Arch Virol (1994) [Suppl] 9: 429–439

Archives of Virology



# Genomic organization and expression of astroviruses and caliciviruses

**M. J. Carter**

School of Biological Sciences, University of Surrey, Guildford, U.K.

**Summary.** Astroviruses and caliciviruses are two families defined initially by their characteristic morphology. Many of these viruses have been difficult to grow in culture. Molecular biology has now provided a valuable insight into the nature of these viruses, and in many respects knowledge of genome structure now outstrips that of more classical virological features. However these advances have allowed a more detailed approach to virus classification and have led to the establishment of the *Astroviridae* as a distinct virus family.

## Introduction

Both astroviruses (family *Astroviridae*) and caliciviruses (family *Caliciviridae*) were initially defined by their morphology under the electron microscope. Caliciviruses are 34–38 nm and covered by cup-like depressions in which stain accumulates (*calix*, a cup). The particle margin is jagged where cups are seen edge-on [3, 4]. Astroviruses are smaller (28 nm), with a smooth margin [22]. A proportion of the virions bear a prominent five or six-pointed star-like motif (*astro*, a star). The morphologies of astroviruses and caliciviruses are each unique (Fig. 1) and until recently there was no satisfactory explanation of how either is formed.

Both virus families are mostly associated with diarrhoea, but some members may cause skin lesions, or respiratory or hepatic infections. Some caliciviruses grow readily in vitro, but despite this they remained relatively uncharacterized. Astroviruses were more difficult to grow and have only recently been recognized as a distinct family. A few strains of a human virus were adapted to growth in continuous cell lines by passage in primary human embryo cells [10]. Caliciviruses remain the better characterized, and it now appears that they may share some features with the astroviruses.

