FAMILY   CORONAVIRIDAE

Taxonomic structure of the family

Family	*Coronaviridae*
Subfamily	*Coronavirinae*
Genus	*Alphacoronavirus*
Genus	*Betacoronavirus*
Genus	*Gammacoronavirus*
Subfamily	*Torovirinae*
Genus	*Torovirus*
Genus	*Bafinivirus*

Distinguishing features

The members of the family *Coronaviridae*, a monophyletic cluster in the order *Nidovirales*, are enveloped, positive stranded RNA viruses of three classes of vertebrates: mammals (corona -and toroviruses), birds (coronaviruses) and fish (bafiniviruses). Virions are spherical, 120–160 nm across (*Coronavirinae*), bacilliform, 170–200 × 75–88 nm (*Bafinivirus*) or found as a mixture of both, with bacilliform particles characteristically bent into crescents (*Torovirus*). The particles are typically decorated with large, club- or petal-shaped surface projections (the "peplomers" or "spikes"), which in electron micrographs of spherical particles create an image reminiscent of the solar corona. This inspired the name of the "true" coronaviruses (now grouped in the subfamily *Coronavirinae*), which was later adopted for the whole family. Nucleocapsids are helical and can be released from the virion by treatment with detergents. Whereas the coronavirus nucleocapsid appears to be loosely-wound, those of the *Torovirinae* are distinctively tubular.

In terms of genome size and genetic complexity, the *Coronaviridae* are the largest RNA viruses identified so far, rivaled only by the okaviruses, large nidoviruses of invertebrates assigned to the family *Roniviridae*. Replication has been studied in detail only for coronaviruses, but the limited data available for toro- and bafiniviruses suggest that the latter viruses use essentially similar strategies. Virions attach to dedicated host cell surface receptors via their spikes (Table 1) and release their genome into the target cell via fusion of the viral envelope with the plasma membrane and/or the limiting membrane of an endocytic vesicle. The entire replication cycle takes place in the cytoplasm and involves the production of full-length and subgenome-sized (sg) minus-strand RNA intermediates with the viral genome serving both as mRNA for the replicase polyproteins and as a template for minus-strand synthesis. RNA synthesis is catalyzed by an as yet poorly characterized replication–transcription complex, composed of viral and host proteins and associated (at least in coronaviruses) with an interconnected network of modified intracellular membranes and double-membrane vesicles that are presumably endoplasmic reticulum (ER)-derived.

The genome contains multiple ORFs. Its 5'-most two-thirds are occupied by the replicase gene, which is comprised of two overlapping ORFs called 1a and 1b (Figure 1). The replicase gene is translated to produce polyprotein pp1a and, subject to programmed −1 ribosomal frameshifting, a C-terminally extended product, pp1ab. The polyproteins are co- and post-translationally processed by a set of virus-encoded proteinases and, thus, are not detectable as full-length proteins in virus-infected cells. The N-termini of pp1a and pp1ab are processed by one or two papain-like proteinases, whereas the C-terminal half of coronavirus pp1a and the ORF1b-encoded part of pp1ab are cleaved at 11 well-conserved sites by the main proteinase (M^{pro} or $3CL^{pro}$), a nidovirus-wide conserved enzyme with a chymotrypsin-like fold, a poliovirus 3C proteinase-like substrate specificity and either a serine (torovirus, bafinivirus) or a cysteine (coronavirus) as active site nucleophile. In coronaviruses, proteolytic processing results in the production of 15 (in viruses belonging to the species *Avian coronavirus*) or 16 mature products, commonly referred to as non-structural proteins (nsp's) and numbered according to their position – from N- to C-terminus – in the viral polyproteins (Figure 1). Many nsp's are unique enzymes involved in one or more essential step(s) in viral replication. Others appear to be exclusively involved in virus–host interactions (including immune evasion) and are dispensable for virus propagation *in vitro* (Table 2). Polyprotein processing in toro- and bafiniviruses has not been studied in detail.

Table 1: Coronavirus primary attachment factors and receptors

Virus species	Host	Attachment factor	Main receptor
Alphacoronavirus 1			
Canine coronavirus type I	Dog		?
Canine coronavirus type II	Dog		APN
Feline coronavirus type I	Cat		?
Feline coronavirus type II	Cat		APN
Transmissible gastroenteritis virus	Pig	Sialic acid	APN
Human coronavirus 229E	Human		APN
Human coronavirus NL63	Human		ACE2
*Betacoronaviru*s 1			
Bovine coronavirus	Cow		9-O-Ac Sia?
Equine coronavirus	Horse		9-O-Ac Sia?
Human coronavirus OC43	Human		9-O-Ac Sia?
Porcine hemagglutinating encephalomyelitis virus	Pig		9-O-Ac Sia?
Murine coronavirus[*]	Mouse	4-O- or 9-O-Ac Sia	CEACAM1a
SARS-related coronavirus	Human		ACE2

Abbreviations: APN, aminopeptidase N; ACE2, angiotensin-converting enzyme 2; CEACAM1a, carcinoembryonic antigen adhesion molecule 1.

[*]Murine coronaviruses occur in two types that use either 4- or 9-*O*-acetylated sialic acid (*O*-Ac Sia) as primary attachment factor and CEACAM1a as main receptor.

The 3′-proximal genes (3 in bafiniviruses and up to at least 12 in some coronaviruses) code for the structural proteins and, in the case of coronaviruses, a variable number of "accessory" or "niche-specific" proteins. These genes are expressed – as is typical for nidoviruses – from a 3′-coterminal nested set of sg mRNAs that are thought to be transcribed not from the full-length minus-strand anti-genome, but from a mirror copy set of sg minus-strand templates.

Members of the family *Coronaviridae* all seem to share two envelope protein species, the membrane (M) and spike (S) proteins. Similarities in size, predicted structures and presumed function(s) suggest a common ancestry, and the remote, but significant sequence similarities observed for toro-, bafini- and (to lesser extent) coronavirus S proteins lend further support to this view. Presumably, progenitors of the S and M proteins were encoded in the last common ancestor of the *Corona-* and *Torovirinae* lineages. Virus assembly involves budding of preformed nucleocapsids at membranes of the endoplasmic reticulum and early Golgi compartment and the completed virions are released via the exocytotic pathway. Nidovirus replication is discussed in more detail in paragraphs below and also in Chapter *Nidovirales*.

All members of the *Coronaviridae* family share the following characteristics:

- Virions: enveloped and decorated with large (15–20 nm) surface projections.
- Nucleocapsid: helical, comprised of genome and multiple copies of a single basic phosphoprotein species (N).
- Envelope: contains a variable number of viral membrane protein species, two of which seem to be conserved family-wide and are essential for virion morphogenesis and/or infectivity (at least in coronaviruses):
  - a 200- to 250-aa triple-spanning $N^{exo}C^{endo}$ integral membrane protein M
  - an extensively N-glycosylated, 1100- to 1600-aa class I fusion protein S which forms peplomers.
- Genome: positive sense RNA, linear, unimolecular, infectious, 26–32 kb in length, capped, polyadenylated and structurally polycistronic.
- General genome organization: 5′-UTR-replicase-S-M-N-UTR-3′ (genes named after their product), with the genome functioning as mRNA for the replicase gene.

Pos. ssRNA

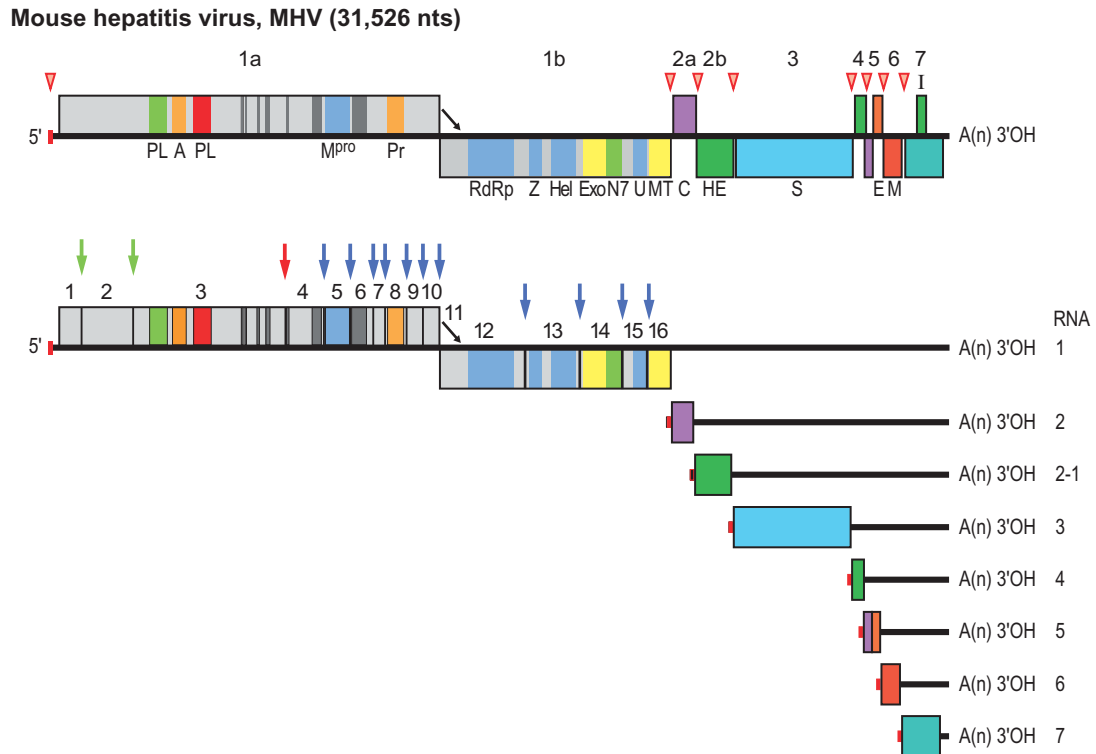**Mouse hepatitis virus, MHV (31,526 nts)**

Figure 1: Coronavirus genome organization and expression. (Upper panel) Schematic representation of the genome of mouse hepatitis virus (MHV) shown as an example. ORFs are represented by boxes, indicated by number (above) and encoded protein (acronyms below). Regions encoding key domains in replicase polyproteins pp1a and pp1ab are colour-coded with hydrophobic segments shown in dark grey. The 5′-leader sequence is depicted by a small red box. The arrow between ORF 1a and 1b represents the ribosomal frameshifting site. The poly(A) tail is indicated by "A(n)". Red arrowheads indicate the locations of transcription-regulating sequences (TRSs). PL (green) papain-like proteinase 1 (PL1^{pro}); PL (red), papain-like proteinase 2 (PL2^{pro}); A, ADP-ribose-1″phosphatase (macrodomain); M^{pro}, 3C-like main protease; Pr, noncanonical RNA-dependent RNA polymerase, putative primase; RdRp, RNA-dependent RNA polymerase; Z, zinc-binding domain; Hel, helicase domain; Exo, 3′ to-5′ exoribonuclease domain; N7, guanine-N7-methyltransferase; U, nidoviral uridylate-specific endoribonuclease (NendoU); MT, ribose-2′-*O*-methyltransferase domain; HE, hemagglutinin-esterase; S, spike protein; E, envelope protein; M, membrane protein, N, nucleocapsid protein; I, internal ORF. (Lower panel) Processing of the replicase polyproteins and structural relationship between the genomic RNA and subgenomic mRNAs of coronaviruses. Arrows indicate cleavage sites for PL1^{pro} (green), PL2^{pro} (red) and M^{pro} (blue). The locations of the non-structural proteins (nsp's) are indicated by their number (see also Table 2). mRNA species are numbered as by convention on the basis of their size, from large to small, with the genome designated as RNA1. For the sg mRNAs only ORF(s) that are translated are shown.

