Molecular characterization of a novel adult diarrhoea rotavirus strain J19 isolated in China and its significance for the evolution and origin of group B rotaviruses

Shengjun Jiang,¹ Shaozhong Ji,² Qing Tang,² Xiaoying Cui,²
Hongyang Yang,² Biao Kan² and Shouyi Gao²

Correspondence

Shaozhong Ji
szji002@yahoo.com.cn

¹Tropical Crops Genetic Resources Institute, Chinese Academy of Tropical Agricultural Sciences, Danzhou city, Hainan province 571737, PR China

²State Key Laboratory for Communicable Disease Control and Prevention, Chinese Center for Disease Control and Prevention, Beijing 102206, PR China

The complete genome of a novel adult diarrhoea rotavirus strain J19 was cloned and sequenced using an improved single-primer sequence-independent method. The complete genome is 17 961 bp and is AU-rich (66.49%). Northern blot analysis and genomic sequence analysis indicated that segments 1–11 encode 11 viral proteins, respectively. Protein alignments with the corresponding proteins of J19 with B219, and groups A, B and C rotaviruses, produced higher per cent sequence identities to B219. Among groups A, B and C rotaviruses, 10 proteins from group B rotaviruses exhibited slightly higher amino acid sequence identity to the J19 proteins, but proteins of J19 showed low amino acid sequence identity with groups A and C rotaviruses. Construction of unrooted phylogenetic trees using a set of known proteins and representatives of three known rotavirus groups revealed that six structural proteins were positioned close to B219 and the basal nodes of groups A, B and C lineages, although with a preferred association with group B lineages. Phylogenetic analysis of the five non-structural proteins showed a similar trend. The results of the serological analysis, protein sequence analysis and phylogenetic analysis suggested that J19 would be a novel rotavirus strain with great significance to the evolution and origin of group B rotaviruses.

Received 8 March 2008

Accepted 20 May 2008

INTRODUCTION

Rotaviruses are important aetiological agents of disease in humans and animals. The viral genome is composed of 11 segments of double-stranded (ds) RNA that encode structural and non-structural proteins. The group A rotavirus strain SA11 genome, sequenced in 1990, is AU-rich and has 18 555 bp (Both *et al.*, 1984; Estes *et al.*, 1984; Mitchell & Both, 1990). Its RNA segments encode six structural proteins (VP1, VP2, VP3, VP4, VP6 and VP7) and five non-structural proteins (NSP1, NSP2, NSP3, NSP4 and

NSP5) (Estes, 2001). VP6 is the group antigen determinant, making up approximately 50 % of the viral protein (Estes *et al.*, 1984; Estes & Cohen, 1989). The antigenic properties of VP6 are used to define seven antigenically different rotavirus groups, named groups A–G. Groups A, B and C rotaviruses infect humans, while other groups have been found only in animal species (Saif & Jiang, 1994). Group A rotaviruses cause severe diarrhoea in infants and young children, but they can also infect adults. Group C rotavirus infection occurs both in children and in adults, usually in sporadic cases or clustered outbreaks (Rodger *et al.*, 1982; Caul *et al.*, 1990; Jiang *et al.*, 1995; Kuzuya *et al.*, 1996; Nilsson *et al.*, 2000; Adah *et al.*, 2002; Chen *et al.*, 2002; Ji *et al.*, 2002; Schnagl *et al.*, 2004).

In China, large waterborne epidemics caused by the human group B rotavirus strain adult diarrhoea rotavirus (ADRV) infected thousands of people aged between 10 and 40 years of age in the 1980s (Hung *et al.*, 1983, 1984; Chen *et al.*, 1985, 1990). Since then, only sporadic ADRV waterborne epidemics have been detected in China. However, two human group B rotavirus strains identified in 1998 in India

The GenBank/EMBL/DDBJ accession numbers for the full genomic sequences and their deduced protein sequence data of the novel adult diarrhoea rotavirus strain J19 are: VP1 gene DQ113897, protein AAZ03485.1; VP2 gene DQ113898, protein AAZ03486.1; VP3 gene DQ113900, protein AAZ03488.1; VP4 gene DQ113899, protein AAZ03487.1; VP6 gene DQ113902, protein AAZ03490.1; VP7 gene DQ113905, protein AAZ03493.1; NSP1 gene DQ113901, protein AAZ03489.1; NSP2 gene DQ113903, protein AAZ03491.1; NSP3 gene DQ113904, protein AAZ03492.1; NSP4 gene DQ113906, protein AAZ03494.1; NSP5 gene DQ113907, protein AAZ03495.1.

Supplementary material is available with the online version of this paper.

(CAL-1) and 2000 in Bangladesh (Bang373), showed high sequence identity with the Chinese ADRV (Krishnan *et al.*, 1999; Sanekata *et al.*, 2003).