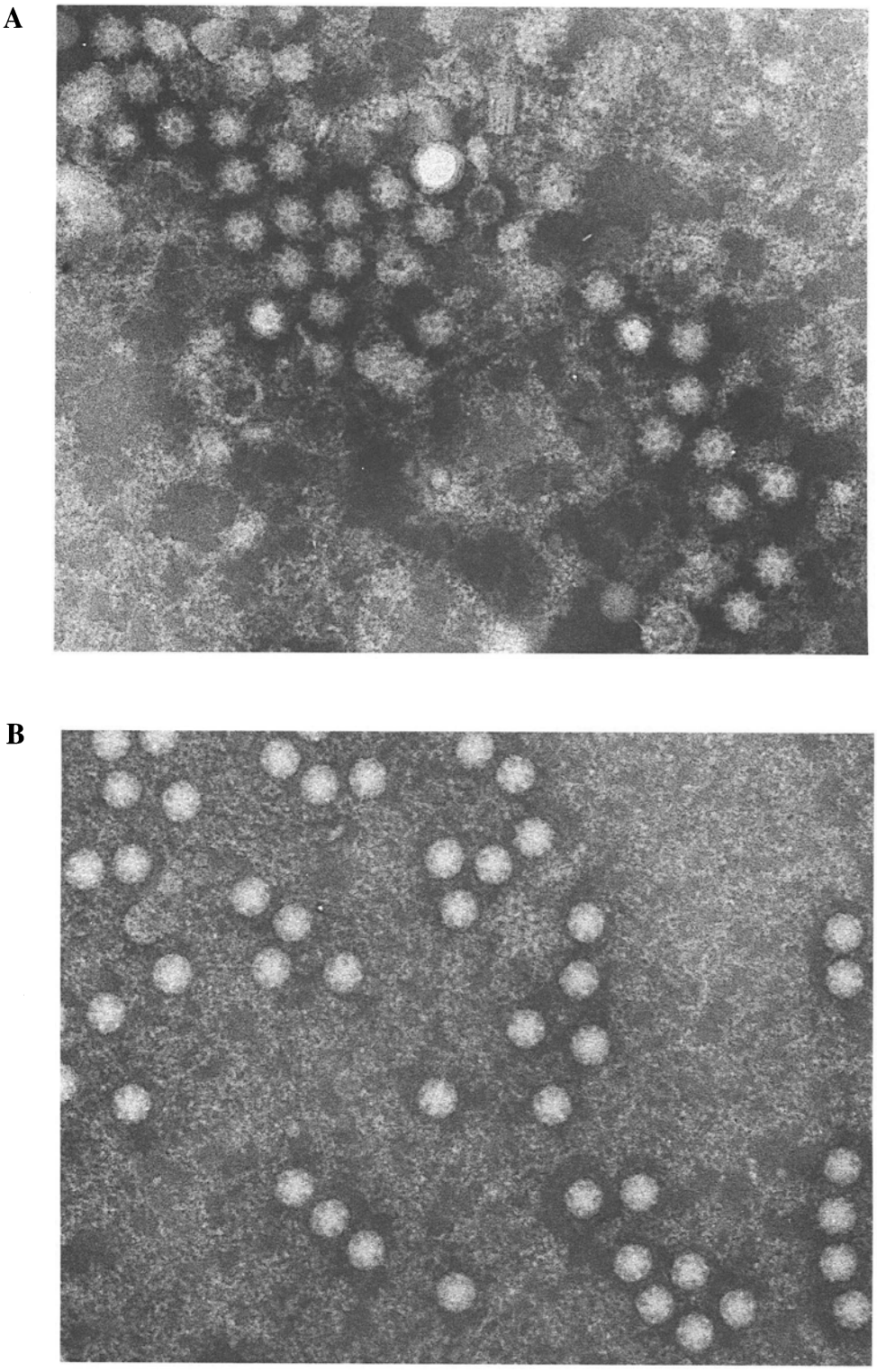


**Fig. 1.** Caliciviruses and astroviruses. Virions stained with phosphotungstic acid, pH 7.0, and are reproduced at a magnification of ×2000 000. **A** Caliciviruses. **B** Astroviruses. Electron micrographs were kindly provided by Prof. C. R. Madeley

## Genome structure of caliciviruses

The best characterized caliciviruses are listed in Table 1. Complete sequences have been determined for feline calicivirus (FCV) [6], for Norwalk virus (NV) [EMBL Database Accession No. M87661], for rabbit haemorrhagic disease virus (RHDV) [11] and for the candidate virus of enterically transmitted human non-A non B hepatitis, hepatitis E virus (HEV) [18]. Partial sequence has been obtained for San Miguel sea-lion virus (SMSV) [15] and for European brown hare syndrome virus (EBHSV) (G. Meyers, pers. comm.). The genomes of these viruses are organized in a similar manner (Fig. 2). The virions contain a single type of capsid protein which is encoded towards the 3′ end [5, 17], non-structural proteins are encrypted by the 5′ end [14]. FCV, NV and HEV carry genes for non-structural proteins in a separate open reading frame (ORF1), and the capsid protein is specified by a second region (ORF2), separated from ORF1 by a termination codon and a frameshift. However in RHDV and also EBHSV, these two genes are fused in frame (ORF1/2) and translation forms a single large polyprotein. Finally, most caliciviruses contain a small, third ORF at the extreme 3′ end (ORF3). This is lacking in HEV, in which a third gene is located internally, spanning the juncton of ORFs 1 and 2.

### *Calicivirus RNA synthesis*

Early reports suggested that up to eight RNA species were synthesized in infected cells and that sequential open reading frames are present in the genome [2, 16]. This has been partially confirmed by sequencing studies, but there are too few ORFs to account for all these mRNAs, and a re-examination of virus-specific mRNA synthesis found only two were present (T. D. K. Brown, pers. comm.). Multiple RNAs were observed in FCV infection, a rapidly cytocidal process, and most

**Table 1.** Best characterized caliciviruses

| Name | | Serotypes |
|---|---|---|
| Vesicular exanthema of swine | (VESV) | 13 |
| San Miguel sealion | (SMSV) | 16 |
| Feline calicivirus | (FCV) | 1 |
| Small round structured viruses e.g. Norwalk virus | (SRSV) (NV) | 5 |
| Rabbit haemorrhagic disease | (RHDV) | 1 |
| European brown hare syndrome | (EBHSV) | 1 |
| Hepatitis E virus | (HEV) | 1 |

