, with less than 3% variation among all isolates (**Table 2**). This is also evident when comparing the individual proteins of two CFTV strains (**Table 4**). Conserved domains of viral envelope glycoprotein, fusion protein, viral serine protease and replicase were identified, with several accessory motifs present. Nevertheless, amino acid variations were also noted in each group as well as inter-group comparisons (**Table 5**). Interestingly, particular variations on 4 residues located in viral capsid, pre-membrane, NS3 and NS5 proteins were identified in strains of each group, suggesting similar selective pressures occurring in all locations. Conversely, some variations, especially found on the putative viral capsid, envelope and non-structural proteins NS1, NS2a, NS3, NS4b and NS5, are detected in strains from either location, which might be reflecting environmental adaptive processes.

We have isolated and propagated the CTFV_{tur} strains using C6/36 cell lines derived from *Aedes albopictus*, and Vero cells failed to support viral growth regardless of strain and passage number. This is a distinguishing feature of ISFs, and several host range and tissue tropism efforts using cells or laboratory animals have been unsuccessful so far. Nevertheless, various cell lines of mosquito origin have been reported to be susceptible to infection with ISFs (Blitvich and Firth, 2015). Formation of a visible CPE on cell lines also seems to vary according to the ISF strain analyzed and passage number, with some of the strains producing no microscopically visible effect on cells (Stollar and Thomas, 1975, Blitvich and Firth, 2015). Up to the fourth passage, we could observe minimal changes in C6/36 cells infected with CTFV_{tur} strains, with very mild rounding and premature detachment. CTFV strains from Portugal have been reported to induce cell growth retardation and

aggregate formation starting with the third blind passage (Parreira et al., 2012). Further passages might elucidate whether a visible or distinguishing feature of CTFV/CTFVtur growth on cells will be identified.

CTFVtur isolations were accomplished from *Cx. theileri* pools that comprise exclusively female mosquitoes and 7–42 individuals (**Table 1**). Same mosquito species were also involved in initial isolation of the virus (Parreira et al., 2012). Moreover, presence of partial CFTV NS5 sequences could also be detected in *Cx. pipiens* pools in Turkey as well in Portugal and Spain on various occasions between 2006 and 2015 (Vazquez et al., 2012; Parreira et al., 2012; Ergunay et al., 2016). In Thailand, the sequence of Wang Thong virus which is closely related to CFTV was identified in *Cx. fuscocephala* mosquitoes. Thus, the host range of CFTV may not be restricted to *Cx. theileri*. Although a grouping of the ISFs according to their host mosquito species seemed plausible in initial analyses (Hoshino et al., 2007), several phylogenetically distinct ISFs have been repeatedly detected in various *Culex* and *Aedes* sp. mosquitoes in different areas (Blitvich and Firth, 2015), suggesting independent introductions and multiple host-switching events (Cook et al., 2012). However, a significantly higher detection rate observed in *Cx. theileri* compared to *Cx. pipiens* in two regions of Turkey (Ergunay et al., 2016) and independent isolations from Portugal and Turkey suggest a more efficient replication and probably higher viral loads in this species. *Cx. theileri* and *Cx. pipiens*, along with several others, are considered as competent WNV vectors among the mosquito species distributed throughout Europe (Calistri et al., 2010). Therefore, an interaction of CFTV and WNV in the mosquito host which might affect WNV propagation and transmission might be possible, which we intend to explore using in vivo systems in the future.

5. Conclusion

We hereby describe the first fully characterized insect-specific flaviviruses in circulation in Turkey. The ends of the CFTV genome have been completed, based on data from three isolates. Major features of the nucleotide and putative amino acid sequences have been investigated which revealed several structural characteristics of the strains.

Acknowledgements

This study was partially supported by The Armed Forces Health Surveillance Center, Global Emerging Infections Surveillance and Response System (AFHSC-GEIS), United States (with Yvonne-Marie Linton as the principal investigator). KE is a recipient of the Georg Forster Research Fellowship (HERMES) for Experienced Researchers by the Alexander von Humboldt Foundation, 2015. This manuscript was prepared whilst YML held a National Research Council (NRC) Research Associateship Award at the Walter Reed Army Institute of Research. This research was performed in part under a Memorandum of Understanding between the Walter Reed Army Institute of Research and the Smithsonian Institution, with institutional support provided by both organizations. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript. The material to be published reflects the views of the authors and should not be construed to represent those of the US Department of the Army or the US Department of Defense.

