Journal of General Virology (1999), **80**, 3181–3187. Printed in Great Britain

Molecular characterization of human group C rotavirus genes 6, 7 and 9

Vivienne L. A. James,[1] Paul R. Lambden,[2] Yu Deng,[2] E. Owen Caul[3] and Ian N. Clarke[2]

[1,2] Public Health Laboratory[1] and Department of Molecular Microbiology, University Medical School[2], Southampton General Hospital, Tremona Road, Southampton SO16 6YD, UK

[3] Regional Virus Laboratory, Public Health Laboratory, Myrtle Road, Bristol BS2 8EL, UK

Genes 6, 7 and 9 of human group C rotavirus 'Bristol' strain, encoding non-structural proteins (NSP) 3, 1 and 2, respectively, were cloned and sequenced. Human group C rotavirus genome segment 6 is 1350 bp and contains a single ORF of 1231 nucleotides (encoding 402 amino acids). Genome segment 7 is 1270 bp and encodes a protein of 394 amino acids and genome segment 9 is 1037 bp and encodes a 312 amino acid protein. The human group C rotavirus genes 6, 7 and 9 showed 78, 67 and 88% sequence identity, respectively, to the corresponding porcine group C rotavirus genes. The derived protein sequences were compared with those of the porcine 'Cowden' group C and mammalian group A rotavirus strains. The human group C rotavirus NSP1 protein sequence is one amino acid longer than the porcine group C equivalent. In common with group A and porcine group C rotaviruses, the human group C rotavirus NSP1 protein has a zinc finger motif. Human group C rotavirus NSP2 has two hydrophobic heptad repeat regions, a basic, RNA-binding domain and a basic, proline-rich region. Human group C rotavirus NSP3 has both single- and double-stranded RNA-binding domains and several hydrophobic heptad repeat regions, one of which forms a leucine zipper. This work completes the molecular characterization of the non-structural proteins of a human group C rotavirus. Phylogenetic analysis of all the non-structural genes of group A, B and C rotaviruses suggests that these viruses have diverged at a constant rate from a common ancestor.

Introduction

Rotaviruses are a genus of the family *Reoviridae* and are an important cause of gastroenteritis in humans and animals. Seven distinct rotavirus groups (A–G) have been described (Kapikian & Chanock, 1996), but only groups A, B and C have been identified in both humans and animals. Group A rotaviruses were recognized as human pathogens in 1973 (Bishop *et al.*, 1973; Flewett *et al.*, 1973). Group A rotaviruses are endemic and are the major cause of severe gastroenteritis in infants under 2 years of age. Human group B rotaviruses were originally described in 1983 (Hung *et al.*, 1983) and have been responsible for several large outbreaks of severe gastroenteritis in adults in China. Although serological studies have suggested that infection with group B rotaviruses has occurred in other countries (Brown *et al.*, 1987; Hung *et al.*, 1987; Nakata *et al.*, 1987), the virus was first reported outside China in 1999 (Krishnan *et al.*, 1999). In this Indian study, group B rotaviruses were confirmed as the cause of five cases of severe adult diarrhoea in Calcutta during 1997–98. Group C rotavirus (Cowden strain) was first described in 1980 in pigs (Saif *et al.*, 1980) and, in 1982, the first human group C rotavirus was described (Rodger *et al.*, 1982). Sporadic cases and limited outbreaks, affecting both adults and children, have since been reported worldwide. Although serological evidence existed previously for group C rotavirus infections in cattle, the bovine 'Shintoku' strain was only isolated in 1991 (Tsunemitsu *et al.*, 1991).

The rotavirus genome consists of 11 segments of double-stranded (ds) RNA that encode six structural and five non-structural proteins. The different rotavirus groups can be identified by the characteristic profile of the dsRNA on SDS–PAGE and defined by terminal fingerprint analysis of the

Author for correspondence: Vivienne James. Present address: Central Public Health Laboratory, 61 Colindale Avenue, London NW9 5HT, UK. Fax +44 20 8205 1488. e-mail vjames@phls.nhs.uk

The EMBL accession numbers for the sequences of the human group C rotavirus 'Bristol' genes 6, 7 and 9 are AJ132203–AJ132205.

genome segments (Pedley *et al.*, 1986). Sequence data are available for all eleven genome segments of the porcine group C rotavirus 'Cowden' strain, although many of these are incomplete and do not have defined 5′ and 3′ termini. Sequences are also available for genome segments 3–6, 8 and 10 of the bovine group C rotavirus 'Shintoku' strain. Genes corresponding to the structural proteins VP1 and VP2 and the non-structural proteins NSP1, 2 and 3 of human group C rotavirus have not been determined. The purpose of this work was to determine the coding assignments and sequences of the non-structural proteins NSP1, 2 and 3 of human group C rotavirus and to compare them with the corresponding porcine group C and mammalian group A rotavirus proteins.

Methods

■ **RNA extraction, cDNA synthesis and coding assignment.** Human group C rotavirus 'Bristol' strain has been described previously (Caul *et al.*, 1990). Rotavirus dsRNA was extracted and purified from a faecal sample using the RNAzolB and Geneclean II procedures as described previously (Lambden & Clarke, 1995). Random-primed cDNA was synthesized and purified by Sephacryl S-400 spin-column chromatography as described previously (James *et al.*, 1998).

