BACKGROUND

The virus taxonomy

The order *Nidovirales* (*Nido*: from Latin *nides*, meaning "nest") was formally established by the International Committee on Taxonomy of Viruses (ICTV) during the Xth International Congress of Virology (Jerusalem, August 1996). The nidoviruses can be clearly distinguished from other enveloped, positive sense RNA viruses on the basis of the mechanism they use to synthesize subgenomic mRNAs by discontinuous transcription. Although nidoviruses differ from each other in terms of virion morphology and size, they share basic features of the genome structure and pattern of gene expression, which will be described later. The order includes three families: *Coronaviridae, Arteriviridae*, and *Roniviridae*.

The *Coronaviridae* (*Corona*: from Latin *corona* for “crown”) family includes two genera, *Coronavirus* and *Torovirus*. The genus *Coronavirus* has been divided into three groups according to genetic and serological properties. Group 2, to which bovine coronavirus (BCoV) belongs, contains mammalian and avian viruses such as mouse hepatitis virus (MHV), porcine hemagglutinating encephalomyelitis virus (HEV), rat coronavirus (RCV), and human coronavirus strain OC43 (HCV-OC43). Another novel coronavirus, severe acute respiratory syndrome virus (SARS-CoV) cannot be included in any group within the family. The complete genome sequence of SARS-CoV has already been reported (Marra *et al*., 2003; Snijder *et al*., 2003b), but little information is available for this virus in most of its molecular aspects.

The *Arteriviridae* (*Arteri*: from equine *arteritis*) family contains equine arteritis virus (EAV), lactate dehydrogenase-elevating virus (LDV), porcine respiratory and reproductive syndrome virus (PRRSV), and simian hemorrhagic fever virus (SHFV). Of four arteriviruses, both EAV and PRRSV are considered to be socially and/or economically important viruses, infecting horse and swine, respectively.

The *Roniviridae* family and detailed taxonomic structure of the order *Nidovirales* can be found on the website of the International Committee on Taxonomy of Viruses (ICTV) (http://www.ncbi.nlm.nih.gov/ICTVdb/Ictv/ICTVindex.htm; 13-March-2004).

Genome organization and properties

The genomes of both BCoV and EAV are very similar in organization. They are single stranded, positive sense RNA molecules with a length of 12,7 kb for EAV and between 27 and 32 kb for BCoV. The latter has the largest genome of all positive sense RNA viruses. The genomes have a methylated cap structure at the 5’ end and a poly (A) tail at the 3’ end. The RNA starts with a leader sequence of 65-98 nucleotide (nt) for coronaviruses (Lai & Cavanagh, 1997) and 170-211 nt for arteriviruses (Snijder & Spaan, 1995). The number of open reading frames

(ORFs) is different, 9 for EAV and 13 for BCoV (Chouljenko *et al*., 2001). Two large overlapping ORFs, 1a and 1b occupy almost two-thirds of the genomes at the 5’ end and encode replicase polyproteins. The genes encoding structural proteins are located downstream of the replicase genes. For BCoV the genes are arranged in the order 5’ replicase-HE-S-E-M-N-3’, interspersed with genes encoding several non-structural proteins and nonessential genes among them (Lai & Cavanagh, 1997). The gene order of EAV is 5’- replicase -E-GP(2b-3-4-5)-M-N-3’. Most of the EAV genes are overlapping between 3’ end of one ORF and 5’ end of the next, providing sequence signal for subgenomic RNA transcription and a mechanism of regulation of gene expression. The coding regions of genomes are both flanked by 5’ and 3’ untranslated regions (UTRs).

Virus life cycle

The nidoviruses share a common replication cycle as displayed in Fig. 1, using EAV as an example. After attachment to the target cell membrane, the virus enters the cell presumably by membrane fusion. The replication cycle starts with translation of ORFs 1a and 1b from the released genomic RNA using the host cellular machinery. The translation products are two large polyproteins, which are cleaved into 12 non-structural proteins by three virus-encoded proteases (Snijder *et al*., 1994; van Dinten *et al*., 1996, 1999; Wassenaar *et al*., 1997). The expression of ORF1b requires a –1 ribosomal frame shifting just before the ORF1a stop codon (den Boon *et al*., 1991). Two signals in the overlapping regions

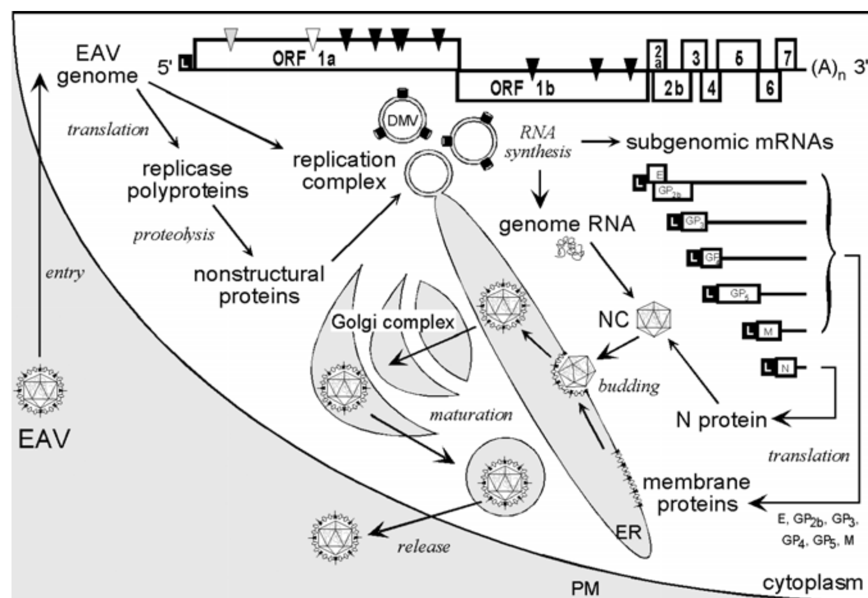

Fig. 1. Overview of Nidovirus life cycle, using EAV as an example (Courtesy of Dr. E. Snijder).

of ORFs 1a and 1b are assumed to facilitate the event (Brierley, 1995): a slippery sequence “5’-GUUAAAC-3’, and a downstream pseudoknot structure. The non-structural proteins, which include an RNA dependent RNA polymerase (RdRp), form the replication complex that promotes copying of the positive sense genomic RNA into a complementary negative sense RNA. The latter serves as a template for synthesis of new genomic RNA molecules, and for synthesis of a nested set of subgenomic mRNAs by a mechanism of discontinuous transcription of the remaining ORFs (Pasternak *et al*., 2001). These subgenomic mRNAs are further translated into seven structural proteins, comprising the major and minor components of the virions. The genomic RNA and structural proteins are assembled at cellular membranes, and the mature viral particles are released into the extracellular medium (Magnusson *et al*., 1970).