- Replicase gene: comprised of overlapping ORFs 1a and 1b that code for two huge polyproteins, pp1a and pp1ab, production of the latter requiring a programmed −1 ribosomal frameshift; pp1a and pp1ab are processed autoproteolytically.
- ORFs downstream of the replicase gene: expression from a 3′ co-terminal nested set of two or more subgenomic mRNAs that are capped and polyadenylated.
- Morphogenesis: virion assembly through budding of preformed nucleocapsids at smooth intracellular membranes of endoplasmic reticulum/early Golgi compartments.

The replicase polyproteins of the *Coronaviridae* comprise a number of characteristic domains arranged in a conserved order (see Chapter *Nidovirales*; see also this Chapter Figures 1, 9 and 12 and Table 2). Two ORF1a-encoded replicase domains, an ADP-ribose-1″-phosphatase (ADRP, also called macrodomain; located in coronavirus nsp3) and a noncanonical "secondary" RdRp with possible primase activity (coronavirus nsp8) may represent diagnostic markers that distinguish members of the family *Coronaviridae* from viruses in other nidovirus taxa.

Demarcation criteria for genera and species

Only viruses for which a complete genome sequence is available (see Supplementary Table 1 available online on Science Direct®, www.sciencedirect.com) are to be considered for taxonomy and the following demarcation criteria are used.

- Established and newly identified members of the family *Coronaviridae* are assigned to a subfamily and genus on the basis of rooted phylogeny and calculation of pair-wise evolutionary distances for the following *Coronaviridae*-wide conserved domains in replicase polyprotein pp1ab: ADRP, nsp5 (3CL^pro), nsp12 (RdRp), nsp13 (Hel), nsp14 (ExoN), nsp15 (NendoU) and nsp16 (O-MT). This procedure, developed by Lauber and Gorbalenya (*in preparation*), at present unambiguously identifies 20 distinct non-overlapping clusters (with the largest intra-cluster distance being smaller than the smallest inter-cluster distance): 17 coronaviruses, 2 toroviruses, 1 bafinivirus). Likewise, the higher-rank clusters corresponding to genus and subfamily levels are recognized.
  - Phylogenetic outliers assigned to the family *Coronaviridae* may be considered representatives of a new genus when they do not cluster with any of the current genera and share less

Table 2: Cleavage products of coronavirus replicase polyproteins pp1a and pp1ab: names, assigned functions and structure

Protein	Assigned function	E/N*	MMDB ID
nsp1†	IFN antagonist	N	ND
	Degradation of host mRNAs		
	Inhibition of translation		
	Cell cycle arrest		
nsp2	Unknown; associates with RTCs	N	ND
nsp3	Papain-like proteinase PL1^pro; polyprotein processing	N	37505; 42180
	Papain-like proteinase PL2^pro; polyprotein processing, DUB	E	
	ADP-ribose-1″phosphatase (macrodomain); RNA-binding	N	
	IFN antagonist		
	DMV formation?		
nsp4	Unknown; DMV formation?		76092
nsp5	Main proteinase M^pro; polyprotein processing	E	20276; 23158
nsp6	Unknown; DMV formation?		ND
nsp7	ssRNA binding		36090
nsp8	Noncanonical "secondary" RdRp with putative primase activity; forms hexadecameric supercomplex with nsp7		36090
nsp9	ssRNA binding; associates with RTCs		26498; 60895
nsp10	Dodecameric zinc finger protein; associates with RTCs, stimulates nsp16 methyltransferase activity		40869; 40904
nsp11	Unknown		ND
nsp12	RdRp	E	ND
nsp13	Helicase	E	ND
	RNA 5′-triphosphatase		
nsp14	3′→5′exoribonuclease (required for RdRp fidelity)	N	ND
	Guanine-N7-methyltransferase (RNA cap formation)	N	
nsp15	Hexameric uridylate-specific endoribonuclease	N	40936
nsp16	Ribose-2′-*O*-methyltransferase (RNA cap formation)	N	ND

MMDB ID numbers are listed for replicase proteins for which crystal structures are available. IFN, interferon; RTC, replicase/transcriptase complex; DUB, deubiquitinating enzyme; DMV, double-membrane vesicles.
*Essential (E) or non-essential (N) for replication in cultured cells.
†Absent in gammacoronaviruses.

Pos. ssRNA

than 46% sequence identity in the aforementioned conserved replicase domains with any other established member of the family.

- Viruses that share more than 90% aa sequence identity in the conserved replicase domains are considered to belong to the same species. This 90% identity threshold serves as the sole species demarcation criterion.

SUBFAMILY   *CORONAVIRINAE*

Taxonomic structure of the subfamily

Subfamily	*Coronavirinae*
Genus	*Alphacoronavirus*
Genus	*Betacoronavirus*
Genus	*Gammacoronavirus*

On the basis of rooted and unrooted phylogenetic trees estimated for different regions of the genome, four coronavirus (CoV) clusters can be distinguished, three of which (corresponding to the former nonofficial "groups" 1, 2 and 3) have been recognized and classified as genera (*Alpha-*, *Beta-* and *Gammacoronavirus*, respectively). The fourth cluster comprises a number of recently identified coronaviruses of birds and by all standards appears to represent a novel (but yet to be approved) genus, provisionally named *Deltacoronavirus*. In the genus *Betacoronavirus*, four separate lineages can be discerned, designated A through D, that correspond to former subgroups 2A through D, respectively (Figure 2).
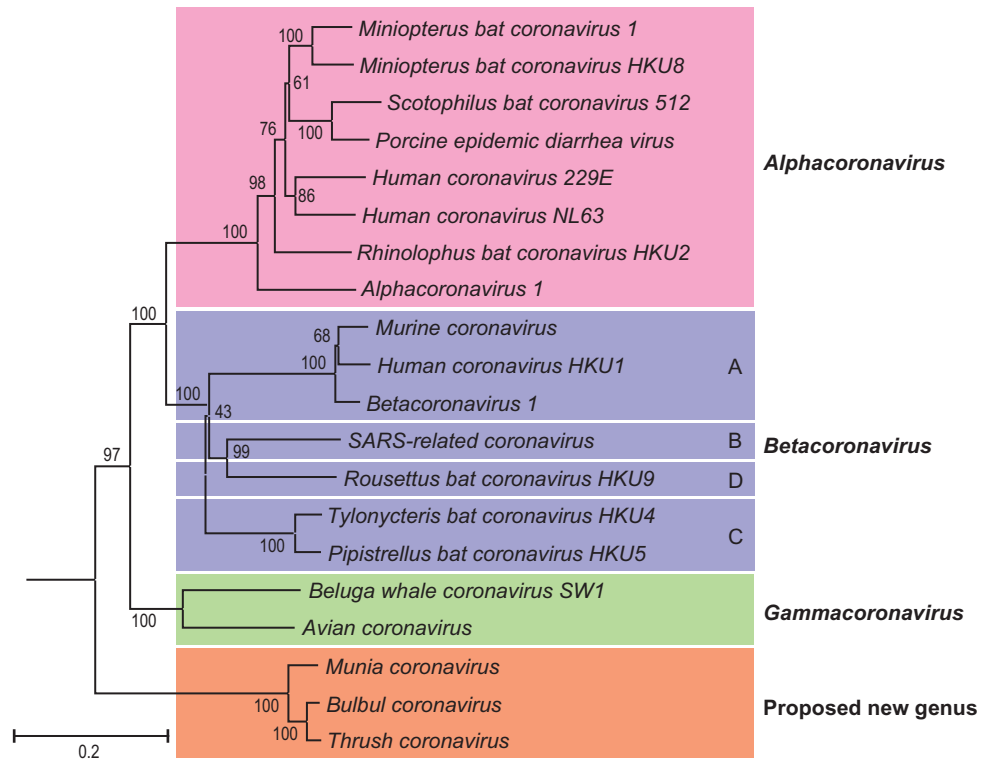

Figure 2: Phylogenetic relationships among the members of the subfamily *Coronavirinae*. A rooted neighbor-joining tree was generated from amino acid sequence alignments of RdRp and helicase domains with equine torovirus Berne as outgroup. The tree reveals four main monophyletic clusters corresponding to genera *Alpha-*, *Beta-* and *Gammacoronavirus* and an envisaged new genus (color-coded), and also shows the distinct betacoronavirus lineages A through D.

Virion properties

MORPHOLOGY

By conventional negative-staining electron microscopy, virions appear pleiomorphic, roughly spherical, 120–160 nm in diameter, with a characteristic fringe of large (ca. 20 nm), petal-shaped surface projections that are comprised of trimers of the spike (S) glycoprotein (Figure 3). Group A betacoronaviruses (Figure 2) display a second type of surface projection, 5–7 nm in length, comprised of the homodimeric hemagglutinin-esterase (HE) glycoprotein. Coronavirions as studied by cryo-electron tomography are homogeneous in size and spherical (envelope outer diameter 85 ± 5 nm). The envelope is exceptionally thick (7.8 ± 0.7 nm) in comparison to typical biological membranes (average thickness ca. 4 nm). The nucleocapsid, a loosely-wound helix, seems to be tightly folded to form a compact core that appears to be separated from the envelope by a gap of about 4 nm (Figure 3).

PHYSICOCHEMICAL AND PHYSICAL PROPERTIES

The estimated Mr of the virion is 400×10^6, its buoyant density in sucrose and CsCl is 1.15–1.20 g cm^{-3} and 1.23–1.24 g cm^{-3}, respectively, and its $S_{20,W}$ is 300 to 500S. Particles are sensitive to heat, lipid solvents, non-ionic detergents, formaldehyde, oxidizing agents and UV irradiation.

NUCLEIC ACID

Members of the subfamily *Coronavirinae* possess a unimolecular, positive stranded RNA genome, which is capped, polyadenylated and infectious. Genome lengths range from 26.4 in the non-assigned Thrush coronavirus to 31.7 kb for the gammacoronavirus Beluga whale coronavirus, the largest RNA virus known to date. At present, complete genomes are available for more than 300 naturally-occurring coronaviruses. (For a complete list refer to Supplementary Table 1 available online on Science Direct®, www.sciencedirect.com.)

