1-1-2001

Comparison of genomic and predicted amino acid sequences of respiratory and enteric bovine coronaviruses isolated from the same animal with fatal shipping pneumonia

V. N. Chouljenko
School of Veterinary Medicine

X. Q. Lin
School of Veterinary Medicine

J. Storz
School of Veterinary Medicine

K. G. Kousoulas
School of Veterinary Medicine

A. E. Gorbalenya
School of Veterinary Medicine

Recommended Citation

Chouljenko, V., Lin, X., Storz, J., Kousoulas, K., & Gorbalenya, A. (2001). Comparison of genomic and predicted amino acid sequences of respiratory and enteric bovine coronaviruses isolated from the same animal with fatal shipping pneumonia. *Journal of General Virology, 82* (12), 2927-2933. https://doi.org/10.1099/0022-1317-82-12-2927

Journal of General Virology (2001), **82**, 2927–2933. Printed in Great Britain SHORT COMMUNICATION

Comparison of genomic and predicted amino acid sequences of respiratory and enteric bovine coronaviruses isolated from the same animal with fatal shipping pneumonia

Vladimir N. Chouljenko,[1] X. Q. Lin,[1] J. Storz,[1] Konstantin G. Kousoulas[1] and Alexander E. Gorbalenya[2]

[1] Department of Pathobiological Sciences, School of Veterinary Medicine, Louisiana State University, Baton Rouge, LA 70803, USA

[2] Advanced Biomedical Computing Center, Science Application International Corporation, National Cancer Institute, Frederick, MD 21702-1201, USA

The complete genome sequences are reported here of two field isolates of bovine coronavirus (BCoV), which were isolated from respiratory and intestinal samples of the same animal experiencing fatal pneumonia during a bovine shipping fever epizootic. Both genomes contained 31028 nucleotides and included 13 open reading frames (ORFs) flanked by 5′- and 3′-untranslated regions (UTRs). ORF1a and ORF1b encode replicative polyproteins pp1a and pp1ab, respectively, that contain all of the putative functional domains documented previously for the closest relative, mouse hepatitis virus. The genomes of the BCoV isolates differed in 107 positions, scattered throughout the genome except the 5′-UTR. Differences in 25 positions were non-synonymous and were located in all proteins except pp1b. Six replicase mutations were identified within or immediately downstream of the predicted largest pp1a-derived protein, p195/p210. Single amino acid changes within p195/p210 as well as within the S glycoprotein might contribute to the different phenotypes of the BCoV isolates.

Coronaviruses are important causes of human and animal diseases that include respiratory infection, gastroenteritis, hepatic and neurological disorders as well as immune-mediated disease such as feline infectious peritonitis, and other persistent infections (reviewed in Spaan *et al.*, 1988; Wege *et al.*, 1982). We investigated two different epizootics of acute respiratory tract disease and bovine 'shipping fever pneumonia', each affecting more than 100 animals. We found that respiratory bovine coronaviruses (RBCoVs) were the most frequent viruses isolated from nasal swabs and lung samples of affected animals (Storz *et al.*, 2000*a*, *b*). Unlike the prototypic BCoV strain L9 and the enteric wild-type strain (EBCoV) LY138, these RBCoV strains exhibited high cell-fusing activities for the G clone of human rectal tumour (HRT-18G) cells and a restricted haemagglutinating pattern (Storz *et al.*, 1996, 2000*b*). A few animals also shed coronaviruses that were phenotypically similar to the wild-type EBCoV isolated from intestinal tracts. One of these animals with fatal pneumonia served as a source for the RBCoV strain 98TXSF-110-LUN (LUN) and the enteric strain 98TXSF-110-ENT (ENT), respectively isolated from lung and intestinal samples.

Previously, we reported nucleotide sequences of approximately 9·5 kbp from the 3′ region downstream of the replicase genes specified by respiratory and enteric BCoV isolates from different animals (Chouljenko *et al.*, 1998). We report here a comparative analysis of the newly determined complete genomes of the LUN and ENT strains of BCoV. Besides being the first report of complete genome sequences of BCoV, this analysis also provides insight into replicase strain-specific differences, which are associated with the evolution of BCoV in the natural host.