M13 clones for each of the 11 gene segments of human group C rotavirus 'Bristol' strain have been constructed and stored as an ordered genomic library (Lambden *et al.*, 1992). Briefly, the cloning method involved ligation of a single amino-linked modified oligonucleotide to the 3′ termini of each dsRNA genome segment by using T_4 RNA ligase. The tailed RNA was then converted to cDNA by using a complementary primer and reverse transcriptase. The resultant cDNA was annealed and repaired and then amplified by PCR by using a single complementary oligonucleotide primer. The PCR products were ligated into dephosphorylated, *Sma*I-cleaved M13 mp8 and transformed into *E. coli* JM101. This method increased the probability that the resultant clones would contain full-length inserts of the genes.

^{32}P-labelled cDNA probes used for gene assignment studies were generated as described previously (Deng *et al.*, 1995) and the coding assignments were determined by Northern blot analysis.

■ **Sequence analysis.** Sequences were compiled from M13 clones and by direct sequencing of RT–PCR amplicons. Recombinant M13 templates were prepared by standard techniques and sequenced initially with universal primer designed to hybridize 49 bp upstream of the *Sma*I cloning site by using the ABI PRISM terminator cycle sequencing kit (Applied Biosystems) and an Applied Biosystems model 373A automated sequencer. The complete sequences were determined by preparing primers and sequencing stepwise along the gene from each previous sequence. Internal primer sequences were designed approximately 150 bp upstream of the previously deduced sequence and resulted in read-lengths of 400–500 bp. The oligonucleotides were synthesized on a Millipore Expedite 8909 automated synthesizer by using *β*-cyanoethyl phosphoramidite chemistry. Computer analyses of the sequence data were performed by using the Lasergene software (DNASTAR).

Results and Discussion

Analysis of the M13 recombinants confirmed that they were full-length clones containing terminal sequences matching the group C rotavirus consensus 5′ and 3′ termini, GGC(A/T)$_4$AA and TGTGGCT, respectively (Lambden & Clarke, 1995). More than 92% of the gene sequences were determined by direct sequencing of the RT–PCR amplicons and included the sequences of the entire ORFs. Northern blot analysis of group C rotavirus dsRNA with probes representing the full-length NSP1, NSP2 and NSP3 inserts confirmed the coding assignments of the clones for genes 6, 7 and 9 (data not shown).

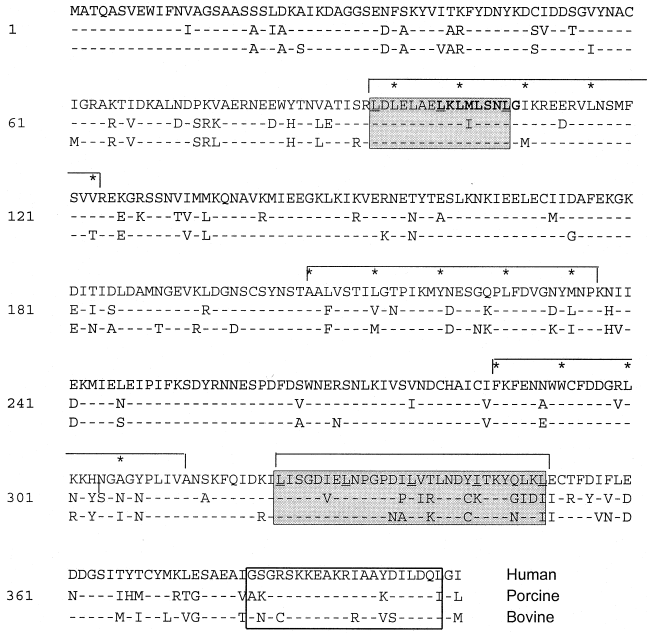

Fig. 1. Group C rotavirus NSP3 protein alignment. NSP3 protein sequences derived from the nucleotide sequences from human group C rotavirus, the porcine 'Cowden' strain and bovine 'Shintoku' strains were aligned using the CLUSTAL W program (Thompson *et al.*, 1994). The sequences are numbered from the first amino acid residue. Identical amino acids are represented by dashes. The following functional domains are indicated: the ssRNA-binding consensus sequence (bold type), hydrophobic heptad repeats (lines above the sequence: asterisks indicate each seventh hydrophobic amino acid), the leucine zipper domains (shaded boxes with leucine/isoleucine residues underlined) and the dsRNA-binding domain (open box).

Gene segment 6

Genome segment 6, encoding the NSP3 (NS34) protein, is 1350 nucleotides long, the same length as the cognate bovine 'Shintoku' gene and two nucleotides longer than the genome sequence for the porcine 'Cowden' strain. The two extra nucleotides are at positions 1250 and 1251 in the human and bovine group C sequences in the 3′ non-coding region. The NSP3 genome segment showed 78% identity to the porcine 'Cowden' equivalent and 79% identity to the bovine 'Shintoku' NSP3 gene. Similarly, the 'Cowden' and 'Shintoku' NSP3 genes share 78% sequence identity. The ATG translation initiation codon is at position 25, in a favourable context for

translation initiation (Kozak, 1991). Computer analysis of the sequence revealed a single ORF of 1206 nucleotides and a 3′ non-coding region of 120 nucleotides.