The structural proteins

a. BCoV. There are five major structural proteins encoded within the 3’ end of BCoV viral genome: spike (S) glycoprotein, hemagglutinin-esterase (HE) protein, small membrane (E) protein, integral membrane (M) protein, and nucleocapsid (N) protein.

The S protein is the largest membrane glycoprotein of the virion. It comprises 1363 amino acid residues and is cleaved at the position between amino acid 768 and 769 to form two subunits (Abraham *et al*., 1990): S1 (the N terminal half of the S protein), and S2 (the C terminal half of the S protein). The S2 is much more conserved than S1. The S protein contains several antigenic sites, which are targeted by both the humoral immune response and cytotoxic T lymphocytes (Van Regenmortel *et al*., 2000). It predominately induces virus neutralizing antibodies thus becomes a candidate antigen for vaccine. Other important biological functions of this protein include binding to receptors of target cells, ability to infect human rectal tumour (HRT-18) cells (Popova & Zhang, 2002), and mediation of membrane fusion in insect cells (Yoo *et al*., 1991).

The HE protein is a 140 kD envelope-associated glycoprotein made of two disulfide bond-linked 65 kD subunits. The HE glycoprotein has hemagglutinating and receptor destroying enzyme activities.

The N protein is a 50 to 60 kD, highly basic phosphoprotein. Its main function is in viral RNA synthesis, binding to the viral RNA and providing the virion with a helical morphology. This protein can induce a protective cellular immune response, however, the most efficient induction of virus neutralizing antibodies has been achieved with a combination of the S and N proteins (van Regenmortel *et al*., 2000).

Both E and M proteins are essential for viral particle assembly. The M protein has the least immunogenicity among the major viral proteins (Lin *et al*., 2000), but it is involved in the production of IFN type 1, which is an early indicator of infection.

b. EAV. The structural proteins of EAV differ from BCoV in two aspects: the number of glycoproteins and formation of di- or trimer complex of these proteins. EAV particles contain one nucleocapsid protein (N), two major envelop proteins, the membrane (M) protein and glycoprotein 5 (GP5). In addition, three minor structural proteins (GP2b, -3, and -4), and a small envelop (E) protein are found in virions. The M and GP5 proteins associate into disulfide-linked heterodimers, a process that is essential for virus infectivity (Snijder *et al*., 2003a). The GP2b, -3, and -4 occur in the virion as heterotrimeric complexes (Wieringa *et al*., 2003).

The nucleocapsid (N) protein is a small basic phosphoprotein specified by ORF7 (de Vries *et al*., 1992). Like its counterpart in BCoV, the main function of N protein is to package the viral genomic RNA into a isometric core.

Two unglycosylated proteins, the membrane (M) and envelop (E) protein, are encoded by ORF6 and ORF2a, respectively (de Vries *et al*., 1992; Snijder *et al*., 1999). The M protein may be critical for viral particle assembly.

Four glycoproteins (GP2b to GP5) are specified by the corresponding ORFs. The GP5 is the largest protein in the viral envelop (de Vries *et al*., 1992). An important function of this protein is the induction of neutralizing antibody during the immune response. Four distinct neutralization sites have been identified in the GP5 ectodomain, which include sites A (aa 49), B (aa 61), C (aa 67-90), and D (aa 99-106) (Balasuriya *et al*., 1997). The latter three sites are located in the variable region 1 (amino acid 61-121) (Balasuriya *et al*., 1995).

Wieringa *et al*. characterized GP3 and -4 as structural glycoproteins in 2002. The detailed functions of these proteins are still unknown. So is that of the GP2b protein.

Genome diversity

A common feature of RNA viruses is a high mutation rate, and short replication times (Domingo & Holland, 1997). Mutations occur more frequently in RNA than in DNA viruses due to less effective proofreading capacity of the RNA dependent RNA polymerase. Mutation rates differ greatly at different genes or different regions of the same gene, in respect to nucleotide sequence context on the template molecule. Other factors, such as immunological selective pressure, serve as evolutionary forces to accelerate occurrence of mutations.

The main interest of studies on genetic diversity of BCoV is to elucidate the relationship between variations in nucleotide sequence and the properties of strains or isolates such as virulence, plaque size and morphology, resistance/susceptibility to neutralization, biochemical characteristics, and tissue predilection in the host. From the same animal with fatal shipping pneumonia, comparison of the enteric and respiratory BCoV isolates showed 107 nucleotide differences throughout the genome except 5’ UTR and 25 amino acid substitutions in almost all proteins when compared at genomic and amino acid level,

respectively (Chouljenko *et al*., 2001). For the HE gene, sequence differences between WD and CD strains revealed a high level of nucleotide identity, few amino acid differences scattered along the protein sequence, suggesting lack of relationship between nucleotide sequences of these isolates and characteristics of the disease, i.e., respiratory or enteric (Kourtesis *et al*., 2001).

Studies on EAV diversity have mainly focused on ORF5 and GP5, since this protein expresses the known neutralization epitopes, and is under immunological pressure thus being good candidate region to express nucleotide diversity (Balasuriya *et al*., 1997; Stadejek *et al*., 1999). It has been shown that during a 10-year persistent infection of stallions, ORF5 along with ORF3 evolved more rapidly than ORFs 2 and 4, whereas ORFs 6 and 7 were highly conserved (Hedges *et al*., 1999), suggesting that ORFs 5 and 3 are subject to strong selection pressure in the host.

