Biochemical and Biophysical Research Communications **287,** 455–461 (2001)
doi:10.1006/bbrc.2001.5610, available online at http://www.idealibrary.com on IDEAL®

Complete Coding Sequence of the Alkhurma Virus, a Tick-Borne Flavivirus Causing Severe Hemorrhagic Fever in Humans in Saudi Arabia

Rémi N. Charrel,*,†,[1] Ali Mohamed Zaki,‡ Houssam Attoui,* Mazen Fakeeh,‡ Frédérique Billoir,* Amany Ibrahim Yousef,‡ Reine de Chesse,* Philippe De Micco,* Ernest A. Gould,§ and Xavier de Lamballerie*,†

**Unité des Virus Emergents EA 3292, Université de la Méditerranée, Faculté de Médecine, 27 boulevard Jean Moulin, Marseille 13005, France; †Maladies virales émergentes & systèmes d'information UR 034, Institut de Recherche pour le Développement, France; ‡Virology Laboratory, Dr. Suliman Fakeeh Hospital, P.O. Box 2537, Jeddah, Saudi Arabia; and §National Environment Research Council Institute of Virology and Environmental Microbiology, Oxford, United Kingdom*

Received August 17, 2001

To date, tick-borne flaviviruses responsible for hemorrhagic fever in humans have been isolated in Siberia (Omsk hemorrhagic fever virus), India (Kyasanur Forest disease virus, KFDV), and in Saudi Arabia (Alkhurma virus, ALKV). Prior to this study, only partial coding sequences of these severe pathogens had been determined. We report here the complete coding sequence of ALK virus, which was determined to be 10,248 nucleotides (nt) long, and to encode a single 3,416 amino acid polyprotein. Independent analyses of the complete polyprotein and the envelope protein provided genetic and phylogenetic evidence that ALKV belongs to the tick-borne flavivirus group, within which it is most closely related to KFDV. Analysis of structural genes, genetic distances, and evolutionary relationship indicate that ALKV and KFDV derived from a common phylogenetic ancestor and constitute two genetic subtypes of the same virus species according to current genetic criteria of classification.

Key Words: **flavivirus; Flaviviridae; phylogeny; tick-borne virus; hemorrhagic fever; Alkhurma virus.**

In 1995, a virus tentatively named Alkhurma virus (ALKV) and related to the tick-borne (TB) flaviviruses (1) was isolated from the blood of several patients with severe hemorrhagic fever in Saudi Arabia. Since that time, a total of 16 cases has been confirmed by virus isolation, of which 4 had a fatal outcome. The discovery of this virus was considered to be an important event because TB flaviviruses responsible for hemorrhagic fever in humans had been isolated previously only in Siberia (Omsk hemorrhagic fever virus, OHFV) and in India (Kyasanur Forest disease virus, KFDV). The sequence determination of ALKV NS5 gene suggested that the virus was closely related to KFDV (1), one of the most pathogenic TB flaviviruses, causing hemorrhagic manifestations with a case fatality rate of 2 to 10% (2). It was first recognized in 1957 in the Kyasanur Forest (Shimoga District, India) (3) and causes annually an average of 500 cases. However, genetic comparison of ALKV with TB flaviviruses (an in particular with OHFV) was hampered by the lack of genetic information in the structural genes, which have been mostly studied in previous studies of TB-flaviviruses.

In this study, we report the sequence determination of the complete open reading frame (ORF) of the ALKV prototype strain. This constitutes the first complete genetic characterization of a TB-complex flavivirus responsible for hemorrhagic manifestations. Comparative analyses with other mammalian TB flaviviruses and phylogenetic studies are presented and discussed.

MATERIALS AND METHODS

Viral strain. The ALKV prototype strain 1176 was recovered in 1995 from the blood of a patient with fever, headache, retro-orbital pain, joint pain, generalized muscle pains, anorexia, and vomiting; the strain was successively passaged twice in suckling mouse brains, three times in Vero cells, once in sheep and finally once again in newborn mice. The current study was performed on viral RNA extracted from these newborn mice brains.

RNA extraction. RNA was extracted using the RNA NOW TC-Kit (Biogentex. Inc., Seabrook, TX) according to the manufacturer's instructions, resuspended into 50 μl of RNase-free sterile water, and stored at −70°C until processed.

[1] To whom correspondence should be addressed at Unité des Virus Emergents, Faculté de Médecine, 27, boulevard Jean Moulin, Marseille 13005, France. Fax: (33) 491 32 44 95. E-mail: rnc-virophdm@gulliver.fr.

