

## Review article

# Toroviruses: replication, evolution and comparison with other members of the coronavirus-like superfamily

Eric J. Snijder<sup>1</sup>\* and Marian C. Horzinek<sup>2</sup>

<sup>1</sup>Department of Virology, Institute of Medical Microbiology, Faculty of Medicine, Leiden University, P.O. Box 320, 2300 AH Leiden and <sup>2</sup>Virology Division, Department of Infectious Diseases and Immunology, Veterinary Faculty, University of Utrecht, Yalelaan 1, 3584 CL Utrecht, The Netherlands

### General introduction

Based on their morphological and physicochemical characteristics, the *Toroviridae* were initially proposed to constitute a new family of enveloped RNA viruses (Horzinek & Weiss, 1984; Horzinek *et al.*, 1987). However, recent analysis of the genetic information and replication strategy of the prototype Berne virus (BEV) (Snijder *et al.*, 1988, 1990*a, c*) has revealed that toroviruses are not unique: they are clearly related to the *Coronaviridae* and, more distantly, to the arteriviruses (den Boon *et al.*, 1991*b*). This information has led to the reclassification of the toroviruses as a new genus in the coronavirus family (Pringle, 1992) and to the introduction of the unofficial term 'coronavirus-like superfamily' to indicate the evolutionary ties between the three virus groups mentioned above.

The history of torovirus research not only illustrates the taxonomic consequences present-day molecular analysis may have; the BEV genome has also turned out to be a showcase for the two driving forces in RNA virus evolution: divergence from a common ancestor and RNA recombination (Snijder *et al.*, 1991*a*). Comparison of the torovirus genome with those of corona- and arteriviruses has contributed to our understanding of the evolution and replication of all three virus groups.

The aim of this review is to present an overview of the data obtained during the molecular characterization of the torovirus prototype BEV. In addition, evolutionary aspects of our results and their relevance for the classification of toroviruses, coronaviruses and arteriviruses will be discussed. For information about the clinical and epidemiological aspects of toroviruses the reader is referred to a recent review by Koopmans & Horzinek (1993).

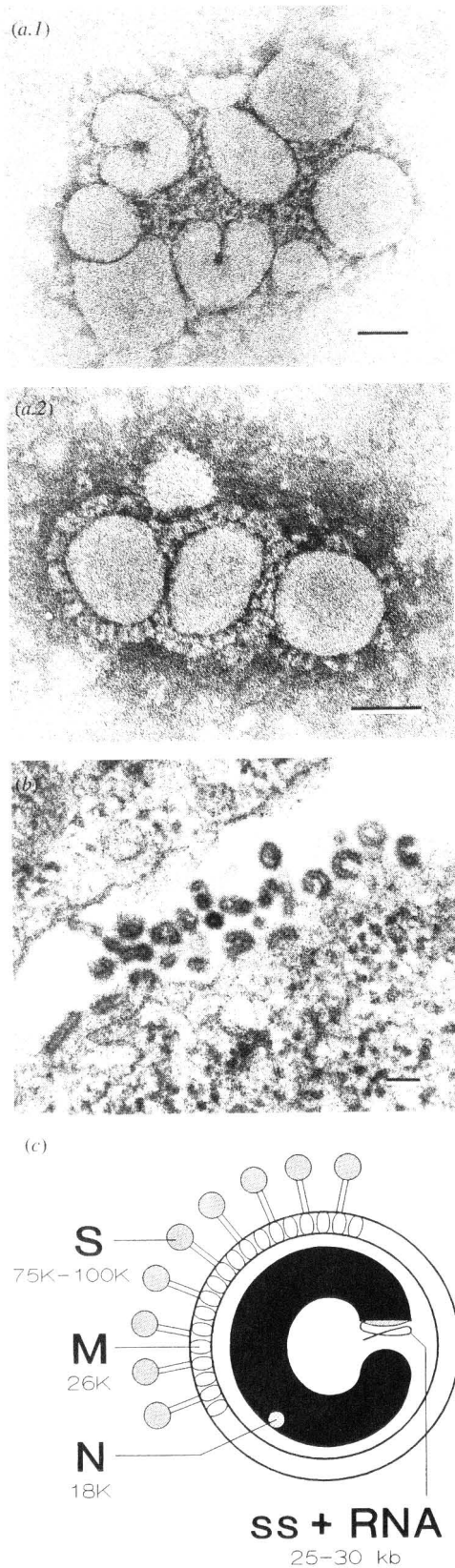
### History of torovirology

The isolation, biochemical characterization and epidemiology of established and possible toroviruses have been reviewed previously (Horzinek *et al.*, 1987; Weiss & Horzinek, 1987). Therefore, we will only briefly summarize the history of torovirus research in this paragraph.

In 1972, BEV was isolated from a diarrhoeal horse during routine diagnostic work at the Veterinary Faculty, University of Berne, Switzerland. The isolate, designated P138/72, displayed an unusual morphology but was not studied in more detail until a similar virus had been isolated from diarrhoeal calves in Breda, Iowa, U.S.A., in 1979 (Weiss *et al.*, 1983; Woode *et al.*, 1982). Berne virus and Breda virus (BRV) were found to be antigenically related (Weiss *et al.*, 1983). In later years, a second BRV 'serotype' was identified (Woode *et al.*, 1985), and similar pleomorphic viruses were found in the stools of children and adults with gastroenteritis in Birmingham, U.K., and Bordeaux, France (Beards *et al.*, 1984). Immunoelectron microscopy indicated that the human viruses were serologically related to BRV and BEV (Beards *et al.*, 1986).

Antibodies against toroviruses appear to be widespread in ungulates (horses, cattle, sheep, goats and pigs; Weiss *et al.*, 1984; Brown *et al.*, 1987; Koopmans *et al.*, 1989). Some wild mice and laboratory rabbits have also been found seropositive, but inconclusive results were obtained with human sera (Weiss *et al.*, 1984; Brown *et al.*, 1987). Intravenous inoculation of foals with BEV induced neutralizing antibodies, but did not cause clinical symptoms (Weiss *et al.*, 1984). However, this may have been due to the non-oral method of infection or to adaptation of the virus to *in vitro* culture. Upon experimental oral infection BRV was found to cause neonatal calf diarrhoea (Woode *et al.*, 1985).

The unique morphology of BEV and BRV (see below)



and the physicochemical properties of BEV (Weiss & Horzinek, 1986a) led to the proposal of a new family of enveloped animal viruses, the *Toroviridae* (Horzinek & Weiss, 1984; Horzinek *et al.*, 1987; Weiss & Horzinek, 1987); BEV, the only torovirus so far which can be grown in cultured cells, was identified as the prototype.

### The torovirus particle

Torovirions are pleomorphic and measure 120 to 140 nm for their largest diameter (Fig. 1a). Spherical, oval, elongated and kidney-shaped particles are observed. Their two most conspicuous features are the spikes on the envelope, which resemble the peplomers of coronaviruses, and the tubular nucleocapsid of helical symmetry (Fig. 1b), which seems to determine the shape of the virion (Weiss *et al.*, 1983). The presence of nucleocapsids in the form of a doughnut, a shape described in Latin by the word 'torus', led to the proposal of the name *Toroviridae* for this group of viruses (Horzinek & Weiss, 1984).

The morphogenesis of BEV [in cultured cells (Weiss & Horzinek, 1986b)] and BRV [in intestinal cells of infected calves (Fagerland *et al.*, 1986)] has been studied by electron microscopy. Preformed tubular nucleocapsids were found to bud at intracellular membranes, predominantly those of the Golgi system. A morphological change seems to take place during virus maturation: prior to budding the nucleocapsids are straight rods whereas the characteristic torus shape is prevalent in the extracellular virus. Nucleocapsids of both BEV and BRV were found to accumulate in the nuclei of infected cells. A nuclear stage or function may be involved in viral replication (Weiss & Horzinek, 1986b; Horzinek *et al.*, 1984), although a more trivial explanation is possible, e.g. late c.p.e.

From its inactivation by organic solvents it was concluded that BEV is an enveloped virus. The insensitivity of BEV replication to 5'-iodo-2'-deoxyuridine indicated that toroviruses contain an RNA genome (Weiss *et al.*, 1983). Buoyant densities of 1.16, 1.18 and 1.14 g/ml in sucrose were determined for BEV (Weiss *et al.*, 1983), BRV serotype 2 (Koopmans *et al.*, 1986) and human toroviruses (Beards *et al.*, 1986), respectively.