The Chinese novel ADRV (NADRV) was first identified in sporadic cases in Huaihua city in 1987 and in Fuzhou city in 1988, identification was based on a specific electrophoresis migration pattern (Wang *et al.*, 1987). The same rotavirus was implicated in gastroenteritis outbreaks in Beijing city in 1994 and in Shijiazhuang city in 1997. Partial genes of NADRV were cloned and sequenced from stool samples of patients and primary results showed that the novel rotavirus did not belong to known groups A, B or C rotaviruses (Yang *et al.*, 1998, 2004a). Later, an unusual human rotavirus strain, B219, was detected in a stool specimen from a 65-year-old patient with diarrhoea in Bangladesh during April 2002, which is genetically closely related to the NADRV strain (Alam *et al.*, 2007). Recently, the novel rotavirus causing adult diarrhoea in a large scale in Shijiazhuang city in China in 1997 was isolated from stool samples and was found to be similar to NADRV. It was propagated successfully in human embryonic kidney cells *in vitro* and was designated a novel ADRV strain, J19 (Ji *et al.*, 2002). Here, we further characterized strain J19 by cloning and analysing its genome segments. The information will provide direction toward understanding the origin and evolution of this novel rotavirus.

METHODS

Virus cultures and dsRNA preparation. The NADRV strain J19 was propagated in human embryonic kidney cells in Dulbecco's modified Eagle's medium with 100 µg trypsin (Promega) ml⁻¹ (Ji *et al.*, 2002). The rotavirus dsRNA was extracted from infected tissue cultures (Gouvea *et al.*, 1990) and treated with RNase T₁ (Promega) to degrade the single-stranded RNA (for details see Supplementary material available in JGV Online).

cDNA preparation, PCR amplification and coding assignment. The complete J19 genome was amplified and cloned using an improved single-primer sequence-independent method (Lambden *et al.*, 1992; Jiang *et al.*, 2006). The cloning method involved ligation of a single-amino-linked modified oligonucleotide to the 3' termini of each dsRNA segment by using T₄ RNA ligase. The tailed RNA was converted to cDNA by using a complementary primer and reverse transcriptase. The resultant cDNA was annealed and repaired and amplified by PCR by using a single complementary oligonucleotide primer. The PCR products were ligated into pMD-18T-vector (TaKaRa) and transformed into JM109 cells. In order to amplify and clone genome segments of J19 larger than 2 kb, pairs of inhibitory primers, which were blocked with NH₂ at the 3' termini designed from the 5' and 3' termini of seven known J19 genes by a single-primer sequence-independent method, were used to prevent amplification of short gene sequences and improve amplification of genome segments larger than 2 kb (Lambden *et al.*, 1992; Jiang *et al.*, 2006). Conditions for these PCR reactions were the same as described above, with the addition of inhibition primers. Inhibition primers would anneal to their corresponding genes during the PCR amplification and prevent their PCR amplification as their 3' termini was blocked by NH₂. Dig-labelled cDNA probes for gene assignment studies were generated as described previously (Ando *et al.*, 1995) and the coding assignments were determined by Northern blot analysis (for details see Supplementary material available in JGV Online).

Sequence analysis and phylogenetic analysis. The nucleotide sequences were analysed using the ORF Finder and BLAST software program (NCBI). Using the known structural and non-structural protein sequences representing the three known rotavirus groups and J19, B219 and NADRV rotaviruses, the phylogenetic trees were generated using the MEGALIGN part of the DNASTAR software package (DNA Star). All viral protein accession numbers were previously published in GenBank.

RESULTS

Complete J19 genomic cloning and sequencing

Using the single-primer sequence-independent PCR amplification method reported by Lambden *et al.* (1992), seven small genes of J19 were amplified, cloned and sequenced. Although four large genes more than 2 kb could be amplified with the single primer, they failed to clone with the T-vector. This meant that the quantity of DNA from the PCR products of large genes was not enough to clone them. In order to increase the yields of PCR products of these large genes, seven pairs of inhibitory primers designed from the small genes were used to amplify the reverse transcribed cDNA. To determine the right combination of inhibitory primers, one pair, two pairs and three pairs of inhibitory primers were tested in the PCR. Results showed that one pair of inhibitory primers from genes 8, 9 or 11 could effectively improve amplification of the four genes that are larger than 2 kb. This means that every pair of inhibitory primers would have different inhibitory effects to improve yields of large genome segments. When two pairs of inhibitory primers were added to one PCR, all combinations significantly improved the amplification of the large genome segments. This means that two pairs of inhibitory primers have synergistic effects to improve yields of the large genome segments. When three pairs of inhibitory primers were added in one PCR, results were no better than those of two pairs of inhibitory primers. Therefore, we chose to use two pairs of inhibitory primers to improve yields of the PCR products of the large genome segments of rotaviruses. Although PCR yields of large genome segments were improved by adding inhibitory primers, PCR yields of small genome segments were still significant. This implied that amplification of small genome segments was only partially inhibited. Although PCR yields of the four large genome segments of J19 were increased significantly with the addition of inhibition primers, we still failed to clone the full-length cDNA of gene 1 of J19. However, a shorter fragment of gene 1 was cloned and the full-length sequencing of gene 1 was completed by genome walking. Finally, the complete genome of J19 was obtained, and the 5' and 3' terminal nucleotides of 11 genes of NADRV-J19 (Table 1) illustrated that they have the same conserved terminal sequence (5'-GGXA.....ACCC-3') as rotaviruses and proved that genomic data of these 11 genes of NADRV-J19 should contain the full-length terminal nucleotide sequences of their corresponding dsRNA genes.