The authors are grateful to Anette Teichmann and Delia Barz for their support during cell culture inoculations and Ursula Erikli for copy-editing. Preliminary findings of this study have been accepted for poster presentation at the International Meeting

on Emerging Diseases and Surveillance (IMED) to be held from November 4 to 7, 2016, in Vienna, Austria.

REFERENCES

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. *J. Mol. Biol.* 215, 403-410.

Bateman, A., Birney, E., Cerruti, L., Durbin, R., Eddy SR., Griffiths-Jones S., Howe K.L., Marshall, M., Sonnhammer, E.L., 2002. The Pfam protein families database. *Nucleic Acids Res.* 30, 276-280.

Blitvich B.J., Firth, A.E., 2015. Insect-specific flaviviruses: a systematic review of their discovery, host range, mode of transmission, superinfection exclusion potential and genomic organization. *Viruses* 7, 1927-1959. doi: 10.3390/v7041927.

Buchfink, B., Xie, C., Huson, D.H., 2015. Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* 12, 59-60.

Calzolari, M., Ze-Ze, L., Ruzek, D., Vazquez, A., Jeffries, C., Defilippo, F., Osorio, H.C., Kilian, P., Ruiz, S., Fooks, A.R., Maioli, G., Amaro, F., Tlustý, M., Figuerola, J., Medlock, J.M., Bonilauri, P., Alves, M.J., Sebesta, O., Tenorio, A., Vaux, A.G., Bellini, R., Gelbic, I., Sanchez-Seco, M.P., Johnson, N., Dottori, M., 2012. Detection of mosquito-only flaviviruses in Europe. *J. Gen. Virol.* 93, 1215-1225.

Calzolari, M., Ze-Ze, L., Vazquez, A., Sanchez Seco, M.P., Amaro, F., Dottori, M., 2016. Insect-specific flaviviruses, a worldwide widespread group of viruses only detected in insects. *Infect. Genet. Evol.* 40, 381-388. doi: 10.1016/j.meegid.2015.07.032.

Calistri, P., Giovannini, A., Hubalek, Z., Ionescu, A., Monaco, F., Savini, G., Lelli, R., 2010. Epidemiology of West Nile in Europe and in the Mediterranean Basin. *Open Virol. J.* 4, 29-37. doi: 10.2174/1874357901004020029.

Cook, S., Holmes, E.C., 2006. A multigene analysis of the phylogenetic relationships among the flaviviruses (Family: Flaviviridae) and the evolution of vector transmission. *Arch. Virol.* 151, 309-325.

Cook, S., Moureau, G., Harbach, R.E., Mukwaya, L., Goodger, K., Ssenfuka, F., Gould, E., Holmes, E.C., de Lamballerie, X., 2009. Isolation of a novel species of flavivirus and a new strain of *Culex flavivirus* (Flaviviridae) from a natural mosquito population in Uganda. *J. Gen. Virol.* 90, 2669-2678.

Cook, S., Moureau, G., Kitchen, A., Gould, E.A., de Lamballerie, X., Holmes, E.C., Harbach, R.E., 2012. Molecular evolution of the insect-specific flaviviruses. *J. Gen. Virol.* 93, 223-234. doi: 10.1099/vir.0.036525-0.

Crochu, S., Cook, S., Attoui, H., Charrel, R.N., De Chesse, R., Belhouchet, M., Lemasson, J.J., de Micco, P., de Lamballerie, X., 2004. Sequences of flavivirus-related RNA viruses persist in DNA form integrated in the genome of *Aedes* spp. mosquitoes. *J. Gen. Virol.* 85, 1971-1980.

Ergunay, K., Whitehouse, C., Ozkul, A., 2011. Current status of human arboviral infections in Turkey. *Vector Borne Zoonotic Dis.* 11, 731-741.

Ergunay, K., Gunay, F., Oter, K., Kasap, O.E., Orsten, S., Akkutay, A.Z., Erdem, H., Ozkul, A., Alten, B., 2013. Arboviral surveillance of field-collected mosquitoes reveals circulation of West Nile virus lineage 1 strains in Eastern Thrace, Turkey. *Vector Borne Zoonotic Dis.* 13, 744-752. doi: 10.1089/vbz.2012.1288.

