Journal of General Virology (2002), **83**, 1941–1951. Printed in Great Britain

Common evolutionary origin of aquareoviruses and orthoreoviruses revealed by genome characterization of Golden shiner reovirus, Grass carp reovirus, Striped bass reovirus and golden ide reovirus (genus *Aquareovirus*, family *Reoviridae*)

Houssam Attoui,[1] Qin Fang,[2] Fauziah Mohd Jaafar,[1] Jean-François Cantaloube,[1] Philippe Biagini,[1] Philippe de Micco[1] and Xavier de Lamballerie[1,3]

[1] Unité des Virus Emergents EA3292, Université de la Méditerranée, Faculté de Médecine de Marseille, EFS Alpes-Méditerranée, 13005 Marseille, France

[2] Hydrobiology Institute, Wuhan Institute of Virology, CAS, Wuchang, 430071, Wuhan, Hubei, China

[3] Maladies virales émergentes et systèmes d'information UR 034, Institut de Recherche pour le Développement, Marseille, France

Full-length and partial genome sequences of four members of the genus *Aquareovirus*, family *Reoviridae* (Golden shiner reovirus, Grass carp reovirus, Striped bass reovirus and golden ide reovirus) were characterized. Based on sequence comparison, the unclassified Grass carp reovirus was shown to be a member of the species *Aquareovirus C*. The status of golden ide reovirus, another unclassified aquareovirus, was also examined. Sequence analysis showed that it did not belong to the species *Aquareovirus A* or *C*, but assessment of its relationship to the species *Aquareovirus B*, *D*, *E* and *F* was hampered by the absence of genetic data from these species. In agreement with previous reports of ultrastructural resemblance between aquareoviruses and orthoreoviruses, genetic analysis revealed homology in the genes of the two groups. This homology concerned eight of the 11 segments of the aquareovirus genome (amino acid identity 17–42%), and similar genetic organization was observed in two other segments. The conserved terminal sequences in the genomes of members of the two groups were also similar. These data are undoubtedly an indication of the common evolutionary origin of these viruses. This clear genetic relatedness between members of distinct genera is unique within the family *Reoviridae*. Such a genetic relationship is usually observed between members of a single genus. However, the current taxonomic classification of aquareoviruses and orthoreoviruses in two different genera is supported by a number of characteristics, including their distinct G+C contents, unequal numbers of genome segments, absence of an antigenic relationship, different cytopathic effects and specific econiches.

Introduction

Viruses belonging to the family *Reoviridae* that infect aquatic animals are classified within the genus *Aquareovirus*. This genus was created by the International Committee on Taxonomy of Viruses (ICTV) in 1991 (Francki *et al.*, 1991). The type is Striped bass reovirus (SBRV, species *Aquareovirus A*), and five other species (*Aquareovirus B* to *F*) are recognized, along with a number of tentative species. The current classification has been established on the basis of three main criteria: RNA–RNA hybridization, electrophoretype analysis and antigenic properties (Mertens *et al.*, 2000). Sequence data are recognized criteria for aquareovirus classification but, until very recently, analysis of the molecular relationships among aquareoviruses has been hampered by the paucity of the available genetic information.

Author for correspondence: Houssam Attoui.
Fax +33 4 91 32 44 95. e-mail virophdm@lac.gulliver.fr

The accession numbers of sequences reported in this paper are AF450318–AF450324, AF403390–AF4034141 and AF418294–AF418304.

Aquareoviruses have been isolated from a wide variety of aquatic animals, including molluscs, finfish and crustaceans. In the past, they have been referred to as reovirus-like or rotavirus-like aquatic viruses. Like members of the genus *Rotavirus*, their genomes are composed of 11 segments of dsRNA. The genome is contained in a core surrounded by a double-layered icosahedral capsid that physically resembles capsids of mammalian orthoreoviruses (MRV), as demonstrated by cryoelectron microscopy (Shaw *et al.*, 1996).

Aquareoviruses grow in fish cell lines and produce large syncytia that represent the typical cytopathic effect of their replication. In fishes, their typical pathogenic effect is haemorrhage, which represents a serious threat to fish breeding. As an example, Grass carp reovirus (GCRV) provokes severe haemorrhage in fingerling and yearling grass carp, leading to 80% mortality (Fang *et al.*, 1989).

Previously, partial genome sequences have been deposited in databases for *Aquareovirus A*, *Aquareovirus B* and GCRV. In this paper, we report the full-length genome sequences of Golden shiner reovirus (GSRV, *Aquareovirus C*) and GCRV (*Aquareovirus C*). We also have characterized genome segments 2, 3, 4, 8 and 10 of SBRV (*Aquareovirus A*) and segments 2 and 5 of golden ide reovirus (GIRV, unclassified). These molecular data provide the opportunity for the first robust analysis of the genetic relationships between aquareoviruses A, B and C and unclassified viruses. In addition, the analysis of complete genome sequences from three different aquareoviruses allowed the reassessment of the evolutionary relationship between viruses belonging to the genera *Aquareovirus* and *Orthoreovirus*.

Methods

■ The viruses and their propagation

Golden shiner reovirus (GSRV). GSRV is a member of the species *Aquareovirus C*. It has been isolated in the USA from the freshwater golden shiner fish [*Notemigonus crysoleucas*, family *Cyprinidae* (the minnow family)] showing petechial haemorrhage in skin (Plumb *et al.*, 1979). The virus was purchased from the ATCC (reference VR-957). GSRV was propagated in fathead minnow (FHM) cells. Confluent monolayers of FHM were infected with 5 p.f.u. GSRV per cell in the presence of Eagle's minimum essential medium with Hanks' salts (EMEM) supplemented with 10% foetal bovine serum. The cells were incubated for 3 days at 30 °C.

Grass carp reovirus (GCRV). GCRV is now considered to belong to *Aquareovirus C*. Four isolates are now identified (Fang *et al.*, 1989) as GCRV-873 (prototype strain), GCRV-875, GCRV-876 and GCRV-991. The virus was isolated from the freshwater grass carp fish (*Ctenopharyngodon idellus*, family *Cyprinidae*) in the People's Republic of China (Chen & Jiang, 1984; Ke *et al.*, 1990) and provokes severe haemorrhage in fingerling and yearling grass carp (Fang *et al.*, 1989). GCRV was propagated in *Ctenopharyngodon idellus* kidney (CIK) cells. The cells were incubated for 3 days at 28 °C. Other experimental conditions were as above.

Golden ide reovirus (GIRV). GIRV is an unclassified aquareovirus. The virus was isolated in Germany (Neukirch *et al.*, 1999) from the freshwater golden ide fish (*Leuciscus idus melanotus*). It differs from other aquareoviruses in its distinct cytopathic effect in cell culture (absence of syncytia and formation of focal aggregates of round cells) and by its sensitivity to chloroform. This virus was provided by Dr M. Neukirch. GIRV was grown in FHM cells at 20 °C. Other experimental conditions were as above.

