Journal of General Virology (1998), **79**, 2481–2489. Printed in Great Britain

Comparative sequence analysis of American, European and Asian isolates of viruses in the genus *Coltivirus*

Houssam Attoui, Rémi N. Charrel, Frédérique Billoir, Jean-François Cantaloube, Philippe de Micco and Xavier de Lamballerie

Laboratoire de Virologie Moléculaire, Tropicale et Transfusionnelle, Faculté de Médecine de Marseille, Université de la Méditerranée, 27 Boulevard Jean Moulin, 13385 Marseille cedex 5, France

In this study, the basis for the classification of virus isolates grouped within the genus *Coltivirus*, family *Reoviridae*, is discussed. Sequences of dsRNA segments from American (segments 9–12), European (segment 12) and Asian (segments 7–12) isolates were characterized and polythetic criteria were defined for their taxonomic classification. These criteria (including sequence analysis) permitted the different species to be distinguished and classified into two groups. In both groups, subgroups were defined according to the degree of homology between the genomic sequences. American and European isolates are classified within group A, which includes subgroups A1 (Colorado tick fever virus species) and A2 (Eyach virus species). Asian isolates are classified in group B, which includes subgroups B1 (JKT-7075 virus species) and B2 (JKT-6423 virus species). The proteins encoded by the sequenced genomic segments were analysed. This allowed the identification of dsRNA binding domains in the proteins encoded by segment 8 of subgroup B1 isolates and segment 12 of subgroup B2 isolates. A conserved pattern of amino acids in segment 7 of group B isolates matched sequences found in the catalytic domains of protein kinases.

Introduction

Viruses which have genomes of 12 segments of dsRNA and have been isolated from vertebrates, ticks and mosquitoes are grouped within the genus *Coltivirus*, family *Reoviridae*. The type species of this genus is Colorado tick fever (CTF) virus (Murphy *et al.*, 1995). Other members of the genus are Eyach virus, isolated in Germany (Rehse-Küpper *et al.*, 1976), and two Eyach virus antigenic variants (AR577 and AR578), isolated in France (Chastel *et al.*, 1984). In previous studies, an antigenic relationship between these American and European isolates was established (Chastel *et al.*, 1984; Karabatsos *et al.*, 1987) and full-length nucleotide sequences of genomic segments 10, 11 and 12 of CTF virus were determined (Attoui *et al.*, 1997). Other viruses are also considered as tentative species of the genus, e.g. Indonesian isolates (Brown *et al.*, 1993) from mosquitoes and Chinese isolates (Xu *et al.*, 1990; Chen & Tao, 1996) from mosquitoes, ticks and vertebrates.

Author for correspondence: Xavier de Lamballerie.
Fax +33 491 18 95 98. e-mail virophdm@lac.gulliver.fr

The GenBank accession numbers of the sequences reported in this paper are AF007173–AF007186, AF019908, AF019909, AF052008–AF052035.

The taxonomic relationships between the viruses grouped within the genus *Coltivirus* have been only poorly explored and do not permit exact classification of these different isolates. In particular, there is an absence of information describing the genetic relationships between these viruses.

In this paper, we present some comparative sequence analyses of the genome segments of American, European and Asian isolates. In order to define the taxonomic status of these isolates, species parameters have been established in agreement with those used for other genera of the family *Reoviridae*. Accordingly, members of a single *Coltivirus* species may be identified according to:

1. Genetic criteria.

(i) Analysis of electropherotypes by agarose gel electrophoresis.

(ii) RNA cross-hybridization assays (Northern or dot blots, with probes made from viral RNA or cDNA).

(iii) RNA sequence analysis (including quantification of variation in genome segments and analysis of conserved RNA terminal sequences).

0001-5580

2. Serological comparisons by complement fixation and neutralization assays.
3. Their ability to exchange genetic material by genome segment reassortment during dual infection, thereby producing viable progeny reassortant viruses.

Application of these criteria to the analysis of American, European and Asian isolates has given some insight into their taxonomic status within the genus. We have produced a classification that is included in the forthcoming Report of the International Committee on Taxonomy of Viruses. We also present here preliminary data concerning the putative proteins encoded by some of the genomic segments of these isolates as deduced from the comparison to sequences deposited in databases.

Methods

■ **Virus strains.** Of the American isolates, Florio strain (N-7180) of CTF virus was purchased from the ATCC and two other CTF virus isolates (R-1575 and 69V28) and a related virus (S6-14-03) were provided by N. Karabatsos (Centers for Disease Control & Prevention, Fort Collins, CO, USA).

European isolates (Eyach and AR578 viruses) were provided by N. Karabatsos and C. Chastel (University of Brest, Brest, France), respectively.

Indonesian isolates (JKT-6423, JKT-6969, JKT-7043 and JKT-7075) and the Chinese isolate (Banna virus) were provided by R. B. Tesh and R. E. Shope (both at University of Texas Medical Branch, Galveston, TX, USA).

■ **Cell lines.** BHK-21 cells were cultured as monolayers or suspensions as previously described (Attoui *et al.*, 1997). C6/36 mosquito cells were cultured as monolayers in Leibovitz's L-15 medium supplemented with 5% FBS, 2% tryptose phosphate broth, penicillin G (100 IU/ml) and streptomycin (100 μg/ml).

■ **Virus propagation.** The Florio strain of CTF virus was propagated in suspensions of BHK-21 cells as described elsewhere (Attoui *et al.*, 1997). S6-14-03 virus was cultured in monolayers of BHK-21 cells. The two CTF strains (R-1575 and 69V28), Eyach and AR578 viruses were propagated by intracerebral inoculations of suckling mice and infected brains were recovered at the seventh day post-infection (p.i.). Indonesian isolates JKT-6423, JKT-6969, JKT-7043 and JKT-7075, as well as Banna virus, were propagated in C6/36 mosquito cells. Infected cells were recovered at the fourth day p.i.