PROTEINS

Coronaviruses all share the following structural protein species:

- the spike protein S, a large (1128–1472 aa), homo-trimeric type I membrane glycoprotein. S is a class I fusion protein that mediates receptor-binding and membrane fusion;
- the membrane glycoprotein M, a 218 to 263-aa integral type III membrane protein with predicted triple-spanning $N^{exo}C^{endo}$ topology. Depending on the virus species, the amino-terminal
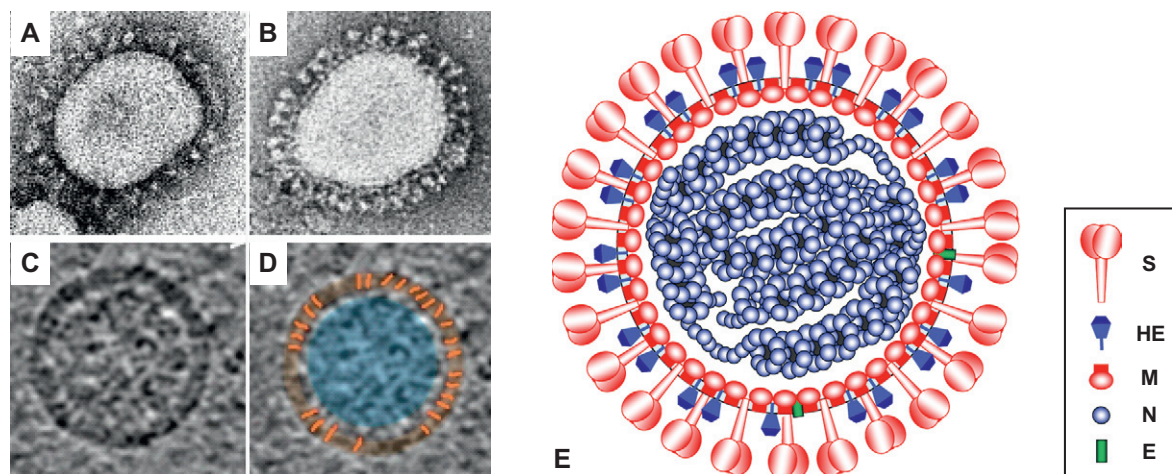
Figure 3: Coronavirus virion morphology. (A,B) Negative staining (2% phosphotungstic acid) electron micrographs of murine coronavirus particles. Shown are (A) a virion of murine coronavirus laboratory strain A59 that lacks HE expression and (B) one of a recombinant MHV-59 virus in which HE expression was restored (B) (courtesy Jean Lepault, Laboratory of Molecular and Structural Virology, Gif-sur-Yvette Cedex, France). (C,D) Cryo-electron tomographs of mouse hepatitis virus. A virtual slice (7.5 nm thick) through a reconstructed MHV particle (left) with highlighted features superimposed (right). The envelope is colored in orange with conspicuous striations highlighted; the nucleocapsid region is colored in blue. Note low-density region (ca. 4 nm) between envelope and nucleocapsid (reprinted with permission from Barcéna *et al.* (2008) *Proc. Natl Acad. Sci., U S A*, **106**, 582-587, © 2008 National Academy of Sciences, USA). (E) Schematic representation of a (lineage A) betacoronavirus virion.

ectodomain is decorated with *N*- or *O*-linked glycans. The long C-terminal endodomain, comprising an amphiphilic region and a hydrophilic tail, is believed to associate with the inner leaflet of the membrane to form a matrix-like lattice, which would explain the remarkable thickness of the coronavirus envelope (Figure 3). In transmissible gasteroenteritis virus of swine (TGEV, sp. *Alphacoronavirus 1*), a second population of tetra-spanning M proteins, adopting an N^{exo}-C^{exo} topology in the viral envelope, has been described;

- the envelope protein (E), a small (74–109 aa) pentameric integral membrane protein with ion channel and/or membrane permeabilizing (viroporin) activities. With around 20 copies per particle, the E protein is only a minor structural component. Although its precise function remains to be defined, the E protein plays a role in virion assembly and morphogenesis and has been identified as a virulence factor for the severe acute respiratory syndrome-coronavirus (SARS-CoV);
- the nucleocapsid protein N, a 349 to 470 aa RNA-binding phosphoprotein. Besides its obvious function in genome encapsidation, the N protein also is involved in RNA synthesis and translation, displays RNA chaperone activity, and acts as a type I interferon antagonist.

Depending on the coronavirus species, additional accessory proteins may be incorporated into the virion. Group A betacoronaviruses (*Betacoronavirus* 1, *Murine coronavirus* and *Human coronavirus HKU1*) code for an accessory homo-dimeric type I envelope glycoprotein, the hemagglutinin-esterase (HE). It mediates reversible virion attachment to *O*-acetylated sialic acids by acting both as a lectin and as a sialate-*O*-acetylesterase. The coronavirus HE shares about 30% aa sequence identity with the torovirus HE protein and is equally related to subunit 1 of the influenza C virus hemagglutinin-esterase fusion protein (HEF). In SARS-CoV, proteins 3a, 6 and 7 have been described as structural proteins and nsp2 through 5 and nsp9 were all detected in purified virion preparations.

In virions of murine coronavirus, the stoichiometric ratio of N, M and HE proteins is approximately 1:2.6:0.4; in TGEV, N and M occur at a ratio of 1:3. There are no reliable estimates for the S protein as it is present in small quantities in virus particles, may occur both in cleaved and uncleaved forms, and is easily lost during virus purification.

LIPIDS
Coronaviruses acquire their lipid envelopes by budding at membranes of the endoplasmic reticulum, intermediate compartment and/or Golgi complex. The S and E proteins are palmitoylated.

CARBOHYDRATES
Coronavirus S and HE proteins are heavily glycosylated and contain multiple *N*-linked glycans (20–35 and 5–11, respectively). The M protein of coronaviruses contains a small number of either *N*- or *O*-linked glycans, depending on the virus species, located near the amino-terminus. Coronavirus E proteins are not glycosylated.

Genome organization and replication

Coronavirus genomes contain 5′ and 3′ UTRs ranging in size from 200 to 600 and from 200 to 500 nt, respectively. Signals for genome replication and encapsidation reside not only in these UTRs, but also in adjacent and more internal coding regions. Six ORFs are conserved subfamily-wide and arranged in a fixed order: (as listed in the 5′ to-3′ direction) ORFs 1a and 1b, together comprising the replicase gene, and the ORFs for the structural proteins S, E, M and N. Downstream of ORF1b and interspersed between the structural protein genes, there may be up to at least eight accessory (also called "group" or "niche-specific") genes, the products of which are generally dispensable for replication *in vitro*, but key to efficient replication during natural infection (Figure 1).

Apparently, these accessory genes were acquired through horizontal gene transfer and occasionally also lost again as the different coronaviruses evolved and diverged while adapting to new hosts and niches. The diversity of accessory genes, most of which are specific only to a distinct CoV lineage species or strain (see also Figures 5–7), attest to the plasticity and highly dynamic nature of the coronavirus genome.

While the genome serves as an mRNA for the replicase polyproteins, the 3′ proximal genes are expressed from a nested set of sg mRNAs the coding regions of which (the "body" sequences) are

3'-coterminal with the genome. Each of these mRNAs is provided with a short 5' leader sequence identical to the 5'-terminal end of the genome. Leader and body sequences are not contiguous on the genome (they may in fact be separated by more than 20,000 nts), but become joined in a process of discontinuous minus-strand RNA synthesis (detailed below). Although all except the smallest mRNAs are structurally polycistronic, translation is restricted to the 5'-proximal ORF(s) not present in the next smaller mRNA of the set (Figure 1).

On the genome, each transcription unit (one or more ORFs expressed from a single RNA species) is preceded by a short conserved sequence element, commonly called the transcription-regulating sequence (TRS). A TRS copy is also found at the 5' end of the genome, immediately downstream of the leader sequence. According to the prevailing model for transcription, leader–body fusion occurs during the synthesis of genome-templated sg minus-strand RNAs by 3'-discontinuous extension via a mechanism resembling homology-assisted RNA recombination. This process apparently is driven by sequence complementarity between the anti-TRS at the 3' end of the nascent minus-strand and the 5' genomic TRS (Figure 4). In support of this model, the production of a 5'-terminal nested set of transcriptionally-active sg minus-strand RNAs with a 3'-terminal anti-leader sequence (in effect a mirror copy set of the mRNAs) has been demonstrated in coronavirus-infected cells. It is believed that each mRNA is transcribed from its corresponding sg minus-strand RNA template via a process of "continuous" RNA synthesis. For more information about other aspects of coronavirus replication, please see the preceding paragraphs and Chapter *Nidovirales*.

Antigenic properties

Cross-reactivity among coronaviruses is limited to (closely-related) species within the same genus. The S protein is the major inducer of virus-neutralizing antibodies that are elicited mainly by epitopes in the amino terminal half of the molecule. The surface-exposed amino-terminus of the M protein induces antibodies that neutralize virus infectivity in the presence of complement, while the HE protein of group A betacoronaviruses induces antibodies that prevent virion binding
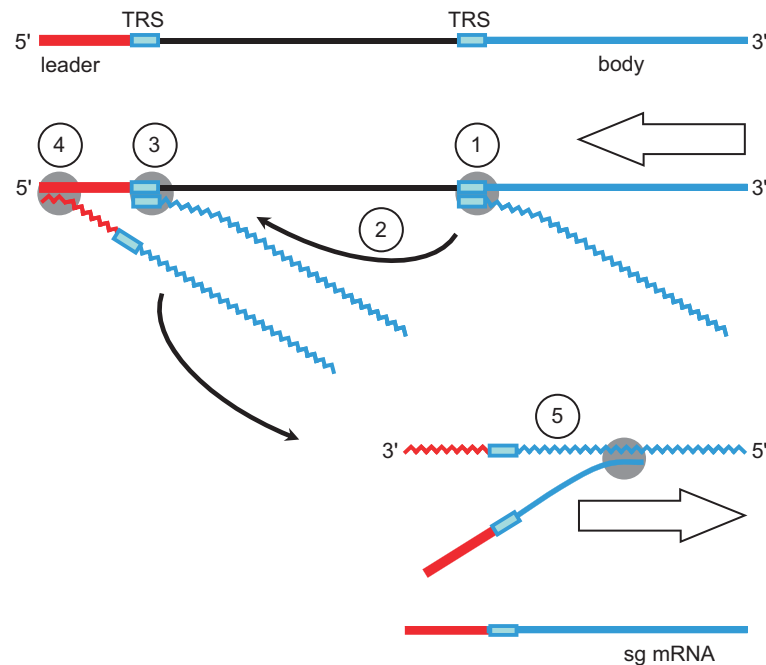

Figure 4: Coronavirus mRNA synthesis: the discontinuous 3'-extension model. Minus-strand synthesis initiates at the 3' end of the genome and proceeds until a TRS is copied (1). The nascent minus-strand RNA may then be transferred to the 5' end of the genome (2). Base complementarity allows the minus-strand RNA to anneal to the leader TRS (3) after which RNA synthesis resumes and body (in blue) and leader sequences (in red) become fused (4). The chimeric sg minus-strand RNA in turn serves as a template for "continuous" synthesis of sg mRNAs (5).

to *O*-acetylated sialic acids or inhibit sialate-*O*-acetylesterase activity. The N protein is a dominant antigen during the natural infection and while N-specific antibodies may provide little immune protection, they are of serodiagnostic relevance.