0006-291X/01 $35.00

TABLE 1

Mammalian Tick-Borne Flaviviruses Used in the Phylogenetic (Fig. 3B) Study and Taxonomic Study (Fig. 4)

Virus species[a]	Subtypes and isolates	Abbreviation[a]	GenBank Accession No.	Sequence	Origin
Alkhurma virus	Human isolate	ALKV	AF331718	Complete	Saudi Arabia
Kyasanur Forest disease virus	Human isolate	KFDV	X74111	Envelope	India
Langat virus	*Ixodes granulatus*	LGTV	M73835	Complete	Malaysia
Omsk hemorrhagic fever virus	*Dermacentor marginatum*	OHFV	X66694	Envelope	Siberia
Powassan virus	Human isolate	POWV	L06436	Complete	Ontario
Deer-tick virus[b]	Deer-tick isolate	DTV	AF135461	Envelope	Wisconsin
Tick-borne encephalitis virus	European subtype	$TBEV^{HYPR}$	U39392	Complete	Eastern Europe
Tick-borne encephalitis virus	European subtype	$TBEV^{NEU}$	U27495	Complete	Eastern Europe
Tick-borne encephalitis virus	Far Eastern subtype	$TBEV^{SOF}$	GNWVTB	Complete[c]	Far-East Russia
Tick-borne encephalitis virus	Siberian subtype	$TBEV^{VAS}$	L40361	Complete	Siberia
Louping ill virus	Irish subtype	LIV^{IR}	CAA60480	Envelope	Ireland
Louping ill virus	British subtype	LIV^{BR}	D12937	Envelope	UK
Louping ill virus	Spanish subtype	LIV^{SSE}	CAA54619	Envelope	Spain
Louping ill virus	Turkish subtype	LIV^{TSE}	S41628	Envelope	Turkey
Louping ill virus	Greek subtype	LIV^{GGE}	X77732	Envelope	Greece
Louping ill virus	Other isolates	LIV^{NEG}	M94956	Envelope	Japan
Louping ill virus	Other isolates	LIV^{WAL}	CAA60481	Envelope	UK (Wales)
Louping ill virus	Other isolates	LIV^{369}	Y07863	Envelope	UK
Louping ill virus	Other isolates	LIV^{NOR}	D12936	Envelope	Norway

[a] As referred in the VIIth Report of the ICTV (11).
[b] Proposed as a genotype of POWV by Telford *et al.* (12).
[c] Amino acid sequence available only.

Reverse transcription. RT was carried out at 42°C in a 20-μl reaction mixture that included 11 μl of RNA extract, 200 U of Superscript II RNase H^- Reverse Transcriptase (Life Technologies, Inc., Grand Island, NY) and 2 pmol of oligonucleotide 3PNC-2R (5′-GCTCAGGGAGAACAAGAACCG-3′) located in the 3′ noncoding region. The reaction mixture was subsequently treated with DNase-free RNase (Roche Diagnostics, Meylan, France). The resulting non-infectious subgenomic cDNA was received at the Unité des Virus Emergents and further processed.

Polymerase chain amplification and sequencing reactions. DNA products were amplified by PCR from 10 overlapping regions (Fig. 1). PCRs were carried out in a volume of 50 μl that included 10 mM Tris–HCl [pH 9.0], 1.5 mM MgCl2, 50 mM KCl, 0.1% Triton X-100, 200 μM each dNTP, 0.2 μM of each primer, 5 μl of cDNA and 1.5 U of *Taq* DNA polymerase (Promega Corp., Madison, WI). The thermocycler profile was 5 min at 95°C, followed by 35 cycles of 30 s at 95°C, 1 min at 50°C, and 2 min at 72°C, with a 7-min final extension at 72°C. PCR products were purified from agarose gel (Wizard PCR Preps DNA Purification System, Promega Corp.) and directly sequenced using the ABI Prism Dye Terminator Cycle Sequencing Ready Reaction Kit (Perkin–Elmer Corp., Foster City, CA) and an ABI Prism 377XL automated sequencer (Perkin–Elmer Corp.).

Sequence data and phylogenetic analysis. Full-length coding sequences of TB flaviviruses were retrieved from the GenBank database (Table 1) together with the following sequences: yellow fever virus strain 17D (YFV, X03700); dengue 1 virus strain Western Pacific (DENV-1, M88535); dengue 2 virus strain New Guinea C (DENV-2, AF038403); dengue 3 virus strain H87 (DENV-3, M93130); dengue 4 virus (DENV-4, M14931); Kunjin virus strain MRM61C (KUNV, D00246); Japanese encephalitis virus strain JaOArS982 (JEV, M18370), West Nile virus strain Eg101 (WNV, AF260968); St. Louis encephalitis virus strain MI-7 (SLEV, AF160194); Murray Valley encephalitis virus strain MVE-1-51 (MVEV, NC 000943); mosquito cell fusing agent (CFAV, M91671); Rio Bravo virus strain M-64 (RBV, AF144692), Apoi virus strain Kitaoka (APOIV, AF160193).

These complete flavivirus amino acid (AA) sequences were aligned using the Clustal W 1.7 program (4). This permitted (i) to deduce the putative cleavage sites of the ALKV polyprotein and (ii) to calculate AA sequence identities using the pairwise distance algorithm of the MEGA software program (5).

Phylogenetic relationships were determined using these pairwise distances and the neighbor-joining (NJ) method implemented in MEGA. The robustness of the resulting branching patterns was tested by bootstrap analysis with 500 replications and the correspondence of phylogenetic groupings with previously described serocomplexes was analyzed.