The amino acid sequence predicted a protein 402 amino acids long, acidic in nature (Fig. 1), with a predicted molecular mass of 45·3 kDa and a calculated isoelectric point of 4·85. Eight cysteine residues were conserved between the human, porcine and bovine NSP3 proteins at positions 60, 171, 200, 280, 284, 294, 353 and 369. Three major functional domains described for the group A rotavirus SA11F NSP3 protein (Mattion *et al*., 1992) were also present on the group C NSP3 protein. The regions are a basic region for single-stranded (ss) RNA binding, a heptad repeat region for oligomerization and a leucine zipper motif. The group A rotavirus NSP3 protein basic region extends from amino acids 81 to 150; this region is also charged in the group C rotaviruses, with a large number of lysine and arginine residues between amino acids 79 and 151. The region between amino acids 89 and 125 is conserved in the group C rotaviruses, with three amino acid changes between human and bovine strains and just two amino acid changes between the human and porcine strains. The consensus sequence for ssRNA-binding, (I/L)XXM(I/L)(S/T)XXG, seen in orbiviruses, reoviruses and group A rotaviruses (Rao *et al*., 1995) is found in this region and is located at amino acids 104–112 of NSP3 of SA11. In the human and bovine group C rotaviruses, the sequence ^{100}LXXMLSXXG^{108} is conserved and is highlighted in Fig. 1. This is at variance with the 'Cowden' sequence, where the methionine residue is replaced by isoleucine (Qian *et al*., 1991).

The group C rotaviruses also appear to have several regions of hydrophobic heptad repeats necessary for oligomerization of the NSP3 protein. The 'Cowden' strain has a leucine-rich region between amino acids 87 and 116, a short leucine zipper between amino acids 93 and 107 and another five hydrophobic heptad repeats between amino acids 95 and 123 (Mattion *et al*., 1992). In the human group C rotavirus, leucine at amino acid 87 is replaced by valine. Further heptad repeats are seen at amino acids 206–236, 286–312 and 323–351. These regions contain the amino acids proline and glycine, which would suggest the absence of an α-helical coil, although the heptad repeats might be sufficient to promote helix formation (Cohen & Parry, 1986). The last heptad repeat conforms to a leucine zipper of five repeats. Following the basic region, there is a short, highly acidic region extending from amino acids 153 to 171, which mirrors the structure found in the cognate group A rotavirus gene product (amino acids 151–169).

Besides ssRNA binding, the NSP3 proteins of the group C rotaviruses have been shown to bind dsRNA (Langland *et al*., 1994). The group C rotavirus NSP3 proteins are 87–89 amino acids longer than their group A counterparts. The carboxy-terminal region of the group C NSP3 protein contains the dsRNA-binding domain, with a consensus sequence LX_{39}(G/A)XGXSKKXAKXXAAXX(A/I)LXXL (Langland *et al*., 1994). This consensus region is also present in human group C rotavirus (amino acids 340–400).

```
1     MANSFREMLYWYRKIIDRKLPCVNVNIWRREIAYKANGICLNCLNECKVCPCDYCGIRHK
      ----Y------FG-T---N--Y--T-G--KQKGR-D*-------D---LYS--H---K--

61    CENCLNSDCFMNTNNEFNSHRWITFDEEPSQMVLFEYWIMYKDYFLSKFNYNYKAQLKIL
      -G--VL-E--LDVK----KY--LV-----D-A--LQH---------Q----RLAT-A---

121   NMNKNRRFHINESKKKALSVPITSQYLKFKFNNKIYIMFGTFLTSKIQPWIQLKSLKVGY
      -----QK-QL--GR-R---------F---RLFG----Q---IM-N--Q--LE-ST--I--

181   IQSLNVDRCAKLIATKGMFATNSFKSSCITEINARRPISECDYLIEACLCNENNEWKFSA
      L-L---E--SE-M--R-Q-T--VA-TA-----KC----YDN-CI---Y-DKNDRG---A-

241   VMGRDKIPVTKSLAMKYFCKNINTELFYYGHSKCHVVSECPRWNQQLRVLNASTLNIIFR
      IL-RR-----QK---E--M-SLRA-----A--R--TL-N-----EG--L--S-----V--

301   RQFMNEVVEWFENFTQLTGMHYDFIKTCVYNKVIISHFRKEIEDYINSGKKISLSSVIPD
      ------I-----I-SQY--S--E--TE--HD-SA-TA-KQ------KE-KQ-T-K--V-E

361   GHALYTNIDILRISLMLAIDVALNRIESQQMDVL      Human
      E--A-RH-LR--E-------A--S--R--S-G--      Porcine
```

Fig. 2. Group C rotavirus NSP1 protein sequence alignment. NSP1 protein sequences derived from the nucleotide sequences from human group C rotavirus and the porcine 'Cowden' strain were aligned using CLUSTAL W. The sequences are numbered from the first amino acid residue. Identical amino acids are represented by dashes. The zinc finger motifs are underlined. *, Gap introduced to maximize the alignment.