Diseases caused by BCoV and EAV

Bovine coronavirus has been considered as an important pathogen, causing calf diarrhoea (CD), winter dysentery (WD) in adult cattle and respiratory disease in calves. The same strain can cause natural outbreaks of both CD and WD (Traven *et al.*, 2001) and the same animal often shed the virus both via the respiratory and the enteric routes (Chouljenko *et al*., 2001). Respiratory isolates used in experimental infection models have surprisingly caused enteric but no respiratory disease in calves (Cho *et al*., 2001). Although isolates are often referred to as causing CD, WD or pure respiratory disease, respiratory and enteric signs often occur simultaneously or one followed by another. In general, the virus can spread to animals of all ages causing high morbidity, but usually low mortality. This virus is repeatedly reported the Swedish cattle population, causing severe problems in various age groups of animals (Alenius *et al*., 1991; Traven *et al*., 1999).

EAV is transmitted primarily through the respiratory and venereal routes. In general, naturally infected horses show little or no clinical signs, but the outcome of infection can vary greatly, i.e. systemic disease, arteritis, abortion, and foetal death, depending on the virus strain and the animal's condition. An important aspect of infection with EAV is that around 60% of stallions may become persistently infected for several years or a lifetime. These stallions serve as a natural reservoir of virus and can disseminate virus to susceptible mares during breeding (Timoney & McCollum, 1993).

The aim of present study

1. Genetic comparison of BCoV isolates from Sweden and Demark and stability of BCoV in enteric and respiratory infections.

2. To study the genetic stability of ORFs encoding EAV glycoproteins during an experimental infection of horses.

General conclusions

A nested RT-PCR method targeting the HE gene of BCoV was developed, which can be successfully applied for the diagnosis of BCoV infection. A conventional PCR was used to amplify the S gene of BCoV to study genetic stability of the virus. Sequence analysis of the S gene showed that this gene is less conserved than the HE gene. S gene sequencing did not reveal differences between viruses from nasal and fecal swabs originating from the same animal. Future studies should address a detailed profile of BCoV diversity. This could be obtained if the whole sequence of S gene, or at least the fragment containing the two known neutralization domains, is amplified and sequenced for analysis.

Analysis of ORFs encoding EAV glycoproteins has shown the occurrence of mutations in the course of an experimental infection. However, while ORFs 2b and 3 had stabilizing mutations, ORFs 4 and 5 were subject to positive selection. Mutations in ORF5 mainly clustered in variable region 1, resulting in new genetic variants with changed neutralization domain in the GP5. The new virus variants, however could not be entirely related to induction of recurrent viraemia in the infected horses, which is presumed to have been caused by other factors. It would be of interest to investigate the occurrence of other forms of persistent infection, rather than carrier stallions, upon infection with EAV.

References

Abraham S. Kienzle. T.E. Lapps W. & Brian D.A. (1990). Deduced sequence of the bovine coronavirus spike protein and identification of the internal proteolytic cleavage site. Virology 176, 296-301.

Alenius S, Niskanen R, Juntti N & Larsson B. (1991). Bovine coronavirus as the causative agent of winter dysentery: serological evidence. Acta Vet Scand 32, 163-170.

Balasuriya UB, Maclachlan NJ, De Vries AA, Rossitto PV & Rottier PJ. (1995). Identification of a neutralization site in the major envelope glycoprotein (GL) of equine arteritis virus. Virology. 207, 518-527.

Balasuriya UB, Patton JF, Rossitto PV, Timoney PJ, McCollum WH & MacLachlan NJ. (1997). Neutralization determinants of laboratory strains and field isolates of equine arteritis virus: identification of four neutralization sites in the amino-terminal ectodomain of the G(L) envelope glycoprotein. Virology 232, 114-128.

Brierley I. (1995). Ribosomal frameshifting viral RNAs. J Gen Virol 76, 1885-1892.

Cho KO, Hasoksuz M, Nielsen PR, Chang KO, Lathrop S & Saif LJ. (2001). Cross-protection studies between respiratory and calf diarrhea and winter dysentery coronavirus strains in calves and RT-PCR and nested PCR for their detection. Arch Virol. 146, 2401-2419.

Chouljenko VN, Lin XQ, Storz J, Kousoulas KG & Gorbalenya AE. (2001). Comparison of genomic and predicted amino acid sequences of respiratory and enteric bovine coronaviruses isolated from the same animal with fatal shipping pneumonia. J Gen Virol. 82, 2927-2933.

de Vries AA, Chirnside ED, Horzinek MC & Rottier PJ. (1992). Structural proteins of equine arteritis virus. J Virol. 66, 6294-6303.

den Boon JA, Snijder EJ, Chirnside ED, de Vries AA, Horzinek MC & Spaan WJ. (1991). Equine arteritis virus is not a togavirus but belongs to the coronaviruslike superfamily. J Virol. 65, 2910-2920.

Domingo E & Holland JJ. (1997). RNA virus mutations and fitness for survival. Annu Rev Microbiol. 51,151-178.

Hedges JF, Balasuriya UB, Timoney PJ, McCollum WH & MacLachlan NJ. (1999). Genetic divergence with emergence of novel phenotypic variants of equine arteritis virus during persistent infection of stallions. J Virol. 73, 3672-3681.

Kourtesis AB, Gelinas AM & Dea S. (2001). Genomic and antigenic variations of the HE glycoprotein of bovine coronaviruses associated with neonatal calf diarrhea and winter dysentery. Arch Virol. 146, 1219-1230.

Lai MM & Cavanagh D. (1997). The molecular biology of coronaviruses. Adv Virus Res. 48,1-100.

Lin XQ, O eilly KL, Storz J, Purdy CW & Loan RW. (2000). Antibody responses to respiratory coronavirus infections of cattle during shipping fever pathogenesis. Arch Virol. 145, 2335-2349.

Magnusson P, Hyllseth B & Marusyk H. (1970). Morphological studies on equine arteritis virus. Arch Gesamte Virusforsch. 30, 105-112.

Marra MA, Jones SJ, Astell CR & other 56 authors. (2003). The genome sequence of the SARS-associated coronavirus. Science 300, 1399-1404.

Pasternak AO, van den Born E, Spaan WJ & Snijder EJ. (2001). Sequence requirements for RNA strand transfer during nidovirus discontinuous subgenomic RNA synthesis. EMBO J. 20, 7220-7228.

Popova R & Zhang X. (2002). The spike but not the hemagglutinin/esterase protein of bovine coronavirus is necessary and sufficient for viral infection. Virology 294, 222-236.

Snijder EJ & Spaan WJM. (1995). The coronaviruslike superfamily. In The Coronaviridae. ed. Siddell, Plenum Press, pp. 239-255.