Fig. 1. (a.1 and a.2) BEV particles from a purified preparation showing their pleomorphism. Bar represents 50 nm. (b) Different forms of BEV particles encountered in ultrathin sections through BEV-infected cells. Bar represents 200 nm. Examples of the characteristic tubular structure of the BEV nucleocapsid can be seen. (c) Schematic representation of the architecture of BEV, the torovirus prototype. The localization and size of the S, M and N proteins and the BEV genome are indicated. The photographs in (a) and (b) were reprinted from Weiss *et al.* (1983).

## The torovirus genome

From sedimentation studies, the BEV genome was estimated to be at least 20 kb in length. The genomic RNA was found to be polyadenylated and infectious in a transfection assay, an activity which was destroyed by RNase treatment (Snijder *et al.*, 1988).

The BEV genome probably contains six open reading frames (ORFs; Fig. 2). As in coronaviruses, the first two reading frames (ORFs 1a and 1b) from the 5' end are translated from genomic RNA and constitute the viral replicase gene. The four remaining reading frames, of which ORFs 2, 3 and 5 have been identified as structural genes, are expressed by the generation of a 3'-coterminally nested set of mRNAs (see below).

By sequence analysis of cDNA derived from BEV genomic RNA (Snijder *et al.*, 1990*a*), subgenomic mRNAs (den Boon *et al.*, 1991*b*; Snijder *et al.*, 1989, 1990*b*, 1991*a*) and defective interfering (DI) RNAs (Snijder *et al.*, 1991*b*), two contiguous sequences were obtained. The first one of 1580 nucleotides (nt) represents the 5' end of the genome. It comprises a remarkably large (824 nt) 5' non-translated region and a downstream sequence encoding the N-terminal 252 amino acids of the replicase ORF 1a product (Snijder *et al.*, 1991*b*). Assuming that the BEV replicase gene consists of only

two ORFs (1a and 1b), as it does in coronaviruses, the second genome region (14519 nt) begins 1051 nt upstream of the ribosomal frameshift site, which connects ORFs 1a and 1b, and continues down to the 3' poly(A) tail. It contains the complete sequence of the ORFs 1b to 5 and a 3' non-translated region of 200 nt. If the size of BEV ORF 1a is similar to that of coronaviruses (12 to 14 kb; Bournnell *et al.*, 1987; Lee *et al.*, 1991; Herold *et al.*, 1993), a genome size of 26 to 28 kb can be extrapolated. This implies that 10 to 12 kb of the ORF 1a sequence remain to be determined. The characteristics of the BEV mRNAs and the various ORFs in the genomic sequence are summarized in Table 1.

Only the sequence of the 269 nt directly upstream of the poly(A) tail of the BRV genome has been determined so far (Koopmans *et al.*, 1991). In this region, which comprises the 3' end of the nucleocapsid protein gene and the 3' non-translated region of the genome, the sequence of the bovine torovirus BRV is 93% identical to that of BEV. Although these stretches probably constitute the more highly conserved regions of the genome, the data indicate that bovine and equine torovirus isolates are quite closely related. This assumption is supported by hybridization experiments where labelled BRV genomic RNA fragments hybridized strongly to a broad range of BEV cDNA clones (Koopmans *et al.*, 1991). An exception may be the 5' part of the spike protein gene, which is highly variable also in coronaviruses (Spaan *et al.*, 1988).

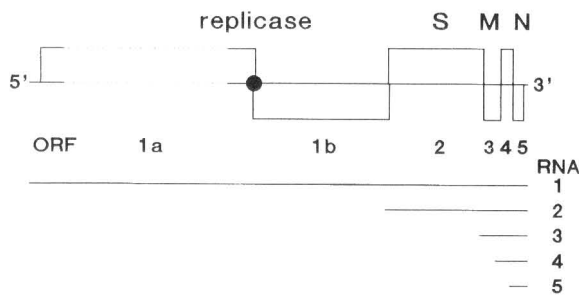


Fig. 2. Genome organization and expression of BEV. The location of ORFs in the BEV genome is indicated. The dashed lines indicate the ORF 1a region which remains to be sequenced (estimated size: 10 to 12 kb). The lower part of the figure shows the nested set of mRNAs generated in infected cells.

## BEV genome transcription and expression

During BEV replication a 3'-coterminally nested set of five mRNAs (including the genome) is generated (Fig. 2). The estimated sizes of these RNAs and their relative abundance late in infection (Snijder *et al.*, 1988) are listed in Table 1. Both coronavirus (for reviews, see Spaan *et al.*, 1988; Lai, 1990) and arterivirus (de Vries *et al.*, 1990) mRNAs contain a common leader sequence at their 5' end. However, no evidence for the presence of such a

Table 1. Characteristics of BEV RNAs and ORFs

RNA	Size (kb)	Relative molarity (%)	Encoded ORF	Position upstream of 3' end	Number of amino acids encoded	Calculated size	Protein
1	25-30?	2	1a 1b	?-13468 13477-6604	? 2291	? 261K	Replicase Replicase
2	6.9	3	2	6684-1941	1581	178K	Spike
3	2.1	30	3	1909-1210	233	26K	Membrane
4	1.4	13	4	1150-724	142	16K	Pseudogene?
5	0.8	52	5	680-200	160	18K	Nucleocapsid

leader in BEV RNAs (Snijder *et al.*, 1990*c*, 1991*b*) has been obtained.

For all three virus groups conserved sequences have been described which are probably involved in subgenomic RNA transcription. For coronaviruses this is the intergenic motif [e.g. 5' AAUCuAuAC 3' for mouse hepatitis virus (MHV)] which has been identified as the site of leader-to-body fusion [reviewed by Spaan *et al.* (1988) and Lai *et al.* (1990)]. Similar but shorter junction sequences (5' UCAACu 3', 5' UnuAACC 3', 5' gnuAACC 3') have been reported for the arteriviruses equine arteritis virus (EAV; de Vries *et al.*, 1990; den Boon *et al.*, 1991*b*), lactate dehydrogenase-elevating virus (LDV; Chen *et al.*, 1993) and Lelystad virus (LV; Meulenber *et al.*, 1993*b*), respectively. Although the torovirus genome also contains conserved AU-rich intergenic sequences (5' uaUcUUUAGa 3'; Fig. 3), there is no evidence for fusion of mRNA bodies to a common leader at this position (Snijder *et al.*, 1990*c*, 1991*b*). BEV mRNAs appear to terminate at or just upstream of the conserved intergenic sequence. In terms of transcription, however, the consequences of this dissimilarity could be limited: direct binding of the polymerase to the various BEV 'core promoters' on a negative-stranded template could simply replace the leader-priming which is assumed to be used by corona- and arteriviruses.

Analysis of BEV subgenomic RNA synthesis by u.v. transcription mapping (Snijder *et al.*, 1990*c*) has indicated that the BEV mRNAs are transcribed independently, from in separate transcription units. However, these experiments were performed late in infection, and alternative transcription mechanisms during the earlier stages of BEV replication cannot be ruled out. In fact, recent data have revealed that the u.v. sensitivity of coronavirus subgenomic RNA synthesis changes in the course of infection (Yokomori *et al.*, 1992), which underscores the complexity of the coronavirus transcription process.

The fact that their subgenomic RNAs contain 5'- and 3'-terminal sequences which are identical to those of the genomic RNA could be important for coronavirus replication: transcriptionally active negative-stranded subgenomic RNAs have been detected in infected cells. This implies that coronavirus subgenomic RNAs may function as replicons (Sawicki & Sawicki, 1990; Sethna *et al.*, 1989), although the possibility of negative-stranded, subgenomic dead-end products (Jeong & Makino, 1992) remains to be excluded.

Until now, attempts to demonstrate negative-stranded viral RNA in torovirus-infected cells have remained unsuccessful, but in EAV-infected cells subgenomic replicative form RNAs appear to be present (E. J. Snijder, unpublished). In view of the many similarities in replication between corona-, toro- and arteriviruses,

these virus groups may very well use similar transcription mechanisms with only minor variations.