■ **Isolation and purification of nucleic acids.** Viral dsRNA was extracted from infected cell cultures or suckling mouse brains using a guanidinium isothiocyanate-derived procedure (RNA NOW; Biogentex). In the case of cell cultures, viral dsRNA was further purified by precipitation of ssRNAs in 2 M LiCl and electrophoresis on a 10% acrylamide gel as previously reported (Attoui *et al.*, 1997). RNA was extracted from the acrylamide matrix by shaking overnight in TNE buffer (10 mM Tris, 100 mM NaCl, 1 mM EDTA, pH 8·0), precipitated from the aqueous phase using 10 vol. butanol and purified using the RNaid kit (BIO 101).

■ **Cloning of genomic segments 7–12 of JKT isolates and Banna virus.** Full-length genomic segments 7–12 of the JKT isolates and Banna virus were cloned and sequenced using a modification of the Lambden method (Lambden *et al.*, 1992; Attoui *et al.*, 1997). Briefly, a 3′ amino-blocked oligodeoxyribonucleotide (PO_4-5′ CCACGTGCCAGA-TGCTCTGGA 3′-amine; 2·5 mM) was ligated to both of the 3′ OH ends of the dsRNA (1 ng) using 10 U T4 RNA ligase. The tailed dsRNA product was purified using the RNaid kit and heat denatured (99 °C, 1 min in 15% DMSO). cDNA was synthesized in the presence of the complementary oligodeoxyribonucleotide primer (5′ GGTGCACGGT-CTACGAGACCT 3′; 1 mM) using 200 U MMuLV Superscript reverse transcriptase (Gibco BRL). Full-length cDNA was amplified by PCR using 2·5 U *Taq* polymerase and 1 μM primer (5′ GGTGCACGGTCTACG-AGACCT 3′). Thermal cycling parameters were as follows: one cycle of denaturation (90 °C, 10 min); and 35 cycles of denaturation (94 °C, 50 s), annealing (65 °C, 50 s) and extension (72 °C, 2 min). The cycling program ended with an extension step at 72 °C for 10 min. PCR products were gel-purified and ligated into a pGEM-T vector (Promega) and recombinant plasmids were used to transform competent *Escherichia coli* XL-Blue cells. Both strands of the cloned PCR products were sequenced using M13 universal primers and the Rhodamine D DNA sequencing kit, and analysed on the ABI prism 377 sequencer (Perkin Elmer).

■ **RT–PCR amplification of genomic segments.** The extracted viral dsRNAs from CTF strains R-1575 and 69V28, Eyach, AR578 and S6-14-03 viruses were resuspended in diethylpyrocarbonate-treated water. Reverse transcription was carried out as previously described in the presence of random hexanucleotides and the MMuLV Superscript

Table 1. Primers used in amplification of CTF virus genomic segments

Primer name	Sequence (5′ → 3′)	Segment	Map position	Orientation
M5720PS	CACTTGTGGCAGATCGGGGAGA	9 (M5)	*	Sense
M5720PA	CACCGGCTTACTACCTCATTTTT	9 (M5)	*	Antisense
M6TGAS	TGAAGATCTGGCGTGGCGAGAG	10 (M6)	970–991	Sense
M6TGAA	GAACACACAGATGTATTCAGACAC	10 (M6)	1165–1188	Antisense
S1S302	GGGAAGACGGACAATGTTCGTG	11 (S1)	367–388	Sense
S1A302	CCATTCACCAACCTCGTCCCG	11 (S1)	646–666	Antisense
S2259S	CTACTCGGATTTGAGTTAAGCCC	12 (S2)	224–246	Sense
COLREV	CCGGGAGAATGATGCTAGG	12 (S2)	640–658	Antisense

* Primers were designed from a partial sequence of genomic segment 9.

reverse transcriptase (Attoui *et al.*, 1998). PCR primers were designed from the sequences genomic segments 9–12 of the Florio strain (N-7180) of CTF virus with the help of the Oligo software program (National Biosciences). Primer sequences are shown in Table 1. The PCR protocol using segment-specific primers is described elsewhere (Attoui *et al.*, 1998). Briefly, the reaction was carried out in a volume of 100 µl and the amplification mixture included: 10 mM Tris–HCl (pH 8·8), 50 mM KCl, 1·5 mM $MgCl_2$, 0·1% Triton X-100, 0·2 mM of each dNTP, 20 µl cDNA solution, 1 µM of each primer and 2·5 U *Taq* DNA polymerase. The thermal cycling parameters were as described above except that annealing was carried out at 55 °C. Amplification products were cloned and sequenced as described above.

■ **Sequence analysis.** Open reading frames and their encoded putative proteins were identified using the DNA Strider 1.1 program (Institut de Recherche Fondamentale, CEA, France). Sequence alignments were generated by the Clustal W software program (Metrowerks). Phylogenetic analyses were performed with the software program MEGA (Kumar *et al.*, 1993) using the *p*-distance determination algorithm; sequence relatedness values were reported as percentage identity. Comparison of our sequences with those available from nucleic acid and protein sequence databases were performed using GENSTREAM's FASTA and BLAST software programs (EERIE, Nîmes, France) and the protein family domain (Pfam) program (Sanger Centre, UK).