Snijder EJ, Bredenbeek PJ, Dobbe JC, Thiel V, Ziebuhr J, Poon LL, Guan Y, Rozanov M, Spaan WJ & Gorbalenya AE. (2003b). Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage. J Mol Biol 331, 991-1004.

Snijder EJ, Dobbe JC & Spaan WJ. (2003a). Heterodimerization of the two major envelope proteins is essential for arterivirus infectivity. J Virol. 77, 97-104.

Snijder EJ, van Tol H, Pedersen KW, Raamsman MJ & de Vries AA. (1999). Identification of a novel structural protein of arteriviruses. J Virol. 73, 6335-6345.

Snijder EJ, Wassenaar AL & Spaan WJ. (1994). Proteolytic processing of the replicase ORF1a protein of equine arteritis virus. J Virol. 68, 5755-5764.

Stadejek T, Bjorklund H, Bascunana CR, Ciabatti IM, Scicluna MT, Amaddeo D, McCollum WH, Autorino GL, Timoney PJ, Paton DJ, Klingeborn B & Belak S. (1999). Genetic diversity of equine arteritis virus. J Gen Virol. 80 ,691-699.

Timoney PJ & McCollum WH. (1993). Equine viral arteritis. Vet Clin North Am Equine Pract 9, 295-309.

Traven M, Bjornerot L & Larsson B. (1999). Nationwide survey of antibodies to bovine coronavirus in bulk milk from Swedish dairy herds. Vet Rec 144, 527-529.

Traven M, Naslund K, Linde N, Linde B, Silvan A, Fossum C, Hedlund KO & Larsson B. (2001). Experimental reproduction of winter dysentery in lactating cows using BCV -- comparison with BCV infection in milk-fed calves. Vet Microbiol. 81,127-151.

van Regenmortel M, Fauguet C, Dave Bishop, Carsten M, Lemon S, Maniloff J, Mayo M, McGeoch D, Pringle C & Wichner R. (2000). Virus Taxonomy. Seventh Report of the International Committee on Taxonomy of Viruses. Academic Press. UK.

van Dinten LC, Rensen S, Gorbalenya AE & Snijder EJ. (1999). Proteolytic processing of the open reading frame 1b-encoded part of arterivirus replicase is mediated by nsp4 serine protease and Is essential for virus replication. J Virol. 73, 2027-2037.

van Dinten LC, Wassenaar AL, Gorbalenya AE, Spaan WJ & Snijder EJ. (1996). Processing of the equine arteritis virus replicase ORF1b protein: identification of cleavage products containing the putative viral polymerase and helicase domains. J Virol. 70, 6625-6633.

Wassenaar AL, Spaan WJ, Gorbalenya AE & Snijder EJ. (1997). Alternative proteolytic processing of the arterivirus replicase ORF1a polyprotein: evidence that NSP2 acts as a cofactor for the NSP4 serine protease. J Virol. 71, 9313-9322.

Wieringa R, de Vries AA & Rottier PJ. (2003). Formation of disulfide-linked complexes between the three minor envelope glycoproteins (GP2b, GP3, and GP4) of equine arteritis virus. J Virol 77, 6216-6226.
Yoo DW, Parker MD & Babiuk LA. (1991). The S2 subunit of the spike glycoprotein of bovine coronavirus mediates membrane fusion in insect cells. Virology 180, 395-359.

Comparative nucleotide sequence analysis of Bovine Coronavirus (BCoV) Isolates from Sweden and Denmark

Lihong Liu[1], Mikhayil Hakhverdyan[2], Sara Hägglund[3], Lars Erik Larsen[4], Stefan Alenius[3], and Sándor Belák[1,2*]

[1]Department of Biomedical Sciences and Veterinary Public Health, Swedish University of Agricultural Sciences, 751 89 Uppsala, Sweden. [2]Department of Virology, National Veterinary Institute, 751 89 Uppsala, Sweden. [3]Department of Ruminant Medicine and Veterinary Epidemiology, Swedish University of Agricultural Sciences, 750 07 Uppsala, Sweden. [4]Department of Veterinary Diagnostic and Research, Danish Institute for Food and Veterinary Research, V-1790, Copenhagen, Denmark.

Keywords: Bovine coronavirus, genetic stability, RT-PCR, sequencing

* Corresponding author. Mailing address: Department of Virology, The National Veterinary Institute, Ulls väg 2 B, S-751 89 Uppsala, Sweden. Phone: (46) 18 67 41 35. Fax: (46) 18 67 46 69. E-mail: Sandor.Belak@sva.se

Abstract

Bovine Coronavirus (BCoV), a member of genus *Coronavirus* in the family *Coronaviridae,* is an important pathogen in cattle of all ages. Known as a cause of diarrhoea in adult animals (winter dysentery), the virus is also frequently reported as an enteric and respiratory pathogen in calves. BCoV infections often result in high morbidity and outbreaks may cause considerable economical losses. In this study, a nested RT-PCR targeting the HE gene was developed for the detection of BCoV in clinical samples from cattle. The assay was shown to be more sensitive than a conventional PCR, targeting the S gene of BCoV. Nucleotide sequence analysis demonstrated that the amplified region of the HE gene is highly conserved, suggesting it is a suitable target for diagnostic purposes. The S gene, which is less conserved, was chosen for comparative nucleotide sequence analyses of BCoV isolates originating from Swedish and Danish outbreaks. The results indicated few or no nucleotide differences between isolates collected from different animals in the same herd on the same occasion, between nasal discharge and faecal samples obtained from the same animal, and between samples collected at different time points during the course of an infection. These data suggest that the viral RNA is relatively stable during infection. A genetic heterogeneity, in contrast, was demonstrated among isolates obtained from different herds and from outbreaks occurring in different years in the same herd, implying that virus tracing is possible. These data will contribute to an extended understanding of BCoV epidemiology.

Introduction

Bovine coronavirus (BCoV) was first reported as a cause of diarrhoea in newborn calves (Mebus *et al*., 1973). It is now considered as an important pathogen in cattle of all ages, causing calf diarrhoea (CD), winter dysentery (WD) in adult animals (Alenius *et al*., 1991) and respiratory disease in calves (Storz *et al*., 2000). BCoV infections often result in high morbidity, but usually in low mortality. The clinical signs are found to be practically indistinguishable from those of other enteric or respiratory infections. The same strain can cause both CD and WD (Traven *et al*., 2001). Furthermore, the same animal can be infected with both enteric and respiratory strains of BCoV (Chouljenko *et al*., 2001). The respiratory and enteric signs of the disease may occur simultaneously or one followed by another.