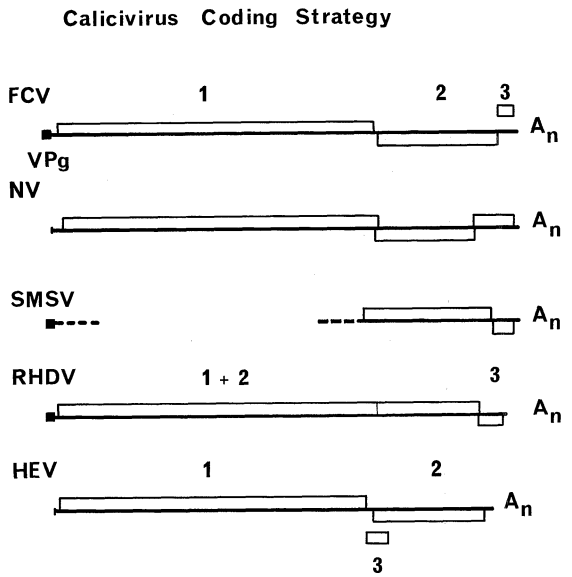


**Fig. 2.** Coding strategy of the caliciviruses. Open reading frames are illustrated as open boxes. The poly A tract at the 3′ end is present in all cases. The 5′ end is covalently linked to a protein VpG ($VP_g$), illustrated where its presence has been demonstrated. VpG is probably also present in NV RNA, although HEV could be capped. Genome lengths (bp) are: *FCV* (feline calicivirus) 7690; *NV* (Norwalk virus) 7644; *RHDV* (rabbit haemorrhagic disease virus) 7437; *HEV* (hepatitis E virus) 7194. *SMSV* San Miguel sea-lion virus

could thus have derived from degradation, or artefactual origin. The prominent bands marked in Fig. 3 result from ribosomal RNA acting as a "snowplough" and concentrating the more diffuse background material into pseudo-bands. These features are removed if mRNA is selected before analysis. This effect was probably compounded [2] by the addition of different RNA as carrier, leading to localized overloading in other positions and an increase in the number of apparent bands.

Both of the RNAs detected in the cell are encapsidated. These comprise the genome (approx. 7.5 kb), and a subgenomic RNA, 2.4–2.7 kb in size. The smaller molecule is 3′ co-terminal with the genome, extending to the 5′ end of ORF2 and serves as an mRNA for the capsid protein. This mRNA is produced by RHDV and EBHSV, in which ORF2 is not separated from ORF1. The subgenomic RNA is presumably the major source of capsid protein made by these viruses, although processing from the ORF1/2 fusion polyprotein may also be possible. Subgenomic mRNA produced to increase expression of a section of a continuous ORF rather than to express a cryptic or modified sequence, is a feature unique to the caliciviruses. It is unclear how the subgenomic RNA is produced. It is likely that the subgenomic RNA is transcribed