Table 1. Non-protein-coding features of 11 gene segments of NADRV J19

Gene	Base	Putative transcription regulatory sequence	Base	3' untranslated region
1	1–6	GGCACT (6) ATGGAA	3511–3538	TAA (3511)ACGTGAGAGTACATCCAT-TAATATACCC
2	1–19	GGCACTTAAGCGCTGCAAG (19) ATGGAA	2942–2969	TAA (2942)ATCACGATCTAGCTTATT-CAATATACCC
3	1–10	GGCACTTAAG (10) ATGTCT	2483–2512	TAG (2483)ATGATATGAAAGAGATCT-TTAAATATACCC
4	1–8	GGCACTTA (8) ATGGCT	2169–2204	TAG (2169)GTCATAAGTAACCTTAATG-AGTTGCTGTAATATACCC
5	1–55	GGAATTTTAGTGGTTTCTCTGATTACA-TTGTCTGGCTTCTAGGACAGTTAGAA-AG (55) ATGATG	1244–1307	TAA (1244)TTTCCTTGTTGCAGCAAG-GATGACTAGTGCATTCAATGTATGA-TTAGCGACCTATCATATACCC
6	1–31	GGCAATTTCTTGCTACAAGTGACTGTG-AAGG (31) ATGGAT	1223–1287	TGA (1223)TTGGTCAGGTTTGGTCAC-GCGCTGCGAAGCGGCGTTCCGTAG-TAGTATAGAAATAAAATATACCC
7	1–52	GGTAAAAAGCTCTTATCTTAGTCACAT-CATCCAGGGTTGCAGCTGGGAGAG-A (52) ATGGTG	947–1004	TAG (947)ATGTTATCAGTAACATTTT-GATCGAAAAAACTAAGATAATCTT-TGATTGATATACCC
8	1–45	GGAATTTTCGATTCTTAAGACGAGCTT-GCTCCGATAGAGTCGGAA (45) ATGGCG	835–932	TGA (835)ATGTAATGAAATACAATC-AATTAAGTACAATTGAATATACTTA-AATCACAAATAATCCGTTAATTACC-CAAGATCT CTCCTTGGGCTAAATACACCC
9	1–14	GGCAATTTGAAGCC (14) ATGTTG	792–820	TGA (792)GAAATCCCGAATGGTGAGC-AAATATACCC
10	1–25	GGCATTTTGTTCATCACAAATCACG (25)- ATGGAA	668–739	TGA (668)TGTGTAGCATGCTCTACCA-TGGATGGAAACATAGCTTGGTAGA-GGCTCTCCCATGCTCAATGCAATAT-ACCC
11	1–56	GGAACCTAAAAACATCAATCGTTGCTGC-AACAGCGGTTGGTTACCTCTTGCCCT-AGA (56) ATGAGC	588–649	TAA (588)TCGAGAGCCGGATACATTC-CCCTATTACTAAGGGTCAGCCCC-TGAgTCAATAGTATACCC

Northern blot analysis of the J19 strain with probes representing these 11 full-length viral proteins confirmed the coding assignments of the clones for these 11 genes (data not shown) and sequencing analysis showed that the 11 genes encoded six structural proteins and five non-

structural proteins (Table 2) corresponding to the equivalents of VP1–VP4, VP6, VP7 and NSP1–NSP5 of group A rotavirus strain SA11. The genome of J19 is AU-rich (66.49%), with a base composition of A (36.45%), G (17.83%), T (30.04%) and C (15.68%), and 66.1%

Table 2. Molecular analysis of the 11 J19 gene segments

RNA segment	Length (bp)	ORF (nt)	Protein	Length (aa)	Molecular mass (Da)	pI
1	3538	7–3510	VP1	1167	132 890	8.94
2	2969	20–2941	VP2	973	111 046	5.81
3	2512	11–2482	VP4	823	92 911	6.99
4	2204	9–2168	VP3	719	84 710	8.58
5	1307	56–1243	NSP1	395	45 649	8.29
6	1287	32–1222	VP6	396	42 921	6.03
7	1004	53–946	NSP2	297	33 432	8.27
8	932	46–834	NSP3	262	30 117	7.26
9	820	15–791	VP7	258	29 264	5.04
10	739	26–667	NSP4	213	25 110	5.17
11	649	57–587	NSP5	176	20 137	5.16

sequence identity with SA11. The complete genome of J19 is 17961 bp in length, which is 51 nt longer than the genome of human group C rotavirus Bristol (17910 bp), but 594 and 884 nt shorter than group A rotavirus strains SA11 (18555 bp) and PO-13 (18845 bp), respectively.