Striped bass reovirus (SBRV). SBRV is a member of the species *Aquareovirus A*. The virus was isolated in the USA from the salt-water striped bass fish (*Morone saxatilis*, family *Centrarchidae*) with haemorrhagic lesions of the skin (Subramanian *et al.*, 1994). This virus was provided by Dr S. K. Samal as a suspension grown in a Chinook salmon embryo cell line (CHSE-214). The virus was studied without further propagation.

■ **Preparation of virus dsRNA.** Clarified supernatants from virus-infected cell cultures were subjected to ultracentrifugation. The pelleted viruses were resuspended in 200 µl EMEM and RNA was extracted with RNA-Now reagent (Biogentex) as described previously (Attoui *et al.*, 2000). The dsRNA segments were separated on a 10% polyacrylamide gel, excised and purified using the RNaid kit (Bio-101).

■ **Cloning and sequence determination of genome segments of *Aquareovirus C*: GCRV and GSRV.** Segments 1, 2 and 3 of GCRV-873 were sequenced in a previous study (Fang *et al.*, 2000; accession nos AF260511–AF260513). Cloning of segments 4–11 of GCRV-873 and segments 1–11 of GSRV was achieved using the single-primer amplification technique (Attoui *et al.*, 2000), and this completed the genome characterization for both viruses.

■ **Partial sequence determination of genes 8 and 10 of GCRV: isolates GCRV-875, GCRV-876 and GCRV-991.** dsRNAs from GCRV-875, GCRV-876 and GCRV-991 were heat-denatured at 99 °C for 1 min in the presence of 15% DMSO and reverse-transcribed under standard conditions using Superscript-II (Invitrogen) and a random hexaprimer mixture.

Primers were designed from the sequence of segments 8 and 10 of GCRV-873 and are reported in Table 1. PCR was performed under standard conditions using *Taq* polymerase (Invitrogen) at an annealing temperature of 55 °C. The PCR products were sequenced directly using the corresponding PCR primers, a D-Rhodamine DNA sequencing kit and an ABI Prism-377 sequence analyser (Perkin Elmer).

■ **Sequence determination of GIRV (unclassified) and SBRV (*Aquareovirus A*).** cDNAs of the genomes of SBRV and GIRV were synthesized by using either random hexaprimers or the single-primer amplification method. Segments 2 of SBRV and GIRV were amplified by PCR using primers designed from the most conserved nucleotide sequence alignments of the polymerases of orthoreoviruses and/or aquareoviruses. These primers are shown in Table 1. Segments 3, 4, 8 and 10 of SBRV and segment 5 of GIRV were gel-purified and cloned by using the single-primer amplification method.

■ **Sequences retrieved from databases.** The nearly complete genome of Chum salmon reovirus (CSRV, *Aquareovirus A*) was retrieved from GenBank. Unpublished sequences of segments 1–11 of CSRV were deposited by S. Rao, G. R. Carner, W. Chen and J. R. Winton under accession numbers AF418294–AF418304. A partial sequence of genome segment 10 of Coho salmon reovirus CSR (*Aquareovirus B*) was retrieved from GenBank (accession number CSU90430).

■ **Sequence analysis methods.** Sequence alignments were performed using CLUSTAL W (Thompson *et al.*, 1994) and the local-BLAST program implemented in the DNATools package (version 5.01.661, S. W. Rasmussen). Phylogenetic analysis was carried out by the neighbour-joining method (Saitou & Nei, 1987) and the p-distance determination algorithm in the MEGA program (Kumar *et al.*, 1993). Sequence relatedness

Table 1. Primers used for the amplification of segments 2 of GIRV and SBRV and segments 8 and 10 of GCRV isolates

Primer	Sequence (5′ → 3′)	Segment (origin)	Map position*	Orientation
GSVseg2S	CCAATACCTGTTAACTGATCTGATTAA	2 (GSRV)	689–719	Sense
SBRseg2rev	TCCTTGACGATGAATTTGAGCGTGC	2 (SBRV)	Partial	Antisense
ReAqSeg2S1	CAAGCSATYATGAGRTCTCAATACGT	2 (aquareovirus + orthoreovirus)	1372–1397	Sense
GIRVSeg2R	GGATGGTCATATCCGAGAGGACACG	2 (GIRV)	Partial	Antisense
ReAqSeg2S2	GCCACCTCYACYGAGCAYACYGCTAATAA	2 (aquareovirus + orthoreovirus)	2086–2114	Sense
AquaSeg2R	CATTCCTTGGTCWCCGGGGGGCATGTA	2 (aquareovirus)	2730–2704	Antisense
Aqua_C8S	GTTATTTTGTGATGGCACAGCGTCA	8 (GCRV)	1–25	Sense
Aqua_C8R	GATGAAAGTCGTGAGGCAGCGGAGACG	8 (GCRV)	1296–1270	Antisense
Aqua_C10S	GCCCCCGATCATCACCACGATG	10 (GCRV)	12–33	Sense
Aqua_C10R	GGGTGGGTAGGCCGGTGCTTA	10 (GCRV)	879–859	Antisense

* With respect to the GSRV sequence.

is reported as percentage identity. Tree drawing was performed with the help of the program TreeView (Page, 1996). Comparisons of GSRV and GCRV sequence data with those available from nucleotide and protein databases were performed by using the NCBI program gapped BLAST (http://www3.ncbi.nlm.gov/blast).

Hydropathy profiles were analysed by using amino acid sequence hydropathy values determined by the method of Kyte & Doolittle (1982) implemented in Microsoft Excel. Sequences aligned with CLUSTAL W were exported while all alignment-generated gaps were maintained (gap hydropathy value = 0). This permitted the comparison of positional hydropathy profiles for amino acid sequences of unequal lengths.

Results

Sequence determination of *Aquareovirus C* viruses GCRV and GSRV

Sequences of genome segments 1–11 of GSRV were deposited in GenBank under accession numbers AF403398–AF403408. Sequences of segments 4–11 of GCRV were deposited under accession numbers AF403390–AF403397. The total RNA segment size, the longest ORF, the sizes of the 5′ and 3′ non-coding regions (NCR) and the size of the putative protein encoded were identified for each segment and are shown in Table 2.

Partial sequence determination of genes 8 and 10 of GCRV isolates GCRV-875, GCRV-876 and GCRV-991

Partial sequences of segments 8 (1296 bp) and 10 (868 bp) of isolates GCRV-875, GCRV-876 and GCRV-991 were determined. The GenBank accession numbers are AF403412–AF403414 (segments 8) and AF403409–AF403411 (segments 10).

Sequence determination of GIRV (unclassified) and SBRV (*Aquareovirus A*)

Primers ReAqSeg2S2/AquaSeg2R amplified a 645 bp sequence of segment 2 of SBRV. Sequencing of this product allowed the design of an SBRV sequence-specific reverse primer designated SBRseg2rev. This primer, together with primer GSVseg2S, allowed PCR amplification of an overlapping sequence of 1491 bp. The final sequence obtained from SBRV segment 2 was 2026 bp long (accession no. AF450318).