An additional analysis was performed using partial AA sequences in the envelope region retrieved from the GenBank database (see Table 2). Alignments and phylogenetic analysis were performed as described above. The observed intra- and interspecies genetic distances were analyzed and used to discuss the taxonomic assignment of ALKV.

RESULTS AND DISCUSSION

The complete coding sequence of ALKV strain1176 was obtained by sequencing 10 overlapping PCR amplification products (Fig. 1).The coding sequence was determined to be 10,248 nucleotides (nt) in length, including a single ORF that encodes a 3,416 amino acid (AA) polyprotein (GenBank Accession No. AF331718). Comparison with the complete AA sequences of other flaviviruses clearly confirmed that ALKV is closely related to mammalian TB flaviviruses. In particular, the analysis of complete sequence data (Fig. 2) showed that the genetic distance between ALKV and non-TB flaviviruses is higher or equal to 57%, whereas it ranged from 21 to 24.3% between ALKV and TB flaviviruses, a value within the range observed between previously characterized members of the TB group (<25.6%). Sim-

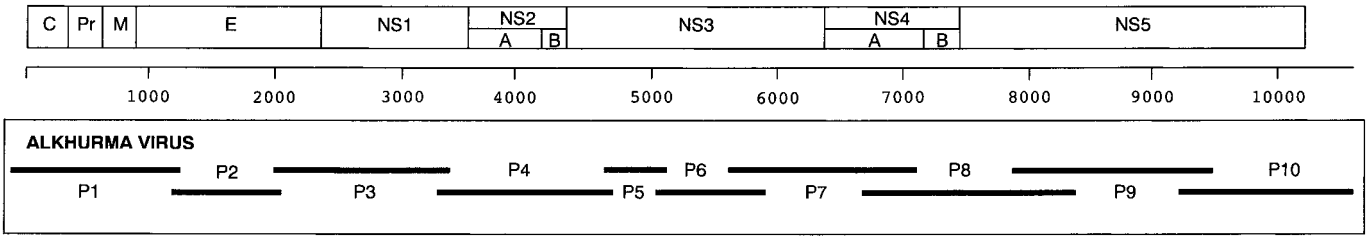
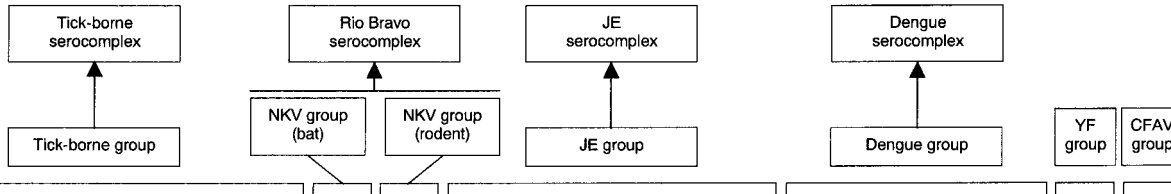


FIG. 1. Strategy used for sequencing the genome of ALKV. The topmost diagram shows a typical flavivirus full-length coding sequence. The different genes and their respective size are indicated. Each graduation corresponds to 1000 nucleotides. Sequencing strategy of ALKV genome was based on the amplification of 10 overlapping PCR products (P1 to P10). Primers sequences are available upon simple request to the corresponding author.