Table 1 shows non-protein-coding features of these 11 gene segments of NADRV J19. From these data, a terminal GG-sequence, common to all rotaviruses, was located at the 5'-end of all the cloned genes, while at the 3'-end of the genes -ATACCC sequence was found in 10 cloned genes (1–7 and 9–11), -ACACCC sequence was found only in the NSP3 gene, which is similar to that of NADRV and B219, but different from those of groups A, B and C rotaviruses (Yang *et al.*, 2004a; Alam *et al.*, 2007). Analysis of the 11 messenger sense RNA fragments with an RNA secondary structure prediction program (GeneBee) demonstrated that all of them could base-pair with themselves for nearly the entire length of the RNA and their 5' and 3' termini could form panhandle structures (Brodsky *et al.*, 1992, 1995).

Nucleotide sequence comparison (data not shown) indicated that three J19 genes have almost the same nucleotide sequences as NADRV, which further proved that J19 was isolated from similar stool samples of NADRV. Results showed that these 11 genes of J19 showed higher nucleotide sequence identities with B219 (more than 90% at nucleotide acid level), but nucleotide sequence of 10 genes of J19 from the VP1 gene showed less than 60% identity to groups A, B and C representative strains.

Deduced protein sequence comparisons between J19 and NADRV, B219 and groups A, B and C

Protein alignments with the corresponding proteins of NADRV and J19 showed that J19 and NADRV have the same protein sequences in NS1, VP6 and NSP3, which further proved that J19 was isolated from the novel

rotavirus that caused an outbreak of gastroenteritis in China. Protein alignments with the corresponding proteins of J19 with B219, and groups A, B and C rotaviruses, produced higher per cent sequence identities to B219 (generally more than 90% at the amino acid level) (Table 3). Among groups A, B and C rotaviruses, 10 proteins from group B rotaviruses exhibited slightly higher amino acid sequence identities to the J19 proteins (13.1–55.7%), but proteins of J19 showed low amino acid sequence identities (less than 30% at the amino acid level) with groups A and C rotaviruses.

VP6 is the major group antigen or serogroup antigen in rotavirus proteins. Rotaviruses that share the group antigen are now classified into groups A–G. VP6 of J19 only showed 8.3, 37.9 and 12.9% identities to groups A, B and C representative rotaviruses, respectively, which agrees with previous findings that J19 failed to react with groups A and B antisera (Ji *et al.*, 2002). However, analysis with the available VP6 proteins in GenBank showed 69.8–99.2, 84.4–99.0 and 39.8–100.0% identities, respectively, among groups A, B and C rotaviruses. Combining with the previous results of serological analysis with groups A, B and C rotaviruses, this implies that J19 would be a representative strain of a new group.

Phylogenetic analysis

Unrooted phylogenetic trees were constructed using a set of known proteins and representatives of NADRV, B219 and three known rotavirus groups A, B and C. It was revealed that six of the structural proteins encoded by J19 (VP1–VP4, VP6 and VP7) were positioned close to B219 and the basal nodes of groups A, B and C lineages, although with a preferred association with group B lineages (Fig. 1). The VP6 tree comprises four distinct lineages: two lineages including only J19, B219 and group B rotaviruses, and the other two lineages including groups A and C

Table 3. Amino acid sequence identity comparison among J19 with NADRV, B219, groups A, B, and C rotavirus representative strains

RNA segment	Protein	NADRV (%)	B219 (%)	Group A representative strain (%)	Group B representative strain (%)	Group C representative strain (%)
1	VP1 (J19)	–	97.2	SA11 (20.4)	IDIR (55.7)	Bristol (19.2)
2	VP2 (J19)	–	98.0	SA11 (11.8)	IDIR (46.9)	Bristol (10.7)
3	VP4 (J19)	–	95.5	WA (11.7)	IDIR (24.5)	Shintoku (13.4)
4	VP3 (J19)	–	93.7	WA (16.3)	–	Shintoku (13.1)
5	NSP1 (J19)	100.0	94.9	TB-Chen (9.6)	KB63 (18.3)	Cowden (10.9)
6	VP6 (J19)	100.0	98.7	SA11 (8.3)	IDIR (37.9)	Cowden (12.9)
7	NSP2 (J19)	–	97.0	PO-13 (10.4)	ADRV (46.1)	Bristol (6.4)
8	NSP3 (J19)	100.0	93.5	WA (10.33)	IDIR (24.4)	Bristol (14.9)
9	VP7 (J19)	–	96.5	SA11 (11.2)	IDIR (18.3)	Bristol (9.3)
10	NSP4 (J19)	–	88.7	WA (13.1)	WH1 (13.1)	Bristol (11.3)
11	NSP5 (J19)	–	93.2	WA (6.8)	IDIR (25.3)	Cowden (1.7)

Fig. 1. Phylogenetic analysis of six structural proteins of J19. Bar, the number of amino acid substitutions ($\times 100$). Group A rotavirus representative strains: SA11, UK, TB-Chen, KU, YM, PO-13, YM1, DS1, WA and DS1; group B rotavirus representative strains: IDIR, Bang373, CAL-1, WH1, ADRV and DB176.