## Structural proteins

The torovirion proteins were initially analysed by metabolic labelling of BEV-infected cells. Possible structural proteins of 19K, 22K, 37K and 75K to 100K were identified in [<sup>35</sup>S]methionine-labelled virions (Horzinek *et al.*, 1984). After sequence analysis of the BEV structural genes and characterization of their products, the 19K, 22K and 75K to 100K polypeptides (Fig. 1*c*) could be assigned to ORFs 5, 3 and 2, respectively, in the BEV sequence (den Boon *et al.*, 1991*a*; Snijder *et al.*, 1989, 1990*b*). However, no evidence for virus specificity was obtained for the 37K band.

Since BRV cannot be grown in cultured cells, its protein composition was studied by means of surface radioiodination of purified virus (Koopmans *et al.*, 1986). Possible virus-specific polypeptide species of 20K, 37K, 85K and 105K were identified in this manner.

### The N protein

Upon detergent treatment, the 22K and 75K to 100K proteins were removed from the BEV particle which indicates that these are membrane-associated (Horzinek *et al.*, 1985). Only the 19K protein was present in purified nucleocapsids and was accordingly named nucleocapsid (N) protein. The N protein is most abundant in the virion, accounting for about 80% of its protein mass. It is a phosphorylated protein possessing RNA-binding properties (Horzinek *et al.*, 1985). The location of the N protein gene was deduced from *in vitro* translation of RNA 5 (Snijder *et al.*, 1988) and sequence analysis of the region upstream of the BEV poly(A) tail (Snijder *et al.*, 1989). Two overlapping potential ORFs, encoding proteins of 160 and 88 amino acids, were identified in the RNA 5 sequence. The larger of the two was concluded to be the BEV N protein gene on the basis of the calculated  $M_r$  value of 18.3K and the amino acid composition of its product. The N protein could not be labelled in the presence of [<sup>35</sup>S]cysteine (Snijder *et al.*, 1989), as predicted from the absence of this amino acid from the 160 residue RNA 5-encoded protein.

Analysis of the second RNA 5 ORF, encoding an 88 amino acid product, revealed this polypeptide to have an  $M_r$  of 9.8K and to be generally hydrophobic. A comparison of the sequences surrounding the AUG codons of the two ORFs showed that the AUG codon of the smaller ORF is in a much less favourable context for use as a translation initiation signal. A probable product of this ORF was found during *in vitro* translation studies

of RNA derived from cDNA constructs (Snijder *et al.*, 1989), but a 9.8K protein has never been observed in BEV virions or in infected cells. Nonetheless we cannot exclude the possibility that this ORF may be a functional gene, especially since a similar situation, that of a smaller hydrophobic protein expressed from an ORF which completely overlaps with the N protein gene, was recently reported for the bovine coronavirus (Senanayake *et al.*, 1992).

#### The M (E) protein

The 22K BEV membrane protein was initially named E (for envelope) protein. However, in view of the inclusion of the torovirus genus in the coronavirus family and the data summarized below, the 22K protein will from now on be referred to as membrane (M) protein, for reasons of consistency.

The BEV M protein, which is translated from mRNA 3 (Snijder *et al.*, 1988), is an unglycosylated polypeptide (Horzinek *et al.*, 1986) accounting for about 13% of the virion protein mass (Horzinek *et al.*, 1985). The M protein gene is 699 nt long and encodes a 26.5K protein that in its N-terminal part contains the three membrane spanning  $\alpha$ -helices which are so characteristic for coronavirus M proteins (den Boon *et al.*, 1991a). The 4.5K difference between the calculated and observed sizes of the M protein is accounted for by aberrant migration in polyacrylamide gels, probably due to the extreme hydrophobicity of the protein.

The membrane topology of the M protein has been inferred from *in vitro* translation studies using the M protein itself and a hybrid protein which contained a C-terminal tag (den Boon *et al.*, 1991a). After *in vitro* translation in the presence of microsomes, about 85% of each protein was resistant to protease K digestion. The disposition of the torovirus M protein in the membrane is very similar to that of coronavirus M proteins: the N terminus is located in the lumen of the endoplasmic reticulum, a large central part of the protein is embedded in the membrane, and the C terminus is located at the cytoplasmic face of the endoplasmic reticulum. Since the BEV M protein does not contain an N-terminal signal sequence, one of the hydrophobic transmembrane domains is assumed to function as an internal signal sequence. Like coronavirus M proteins, the BEV M hybrid carrying the C-terminal tag accumulated in intracellular membranes during transient expression experiments. Thus the torovirus M protein may play a role in the intracellular budding process, as suggested for its coronavirus counterpart (Dubois-Dalcq *et al.*, 1982; Holmes *et al.*, 1981; Rottier *et al.*, 1981).

Despite the striking similarities in size, structure and functional characteristics, no significant sequence simi-

larities could be detected between the M protein sequences of BEV and various coronaviruses (den Boon *et al.*, 1991a). Yet the observed similarities probably reflect common ancestry, particularly since viral structural proteins are known to evolve at a higher rate than non-structural proteins.

#### The S (P) protein

The second torovirus envelope protein, originally known as the peplomer (P) protein, has also recently been renamed, to follow the example of coronavirus protein nomenclature: it will be referred to as the spike (S) protein. The reasons for this change are the same as discussed above for the M protein: common structural characteristics, function and ancestry, but no convincing primary sequence similarities.

The BEV envelope is studded with drumstick-shaped projections which measure about 20 nm in length (Weiss *et al.*, 1983; Fig. 1). In early torovirus studies, the heterogeneous 75K to 100K protein material had already been assumed to represent the viral spike protein(s) since it was recognized by both neutralizing and haemagglutination-inhibiting monoclonal antibodies (Horzinek *et al.*, 1986; Kaeffer *et al.*, 1989). Pulse-chase labelling experiments have revealed that the 75K to 100K material is derived from a 200K precursor present in infected cells, but not in virus particles (Horzinek *et al.*, 1986). Extensive *N*-glycosylation and proteolytic cleavage of the precursor are part of the post-translational processing of the torovirus S protein. From cross-reactivity between different BRV 'serotypes' in radioimmuno-precipitation, it was concluded that the 85K and 105K protein species probably represent surface structures (Koopmans *et al.*, 1986). Rabbit antisera raised against purified BRV were shown to recognize the BEV S protein in immunoprecipitation (E. J. Snijder & M. Koopmans, unpublished).

Berne virus ORF 2 has been identified as the S protein gene. It encodes an apoprotein of 1581 amino acids ( $M_r$  approximately 178K) containing 18 potential *N*-glycosylation sites. Among the common structural characteristics of the BEV and coronavirus S proteins are an N-terminal signal sequence, a putative C-terminal transmembrane anchor, two heptad repeat domains and a possible 'trypsin-like' cleavage site. The mature S protein consists of two subunits; their electrophoretic mobility upon endoglycosidase F treatment strongly suggests that the predicted cleavage site, consisting of five consecutive arginine residues, is functional *in vivo* (Snijder *et al.*, 1990b). As postulated previously for the coronavirus S protein (de Groot *et al.*, 1987), the heptad repeats are assumed to be involved in the generation of an intra-chain coiled-coil secondary structure. Inter-chain inter-

actions of the same kind may be involved in protein oligomerization, which is supported by the fact that in a sucrose gradient assay the BEV S protein was shown to form dimers (Snijder *et al.*, 1990*b*). The intra- and inter-chain coiled-coil interactions may stabilize the elongated BEV spikes.

The apparent absence of cell-to-cell fusion during BEV infection (in culture) may indicate an important functional difference between toro- and coronavirus S proteins. The formation of syncytia in infected cultures is one of the most striking features of coronavirus infection. The S proteins of several coronaviruses have been shown to possess fusion activity at neutral pH (Collins *et al.*, 1982; Vennema *et al.*, 1990); expression of this surface glycoprotein alone leads to fusion. Although BEV causes a pronounced c.p.e., the formation of syncytia was never observed, nor could it be induced by lowering the pH of the culture medium (E. J. Snijder, unpublished).

### The 'coronavirus-like' replicase

Viral replicase genes are at the hub of viral biology. Several domains in the replicase proteins are highly conserved among disparate groups of plant and animal

```

RNA 2  GUGCCUAAUUUUUUAAAAGUGUUUAGUACUAGUUUUUA
RNA 3  AGAAAAGUUAGUCACUUUCUUUAGAAGAAGGUUUGC
RNA 4  GUUUGAGUAGCCACUUAUCUUUAGAAGAUGUUGAU
RNA 5  GUUAGUGAGAGACACUAUCUUUAGAGAAAGAGCCA
DI      NNAACGUUAUCUUUAGAAGUUUAUGUU
  
```

Fig. 3. Alignment of putative BEV 'core promoter' sequences (underlined) and their context. Fully conserved nucleotides are printed in bold. The sequences are derived from sequence analysis of the genome (Snijder *et al.*, 1990*c*) and of DI RNAs (Snijder *et al.*, 1991*b*) of BEV.