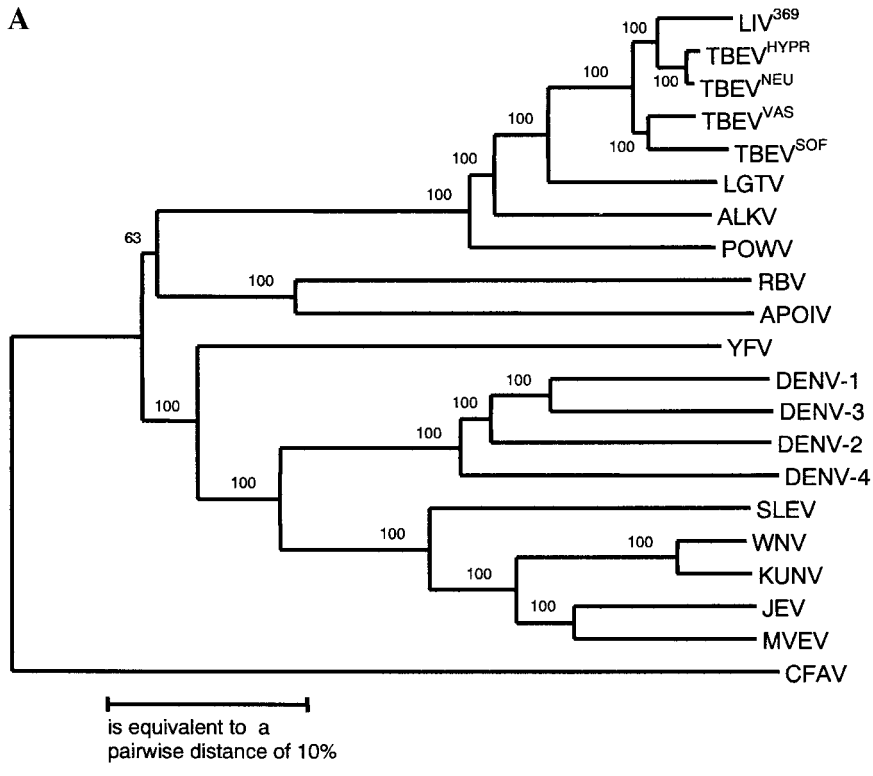
ilarly, the phylogenetic analysis performed with complete coding sequences (Fig. 3A) permitted to include ALKV in the TBE group with a 100% bootstrap value. Finally, the analysis of AA alignments permitted unambiguous identification of the putative cleavage sites of ALKV polyprotein (Table 2A), by reference to previously determined cleavage sites (6). The AnchC, PrM, M, and NS2A, NS2B, NS3, NS4A, 2K, NS4B and NS5 proteins are identical in length in all the TB flaviviruses included in this study (Table 2B). [The NS4B and NS5 proteins of TBEV^{SOF} reported here are slightly different in length from those of the sequence deposited in databases (GNWVTB). They were deduced from a corrected partial sequence of this strain (kindly provided by Dr. Pletnev and Dr. T. Gritsun, personal data), and conform with those reported for all other TBEV strains sequenced to date. Accordingly, all TBEV, LIV and LGTV strains have putative structural and nonstructural proteins that are identical in length.] The 3 genes for which differences were observed are VirC, E and NS1. The E protein is 496 AA long for all TB flaviviruses, including ALKV and KFDV; the only exception is POWV (497 AA). VirC is 96 AA long in TBEV, LIV, LGTV, 94 AA long in POWV, and 97 AA long for both ALKV and KFDV. NS1 protein is 353 AA long for TBEV, LIV and LGTV, and 354 AA long for ALKV and POWV (not determined for KFDV) (Table 2B). Therefore, ALKV possesses the longest polyprotein of all TB-flaviviruses characterized so far.

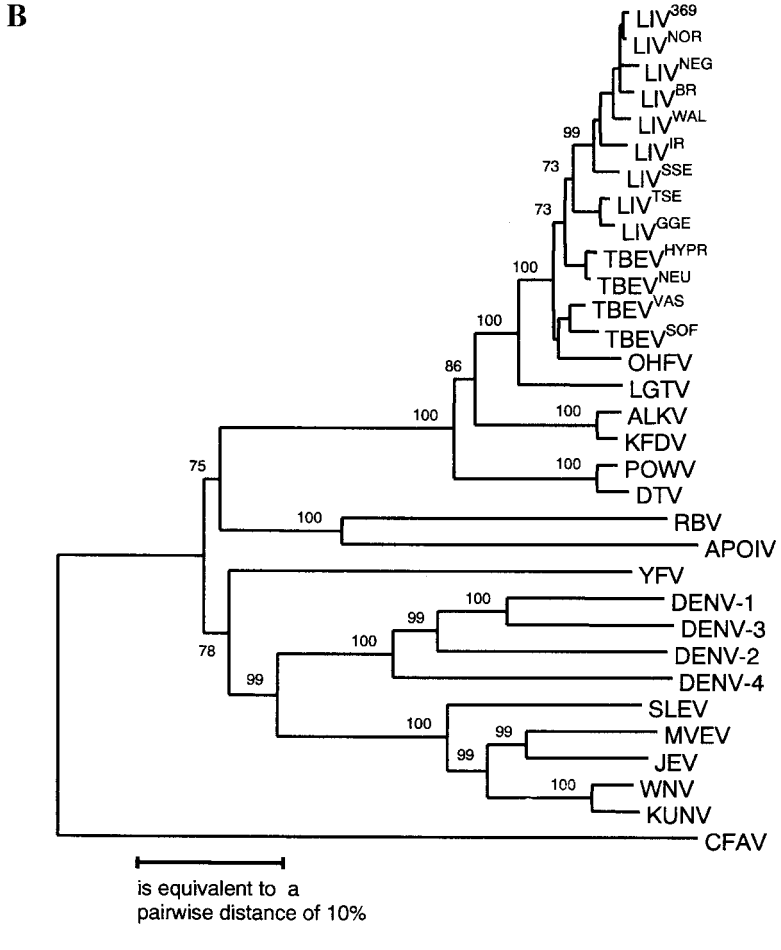
The relatedness of ALKV with TB flaviviruses was also demonstrated by the analysis of the E protein (7). A number of conserved patterns could be identified including: (i) the 12 cysteine residues involved in the intramolecular disulfide bonds; (ii) the 3 potential N-glycosylation sites (positions 154–156, 361–363, and 473–475); (iii) the TB-flavivirus-specific pentapeptide DSGHD previously described (8) at positions 320-324; (iv) the sequence of the ALKV fusion peptide (positions 98–111) was identical to that of TBEV, LIV, LGTV, and KFDV, and 1 AA different from that of POWV.