Ergunay, K., Gunay, F., Erisoz Kasap, O., Oter, K., Gargari, S., Karaoglu, T., Tezcan, S., Cabalar, M., Yildirim, Y., Emekdas, G., Alten, B., Ozkul, A., 2014. Serological, molecular and entomological surveillance demonstrates widespread circulation of West Nile virus in Turkey. *PLoS Negl. Trop. Dis.* 8, e3028. doi: 10.1371/journal.pntd.0003028.

Ergünay, K., Günay, F., Sarıkaya, Y., Kar, S., Örsten, S., Öter, K., Litzba, N., Domingo, C., Erisoz Kasap, Ö., Özkul, A., Mitchell, L., Nitsche, A., Alten, B., Linton, Y.M., 2016. Active biosurveillance reveals first records of Chikungunya virus from Turkey: implications for public health in Turkey and Eastern Europe. *Parasit. Vectors* (submitted).

Folmer, O., Black, M., Hoeh, W., Lutz, R., Vrijenhoek, R., 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Mol. Mar. Biol. Biotechnol.* 3, 294-299.

Gritsun, D.J., Jones, I.M., Gould, E.A., Gritsun, T.S., 2014. Molecular archaeology of Flaviviridae untranslated regions: duplicated RNA structures in the replication enhancer of flaviviruses and pestiviruses emerged via convergent evolution. *PLoS One* 9, e92056. doi: 10.1371/journal.pone.0092056.

Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* 41, 95-98.

Hall-Mendelin, S., McLean, B.J., Bielefeldt-Ohmann, H., Hobson-Peters, J., Hall, R.A., van den Hurk, A.F., 2016. The insect-specific Palm Creek virus modulates West Nile virus infection in and transmission by Australian mosquitoes. *Parasit. Vectors* 9, 414. doi: 10.1186/s13071-016-1683-2.

Hoshino, K., Isawa, H., Tsuda, Y., Yano, K., Sasaki, T., Yuda, M., Takasaki, T., Kobayashi, M., Sawabe, K., 2007. Genetic characterization of a new insect flavivirus isolated from *Culex pipiens* mosquito in Japan. *Virology* 359, 405-414.

Huson, D.H., Beier, S., Flade, I., Gorska, A., El-Hadidi, M., Mitra, S., Ruscheweyh, H.J., Tappu, R., 2016. MEGAN community edition - interactive exploration and analysis of large-scale microbiome sequencing data. *PLoS Comput. Biol.* 12, e1004957.

Kuno, G., Chang, G.J., Tsuchiya, K.R., Karabatsos, N., Cropp, C.B., 1998. Phylogeny of the genus *Flavivirus*. *J. Virol.* 72, 73-83.

Marchler-Bauer, A., Derbyshire, M.K., Gonzales, N.R., Lu, S., Chitsaz, F., Geer, L.Y., Geer, R.C., He, J., Gwadz, M., Hurwitz, D.I., Lanczycki, C.J., Lu, F., Marchler, G.H., Song, J.S., Thanki, N., Wang, Z., Yamashita, R.A., Zhang, D., Zheng, C., Bryant, S.H., 2015. CDD: NCBI's conserved domain database. *Nucleic Acids Res.* 43, D222-D226.

Moureau, G., Ninove, L., Izri, A., Cook, S., de Lamballerie, X., Charrel, R.N., 2010. *Flavivirus* RNA in phlebotomine sandflies. *Vector Borne Zoonotic Dis.* 10, 195-197.

Moureau, G., Cook, S., Lemey, P., Nougairede, A., Forrester, N.L., Khasnatinov, M., Charrel, R.N., Firth, A.E., Gould, E.A., de Lamballerie, X., 2015. New insights into flavivirus evolution, taxonomy and biogeographic history, extended by analysis of canonical and alternative coding sequences. *PLoS One* 10, e0117849.

Newman, C.M., Anderson, T.K., Goldberg, T.L., 2016. Decreased flight activity in *Culex pipiens* (Diptera: Culicidae) naturally infected with *Culex flavivirus*. *J. Med. Entomol.* 53, 233-236. doi: 10.1093/jme/tjv161.

Parreira, R., Cook, S., Lopes, A., de Matos, A.P., de Almeida, A.P., Piedade, J., Esteves, A., 2012. Genetic characterization of an insect-specific flavivirus isolated from *Culex theileri* mosquitoes collected in southern Portugal. *Virus Res.* 167, 152-161.