Results

Analysis of nucleic acid sequences from American and European isolates

A set of primers (S2259S/COLREV) was designed from the previously determined nucleotide sequence of segment 12 of the Florio strain that permitted amplification of genomic segment 12 from all American and European strains (the Asian isolates could not be amplified using this set of primers). The generated amplicons were 437 nt for the American isolates and 440 nt for the European isolates. The length difference was due to an additional non-coding 'GAG' trinucleotide at position 585 of the corresponding CTF virus sequence. The sequences of the amplicons were determined (GenBank accession numbers: CTF viruses, AF007173–AF007175; Eyach viruses, AF007185 and AF007186).

Primers based on the sequences of segments 9–11 of the Florio strain of CTF virus allowed amplification of viral RNAs from all American isolates, but not of RNAs from the European and Asian isolates. The generated PCR products were 584 nt (segment 9), 221 nt (segment 10) and 302 nt (segment 11) and their sequences were determined (GenBank accession numbers: segment 9, AF007176–AF007178; segment 10, AF007182–AF007184; and segment 11, AF007179–AF007181).

Segment 12 of both American and European viruses could be amplified by the same set of primers. Accordingly these isolates were placed in the same group (group A). A first subgroup (designated A1) encompassing all the American isolates was identified. Segments 9–11 of all isolates from this subgroup could also be amplified by common sets of primers. The second subgroup (designated A2) included the European isolates. Segments 9–11 of isolates from this subgroup could not be amplified by the primers successfully used for the A1 subgroup.

When the nucleotide sequences of segment 12 of group A strains were compared, the percentage identity was 94·37–96·78% within subgroup A1, 92·29% within subgroup A2, and 65·68%–67·56% between subgroups A1 and A2. Detailed values are displayed in Table 2.

Within subgroup A1, similar degrees of nucleic acid identity were observed between genomic segments 9–12; these were 97·69–100% (segment 9), 96·84–98·95% (segment 10), and 90·6–94·74% (segment 11). No significant homology was found when the sequences of group A isolates were compared to those of the Asian isolates. The calculated G + C content for sequences of group A coltiviruses was 48·4–51·6%.

Analysis of the putative amino acid sequences of American and European isolates

A sequence of 170 aa from viral protein (VP) 9 of the Florio strain was compared to homologous sequences of strains R-1575, 69V28 and S6-14-03; sequence identity was 99·22–100%. Database searching revealed the presence of the motif $hhhhGx_4GKSx_nhhhhDD$ (where h indicates a bulky hydrophobic residue) that corresponded to the P-loop in the ATP binding site of archeal ATPases.

A 65 aa sequence from VP10 of the Florio strain was compared to homologous sequences of other American isolates. Amino acid identity was 96·84–98·95%. Comparison of full-length VP10 to database proteins revealed a partial match (42%) between a 270 aa hypothetical protein (MJ1287) of unknown function from the archeon *Methanococcus jannaschii* (aa 42–102) and VP10 (aa 40–100).

A 90 aa sequence from VP11 of the Florio strain was compared to the corresponding VP11 sequences of other strains of subgroup A1. The amino acid identity between these proteins was 95·45–98·86%.

The percentage identity of VP12 was 87·74–92·45% within subgroup A1 isolates and 87·74% within subgroup A2 isolates. It was 51·89–55·66% between American and European isolates (Table 2). VP11 and VP12 showed no significant identity with proteins in the database.

Analysis of the nucleic acid sequences from Asian isolates

Full-length segments 7–12 from the JKT-6423, JKT-6969, JKT-7043, JKT-7075 isolates and Banna virus were sequenced (GenBank accession numbers AF019908, AF019909 and AF052008–AF052035).

All Asian strains were assigned to a single group (group B). Sequence analysis permitted the identification of two subgroups, B1 and B2, corresponding to different virus species. The electropherotype of subgroup B1 viruses (JKT-7075) corresponds to a 6-5-1 pattern and that of subgroup B2 viruses

Table 2. Identity between genome sequences from *Coltivirus* isolates

Nucleic acid identities (%) are given and amino acid identities (%) are shown in parentheses.

(*a*) Values calculated for segments 7–12 of group B viruses.

Total sequence	Banna	JKT-6969	JKT-7043	JKT-7075
JKT-6423	87·25 (91·42)	84·39 (81·90)	84·22 (81·08)	47·95 (21·15)
Banna		80·11 (77·61)	79·74 (76·74)	47·45 (20·81)
JKT-6969			98·82 (98·48)	49·16 (21·80)
JKT-7043				49·20 (21·34)

(*b*) Values calculated for segment 9 of group B viruses.

Segment 9	Banna	JKT-6969	JKT-7043	JKT-7075
JKT-6423	83·82 (89·75)	54·48 (41·79)	54·39 (41·79)	41·67 (17·09)
Banna		55·12 (41·07)	55·21 (41·07)	40·60 (17·45)
JKT-6969			99·71 (99·64)	46·57 (18·68)
JKT-7043				46·67 (18·68)

(*c*) Values calculated for segment 7 of group B viruses.

Segment 7	Banna	JKT-6969	JKT-7043	JKT-7075
JKT-6423	85·64 (91·50)	73·22 (72·88)	72·95 (71·57)	49·47 (28·20)
Banna		73·99 (72·22)	73·28 (70·92)	48·93 (27·54)
JKT-6969			98·10 (97·39)	51·22 (28·85)
JKT-7043				51·39 (27·87)

(*d*) Values calculated for segment 12 of group A viruses.