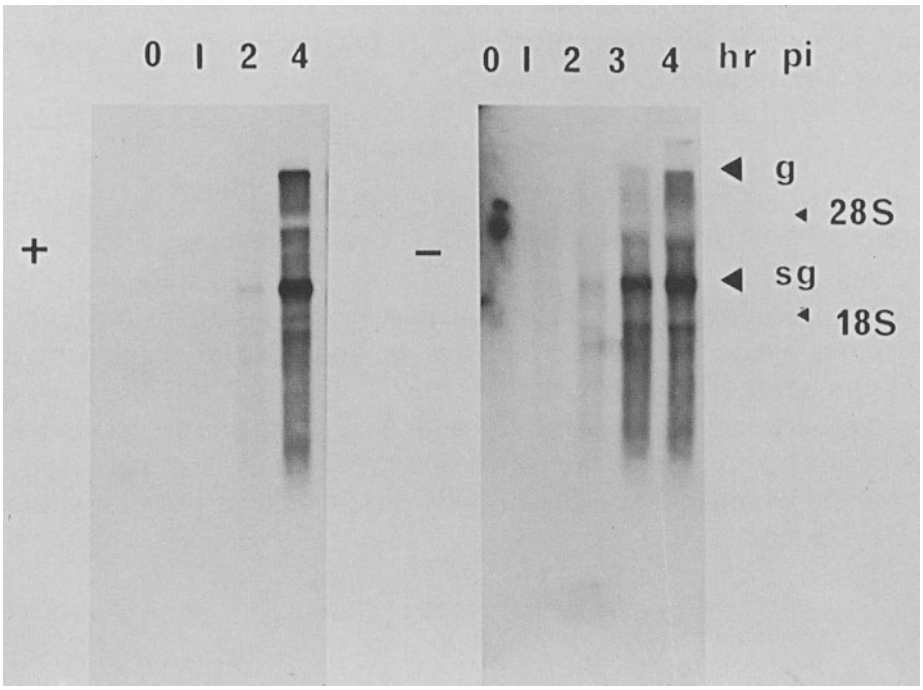


**Fig. 3.** Calicivirus RNA synthesis. Total cell RNA was extracted from cells at the times shown following infection, Northern blotted and probed with strand-specific cDNA probes from the 3′ terminal 300 bases. Designation of (+) and (−) given for each panel refers to the polarity of the probe used. Bands present between the genomic (*g*) and subgenomic (*sg*) RNAs are caused by ribosomal RNA (*18S*, *28S*)

from a genome-sized RNA template, but since both RNAs are encapsidated, and both have been found inside the cell in their negative forms (Fig. 3), and each may be replicated independently [2]. However the existence of the negative form of subgenomic RNA has been disputed and an alternative derivation of this species by partial degradation has been suggested [17].

### *Protein translation and maturation*

#### Non-structural proteins

Proteins specificed by ORF1 resemble picornavirus non-structural proteins, regions have thus been designated 3C-like (cysteine protease), 2C-like (helicase) and 3D-like (polymerase) [14]. However, this leaves about 80 Kd of protein for which no function has yet been assigned. Nomenclature used is assigned by sequence comparison and does not imply maturational relationships. ORF1 is translated as a polyprotein which is subsequently cleaved to produce the final virus products. There

is general correspondence between the numbers and sizes of the proteins found in many calicivirus infections [1] but the proteolytic maturation process is not understood.

### The capsid protein

In the case of NV and RHDV, the subgenomic mRNA specifies a protein of the size observed in mature virus. However, ORF2 of both FCV and SMSV are larger than the virion protein and maturational cleavage is required [1, 7]. Most calicivirus capsid proteins have blocked N termini; but in FCV strain F9, this was not so and sequencing identified the sequence ADDGSIT at the N terminus [5]. This indicates cleavage between residues E(123) and A(124) of the precursor leaving 63 Kd, which agrees with the 62 Kd observed. The C terminus cannot be extensively trimmed since an antibody binding site is present within the last 34 residues.

### The ORF3 gene product

There is no information on the product of the 3′ terminal ORF3 gene. The protein specified by this gene is the best conserved among different strains of FCV (92–98%) and shows 27% identity between FCV and RHDV [6], rising to 47% similarity when conservation of amino acid character is considered [12]. Although there is more variation between the ORF3 protein specified by NV and SMSV, all such calicivirus proteins are relatively small and have a predominance of basic residues. This high degree of conservation argues strongly for a role at the protein level, but no product from this area has been identified. Sera from FCV-vaccinated cats do not recognize the product of the ORF3 gene when it is expressed as a beta-galactosidase fusion protein in *E. coli* (I. D. Milton, unpubl. obs.). ORF3 expression by frameshifting has been suggested, fusing ORF3 to the capsid protein [17]. However no such chimaeric protein has been found and it may be that a small RNA detected in the cell [2], but not confirmed [17], could serve as mRNA for this ORF.