Petersen, T.N., Brunak, S., von Heijne, G., Nielsen, H., 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods* 8, 785-786. doi: 10.1038/nmeth.1701.

Ramsdale, C.D., Alten, B., Caglar, S.S., Ozer, N., 2001. A revised, annotated checklist of the mosquitoes (Diptera, Culicidae) of Turkey. *Eur. Mosq. Bull.* 9, 18-28.

Sanchez-Seco, M.P., Vazquez, A., Collao, X., Hernandez, L., Aranda, C., Ruiz, S., Escosa, R., Marques, E., Bustillo, M.A., Molero, F., Tenorio, A., 2010. Surveillance of arboviruses in Spanish wetlands: detection of new flavi- and phleboviruses. *Vector Borne Zoonotic Dis.* 10, 203-206.

Simmonds, P., Becher, P., Collett, M.S., Gould, E.A., Heinz, F.X., Meyers, G., Monath, T., Pletnev, A., Rice, C.M., Stiasny, K., Thiel, H.-J., Weiner, A., Bukh, J., 2011. Family *Flaviviridae*. in: *Virus taxonomy: classification and nomenclature of viruses. Ninth Report of the International Committee on Taxonomy of Viruses.* Elsevier, San Diego, pp. 1003-1020.

Song, J., Tan, H., Perry, A.J., Akutsu, T., Webb, G.I., Whisstock, J.C., Pike, R.N., 2012. PROSPER: an integrated feature based tool for predicting protease substrate cleavage sites. *PLoS One* 7, e50300.

Stollar, V., Thomas, V.L., 1975. An agent in the *Aedes aegypti* cell line (Peleg) which causes fusion of *Aedes albopictus* cells. *Virology* 64, 367-377.

Vasilakis, N., Tesh, R.B., 2015. Insect-specific viruses and their potential impact on arbovirus transmission. *Curr. Opin. Virol.* 15, 69-74. doi: 10.1016/j.coviro.2015.08.007.

Vazquez, A., Sanchez-Seco, M.P., Palacios, G., Molero, F., Reyes, N., Ruiz, S., Aranda, C., Marqués, E., Escosa, R., Moreno, J., Figuerola, J., Tenorio, A., 2012. Novel flaviviruses detected in different species of mosquitoes in Spain. *Vector Borne Zoonotic Dis.* 12, 223-229.

Villordo, S.M., Carballeda, J.M., Filomatori, C.V., Gamarnik, A.V., 2016. RNA structure duplications and flavivirus host adaptation. *Trends Microbiol.* 24, 270-283. doi: 10.1016/j.tim.2016.01.002.

Tamura, K., Peterson, D., Peterson, N., Filipski, A., Kumar, S., 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol. Biol. Evol.* 30, 2725-2729.

Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22, 4673-4680.

FIGURE LEGENDS

Figure 1: Maximum likelihood analysis of the complete flavivirus nucleotide sequences. Virus strains included in the analysis are indicated with GenBank accession number, virus and isolate name. Bootstrap values lower than 60 are not shown. Viruses characterized in this study are labelled.

Figure 2: Maximum likelihood analysis of the putative viral polyprotein sequences. Virus strains included in the analysis are indicated with GenBank accession number, virus and isolate name. Bootstrap values lower than 60 are not shown. Viruses characterized in this study are labelled.

Figure 3: Maximum likelihood analysis of the flaviviral NS5 sequences. Virus strains included in the analysis are indicated with GenBank accession number, virus and isolate name. Bootstrap values lower than 75 are not shown. Viruses characterized in this study are labelled.

Table 1: Features of the culex theileri virus Turkey (CTFVtur) isolates characterized in the study

Isolate	Source	Size*	Species	Location	Coordinates	Date	Accession No.
CTFVtur1	Mosquito.pool	20	<i>Cx. theileri</i>	Canakkale - Ayvacik	39°29'36.1" N, 26°19'26.7" E	August, 2015	KX652375
CTFVtur2	Mosquito pool	20	<i>Cx. theileri</i>	Canakkale - Ayvacik	39°29'36.1" N, 26°19'26.7" E	August, 2015	KX652376
CTFVtur3	Mosquito pool	43	<i>Cx. theileri</i>	Kirklareli – Bostanli	41°36'48.6" N, 26°57'56.3" E	August, 2015	KX652378
CTFVtur4	Mosquito pool	7	<i>Cx. theileri</i>	Kirklareli - Karahamza	41°52'12.4"N, 27°00'44.2"E	August, 2014	KX652377

* all pools comprise female mosquitoes

Table 2: Pairwise comparison of the nucleotide and putative amino acid (in parantheses) sequences of the complete polyprotein coding region of the culex theileri flaviviruses isolated in Turkey (CTFVtur 1-4) and related culex theileri flavivirus (CTFV) strains (153 and 178) from Portugal. Similarity values were indicated as percent.