Segment 12	69V28	R-1575	S6-14-03	Eyach	AR578
N-7180	95·98 (92·45)	96·78 (92·45)	94·37 (87·74)	66·22 (51·89)	66·76 (53·77)
69V28		95·44 (91·51)	95·17 (89·62)	66·76 (55·66)	67·56 (55·66)
R-1575			95·44 (92·45)	65·95 (54·72)	66·49 (55·66)
S6-14-03				65·68 (53·77)	65·95 (54·72)
Eyach					92·29 (87·74)

(isolates JKT-6423, JKT-6969, JKT-7043 and Banna virus) to a 6-6 pattern. The electrophoretic profiles are displayed in Fig. 1 and the lengths of segments 7–12 are reported in Table 3.

Two genotypes could be distinguished in subgroup B2 according to the length of segments: genotype 2a, including JKT-6423 and Banna virus; and genotype 2b, including JKT-6969 and JKT-7043 (Tables 3 and 4).

Within subgroup B2, the identity between homologous segments 8, 10, 11 and 12 was 83–98%. By contrast, identity in segment 7 was 72–98% and that in segment 9 was 54–99%. The comparison of the segments of subgroup B1 to corresponding segments of subgroup B2 showed that the highest degree of identity ranged between 46·67 (segment 9) and 52·16% (segment 12). The G+C content calculated from the sequenced segments of group B isolates was 37–39%.

Analysis of the 5′ and 3′ non-coding regions of Asian isolates

Analysis of the non-coding regions (NCR) of the cloned segments allowed identification of conserved motifs located at the termini. In all positive strands of the genome segments from the subgroup B1 viruses, the motifs 5′ GUAU$^{A}/_{U}{}^{A}$/

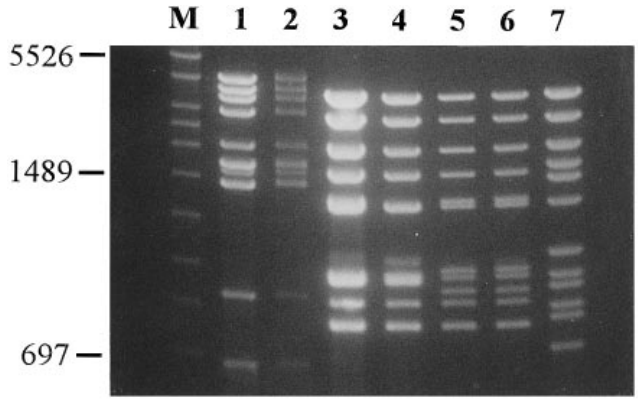

Fig. 1. Electrophoretic profiles of the genomes of cultivable coltiviruses. Lane M: size marker (bp). Lanes 1–7: genomes of CTF, S6-14-03, JKT-6423, Banna, JKT-6969, JKT-7043 and JKT-7075 viruses, respectively.

$_{U}AA^{A}/_{U}{}^{A}/_{U}U$ 3′ and 5′ $^{A}/_{G}C^{C}/_{U}GAC$ 3′ were found in the 5′ NCR and the 3′ NCR, respectively. In all positive strands of segments from subgroup B2 viruses, the motif 5′ GUAGAA $^{A}/_{U}{}^{A}/_{U}{}^{A}/_{U}U$ 3′ was identified in the 5′ NCR and the motif 5′ $A^{A}/_{C}{}^{C}/_{U}GAC$ 3′ was observed in the 3′ NCR. Segments of all group B isolates have in common the first two nucleotides in the 5′ NCRs (GU) and the last two nucleotides in the 3′ NCRs (AC). These dinucleotides are inverted complements.

Segment-specific inverted terminal repeats (ITRS), involving the conserved motifs, were detected in all segments of the different isolates. These ITRS could interact by base pairing, thus forming secondary structures. Such possible secondary structures were generated with the help of the MacStan software program (Gast, 1994) and are displayed in Fig. 2 for segments 7–12 of isolate JKT-6423.

Table 3. Comparison of segments 7–12 of group B coltiviruses: lengths of nucleotide and protein sequences

Lengths of nucleotide sequences (bp) are shown. Values given in parentheses indicate the number of amino acid residues in the protein encoded by the segment.

Gene	JKT-6423	Banna	JKT-6969	JKT-7043	JKT-7075
7	1136 (307)	1137 (307)	1136 (307)	1136 (307)	1259 (347)
8	1119 (302)	1119 (302)	1119 (302)	1119 (302)	1140 (302)
9	1101 (283)	1100 (283)	1141 (280)	1141 (280)	1054 (303)
10	977 (249)	978 (249)	977 (241)	977 (249)	946 (260)
11	867 (180)	867 (180)	867 (178)	867 (178)	894 (263)
12	862 (207)	861 (207)	862 (207)	862 (207)	756 (190)

Table 4. List of species in the genus *Coltivirus*

Species in the genus (vector species: host)	Sequence accession number	Abbreviation
***Coltivirus* group A**		
Colorado tick fever virus species, subgroup A1 (*Ixodidae* ticks: rodents, humans); over 22 isolates reported; two serotypes (CTF virus and S6-14-03)	Segment 9, AF007172; segment 10, AF000720; segment 11, U72694; segment 12, U5322	CTFV
Eyach virus species, subgroup A2 (*Ixodidae* ticks: possibly humans); includes Eyach virus, AR577, AR578	Segment 12, AF007185	EYAV
***Coltivirus* group B**		
JKT-7075 species, subgroup B1 (*Culex* mosquitoes: no known host); includes JKT-7075	Segment 7, AF052023; segment 8, AF052022; segment 9, AF052021; segment 10, AF052020; segment 11, AF052019; segment 12, AF019909	JKT-7075
JKT-6423 species*, subgroup B2 (*Culex* and *Anopheles* mosquitoes: humans, other vertebrates); includes JKT-6423, JKT-6969, JKT-7043, Banna virus	Segment 7, AF052018; segment 8, AF052017; segment 9, AF052016; segment 10, AF052015; segment 11, AF052014; segment 12, AF019908	JKT-6423

* Chinese isolates HN59, HN131, HN191 and HN295 most probably belong to JKT-6423 species based on serological and electrophoretic profile analyses (Xu *et al.*, 1990; Chen & Tao, 1996).