Gene segment 7

The non-structural protein NSP1 (NS53) is encoded by gene 7 in group C rotaviruses. The gene sequence is 1270 nucleotides long. Computer analysis revealed a single ORF of 1182 nucleotides with a 5′ non-coding region of 37 bp and a 3′ non-coding region of 48 bp. The nucleotide sequence GTCAACATGG in the region of the putative ATG codon (underlined), with A at −3 and G at +4, constitutes a strong signal for initiation of translation (Kozak, 1991). The human group C rotavirus gene 7 sequence is 35 nucleotides longer than the porcine equivalent (Bremont *et al*., 1993), with an extra 31 nucleotides at the 5′ terminus, one extra base at the 3′ terminus and a novel extra codon at positions 135–137. The sequence has 67% identity to the equivalent 'Cowden' sequence.

The predicted protein sequence contains 394 amino acid residues, with a molecular mass of 46·6 kDa (Fig. 2). It is a basic protein with a calculated isoelectric point of 8·9. Human group C rotavirus NSP1 has 61% amino acid sequence identity to the equivalent 'Cowden' sequence, which contains 393 amino acids. This is consistent with other reports that suggest that NSP1 is the most variable rotavirus protein and that diversity in sequences is evident between strains from different species (Hua *et al*., 1993; Dunn *et al*., 1994). Studies examining reassortants between rotavirus strains have implicated NSP1 in host-range selection (Graham *et al*., 1987; Palombo & Bishop, 1994) and pathogenic phenotype (Broome *et al*., 1993).

The group C rotavirus NSP1 proteins (393–394 amino acids) are much smaller than the corresponding group A rotavirus NSP1 proteins, which are 486–495 amino acids in length. No function has been found for the C-terminal 300 amino acids of the group A rotaviruses (Hua *et al*., 1994).

Group A rotavirus mutants lacking either the central or carboxy half of the protein were viable and therefore this region is not required for virus replication *in vitro* (Tian *et al.*, 1993; Hua & Patton, 1994). The functional regions of the group A rotaviruses, demonstrated by deletion mapping (Hua *et al.*, 1994), are the RNA-binding domain and the intracellular localization domain. The RNA-binding domain is cysteine rich and forms a zinc finger (Bremont *et al.*, 1987; Mitchell & Both, 1990; Hua *et al.*, 1993, 1994; Xu *et al.*, 1994). Analysis of the human group C NSP1 amino acid sequence shows the zinc finger motif, $CX_2CX_3CX_4CX_2CX_3HXCX_2CX_4C$, between amino acids 40 and 69. An identical motif is also found in the 'Cowden' strain NSP1. Both the human and porcine group C sequences have only four residues before the final cysteine of the motif at position 69, compared with the group A NSP1 proteins, which have five amino acids at this position. This region is not required for growth in cell culture, as viable mutants lacking the zinc finger motif coding region have been isolated (Taniguchi *et al.*, 1996*a*). Additionally, low frequency reassortment occurs between strains lacking the cysteine-rich region and heterologous virus: thus, this region is not essential for genome segment reassortment, although the intact NSP1 enhances the efficiency of genome segment assortment (Okada *et al.*, 1999). A second motif of $HX_2(C/H)X_6CX_2C$ has been noted in a few group A rotavirus strains (Bremont *et al.*, 1987; Taniguchi *et al.*, 1996*b*). A corresponding sequence, HX_2CHX_4C, at amino acids 271–280 is conserved between the NSP1 of human and porcine group C rotaviruses.

Gene segment 9

Gene segment 9 encodes the non-structural protein NSP2 (NS35). The nucleotide sequence is 1037 bp long and shares 88% identity (between nucleotides 21 and 1015) to the gene 9 sequence of the porcine 'Cowden' strain (Bremont *et al.*, 1993). The sequence contains a single ORF of 936 nucleotides with a 5′ untranslated leader sequence of 42 bp and an initiator codon in a favourable context for translation initiation (Kozak, 1991). The predicted protein sequence of 312 amino acids has a predicted molecular mass of 35·9 kDa and has 93·9% identity to the 'Cowden' NSP2 protein (Fig. 3).

The NSP2 proteins have been shown to bind ssRNA in a sequence-independent manner, with binding confined to oligomerized states of NSP2 (Kattoura *et al.*, 1992). Patton *et al.* (1993) postulated that the conserved and highly basic region from amino acids 205 to 241 of group A rotaviruses was the RNA-binding domain. This region is also highly basic in the group C rotaviruses and most of the basic amino acids are conserved. Preceding this basic region are two short regions of hydrophobic heptad repeats (amino acids 124–138 and 157–171) that are conserved between all the sequences and could form the oligomerization domain.

The first seven amino acids are absolutely conserved between all group A and group C rotavirus NSP2 proteins. Only one cysteine residue is absolutely conserved between the group A and group C NSP2 proteins, at position 6. Truncated NSP2 loses its ability to bind RNA when more than four amino acids are absent from the termini (Kattoura *et al.*, 1992): therefore the conserved cysteine at position 6 could be involved in stabilizing the oligomerized NSP2 proteins.