	CTFV153	CTFV178	CTFVtur1	CTFVtur2	CTFVtur3	CTFVtur4
CTFV153	-	99.04 (99.31)	94.71 (99.01)	94.74 (98.92)	94.70 (98.89)	94.93 (98.98)
CTFV178	99.04 (99.31)	-	94.56 (98.80)	94.59 (98.65)	94.51 (98.68)	94.76 (98.77)
CTFVtur1	94.71 (99.01)	94.56(98.80)	-	99.14 (99.73)	99.09 (99.73)	99.46 (99.85)
CTFVtur2	94.74 (98.92)	94.59 (98.65)	99.14 (99.73)	-	99.27 (99.73)	99.64 (99.85)
CTFVtur3	94.70 (98.89)	94.51 (98.68)	99.09 (99.73)	99.27 (99.73)	-	99.61 (99.88)
CTFVtur4	94.93 (98.98)	94.76 (98.77)	99.46 (99.85)	99.64 (99.85)	99.61 (99.88)	-

Table 3: Nucleotide alignment and comparison of the tandem repeats observed in insect-specific flaviviruses. The strains involved are as follows, indicated with isolate name and GenBank accession number: *Culex theileri* flavivirus Turkey: CTFVtur1 (KX652375), CTFVtur2 (KX652376), CTF Vtur3 (KX652378), CTF Vtur4 (KX652377); *Culex theileri* flavivirus: 153 (HE574573), 178 (HE574574); Kamiti River virus: SR-82 (AY149905), SR-75 (AY149904); Cell fusing agent virus: Culebra (AH015271), Galveston (KJ741267), YFVCFAPP (M91671); *Culex* flavivirus: Toyama2627 (AB701776), Toyama1431 (AB701775), Toyama861 (AB701774), Toyama791 (AB701773), Toyama740 (AB701772) Toyama75 (AB701771) Toyama41 (AB701770), Toyama1849 (AB701769), Toyama1701 (AB701768) Toyama734 (AB701767) Toyama71 (AB701766), DG1007 (JQ308190), DG1068 (JQ308189), DG1064 (JQ308188), HLD105 (JQ308187), HLD102 (JQ308186), H0901 (HQ678513), Iowa07 (FJ663034), HOU24518 (FJ502995), NIID-21-2 (AB377213), Tokyo (AB262759) Mex07 (EU879060).

Insect Flavivirus	Repeated Sequence (5'-3')
<i>Culex theileri</i> flavivirus Turkey	TGACGACA --- CCCCCTCCCCAGTCCCTTAGGTC- TGACGACA --- CCCC GGCCG CAGTCCCCAAGGTC TGACGACA --- CCCC GGCCCCAGTCCCCAAGGTC
<i>Culex theileri</i> flavivirus	TGACTACT --- CCCCCTCCCC -GTCCCTTAGGTC- TGACGACA --- CCCC G CCCCCA -GTCCCTTAGGTC- TGACGACA --- CCCC GGCCCCAGTCCCCAAGGTC
<i>Culex</i> flavivirus	TGAC - AA - CGCTCCGGCCCCAGTCCCC----- TGAC - CA - TACCCCGACCCCAGTCCGA----- GAACG - A - CACCCCGGCCCCAGTTCTC----- TAACG - A - CACTCCGGCCCCAGTTCCC-----
Cell fusing agent virus	TGACG - CACCCGCTCC - CCTCAGTCCCC----- TGACG - AACCCGCTCC - CCTCAGTCCCC-----
Kamiti River virus	TGACG - CACCCGCTCC - CCCGAGTCCCC-----