	ALKV	LIV	TBEV^{HYPR}	TBEV^{NEU}	LGTV	POWV	RBV	APOIV	JEV	MVEV	KUNV	WNV	SLEV	DEN4	DEN2	DEN1	DEN3	YFV	CFAV
ALKV	-	0.221	0.214	0.210	0.220	0.243	0.572	0.575	0.593	0.593	0.590	0.589	0.591	0.604	0.597	0.599	0.600	0.574	0.736
LIV	0.296	-	0.058	0.056	0.174	0.247	0.580	0.573	0.596	0.598	0.588	0.592	0.592	0.607	0.608	0.606	0.603	0.571	0.734
TBEV^{HYPR}	0.294	0.121	-	0.012	0.162	0.239	0.577	0.571	0.596	0.594	0.586	0.589	0.590	0.605	0.604	0.603	0.598	0.570	0.732
TBEV^{NEU}	0.294	0.119	0.027	-	0.160	0.238	0.577	0.570	0.594	0.592	0.584	0.587	0.589	0.605	0.604	0.602	0.597	0.569	0.733
LGTV	0.294	0.257	0.253	0.254	-	0.245	0.575	0.578	0.598	0.600	0.589	0.592	0.591	0.604	0.597	0.603	0.603	0.575	0.735
POWV	0.313	0.307	0.299	0.302	0.305	-	0.580	0.572	0.587	0.591	0.588	0.590	0.588	0.601	0.599	0.598	0.595	0.572	0.733
RBV	0.492	0.486	0.482	0.482	0.484	0.485	-	0.458	0.621	0.622	0.622	0.622	0.619	0.626	0.623	0.617	0.623	0.611	0.736
APOIV	0.487	0.480	0.476	0.477	0.479	0.478	0.392	-	0.630	0.627	0.625	0.628	0.626	0.628	0.628	0.621	0.622	0.614	0.738
JEV	0.501	0.495	0.493	0.494	0.499	0.499	0.521	0.523	-	0.182	0.234	0.239	0.328	0.490	0.484	0.486	0.483	0.538	0.740
MVEV	0.502	0.505	0.502	0.504	0.504	0.501	0.519	0.523	0.276	-	0.231	0.236	0.311	0.492	0.484	0.488	0.486	0.538	0.744
KUNV	0.502	0.498	0.493	0.495	0.498	0.499	0.512	0.518	0.308	0.301	-	0.072	0.323	0.487	0.475	0.479	0.481	0.542	0.741
WNV	0.498	0.495	0.494	0.495	0.499	0.494	0.521	0.523	0.300	0.299	0.208	-	0.323	0.489	0.478	0.486	0.486	0.543	0.745
SLEV	0.507	0.502	0.504	0.504	0.505	0.506	0.525	0.522	0.348	0.344	0.345	0.337	-	0.482	0.471	0.479	0.475	0.535	0.737
DEN4	0.506	0.508	0.507	0.508	0.509	0.509	0.522	0.521	0.436	0.432	0.434	0.437	0.420	-	0.313	0.321	0.308	0.556	0.750
DEN2	0.502	0.505	0.506	0.507	0.507	0.496	0.512	0.523	0.428	0.433	0.430	0.424	0.419	0.327	-	0.277	0.279	0.543	0.750
DEN1	0.508	0.509	0.505	0.507	0.512	0.504	0.512	0.520	0.431	0.430	0.425	0.428	0.423	0.336	0.314	-	0.219	0.542	0.744
DEN3	0.504	0.504	0.504	0.503	0.505	0.502	0.514	0.522	0.433	0.425	0.432	0.432	0.426	0.330	0.315	0.282	-	0.549	0.745
YFV	0.491	0.482	0.486	0.486	0.485	0.491	0.510	0.513	0.459	0.462	0.458	0.460	0.457	0.477	0.465	0.463	0.465	-	0.745
CFAV	0.577	0.565	0.565	0.565	0.567	0.568	0.588	0.591	0.578	0.582	0.576	0.578	0.581	0.583	0.580	0.575	0.575	0.571	-

FIG. 2. Matrix of genetic distances. The upper-right matrix represents pairwise distances between standard amino acids alignments. The lower-left matrix represents pairwise distances between standard nucleotides alignments. Groupings at the top of the figure were made using 0.450 and 0.380 cutoff values for amino acids and nucleotides respectively. Correspondence with serocomplexes is indicated.


A
LIV369
TBEVHYPR
TBEVNEU
TBEVVAS
TBEVSOF
LGTV
ALKV
POWV
RBV
APOIV
YFV
DENV-1
DENV-3
DENV-2
DENV-4
SLEV
WNV
KUNV
JEV
MVEV
CFAV
is equivalent to a
pairwise distance of 10%


B
LIV369
LIVNOR
LIVNEG
LIVBR
LIVWAL
LIVIR
LIVSSE
LIVTSE
LIVGGE
TBEVHYPR
TBEVNEU
TBEVVAS
TBEVSOF
OHFV
LGTV
ALKV
KFDV
POWV
DTV
RBV
APOIV
YFV
DENV-1
DENV-3
DENV-2
DENV-4
SLEV
MVEV
JEV
WNV
KUNV
CFAV
is equivalent to a
pairwise distance of 10%