rotaviruses. The group B rotavirus lineage is further divided into infectious diarrhoea of infant rats (IDIR) and two sister groups: the Chinese lineage containing human group B rotavirus strains ADRV and WH1 and the Indian-Bangladeshi lineage containing human group B rotavirus strains CAL-1 and Bang373. Rooted with NSP1–NSP5, the trees for NSP1–NSP5 (Fig. 2) are topologically consistent with trees constructed by using six structural proteins. From these results and consulting the demarcation criteria used for rotavirus species assignment (Ramig *et al.*, 2005), J19 would be genetically related to NADRV, B219 and group B rotaviruses, which would be proposed as the prototype of a new species with great significance for the evolution and origin of group B rotaviruses.

DISCUSSION

Rotaviruses are important aetiological agents for diseases in humans and animals. Amplification and cloning of

rotavirus genes are important to monitor infection, variation and recombinant vaccine development. However, it is difficult to clone genome segments larger than 2 kb for newly discovered rotaviruses. Two kinds of methods were used to clone rotavirus genes. The first kind was based on polyadenylation of genomic dsRNA, oligo(dT)-primed reverse transcription, followed by blunt-ended cloning or dC-tailing and cloning into dG-tailored pBR322 (Imai *et al.*, 1983; Cashdollar *et al.*, 1984), or directly cloned into the TA vector pCR2.1 (Zhang & Rowhani, 2000), which is considered technically complicated, inefficient and also limited by the starting material (Potgieter *et al.*, 2002). Other methods have been improved on the basis of a single-primer dsRNA cloning sequence-independent method (Lambden *et al.*, 1992). The method was initially used to clone small fragments and was improved further by extending the primer that ligates to dsRNA with a poly(A) tail to facilitate oligo(dT) priming of poly(dA)-oligonucleotide-ligated dsRNA (Vreede *et al.*, 1998). However, the effectiveness of the improved method

Fig. 2. Phylogenetic analysis of five non-structural proteins of J19. Bar, the number of amino acid substitutions ($\times 100$). Group A rotavirus representative strains: SA11-both, NR1, TB-Chen, KU, SA11, PO-13, WA, RMC321, DS1 and D; group B rotavirus representative strains: IDIR, Bang373, CAL-1, WH1, ADRV and KB63.

is also limited by the amount of starting material. Recently, it was shown that the use of methylmercuric hydroxide for denaturing dsRNA instead of the commonly used combination of heat and DMSO also improved the method. The improved method was used to amplify complete genomes of six dsRNA viruses using a series of one-tube reactions with a small amount of starting material (Potgieter *et al.*, 2002; Maan *et al.*, 2007). However, results did not report obtaining the complete genome sequences of those viruses.

Here, is the first study to put forward the concept of inhibitor priming. Because pairs of inhibitory primers originate from small gene fragments, small gene fragments could be amplified and cloned by the single-primer amplification sequence-independent dsRNA technique described by Lambden *et al.* (1992). PCR amplification of genome segments larger than 2 kb was improved by the use of two pairs of inhibitory primers and the product could

then be cloned and sequenced. Based on the efficiency of amplifying and cloning large genome segments of J19, it is proposed that the improved method would be effective in amplifying and cloning large genome segments of other rotaviruses.

The largest reported rotavirus epidemic affected over one million people in China in 1980s and was caused by ADRV. As a result of the outbreak, significant effort has been directed towards discovering ADRV's origin. The percentage of amino acid identity between the VP6 equivalent proteins of various groups A, B and C rotaviruses and different group A subgroups suggested that the group B (ADRV) VP6 protein has diverged from groups A and C proteins in the distant past, most probably from a common ancestral VP6 precursor protein (Chen *et al.*, 1991). It was presumed that there must be an ancestral strain for group B rotaviruses in China or other areas. This novel ADRV has

since caused several outbreaks of adult diarrhoea in China (Hung *et al.*, 1983, 1984; Chen *et al.*, 1985), while sporadic cases caused by two group B rotavirus strains, CAL-1 and Bang373, were detected in India and Bangladesh (Kobayashi *et al.*, 2001; Ahmed *et al.*, 2004), respectively. CAL-1 and Bang373 were found to be genetically related and showed high nucleotide sequence identities (91–92 %) to ADRV, suggesting that these two human group B rotaviruses may have the same origin (Kobayashi *et al.*, 2001; Ahmed *et al.*, 2004). Protein sequence analyses and phylogenetic analyses of NSP2 from group B rotavirus strain CAL-1 showed that the strain may be the progenitor of ADRV (Sen *et al.*, 2001). However, repeated sequencing of the ADRV NSP2 and NSP3 genes combined with phylogenetic analyses indicated that the ADRV and CAL strains evolved from a common ancestor (Jiang *et al.*, 2005). Another human group B rotavirus strain WH1 was detected in an adult sporadic case of diarrhoea in Wuhan city, China, 20 years after the identification of ADRV. Although WH1 was found to be genetically close to ADRV, some WH1 protein genes had notable characteristics that are not found in ADRV, but are commonly observed in CAL-1 and Bang373 (Yang *et al.*, 2004b).