Primers ReAqSeg2S2/AquaSeg2R also amplified a 645 bp sequence of segment 2 of GIRV. Sequencing of this product allowed the design of a GIRV sequence-specific reverse primer designated GIRVseg2rev. This primer, together with primer ReAqSeg2S1, allowed PCR amplification of an overlapping sequence of 836 bp. The final sequence of GIRV segment 2 was 1360 bp long (accession no. AF450323).

Segments 3, 4, 8 and 10 of SBRV and segment 5 of GIRV were separated and cloned by single-primer amplification. PCR amplification of cDNA from segment 3 using primer B (Attoui *et al.*, 2000) has generated an amplicon that corresponded to a specific sequence of 1186 bp (accession no. AF450319). PCR from segment 4 resulted in a 1116 bp sequence (accession no. AF450320) located at the 3′ terminus. Segments 8 and 10 were cloned as full-length sequences and were found to be respectively 1317 bp long (accession no. AF450321) and 987 bp long (accession no. AF450322). Segment 5 of GIRV was cloned as a full-length product and was found to be 2238 bp long (accession no. AF450324).

Sequence analysis of *Aquareovirus* C viruses GCRV and GSRV

Comparison of the genome electrophoretypes. The genomes of GCRV and GSRV showed identical electrophoretypes on a 1·2% agarose gel. Slight variations were detected in the PAGE profiles. This could be explained by sequence variations, as observed with other viruses belonging to a single species within the family *Reoviridae* (Mertens *et al.*, 2000).

Comparison of GSRV with GCRV. The comparison of genome segments 1–11 of GCRV and GSRV showed high degrees of identity (nucleotide sequences, 90·56–97·68%; amino acid sequences, 96–99·75%). Cognate segments in the genomes of

Table 2. Properties of dsRNA segments of GSRV, GCRV and CSRV

Segment	Length (bp)	Putative encoded protein		5′ NCR		3′ NCR	
		Length (aa)	Mass (Da)	Length (bp)	Terminal sequence	Terminal sequence	Length (bp)
GSRV							
1	3949	1299	141 266	12	5′ GUUAUUU	UUCAUC 3′	37
2	3877	1274	141 585	12	5′ GUUAUUU	UUCAUC 3′	40
3	3702	1214	132 058	12	5′ GUUAUUU	UUCAUC 3′	45
4	2320	742	79 463	25	5′ GUUAUUG	UUCAUC 3′	66
5	2239	728	80 249	17	5′ GUUAUUU	AUCAUC 3′	35
6	2039	648	68 557	30	5′ GUUAUUU	UUCAUC 3′	62
7	1414	146, 274*	15 705, 31 171	13	5′ GUUAUUU	UUCAUC 3′	70
8	1297	412	44 594	12	5′ GUUAUUU	UUCAUC 3′	46
9	1130	352	37 695	31	5′ GUUAUUU	AUCAUC 3′	40
10	909	276	29 790	30	5′ GUUAUUU	UUCAUC 3′	48
11	820	244	26 491	42	5′ GUUAUUG	UUCAUC 3′	43
				Consensus	5′ GUUAUUu/g	a/uUCAUC 3′	
GCRV							
1	3949	1299	141 406	12	5′ GUUAUUU	UUCAUC 3′	37
2	3877	1274	141 536	12	5′ GUUAUUU	UUCAUC 3′	40
3	3702	1214	132 104	12	5′ GUUAUUU	UUCAUC 3′	45
4	2320	742	79 642	25	5′ GUUAUUG	UUCAUC 3′	66
5	2239	728	80 243	17	5′ GUUAUUU	AUCAUC 3′	35
6	2039	648	68 600	30	5′ GUUAUUU	UUCAUC 3′	62
7	1414	146, 274*	15 706, 31 239	13	5′ GUUAUUU	UUCAUC 3′	70
8	1296	412	44 580	11	5′ GUUAUUU	UUCAUC 3′	46
9	1130	352	37 695	31	5′ GUUAUUU	AUCAUC 3′	40
10	909	276	29 805	30	5′ GUUAUUU	UUCAUC 3′	48
11	820	244	26 419	42	5′ GUUAUUG	UUCAUC 3′	43
				Consensus	5′ GUUAUUu/g	a/uUCAUC 3′	
CSRV							
1	3947	1297	140 930	13	5′ GUUUUAU	AUCAUC 3′	40
2	3867	1240	137 579	115	5′ GUUUUAU	AUCAUC 3′	29
3	3690	1210	131 949	18	5′ GUUUUAU	UUCAUC 3′	39
4	Partial	–	–	ND	ND	AUCAUC 3′	117
5	2242	723	80 151	21	5′ GUUUUAU	UUCAUC 3′	49
6	2052	643	68 878	53	5′ GUUUUAU	UUCAUC 3′	67
7	1395	452	49 842	5	5′ GUUUUAU	UUCAUC 3′	31
8	1317	417	45 335	12	5′ GUUUUAU	UUCAUC 3′	51
9	1118	350	38 057	25	5′ GUUUUAG	UUCAUC 3′	28
10	985	298	32 410	27	5′ GUUUUAG	UUCAUC 3′	61
11	781	144, 154, 118†	15 117, 16 901, 13 041	24	5′ GUUUUAG	UUCAUC 3′	52
				Consensus	5′ GUUUUAu/g	a/uUCAUC 3′	

* Segment is bicistronic.
† Segment is tricistronic.
ND, Not determined.

the two viruses had the same electrophoretic mobility. The detailed nucleotide and amino acid sequence identity values between cognate genes of these two viruses are shown in Table 3.

Comparison of GSRV and GCRV with MRV. BLAST comparison of the genomes of GCRV and GSRV with a local *Reoviridae* sequence database showed that they have remarkable similarity to the genome of MRV. Of the 11 segments of either GSRV or GCRV, eight segments showed significant identity (17–42%: 42% within the polymerase) to segments of MRV serotypes 1, 2 and 3 (Fig. 1). The correspondence was not always a function of the electrophoretic mobility, as shown in Fig. 1. Only segments 7, 10 and 11 did not show clear homology to genes

Table 3. Comparison of segments 1–11 of the different aquareoviruses

Values are percentage nucleotide (NUC) or amino acid (PROT) sequence identities. –, Sequence not available for GIRV or SBRV.