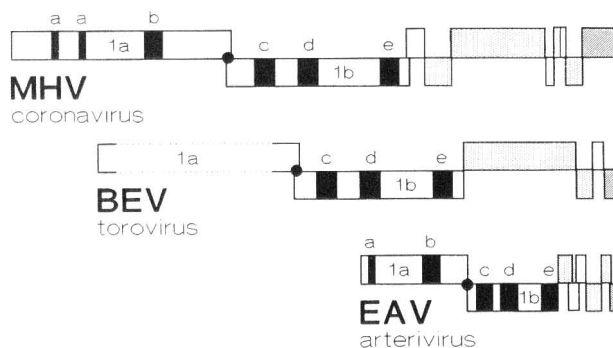


Fig. 4. Genome organization of MHV (coronavirus), BEV (torovirus) and EAV (arterivirus). The genes encoding structural proteins are indicated by dotted (envelope proteins) and hatched (N proteins) boxes. The ORFs that form the replicase gene (1a and 1b) are indicated for each virus. Filled boxes represent homologous (c, d, e) and possibly homologous (a, b) replicase domains: a, papain-like protease; b, chymotrypsin/3C-like protease; c, polymerase; d, helicase; e, C-terminal 1b domain.

RNA viruses (Strauss & Strauss, 1988; Goldbach & Wellink, 1988). Viral structural proteins, on the other hand, tend to be less conserved in evolution. Among positive-stranded RNA viruses a number of successful replicase classes has been recognized, which are all associated with a specific type of genome organization and replication strategy. Superfamilies of picornavirus-like, alphavirus-like, flavivirus-like (Strauss & Strauss, 1988; Goldbach & Wellink, 1988) and coronavirus-like (CVL; Snijder *et al.*, 1990*a*; den Boon *et al.*, 1991*b*; Snijder *et al.*, 1993) positive-stranded RNA viruses have been discriminated.

The comparison of the replication strategy and replicase properties of corona-, toro- and arteriviruses has clearly distinguished the CVL replicase module from its alpha-, picorna-, and flavivirus-like equivalents. The organization (two ORFs) and expression (using ribosomal frameshifting) of the gene as well as the arrangement of conserved domains within its product are unique. Moreover, the basic CVL genome organization (replicase gene-envelope protein genes-nucleocapsid protein gene) and expression strategy (using a 3'-coterminal nested set of five to eight mRNAs) are unparalleled (Fig. 2 and 4).

The concept of a CVL replicase dictating a unique replication strategy originated from the molecular characterization of the Berne torovirus, particularly from the detection of striking homologies between toro- and coronavirus replicase sequences (Snijder *et al.*, 1990*a*). Subsequently, the CVL group was expanded to include the arteriviruses, which also produce a 3'-coterminal nested set of mRNAs. The protein-coding regions of the genomes of EAV (den Boon *et al.*, 1991*b*), LDV (Kuo *et al.*, 1991; Godeny *et al.*, 1990, 1993; Chen *et al.*, 1993) and LV (Meulenberg *et al.*, 1993*a*; Conzelmann *et al.*, 1993) have now been sequenced completely. Although arterivirus replicase genes are considerably smaller (9.5 to 12 kb) than their toro- and coronavirus counterparts, they also contain two ORFs, the downstream ones of which are expressed by ribosomal frameshifting. In addition, the conserved domains first identified in toro- and coronavirus ORF 1b products have been identified in the same relative positions in the arterivirus replicase (see below and Fig. 4).

Unlike most other ORFs of CVL genomes, replicase ORF 1b is not expressed from a separate subgenomic RNA but by ribosomal frameshifting during the translation of genomic RNA (Brierley *et al.*, 1989; Snijder *et al.*, 1990*a*; den Boon *et al.*, 1991*b*). The ORF 1b product contains a number of domains which probably are indispensable during the early stages of viral replication, e.g. the synthesis of negative-stranded RNA and the onset of subgenomic RNA synthesis. The frameshifting mechanism and the RNA structures involved in this

process are remarkably conserved in all three CVL groups. In particular, its occurrence in the distantly related arteriviruses identifies translational frameshifting as an ancient and probably essential regulating step in replicase gene expression. Apparently the ratio between the levels of ORF 1a- and ORF 1b-derived proteins in infected cells has to be regulated. Using reporter gene constructs, frameshifting efficiencies in expression systems of 25 to 40%, 20 to 30% and 15 to 20% have been observed for corona-, toro- and arteriviruses, respectively (Brierley *et al.*, 1989; Bredenbeek *et al.*, 1990a; Snijder *et al.*, 1990a; den Boon *et al.*, 1991b).

### Conserved domains in CVL replicases

The ORF 1b protein sequence contains the two most highly conserved domains of the CVL replicase, the putative RNA-dependent RNA polymerase (Kamer & Argos, 1984; Poch *et al.*, 1989) and the NTP-dependent helicase (Hodgman, 1988; Gorbalenya & Koonin, 1989). The presence of these domains is not that surprising since both are found in all positive-stranded RNA viruses and attest that the entire group may have descended from the same ancestor. However, the percentages of identical amino acid residues in these domains, up to 50% between toro- and coronaviruses (Snijder *et al.*, 1990a) and about 30% between toro- and arteriviruses, indicate that the CVL viruses are more closely related to each other than to other groups of positive-stranded RNA viruses (Snijder *et al.*, 1990a; den Boon *et al.*, 1991b; Godeny *et al.*, 1993; E. J. Snijder & A. E. Gorbalenya, unpublished). Furthermore, only in the CVL replicases is the helicase domain located downstream of the polymerase motif, which carries another CVL trademark: the substitution of the classic GDD in the core of the motif by an SDD.

A third, 'unique' domain has been identified in the C-terminal region of the ORF 1b protein; its sequence is less conserved than in the two 'universal' domains discussed above. The significance of this domain is emphasized by the absence of a homologue in other viral replicases. It may be unique for CVL viruses and essential for their replication.

Ribosomal frameshifting during replicase gene expression leads to the production of an ORF 1a-ORF 1b fusion protein which is large in the case of arteriviruses (345K to 420K) (den Boon *et al.*, 1991b; Meulenberg *et al.*, 1993a; Godeny *et al.*, 1993), and very large (740K to 810K) in coronaviruses (Bournsnel *et al.*, 1987; Lee *et al.*, 1991; Herold *et al.*, 1993) and probably also in toroviruses. As described for other viral replicases, the CVL replicases are proteolytically cleaved into smaller

active units. A number of cleavage products has recently been detected in coronavirus- (Brierley *et al.*, 1990; Denison *et al.*, 1991, 1992) and arterivirus-infected (Snijder *et al.*, 1992; E. J. Snijder, A. L. M. Wassenaar & W. J. M. Spaan, unpublished) cells. The proteases involved in replicase protein processing are encoded by the ORF 1a sequence. Although experimental data on some of them have been obtained recently (Snijder *et al.*, 1992; Baker *et al.*, 1990; E. J. Snijder, A. L. M. Wassenaar & W. J. M. Spaan, unpublished), most protease domains are still putative and based on sequence comparison only (Gorbalenya *et al.*, 1989a; Lee *et al.*, 1991; den Boon *et al.*, 1991b; Meulenberg *et al.*, 1993a; Godeny *et al.*, 1993; Chen *et al.*, 1993; Herold *et al.*, 1993). Unfortunately, only the 5' and 3' sequences of BEV ORF 1a (approximately 1 kb from each end) have been determined so far and no putative proteases have been identified in these regions (Snijder *et al.*, 1990a, b).

From a comparative analysis of the available ORF 1a protein sequences a pattern emerges which may apply also to the torovirus ORF 1a protein. At about two-thirds down in their ORF 1a protein sequences, arteriviruses and coronaviruses possess a domain belonging to the protease superfamily which comprises the chymotrypsin-like and picornavirus 3C-like proteolytic enzymes (Gorbalenya *et al.*, 1989b; Bazan & Fletterick, 1988). Though the predicted catalytic nucleophile of the coronavirus protease (Cys) differs from that in the arterivirus proteases (Ser), this domain may still be a remnant from a common ancestor of both virus groups. The exchange of Cys for Ser at the active site of the enzyme is considered to be feasible (Gorbalenya *et al.*, 1989b; Bazan & Fletterick, 1988, 1990).