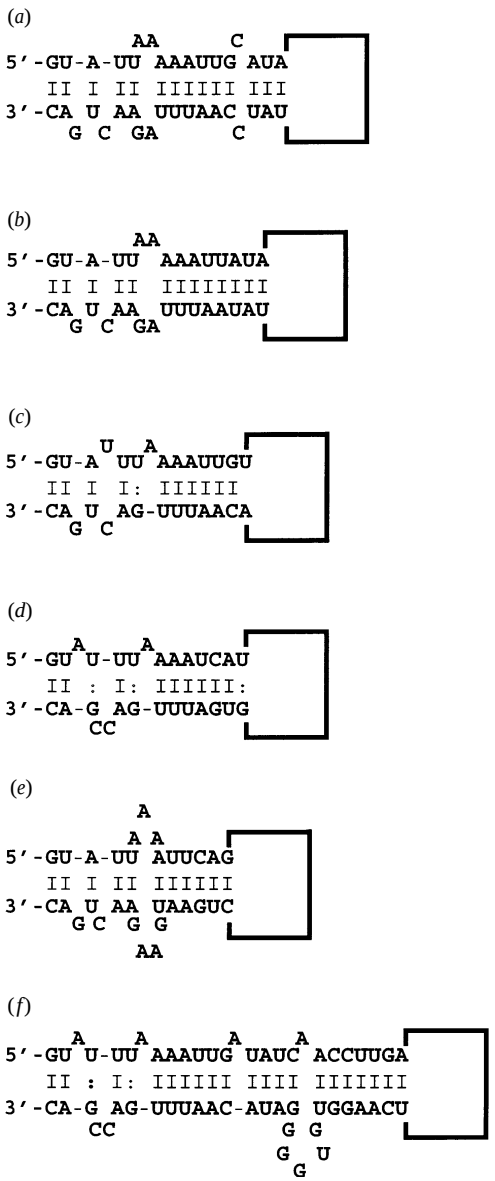

Fig. 2. Possible interaction between terminal sequences of the sense strands of segment 12 (*a*), 11 (*b*), 10 (*c*), 9 (*d*), 8 (*e*) and 7 (*f*) of JKT-6423 virus. The structures were generated with the help of the MacStan software program. 'I' denotes canonical base pairing while ':' denotes non-canonical base pairing.

Analysis of the putative amino acid sequences of Asian isolates

The putative VP encoded by segments 7–12 of group B isolates were determined. As with the nucleic acid sequences, analysis of the putative proteins allowed two different species to be distinguished: species JKT-7075 and species JKT-6423. The lengths of the encoded proteins are reported in Table 3.

Amino acid sequence identity was 85–100% for VP8, VP10, VP11 and VP12 of subgroup B2 isolates. The VP7 sequence identities were 70·9–97·39%. VP9 was the most variable, with 41·79–99·64% identity. Comparison of the

```
7-7075    174-LAGLKALMLFHEHTNCIHGDCNPSNIMCDKMGNVKIVD-211
MAPK      257-LATVKALNHLKENLKIIHRDIKPSNILLDTNGNIKLCD-294
           LA +KAL    E+    IH D  PSNI+ D  GN+K+ D
7-Banna             156-HSSNPHLLHGDANPDNIMSDSNGYLKLVD-184
7-6423              156-HSSNQHLLHGDANPDNIMSDSEGYLKLVD-184
P34-PK              119-YSHGHRILHRDLKPQNLLIDRQGALKLAD-147
                       +S    +LH D  P N++ D +G LKL D
7-7043                 162-YLHGDVNPENLMSDPNGYLKLVD-184
7-6969                 162-YLHGDVNPENIMSDANGYLKLVD-184
Ser./Threo.PK          138-VAHRDIKPENILLDKNGNLKLAD-160
                          H D+ PENI+ D NG LKL D

P68-kinase                IHRDLKPSNIFLVDTKQVKIGD

Consensus                 ΔH-D--P-NΔ--D--G-ZKΔ-D
```

Fig. 3. Match between protein kinases and putative proteins encoded by segment 7 of different group B *Coltivirus* isolates. 7-7075, VP7 from JKT-7075; MAPK, human dual specificity mitogen-activated protein kinase (accession number P45985); 7-Banna, VP7 from Banna virus; 7-6423, VP7 from JKT-6423; P34-PK, cell division control protein-2 homologue (P34 protein kinase, accession number P34112); 7-7043, VP7 from JKT-7043; 7-6969, VP7 from JKT-6969; Ser./Threo.PK, yeast probable serine/threonine-protein kinase YBR274w (accession number P38147); P68-kinase, human double-stranded RNA-dependent protein kinase (accession number P19525). Shadowed sequences represent identical and similar (+) amino acids. Δ represents I or L; Z represents I, L or V; and '—' in the consensus sequence represents any amino acid.

homologous proteins of isolates of subgroups B1 and B2 showed that the highest degrees of identity ranged between 17 (VP9) and 28% (VP7). However, alignments of the different amino acid sequences from group B isolates revealed partially conserved amino acid patterns in VP7 (35–93% identity) and VP10 (39–100% identity).