BCoV is a member of the genus *Coronavirus*, family *Coronaviridae*. This genus can be divided into three groups according to genetic and serological properties. Group 2, to which BCoV belongs, contains mammalian and avian viruses, such as mouse hepatitis virus (MHV), turkey coronavirus (TCV), sialodacryadenitis virus (SDAV), porcine hemagglutinating encephalomyelitis virus (HEV), and human coronavirus OC43 (HCoV-OC43). One character feature of the group 2 is the

presence of a gene encoding hemagglutinin-esterase protein, which is absent in both groups 1 and 3.

The BCoV virion is enveloped and spherical in shape. The genome is a single-stranded, positive-sense RNA molecule of 27 to 32 kb in size, the largest among all RNA viruses. It includes 13 open reading frames (ORF) flanked by 5'- and 3'-untranslated regions (UTRs). Some of the ORFs are overlapping, whereas others are separated by intergenic sequences (Chouljenko *et al*., 2001). The ORF 1a and 1b encode polyproteins, which are further cleaved to form an active mature RNA polymerase and other non-structural proteins. There are five major structural proteins encoded within the genomic RNA: spike (S) glycoprotein (ORF4), membrane (M) protein (ORF9), nucleocapsid (N) protein (ORF10), hemagglutinin-esterase (HE) protein (ORF3) and small membrane protein (E) (ORF8). The rest of the ORFs encode unknown or not well characterized non-structural proteins.

The HE protein is a 140 kD envelope-associated glycoprotein made of two disulfide bond-linked 65 kD subunits. The HE glycoprotein is responsible for hemagglutination and receptor destroying enzymatic activities.

The S protein, the largest membrane glycoprotein of BCoV, forms club-shaped structures on the surface of the virion (Stair *et al*., 1972). It comprises 1363 amino acid and is cleaved at the position between amino acid 768 and 769 to form two subunits (Abraham *et al*., 1990): S1 (the N terminal half of the S protein), and S2 (the C terminal half of the S protein). The S protein is widely investigated because it predominately induces virus neutralizing antibodies and thus becomes a candidate antigen for vaccine development. It also contains several antigenic sites, which are targeted by both of the humoral immune response and cytotoxic T lymphocytes (CTL; Van Regenmortel *et al*., 2000).

Tsunemitsu & Saif (1995) have shown that there are differences in antigenic and biological properties among BCoV CD and WD strains, but that these differences are unrelated to the clinical source of strains. The same situation was observed among bovine enteric and respiratory coronaviruses by using monoclonal antibodies against the HE and S proteins (Hasoksuz *et al*., 1999). The molecular analysis of S1 subunit has shown only a few amino acid changes among paired isolates obtained from the enteric and the respiratory tracts from the same animal (Hasoksuz *et al*., 2002). Comparison of the HE amino acid sequences revealed 18 amino acid differences between the reference Mebus strain and BCoV isolates associated with outbreak of respiratory disease in Canadian diary farms (Gelinas *et al*., 2001). Genomic and predicted amino acid sequences have been compared between respiratory and enteric strains, isolated from the same animal with fatal shipping pneumonia. These studies identified 107 nucleotide substitutions throughout the genome, which led to 24 amino acid changes (Chouljenko *et al*., 2001).

BCoV is repeatedly reported in the Swedish cattle population, causing severe problems in animals of various age groups (Alenius *et al*., 1991; Traven *et al*.,

1999). The goal of this study was to develop a sensitive diagnostic tool for the detection of BCoV and to perform a genetic comparison between isolates collected during outbreaks of disease in Sweden and Demark. Virus stability was assessed by comparing isolates originating from different animals within and between herds, from the same animal during the course of the infection and from the respiratory and enteric tract of the same animal.

Materials and Methods

Bovine clinical samples

A total of 80 samples were collected from different herds with diarrhoea and/or respiratory clinical signs in Sweden and Denmark. The samples, comprising nasal swabs and faeces, were stored at -70 °C until analysed.

RNA extraction

RNA was extracted with TRIzol LS Reagent (Invitrogen, USA), according to the manufacturer's instructions. Briefly, 250 µl of samples were mixed with 750 µl of TRIzol and incubated for 5 min at room temperature. Afterwards, 200 µl of chloroform were added, and the tubes were shaken vigorously for 15 sec and centrifuged at 12 300 rpm (Sigma, Germany) for 15 min. RNA was collected from the supernatant and precipitated overnight with 500 µl of isopropanol at -20°C. The precipitated RNA was centrifuged at 12 300 rpm for 20 min at 4°C and washed with 1 ml of cold 75% ethanol. The pellets were air dried, resuspended in 30 µl of DMPC water and stored at -20°C for further use.

Synthesis of cDNA

A mix of 5 µl RNA, 1 µl of random hexamers (Pharmacia) and 5 µl of DMPC water was denatured at 65°C for 10 min and then chilled on ice. The reverse transcription mix comprised 4.5 µl of DMPC water, 5 µl of 5X First standard buffer (Invitrogen), 2.5 µl of 2 mM dNTP mix (Amersham Biosciences), 32U of RNAguard (Amersham Biosciences) and 200U of Moloney murine leukaemia virus reverse transcriptase (Invitrogen). The total volume of reaction mix was 25 µl. The reverse transcription was performed at 37°C for 90 min followed by the inactivation of the enzyme at 95°C for 5 min.

Primers and PCR amplification of HE and S genes

Primers for BCoV nested PCR targeting HE gene were selected by using Primer Express software (version 1.0, Applied Biosystems, Foster City, CA, USA). Published primer sequences (Hasoksuz *et al.*, 2002) were used for the conventional PCR to amplify S gene. The nucleotide sequence and positions of the primers used were presented in Table 1.