The replicase part that starts with the central region of ORF 1a and continues down to the 3' end of ORF 1b appears to be the core of the CVL replicase polyprotein: the well conserved domains (protease-polymerase-helicase-C-terminal 'unique' ORF 1b domain) are within this bracket, and only small insertions and deletions can be detected within the coronavirus or arterivirus groups.

The N-terminal half of the ORF 1a protein, on the other hand, is quite heterogeneous: the variability in the 5' half of ORF 1a is largely responsible for the size differences in corona- and arterivirus genomes. Genome sizes of 27.3 kb (human coronavirus 229E; Herold *et al.*, 1993), 27.6 kb [infectious bronchitis virus (IBV); Bournsnel *et al.*, 1987] and 31.3 kb (MHV; Lee *et al.*, 1991) have been reported for coronaviruses. In the arterivirus group, EAV possesses the smallest genome (12.7 kb; den Boon *et al.*, 1991b), followed by LDV (14.2 kb; Godeny *et al.*, 1993) and LV (15.1 kb; Meulenberg *et al.*, 1993a). A comparison between the coronavirus and arterivirus N-terminal ORF 1a protein sequences does not reveal

any significant similarities, and even within the coronavirus and arterivirus groups there is little resemblance in this region.

However, one striking observation can be made: both corona- and arteriviruses contain one or two domains in the N-terminal half of their ORF 1a proteins which are predicted to belong to the papain-like protease superfamily. Two of these have recently been characterized: in EAV (Snijder *et al.*, 1992) and MHV (Baker *et al.*, 1990), a papain-like protease is responsible for the production of an N-terminal ORF 1a cleavage product. The location of other predicted papain-like protease domains suggests that the N-terminal ORF 1a part contains a number of self-cleaving subunits. The chymotrypsin/3C-like protease would in this case be responsible for the processing of the core replicase. Therefore, determination of the remaining part of the BEV ORF 1a sequence would be very useful, particularly for a comparative analysis of the chymotrypsin/3C-like proteases of coronaviruses and other virus groups (Bazan & Fletterick, 1990).

## Recombination

In addition to divergence from a common ancestor, recombination is considered an important factor in RNA virus evolution (Strauss & Strauss, 1988; Goldbach & Wellink, 1988). Homologous recombination between highly similar RNA sequences has been found to occur during the multiplication of many plant and animal RNA viruses (for a recent review, see Lai, 1992). The MHV genome is known for its high homologous recombination frequency (Makino & Lai, 1989; Spaan *et al.*, 1988; Lai, 1990, 1992), which may be a consequence of the coronavirus replication strategy and of replicase properties. The variability of the coronavirus genome,

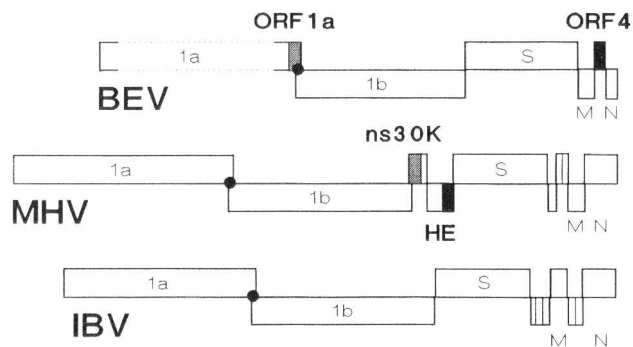


Fig. 5. Schematic representation of the genome organization of the torovirus BEV and the coronaviruses MHV and IBV. The replicase gene (1a and 1b) and the three basic structural genes S, M and N are indicated. Filled (ORF 4/HE) and cross-hatched (ORF 1a/ns30K) boxes indicate homologous sequences of BEV and MHV which are thought to be derived from independent recombination events.

particularly the changing number of reading frames in its 3' end (Fig. 5), can partly be explained by recombination. Sequence analysis of the MHV genome has produced an undisputed example of (non-homologous) recombination (Luytjes *et al.*, 1988): the haemagglutinin/esterase (HE) protein of the negative-stranded influenza virus type C (IVC) and the HE protein present in certain coronaviruses (MHV, bovine coronavirus, human coronavirus OC43) are related by common ancestry. Since the HE gene is absent in coronaviruses from other antigenic clusters (e.g. IBV; Fig. 5), recombination involving an IVC-like virus and an ancestral coronavirus was postulated: a heterologous recombination event is required to explain the presence of the same gene module in viruses of fundamentally different families (Luytjes *et al.*, 1988).

It is too early to make general statements about the genetic variability of toroviruses, but recombination has probably played an important role also in their evolution. Two non-homologous RNA recombinations have been pre-supposed to explain similarities which were detected in the BEV genome (Snijder *et al.*, 1990*c*). In addition to the evidence for common ancestry presented above, these recombinations associate toroviruses with coronaviruses, particularly with the antigenic cluster to which MHV belongs.

The first sequence homology involves BEV ORF 4 and the coronavirus HE gene (Fig. 5). The 142 amino acid sequence of BEV ORF 4 shares 30 to 35% identical residues with the C-terminal parts of both the IVC and the coronavirus HE sequences. The 5' two-thirds of the HE gene, which contain the catalytic centre of the acetylsterase activity (Vlasak *et al.*, 1989), are lacking in BEV; this has led to the hypothesis that BEV ORF 4 is a pseudogene (Snijder *et al.*, 1991*a*). The ORF 4 sequence does not indicate a closer relationship to the homologous gene of either coronaviruses or IVC. Although independent recombination events cannot be excluded, the presence of the same C-terminal adaptation, a membrane anchor which is not present in the IVC HE protein (Luytjes *et al.*, 1988), lends some credibility to a recombination between ancestral toro- and coronaviruses. Thus either a coronavirus or a torovirus may have been involved in the initial recombination with an IVC ancestor.

The second non-homologous recombination during BEV evolution is quite similar to the case of ORF 4. The C terminus of the BEV ORF 1a protein contains 31 to 36% identical amino acid residues when compared to the N-terminal 190 amino acids of the non-structural coronavirus protein ns30/32K (Bredenbeek *et al.*, 1990*b*; Cox *et al.*, 1989); again this protein is found only in the coronaviruses closely related to MHV (Snijder *et al.*, 1991*a*), where it is located between the polymerase and



HE genes (Fig. 5). A sequence related to the 5' two-thirds of this coronavirus non-structural gene, which is expressed from a separate subgenomic mRNA, appears to have been integrated into ORF 1a of BEV and is now expressed as part of its replicase.

It is clear that the sequence similarities described above must be the result of non-homologous RNA recombination events. Despite the remarkable coincidence, however, they do not imply that toroviruses are more closely related to the MHV cluster than to other coronaviruses. First, the homologous sequences are located at different positions in the genome (Fig. 5). Second, the BEV replicase is not more closely related to the MHV replicase than to that of IBV, for example. Third, the sequence similarities between the S and M proteins, respectively, of corona- and toroviruses are so low (see above) that the high similarities (> 30% identical residues) discussed above can lead only to this conclusion: divergence between BEV ORF 4 and the coronavirus HE gene and between the BEV ORF 1a fragment and the coronavirus ns30/32K gene must have occurred more recently than the divergence of the other homologous genes of toro- and coronaviruses.

Considering the fact that several extant representatives of both virus groups cause enteric infections, direct recombination between toro- and coronaviruses during co-infection of the same cell is feasible. However, the involvement of a third party of viral or cellular origin cannot be excluded.

## Evolution and taxonomy

Features shared by members of the CVL superfamily include the basic genome organization, the production of 3'-coterminal nested sets of mRNAs, and the presence of homologous replicase domains (Fig. 4). Noticeable differences are dissimilar N proteins and nucleocapsid architecture, the (probable) absence of a common 5' leader sequence in the BEV mRNAs, and the much smaller genome size and absence of a large spike glycoprotein in arteriviruses.