	GIRV		SBRV		CSRV		GCRV	
Segment	NUC	PROT	NUC	PROT	NUC	PROT	NUC	PROT
GSRV								
1	–	–	–	–	53·9	44·1	94·7	98·9
2	71·7	81·3	64·8	66·3	64·0	65·3	94·4	99·5
3	–	–	62·5	61·1	61·1	55·1	90·6	99·8
4	–	–	49·3	24·2	50·9	25·9	96·2	99·1
5	59·4	56·6	–	–	47·6	34·6	91·5	97·0
6		–	–	–	57·7	50·2	94·2	99·2
7	–	–	–	–	52·3	26 (NS31), 30 (NS16)	95·4	98 (NS31), 96(NS16)
8	–	–	55·4	44·9	55·6	45·2	94·8	98·8
9	–	–	–	–	52·1	38·5	92·2	100
10	–	–	50·1	19·7	48·7	20·5	91·5	96·4
11	–	–	–	–	49·8	24 (ORF1), 25 (ORF2)	97·7	98·8
GCRV								
1	–	–	–	–	54·0	44·0		
2	71·7	81·3	65·3	66·5	63·9	65·5		
3	–	–	62·5	61·1	62·5	55·1		
4	–	–	49·8	24·5	50·9	26·2		
5	59·2	56·7	–	–	47·5	34·4		
6	–	–	–	–	56·9	50·2		
7	–	–	–	–	52·2	26 (NS31), 30 (NS16)		
8	–	–	54·7	45·2	55·2	44·9		
9	–	–	–	–	52·8	38·5		
10	–	–	48·9	19·7	48·6	20·5		
11	–	–	–	–	50·5	24 (ORF1), 25 (ORF2)		
CSRV								
1	–	–	–	–				
2	62·0	64·8	78·7	95·9				
3	–	–	78·0	84·4				
4	–	–	70·8	71·2				
5	48·2	33·9	–	–				
6	–	–	–	–				
7	–	–	–	–				
8	–	–	77·9	86·8				
9	–	–	–	–				
10	–	–	74·1	80·9				
11	–	–	–	–				
SBRV								
1	–	–						
2	63·0	65·0						

of MRV. Such high values of amino acid sequence identity also help in the prediction of the function (and possibly the virus architecture) of the putative proteins of GSRV and GCRV (Fig. 1).

ORF analysis has shown that segments 7 of GSRV and GCRV are bicistronic, encoding two proteins. The first protein (designated NS16, based on its theoretical molecular mass), encoded by the sequence between bases 14 and 454, was found to show relatedness to VP12 of Colorado tick fever virus (amino acid sequence identity 28%, calculated from alignment generated by CLUSTAL W), while the second protein (NS31) showed no relatedness to any previously reported reovirus protein sequence.

The orthoreovirus genome segment 7 (segment S1) is also bicistronic, encoding the outer capsid cell-attachment protein and a basic protein of unknown function. Based on their similar organization, it is possible that genome segments 7 of *Aquareovirus C* and MRV are equivalent.

By analogy to *Aquareovirus A* and *B*, segment 10 of *Aquareovirus C* encodes outer capsid proteins. This is also true

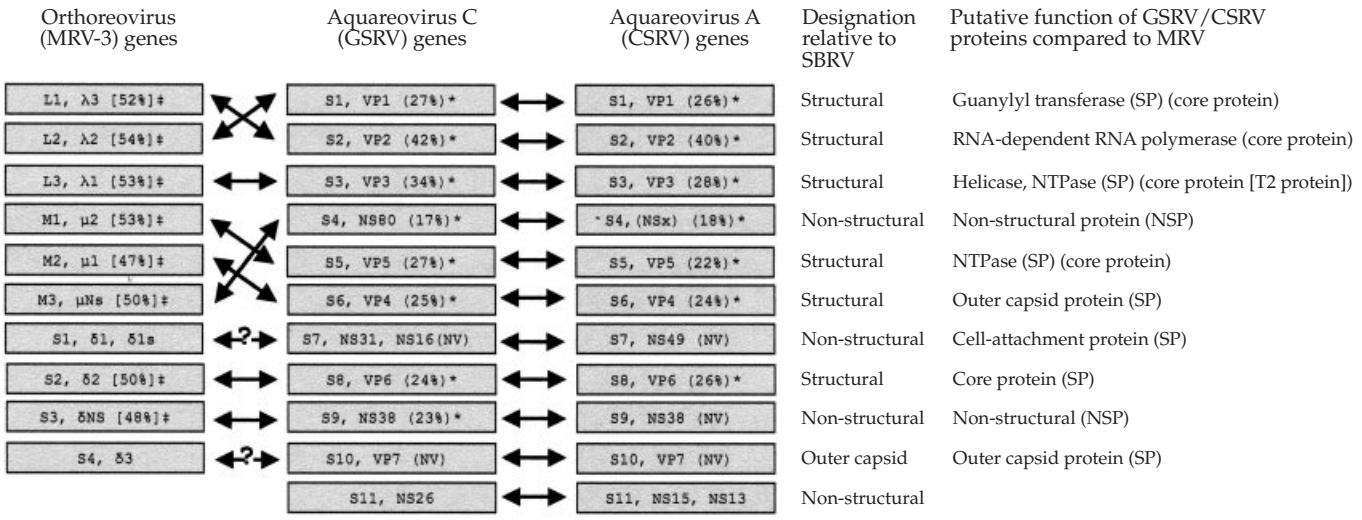

Fig. 1. Identities between *Aquareovirus C* (GSRV), *Aquareovirus A* (CSRV) and MRV serotype 3 (MRV-3). Bold double-headed arrows indicate homology between genes; arrows carrying question marks indicate that assignment was based on aspects of organization. Percentage amino acid identities between homologous proteins of either GSRV or CSRV and MRV-3 are indicated by asterisks (*); (NV), no valid amino acid sequence identity; ‡, genetic distances calculated from alignments generated with CLUSTAL W; (NSx), theoretical size of the protein unknown because of the sequence is partial; NSP, non-structural protein; SP, structural protein. The putative functions of proteins from *Aquareovirus A* and *Aquareovirus C* were predicted by comparison to the already-established functions of MRV-3 proteins (Mertens *et al.*, 2000). Designations relative to SBRV were taken from Subramanian *et al.* (1994).

for segment 10 of MRV. Analysis of the hydropathy plot of the proteins encoded by this segment of GSRV or GCRV and MRV showed similar hydropathy profiles in the amino-terminal part of the protein, with four domains in the order hydrophobic–hydrophilic–hydrophobic–hydrophilic (Fig. 2). Alignment of these protein sequences also showed numerous similar motifs (positions 21–27, GRLTLYT/GRVSIYS; 45–55, CGRYTICAFCL/CGGAVVCMHCL; 128–131, IVEL/LVEL; 248–255, DDGHQARSA/DFGHFGLSH) with respect to the MRV sequence. Accordingly, it is likely that segments 10 of *Aquareovirus C* and MRV are equivalent. If this is true, segment 11 of *Aquareovirus C* would have no equivalent in the MRV genome.

Analysis of the NCRs of GSRV and GCRV segments. Segments 1–11 of GCRV and GSRV were found to have conserved terminal sequences. All positive strands of each dsRNA segment had the motif 5′ GUUAUUU/G 3′ in common at the 5′ end and the motif 5′ A/UUCAUC 3′ in common at the 3′ end. The first and last nucleotides of each segment are inverted complements. In previous studies of reovirus genomes, comparable conserved motifs have been reported (Mertens *et al.*, 2000). They are supposed to act as sorting signals, bringing a single copy of each genome segment into the nascent virus capsid (Anzola *et al.*, 1987; Xu *et al.*, 1989).