Table 4: Comparison of the putative amino acid sequences of the culex theileri flavivirus Turkey (CTFVtur) isolates with related culex theileri flavivirus (CTFV) strain 153 and other mosquito-specific and mosquito-borne flaviviruses. Viruses included in the analyses and GenBank accession numbers are: culex flavivirus strain Tokyo (CxFTokyo, AB262759), Quang Binh virus (QBV, NC012671), cell fusion agent virus (CFAV, NC001564), West Nile virus strain T2 from Turkey (WNVT2, KJ958922), Japanese encephalitis virus (JEV, NC001437)

	CTFVtur1		CTFV153		CxFTokyo		QBV		CFAV		WNVT2		JEV	
	Size	ID%	Size	ID%	Size	ID%	Size	ID%	Size	ID%	Size	ID%	Size	ID%
C+AC	136	136	97.7	139	43.8	136	47.8	136	51.7	123	12.3	127	17.3	
PrM+M	142	142	100	143	76.2	142	76.7	142	77.4	167	15.7	167	14.2	
E	427	427	98.8	427	74.4	427	77.5	427	81.0	501	15.3	500	16.8	
NS1	369	369	99.4	369	70.7	369	73.9	390	37.4	352	22.0	352	21.2	
NS2a	228	228	98.2	230	63.0	229	63.7	232	18.7	231	12.9	227	12.1	
NS2b	143	143	100	142	62.9	143	53.1	124	15.7	131	12.4	131	15.1	
NS3	577	577	99.4	578	70.9	578	77.5	577	44.7	619	29.5	620	30.0	
NS4a	189	189	100	189	53.4	188	59.7	145	16.4	126	9.4	125	8.9	
NS4b	257	257	97.6	257	62.2	258	60.0	280	17.1	278	12.2	278	13.6	
NS5	889	889	98.8	889	82.0	889	84.4	888	60.1	905	43.3	905	43.2	
ORF	3357	3357		3363		3359		3341		3433		3432		

Table 5: Amino acid variations observed in the putative polyprotein of culex theileri flavivirus Turkey isolates (CTFVtur 1-4) and related culex theileri flavivirus (CTFV) strains (153 and 178) from Portugal.