Further analysis of AA motifs showed that ALKV is more closely related to KFDV (7): (i) the KFDV specific AKG motif at positions 2–4 of the VirC protein is also present in the ALKV polyprotein; (ii) the insertion of 1 AA residue at position 93 of the VirC protein is found only in ALKV and KFDV polyproteins; this residue is an arginine for both viruses; (iii) the KFDV specific EHLPKA hexapeptide at positions 207–212 of the E protein is present in the ALKV polyprotein (8); (iv) the KFDV specific EGSK motif (position 308–311 of the E protein) that was related to the non reactivity of Mab 4.2 with KFDV E protein is identical in ALKV (9); (v) ALKV and KFDV share the same AA substitutions at positions 76 and 489 of the E protein.

However, it is possible to distinguish between ALKV and KFDV in the hypervariable region of the E protein (positions 232–234) (10) where ALKV encodes the sequence AHE distinct from the AQE motif of KFDV (7). Therefore, whether ALKV represents a species distinct from KFDV or a new subtype within the KFDV species required further investigation.

To address this question, phylogenetic analysis was performed within a 496 AA region of the envelope gene where sequences are available for 19 TB flaviviruses. The topology of the tree obtained (Fig. 3B) is consistent with that observed using complete ORF sequences but brings the additional information that ALKV and KFDV derive from a common ancestor (assessed by a 100% bootstrap value). AA distances were used to study the distribution of evolutionary distances upon pairwise comparison (Fig. 4A). The bimodal shape of the curve revealed a cutoff value at 14%, above which AA distances correspond to strains that belong to distinct TB-virus species according to the ICTV classification (11). The only noticeable exception was represented by LIV and TBEV strains that could not be differentiated using this cutoff value. The distance ranges within TBEV strains (up to 7.3%, Fig. 4B) or LIV strains (up to 4.8%, Fig. 4C), and between TBEV and LIV strains (3.8–9.3%, Fig. 4D) are widely overlapping. There is therefore no evidence from the distribution of genetic distances that TBEV and LIV constitute distinct species. However, in the case of ALKV and KFDV, the observed 3.0% distance is compatible only with these viruses belonging to the same virus species suggesting that ALKV may be considered a genetic subtype of KFDV.

The ecological conditions that determine virus survival and propagation are very different for KFDV in India (3) compared with ALKV in Saudi Arabia. Hu-

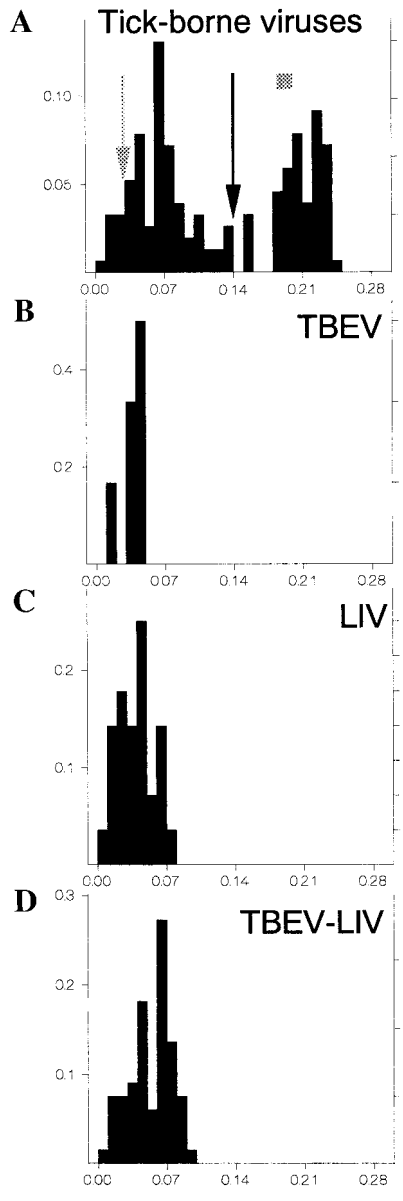


FIG. 4. Distribution of evolutionary distances upon pairwise comparison. The distances were calculated from a 496-AA region located in the E protein (position 1–496) The genetic distance is reported on the x-axis. Frequency of genetic distances is recorded on the y-axis. (A) Distribution of the distances observed between the 19 strains of tick-borne flaviviruses presented in Table 1. The shaded square represents AA distances (18.7–20.8%) observed between ALKV and LIV, TBEV, LGTV strains. The gray arrow corresponds to the AA distance observed between ALKV and KFDV. The black arrow represents the cutoff for species identification. (B) Distribution of the distances observed within the TBEV strains. (C) Distribution of the distances observed within the LIV strains. (D) Distribution of the distances observed between TBEV and LIV strains.

FIG. 3. Phylogenetic analysis of flaviviruses based on complete polyprotein (A) and a 496-AA region of the envelope (B). Distances and groupings were determined by the pairwise distance algorithm and the neighbor-joining method using the MEGA software program (5). Bootstrap values are indicated and correspond to 500 replications. In A, the European subtype TBEV strains group with LIV[369], while the Siberian and Far Eastern subtype TBEV strains form a distinct subgroup. The ALKV branch is located between that of POWV and LGTV, as previously reported for KFDV from the phylogenetic analysis of envelope sequences (7).