Comparison of sequences of segments 8 and 10 among GCRV isolates. Sequence comparison of amplicons from different virus isolates showed that, for segments 8 and 10, nucleotide and amino acid sequence identities were nearly 100%.

Sequence analysis of *Aquareovirus A* viruses SBRV and CSRV

Comparison of SBRV to CSRV. Sequences from segments 2, 3, 4, 8 and 10 of SBRV were characterized in this study. Sequence comparison of these segments to their cognates in CSRV revealed nucleotide sequence identities between 70·77 and 78·65% and amino acid sequence identities between 71·2 and 95·87% (the highest values being those of polymerase sequences; Table 3).

Comparison of *Aquareovirus A* with *Aquareovirus C*. The nucleotide and amino acid sequence identities between SBRV or CSRV and GSRV or GCRV respectively ranged from 49·45 to 63·89% and from 20·45 to 65·53%. All segments of *Aquareovirus C* genomes have cognates in the genomes of *Aquareovirus A* viruses. The correspondence reflects the order of the electrophoretic mobility of the genomes perfectly, as shown in Fig. 1.

Three obvious differences between *Aquareovirus A* and *Aquareovirus C* genomes were noticed, the first being the monocistronic character of segment 7 of CSRV, while those of GCRV and GSRV are bicistronic. Remarkably, both proteins NS16 and NS31 from *Aquareovirus* C (encoded by the two distinct ORFs of segment 7) showed similarity to the protein encoded by segment 7 of CSRV (NS49, encoded by a unique ORF) (sequence identities shown in Table 3). The second difference is the lack of similarity between the NS38 protein (segment 9) of CSRV and the σNS protein (segment 9) of

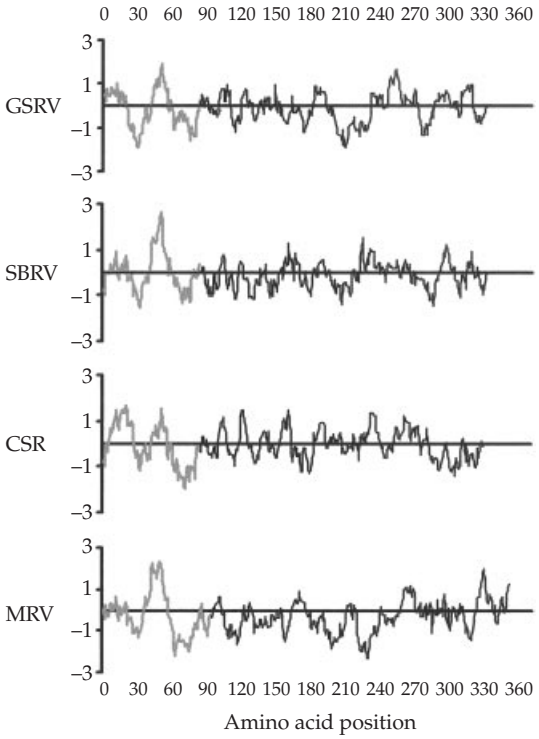

Fig. 2. Comparison of hydropathy profiles of aligned VP7 sequences (segment 10) of aquareoviruses (GSRV, SBRV, CSR) and σ3 sequence of MRV. The parts of the profiles represented by grey lines indicate the succession of the four domains: hydrophobic–hydrophilic–hydrophobic–hydrophilic.

MRV, while NS38 of GCRV and GSRV showed obvious similarity to this protein. The third difference is that segment 11 of CSRV is tricistronic (ORF1, nt 25–435; ORF2, nt 89–553; ORF3, nt 375–731). Only the proteins encoded by ORF1 (NS13) and ORF3 (NS15) showed similarity to the NS26 protein (segment 11, monocistronic) of *Aquareovirus C*.

Comparison of SBRV and CSRV with MRV. BLAST comparison of the genomes of SBRV and CSRV with a local *Reoviridae* sequence database revealed similarity to the genome of MRV. Values of calculated amino acid sequence identity between CSRV or SBRV and MRV ranged between 18 and 40%.

Similarly to *Aquareovirus C* species, segments 1, 2, 3, 4, 5, 6 and 8 of CSRV exhibited similarity to segments 1, 2, 3, 4, 5, 6 and 8 in the order given in Fig. 1. The relationships between segments 7, 10 and 11 of CSRV and those of MRV were comparable to those between *Aquareovirus C* and MRV. As mentioned above, and in contrast to *Aquareovirus C*, the NS38 protein (segment 9) of CSRV did not exhibit similarity to the σNS protein (segment 9) of MRV.

Analysis of NCRs of CSRV and SBRV segments. Segments 1–11 of CSRV and the full-length segments 8 and 10 of SBRV were found to have conserved terminal sequences. All positive strands of the sequenced dsRNA segments had the motif 5′ GUUUUAU/G 3′ in common at the 5′ end and the motif 5′ A/UUCAUC 3′ in common at the 3′ end. Again, the first and last nucleotides of each segment are inverted complements.

Comparison of terminal sequences of the NCRs of *Aquareovirus A* and *Aquareovirus C*. The conserved terminal sequences of *Aquareovirus A* are highly similar to those of *Aquareovirus C*, with one obvious difference. In the 5′ motif, the adenine base is located at position 4 (5′ GUUAUUU/G 3′) from the terminus for *Aquareovirus* C and at position 6 for *Aquareovirus* A (5′ GUUUUAU/G 3′).

Sequence analysis of GIRV segments 2 and 5

Comparison of GIRV with *Aquareovirus A* and *Aquareovirus C*. The nucleotide and amino acid sequence identities between GIRV and *Aquareovirus A* viruses respectively ranged from 48·24 to 62·97% and from 33·89 to 64·98%. The nucleotide and amino acid sequence identities between GIRV and *Aquareovirus C* viruses respectively ranged from 59·24 to 71·69% and from 56·73 to 81·34%.

Comparison of GIRV with MRV. BLAST comparison of GIRV genome with the local *Reoviridae* database showed similarity to the genome of MRV. Values of amino acid sequence identities between GIRV and MRV were 26% in segment 5 and 41% in segment 2. These values are comparable to those between *Aquareovirus A* or *Aquareovirus C* and MRV.

Analysis of NCRs of GIRV segments: comparison with *Aquareovirus A* and *Aquareovirus C*. Only segment 5 of GIRV was sequenced fully. The 5′- and 3′-terminal nucleotides of the positive strand of segment 5 were identical to those of *Aquareovirus A*.