Analysis of the BEV genome has revealed traces of all the processes that make RNA virus evolution such a fascinating field. First, there are the unambiguous primary protein sequence homologies in the replicase. They are the most convincing evidence for common ancestry and for the divergent evolution which has resulted in the present-day replicase genes of toro-, corona- and arteriviruses. Second, the same common ancestry probably connects the BEV and coronavirus S and M proteins. Here primary sequence similarities are no longer detectable, but the (predicted) structural characteristics of the proteins and their linkage to homologous replicase genes are highly suggestive. The

observations made for the S and M proteins are in agreement with the general perception that structural proteins evolve at a higher rate than non-structural ones. Third, the recombination events discussed in the previous paragraphs underline the role of recombination in RNA virus evolution in general, and in the evolution of corona- and toroviruses in particular. Fourth, CVL genomes may show the results of a process for which the name 'modular evolution' has been coined (Strauss & Strauss, 1988; Goldbach & Wellink, 1988; Zimmern, 1987): the exchange by recombination of complete genes or gene sets (modules). Modular evolution provides a plausible explanation for the presence of a CVL replicase and for the definitely coronavirus-unlike set of structural proteins in the arterivirus group (de Vries *et al.*, 1992). However, the BEV N protein is also difficult to account for: its size, sequence and quaternary structure in association with the genome are unrelated to both corona- and arteriviruses.

A meaningful classification of the members of the CVL superfamily clearly requires four hierarchical levels (Fig. 6). The coronavirus and torovirus species have now been classified into two genera within the family *Coronaviridae*. The obvious evolutionary link of this family to the arteriviruses would be reflected most accurately by accommodating the present genus arterivirus in a family of its own and by establishing an order (to replace the 'superfamily') comprising the *Coronaviridae* and *Arteriviridae* families.

The above implies that the weight of traditional

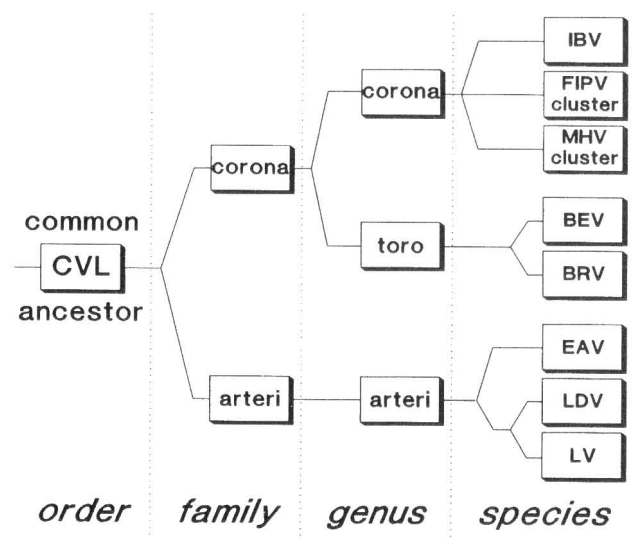


Fig. 6. Pedigree of the CVL superfamily. The distances between the various CVL viruses and groups have (largely) been chosen arbitrarily. The four hierarchical taxonomic levels which can be discerned are indicated. FIPV, feline infectious peritonitis virus. For simplicity, the coronavirus clusters which include MHV and FIPV are shown as a single species.

taxonomic criteria, e.g. of virion and nucleocapsid structure, will necessarily diminish. At the same time the importance of 'modern' virus characteristics like genome organization, replication strategy and sequence homologies will increase. Thus molecular virology has provided the basis for the rather drastic measure of reassignment and relegation of *Togaviridae* family members. In most cases, however, the results of traditional and modern taxonomic initiatives will be synergistic rather than opposed; this, too, is exemplified by the data discussed in this review. Both practices of classification lead to the same inevitable conclusions: the three virus groups are (at least partially) derived from a common ancestor, the corona- and toroviruses being more closely related to each other than to the arteriviruses (Fig. 6).

Evidently the continuous nature of sequence variation will pose a taxonomic problem. What percentage of identical amino acids is required to define a genus or family? It is even more difficult to integrate the results of genetic recombination, the exchange of entire modules of information, into virus systematics. On the other hand, these two issues confront virologists most cogently with the genetic basis of the problem. Virus evolution is also governed by heredity, and the processes of mutation, recombination and selection will have to be accommodated in future virus taxonomy.

From the Virology Division at the Utrecht Veterinary Faculty a large number of people contributed (one way or the other) to the Berne virus project. We especially acknowledge the valuable contributions of Joke Ederveen, Johan den Boon and Willy Spaan. We are grateful to Marianne Weiss for permission to use the electron microscope pictures in Fig. 1. We would like to thank the authors of the following references for sharing unpublished data and preprints of their publications with us: Chen *et al.* (1993), Godeny *et al.* (1993), Herold *et al.* (1993) and Meulenberg *et al.* (1993*b*). We thank Thea Kuipers and Mareen de Best for assistance in preparation of the manuscript. A substantial part of the Berne virus research described in this review was supported by the Division for Health Research TNO in cooperation with the Dutch Organization for Scientific Research (NWO/Medigon; project 900-502-081).

## References

- BAKER, S. C., LA MONICA, N., SHIEH, C.-K. & LAI, M. M. C. (1990). Murine coronavirus gene 1 polyprotein contains an autoproteolytic activity. *Advances in Experimental Medicine and Biology* **286**, 283–289.
- BAZAN, J. F. & FLETTERICK, R. J. (1988). Viral cysteine proteases are homologous to the trypsin-like family of serine proteases: structural and functional implications. *Proceedings of the National Academy of Sciences, U.S.A.* **85**, 7872–7876.
- BAZAN, J. F. & FLETTERICK, R. J. (1990). Structural and catalytic models of trypsin-like viral protease. *Seminars in Virology* **1**, 311–322.
- BEARDS, G. M., HALL, C., GREEN, J., FLEWETT, T. H., LAMOULIATTE, F. & DU PASQUIER, P. (1984). An enveloped virus in stools of children and adults with gastroenteritis that resembles the Breda virus of calves. *Lancet* **ii**, 1050–1052.
- BEARDS, G. M., BROWN, D. W. G., GREEN, J. & FLEWETT, T. H. (1986). Preliminary characterisation of torovirus-like particles of humans: comparison with Berne virus of horses and Breda virus of calves. *Journal of Medical Virology* **20**, 67–78.
- BOURNSELL, M. E. G., BROWN, T. D. K., FOULDS, I. J., GREEN, P. F., TOMLEY, F. M. & BINNS, M. M. (1987). Completion of the sequence of the genome of the coronavirus avian infectious bronchitis virus. *Journal of General Virology* **68**, 57–77.
- BREDENBEEK, P. J., PACHUK, C. J., NOTEN, J. F. H., CHARITÉ, J., LUYTJES, W., WEISS, S. R. & SPAAN, W. J. M. (1990*a*). The primary structure and expression of the second open reading frame of the polymerase gene of the coronavirus MHV-A59. *Nucleic Acids Research* **18**, 1825–1832.
- BREDENBEEK, P. J., NOTEN, J. F. H., HORZINEK, M. C. & SPAAN, W. J. M. (1990*b*). Identification and stability of a 30 Kd non-structural protein encoded by mRNA 2 of mouse hepatitis virus in infected cells. *Virology* **175**, 303–306.
- BRIERLEY, I., DIGGARD, P. & INGLIS, S. (1989). Characterization of an efficient coronavirus ribosomal frameshifting signal: requirement for an RNA pseudoknot. *Cell* **57**, 537–547.
- BRIERLEY, I., BOURNELL, M. E. G., BINNS, M. M., BILIMORIA, B., ROLLEY, N. J., BROWN, T. D. K. & INGLIS, S. C. (1990). Products of the polymerase-encoding region of the coronavirus IBV. *Advances in Experimental Medicine and Biology* **286**, 275–282.
- BROWN, D. W. G., BEARDS, G. M. & FLEWETT, T. H. (1987). Detection of Breda virus antigen and antibody in humans and animals by enzyme immunoassay. *Journal of Clinical Microbiology* **25**, 637–640.
- CHEN, Z., KUO, L., ROWLAND, R. R. R., EVEN, C., FAABERG, K. S. & PLAGEMANN, P. G. W. (1993). Sequence of 3' end of genome and of 5' end of open reading frame 1a of lactate dehydrogenase-elevating virus and common junction motifs between 5' leader and bodies of seven subgenomic mRNAs. *Journal of General Virology* **74**, 643–660.
- COLLINS, A. R., KNOBLER, R. L., POWELL, H. & BUCHMEIER, M. J. (1982). Monoclonal antibodies to murine hepatitis virus-4 (strain JHM) define the viral glycoprotein responsible for attachment and cell fusion. *Virology* **119**, 358–371.
- CONZELMANN, K. K., VISSER, N., VAN WOENSEL, P. & THIEL, H. J. (1993). Molecular characterization of porcine reproductive and respiratory syndrome virus (PRRSV), a member of the arterivirus group. *Virology* **193**, 329–339.
- COX, G. J., PARKER, M. D. & BABIUK, L. A. (1989). The sequence of cDNA of bovine coronavirus 32K nonstructural gene. *Nucleic Acids Research* **17**, 5847.
- DE GROOT, R. J., LUYTJES, W., HORZINEK, M. C., VAN DER ZEIJST, B. A. M., SPAAN, W. J. M. & LENSTRA, J. A. (1987). Evidence for a coiled-coil structure in the spike proteins of coronaviruses. *Journal of Molecular Biology* **196**, 963–966.
- DEN BOON, J. A., SNIJDER, E. J., KRIJNSE LOCKER, J., HORZINEK, M. C. & ROTTIER, P. J. M. (1991*a*). Another triple-spanning envelope protein among intracellularly budding RNA viruses: the torovirus E protein. *Virology* **182**, 655–663.
- DEN BOON, J. A., SNIJDER, E. J., CHIRNSIDE, E. D., DE VRIES, A. A. F., HORZINEK, M. C. & SPAAN, W. J. M. (1991*b*). Equine arteritis virus is not a togavirus but belongs to the coronavirus-like superfamily. *Journal of Virology* **65**, 2910–2920.
- DENISON, M. R., ZOLTICK, P. W., LEIBOWITZ, J. L., PACHUK, C. J. & WEISS, R. (1991). Identification of polypeptides encoded in open reading frame 1b of the putative polymerase gene of the murine coronavirus mouse hepatitis virus A59. *Journal of Virology* **65**, 3067–3082.
- DENISON, M. R., ZOLTICK, P. W., HUGHES, A., GIANGRECO, B., OLSEN, A. L., PERLMAN, S., LEIBOWITZ, J. L. & WEISS, S. R. (1992). Intracellular processing of the N-terminal ORF1a proteins of the coronavirus MHV-A59 requires multiple proteolytic events. *Virology* **189**, 274–284.
- DE VRIES, A. A. F., CHIRNSIDE, E. D., BREDENBEEK, P. J., GRAVESTIEN, L. A., HORZINEK, M. C. & SPAAN, W. J. M. (1990). All subgenomic mRNAs of equine arteritis virus contain a common leader sequence. *Nucleic Acids Research* **18**, 3241–3247.
- DE VRIES, A. A. F., CHIRNSIDE, E. D., HORZINEK, M. C. & ROTTIER, P. J. M. (1992). Structural proteins of equine arteritis virus. *Journal of Virology* **66**, 6294–6303.
- DUBOIS-DALCO, M. E., DOLLER, E. W., HASPEL, M. V. & HOLMES, K. V. (1982). Cell tropism and expression of mouse hepatitis virus (MHV) in mouse spinal cord cultures. *Virology* **119**, 317–331.
- FAGERLAND, J. A., POHLENZ, J. F. L. & WOODE, G. N. (1986). A