Database searching showed that the partially conserved amino acid motif in the VP7 proteins corresponded to the pattern found in the catalytic domains of different protein kinases (ΔH-D--P-NΔ--D--G-ZKΔ-D, where Δ represents I or L; Z represents I, L or V; and '-' represents any amino acid), including the dsRNA-dependent protein kinase (Fig. 3).

Database searching using the BLASTP program showed that the sequences of JKT-7075 VP8 (aa 5–63) partially matched a sequence of ribonuclease III (Fig. 4). Similarly, a partial match was found between the sequence of VP12 (aa 3–63) from JKT-6423, JKT-6969, JKT-7043 and Banna virus, and that of the maternal effect protein staufen (Fig. 4). These sequences corresponded to dsRNA binding domains (DRBDs) as deduced from the analysis by the Pfam program.

VP8 of isolates other than JKT-7075, and VP9 and VP11 proteins did not show a significant matching with proteins from databases. Despite the conserved pattern in the VP10 protein, similarity with database proteins was not significant and homologies between putative VP of groups A and B isolates were insignificant.

Discussion

The family *Reoviridae* includes members that have polyphyletic origins (Koonin, 1992). Among viruses exhibiting a dsRNA 12 segment genome, those that infect plants are

```
8-7075      5-  KGMLQELCVKRGLELPVYEKLSKVGPDHAPTITVKLTANGIEVIEAASSRAQAEKLAAA-63
RNAseIII  159-KTALQEWAQARGLPPPRYETLGRDGPDHAPQFRIAVVLASGETEEAQAGSKRNAEQAAA-217
              K  LQE    RGL  P YE L + GPDHAP   + +     E  EA +   +  + AAA

12-B2       3-  PVSVVHSFARSQGLPLNFETVGCEGPSHDPRFVIECKLLDFQHQCT-D---SSKKRAIQKICVLI-63
MEPS      579-PISQVHEIGIKRNMTVHFKVLREEGPAHMKNFITACIVGSIVTEGEGNGKKVSKKRAAEKMLVEL-643
              P+S VH     + + ++F+ +  EGP+H   F+  C +     +   +    SKKRA +K+ V +
```

Fig. 4. Match of proteins of group B coltiviruses to proteins harbouring dsRNA binding domains. RNAse III, *Rhodobacter capsulatus* ribonuclease III (accession number 52698); MEPS, *Drosophila melanogaster* maternal effect protein staufen (accession number P25159); 8-7075, VP8 from JKT-7075; 12-B2, VP12 of subgroup B2 isolates. Shadowed sequences represent identical and similar (+) amino acids.

grouped within the genus *Phytoreovirus* and those isolated from vertebrates, ticks and mosquitoes are classified in the genus *Coltivirus*. The antigenic relationship between CTF (American) and Eyach (European) virus isolates has been analysed (Chastel *et al*., 1984; Karabatsos *et al*., 1987; Rehse-Küpper *et al*., 1976). Neutralization of both American and European isolates by antibodies to CTF virus (Karabatsos *et al*., 1987) suggests that these viruses are antigenically related. In our experience, European strains can be detected in cell cultures by an indirect immunofluorescence assay using anti-CTF virus antibodies, which is consistent with these data.

Primers designed from the sequences of segments 9–12 of the Florio strain (Attoui *et al*., 1997, 1998) of CTF virus permitted PCR amplification of these segments from the R-1575, 69V28 and S6-14-03 viruses. The amplified sequences had the same lengths as those of the Florio strain and exhibited high degrees of nucleic acid and deduced protein identities. The primers used for the amplification of segments 9–11 did not amplify the homologous segments from the European viruses. However, primers S2259S/COLREV, designed from the sequence of segment 12 of CTF virus, amplified segment 12 of Eyach virus and one of its variants (AR578). Identity to the homologous CTF virus sequences was 65·68–67·56% for nucleic acids and 51·99–55·66 % for amino acids.

These molecular findings are in agreement with previously reported serological conclusions which suggest a common phylogenetic origin of American and European strains, and corroborate the geographical clusterization. Accordingly, American and European viruses were assigned to a single group designated *Coltivirus* group A. The nucleic and amino acid identities, the genome electropherotype of Eyach virus and the geographical origin of the isolates suggest that group A viruses belong to two different species: the CTF virus species (subgroup A1) which includes the American strains (virus S6-14-03 being a distinct serotype); and the Eyach virus species (subgroup A2), which includes European viruses.

The molecular relationship between American coltiviruses and the Indonesian JKT viruses was previously analysed by Brown *et al*. (1993) using RNA–RNA hybridization which did not produce any cross-hybridization signal. Furthermore, genomes of strains JKT-6423, JKT-6969, JKT-7043 did not cross-hybridize with that of strain JKT-7075.

In this study, we cloned segments 7–12 of JKT-7075, JKT-6423, JKT-6969, JKT-7043 and Banna viruses and determined their complete nucleotide sequences. Sequence analysis did not reveal any significant homology to group A coltiviruses, which is in agreement with the findings of Brown *et al*. (1993). Moreover, the G + C content calculated from the sequences of segments 7–12 of Asian isolates was 37–39%, which is clearly lower than that of group A coltiviruses (48·4–51·6%).

Analysis of the sequences of Asian isolates showed that nucleic acids exhibit 47·75–98·82% overall identity and revealed the presence of common amino acid patterns in segments 7 and 10. Asian viruses could thus be assigned to a second group (*Coltivirus* group B). Further investigations revealed the existence of two species represented by isolates of subgroups B1 and B2, the latter including two distinct genotypes.