Global analysis of the sequenced genomes of aquareoviruses

Sequence comparison of the putative RNA-dependent RNA polymerases of GCRV, GSRV, SBRV and GIRV and other members of the family *Reoviridae*. The amino acid sequences of the polymerases of GCRV, GSRV, SBRV and GIRV were compared with the sequences of putative RNA-dependent RNA polymerases of representative viruses from nine genera of the family *Reoviridae*: *Seadornavirus* (12 segments), species *Banna virus* (isolate BAV-In6423; accession no. AF133430) and *Kadipiro virus* (isolate KDV-Ja7075; AF133429); *Coltivirus* (12 segments), species *Colorado tick fever virus* (isolate CTFV-Fl; AF134529); *Orthoreovirus* (10 segments), species *Mammalian orthoreovirus* serotypes 1 (MRV-1; M24734), 2 (MRV-2; M31057) and 3 (MRV-3; M31058) and Ndelle virus (NDEV; AF368033); *Orbivirus* (10 segments), species *African horse sickness virus* serotype 9 (AHSV-9; U94887), *Bluetongue virus* serotypes 2 (BTV-2; L20508), 10 (BTV-10; X12819), 11 (BTV-11; L20445), 13 (BTV-13; L20446) and 17 (BTV-17; L20447) and *Palyam virus* isolate CHUV (Baa76549); *Rotavirus* (11 segments), species *Rotavirus A* (RV-A) strains BoRV-A/RF

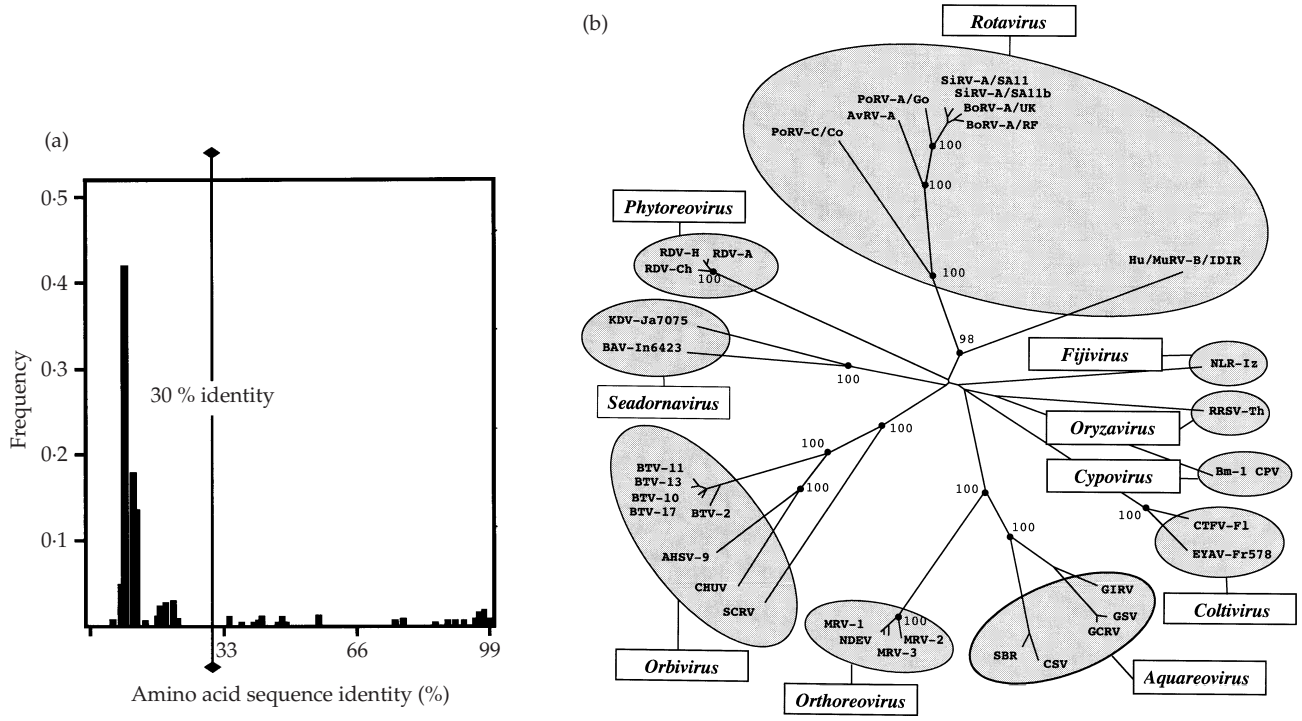

Fig. 3. (a) Frequency distribution histogram of amino acid sequence identities between polymerases of members of the *Reoviridae*. Sequence analysis included polymerase sequences of 34 viruses, including four isolates of MRV and five aquareoviruses. The vertical line at 30% amino acid sequence identity indicates the limits of distinction between genera except for *Rotavirus B*, which is 22% identical to other rotaviruses, and aquareoviruses, which show 40–42% amino acid sequence identity to orthoreoviruses. (b) Radial neighbour-joining tree displaying relationships between the polymerase sequences of members of the *Reoviridae*. Bootstrap values of 500 replications are indicated at the nodes (●) within each genus.

(J04346), BoRV-A/UK (X55444), SiRV-A/SA11b (X16830), SiRV-A/SA11 (AF015955), PoRV-A/Go (M32805) and AvRV-A (Baa24146), *Rotavirus B* strain Hu/MuRV-B/IDIR (M97203) and *Rotavirus C* strain PoRV-C/Co (M74216); *Fijivirus* (10 segments), species *Nilaparvata lugens reovirus* strain NLRV-Iz (D49693); *Phytoreovirus* (12 segments), species *Rice dwarf virus* isolates RDV-Ch (U73201), RDV-H (D10222) and RDV-A (D90198); *Oryzavirus* (10 segments), species *Rice ragged stunt virus* strain RRSV-Th (U66714); and *Cypovirus* (10 segments), species *Bombyx mori cytoplasmic polyhedrosis virus*-1 strain Bm-1 CPV (AF323781).

It was found that all members of a single genus exhibited amino acid sequence identities of over 30% (Fig. 3a). The only exception was *Rotavirus B*, which was only 22% identical to other rotaviruses. Between members of the genera *Aquareovirus* and *Orthoreovirus*, the amino acid sequence identity ranged from 40 to 42%. This value is therefore comparable to the amino acid sequence identity observed between members of a single genus. The results of this analysis are illustrated by a radial neighbour-joining tree (Fig. 3b).

Comparison of deduced amino acid sequences of segment 10 for isolates of viruses belonging to *Aquareovirus A*, *Aquareovirus B* and *Aquareovirus C*. In aquareovirus genomes, segment 10 encodes the major outer capsid protein VP7. VP7 sequences from different isolates of SBRV and CSRV (*Aquareovirus A*), CSR (*Aquareovirus B*), GCRV and GSRV (*Aquareovirus C*) were compared. The results showed that amino acid sequence identities ranged from 19·7 to 100% and that, within a species, amino acid sequence identities were greater than 80%. The amino acid sequence identities were found to be 80·87% among *Aquareovirus A* isolates, 95·77–99·62% among *Aquareovirus C* isolates, 19·53–21·29% between *Aquareovirus A* and *Aquareovirus C* isolates and 20·91–21·43% between *Aquareovirus B* and *Aquareovirus C* isolates.