TABLE 2

Comparative Analysis of the Polyprotein of ALKV with Other Mammalian Tick-Borne Flaviviruses (A, Tick-Borne Flavivirus Cleavage Sites; B, Putative Processing of Viral Polyproteins. Sizes of the Proteins Are Given (AA))

A

Cleavage sites (protease)	VirC/CTHD (VSP)	AnchC/prM (HS)	Pr/M (Furin)	M/E (HS)	E/NS1 (HS)	NS1/NS2A (unknown)
TBEV^{SOF}	RGKRR/SAVDW	GVTLA/ATVRK	SRTRR/SVLIP	APVYA/SRCTH	LGVGA/DVGCA	MVVAD/NGELL
TBEV^{VAS}	RGKKR/STTDW	GVTLA/ATVRK	SRTRR/SVLIP	APVYA/SRCTH	LGVGA/DVGCA	MVVAD/NGELL
TBEV^{HYPR}	RGKRR/SATDW	GMTIA/ATVRK	SRTRR/SVLIP	APVYA/SRCTH	LGVGA/DVGCA	MVVAD/NGELL
TBEV^{NEU}	RGKRR/SATDW	GMTLA/ATVRK	SRTRR/SVLIP	APVYA/SRCTH	LGVGA/DVGCA	MVVAD/NGELL
LIV	RGKRR/SVTNW	GMTLA/ATVRK	SRTRR/SVLIP	APVYA/SRCTH	LGVGA/DVGCA	MVVAD/NGELL
LGTV	RGSRR/TTIDW	GMCLT/ATVRR	SRSRR/SVLIP	VPAYA/SRCTH	LGVGA/DVGCA	MVVAD/NGALL
POWV	RGRRR/SGVDW	TMAMA/TTIHR	SRGKR/SVVIP	GPVYA/TRCTH	MGVGA/DYGCA	MVVAD/NGALL
ALKV	RGKRR/STTGL	TLVIS/ATIRR	GRSRR/SVSIP	APTYA/TRCTH	LGVGA/DMGCA	MVLAD/NGAML
KFDV	RGKRR/STTGS	TLVFS/ATVRR	GRNRR/SVSIP	APTYA/TRCTH	LGVGA/?????	ND
	**.:*/.....	...../.*..:	.*.:*/**.**	.*.**/.****	.****/*.***	**.**/**..*

Cleavage sites (protease)	NS2A/NS2B (VSP)	NS2B/NS3 (VSP)	NS3/NS4A (VSP)	NS4A/2K (VSP)	2K/NS4B (HS)	NS4B/NS5 (VSP)
TBEV^{SOF}	HGRRR/SFSEP	RTARR/SGLVF	ASGRR/SFGDV	AGKQR/SSDDN	GLVAA/NEMGF	SGSRR/GGSEG
TBEV^{VAS}	HRGRR/SFSEP	RSARR/SDLVF	ASGRR/SIGDV	AGKQR/SSDDN	GLVAA/NEMGF	SGSRR/GGAEG
TBEV^{HYPR}	HRGRR/SFSEP	RSSRR/SDLVF	ASGRR/SFGDV	AGKQR/SSDDN	GLVAA/NEMGF	SGGRR/GGSEG
TBEV^{NEU}	HRGRR/SFSEP	RSSRR/SDLVF	ASGRR/SFGDV	AGKQR/SSDDN	GLVAA/NEMGF	SGGRR/GGSEG
LIV	HRGRR/SFSEP	RSSRR/SDLVY	ASGRR/SFGDV	TGKQR/SSDDN	GLVAA/NEMGF	SGGRR/GGSDG
LGTV	SRGRR/SFNEP	GSPRR/TDLVF	ASGRR/SVGDV	TGKQR/SSDDN	GMVAA/NEMGL	TGTRR/GGSEG
POWV	GRGRR/SLSEP	SSTRR/TDLVF	ASGRR/SAVDI	PGKQR/SGEDN	GLVAA/NELGY	QGARR/GGAEG
ALKV	RRNRR/SFSEP	SSGRR/SELVF	ASGRR/SVGDV	PGKQR/SSDDN	GLVTA/NEMGM	TGTRR/GGADG
KFDV	ND	ND	ND	ND	ND	ND
	...**/*..**	...**/..**.	*****/*..*.	.****/*..**	*.*.*/**.*.	.*.**/**..*