Two other regions are well conserved between the group A and group C rotavirus NSP2 proteins. The first is a basic region (amino acids 47–60) with a run of three proline residues with the consensus sequence IXYGXAPPPXF(K/N/R)(K/N/R)R. Proline-rich domains have been identified in several transcription regulators that both activate and repress gene transcription (Hanna-Rose & Hansen, 1996). NSP2 localizes in the viroplasm (Petrie *et al.*, 1984) and NSP2 has been shown *in vivo* to form part of a viral enzyme complex with replicase activity (Aponte *et al.*, 1996). Proline-rich activation motifs are frequently associated with serine and threonine residues (Hanna-Rose & Hansen, 1996). NSP5, which also localizes in the viroplasm, interacts with NSP2 (Poncet *et al.*, 1997) and has a high serine and threonine content. The interaction between NSP2 and NSP5 could regulate rotavirus RNA synthesis.

Between amino acid residues 109 and 139, a second, conserved, basic region is located followed by the first of the hydrophobic heptad repeats with consensus sequence $(V/I)RHLENLX_2RX_5D(V/I)LX_5LX_5(L/M)I$. NSP2 has been shown to form a complex with VP1, the putative RNA polymerase, in infected cells (Kattoura *et al.*, 1994). Furthermore, NSP2 has an affinity for dsRNA (Kattoura *et al.*, 1992) and associates with all 11 dsRNA segments, although Aponte *et al.* (1996) showed that, in the replicase complexes, some regions of the RNAs were single stranded, supporting the theory of a regulatory role in RNA replication.

Phylogenetic analyses

Genomic reassortment during co-infection *in vivo* and *in vitro* with distinct strains of group A rotaviruses is well documented (Ramig, 1997). Evolution of the different rotavirus serogroups could have arisen because of reassortment events and divergent evolution. Alternatively, the different rotavirus serogroups may have arisen from a common ancestor, in which case all genes would display equal divergence. To test these hypotheses, unrooted phylogenetic trees were constructed for groups A, B and C rotavirus NSP1, 2, 3, 4 and 5 gene sequences. Multiple alignments were performed with CLUSTAL X (Thompson *et al.*, 1997) and unrooted trees were generated by using the neighbour-joining method (Saitou & Nei, 1987). Trees were subjected to a bootstrap analysis (Felsenstein, 1985) using 1000 data sets and output as a graphic representation by using DRAWTREE in the PHYLIP package (Felsenstein, 1993).

Non-structural proteins have been considered to be more useful than structural proteins in determining the ancestral relationships of different viruses (Clewley, 1998). Expression of the non-structural proteins is confined within the host cell

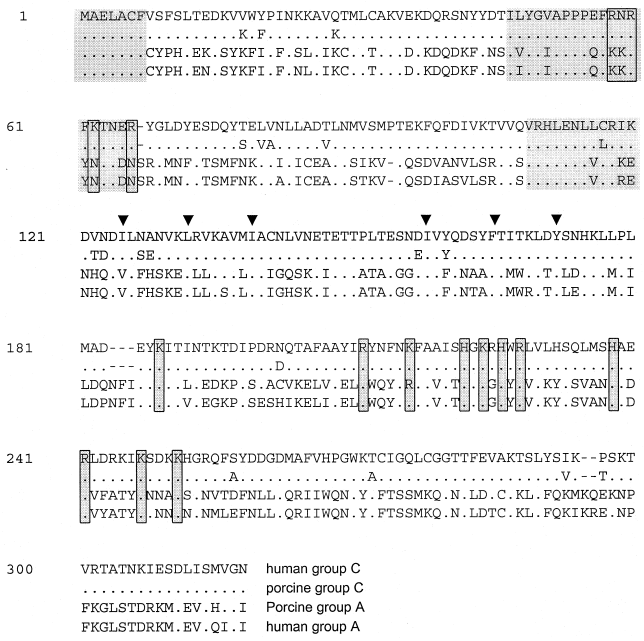

Fig. 3. Rotavirus NSP2 protein alignments. NSP3 protein sequences derived from the nucleotide sequences from human group C rotavirus, the porcine 'Cowden' strain and group A porcine (OSU) and human (DS1) strains were aligned using CLUSTAL W. The sequences are numbered from the first amino acid residue. Identical amino acids are represented by dots. The following features are indicated: basic amino acids (shaded and boxed), conserved regions (shaded) and hydrophobic heptad repeats (▼). -, Gap introduced to maximize the alignment.