Analysis of the G + C contents of the genomes of *Aquareovirus A*, *Aquareovirus B*, *Aquareovirus C* and GIRV. The G + C contents of the genomes of GCRV and GSRV ranged between 53 and 60 mol%; the highest value was calculated from segments 4 and 11. The G + C content of the genomes of GSRV and SBRV ranged between 53 and 57 mol%; the highest value was calculated from segment 4. The G + C content of segment 2 of GIRV was 52 mol% and that of segment 5 was 54 mol%. The G + C content of segment 10 of CSR was found to be 56 mol%.

Discussion

Aquareoviruses and orthoreoviruses share a number of common structural characteristics, in particular a similar ultrastructure in electron microscopy (Shaw *et al.*, 1996).

Besides having different numbers of genome segments, they are antigenically distinct and occupy distinct econiches. The first fish-virus isolates with polysegmented dsRNA genomes were obtained in the late 1970s and early 1980s and included GSRV, American oyster reovirus 13p2, CSRV and Channel catfish reovirus. These isolates were the first to be designated aquareoviruses (Winton *et al.*, 1987). Numerous viruses belonging to the family *Reoviridae* have been isolated from aquatic animals since these initial isolations. The majority of these viruses were found to possess genomes of 11 dsRNA segments and, hence, to belong to the genus *Aquareovirus*. However, virus isolates such as the W and P2 viruses, isolated from Mediterranean crab, were found to possess genomes of 12 dsRNA segments and were hence excluded from classification within the genus *Aquareovirus* (Montanie *et al.*, 1993). To date, aquareoviruses have been isolated from fish, molluscs and crustaceans (Mertens *et al.*, 2000), while orthoreoviruses have been isolated from reptiles, birds and mammals.

One of the first important findings of the current study pertains to the taxonomy of unclassified isolates. Firstly, the genome of GCRV was sequenced completely and genetic analyses have shown that it is almost identical to that of GSRV (*Aquareovirus C*): (i) the genomes of the two viruses are of comparable sizes (GCRV, 23 695 bp; GSRV, 23 696 bp); (ii) cognate segments of the viruses are of comparable sizes; (iii) each genome segment is flanked by identical 5′ and 3′ conserved NCRs; (iv) nucleotide and amino acid sequence identities between the two viruses are very high (respectively 90–97 % and 96–100 %) and the nucleotide sequence variation is mainly a function of the third codon position. These findings show clearly that GSRV and GCRV should be considered as isolates of the same species. Consequently, the hitherto unclassified GCRV belongs to the species *Aquareovirus C*.

Secondly, segments 2 and 5 of the unclassified GIRV were sequenced. Assessment of the genetic relationship of GIRV to the species *Aquareovirus B*, *D*, *E* and *F* was impossible because of the absence of sequence data from these species, but sequence comparison permitted us to exclude GIRV from *Aquareovirus A* and *Aquareovirus C*.

Besides the study of members of the genus *Aquareovirus*, the current study sheds new light on the relationship between orthoreoviruses and aquareoviruses. It is a general rule within the family *Reoviridae*, that the genetic relatedness of viruses belonging to different genera is very low. In the polymerase gene, the only one that allows sequence comparison between different genera, analysis of amino acid sequence identity frequency distribution (Fig. 3a) shows that viruses belonging to different genera have low amino acid sequence identity (< 20 %), denoting a very distant phylogenetic relationship. This is more evident in other genes for which comparison is practically impossible. The situation of orthoreoviruses and aquareoviruses is therefore a remarkable exception: (i) in the polymerase gene, the amino acid sequence identity between members of the two genera is up to 42 % (a value usually observed between members of a single genus); (ii) a clear genetic relationship can be observed between members of the two genera in seven other genes (see Fig. 1) where amino acid sequence identities range between 17 and 42 %; (iii) the T2 protein (the innermost shell of the capsid) of members of the family *Reoviridae* is a protein that is structurally conserved between viruses belonging to different genera; however, it is not possible to perform a significant sequence comparison. This protein is identified as the λ1 protein in the orthoreovirus genome. The equivalent aquareovirus protein, VP3, shows significantly high (37 %) amino acid sequence identity to λ1. This means that the genetic relationship between orthoreoviruses and aquareoviruses is comparable to that between viruses such as St Croix River virus and *Bluetongue virus* (belonging to the genus *Orbivirus*) or *Kadipiro virus* and *Banna virus* (both belonging to the genus *Seadornavirus*). Homology was also identified in the conservation of terminal sequences. For instance, the 5′ and 3′ conserved terminal sequences of GCRV segments are 5′ GUUAUU 3′ and 5′ A/UUCAUC 3′, compared to 5′ GCUAUU 3′ and 5′ A/UUCAUC 3′ in genome segments 4, 7, 8 and 10 of MRV.

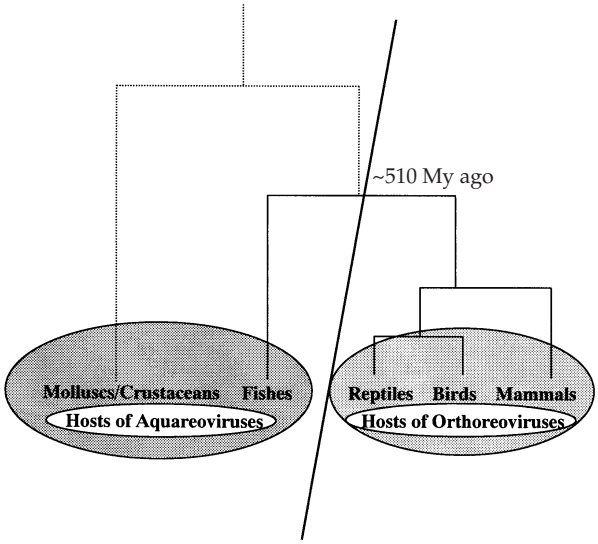

Fig. 4. Simplified scheme of the evolution of the main hosts of orthoreoviruses and aquareoviruses. The separation between the hosts of aquareoviruses and orthoreoviruses about 510 My ago is shown by the diagonal line.

Altogether, these data are undoubtedly an indication of the common evolutionary origin of these viruses. Fig. 4 shows a simplified scheme of the evolution of the main hosts of orthoreoviruses and aquareoviruses. The common ancestor of fish and the group 'reptiles + birds + mammals' existed around 510 million years (My) ago. It has been proposed that the molecular evolutionary rate of genomes of related dsRNA viruses is 10^{-8} to 10^{-9} mutations/nucleotide/year, which is equivalent to that of dsDNA (Attoui *et al.*, 2002). If this is applied to the polymerases of orthoreoviruses and aquareo-

viruses, it appears that divergence between these two groups occurred 49–520 My ago. Despite this imprecision in the evaluation of divergence, it cannot be excluded that orthoreoviruses appeared following the emergence of the evolutionary group that eventually gave rise to reptiles, birds and mammals.