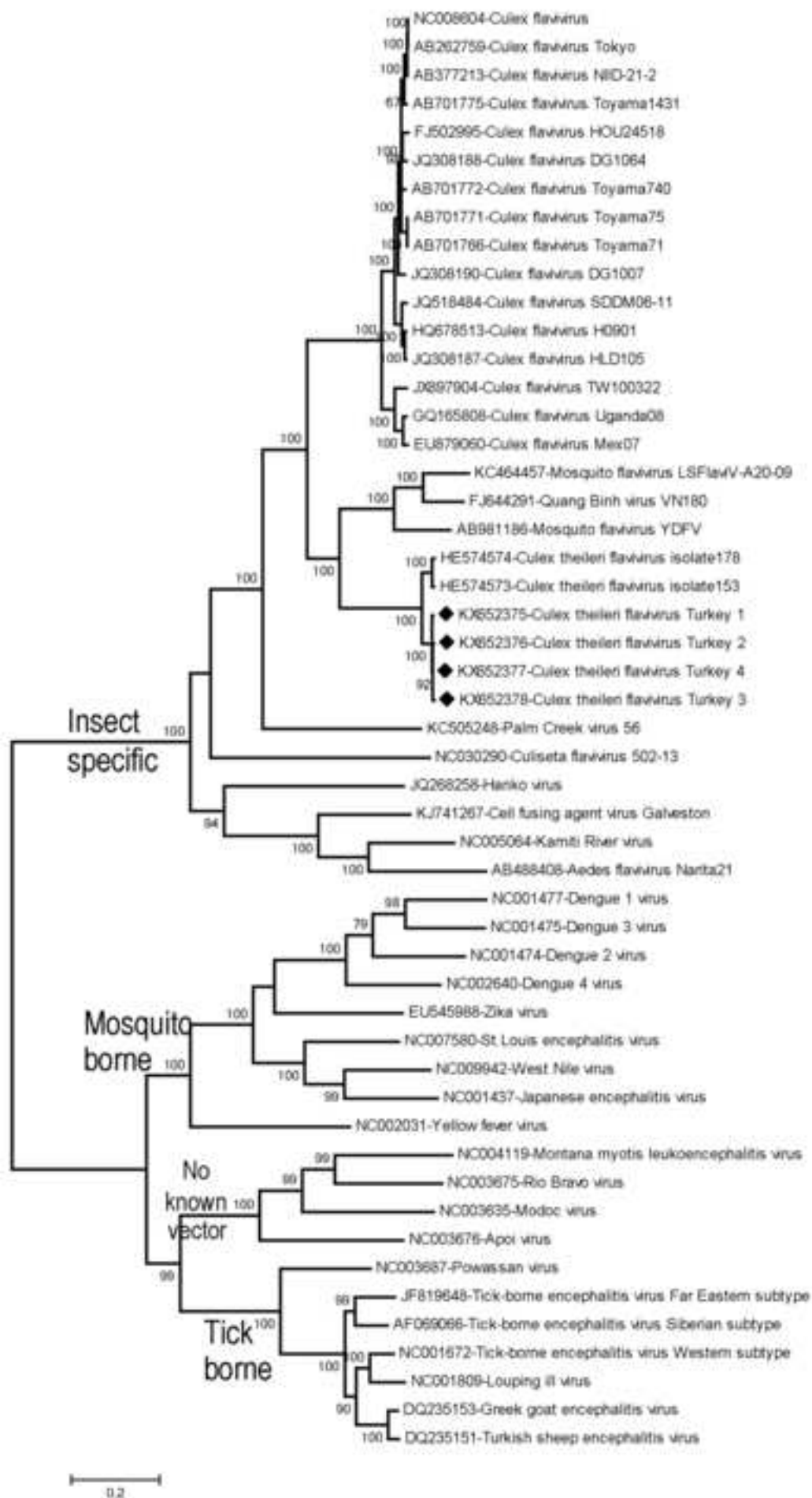
Gene	Position	Strain					
		CTFV153	CTFV178	CTFVtur1	CTFVtur2	CTFVtur3	CTFVtur4
C	39	K	K	R	R	R	R
	44	H	H	H	S	N	N
Anchor C	121	L	L	I	L	L	L
	123	I	I	M	M	M	M
PrM	149	I	I	I	M	M	M
M	258	A	S	A	A	A	A
E	416	F	Y	F	F	F	F
	490	D	D	E	E	E	E
	523	N	D	D	D	D	E
	543	V	A	A	A	A	A
	604	R	R	K	K	K	A
	667	F	F	L	L	L	L
	715	R	Q	R	R	R	R
NS1	853	A	A	S	S	S	S
	973	F	F	Y	Y	Y	Y
	995	Y	H	Y	Y	Y	Y
	1013	S	P	S	S	S	S
	1020	R	G	R	R	R	R
	1036	C	R	C	C	C	C
	1078	T	T	P	P	P	P
NS2a	1090	I	I	V	I	I	I
	1104	N	N	N	N	H	N
	1239	T	T	M	M	M	M
	1269	V	V	V	M	V	V
	1272	A	A	T	T	T	T
NS3	1572	S	N	S	S	S	S
	1661	K	K	K	R	R	R
	1750	A	A	T	T	T	T
	1769	N	N	S	S	S	S
	1789	S	T	S	S	S	S
	1796	D	D	D	D	V	D
	1835	R	R	R	W	R	R
	1836	P	P	P	S	P	P
1872	S	G	G	G	G	G	
NS4b	2230	S	S	T	T	T	T
	2244	A	A	T	T	T	T
	2248	L	L	P	P	P	P
	2249	T	T	M	M	L	M
	2355	K	K	R	R	R	R
NS5	2408	I	I	V	V	V	V
	2561	L	P	L	L	L	L
	2604	G	R	G	G	G	G
	2658	S	N	N	S	N	N
	2709	R	R	L	L	L	L
	2714	K	K	K	K	R	K
	2742	L	V	L	L	L	L
	2808	P	S	S	S	S	S
	2836	T	T	I	I	I	I
	2888	L	L	V	V	V	V
	2953	N	D	D	D	D	D
2992	A	V	V	V	V	V	
3233	T	A	A	A	A	A	
3274	M	T	M	M	M	M	