The reaction mix was same for both the nested and conventional PCR, which contained 24 µl of sterile distilled water, 5 µl of 10X PCR buffer, 1 µl of 10 mM dNTP mix, 5 µl of 1 mg/ml BSA, 1.5 µl of each 10 µM primer, 5 µl of 25 mM $MgCl_2$, 1U of *Taq* DNA polymerase (Ampli Taq, Perkin-Elmer) and 5 µl of cDNA or first PCR product. Two drops of mineral oil were added to overlay the reaction mix. Thermocyling profile included initial denaturation at 94°C for 2 min, followed by 35 cycles of denaturation at 94°C for 45 sec, annealing at 50°C for 60 sec, and extension at 72°C for 90 sec and a final extension at 72°C for 7 min. The PCR products (8 µl) were run on 2% agarose gel stained with ethidium bromide and visualized under UV light.

Table 1. Primer sequences for the nested and conventional PCR of BCoV HE and S genes.

Name	Position[a]	Nucleotide sequence (5'-3')	Targeting gene
BCV-727F[b]	727-751	ACTGAAACCATTACCACTGGTTTTG	HE, PCR I
BCV-1286R[c]	1286-1264	GCATCATGCAGCCTAGTACCATT	HE, PCR I
BCV-819F	819-843	AACTGTTCCTACKAAAGCAATCTGT	HE, PCR II
BCV-1225R	1225-1208	TGACCGCRACACCCAAAA	HE, PCR II
S1AF[d]	1-21	ATGTTTTTGATACTTTTAATT	S
S1AR[d]	654-635	AGTACCACCTTCTTGATAAA	S

[a] Nucleotide positions are based on BCoV Quebec isolate BCQ 7373 (HE gene) and Mebus (S gene) strains (GenBank accession No. AF239307 and M31053, respectively).
[b] F – forward primer
[c] R – reverse primer
[d] S1AF and S1AR were taken from published primer sequences (Hasoksuz *et al*., 2002)

Purification of PCR products

The PCR products were purified by using Qiaquick PCR purification kit (Qiagen), if only one band was seen under UV light, or Gel Extraction Kit (Qiagen), if two or more bands appeared in the gel, according to manufacturer's instructions. Briefly, five volumes of binding buffer PB were added to one volume of PCR samples and applied to spin column. The samples were centrifuged at 13 000 rpm for 1 min to bind PCR product into the silica gel membrane in the presence of high salt concentration. The membrane was washed once with 750 µl of buffer PE by centrifuging for 1 min at 13 000 rpm. After drying the membrane by centrifugation, the DNA was eluted in 30 µl of sterile water.

For purification with Gel Extraction Kit, the PCR product was first separated on agarose gel, and the band of correct size was excised using a clean scalpel and put in a microcentrifuge tube. Three volumes of buffer QG was added to one volume of gel slice, and incubated at 50 °C until the gel has completely dissolved. One volume of isopropanol was added to the tube. The samples was applied to the column and centrifuged at 13 000 rpm for 1 min to bind DNA in the membrane. The washing and elution steps were same as above.

Sequencing and sequence analysis

An ABI PRISM BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems) was used for sequencing reactions, which were performed in a 20 µl volume containing 2 µl of BigDye, 1.8 µl of forward or reverse primer, 6 µl of half-Buffer, appropriate amount of DNA template, and filtered Super-Q water. The thermo profile was 25 cycles of denaturation at 96°C for 15 sec, annealing at 50°C for 10 sec, and extension at 60°C for 4 min. The extension products were precipitated with 2 µl of 3 M sodium acetate (pH 4.6), and 50 µl of cold 95% ethanol. After incubation at room temperature for 15 min, the tube was centrifuged at 13 000 rpm for 20 min, and pellets were washed once with 250 µl of 70% ethanol. The pellets were dried at 37°C under a hood. Afterwards, 12 µl of formamide was added to the tube and vortexed just before capillary electrophoresis. Automated sequencing was carried out in an ABI 3100 genetic analyzer (Applied Biosystems). Sequence data were collected automatically with the software provided by the manufacturer. Sequences were analysed with multiple programs of the DNASTAR package (DNASTAR v5.0.3, Madison, WI, USA).

Results

PCR amplification of HE and S genes

The first and second round of amplification of the nested PCR (HE gene) targeted a fragment of 560 bp and 407 bp, respectively (Fig. 1A, B). Sixty out of 80 samples were detected positive by this method, including 38 nasal swabs and faecal samples from Sweden and 22 fecal samples from eight Danish herds. All samples were used to amplify the S gene for sequencing. Fifty-two samples were detected positive by targeting the S gene, eight samples less than that by targeting the HE gene. The nested PCR was more sensitive than conventional PCR.

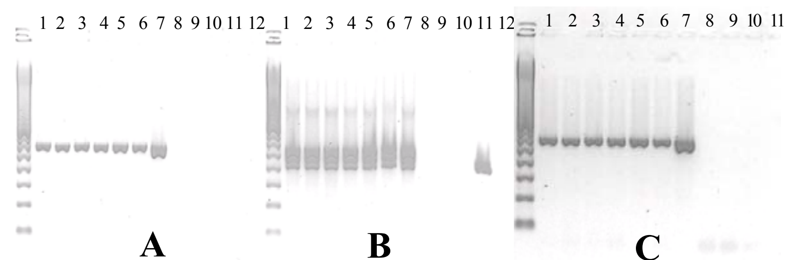

Fig. 1. Representative examples of PCR amplification of BCoV HE and S genes from Sweden. HE gene PCR I (panel A), PCR II (B), and S gene (C) were amplified with positive control (pool 7) and negative control (pool 12; included only in panel A and B). Samples in pool 1-6 were detected positive for both HE gene and S gene tests, and three (pool 8-10) were negative. One sample (pool 11) was negative in first PCR test for HE gene and S gene, but positive in the second PCR test for HE gene.

S gene sequence clustering in relation to date and origin of isolates

An alignment of sequences of the 624 bp fragment of the S gene was performed by the Clustal W method, provided within the MegAlign program. Alignment results showed that there were less than 5% nucleotide substitutions among field isolates from Denmark and Sweden, indicating that the amplified part of the S gene was also relatively conserved. The sequences clustered together, as shown in Fig. 2.