Analyses of the mutation rate of human group B rotaviruses indicated that the Indian–Bangladeshi lineage must have diverged from the Chinese lineage several decades ago (Yang *et al.*, 2004b). In this study, phylogenetic analyses of four structural proteins (VP2, VP4, VP6 and VP7) and five non-structural proteins (NSP1–NSP5) also suggest that the Indian–Bangladeshi lineage and the Chinese lineage diverged from the same progenitor during the evolution of group B rotaviruses. We examined 11 rotavirus proteins and showed that the strain J19 could be one prototype of group B rotaviruses, including ADRV, CAL-1 and Bang373, as well as others. Therefore, we offer a plausible explanation that China may be an epicentre for human group B rotaviruses, in contrast with earlier research (Sen *et al.*, 2001). Our hypothesis could explain the sudden appearance of ADRV or NADRV that caused outbreaks of adult diarrhoea in China.

Phylogenetic analyses suggested that human group B rotaviruses in Asia originated from NADRV J19 and B219; however, we also found that these two strains cluster close to IDIR, a group B rotavirus causing diarrhoea in infant rats and later discovered in humans in the USA (Eiden *et al.*, 1985). IDIR differs significantly from any other human group B rotaviruses found in China and India, such as ADRV and CAL-1. Therefore, group B adult diarrhoea rotavirus strains, such as ADRV, WH1, CAL-1 and Bang373 may have evolved from other ancestral strains in Asia that have not yet been isolated or identified. Recently, an unusual human rotavirus strain B219 was detected in Bangladesh. Cloning and sequence analysis of B219 strain indicated that this virus is genetically closely related to the NADRV strain (J19), which suggests that the NADRV (J19)-like novel human rotaviruses would be distributed to a geographically wider area (Alam *et al.*, 2007). Further extensive surveillance of rotaviruses of

humans and animals in China will help to understand better the origin and evolution of human group B rotavirus. Now, the availability of the full genome sequence of J19, in conjunction with the fact that the J19 strain can apparently be propagated in cell culture (Ji *et al.*, 2002), will allow a more detailed functional analysis of the viral replication cycle.

ACKNOWLEDGEMENTS

We are grateful to Baocheng Shi and Runsheng Chen at Bioinformatics Laboratory, Institute of Biophysics, Chinese Academy of Sciences, for their expert technical assistance in phylogenetic analyses. This study was supported by grants from the National Natural Science Foundation of China (no. 30371278) and the Beijing Natural Science Foundation (no. 5042020).

REFERENCES

- Adah, M. I., Wade, A., Oset, M., Kuzuya, M. & Taniguchi, K. (2002). Detection of human group C rotaviruses in Nigeria and sequence analyses of their genes encoding VP4, VP6 and VP7 proteins. *J Med Virol* **66**, 269–275.
- Ahmed, M. U., Kobayashi, N., Wakuda, M., Sanekata, T., Taniguchi, K., Kader, A., Naik, T. N., Ishino, M., Alam, M. M. & other authors (2004). Genetic analysis of group B human rotaviruses detected in Bangladesh in 2000 and 2001. *J Med Virol* **72**, 149–155.
- Alam, M. M., Kobayashi, N., Ishino, M., Ahmed, M. S., Ahmed, M. U., Paul, S. K., Muzumdar, B. K., Hussain, Z., Wang, Y. H. & Naik, T. N. (2007). Genetic analysis of an ADRV-N like novel rotavirus strain B219 detected in a sporadic case of adult diarrhea in Bangladesh. *Arch Virol* **152**, 199–208.
- Ando, T., Monroe, S. S., Gentsch, J. R., Jin, Q., Lewis, D. C. & Glass, R. I. (1995). Detection and differentiation of antigenically distinct small round structured viruses (Norwalk-like viruses) by reverse transcription-PCR and Southern hybridization. *J Clin Microbiol* **33**, 64–71.
- Both, G. W., Bellamy, A. R. & Siegman, L. J. (1984). Nucleotide sequence of the dsRNA genomic segment 7 of simian 11 rotavirus. *Nucleic Acids Res* **12**, 1621–1626.
- Brodsky, L. I., Vasiliev, A. V., Kalaidzidis, Ya. L., Osipov, Yu. S., Tatuzov, R. L. & Feranchuk, S. I. (1992). GeneBee: the program package for biopolymer structure analysis. *Dimacs* **8**, 127–139.
- Brodsky, L. I., Ivanov, V. V., Kalaidzidis, Ya. L., Leontovich, A. M., Nikolaev, V. K., Feranchuk, S. I. & Drachev, V. A. (1995). GeneBee-NET: internet-based server for analyzing biopolymers structure. *Biochemistry* **60**, 923–928.
- Cashdollar, L. W., Chemelo, R., Esparza, J., Hudson, G. R. & Joklik, W. K. (1984). Molecular cloning of the complete genome of reovirus serotype 3. *Virology* **133**, 191–196.
- Caul, E. O., Ashley, C. R., Darville, J. M. & Bridger, J. C. (1990). Group C rotavirus associated with fatal enteritis in a family outbreak. *J Med Virol* **30**, 201–205.
- Chen, C. M., Huang, T., Bridger, J. C. & McCrae, M. A. (1985). Chinese adult rotavirus is a group B rotavirus. *Lancet* **2**, 1123–1124.
- Chen, G. M., Hung, T. & Mackow, E. R. (1990). cDNA cloning of each genomic segment of the group B rotavirus ADRV: molecular characterization of the 11th RNA segment. *Virology* **175**, 605–609.
- Chen, G. M., Werner-Eckert, R., Hung, T. & Machow, E. R. (1991). Expression of the major inner capsid protein of the group B rotavirus ADRV: primary characterization of genome segment 5. *Virology* **182**, 820–829.