The ectodomains of the S and HE proteins are highly variable, suggestive of extensive antigenic drift. There are also indications for the occurrence of antigenic shifts as there are several examples of intra- and possibly interspecies exchange through RNA recombination of coding sequences of S (for *Avian coronavirus*, *Murine coronavirus* and the *Alphacoronavirus 1* subspecies feline and canine coronavirus) and HE ectodomains (*Murine coronavirus*) sometimes with as yet unidentified coronaviruses serving as donors. Studies performed with murine and feline coronaviruses indicate that both structural and non-structural (replicase) proteins serve as $CD4^+$ and $CD8^+$ T cell antigens.

There is no serologic cross-reactivity between corona-, toro- and bafiniviruses.

Biological properties

Coronaviruses infect birds and mammals and include several pathogens of clinical, veterinary and economic interest. Transmission is not by biological vectors, but – depending on the virus species – via fomites or via aerogenic and/or fecal–oral routes. As CoVs primarily target epithelial cells, they are generally associated with gastrointestinal and respiratory infections that may be acute or become chronic with prolonged shedding of virus. In general, these infections are mild and often asymptomatic. Some coronaviruses, however, cause severe, even lethal disease. Murine coronavirus (genus *Betacoronavirus*) may cause hepatitis and severe neurologic infection, resulting in paralysis and demyelination, providing a rodent model for the study of the neuropathogenesis of human multiple sclerosis. Some members of the species *Alphacoronavirus 1* (feline, canine and ferret coronavirus) cause fatal immune-mediated systemic infections in their respective hosts, presumably through the infection of cells of the macrophage/monocyte lineage, with widespread inflammatory lesions in multiple organs. The human coronaviruses that were identified early on (*Betacoronavirus-1* subspecies HCoV-OC43 and *Alphacoronavirus* HCoV-229E) mostly cause common colds and have long been considered of modest clinical importance. It is now recognized that these viruses may also cause severe lower respiratory tract infections (LRTI) in infants and elderly, and apparently are responsible for about 5% of infant hospitalizations from LRTI, globally.

In 2002–2003, a previously unknown coronavirus, SARS-CoV, caused an epidemic in human populations of a severe pulmonary disease with a mortality rate of 10% that rapidly spread to four continents, infecting 8,096 individuals and claiming 774 victims before it was contained. Epidemiological evidence indicates that this novel human virus originated in bats, spread to Himalayan palm civets, Chinese ferret badgers and raccoon dogs at the wet markets of Guangdong, China, to enter the human population through handling or consumption of these exotic species. Although SARS has since vanished, the episode does underline the pathogenic potential of coronaviruses and the possibility of novel emerging coronavirus infections arising from cross-species transmissions. Similar incidents, though with a less dramatic outcome, seem to have given rise to human coronavirus OC43 (a single cross-species transmission of bovine coronavirus from cattle to humans), to human coronavirus 229E (transmitted from bats?) and, more recently, to canine respiratory coronavirus (transmission of bovine coronavirus to dogs). In the wake of the SARS epidemic, molecular surveillance and virus discovery studies have yielded evidence for at least 60 novel coronaviruses among which are two new human respiratory coronaviruses, HCoV-HKU1 and HCoV-NL63. The latter is considered an important cause of (pseudo)croup and bronchiolitis in children. These studies also revealed a new lineage of predominantly avian viruses (Thrush, Bulbul and Munia coronavirus), with possible relatives in mammals (Asian leopard cat, Chinese ferret badger), that on the basis of rooted phylogeny appear to belong to a new genus (Figure 2). Bats harbor an exceptionally wide diversity of coronaviruses and have been proposed to play a vital role in coronavirus ecology and evolution, maybe even as the original hosts from which many if not all alpha- and betacoronavirus lineages were derived. Bat population densities and their roosting and migration habits would all favor such a role. Although this hypothesis has its merits and the recent virus discovery studies that prompted this view have been of truly Herculean proportions, it is of note that the actual coronavirus sampling size remains in fact limited and as efforts so far focused mainly on bats, our present perceptions may be biased. Further surveillance studies of similar extent must be performed in other host species (rodents, birds) before final conclusions can be drawn.

GENUS ALPHACORONAVIRUS

Type species   ***Alphacoronavirus 1***

Distinguishing features

The viruses in this genus form a distinct monophyletic group within the *Coronavirinae* subfamily. Apart from their relatively close phylogenetic relationship, the only general characteristics that would set them apart from other coronaviruses are (i) a unique type of nsp1, distinct in size and sequence from betacoronavirus nsp1 and without apparent counterpart in the gammacoronaviruses, and (ii) the presence of a commonly-shared accessory gene (designated ORF3 in most alphacoronavirus species, ORF3b and 3c in TGEV and in FCoV/CCoV, respectively) for a dispensable multi-spanning alphacoronavirus membrane protein (αmp). While for some alphacoronaviruses, αmp is the only accessory protein, others may carry up to at least six accessory genes (e.g. members of the subspecies canine coronavirus in the species *Alphacoronavirus 1*; note that "subspecies" is not an officially recognized level in virus taxonomy; the term is used here and throughout this chapter to indicate well-defined monophyletic groups of viruses within a coronavirus species that are genetically and biologically distinct from other members of the same species). A comparison of the genome organization of alphacoronaviruses is presented in Figure 5.
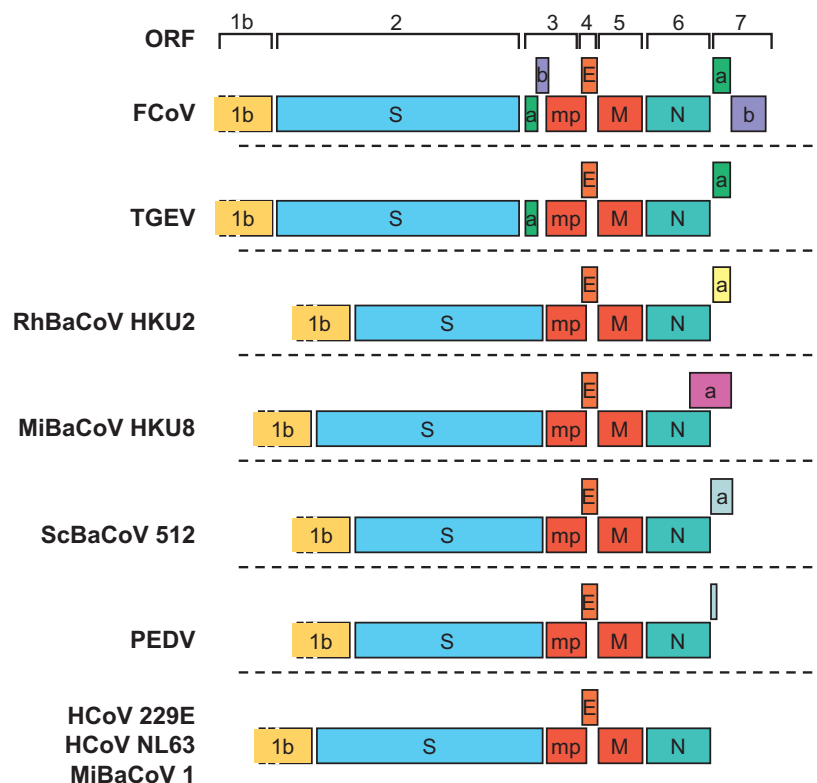

Figure 5: Alphacoronavirus genome organization. Comparison of the 3'-terminal genomic regions downstream of ORF1b of alphacoronaviruses representative of the different species and subspecies. ORFs are depicted as coloured boxes and indicated by number (above) and encoded protein. ORFs for accessory proteins are named as by convention according to number (referring to the mRNA species from which they are expressed) and, in the case of multiple ORFs in one transcription unit, alphabetically. Conservation of genes is indicated by identical colouring. Accessory genes of different viruses that are located in the same genomic location but believed to encode non-related products are coloured differently. For the abbreviations of virus names, please see list of species in the genus *Alphacoronavirus* below. 1b, ORF1b; mp, alphacoronavirus-specific accessory membrane protein αmp; all other acronyms as in Figure 1.

Pos. ssRNA

List of species in the genus *Alphacoronavirus*

Alphacoronavirus 1		
Canine coronavirus		
Canine coronavirus type I		
Canine coronavirus strain Elmo/02	[AY426983]	(CCoV Elmo/02)
Canine coronavirus type II		
Canine coronavirus strain NTU336/F/2008	[GQ477367]	(CCoV NTU336/F/2008)
Feline coronavirus		
Feline coronavirus type I		
Feline coronavirus C1Je	[DQ848678]	(FCoV C1Je)
Feline coronavirus type II		
Feline infectious peritonitis virus WSU 79-1146	[AY994055]	(FIPV 79-1146)
Porcine respiratory coronavirus		
Porcine respiratory coronavirus ISU-1	[DQ811787]	(PRCV ISU-1)
Transmissible gastroenteritis virus		
Transmissible gastroenteritis virus virulent Purdue	[AJ271965]	(TGEV virulent Purdue)
Human coronavirus 229E		
Human coronavirus 229E	[AF304460 = NC_002645]	(HCoV 229E)
Human coronavirus NL63		
Human coronavirus NL63 Amsterdam 1	[AY567487]	(HCoV NL63)
Miniopterus bat coronavirus 1		
Miniopterus bat coronavirus 1A		
Miniopterus bat coronavirus 1A AFCD62	[EU420138 = NC_010437]	(Mi-BatCoV 1A AFCD62)
Miniopterus bat coronavirus 1B		
Miniopterus bat coronavirus 1B AFCD307	[EU420137 = NC_010436]	(Mi-BatCoV 1B AFCD307)
Miniopterus bat coronavirus HKU8		
Miniopterus bat coronavirus HKU8 AFCD77/08/05 Mm	[EU420139 = NC_010438]	(Mi-BatCoV HKU8 AFCD77/08/05 Mm)
Porcine epidemic diarrhea virus		
Porcine epidemic diarrhea virus CV777	[AF353511 = NC_003436]	(PEDV CV777)
Rhinolophus bat coronavirus HKU2		
Rhinolophus bat coronavirus HKU2/HK/46/2006	[EF203065]	(Rh-BatCoV HKU2/HK/46/2006)
Scotophilus bat coronavirus 512		
Scotophilus bat coronavirus 512/2005	[DQ648858 = NC_009657]	(Sc-BatCoV 512/2005)

Species names are in italic script; names of subspecies and isolates are in roman script. Sequence accession numbers [ ] and assigned abbreviations ( ) are also listed.