- morphological study of the replication of Breda virus (proposed family *Toroviridae*) in bovine intestinal cells. *Journal of General Virology* **67**, 1293–1304.
- GODENY, E. K., SPEICHER, D. W. & BRINTON, M. A. (1990). Map location of lactate dehydrogenase-elevating virus (LDV) capsid protein (Vp1) gene. *Virology* **177**, 768–771.
- GODENY, E. K., CHEN, L., KUMAR, S. N., METHVEN, S. L., KOONIN, E. V. & BRINTON, M. A. (1993). Complete genomic sequence and phylogenetic analysis of the lactate dehydrogenase-elevating virus. *Virology* **194**, 585–596.
- GOLDBACH, R. & WELLINK, J. (1988). Evolution of plus-strand RNA viruses. *Intervirology* **29**, 260–267.
- GORBALENYA, A. E. & KOONIN, E. V. (1989). Viral proteins containing the purine NTP-binding sequence pattern. *Nucleic Acids Research* **17**, 8413–8440.
- GORBALENYA, A. E., KOONIN, E. V., DONCHENKO, A. P. & BLINOV, V. M. (1989a). Coronavirus genome: prediction of putative functional domains in the non-structural polyprotein by comparative amino acid sequence analysis. *Nucleic Acids Research* **17**, 4847–4861.
- GORBALENYA, A. E., DONCHENKO, A. P., BLINOV, V. M. & KOONIN, E. V. (1989b). Cysteine proteases of positive strand RNA viruses and chymotrypsin-like serine proteases: a distinct protein superfamily with a common structural fold. *FEBS Letters* **243**, 103–114.
- HEROLD, J., RAABE, T., SCHIELLE-PRINZ, B. & SIDDELL, S. G. (1993). Nucleotide sequence of the human coronavirus 229E RNA polymerase locus. *Virology* **195**, 680–691.
- HODGMAN, T. C. (1988). A new superfamily of replicative proteins. *Nature, London* **333**, 22–23.
- HOLMES, K. V., DOLLER, E. W. & STURMAN, L. S. (1981). Tunicamycin resistant glycosylation of a coronavirus glycoprotein: demonstration of a novel type of viral glycoprotein. *Virology* **115**, 334–344.
- HORZINEK, M. C. & WEISS, M. (1984). *Toroviridae*: a taxonomic proposal. *Zentralblatt für Veterinärmedizin [B]* **31**, 649–659.
- HORZINEK, M. C., WEISS, M. & EDERVEEN, J. (1984). Berne virus is not 'coronavirus-like'. *Journal of General Virology* **65**, 645–649.
- HORZINEK, M. C., EDERVEEN, J. & WEISS, M. (1985). The nucleocapsid of Berne virus. *Journal of General Virology* **66**, 1287–1296.
- HORZINEK, M. C., EDERVEEN, J., KAEFFER, B., DE BOER, D. & WEISS, M. (1986). The peplomers of Berne virus. *Journal of General Virology* **67**, 2475–2483.
- HORZINEK, M. C., FLEWETT, T. H., SAIF, L. J., SPAAN, W. J. M., WEISS, M. & WOODE, G. N. (1987). A new family of vertebrate viruses: *Toroviridae*. *Intervirology* **27**, 17–24.
- JEONG, Y. S. & MAKINO, S. (1992). Mechanism of coronavirus transcription – duration of primary transcription initiation activity and effects of subgenomic RNA transcription on RNA replication. *Journal of Virology* **66**, 3339–3346.
- KAEFFER, B., VAN KOOIJEN, P., EDERVEEN, J., VAN EDEN, W. & HORZINEK, M. C. (1989). Properties of monoclonal antibodies against Berne virus (*Toroviridae*). *American Journal of Veterinary Research* **50**, 1131–1137.
- KAMER, G. & ARGOS, P. (1984). Primary structural comparison of RNA-dependent polymerases from plant, animal and bacterial viruses. *Nucleic Acids Research* **18**, 7269–7282.
- KOOPMANS, M. & HORZINEK, M. C. (1993). Toroviruses of animals and humans: a review. *Advances in Virus Research* (in press).
- KOOPMANS, M., EDERVEEN, J., WOODE, G. N. & HORZINEK, M. C. (1986). Surface proteins of Breda virus. *American Journal of Veterinary Research* **47**, 1896–1900.
- KOOPMANS, M., VAN DEN BOOM, U., WOODE, G. N. & HORZINEK, M. C. (1989). Seroepidemiology of Breda virus in cattle using ELISA. *Veterinary Microbiology* **19**, 233–243.
- KOOPMANS, M., SNIJDER, E. J. & HORZINEK, M. C. (1991). cDNA probes for the diagnosis of bovine torovirus (Breda virus) infection. *Journal of Clinical Microbiology* **29**, 493–497.
- KUO, L. L., HARTY, J. T., ERICKSON, L., PALMER, G. A. & PLAGEMANN, P. G. W. (1991). A nested set of eight RNAs is formed in macrophages infected with lactate dehydrogenase-elevating virus. *Journal of Virology* **65**, 5118–5123.
- LAI, M. M. C. (1990). Coronavirus: organization, replication and expression of genome. *Annual Review of Microbiology* **44**, 303–333.
- LAI, M. M. C. (1992). RNA recombination in animal and plant viruses. *Microbiological Reviews* **56**, 61–79.
- LEE, H. J., SHIEH, C. K., GORBALENYA, A. E., KOONIN, E. V., LAMONICA, N., TULER, J., BAGDZHADZHUYAN, A. & LAI, M. M. C. (1991). The complete sequence (22 kilobases) of murine coronavirus gene-1 encoding the putative proteases and RNA polymerase. *Virology* **180**, 567–582.
- LUYTJES, W., BREDENBEEK, P. J., NOTEN, J. F. H., HORZINEK, M. C. & SPAAN, W. J. M. (1988). Sequence of mouse hepatitis virus A59 mRNA2: indications for RNA recombination between coronaviruses and influenza C virus. *Virology* **164**, 415–422.
- MAKINO, S. & LAI, M. M. C. (1989). High-frequency leader sequence switching during coronavirus defective interfering RNA replication. *Journal of Virology* **63**, 5285–5292.
- MEULENBERG, J. J. M., HULST, M. M., DE MEIJER, E. J., MOONEN, P. L. J. M., DEN BESTEN, A., DE KLUYVER, E. P., WENSVOORT, G. & MOORMANN, R. J. M. (1993a). Lelystad virus, the causative agent of porcine epidemic abortion and respiratory syndrome (PEARS), is related to LDV and EAV. *Virology* **192**, 62–72.
- MEULENBERG, J. J. M., DE MEIJER, E. & MOORMANN, R. J. M. (1993b). Subgenomic RNAs of Lelystad virus contain a conserved leader-body junction sequence. *Journal of General Virology* **74**, 1697–1701.
- POCH, O., SAUVAGET, L., DELARUE, M. & TORDO, N. (1989). Identification of four conserved motifs among the RNA dependent polymerase encoding elements. *EMBO Journal* **8**, 3867–3874.
- PRINGLE, C. R. (1992). Committee pursues medley of virus taxonomic issues. *ASM News* **58**, 475–476.
- ROTTIER, P. J. M., HORZINEK, M. C. & VAN DER ZEIJST, B. A. M. (1981). Viral protein synthesis in mouse hepatitis virus strain A59-infected cells: effect of tunicamycin. *Journal of Virology* **40**, 350–357.
- SAWICKI, S. G. & SAWICKI, D. L. (1990). Subgenomic mouse hepatitis virus replicative intermediates function in RNA synthesis. *Journal of Virology* **64**, 1050–1056.
- SEANAYAKE, S. D., HOFMANN, M. A., MAKI, J. L. & BRIAN, D. A. (1992). The nucleocapsid protein gene of bovine coronavirus is bicistronic. *Journal of Virology* **66**, 5277–5283.
- SETHNA, P. B., HUNG, S. L. & BRIAN, D. A. (1989). Coronavirus subgenomic minus-strand RNAs and the potential for mRNA replicons. *Proceedings of the National Academy of Sciences, U.S.A.* **86**, 5626–5630.
- SNIJDER, E. J., EDERVEEN, J., SPAAN, W. J. M., WEISS, M. & HORZINEK, M. C. (1988). Characterization of Berne virus genomic and messenger RNAs. *Journal of General Virology* **69**, 2135–2144.
- SNIJDER, E. J., DEN BOON, J. A., SPAAN, W. J. M., VERJANS, G. M. G. M. & HORZINEK, M. C. (1989). Identification and primary structure of the gene encoding the Berne virus nucleocapsid protein. *Journal of General Virology* **70**, 3363–3370.
- SNIJDER, E. J., DEN BOON, J. A., BREDENBEEK, P. J., HORZINEK, M. C., RIJNBRAND, R. & SPAAN, W. J. M. (1990a). The carboxyl-terminal part of the putative Berne virus polymerase is expressed by ribosomal frameshifting and contains sequence motifs which indicate that toro- and coronaviruses are evolutionarily related. *Nucleic Acids Research* **18**, 4535–4542.
- SNIJDER, E. J., DEN BOON, J. A., SPAAN, W. J. M., WEISS, M. & HORZINEK, M. C. (1990b). Primary structure and post-translational processing of the Berne virus peplomer protein. *Virology* **178**, 355–363.
- SNIJDER, E. J., HORZINEK, M. C. & SPAAN, W. J. M. (1990c). A 3'-coterminal nested set of independently transcribed messenger RNAs is generated during Berne virus replication. *Journal of Virology* **64**, 331–338.
- SNIJDER, E. J., DEN BOON, J. A., HORZINEK, M. C. & SPAAN, W. J. M. (1991a). Comparison of the genome organization of toro- and coronaviruses: both divergence from a common ancestor and RNA recombination have played a role in Berne virus evolution. *Virology* **180**, 448–452.
- SNIJDER, E. J., DEN BOON, J. A., HORZINEK, M. C. & SPAAN, W. J. M. (1991b). Characterization of defective interfering RNAs of Berne virus. *Journal of General Virology* **72**, 1635–1643.
- SNIJDER, E. J., WASSENAAR, A. L. M. & SPAAN, W. J. M. (1992). The 5' end of the equine arteritis virus genome encodes a papainlike cysteine protease. *Journal of Virology* **66**, 7040–7048.