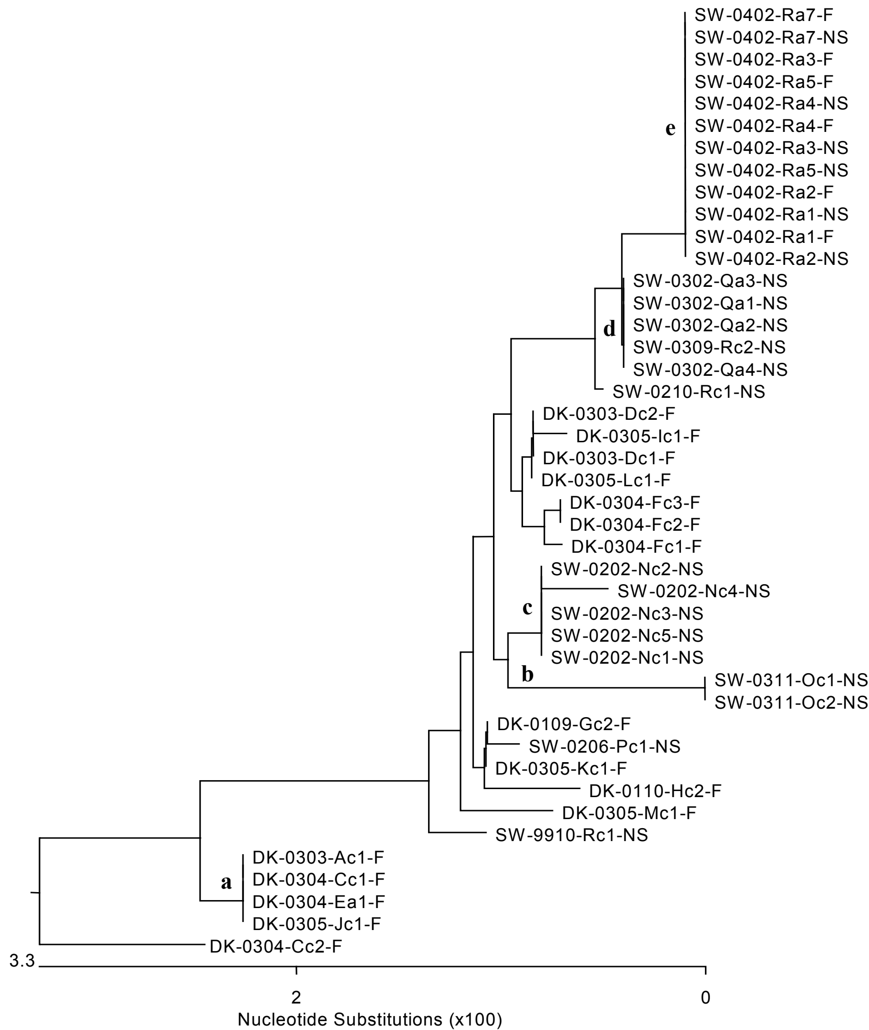

Fig. 2. Phylogenetic tree based on a 624 bp fragment of the S gene. BCoV isolates originated from herds associated with respiratory and/or enteric disease in Sweden and Denmark. The tree was generated from sequence alignments by the clustal W method within the multiple program package DNASTAR. Bold **a** to **e** indicates different clusters. SW and DK refer to the country (Sweden and Denmark) of origin. Year and month of isolation are preceding a capital letter for the herd (A to R), the age (c; calf, a; adult), identity of the sample (1-7) and type of sample (F; fecal samples, NS; nasal swabs).

Members of cluster **a** were obtained from four different herds in Denmark and grouped far from the others. The samples were collected from three calves and an adult animal within a three month period in 2003. All of the samples in cluster **a** had identical nucleotide sequences along the analysed region, suggesting that these four animals were probably infected with the same virus.

Cluster **b** contained two identical sequences from a Swedish herd associated with severe respiratory and enteric signs and high mortality in calves. Sequence divergence of these two viruses was higher within enteric than respiratory strains, suggesting that they were closer to a respiratory than to a enteric strain.

Five sequences were grouped in cluster **c**. The samples (nasal swabs) were collected from four naturally infected animals with mild enteric and/or respiratory disease in one Swedish herd. Two samples (SW-0202-Nc1-NS and SW-0202-Nc5-NS) were obtained from the same animal five days apart. SW-0202-Nc2-NS and SW-0202-Nc4-NS were collected on day 0, SW-0202-Nc1-NS on day 3 and SW-0202-Nc5-NS on day 8. Except for one (SW-0202-Nc4-NS), the other sequences were identical.

Five nucleotide sequences in cluster **d** originated from two Swedish herds. In herd Q, severe enteric disease was the main clinical sign and faecal samples were taken from adult cattle. The virus SW-0309-Rc1-NS was found seven months later in another herd (R) from the same area of Sweden and derived from a calf with respiratory disease and fever. The overall clinical signs observed in the latter herd (consisting only of young animals) were soft faeces, cough and nasal discharge. All five sequences had a 100% homology, indicating that the same virus was likely to be transmitted between the two herds and that this virus gave rise to different clinical signs in different age groups of animals. Another virus (SW-0210-Rc1-NS), which were obtained from herd R two years earlier, showed a diverging sequence.

Cluster **e** consisted of twelve paired respiratory and enteric isolates form 6 adult cattle in the same herd in Sweden. All of the sequences shared a 100% identity.

Discussion

Different methods have been developed for the detection of BCoV recently (Fukutomi *et al.*, 1999; Gaber & Kapil, 1999; Naslund *et al.*, 2000; Cho *et al.*, 2001). Of those, RT-PCR and nested PCR targeting the N gene have been developed for the detection of BCoV (Cho *et al.*, 2001). However, the nested RT-PCR method targeting the HE gene of BCoV has an apparent advantage over that targeting the N gene: the HE gene is only presented in the members of the group 2 coronavirus. Thus, a higher specificity is expected by targeting the HE gene rather than the N gene. The sequence alignment of 31 samples, based on a 407 bp fragment of HE gene, has shown that this part of genome is so highly conserved that there were only a few nucleotide substitutions scattered throughout HE gene

(data not shown). The similar results have been shown by other groups (Kourtesis *et al*., 2001; Gelinas *et al*., 2001). Thus the HE gene, we suggest, is not a appropriate target for the genetic comparison of different isolates, but it can be used for diagnostic purposes. In contrast, the S gene is less conserved than the HE gene and should be suitable for genetic comparison. The sequence alignment, based on a 624 bp fragment of the S gene of BCoV, showed that the analysed BCoV differ in 39 positions.