List of other related viruses which may be members of the genus *Alphacoronavirus* but have not been approved as species

Carollia bat coronavirus 1FY2BA/Trinidad/2007		(Ca-BatCoV 1FY2BA/Trinidad/2007)
Chaerophon bat coronavirus 40/Kenya/2006		(Ch-BatCoV 40/Kenya/2006)
Chaerophon bat coronavirus 22/Kenya/2006		(Ch-BatCoV 22/Kenya/2006)
Chinese ferret badger coronavirus DM95/03	[EU769560]	(CFBCoV DM95/03)
Eptesicus bat coronavirus 65/RM/2006	[EF544566]	(Ep-BatCoV 65/RM/2006)
Ferret coronavirus	[GU338456; GU338457]	(FerCoV)
Glossophaga bat coronavirus 1CO7BA/Trinidad/2007		(Gl-BatCoV 1CO7BA/Trinidad/2007)
Kenya bat coronavirus BtKY12	[GQ920811]	(BatCoV BtKY12)
Kenya bat coronavirus BtKY21	[GQ920819]	(BatCoV BtKY21)
Myotis bat coronavirus HKU6/HK/21/2005	[DQ249224; DQ249247]	(My-BatCoV HKU6/HK/21/2005)
Myotis bat coronavirus D2.2/Germany/2007		(My-BatCoV D2.2/Germany/2007)
Myotis bat coronavirus D8.38/Germany/2007		(My-BatCoV D8.38/Germany/2007)
Myotis bat coronavirus 3/RM/2006	[EF544567]	(My-BatCoV 3/RM/2006)
Myotis bat coronavirus 48/RM/2006	[EF544565]	(My-BatCoV 48/RM/2006)

Myotis Bat coronavirus M.mac/Australia/ CoV034/2008	[EU834951]	(My-BatCoV M.mac/Aus/ CoV034/2008)
Miniopterus bat coronavirus 088/Australia/2007	[EU834952]	(Mi-BatCoV 088/Australia/2007)
Miniopterus bat coronavirus HKU7/HK/13/2005	[DQ249226; DQ249249]	(Mi-BatCoV HKU7/HK/13/2005)
Nyctalus bat coronavirus VM366/NLD/2008		(Ny-BatCoV VM366/NLD/2008)
Pipistrellus bat coronavirus D5.16/Germany/2007		(Pi-BatCoV D5.16/Germany/2007)
Pipistrellus bat coronavirus D5.71/Germany/2007		(Pi-BatCoV D5.71/Germany/2007)
Pipistrellus bat coronavirus VM312/NLD/2008		(Pi-BatCoV VM312/NLD/2008)
Raccoon dog coronavirus GZ43/03	[EU769559; EF192159]	(RDCoV GZ43/03)
Rhinolophus bat coronavirus A970/SD/2005		Rh-BatCoV A970/SD/2005
Rousettus bat coronavirus HKU10/GD/183/2005		(Ro-BatCoV HKU10/GD/183/2005)
Yellow-bellied weasel coronavirus GX/D726/2005	[ABQ39953.1|]	(YWCoV GX/D726/05)

GENUS BETACORONAVIRUS

Type species   *Murine coronavirus*

Distinguishing features

Betacoronaviruses form a distinct monophyletic group in the *Coronavirinae* subfamily. Except for their relatively close phylogenetic relationship, the only known general characteristic that would set them apart from other coronaviruses is their unique nsp1, distinct in size and sequence from alphacoronavirus nsp1 and without obvious counterpart in the gammacoronaviruses. Four beta-coronavirus lineages can be distinguished (A through D; Figure 2) each with a unique set of accessory genes (Figure 6).

List of species in the genus *Betacoronavirus*

Betacoronavirus 1		
Bovine coronavirus		
Bovine coronavirus Mebus	[U00735]	(BCoV Mebus)
Equine coronavirus		
Equine coronavirus NC99	[EF446615 = NC_010327]	(ECoV NC99)
Human coronavirus OC43		
Human coronavirus OC43 ATCC VR-759	[AY585228]	(HCoV OC43 ATCC VR-759)
Porcine hemagglutinating encephalomyelitis virus		
Porcine hemagglutinating encephalomyelitis virus VW572	[DQ011855 = NC_007732]	(PHEV VW572)
Human coronavirus HKU1		
Human coronavirus HKU1 N1	[AY597011 = NC_006577]	(HCoV HKU1 N1)
Murine coronavirus		
Mouse hepatitis virus		
Murine hepatitis virus JHM	[NC_006852]	(MHV JHM)
Rat coronavirus		
(Rat sialodacryoadenitis coronavirus)		
Rat coronavirus Parker	[FJ938068]	(RCoV Parker)
Pipistrellus bat coronavirus HKU5		
Pipistrellus bat coronavirus HKU5/ HK/03/2005	[EF065509 = NC_009020]	(Pi-BatCoV HKU5/ HK/03/2005)
Rousettus bat coronavirus HKU9		
Rousettus bat coronavirus HKU9/ GD/005/2005	[EF065513 = NC_009021]	(Ro-BatCoV HKU9/ GD/005/2005)
Severe acute respiratory syndrome-related coronavirus		
SARS-related human coronavirus		
SARS-related human coronavirus Urbani	[AY278741]	(SARS CoV Urbani)
SARS-related Rhinolophus bat coronavirus RF1		

Pos. ssRNA

SARS-related Rhinolophus bat coronavirus Rf1/2004	[DQ412042 = NC_009695]	(SARSr-Rh-BatCoV RF1/2004)
SARS-related Rhinolophus bat coronavirus Rm1		
SARS-related Rhinolophus bat coronavirus Rm1/2005	[DQ412043 = NC_009696]	(SARSr-Rh-BatCoV Rm1/2005)
SARS-related Rhinolophus bat coronavirus Rp3		
SARS-related Rhinolophus bat coronavirus Rp3/2004	[DQ071615 = NC_009693]	(SARSr-Rh-BatCoV Rp3/2004)
SARS-related Rhinolophus bat coronavirus HKU3		
SARS-related Rhinolophus bat coronavirus HKU3/HK/24/2005	[DQ022305 = NC_009694]	(SARSr-Rh-BatCoV HKU3/HK/24/2005)
SARS-related palm civet coronavirus		
SARS-related palm civet coronavirus SZ3/2003	[AY304486]	(SARSr-CiCoV SZ3/2003)
SARS-related chinese ferret badger coronavirus		
SARS-related chinese ferret badger coronavirus CFB/SZ/94/03	[AY545919]	(SARSr CoV CFB/ SZ/94/03)
Tylonycteris bat coronavirus HKU4		
Tylonycteris bat coronavirus HKU4/ HK/04/2005	[EF065505 = NC_009019]	(Ty-BatCoV HKU4/ HK/04/2005)

Species names are in italic script; names of subspecies and isolates are in roman script. Sequence accession numbers [ ] and assigned abbreviations ( ) are also listed.
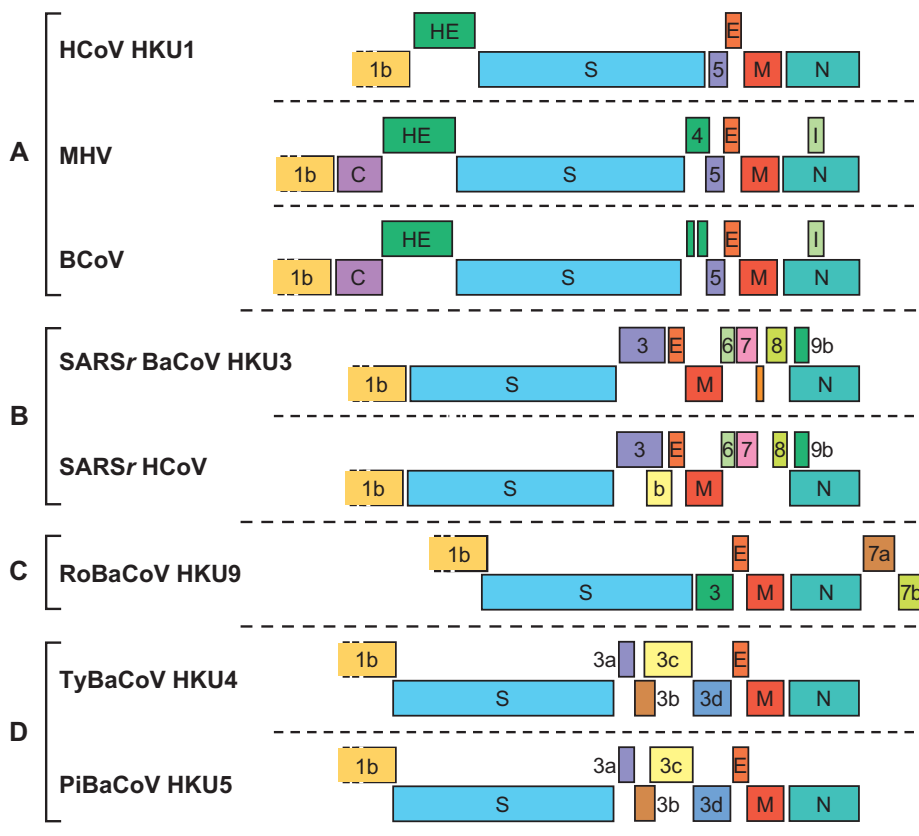

Figure 6: Betacoronavirus genome organization. Comparison of the 3'-terminal genomic regions downstream of ORF1b of betacoronaviruses representative for the different species and subspecies. ORFs are depicted as boxes, color-coded and indicated by number and gene product as in Figure 5. For the abbreviations of virus names, please see the list of species in the genus *Betacoronavirus*.

List of other related viruses which may be members of the genus *Betacoronavirus* but have not been approved as species

      Rhinonicteris bat coronavirus 000/Australia/2006   [EU834950]   (Rh-BatCoV 000/Australia/2006)

GENUS   *GAMMACORONAVIRUS*

Type species   *Avian coronavirus*

Distinguishing features

Gammacoronaviruses form a distinct monophyletic group in the *Coronavirinae* subfamily. Except for their relatively close phylogenetic relationship, there are no known common characteristics in terms of virion morphology, genome organization and gene composition, replication or biology that would set them apart from other coronaviruses. Viruses of the species *Avian coronavirus* lack an nsp1 moiety. Whether this is also the case for members of the other gammacoronavirus species, *Beluga whale coronavirus*, remains to be determined. For the genome organization of gammacoronaviruses see Figure 7.
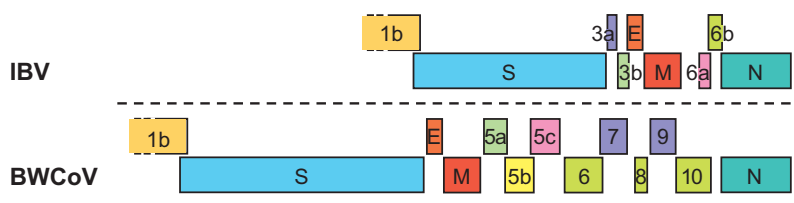

Figure 7: Gammacoronavirus genome organization. Comparison of the 3′-terminal genomic regions downstream of ORF1b of infectious bronchitis virus (IBV; sp. *Avian coronavirus*) and Beluga whale coronavirus SW1 (BWCoV). ORFs are depicted as boxes, colour-coded and indicated by number and gene product as in Figures 5 and 6.

List of species in the genus *Gammacoronavirus*

Avian coronavirus
   Infectious bronchitis virus
      Infectious bronchitis virus Beaudette   [M95169 = NC_001451]   (IBV Beaudette)
   Turkey coronavirus
      Turkey coronavirus ATCC   [EU022526]   (TCoV ATCC)
Beluga whale coronavirus SW1
      Beluga whale coronavirus SW1   [EU111742 = NC_010646]   (BWCoV SW1)

Species names are in italic script; names of subspecies and isolates are in roman script. Sequence accession numbers [ ] and assigned abbreviations ( ) are also listed.