- Chen, Z., Lambden, P. R., Lau, J., Caul, E. O. & Clarke, L. N. (2002). Human group C rotavirus: completion of the genome sequence and gene coding assignments of a non-cultivable rotavirus. *Virus Res* **83**, 179–187.
- Eiden, J., Vonderfecht, S. & Yolken, R. H. (1985). Evidence that a novel rotavirus-like agent of rats can cause gastroenteritis in man. *Lancet* **2**, 8–11.
- Estes, M. (2001). Rotaviruses and their replication. In *Fields Virology*, 4th edn, pp. 1747–1786. Edited by D. M. Knipe & P. M. Howley. Philadelphia, PA: Lippincott Williams & Wilkins.
- Estes, M. K. & Cohen, J. (1989). Rotavirus gene structure and function. *Microbiol Rev* **53**, 410–449.
- Estes, M. K., Mason, B. B., Crawford, S. & Cohen, J. (1984). Cloning and nucleotide sequence of the simian rotavirus gene 6 that codes for the major inner capsid protein. *Nucleic Acids Res* **12**, 1875–1887.
- Gouvea, V., Glass, R. I., Woods, P., Taniguchi, K., Clark, H. F., Forrester, B. & Fang, Z. Y. (1990). Polymerase chain reaction amplification and typing of rotavirus nucleic acid from stool specimens. *J Clin Microbiol* **28**, 276–282.
- Hung, T., Chen, G. M., Wang, C. G., Chou, Z. Y., Chao, T. X., Ye, W. W., Yao, H. L. & Meng, K. H. (1983). Rotavirus-like agent in adult nonbacterial diarrhea in China. *Lancet* **2**, 1078–1079.
- Hung, T., Chen, G. M., Wang, C. G., Yao, H. L., Fang, Z. Y., Chao, T. X., Chou, Z. Y., Ye, W. & Chang, X. J. (1984). Waterborne outbreak of rotavirus diarrhea in adults in China caused by a novel rotavirus. *Lancet* **1**, 1139–1142.
- Imai, M., Richardson, M. A., Ikegami, N., Shaiki, A. J. & Furuichi, Y. (1983). Molecular cloning of double-stranded RNA viruses genomes. *Proc Natl Acad Sci U S A* **80**, 373–377.
- Ji, S., Bi, Y., Yang, H. Y., Yang, F. R., Song, J. Q., Tao, X. X. & Cui, X. Y. (2002). Cultivation and serial propagation of a new rotavirus causing adult diarrhea in primary human embryo kidney cells. *Zhonghua Yi Xue Za Zhi* **82**, 14–18. (in Chinese)
- Jiang, B., Dennehy, P. H., Spangenberg, S., Gentsch, J. R. & Glass, R. L. (1995). First detection of group C rotavirus in fecal specimens of children with diarrhea in the United States. *J Infect Dis* **172**, 45–50.
- Jiang, B., Wang, Y., Glass, R. I. & Fang, Z. Y. (2005). The evolution of human group B rotaviruses: correction and an update. *J Clin Virol* **34**, 158–159.
- Jiang, S. J., Ji, S. Z., Tang, Q., Yang, H. Y., Cui, X. Y., Bi, Y., Kan, B. & Gao, S. Y. (2006). An improved method to amplify and clone large genome segments of a novel adult diarrhea rotavirus strain J19. *Chinese Journal of Microbiology and Immunology* **26**, 781–785.
- Kobayashi, N., Naik, T. N., Kushhara, Y., Krishnan, T., Sen, A., Bhattacharya, S. K., Taniguchi, K., Alam, M. M., Urassawa, T. & Urasawa, S. (2001). Sequence analysis of genes encoding structural and nonstructural proteins of a human group B rotavirus detected in Calcutta. *J Med Virol* **64**, 583–588.
- Krishnan, T., Sen, A., Choudhury, J. S., Das, S., Naik, T. N. & Bhattacharya, S. K. (1999). Emergence of adult diarrhea rotavirus in Calcutta, India. *Lancet* **353**, 380–381.
- Kuzuya, M., Fuji, R., Hamano, M., Nakamura, J., Yamada, M., Nii, S. & Mori, T. (1996). Molecular analysis of outer capsid glycoprotein (VP7) genes from two isolates of human group C rotaviruses with different genome electrophoretotypes. *J Clin Microbiol* **34**, 3185–3189.