BCoV strains were propagated in HRT-18G cells (Storz *et al.*, 1996). After the second passage, viruses were further plaque-purified three times and genomic RNA was extracted from gradient-purified virus to produce cDNAs, as described previously (Chouljenko *et al.*, 1998). In order to ensure that sequences were derived from the predominant virus species, only direct RT–PCR products were used as templates for rTth DNA polymerase-XL enzyme (PE Applied Biosystems). To minimize the potential damage to nucleotides caused by UV irradiation and ethidium bromide, all templates were purified by the crystal violet method (Invitrogen) and DNA was recovered from the gel by using the Zymoclean gel DNA recovery kit (Zymo Research). PCRs were performed by using

Author for correspondence: Konstantin Kousoulas.
Fax +1 225 578 9701. e-mail vtgusk@lsu.edu

The GenBank accession numbers of sequences reported in this paper are AF391541 and AF391542.

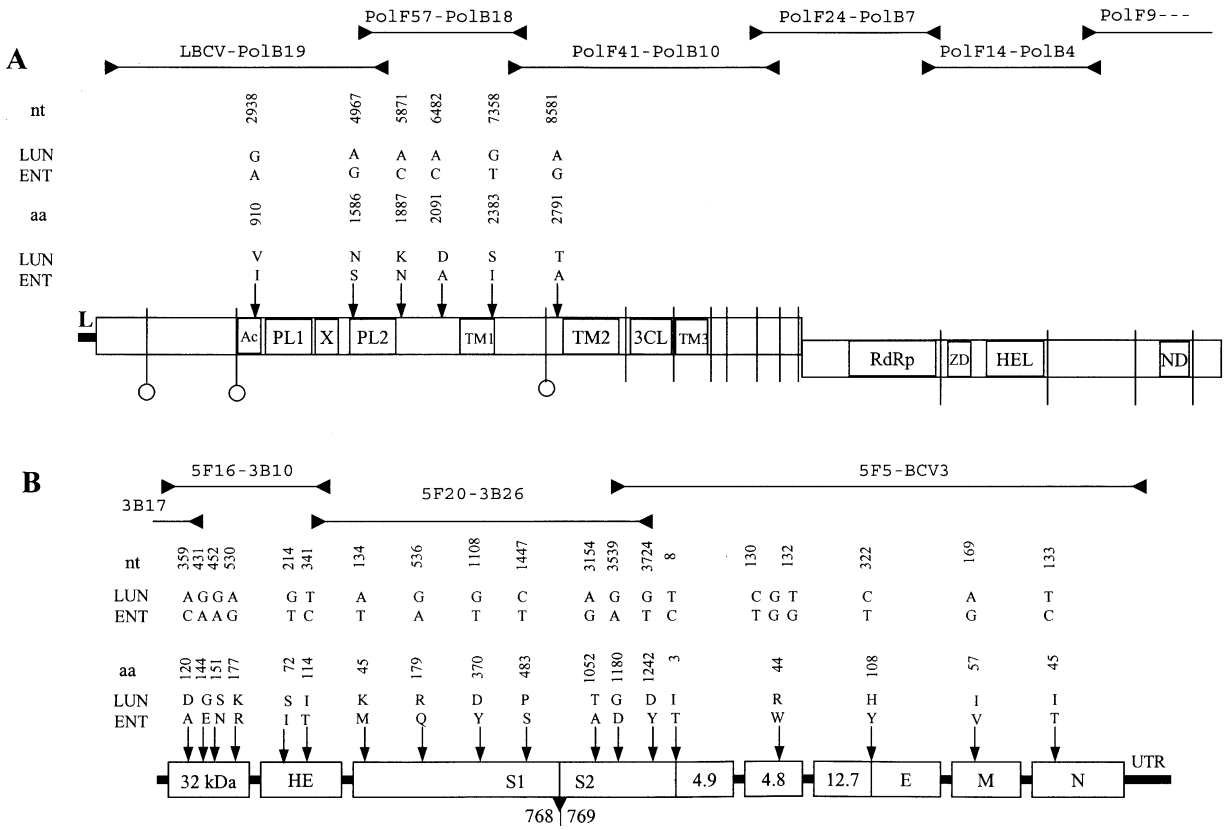

Fig. 1. Sequencing strategy and amino acid differences between the LUN and ENT isolates. The primer for the 5′ end of the BCoV genome (18 nt; primer LBCV) was synthesized on the basis of available sequence (Hoffman *et al.*, 1993). The 3′-end primer sequence (BCV3) was as reported previously (Chouljenko *et al.*, 1998). The exact location and actual primer sequences used to generate the overlapping RT–PCR products that produced the complete, 31 028 nt genomic sequences specified by LUN and ENT isolates are available in Supplementary Table 1 at JGV Online (http://vir.sgmjournals.org/). The positions of non-synonymous nucleotide changes in the genomes of BCoV isolates LUN and ENT and the amino acids specified are shown for the replicase polyproteins (A) and other proteins located downstream (B). Residues in each protein are numbered independently. (A) The replicase pp1ab polyprotein encoded by overlapping ORF1a (nt 211–13362) and ORF1b (nt 13332–21494). The approximate positions of predicted functional domains and protease cleavage sites in pp1ab are shown, as listed in Table 1: 3CL, main proteinase domain; TM1, TM2 and TM3, membrane-spanning domains; PL1 and