List of other related viruses which may be members of the genus *Gammacoronavirus* but have not been approved as species

      Duck coronavirus   [AJ854130]   (DCoV)
      Goose coronavirus   [AJ854145-AJ854159]   (GCoV)
      Pigeon coronavirus   [AJ854131]   (PCoV)

List of unassigned species in the subfamily *Coronavirinae*

      None reported.

List of other related viruses which may be members of the subfamily *Coronavirinae* but have not been approved as species

      Asian leopard cat coronavirus   [EF584908]   (ALCCoV)
      Bulbul coronavirus HKU11   [FJ376619]   (BuCoV HKU11)

Pos. ssRNA

Chinese ferret badger coronavirus GX/247/06	[ABQ39964]	(CFBCoV GX/247/06)
Munia coronavirus HKU13	[FJ376622 = NC_011550]	(MunCoV HKU13)
Parrot coronavirus AV71/99	[ABB45386]	(PaCoV)
Thrush coronavirus HKU12	[FJ376621 = NC_011549]	(ThCoV HKU12)

SUBFAMILY   TOROVIRINAE

Taxonomic structure of the subfamily

Subfamily	*Torovirinae*
Genus	*Torovirus*
Genus	*Bafinivirus*

Distinguishing features

The members of the bigeneric subfamily *Torovirinae* (family *Coronaviridae*) form a distinct mono-phyletic cluster that is well-separated from the "true" coronaviruses united in the subfamily *Coronavirinae* (for a phylogram depicting the relationships among the *Coronaviridae*, see Chapter *Nidovirales*, Figure 4). Apart from their relatively close phylogenetic relationship, bafini- and toro-viruses can be distinguished from their closest relatives, the true coronaviruses, by the following common features:

- Bacilliform virion particles
- Tubular nucleocapsids of presumably helical symmetry
- Small nucleocapsid proteins (ca. 160 aa), less than half the size of that of the coronavirus N protein
- The apparent lack of an equivalent of coronavirus envelope protein E
- A relatively simple genome organization (no accessory genes with the debatable exception of the torovirus HE gene) and exceptionally long (>800 nt) 5'-terminal UTRs.

Toro- and bafiniviruses can be distinguished from each other by:

- their transcription mechanism (detailed below)
- an ORF1a-encoded cyclic nucleotide phosphodiesterase domain (CPD) unique to toroviruses
- the HE gene, present only in toroviruses.

Remarkably, in coronaviruses, related CPD and HE proteins have been identified, but exclusively in one subset of betacoronaviruses (group A) and in completely different genome locations; here, the CPD protein is encoded by an accessory gene, located immediately downstream of ORF1b (see Figure 6). The torovirus CPD and HE coding sequences are believed to have been acquired by hori-zontal gene transfer from as yet unknown donors presumably after the toro-bafinivirus split.

GENUS   TOROVIRUS

Type species   *Equine torovirus*

Virion properties

MORPHOLOGY

In conventional negative-staining electron micrographs, toroviruses appear as a mixture of spheri-cal, rod- and kidney-shaped particles (Figure 8A). Native torovirus particles presumably are bacil-liform with rounded ends, measuring 100–140 nm in length and 35–42 nm in width (envelope outer rim). Virions carry two types of surface projections that in size and shape resemble those of the (beta) coronaviruses: multimers of the S protein comprising large 15-20 nm peplomers and homo-dimers of the HE protein forming an inner ring of smaller (5-7 nm) spikes. The most distinctive virion ele-ment, the core, is a flexible and seemingly hollow tube of helical symmetry (periodicity ca. 4.5 nm),
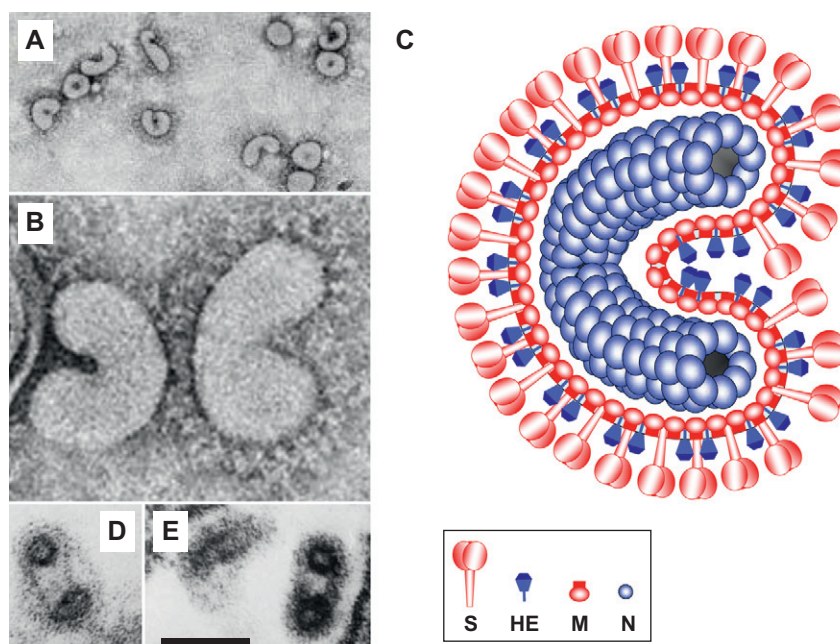
Figure 8: Virion morphology of equine torovirus Berne. (A) Negative staining electron micrograph of extracellular EToV particles (2% phosphotungstic acid). (B) Close-up of negatively-stained EToV virions. (Courtesy of Dolores Rodriguez Aguirre, Department of Molecular and Cell Biology, National Centre of Biotechnology, Madrid, Spain.) Note that in EToV strain Berne, the HE gene is inactivated and that virions consequently display only one type of spike, the peplomers comprised of the S protein. (C) Schematic representation of a torovirus virion. (D and E) Cross-sections of intracellular EToV virions showing the tubular nucleocapsid with central cavity and the viral envelope. The bar represents 100 nm.

about 100 nm in length and about 23 nm across with a central channel of about 10 nm in diameter. In crescent-shaped and spherical (disk-shaped?) particles, the nucleocapsid is bent into an open toroid, from which the name "torovirus" was derived (Figure 8).

PHYSICOCHEMICAL AND PHYSICAL PROPERTIES

Virions have a buoyant density in sucrose of 1.14–1.18 g cm^{-3}. Particles are sensitive to heat, lipid solvents, non-ionic detergents, formaldehyde, oxidizing agents and UV irradiation, but highly resistant to bile salts (0.1% deoxycholate) and extreme pH conditions (infectivity not affected by exposure to pH values between 2.5 and 10.3).

NUCLEIC ACID

The torovirus genome is a positive-stranded, capped and polyadenylated RNA molecule of about 28 kb in length that is infectious when transfected into mammalian cells. At present, complete genomes are available only for bovine torovirus (BToV) strain Breda and equine torovirus (EToV) strain Berne. Partial sequences, mostly for the genes for the structural proteins, are available for various Eurasian BToV and porcine torovirus (PToV) field strains.

PROTEINS

Virions of torovirus field strains contain the following protein species: (i) the spike protein S, a large (1562-1584 aa), presumably trimeric envelope glycoprotein with features typical for class I fusion proteins (bioinformatical analysis revealed heptad repeat regions and a putative fusion peptide); (ii) the hemagglutinin-esterase protein HE, a homo-dimeric 416 to 430 aa type I membrane glycoprotein that mediates reversible virion attachment to *O*-acetylated sialic acids by acting both as a lectin and as a receptor-destroying enzyme; (iii) the membrane protein M, a highly conserved 233-aa nonglycosylated integral membrane protein with three predicted transmembrane regions and a $N^{exo}C^{endo}$ topology; (iv) the nucleocapsid protein N, a 159 to 167-aa basic RNA-binding phosphoprotein. The HE protein is dispensable for replication *in vitro* and its expression has been lost in EToV strain Berne presumably as a result of adaptation to replication in cultured cells.

Pos. ssRNA

LIPIDS
Torovirus virions acquire their lipid envelope by budding at smooth intracellular membranes of ER and Golgi complex.

CARBOHYDRATES
Torovirus S and HE proteins carry multiple *N*-glycans (19–25 and 7–13, respectively).

Genome organization and replication

During natural infection, toroviruses presumably attach to their host cells by binding to 9-mono-*O*-(PToV) or 7,9-di-*O*-acetylated sialic acids (BToV) via their HE protein. Entry, however, would require the S protein to bind to a main receptor, most likely a specific glycoprotein, and to mediate fusion between the viral envelope and a cellular membrane. Whether entry occurs at the plasma membrane or via endocytosis is not known nor has any torovirus main receptor been identified so far.

The viral genome contains 5′ and 3′ UTRs of 821–857 and 200 nt, respectively, and six ORFs called ORF1a and 1b (together comprising the replicase gene), -2, -3, -4 and -5, the latter four encoding the structural proteins (S, M, HE and N, respectively; Figure 9). Signals for EToV genome replication and possibly also for encapsidation apparently reside in the 5′-terminal 604 and 3′-terminal 200 residues as suggested by studies with defective interfering RNAs.

Replication is believed to occur largely as described for coronaviruses with the genes for the structural proteins being expressed from a nested set of four sg mRNAs (designated (m)RNAs 2 through 5) that are 3′-coterminal with the genome (RNA 1). However, in striking contrast to corona- and bafiniviruses, toroviruses employ a mixed transcription strategy and combine discontinuous and continuous RNA synthesis to produce their complement of mRNAs. mRNAs 3, 4 and 5 lack a common leader sequence and are fully co-linear with the viral genome. The genes for M, HE and N, expressed from these RNAs, are each preceded by a conserved 13–14 nt sequence element, conforming to consensus $(C)ACN_{3-4}CUUUAGA$, a copy of which (but without the 5′-terminal C residue) is also found at the extreme 5′ end of the genome. This sequence element is thought to act as a premature termination signal of genome-templated minus-strand synthesis and, in the resulting sg minus-strand RNAs, as a promoter for mRNA synthesis with transcription initiating at the 5′-most adenine residue. The S gene lacks such an internal putative terminator/promoter (TP) element and apparently is expressed via a process similar to, yet distinct from, discontinuous RNA synthesis in coronaviruses. Its mRNA (mRNA 2) is the only sg mRNA species to carry a short 15–18 nt leader identical to the genomic 5′ end (i.e. the genomic TP). A conserved hairpin structure in ORF1b is believed to attenuate minus-strand synthesis to allow a subsequent similarity-assisted template switching event facilitated by sequence complementarity between the 3′ end of the nascent minus strand RNA and residues 16 through 38 of the genome that are located immediately downstream of the 5′-terminal genomic TP copy. This would result in a chimeric sg minus-strand RNA that, because of its acquisition of a complementary copy of the genomic TP, can serve as template to direct 'continuous' synthesis of mRNA 2 (Figure 9).

Limited information is available about torovirus morphogenesis. The M and S proteins are produced in the endoplasmic reticulum and largely retained in premedial-Golgi compartments. Assembly of torovirus virions occurs through budding of preformed nucleocapsids at smooth membranes consistent with the ER-to-Golgi intermediate compartment and/or Golgi complex. Mature particles egress by exocytosis.