- Lambden, P. R., Cooke, S. J., Caul, E. O. & Clarke, I. N. (1992). Cloning of noncultivable human rotaviruses by single primer amplification. *J Virol* **66**, 1817–1822.
- Maan, S., Rao, S., Maan, N. S., Anthony, S. J., Attoui, H., Samuel, A. R. & Mertens, P. P. C. (2007). Rapid cDNA synthesis and sequencing techniques for the genetic study of bluetongue and other dsRNA viruses. *J Virol Methods* **143**, 132–139.
- Mitchell, D. B. & Both, G. W. (1990). Completion of the genomic sequence of the simian rotavirus SA11: nucleotide sequences of segments 1, 2, and 3. *Virology* **177**, 324–331.
- Nilsson, M., Svenungsson, B., Hedlund, K. O., Uhnoo, I., Lagergren, A., Akre, T. & Svensson, L. (2000). Incidence and genetic diversity of group C rotavirus among adults. *J Infect Dis* **182**, 678–684.
- Potgieter, A. C., Steele, A. D. & Van Dijk, A. A. (2002). Cloning of complete genome sets of six dsRNA viruses using an improved cloning method for large dsRNA genes. *J Gen Virol* **83**, 2215–2223.
- Ramig, R. F., Ciarlet, M., Mertens, P. P. C. & Dermody, T. S. (2005). Rotavirus. In *Virus Taxonomy, Eighth Report of the ICTV*, pp. 484–496. Edited by C. M. Fauquet, M. A. Mayo, J. Maniloff, U. Desselberger & L. A. Ball. Amsterdam: Elsevier Academic Press.
- Rodger, S. M., Bishop, R. F. & Holmes, I. H. (1982). Detection of a rotavirus-like agent associated with diarrhea in an infant. *J Clin Microbiol* **16**, 724–726.
- Saif, L. J. & Jiang, B. (1994). Nongroup A rotaviruses of humans and animals. *Curr Top Microbiol Immunol* **185**, 339–371.
- Sanekata, T., Ahmed, M. U., Kaderet, A., Taniguchi, K. & Kobayashi, N. (2003). Human group B rotavirus infectious cause severe diarrhea in children and adults in Bangladesh. *J Clin Microbiol* **41**, 2187–2190.
- Schnagl, R. D., Boniface, K., Cardwell, P., McCarthy, D., Ondracek, C., Coulson, B., Erlich, J. & Morey, F. (2004). Incidence of group C human rotavirus in central Australia and sequence variation of the VP7 and VP4 genes. *J Clin Microbiol* **42**, 2127–2133.
- Sen, A., Kobayashi, N., Das, S., Krishnan, T., Bhattacharya, S. K. & Naik, T. N. (2001). The evolution of human group B rotaviruses. *Lancet* **357**, 198–199.
- Vreede, F. T., Clote, M., Napier, G. B., Vand Dijk, A. A. & Viljoen, G. J. (1998). Sequence-independent amplification and cloning of large dsRNA virus genome segments by poly(dA)-oligonucleotide ligation. *J Virol Methods* **72**, 243–247.
- Wang, C., Hu, C., Hung, F., Chen, X. & Huang, T. (1987). A novel discovered rotavirus from adult acute diarrhea particles in China. *Chinese Journal of Virology* **3**, 321–325.
- Yang, H., Chen, S., Ji, S. Z., Meng, Z., Kang, B., Zhang, Z. & Qi, S. (1998). A novel rotavirus causing a large scale of adult diarrhea in Shi Jiazhuang. *Zhonghua Liu Xing Bing Xue Za Zhi* **19**, 336–338. (in Chinese)
- Yang, H., Makeyev, E. V., Kang, Z., Ji, S., Bamford, D. H. & van Dijk, A. A. (2004a). Cloning and sequence analysis of dsRNA segments 5, 6 and 7 of a novel non-group A, B, C adult rotavirus that caused an outbreak of gastroenteritis in China. *Virus Res* **106**, 15–26.
- Yang, J. H., Kobayashi, N., Wang, Y. H., Zhou, X., Li, Y., Zhou, D. J., Hu, Z. H., Ishino, M., Alam, M. M. & other authors (2004b). Phylogenetic analysis of a human group B rotavirus WH-1 detected in China in 2002. *J Med Virol* **74**, 662–667.
- Zhang, Y. P. & Rowhani, A. (2000). A strategy for rapid cDNA cloning from double-stranded RNA templates isolated from plants infected with RNA viruses by using Taq DNA polymerase. *J Virol Methods* **84**, 59–63.