The ORF3 product specified by the internal gene of HEV is distinct from that encoded by the other caliciviruses. It is recognized by antibodies from infected animals. A third mRNA has been detected in infected liver, slightly larger than that encoding the capsid protein, and possibly functions as mRNA for ORF3 [23].

## *Classification*

The caliciviruses were first classified on the basis of their distinctive morphology. This limited the number of firmly established members,

because classical morphological features are not always clear. NV for instance has remained only a candidate calicivirus because of the "fuzzy" appearance of the virion. However, despite differences in ORF1/2 structure RHDV, EBHSV and NV can now be considered caliciviruses because of their gene order, and the similarities of their non-structural polypeptides. However, the status of HEV is questionable. HEV virions do display (albeit poorly) many of the morphological features of the caliciviruses [19], and the gene order is largely calicivirus-like; but there is no convincing similarity between the sequence of HEV ORF1 proteins and those of other caliciviruses (10–20%), and the internal ORF3 has no counterpart in other caliciviruses. The product of an internal ORF in FCV [5] was not recognized by immunize cat serum, and there is no evidence that such a gene is real (I. D. Milton, unpubl. obs.). Furthermore, recent computer analysis has grouped functional motifs in the ORF1 sequence of HEV away from other caliciviruses, and motifs indicative of capping functions have been identified [8]. If HEV RNA is capped rather than protein-linked, such differences may imply that HEV belongs to a separate taxonomic grouping, irrespective of the morphology of the virion.

## Astroviruses: genome structure

The best characterized astroviruses are human astrovirus (HuAV, 5–6 serotypes), bovine astrovirus (BAV, at least 2 serotypes), porcine astrovirus (PAV), and ovine astrovirus (OAV). The best characterized of all is HuAV, a complete genome sequence should be available soon, but information is already available from the 3′ end and from some internal regions [20]. These show few similarities to members of any other virus family, and no small 3′ ORF equivalent to ORF3 of the caliciviruses is present. However, the genome does not comprise a single unbroken ORF and a separate gene at the 3′ end, could encode a protein of 90 Kd (S. Monroe, pers. comm. 1992). A second ORF extends towards the 5′ end and contains both a polymerase motif (T. Lewis, pers. comm.) and a nuclear addreassing sequence (EMBL Database Accession No. Z 16420).

### *RNA synthesis*

Astrovirus genome expression is similar to that of the caliciviruses and two RNAs are synthesized [13], the genome and a subgenomic species of a size suitable for the expression of the 3′ ORF. At first this was not confirmed by Northern blotting [20] but this discreapancy has now been resolved: Unlike the caliciviruses, the astrovirus subgenomic RNA may not be packaged, and does not persist at the late times studied. Blotting studies easily demonstrated both RNAs when samples were analysed

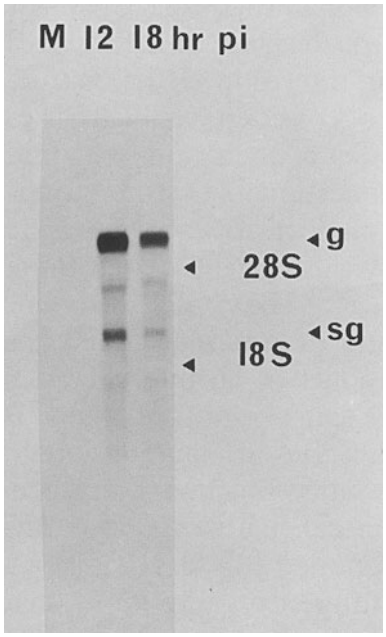


**Fig. 4.** Astrovirus RNA synthesis. RNA was extracted from CaCo-2 cells infected with human astrovirus serotype 1 at 12 and 18 h pi for Northern blotting and detection with a cDNA probe prepared from the 3′ end of the genome. *g* Genomic and *sg* subgenomic RNAs. The position of ribosomal RNA causing "snowploughing" between the two virus RNAs is indicated

earlier in infection (Fig. 4). Both RNAs may exist in both positive and negative senses (S. Monroe, pers. comm.).