- SNIJDER, E. J., HORZINEK, M. C. & SPAAN, W. J. M. (1993). The coronavirus-like superfamily. In *Coronaviruses: Molecular Biology and Pathogenesis*. Edited by H. Laude & J. F. Vautherot. New York: Plenum Publishing Corporation (in press).
- SPAAN, W., CAVANAGH, D. & HORZINEK, M. C. (1988). Coronaviruses: structure and genome expression. *Journal of General Virology* **69**, 2939–2952.
- STRAUSS, J. H. & STRAUSS, E. G. (1988). Evolution of RNA viruses. *Annual Review of Microbiology* **42**, 657–683.
- VENNEMA, H., HEIJNEN, L., ZIJDERVELD, A., HORZINEK, M. C. & SPAAN, W. J. M. (1990). Intracellular transport of recombinant coronavirus spike proteins – implications for virus assembly. *Journal of Virology* **64**, 339–346.
- VLASAK, R., MUSTER, T., LAURO, A. M., POWERS, J. & PALESE, P. (1989). Influenza C virus esterase: analysis of catalytic site, inhibition, and possible function. *Journal of Virology* **63**, 2056–2062.
- WEISS, M. & HORZINEK, M. C. (1986*a*). Resistance of Berne virus to physical and chemical treatment. *Veterinary Microbiology* **11**, 41–49.
- WEISS, M. & HORZINEK, M. C. (1986*b*). Morphogenesis of Berne virus (proposed family Toroviridae). *Journal of General Virology* **67**, 1305–1314.
- WEISS, M. & HORZINEK, M. C. (1987). The proposed family *Toroviridae*: agents of enteric infections. *Archives of Virology* **92**, 1–15.
- WEISS, M., STECK, F. & HORZINEK, M. C. (1983). Purification and partial characterization of a new enveloped RNA virus (Berne virus). *Journal of General Virology* **64**, 1849–1858.
- WEISS, M., STECK, F. & KADERLI, R. (1984). Antibodies to Berne virus in horses and other animals. *Veterinary Microbiology* **9**, 523–531.
- WOODE, G. N., REED, D. E., RUNNELS, P. L., HERRIG, M. A. & HILL, H. T. (1982). Studies with an unclassified virus isolated from diarrhoeic calves. *Veterinary Microbiology* **7**, 221–240.
- WOODE, G. N., SAIF, L. J., QUESADA, M., WINAND, N. J., POHLENZ, J. F. L. & GOURLEY, N. (1985). Comparative studies on three isolates of Breda virus of calves. *American Journal of Veterinary Research* **46**, 1003–1010.
- YOKOMORI, K., BANNER, L. R. & LAI, M. M. C. (1992). Coronavirus messenger RNA transcription – UV light transcriptional mapping studies suggest an early requirement for a genomic-length template source. *Journal of Virology* **66**, 4671–4678.
- ZIMMERN, D. (1987). Evolution of RNA viruses. In *RNA Genetics*, vol. 2, pp. 211–240. Edited by J. J. Holland, E. Domingo & P. Ahlquist. Boca Raton: CRC Press.