The genetic heterogeneity of BCoV was marked in the Danish samples. Some of viruses from different herds clustered together and showed the same nucleotide sequence, whereas the others from the same herd fall into different clusters. For example, one calf in herd C was clustered with three animals from different herds; however, the second calf in herd C was infected with a BCoV showing the largest divergence of S gene from others. The reason for this heterogeneity may be the lack of strict bio-security control within farms, which led to virus spread and transmission among different farms. Further investigation are needed to understand the BCoV epidemiology in Denmark.

A study on genetic stability of BCoV was also carried out by comparing the nucleotide sequence of the S gene from 12 paired nasal and faecal samples from 6 animals in Sweden. Our results failed to demonstrate any nucleotide differences between nasal and faecal samples. Since the S glycoprotein is a major neutralizing antigen, it is reasonable that there should be some mutations in the S gene of BCoV under immunological selective pressure of the host. Indeed, this has been confirmed *in vitro* by Yoo & Deregt (2001), who found that an amino acid change within antigenic domain II of the BCoV S protein, caused by a single nucleotide change in the S gene, confers virus resistance to monoclonal antibodies' neutralization. It is of interest to analyses this region of the S gene for a better understanding of BCoV stability *in vivo*, especially in enteric and respiratory systems of the same animal.

In conclusion, the nested PCR targeting HE gene is suitable for diagnosis of BCoV rather than genetic diversity studies. The S gene is less conserved than the HE gene and therefore more appropriate for analysis genetic stability and evolutionary aspects in BCoV. Using S gene analysis we have shown that isolates of BCoV occurring in Denmark are rather heterogeneous. Also, comparison of S gene sequences from BCoV origination from nasal or faecel samples did not show strain specificity with regards to the clinical signs.

Acknowledgments

This project was supported by the Swedish Farmers' Foundation for Agricultural Research (SLF). We are grateful to Alia Yacoub for technical assistance, and Dr. Claudia Baule for her critical review of this manuscript.

References

Abraham S, Kienzle TE, Lapps W & Brian DA. (1990). Deduced sequence of the bovine coronavirus spike protein and identification of the internal proteolytic cleavage site. Virology 176, 296-301.

Alenius S, Niskanen R, Juntti N & Larsson B. (1991). Bovine coronavirus as the causative agent of winter dysentery: serological evidence. Acta Vet Scand 32, 163-170.

Chouljenko VN, Lin XQ, Storz J, Kousoulas KG & Gorbalenya AE. (2001). Comparison of genomic and predicted amino acid sequences of respiratory and enteric bovine coronaviruses isolated from the same animal with fatal shipping pneumonia. J Gen Virol 82, 2927-2933.

Fukutomi T, Tsunemitsu H & Akashi H. (1999). Detection of bovine coronaviruses from adult cows with epizootic diarrhea and their antigenic and biological diversities. Arch Virol 144, 997-1006.

Gaber F & Kapil S. (1999). Development of an antigen spot test for detection of coronavirus in bovine fecal samples. Clin Diagn Lab Immunol 6, 542-544.

Gelinas AM, Boutin M, Sasseville AM & Dea S. (2001). Bovine coronaviruses associated with enteric and respiratory diseases in Canadian dairy cattle display different reactivities to anti-HE monoclonal antibodies and distinct amino acid changes in their HE, S and ns4.9 protein. Virus Res 76, 43-57.

Hasoksuz M, Lathrop S, Al-dubaib MA, Lewis P & Saif LJ. (1999). Antigenic variation among bovine enteric coronaviruses (BECV) and bovine respiratory coronaviruses (BRCV) detected using monoclonal antibodies. Arch Virol 144, 2441-2447.

Hasoksuz M, Sreevatsan S, Cho KO, Hoet AE & Saif LJ. (2002). Molecular analysis of the S1 subunit of the spike glycoprotein of respiratory and enteric bovine coronavirus isolates. Virus Res 84, 101-109.

Mebus CA, Stair EL, Rhodes MB & Twiehaus MJ. (1973). Neonatal calf diarrhea: propagation, attenuation, and characteristics of a coronavirus-like agent. Am J Vet Res 34, 145-50.

Naslund K, Traven M, Larsson B, Silvan A & Linde N. (2000). Capture ELISA systems for the detection of bovine coronavirus-specific IgA and IgM antibodies in milk and serum. Vet Microbiol 72, 183-206.

Stair EL, Rhodes MB, White RG & Mebus CA. (1972). Neonatal calf diarrhea: purification and electron microscopy of a coronavirus-like agent. Am J Vet Res 33, 1147-1156.

Storz J, Purdy CW, Lin X, Burrell M, Truax RE, Briggs RE, Frank GH & Loan RW. (2000). Isolation of respiratory bovine coronavirus, other cytocidal viruses, and Pasteurella spp from cattle involved in two natural outbreaks of shipping fever. J Am Vet Med Assoc 216, 1599-1604.

Traven M, Bjornerot L & Larsson B. (1999). Nationwide survey of antibodies to bovine coronavirus in bulk milk from Swedish dairy herds. Vet Rec 144, 527-529.

Traven M, Naslund K, Linde N, Linde B, Silvan A, Fossum C, Hedlund KO & Larsson B. (2001). Experimental reproduction of winter dysentery in lactating cows using BCV -- comparison with BCV infection in milk-fed calves. Vet Microbiol 81, 127-151.

Tsunemitsu H & Saif LJ. (1995). Antigenic and biological comparisons of bovine coronaviruses derived from neonatal calf diarrhea and winter dysentery of adult cattle. Arch Virol 140, 1303-1311.

van Regenmortel M, Fauguet C, Dave Bishop, Carsten M, Lemon S, Maniloff J, Mayo M, McGeoch D, Pringle C & Wichner R. (2000). Virus Taxonomy. Seventh Report of the International Committee on Taxonomy of Viruses. Academic Press. UK.

Yoo D & Deregt. (2001). A single amino acid change within antigenic domain II of the spike protein of bovine coronavirus confers resistance to virus neutralization. Clin Diagn Lab Immunol. 8, 297-302.

Zhang XM, Kousoulas KG & Storz J. (1991). Comparison of the nucleotide and deduced amino acid sequences of the S genes specified by virulent and avirulent strains of bovine coronaviruses. Virology. 183, 397-404.