### *Protein translation and maturation*

#### Non structural proteins

Little is known about astrovirus non-structural proteins but these are probably encoded as a polyprotein towards the 5′ which contains a polymerase motif and a nuclear targeting signal. It is not known if this is functional, but a discrete immune-fluorescence has been observed in the nucleus [22].

#### Structural proteins

The large ORF at the 3′ terminus is thought to specify the structural proteins. In support of this hypothesis we have identified a short run of basic amino acids in this gene which is similar to basic stretches found in the capsid proteins of two coronaviruses (Table 2). All astroviruses contain at least two proteins of approximately 30 Kd, as well as a smaller molecule whose size varies between viruses 13–27 Kd [22]. A fourth small protein (5.2 Kd) has been observed but not confirmed [9]. These account for 80–90 Kd and an unstable protein of this size has been observed in infection [13]. This could be a precursor to the virion

**Table 2.** Comparison of basic residues in the presumptive structural gene of human astrovirus type 1 with similar sequences in coronavirus capsid proteins

| | |
|---|---|
| GRSRSKSRARSQSRGR | Human astrovirus type 1 |
| ***.** ******* | |
| SRSRSRNRSQSRGR | Transmissible gastroenteritis virus |
| *** *** * | |
| SRSTSRASS | Bovine coronavirus nucleocapsid |

* Residues identical to those in astrovirus type 1 sequence; (.) residues conserved in character with astrovirus type 1

proteins, and its size is consistent with the coding capacity of the 3′ ORF.

### *Classification*

Astroviruses have been variously suggested as possible members of the family *Picornaviridae*, based on their structural proteins, or the family *Caliciviridae*, based on intracellular RNA synthesis. However, their unique morphology, and the observation of protein migration to the nucleus, suggest that they are distinct from members of both these virus families. Accordingly the astroviruses are now considered as representing a distinct family, the *Astroviridae* (Monroe et al., in press). Perhaps both astroviruses and caliciviruses belong to a larger grouping of non-enveloped positive-stranded RNA viruses with distinctive gene order and mode of expression.

## Note added in proof

The complete sequences of two serotypes of human astrovirus (types 1 and 2) have now been determined. These both show that the non-structural gene (ORFf1) is present as two sections (ORF 1a and 1b), linked by motifs indicative of ribosomal frameshifting. Furthermore the viral protease has a serine at the active site, rather than a cysteine as in the picorna- and caliciviruses. These features are similar to those observed in the luteoviruses of plants and further support the classification of astroviruses as a novel family of animal viruses.

## Acknowledgements

I am grateful to those workers who allowed me to present their data at Clearwater and in this review. In particular, Dr. J. Neill at Pioneer Hibred International, Iowa, U.S.A.; Dr. G. Meyers and Prof. J. Thiel at the Bundesforschungsanstalt für Viruskrankheiten der Tiere, Tübingen, Germany; Dr. F. Parra at the University of Oviedo,

Spain, and Dr. C. Rossi at the Instituto Zooprofilattico Sperimentale, Brescia, Italy, for their help with the calicivirus section of this presentation. Also Dr. S. Monroe at U.S. Centres for Disease Control, Atlanta, U.S.A. for his help with the astroviruses. Finally, I thank my own colleagues at Newcastle, Dr. M. Willcocks and Dr. I. Milton for their assistance.