Antigenic properties

EToV, PToV and BToV are serologically related. During natural infection, antibodies arise against each of the four structural proteins (S, HE, M and N). The spike (S) protein induces virus-neutralizing antibodies; sera from BToV- or PToV-infected animals cross-neutralize EToV.

Biological properties

So far, torovirus infection has been conclusively demonstrated only in ungulates: horse (EToV), bovine (BToV) and swine (PToV). Evidence is based on classical virological studies (isolation and
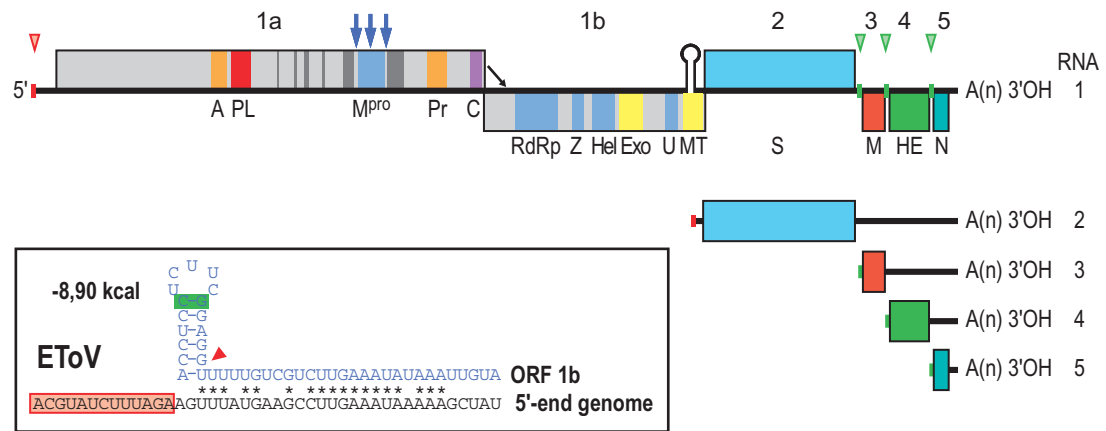
Figure 9: Organization and expression of the torovirus genome. (Upper panel) Schematic representation of the equine torovirus genome. ORFs and other genome elements are indicated as in Figure 1. The 5'-leader sequence present in the genome and sg mRNA 2 is depicted by a small red box. Green and red arrowheads/boxes indicate the locations of the internal and the 5'-terminal putative terminator/promoter (TP) elements, respectively. Blue arrows indicate established M^{pro} cleavage sites. The location of the discontinuous transcription element (DTE) driving mRNA 2 synthesis is shown by a hairpin. PL, papain-like proteinase; C, torovirus-specific ORF1a-encoded cyclic nucleotide phosphodiesterase domain. All other acronyms as in Figure 1. (inset) Structure of the mRNA 2 discontinuous transcription element, showing the hairpin structure and downstream "homology region" with sequence identity to the 5' end of the genome indicated by asterisks. A hairpin residue-pair displaying co-variation in BToV and PToV is highlighted in green. The site of mRNA 2 leader–body fusion is indicated by an arrowhead. The 5'-terminal genomic TP copy is highlighted by a red box. (Lower panel) Models for discontinuous (left) and non-discontinuous sg RNA synthesis (right) in toroviruses. The hairpin indicates the mRNA 2 DTE. Red boxes correspond to the 5'-terminal genomic TP copy and complementary sequences, blue boxes to the DTE homology region and the corresponding 5' genomic acceptor sequence. The TP consensus sequence is shown and highlighted by a green box. Internal TPs and complementary sequences are shown in green. The models show (from top to bottom) synthesis of genome-templated minus-strand RNA (minus-strand RNAs indicated by a wiggly line), attenuation and 3'-discontinuous extension directed by the DTE element and premature termination directed by internal TPs, and subsequent mRNA synthesis from sg minus-strand templates. Details are described in the text.

propagation *in vitro* of EToV and BToV strains; experimental infection of cattle with BToV Breda), serology and molecular genetic analysis (RT-PCR amplification of torovirus sequences from fecal samples of infected animals). There is serological evidence for torovirus infections also in goats and sheep. Among the non-ungulate species that have been proposed as potential hosts for toroviruses are human (HToV), turkey and carnivores, including dog, cat, mustelids, but these claims are supported only by EM detection of torovirus-like particles and/or limited genetic analysis and would require further experimental confirmation.
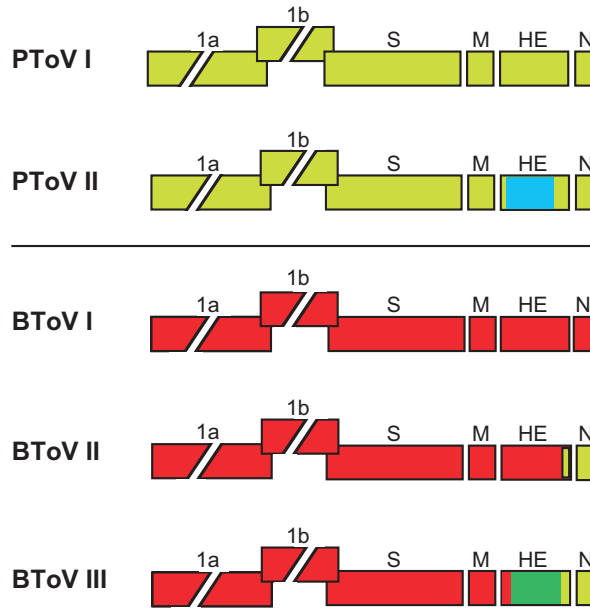
Figure 10: Genome organization of torovirus genotypes reveals evidence for multiple interspecies RNA recombination events. The genes for the replicase polyproteins (1a, 1b) and for the structural proteins S, M, HE, and N are depicted as boxes. For PToV, HE coding sequences are shown in different colors to indicate that one of the genotypes exchanged part of its HE gene through homologous RNA recombination with an as yet unknown torovirus. For BToV, sequences acquired from PToV (in yellow) and from a hitherto unidentified torovirus donor (in green) are also indicated.

Toroviruses of horses, swine and cattle have a world-wide distribution and are evidently ubiquitous as seroprevalence in host populations may exceed 80%. Transmission is probably via the oral/nasal route through contact with contaminated feces or nasopharyngeal secretions. Infected animals shed virus in the feces; in the case of BToV, nasal shedding has also been reported. Toroviruses are likely to cause both acute and chonic infections. Equine and porcine toroviruses are associated with asymptomatic enteric infections and remain viruses in search of a disease. Bovine torovirus, an established respiratory and enteric pathogen of cattle, may cause mild to profuse diarrhoea. The virus, originally designated Breda virus, was first isolated during an outbreak of severe neonatal gastroenteritis with 56.5% morbidity and 8.7% mortality in cattle from dairy farms round the township Breda, Iowa, and duly identified as the etiological agent. In experimentally-infected animals, BToV infects the epithelial cells lining the small and large intestine, with progression from areas of the mid jejunum down to the ileum and colon. Within the small intestine, cells of the upper third of the crypt and the epithelium overlying the Peyer's patches, including M cells, also become infected. Neonatal calves appear to be most susceptible to clinical infection. Maternal antibodies do not prevent infection, but modify the outcome of the disease as colostrum-deprived animals are more prone to develop severe diarrhoea.

Bovine and porcine toroviruses display host species preference at least to a certain degree. In phylogenetic analyses, all PToVs cluster, while extant BToVs mostly resemble the New World BToV isolate Breda, identified 30 years ago. However, there is evidence for recurring intergenotypic/interspecies recombination, suggesting that cross-species transmission may occur at least incidentally. Currently circulating Eurasian BToVs seem to have arisen from a genetic exchange, during which the 3' end of the HE gene, the N gene, and the 3' UTR of a Breda virus-like parent were exchanged for those of PToV. Moreover, some PToV and BToV variants carry chimeric HE genes, which apparently resulted from recombination events involving hitherto unknown toroviruses as donors. For the provisional nomenclature of BToV and PToV genotypes and a comparison of their genome organization, see Figure 10.

Species demarcation criteria in the genus

Only a modest number of toroviruses has been characterized and complete genome sequences are available solely for BToV strain Breda and EToV strain Berne. Thus far, host preference has been the main criterion for torovirus species demarcation, but future taxonomic classifications should follow the general criteria as outlined at the beginning of this chapter. According to these criteria, bovine and equine toroviruses justify their current status as distinct species.

For PToV, sequences are available only for the 3′-terminal region of the genome, containing the genes for the structural proteins. In phylogenetic trees, constructed for S, HE and M genes, all PToV field strains cluster and appear to be separated from EToV and BToV by an evolutionary distance larger than that between the latter two viruses, supporting the notion that PToV represents a distinct species. The limitations of taxonomy based solely upon phylogenetic analysis of genes for the structural proteins, however, are poignantly illustrated by the occurrence of toroviruses that have acquired novel HE and N genes via interspecies recombination.

Although HToV is designated as a species, evidence for the existence of the virus remains tenuous. Available sequence data are open to varying interpretation and, in any case, are insufficient to justify classification. Future taxonomic proposals are planned to resolve the issue.

List of species in the genus *Torovirus*

Bovine torovirus
    Bovine torovirus Breda I   [AY427798]   (BToV Breda)
Equine torovirus
    Equine torovirus strain Berne   [DQ310701; X52374]   (EToV Berne)
Human torovirus
    Human torovirus   (HToV)
Porcine torovirus
    Porcine torovirus Markelo   (PToV)

Species names are in italic script; names of isolates are in roman script. Sequence accession numbers [ ] and assigned abbreviations ( ) are also listed.

List of other related viruses which may be members of the genus *Torovirus* but have not been approved as species

None reported.

GENUS   *BAFINIVIRUS*

Type species   *White bream virus*

Virion properties

MORPHOLOGY

Virions are enveloped and bacilliform in shape (130–160×37–45 nm, excluding the surface projections; Figure 11). The most conspicuous virion elements are the large coronavirus-like spikes (20–25 nm) and a seemingly rigid tubular nucleocapsid of presumably helical symmetry (120–150×19–22 nm, with a central channel of 2–5 nm in diameter).

PHYSICOCHEMICAL AND PHYSICAL PROPERTIES

Virions have a buoyant density in sucrose of 1.17–1.19 g cm^{-3} and are sensitive to lipid solvents.