Finally, these results raise interesting questions concerning the taxonomic classification of orthoreoviruses and aquareoviruses in different genera on the one hand, and the relevance of quantitative taxonomy based on polymerase sequences in the *Reoviridae* on the other hand.

Concerning the first question, it should be noted that the ICTV defines taxa as members of a polythetic class. Therefore, despite the unusual genetic relatedness between orthoreoviruses and aquareoviruses, there are strong arguments that justify their classification within two separate genera: (i) as reported above, the viruses originate from distinct econiches; (ii) aquareoviruses have genomes composed of 11 segments and orthoreovirus genomes are composed of 10 segments of dsRNA; (iii) the G + C content of orthoreoviruses ranges between 44 and 48 mol%, while that of aquareoviruses ranges between 52 and 60 mol%; (iv) orthoreoviruses are non-syncytializing viruses, in contrast to the majority of aquareoviruses; and (v) there is no antigenic relationship between the two groups.

Accordingly, the authors consider that the maintenance of the classification of these viruses in two different genera is justified and that orthoreoviruses and aquareoviruses constitute an interesting, but isolated, example of two genera that undoubtedly originate from a common evolutionary ancestor.

Concerning quantitative taxonomy using polymerase sequences, analysis of amino acid sequence identity frequency distribution (Fig. 3a) shows that all members of a single genus have amino acid sequence identities of $> 30\%$. There is, however, one exception, *Rotavirus B*, which is only 22% identical to other rotaviruses. Similarly, it can be observed that all viruses belonging to different genera have amino acid sequence identities of $< 30\%$, the only exceptions being orthoreoviruses and aquareoviruses, as discussed above. Therefore, the only criterion that remains indisputable is the assignment of two viruses to different genera if their polymerase amino acid sequence identity is $< 20\%$. Clearly, this is a modest contribution to phylogenetic classification. However, it is important to pursue the sequence characterization of representative members of the family *Reoviridae* to allow better definition of the quantitative basis of species delineation in the different genera and to try to improve the genetic criteria for the definition of genera.

The authors wish to thank Dr S. K. Samal for providing SBRV and Dr M. Neukirch for providing GIRV. The authors also thank Dr S. Rao and colleagues for releasing the sequence of CSRV to the database. This work was supported by EU grant 'Reo ID' no. QLK2-2000-00143. The 'Unité des Virus Emergents' is an associated research unit of the Institut de Recherche pour le Développement (IRD). This study was supported in part by the IRD and the EFS Alpes-Méditerranée.

References

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

Attoui, H., Billoir, F., Cantaloube, J.-F., Biagini, P., de Micco, P. & de Lamballerie, X. (2000). Strategies for the sequence determination of viral dsRNA genomes. *Journal of Virological Methods* **89**, 147–158.

Attoui, H., Mohd Jaafar, F., Biagini, P., Cantaloube, J.-F., de Micco, P., Murphy, F. A. & de Lamballerie, X. (2002). Genus *Coltivirus* (family *Reoviridae*): genomic and morphologic characterization of Old World and New World viruses. *Archives of Virology* **147**, 533–561.

Chen, B. S. & Jiang, Y. (1984). Morphological and physicochemical characterization of the hemorrhagic virus of grass carp. *Kexue Tongbao* **29**, 832–835 (in Chinese).

Fang, Q., Ke, L. H. & Cai, Y. Q. (1989). Growth characterization and high titre culture of GCHV. *Virologica Sinica* **3**, 315–319.

Fang, Q., Attoui, H., Cantaloube, J. F., Biagini, P., Zhu, Z., de Micco, P. & de Lamballerie, X. (2000). Sequence of genome segments 1, 2, and 3 of the grass carp reovirus (genus *Aquareovirus*, family *Reoviridae*). *Biochemical and Biophysical Research Communications* **274**, 762–766.

Francki, R. I. B., Fauquet, C. M., Knudson, D. L. & Brown, F. (1991). *Classification and Nomenclature of Viruses. Fifth Report of the International Committee on Taxonomy of Viruses*. Wien: Springer-Verlag.

Ke, L. H., Fang, Q. & Cai, Y. Q. (1990). Characteristics of a novel isolate of grass carp hemorrhage virus [J]. *Acta Hydrobiologica Sinica* **14**, 153–159.

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

Kyte, J. & Doolittle, R. F. (1982). A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology* **157**, 105–132.

Mertens, P. P. C., Arella, M., Attoui, H., Belloncik, S., Bergoin, M., Boccardo, G., Booth, T. F., Chiu, W., Diprose, J. M., Duncan, R. and 34 others (2000). Family *Reoviridae*. In *Virus Taxonomy. Seventh Report of the International Committee for the Taxonomy of Viruses*, pp. 395–480. Edited by M. H. V. van Regenmortel, C. M. Fauquet, D. H. L. Bishop, E. B. Carstens, M. K. Estes, S. M. Lemon, J. Maniloff, M. A. Mayo, D. J. McGeoch, C. R. Pringle & R. B. Wickner. San Diego: Academic Press.

Montanie, H., Bossy, J.-P. & Bonami, J.-R. (1993). Morphological and genomic characterization of two reoviruses (P and W2) pathogenic for marine crustaceans; do they constitute a novel genus of the Reoviridae family? *Journal of General Virology* **74**, 1555–1561.

Neukirch, M., Haas, L., Lehmann, H. & von Messeling, V. (1999). Preliminary characterization of a reovirus isolated from golden ide *Leuciscus idus melanotus*. *Diseases of Aquatic Organisms* **35**, 159–164.

Page, R. D. (1996). TreeView: an application to display phylogenetic trees on personal computers. *Computer Applications in the Biosciences* **12**, 357–358.

Plumb, J. A., Bowser, P. R., Grizzle, J. M. & Mitchell, A. J. (1979). Fish viruses: a double stranded RNA icosahedral virus from a North American cyprinid. *Journal of the Fisheries Research Board of Canada* **36**, 1390–1394.

Saitou, N. & Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* **4**, 406–425.

Shaw, A. L., Samal, S. K., Subramanian, K. & Prasad, B. V. V. (1996). The structure of aquareovirus shows how the different geometries of the two layers of the capsid are reconciled to provide symmetrical interactions and stabilization. *Structure* **4**, 957–967.

Subramanian, K., McPhillips, T. H. & Samal, S. K. (1994). Charac-

terization of the polypeptides and determination of the genome coding assignments of an aquareovirus. *Virology* **205**, 75–81.

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 Research* **22**, 4673–4680.

Winton, J. R., Lannan, C. N., Fryer, J. L., Hedrick, R. P., Meyers, T. R., Plumb, J. A. & Yamamoto, T. (1987). Morphological and biochemical properties of four members of a novel group of reoviruses isolated from aquatic animals. *Journal of General Virology* **68**, 353–364.

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

Received 16 January 2002; Accepted 24 March 2002