## References

1. Carter MJ (1989) Feline calicivirus protein synthesis investigated by Western blotting. Arch Virol 108: 69–79
2. Carter MJ (1990) Transcription of feline calicivirus RNA. Arch Virol 114: 143–152
3. Carter MJ, Madeley CR (1987) Caliciviridae. In: Nermut MV, Steven AC (eds) Animal virus structure. Elsevier, Amsterdam, pp 121–128 (Perspectives in Medicine and Virology, vol 3)
4. Carter MJ, Milton ID, Madeley CR (1991) Caliciviruses. Rev Med Virol 1: 177–186
5. Carter MJ, Milton ID, Turner PC, Meanger J, Bennett M, Gaskell RM (1992) Identification and sequence determination of the capsid protein gene of feline calicivirus. Arch Virol 122: 223–235
6. Carter MJ, Milton ID, Meanger J, Bennett M, Gaskell RM, Turner PC (1992) The complete sequence of a feline calicivirus. Virology 190: 443–448
7. Fretz M, Schaffer FL (1978) Calicivirus proteins in infected cells: evidence for a capsid polypeptide precursor. Virology 89: 318–321
8. Koonin EV, Gorbalenya AE, Purdy MA, Rozanov MN, Reyes GR, Bradley DW (1992) Computer-assisted alignment of functional domains in the non-structural polyprotein of hepatitis E virus: delineation of an additional group of positive-strand RNA plant and animal viruses. Proc Natl Acad Sci USA 89: 8259–8263
9. Kurtz JB (1989) Astroviruses. In: Farthing MJG (ed) Viruses and the gut. Proceedings of the Ninth BSG Smith Kline and French International Workshop. Smith Kline and French Laboratories Ltd, Welwyn Garden City, pp 84–87
10. Lee TW, Kurtz JB (1981) Serial propagation of astrovirus in tissue culture with the aid of trypsin. J Gen Virol 57: 421–424
11. Meyers G, Wirblich C, Thiel H-J (1991) Rabbit haemorrhagic disease virus – molecular cloning and sequencing of a calicivirus genome. Virology 184: 664–676
12. Milton ID, Vlasak R, Nowotny N, Rodak L, Carter MJ (1992) Genomic 3′ terminal sequence comparison of three isolates of rabbit haemorrhagic disease virus. FEMS Microbiol Lett 93: 37–42
13. Monroe SS, Stine SE, Gorelkin L, Herrmann JE, Blacklow NR, Glass RI (1991) Temporal synthesis of proteins and RNAs during human astrovirus infection in cultured cells. J Virol 65: 641–648
14. Neill JD (1990) Nucleotide sequence of a region of the feline calicivirus genome which encodes picorna virus-like RNA dependent RNA polymerase, cysteine protease and 2C polypeptides. Virus Res 17: 145–160
15. Neill JD (1992) Nucleotide sequence of the capsid protein gene of two serotypes of San Miguel sea-lion virus: identification of conserved and non-conserved amino acid sequences among calicivirus capsid proteins. Virus Res 24: 211–222
16. Neill JD, Mengeling WL (1988) Further characterization of the virus-specific RNAs in feline calicivirus infected cells. Virus Res 11: 143–152
17. Neill JD, Reardon IM, Heinrikson RL (1991) Nucleotide sequence and expression of the capsid protein gene of feline calicivirus. J Virol 65: 5400–5447

18. Tam AW, Smith MM, Guerra ME, Huang C-C, Bradley DW, Fry KE, Reyes GR (1991) Hepatitis E virus (HEV): molecular cloning and sequencing of the full-length viral genome. Virology 125: 120–131
19. Ticehurst J, Rhodes LL, Krawczynski K, Asher LVS, Engler WF, Mensing TL, Caudill JD, Sjogren MH, Hoke CH, LeDuc JW, Bradley DW, Binn LN (1992) Infection of owl monkeys (*Aotus trivirgatus*) and cynomolgus monkeys (*Macaca fascicularis*) with hepatitis E virus from Mexico. J Infect Dis 165: 835–845
20. Willcocks MM, Carter MJ (1992) The 3′ terminal sequence of a human astrovirus. Arch Virol 124: 279–289
21. Willcocks MM, Carter MJ, Laidler FR, Madeley CR (1990) Growth and characterization of human faecal astroviruses in a continuous cell line. Arch Virol 113: 73–82
22. Willcocks MM, Carter MJ, Madeley CR (1992) Astroviruses. Rev Med Virol 2: 97–106
23. Yarbough PO, Tam AW, Fry KE, Krawczynski K, McCaustland KA, Bradley Dw, Reyes GR (1991) Hepatitis E virus: identification of type-common epitopes. J Virol 65: 5790–5797

Author's address: Dr. M. J. Carter, School of Biological Sciences, University of Surrey, Guildford GU2 5XH, U.K.