3287	A	G	A	A	A	A
3327	K	K	R	R	R	R
3335	A	P	A	A	A	A
3355	I	I	V	V	V	V

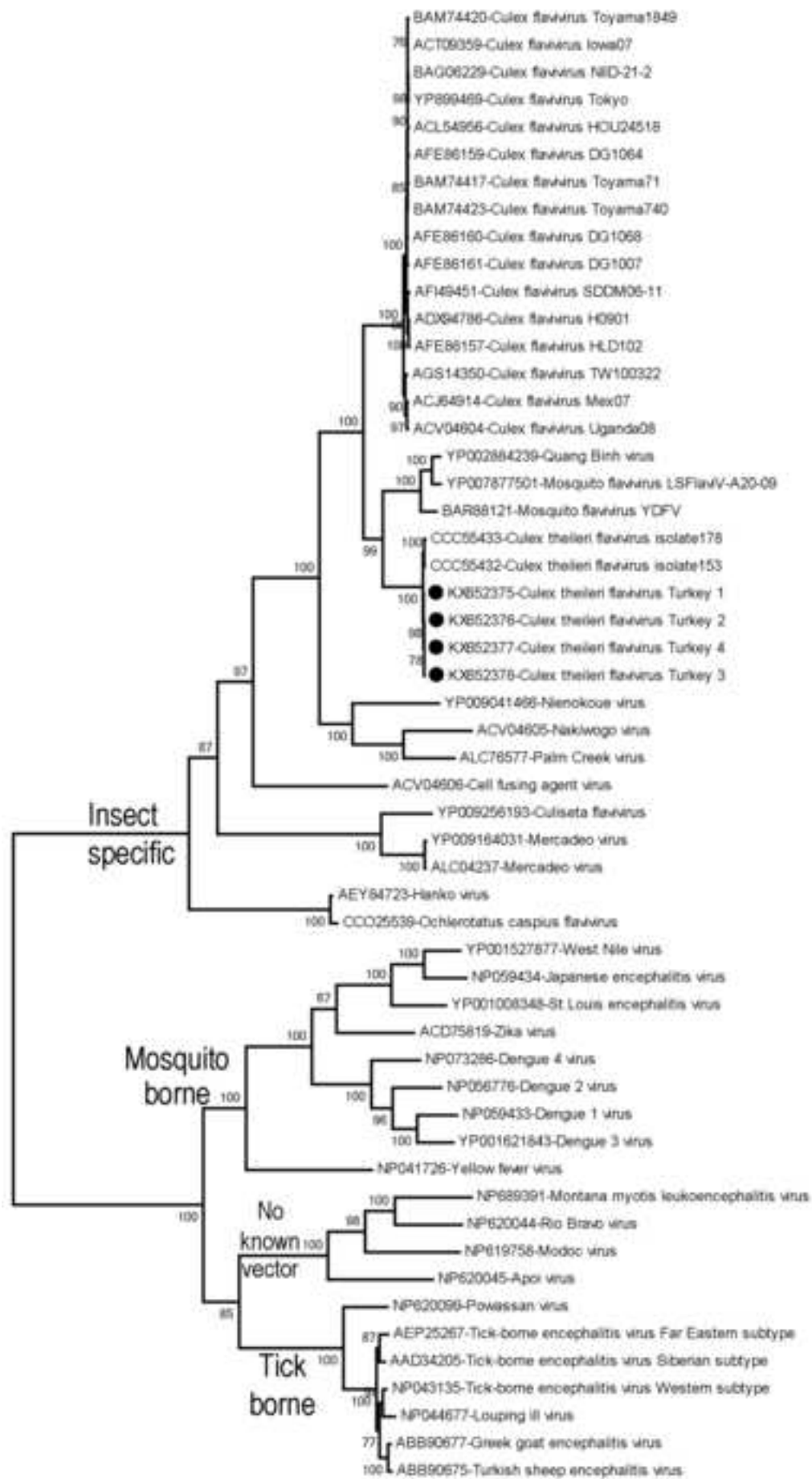
Table 6: Comparison of the fusion peptide motif in the envelope protein of culex theileri flavivirus Turkey (CFTVtur) with related culex theileri flavivirus (CTFV) and other major insect-specific flaviviruses. The motif encompasses 371-384. residues of the CFTVtur/CTFV genome. Conserved residues observed in the genus are indicated in bold.

Flavivirus	Fusion peptide motif
Kamithi River virus	NRGWGTGCFEWGLG
Aedes flavivirus	NRGWGTGCFEWGLG
Nakiwogo virus	NRGWGTGCLEWGLG
Cell fusing agent virus	NRGWGTGCFKWGLG
Culex flavivirus	NRGWGTGCFKWGLG
Culex theileri flavivirus	NRGWGTGCFKWGI G
Culex theileri flavivirus Turkey	NRGWGTGCFKWGI G

Figure(s)



Figure(s)



Figure(s)