Subgroup B1 is represented by JKT-7075 virus. Its nucleic acid sequence identity to sequences of isolates belonging to subgroup B2 (JKT-6423, JKT-6969, JKT-7043 and Banna viruses) is about 49%. Comparison of the proteins of JKT-7075 to those of subgroup B2 viruses showed similarities only in VP7 and VP10.

For isolates of subgroup B2, the overall nucleic acid sequence identity in the conserved segments (8, 10, 11 and 12) is 83·3–98·84%; the sequence identity is 54·39–99·7% in the variable segments (7 and 9). Analysis of the nucleic and amino acid sequences of subgroup B2 viruses reveals the existence of two genotypes. The first one (genotype 2a) is represented by JKT-6423 and Banna viruses and the second one (genotype 2b) is represented by JKT-6969 and JKT-7043 viruses. In each genotype, the variable segments (7 and 9) are similar in terms of their nucleic and amino acid sequences.

Amino acid sequence analysis is in agreement with the nucleic acid sequence results. In each genotype, identity between homologous proteins is very high (over 85%). However, in the case of VP10 of genotype 2b, a difference of 8 aa in the carboxyl terminus was noticed between isolates JKT-6969 and JKT-7043. The VP10 of JKT-7043 is 249 aa (i.e. the same length as VP10 from isolates of genotype 2a), while that of JKT-6969 is 241 aa. However, the nucleic and amino acid sequence identity values are higher between segment 10 of JKT-6969 and JKT-7043 than between that of JKT-6969 and genotype 2a viruses.

The NCRs of segments 10–12 of CTF virus were previously analysed and conserved motifs were identified in their 5′ and 3′ NCRs. The conserved terminal motifs in the NCRs of the

full-length segments of JKT isolates are different from those of CTF virus. Such conserved motifs can be detected in many viruses possessing multi-segmented genomes and would probably play an important role in replication and sorting of RNA segments as well as in virus maturation (Anzola *et al.*, 1987; Xu *et al.*, 1989).

The first and last two nucleotides as well as the terminal repeats in the 5′ and 3′ NCRs of segments of group B isolates were found to be inverted complements. This suggests that the RNA transcript could be held in a circular form either by itself or via protein components (Anzola *et al.*, 1987; Theron & Nel, 1997). This finding has been previously reported for viruses such as CTF virus (Attoui *et al.*, 1997), infectious bursal disease virus (Mundt & Müller, 1995), wound tumour virus (Anzola *et al.*, 1987) and influenza virus (Hsu *et al.*, 1987), where a 'panhandle' structure was described that could possibly function as a guiding site for the virus-specific RNA-dependent RNA polymerase.

Altogether, these data suggest a different phylogenetic origin for group A and B isolates. Group A isolates share a number of characteristics that permit their assignment to this group: (i) the degree of identity between nucleic acid sequences of segment 12 (65·68–96·78%); (ii) common amino acid patterns; (iii) common vectors (*Ixodidae* ticks; *Ixodes* and *Dermacentor* species) and hosts (vertebrates); and (iv) similar G + C content (48·4–51·6%). All subgroup A1 isolates had the same geographical origin and similar genomic electrophoretic pattern.

Group B isolates also share a number of common characteristics: (i) the overall identity between nucleotide sequences is 47·75–98·82% (51·42–98·96% for segment 12); (ii) common amino acid patterns; (iii) common vectors (*Culex* and *Anopheles* mosquitoes); (iv) similar G + C content (37–39%); and (v) a common geographical origin (south-eastern Asia). Within subgroup B2, all isolates also have similar conserved terminal sequences and genomic electrophoretic profiles. The conserved terminal sequences in subgroups B1 and B2 are similar and identical terminal dinucleotides are common to all group B isolates.

The genetic variability of the American isolates of CTF virus was analysed by Bodkin & Knudson (1987) using RNA–RNA hybridization methodology. These authors found that genes 4 and 6 are variable and reported that this finding is the consequence of a genetic reassortment phenomenon. Accordingly, as stated in the Introduction, the ability of virus strains to exchange genetic material by genome segment reassortment during dual infection may also be used as a parameter for the definition of species in the genus *Coltivirus*.

The possible functions of the putative proteins of coltiviruses were investigated. Regarding group A viruses, partial matching with proteins from databases was only found in VP9 and VP10 of CTF virus. An amino acid pattern described as being the ATP binding site in archeal ATPases (Koonin, 1997) was found in VP9, while VP10 partially matched an amino acid sequence from a hypothetical protein of the archeon *Methanococcus jannaschii*.

In the group B isolates, DRBDs were detected in VP8 (subgroup B1) and VP12 (subgroup B2). Within the family *Reoviridae*, proteins exhibiting a dsRNA binding activity have been identified such as the sigma-3 protein in reovirus type 3 (Schiff *et al.*, 1988) and the NSP3 protein in porcine rotavirus group C (Langland *et al.*, 1994). Despite the fact that the DRBD of sigma-3 protein is unrelated to that of NSP3 DRBD (which conforms to those of the group B VP8 and VP12), similar functions have been attributed to these two proteins. They were found to stimulate translation of viral mRNAs by blocking activation of the dsRNA-dependent protein kinase (Langland *et al.*, 1994; Yue & Shatkin, 1997). It is therefore possible that the JKT proteins with DRBDs would have a similar function.

Finally, this study represents the first attempt to determine the taxonomic status of the existing virus isolates belonging to the genus *Coltivirus*, as well as a means for the assignment of viruses having dsRNA genomes of 12 segments that might be isolated in the future. The characterized sequences of isolates of the genus *Coltivirus* can also be utilized for the development of molecular procedures for detection of these viruses.