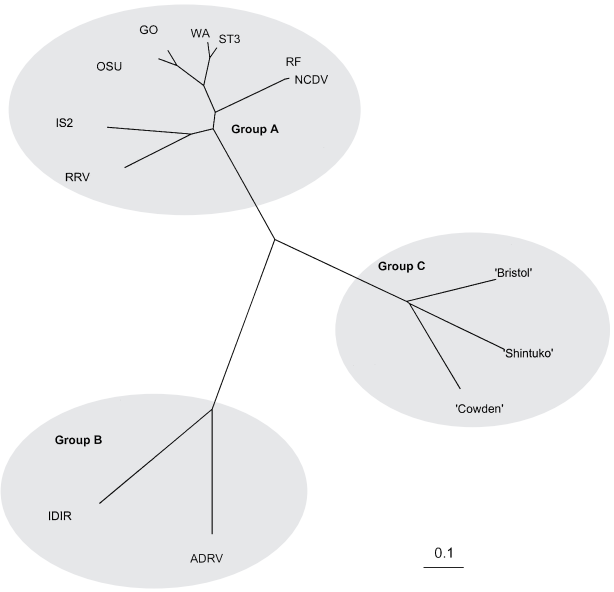

Fig. 4. Unrooted phylogenetic tree of the NSP3 gene (also representative of the NSP 1, 2, 4 and 5 genes) showing the relationship between the mammalian rotavirus groups. Trees were subjected to a bootstrap analysis using 1000 data sets. Shaded ellipses have been added to highlight the individual groups. Accession numbers (in parentheses) for rotaviruses are as follows: RRV (X81426), IS2 (X76645), OSU (X81431), GO (X81430), WA (X81434), ST3 (X81436), RF (Z21639), NCDV (X81429), IDIR (L09722), ADRV (M91436), 'Bristol' (AJ132203), 'Cowden' (M69115) and 'Shintoku' (L12390).

and non-structural proteins are not subject to the conformational structural constraints that affect the structural proteins, suggesting they would be subject to less evolutionary pressure than the structural proteins. However, Xu *et al.* (1994) demonstrated that divergence of gene 5 of group A rotaviruses encoding NSP1 exceeded that of the divergence of the major neutralizing antigen, VP7. Phylogenetic analysis of NSP3 proteins (Rao *et al.*, 1995) revealed that the rodent group B rotavirus (IDIR), porcine group C rotavirus (Cowden) and various mammalian group A rotavirus strains belonged to three different groups with a distinct ancestral origin. In the study presented here, the phylogenetic trees for each of the five genes encoding the non-structural proteins were essentially similar (for NSP3 comparison see Fig. 4), with three major branches, of similar length, for the three serogroups analysed. This branching pattern suggests that the groups may have diverged at the same time, possibly from a common ancestor, rather than as a result of reassortment of the genes. No reassortants have been detected between strains of different rotavirus groups (A–G) (Taniguchi & Urasawa, 1995) and divergence of the rotavirus serogroups is so great that reassortment between viruses from different serogroups does not appear to occur (Yolken *et al.*, 1988). Comparison of the group C branch of the phylogenetic trees suggests that the human, porcine and bovine group C rotaviruses also developed from a common ancestor at a later time. It may still be possible for reassortment to occur between group C rotaviruses from different animal species *in vitro*, although there is no evidence for naturally occurring reassortment (Grice *et al.*, 1994; Jiang *et al.*, 1996).

We thank Dr M. Pickett for the construction of the oligonucleotide primers. V.L.A.J. was funded for this work by the Public Health Laboratory Service.

References

Aponte, C., Poncet, D. & Cohen, J. (1996). Recovery and characterization of a replicase complex in rotavirus-infected cells by using a monoclonal antibody against NSP2. *Journal of Virology* **70**, 985–991.

Bishop, R. F., Davidson, G. P., Holmes, I. H. & Buck, B. J. (1973). Virus particles in epithelial cells of duodenal mucosa from children with acute non-bacterial gastroenteritis. *Lancet* ii, 1281–1283.

Bremont, M., Charpilienne, A., Chabanne, D. & Cohen, J. (1987). Nucleotide sequence and expression in *Escherichia coli* of the gene encoding the nonstructural protein NCVP2 of bovine rotavirus. *Virology* **161**, 138–144.

Bremont, M., Chabanne-Vautherot, D. & Cohen, J. (1993). Sequence analysis of three non structural proteins of a porcine group C (Cowden strain) rotavirus. *Archives of Virology* **130**, 85–92.

Broome, R. L., Vo, P. T., Ward, R. L., Clark, H. F. & Greenberg, H. B. (1993). Murine rotavirus genes encoding outer capsid proteins VP4 and VP7 are not major determinants of host range restriction and virulence. *Journal of Virology* **67**, 2448–2455.

Brown, D. W. G., Beards, G. M., Chen, G. M. & Flewett, T. H. (1987). Prevalence of antibody to group B (atypical) rotavirus in humans and animals. *Journal of Clinical Microbiology* **25**, 316–319.

Caul, E. O., Ashley, C. R., Darville, J. M. & Bridger, J. C. (1990). Group C rotavirus associated with fatal enteritis in a family outbreak. *Journal of Medical Virology* **30**, 201–205.

Clewley, J. P. (1998). A user's guide to producing and interpreting tree diagrams in taxonomy and phylogenetics. *Communicable Disease and Public Health* **1**, 64–66.

Cohen, C. & Parry, D. A. D. (1986). Alpha-helical coiled coils – a widespread motif in proteins. *Trends in Biochemical Sciences* **11**, 245–248.