B

	TBEV^{SOF}	TBEV^{VAS}	TBEV^{HYPR}	TBEV^{NEU}	LIV	LGTV	POWV	ALKV	KFDV
VirC	96	96	96	96	96	96	94	97	97
CTHD	20	20	20	20	20	20	20	20	20
PrM	89	89	89	89	89	89	89	89	89
M	75	75	75	75	75	75	75	75	75
E	496	496	496	496	496	496	497	496	496
NS1	353	353	353	353	353	353	354	354	ND
NS2A	229	229	229	229	229	229	229	229	ND
NS2B	131	131	131	131	131	131	131	131	ND
NS3	621	621	621	621	621	621	621	621	ND
NS4A	126	126	126	126	126	126	126	126	ND
2K	23	23	23	23	23	23	23	23	ND
NS4B	252	252	252	252	252	252	252	252	ND
NS5	903	903	903	903	903	903	903	903	ND
TOTAL	3414	3414	3414	3414	3414	3414	3414	3416	

Note. VirC, mature virion C protein; CTHD, C-terminal hydrophobic domain; AnchC, anchored C protein (mature virion C protein + CTHD); prM, membrane precursor; E, envelope; NS, non structural protein; VSP, viral serine protease; HS, host signalase; ????? indicates that the sequence is not determined yet.

man infection by KFDV generally results from contact with infected ticks or monkeys living in the Kyasanur forest. In contrast, preliminary investigations suggest that human infections by ALKV result from contact with infected sheep or with ticks feeding on sheep, encountered in a semidesert environment. The close genetic relationship observed between the two viruses raises the question of the geographical origin of their common ancestor.

ALKV is the first tick-borne flavivirus causing hemorrhagic fever in humans for which the complete genome sequence has been determined. Identification of the specific viral determinants of hemorrhagic fever will require genetic characterization of the other tick-borne and mosquito-borne flaviviruses responsible for the same kind of clinical manifestations, e.g., KFDV, OHFV, yellow fever, and the dengue viruses.

ACKNOWLEDGMENTS

The "Unité des Virus Emergents" is an associated research unit of the Institut de Recherche pour le Développement (IRD). This study was partly funded by IRD.

REFERENCES

1. Zaki, A. M. (1997) Isolation of a flavivirus related to the tick-borne encephalitis complex from human cases in Saudi Arabia. *Trans. R. Soc. Trop. Med. Hyg.* **91,** 179–181.
2. Work, T. H., and Trapido, H. (1957) Kyasanur Forest disease, a new virus disease in India. *Indian J. Med. Sci.* **11,** 341–345.
3. Monath, T. P, and Heinz, F. X. (1996) Flaviviruses. *In* Fields Virology (Fields, B. N., Knipe, D. M., Howley, P. M., Chanock, R. M., Melnick, J. L., Monath, T. P., Roizman, B., and Straus, S. E., Eds.), 3rd ed., pp. 961–1034, Lippincott–Raven Publishers, Philadelphia, PA.
4. Thompson, J. D., Higgins, D. G., and 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.
5. Kumar S., Tamura, K., and Nei, M. (1993) MEGA: Molecular Evolutionary Genetics Analysis, 1.02 ed., Pennsylvania State University, University Park.
6. Chambers, T. J., Hahn, C. S., Galler, R., and Rice, C. M. (1990) Flavivirus genome organization, expression and replication. *Annu. Rev. Microbiol.* **44,** 649–688.
7. Venugopal, K., Gritsun, T., Lashkevitch, V. A., and Gould, E. A. (1994). Analysis of the structural protein gene sequence shows Kyasanur Forest disease virus as a distinct member in the tick-borne encephalitis virus serocomplex. *J. Gen. Virol.* **75,** 227–232.
8. Shiu, S. Y. W., Ayres, M. D., and Gould, E. A. (1991) Genomic sequence of the structural proteins of louping ill virus: Comparative analysis with tick-borne encephalitis virus. *Virology* **180,** 411–415.
9. Venugopal, K., Buckley, A., Reid, H. W., and Gould, E. A. (1992) Nucleotide sequence of the envelope glycoprotein of Negishi virus shows close homology to louping ill virus. *Virology* **190,** 515–521.
10. Shiu, S. Y. W., Jiang, W. R., Porterfield, J. S., and Gould, E. A. (1992) Envelope protein sequence of dengue virus isolates TH-36 and TH-Sman, and identification of type-specific genetic markers for dengue and tick-borne flaviviruses. *J. Gen. Virol.* **73,** 207–212.
11. Heinz, F. X., Collett, M. S., Purcell, R. H., Gould, E. A., Howard, C. R., Houghton, M., Moormann, R. J. M., Rice, C. M., and Thiel, H.-J. (2000) *Flaviviridae. In* Virus Taxonomy. Seventh Report of the International Committee for the Taxonomy of Viruses (Van Regenmortel, M. H. V., Fauquet, C. M., Bishop, D. H. L., Calisher, C. H., Carsten, E. B., Estes, M. K., Lemon, S. M., Maniloff, J., Mayo, M. A., McGeoch, D. J., Pringle, C. R., and Wickner, R. B., Eds.), Academic Press, San Diego.
12. Telford, S. R., III, Armstrong, P. M., Katavolos, P., Foppa, I., Olmeda Garcia, A. S., Wilson, M. L., and Spielman, A. (1997) A new tick-borne encephalitis-like virus infecting New England deer tick, *Ixodes dammini. Emerg. Infect. Dis.* **3,** 165–170.