NUCLEIC ACID

White bream virus DF24/00, the sole bafinivirus described to date, possesses a 26.6 kb RNA genome, which is capped, polyadenylated and infectious.
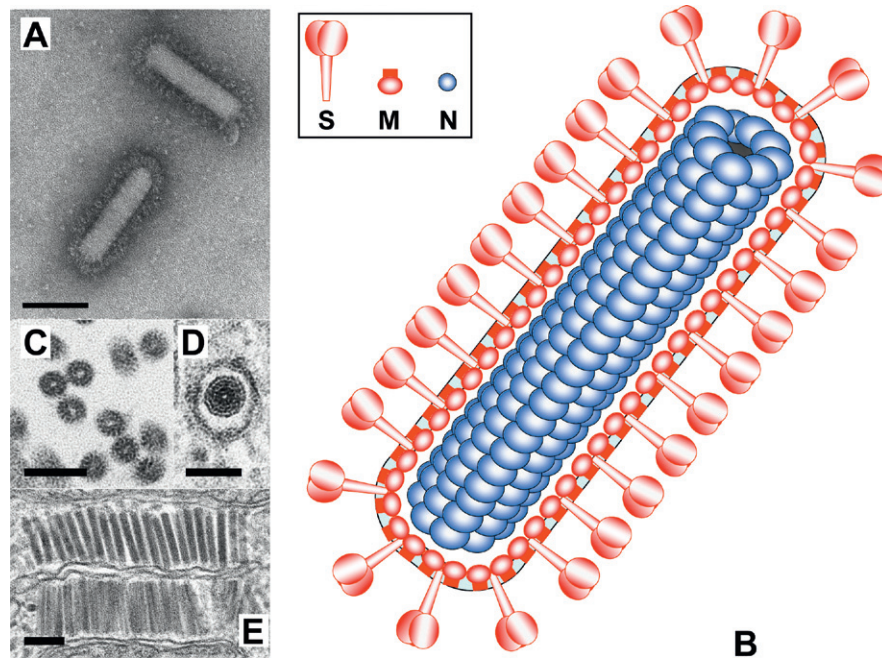
Figure 11: Virion morphology and morphogenesis of white bream virus. (A) Negative staining electron micrograph of extracellular WBV particles (2% phosphotungstic acid, pH 7.4) (courtesy Harald Granzow and Thomas C. Mettenleiter, Friedrich Loeffler Institut, Bundesforschungsinstitut für Tiergesundheit, Greifswald Insel Riems, Germany). (B) Schematic representation of the WBV virion. (C and D) Cross-sections of (C) intra-cytoplasmic nucleocapsids and (D) a virion, with the nucleocapsid seemingly organized by subunits arranged in helical symmetry. (E) Preformed WBV nucleocapsids in the cytoplasm arranged side by side at smooth membranes. All bars represent 100 nm. (C, D and E from Granzow *et al*. (2001). Identification and ultrastructural characterization of a novel virus from fish. *J. Gen. Virol.*, **82**, 2849-2859; with permission.)

PROTEINS

Virions contain the following protein species: (i) the spike protein S, a 1220-aa type I membrane glycoprotein, with features typical for class I fusion proteins (bioinformatical analysis revealed heptad repeat regions and a putative fusion peptide); (ii) a 227-aa integral membrane protein M with three predicted transmembrane regions; (iii) a 161-aa basic nucleocapsid protein.

LIPIDS

WBV acquires its lipid envelope primarily by budding at intracellular membranes and, only rarely, at the plasma membrane.

CARBOHYDRATES

From carbohydrate-specific labelling experiments, the S and M proteins appear to be glycosylated. Glycans on the S protein are recognized by the lectin concanavalin-A and thus are likely to contain α-mannose.

Genome organization and replication

The WBV genome contains five ORFs called ORF 1a, 1b (together comprising the replicase gene) and -2, -3 and -4 (for the spike (S), membrane (M) and nucleocapsid protein (N), respectively). The structural proteins are expressed from three sg mRNA species that are 3′ co-terminal with the genome and believed to be produced via a process of discontinuous minus-strand RNA synthesis similar to that of coronaviruses (Figure 12). Each sg mRNA carries a 42-nt leader sequence identical to the 5′-terminal end of the genome. A conserved nonanucleotide sequence, CA(G/A)CACUAC, located upstream of each structural protein gene and immediately downstream of the genomic 5′-leader sequence, presumably represents the WBV equivalent of the coronavirus TRS core element. Bafinivirus virions assemble through the budding of preformed nucleocapsids at membranes of the ER and/or Golgi complex (Figure 11E) and subsequently egress via exocytosis.
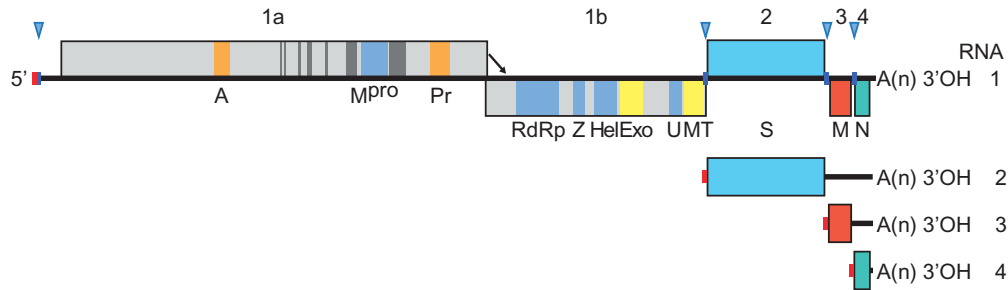
Figure 12: Organization and expression of the WBV genome. Schematic representation of the genome of white bream virus and structural relationship between the genomic RNA and subgenomic mRNAs. 5′ leader sequences are indicated by red boxes, TRSs by blue arrowheads and boxes. Other genome elements and acronyms as in Figure 1.

Antigenic properties

None reported.

Biological properties

White bream virus is the first nidovirus to be isolated from a teleost, white bream (*Blicca bjoerkna L.*), a species of fresh water fish (family *Cyprinidae)*. No information is available on its ecology, biology and pathogenic properties.

Species demarcation criteria in the genus

Newly identified viruses are to be assigned to (or excluded from) the genus *Bafinivirus* on the basis of rooted phylogeny and pair-wise comparisons of *Coronaviridae*-wide conserved domains in replicase polyprotein pp1ab as outlined at the beginning of this chapter.

List of species in the genus *Bafinivirus*

White bream virus
  White bream virus DF 24/00   [DQ898157]   (WBV DF24/00)

Species names are in italic script; names of isolates are in roman script. Sequence accession numbers [ ] and assigned abbreviations ( ) are also listed.

List of other related viruses which may be members of the genus *Bafinivirus* but have not been approved as species

None reported.

List of unassigned species in the subfamily *Torovirinae*

None reported.

List of other related viruses which may be members of the subfamily *Torovirinae* but have not been approved as species

None reported.

Phylogenetic relationships within the family *Coronoviridae*

In rooted and unrooted phylogenetic trees constructed for the main replicative enzymes, members of the family *Coronaviridae* consistently form a monophyletic cluster that is separate from the *Arteri-* and *Roniviridae* (see Chapter *Nidovirales*, Figure 4). The relatively close relationship between members of the family *Coronaviridae* is supported (i) by the presence of unique ORF1a-encoded enzyme

Pos. ssRNA

domains in the replicase polyproteins that might be considered diagnostic molecular markers, i.e. the ADRP and the non-canonical RdRp/putative primase (coronavirus nsp8), and (ii) by similarities in their structural proteins (S and M for all *Coronaviridae*, S, M and N in the *Torovirinae*). On the basis of rooted phylogeny and pair-wise comparisons of *Coronaviridae*-wide conserved replicase domains, four well-separated monophyletic clusters can be distinguished within the subfamily *Coronavirinae*, three of which are established genera (*Alpha-*, *Beta-* and *Gammacoronavirus*). It is anticipated that the remaining cluster, comprised of recently identified avian coronaviruses (Thrush, Bulbul and Munia coronavirus) and related viruses in mammals, will be classified as a new genus (Figure 2). Viruses in the subfamily *Torovirinae* (genera *Bafini-* and *Torovirus*) are phylogenetically more related to each other than to those in the subfamily *Coronavirinae*.

Similarity with other taxa

For features shared with other members of the order *Nidovirales* and with non-nidovirus taxa, please see Chapter *Nidovirales*.

Derivation of names

Corona: from Latin *corona*, "halo"; refers to the characteristic appearance of surface projections that create an image reminiscent of the solar corona.

Toro: from Latin *torus*, a term used in architecture for the convex molding at the base of a column and in geometry for a three-dimensional structure in the shape of a hollow donut; refers to the nucleocapsid morphology in a subset of particles.

Bafini: from *ba*cilliform *fish ni*doviruses, refers to the virion morphology and host tropism.

Further reading

Journals and books

Bárcena, M., Oostergetel, G.T., Bartelink, W., Faas, F.G., Verkleij, A., Rottier, P.J., Koster, A.J. and Bosch, B.J. (2009). Cryo-electron tomography of mouse hepatitis virus: insights into the structure of the coronavirion. *Proc. Natl Acad. Sci., U S A*, **106**, 582–587.

Enjuanes, L., Almazan, F., Sola, I. and Zuñiga, S. (2006). Biochemical aspects of coronavirus replication and virus host-interaction. *Ann. Rev. Microbiol.*, **60**, 211–230.

Gorbalenya, A.E., Enjuanes, L., Ziebuhr, J. and Snijder, E.J. (2006). Nidovirales: evolving the largest RNA virus genome. *Virus Res.*, **117**, 17–37.

Nga, P.T., Parquet, M.D.C., Lauber, C., Parida, M., Nabeshima, T., Yu, F., Thuy, N.T., Inoue, S., Ito, T., Okamoto, K., Ichinose, A., Snijder, E.J., Morita, K. and Gorbalenya, A.E. (2011). Discovery of first insect nidovirus, a missing evolutionary link in the emergence of the largest RNA virus genomes. *PLoS Pathogens*, **7**, in press

Pasternak, A.O., Spaan, W.J. and Snijder, E.J. (2006). Nidovirus transcription: how to make sense …? *J. Gen. Virol.*, **87**, 1403–1421.

Perlman, S., Gallagher, T. and Snijder, E.J. (2008). *Nidoviruses*. Washington, DC, USA: ASM Press.

Perlman, S. and Netland, J. (2009). Coronaviruses post-SARS: update on replication and pathogenesis. *Nat. Rev. Microbiol.*, **7**, 439–450.

Sawicki, S.G., Sawicki, D.L. and Siddell, S.G. (2007). A contemporary view of coronavirus transcription. *J. Virol.*, **81**, 20–29.

Schütze, H., Ulferts, R., Schelle, B., Bayer, S., Granzow, H., Hoffmann, B., Mettenleiter, T.C. and Ziebuhr, J. (2006). Characterization of White bream virus reveals a novel genetic cluster of nidoviruses. *J. Virol.*, **80**, 11598–11609.

Thiel, V. (2007). *Coronaviruses: Molecular and Cellular Biology*. Norfolk, UK: Caister Academic Press.

Woo, P.C.Y., Lau, S.K.P., Huang, Y. and Yean, K.Y. (2009). Coronavirus diversity, phylogeny and interspecies jumping. *Exp. Biol. Med.*, **234**, 1117–1127.

Websites

http://www.viprbrc.org/brc/ (VIPR Virus Pathogen Resource)
http://veb.lumc.nl/SNAD/
http://veb.lumc.nl/SARGENS/

Contributed by

de Groot, R.J., Baker, S.C., Baric, R., Enjuanes, L., Gorbalenya, A.E., Holmes, K.V., Perlman, S., Poon, L., Rottier, P.J.M., Talbot, P.J., Woo, P.C.Y. and Ziebuhr, J.