Deng, Y., Fielding, P. A., Lambden, P. R., Caul, E. O. & Clarke, I. N. (1995). Molecular characterization of the 11th RNA segment from human group C rotavirus. *Virus Genes* **10**, 239–243.

Dunn, S. J., Cross, T. L. & Greenberg, H. B. (1994). Comparison of the rotavirus nonstructural protein NSP1 (NS53) from different species by sequence analysis and northern blot hydridization. *Virology* **203**, 178–183.

Felsenstein, J. (1985). Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**, 783–791.

Felsenstein, J. (1993). PHYLIP (phylogeny inference package), version 3.5c. Distributed by the author. Department of Genetics, University of Washington, Seattle, WA, USA.

Flewett, T. H., Bryden, A. S. & Davies, H. (1973). Virus particles in gastroenteritis. *Lancet* ii, 1497.

Graham, A., Kudesia, G., Allen, A. M. & Desselberger, U. (1987). Reassortment of human rotavirus possessing genome rearrangements with bovine rotavirus: evidence for host selection. *Journal of General Virology* **68**, 115–122.

Grice, A. S., Lambden, P. R., Caul, E. O. & Clarke, I. N. (1994). Sequence conservation of the major outer capsid glycoprotein of human group C rotaviruses. *Journal of Medical Virology* **44**, 166–171.

Hanna-Rose, W. & Hansen, U. (1996). Active repression mechanisms of eukaryotic transcription repressors. *Trends in Genetics* **12**, 229–234.

Hua, J. & Patton, J. T. (1994). The carboxyl-half of the rotavirus nonstructural protein NS53 (NSP1) is not required for virus replication. *Virology* **198**, 567–576.

Hua, J., Mansell, E. A. & Patton, J. T. (1993). Comparative analysis of the rotavirus NS53 gene: conservation of basic and cysteine-rich regions in the protein and possible stem–loop structures in the RNA. *Virology* **196**, 372–378.

Hua, J., Chen, X. & Patton, J. T. (1994). Deletion mapping of the rotavirus metalloprotein NS53 (NSP1): the conserved cysteine-rich region is essential for virus-specific RNA binding. *Journal of Virology* **68**, 3990–4000.

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 non-bacterial diarrhoea in China. *Lancet* ii, 1078–1079.

Hung, T., Chen, G. M., Wang, C. A., Fan, R., Yong, R., Chang, J., Dan, R. & Ng, M. H. (1987). Seroepidemiology and molecular epidemiology of the Chinese rotavirus. In *Novel Diarrhoea Viruses*, pp. 49–62. Edited by G. Bock & J. Whelan. Chichester: John Wiley.

James, V. L. A., Lambden, P. R., Caul, E. O. & Clarke, I. N. (1998). Enzyme-linked immunosorbent assay based on recombinant human group C rotavirus inner capsid protein (VP6) to detect human group C rotaviruses in fecal samples. *Journal of Clinical Microbiology* **36**, 3178–3181.

Jiang, B., Tsunemitsu, H., Dennehy, P. H., Oishi, I., Brown, D., Schnagl, R. D., Oseto, M., Fang, Z. Y., Avendano, L. F., Saif, L. J. & Glass, R. I. (1996). Sequence conservation and expression of the gene encoding the outer capsid glycoprotein among human group C rotaviruses of global distribution. *Archives of Virology* **141**, 381–390.

Kapikian, A. Z. & Chanock, R. M. (1996). Rotaviruses. In *Fields Virology*, 3rd edn, pp. 1657–1708. Edited by B. N. Fields, D. M. Knipe & P. M. Howley. Philadelphia: Lippincott–Raven.

Kattoura, M. D., Clapp, L. L. & Patton, J. T. (1992). The rotavirus nonstructural protein, NS35, possesses RNA-binding activity *in vitro* and *in vivo*. *Virology* **191**, 698–708.

Kattoura, M. D., Chen, X. & Patton, J. T. (1994). The rotavirus RNA-binding protein NS35 (NSP2) forms 10S multimers and interacts with the viral RNA polymerase. *Virology* **202**, 803–813.

Kozak, M. (1991). Structural features in eukaryotic mRNAs that modulate the initiation of translation. *Journal of Biological Chemistry* **266**, 19867–19870.

Krishnan, T., Sen, A., Choudhury, J. S., Das, S., Naik, T. N. & Bhattacharya, S. K. (1999). Emergence of adult diarrhoea rotavirus in Calcutta, India. *Lancet* **353**, 380–381.

Lambden, P. R. & Clarke, I. N. (1995). Cloning of viral double-stranded RNA genomes by single primer amplification. In *Methods in Molecular Genetics*, pp. 359–372. Edited by S. D. Aldolph. London: Academic Press.

Lambden, P. R., 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). Products of the porcine group C rotavirus NSP3 gene bind specifically to double-stranded RNA and inhibit activation of the interferon-induced protein kinase PKR. *Journal of Virology* **68**, 3821–3829.

Mattion, N. M., Cohen, J., Aponte, C. & Estes, M. K. (1992). Characterization of an oligomerization domain and RNA-binding properties on rotavirus nonstructural protein NS34. *Virology* **190**, 68–83.