The authors wish to thank Dr R. B. Tesh and Dr R. E. Shope for kindly providing the Indonesian and Chinese virus strains.

References

Anzola, J. V., Xu, Z., Asamizu, T. & Nuss, D. L. (1987). Segment-specific inverted repeats found adjacent to conserved terminal sequences in wound tumour virus genome and defective interfering RNAs. *Proceedings of the National Academy of Sciences, USA* **84**, 8301–8305.

Attoui, H., Billoir, F., Bruey, J. M., De Micco, P. & de Lamballerie, X. (1998). Serological and molecular diagnosis of Colorado tick fever viral infections. *American Journal of Tropical Medicine and Hygiene* (in press).

Attoui, H., De Micco, P. & de Lamballerie, X. (1997). Complete nucleotide sequence of Colorado tick fever virus segments M6, S1 and S2. *Journal of General Virology* **78**, 2895–2899.

Bodkin, D. K. & Knudson, D. L. (1987). Genetic relatedness of Colorado tick fever virus isolates by RNA–RNA blot hybridization. *Journal of General Virology* **68**, 1199–1204.

Brown, S. E., Gorman, M., Tesh, R. B. & Knudson, D. L. (1993). Coltiviruses isolated from mosquitoes collected in Indonesia. *Virology* **196**, 363–367.

Chastel, C., Main, A. J., Couatarmanac'h, A., Le Lay, G., Knudson, D. L., Quillien, M. C. & Beaucournu, J. C. (1984). Isolation of Eyach virus (*Reoviridae*, Colorado tick fever group) from *Ixodes ricinus* and *I. ventalloi* ticks in France. *Archives of Virology* **82**, 161–171.

Chen, B. & Tao, S. (1996). Arbovirus survey in China in recent ten years. *Chinese Medical Journal* **109**, 13–15.

Gast, F. U. (1994). A Macintosh program for the versatile generation of random nucleic acid sequences and their structural analysis. *Computer Applications in the Biosciences* **10**, 49–51.

Hsu, M. T., Parvin, J. D., Gupta, S., Krystal, M. & Palese, P. (1987). Genomic RNA of influenza viruses are held in a circular conformation in virions and in infected cells by a terminal panhandle. *Proceedings of the National Academy of Sciences, USA* **84**, 8140–8144.

Karabatsos, N., Poland, J. D., Emmons, R. W., Mathews, J. H., Calisher, C. H. & Wolff, K. L. (1987). Antigenic variants of Colorado tick fever virus. *Journal of General Virology* **68**, 1463–1469.

Koonin, E. V. (1992). Evolution of double-stranded RNA viruses: a case for polyphyletic origin from different groups of positive-stranded RNA viruses. *Seminars in Virology* **3**, 327–339.

Koonin, E. V. (1997). Evidence for a family of archaeal ATPases. *Science* **275**, 1489–1490.

Kumar, S., Tamura, K. & Nei, M. (1993). Molecular Evolutionary Genetics Analysis (Pennsylvania State University), version 1.01.

Lambden, P., Cooke, S. J., Caul, E. O. & Clarke, I. N. (1992). Cloning of noncultivatable human rotavirus by single primer amplification. *Journal of Virology* **66**, 1817–1822.

Langland, J. O., Pettiford, S., Jiang, B. & Jacobs, B. L. (1994). Product of the porcine group C rotavirus NSP3 gene binds specifically to double-stranded RNA and inhibits activation of the interferon-induced protein kinase PKR. *Journal of Virology* **68**, 3821–3829.

Mundt, E. & Müller, H. (1995). Complete nucleotide sequences of 5′- and 3′-noncoding regions of both genome segments of different strains of infectious bursal disease virus. *Virology* **209**, 10–18.

Murphy, F. A., Fauquet, C. M., Bishop, D. H. L., Ghabrial, S. A., Jarvis, A. W., Martelli, G. P., Mayo, M. A. & Summers, M. D. (editors) (1995). Virus Taxonomy. Sixth Report of the International Committee on Taxonomy of Viruses. Vienna & New York: Springer–Verlag.

Rehse-Küpper, B., Casals, J., Rehse, E. & Ackermann, R. (1976). Eyach, an arthropod-borne virus related to Colorado tick fever virus in the Federal Republic of Germany. *Acta Virologica* **20**, 339–342.

Schiff, L. A., Nibert, M. L., Co, M. S., Brown, E. A. & Fields, B. N. (1988). Distinct binding site for zinc and double-stranded RNA in the reovirus outer capsid protein $\sigma 3$. *Molecular and Cellular Biology* **8**, 273–283.

Theron, J. & Nel, L. H. (1997). Stable protein–RNA interaction involves the terminal domains of bluetongue virus mRNA but not the terminally conserved sequences. *Virology* **229**, 134–142.

Xu, Z., Anzola, J. V., Nalin, C. M. & Nuss, D. L. (1989). The 3′-terminal sequence of wound tumor virus transcripts can influence conformational and functional properties associated with the 5′ terminus. *Virology* **170**, 511–522.

Xu, P., Wang, Y., Zuo, J. & Lin, J. (1990). New orbiviruses isolated from patients with unknown fever and encephalitis in Yunnan province. *Chinese Journal of Virology* **6**, 27–33.

Yue, Z. & Shatkin, A. J. (1997). Double-stranded RNA-dependent protein kinase (PKR) is regulated by reovirus structural proteins. *Virology* **234**, 364–371.

Received 24 March 1998; Accepted 30 June 1998