Mitchell, D. B. & Both, G. W. (1990). Conservation of a potential metal binding motif despite extensive sequence diversity in the rotavirus nonstructural protein NS53. *Virology* **174**, 618–621.

Nakata, S., Estes, M. K., Graham, D. Y., Wang, S. S., Gary, G. W. & Melnick, J. L. (1987). Detection of antibody to group B adult diarrhea rotaviruses in humans. *Journal of Clinical Microbiology* **25**, 812–818.

Okada, J., Kobayashi, N., Taniguchi, K. & Urasawa, S. (1999). Analysis on reassortment of rotavirus NSP1 genes lacking coding region for cysteine-rich zinc finger motif. *Archives of Virology* **144**, 345–353.

Palombo, E. A. & Bishop, R. F. (1994). Genetic analysis of NSP1 genes of human rotaviruses isolated from neonates with asymptomatic infection. *Journal of General Virology* **75**, 3635–3639.

Patton, J. T., Salter-Cid, L., Kalbach, A., Mansell, E. A. & Kattoura, M. (1993). Nucleotide and amino acid sequence analysis of the rotavirus nonstructural RNA-binding protein NS35. *Virology* **192**, 438–446.

Pedley, S., Bridger, J. C., Chasey, D. & McCrae, M. A. (1986). Definition of two new groups of atypical rotaviruses. *Journal of General Virology* **67**, 131–137.

Petrie, B. L., Greenberg, H. B., Graham, D. Y. & Estes, M. K. (1984). Ultrastructural localization of rotavirus antigens using colloidal gold. *Virus Research* **1**, 133–152.

Poncet, D., Lindenbaum, P., L'Haridon, R. & Cohen, J. (1997). *In vivo* and *in vitro* phosphorylation of rotavirus NSP5 correlates with its localization in viroplasms. *Journal of Virology* **71**, 34–41.

Qian, Y., Jiang, B. M., Saif, L. J., Kang, S. Y., Ojeh, C. K. & Green, K. Y. (1991). Molecular analysis of the gene 6 from a porcine group C rotavirus that encodes the NS34 equivalent of group A rotaviruses. *Virology* **184**, 752–757.

Ramig, R. F. (1997). Genetics of the rotaviruses. *Annual Review of Microbiology* **51**, 225–255.

Rao, C. D., Das, M., Ilango, P., Lalwani, R., Rao, B. S. & Gowda, K. (1995). Comparative nucleotide and amino acid sequence analysis of the sequence-specific RNA-binding rotavirus nonstructural protein NSP3. *Virology* **207**, 327–333.

Rodger, S. M., Bishop, R. F. & Holmes, I. H. (1982). Detection of a rotavirus-like agent associated with diarrhea in an infant. *Journal of Clinical Microbiology* **16**, 724–726.

Saif, L. J., Bohl, E. H., Theil, K. W., Cross, R. F. & House, J. A. (1980). Rotavirus-like, calicivirus-like, and 23-nm virus-like particles associated with diarrhea in young pigs. *Journal of Clinical Microbiology* **12**, 105–111.

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

Taniguchi, K. & Urasawa, S. (1995). Diversity in rotavirus genomes. *Seminars in Virology* **6**, 123–131.

Taniguchi, K., Kojima, K. & Urasawa, S. (1996*a*). Nondefective rotavirus mutants with an NSP1 gene which has a deletion of 500 nucleotides, including a cysteine-rich zinc finger motif-encoding region (nucleotides 156 to 248), or which has a nonsense codon at nucleotides 153–155. *Journal of Virology* **70**, 4125–4130.

Taniguchi, K., Kojima, K., Kobayashi, N., Urasawa, T. & Urasawa, S. (1996*b*). Structure and function of rotavirus NSP1. *Archives of Virology Supplementum* **12**, 53–58.

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.

Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. (1997). The CLUSTAL X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research* **25**, 4876–4882.

Tian, Y., Tarlow, O., Ballard, A., Desselberger, U. & McCrae, M. A. (1993). Genomic concatemerization/deletion in rotaviruses: a new mechanism for generating rapid genetic change of potential epidemiological importance. *Journal of Virology* **67**, 6625–6632.

Tsunemitsu, H., Saif, L. J., Jiang, B. M., Shimizu, M., Hiro, M., Yamaguchi, H., Ishiyama, T. & Hirai, T. (1991). Isolation, characterization, and serial propagation of a bovine group C rotavirus in a monkey kidney cell line (MA104). *Journal of Clinical Microbiology* **29**, 2609–2613.

Xu, L., Tian, Y., Tarlow, O., Harbour, D. & McCrae, M. A. (1994). Molecular biology of rotaviruses. IX. Conservation and divergence in genome segment 5. *Journal of General Virology* **75**, 3413–3421.

Yolken, R., Arango-Jaramillo, S., Eiden, J. & Vonderfecht, S. (1988). Lack of genomic reassortment following infection of infant rats with group A and group B rotaviruses. *Journal of Infectious Diseases* **158**, 1120–1123.

Received 29 March 1999; Accepted 4 August 1999