PL2, accessory proteinase domains; X, conserved domain of unknown function; RdRp, RNA-dependent RNA polymerase; ZD, putative zinc finger; HEL, NTPase RNA helicase; ND, domain conserved exclusively in nidoviruses. Sites predicted to be cleaved by 3CL protease and PL1/PL2 are marked by vertical lines and vertical lines with circles, respectively. L represents the leader sequence at the 5′ end of the genome. (B) Rectangles depict the putative 32 kDa protein, HE, the S1 and S2 subunits of S, N_s 4·9 kDa (4·9), N_s 4·8 kDa (4·8), N_s 12·7 kDa (12·7), E, M and N. The 3′-UTR of BCoV is also shown. The location of the cleavage site within S that separates the S1 and S2 subunits is indicated. All nucleotide differences between isolates LUN and ENT are listed in Supplementary Table 2 at JGV Online (http://vir.sgmjournals.org/).

the GeneAmp XL PCR kit, AmpliWax PCR Gem 100 (providing hot-start PCR) and the GeneAmp PCR System 9600 instrument (all from PE Applied Biosystems). Sequencing was performed on an ABI Prism 377 DNA sequencer (PE Applied Biosystems) according to the strategy outlined in Fig. 1. Oligonucleotide primers were based on either the replicase sequence of the mouse hepatitis virus (MHV) A59 c12 mutant (GenBank accession no. AF029248) or an available BCoV nucleotide sequence (AF058944). After the respective BCoV sequences were elucidated, BCoV-based primers (Fig. 1) were used and overlapping cDNA fragments of 2–6 kbp were re-amplified and re-sequenced. Sequences of the cDNA fragments were aligned to assemble the entire BCoV genome of LUN and ENT strains using the Sequencher 3.1.1 sequence analysis software (Gene Code Corporation).

The genomes of both BCoV isolates contained 31028 nt and included 13 open reading frames (ORFs) flanked by 5′- (nt 1–210) and 3′- (nt 30740–31028) untranslated regions (UTRs). Some of these ORFs overlapped [ORF1a (nt 211–13362, including stop codon) and ORF1b (13332–21494); ORF4 (S, 23641–27732) and ORF5 (N_s 4·9 kDa, 27722–27811); ORF7 (N_s 12·7 kDa, 28106–28435) and ORF8 (E, 28422–28676); ORF10 (N, 29393–30739) and ORF11 (I, 29454–30077)] whereas others were separated by intergenic sequences

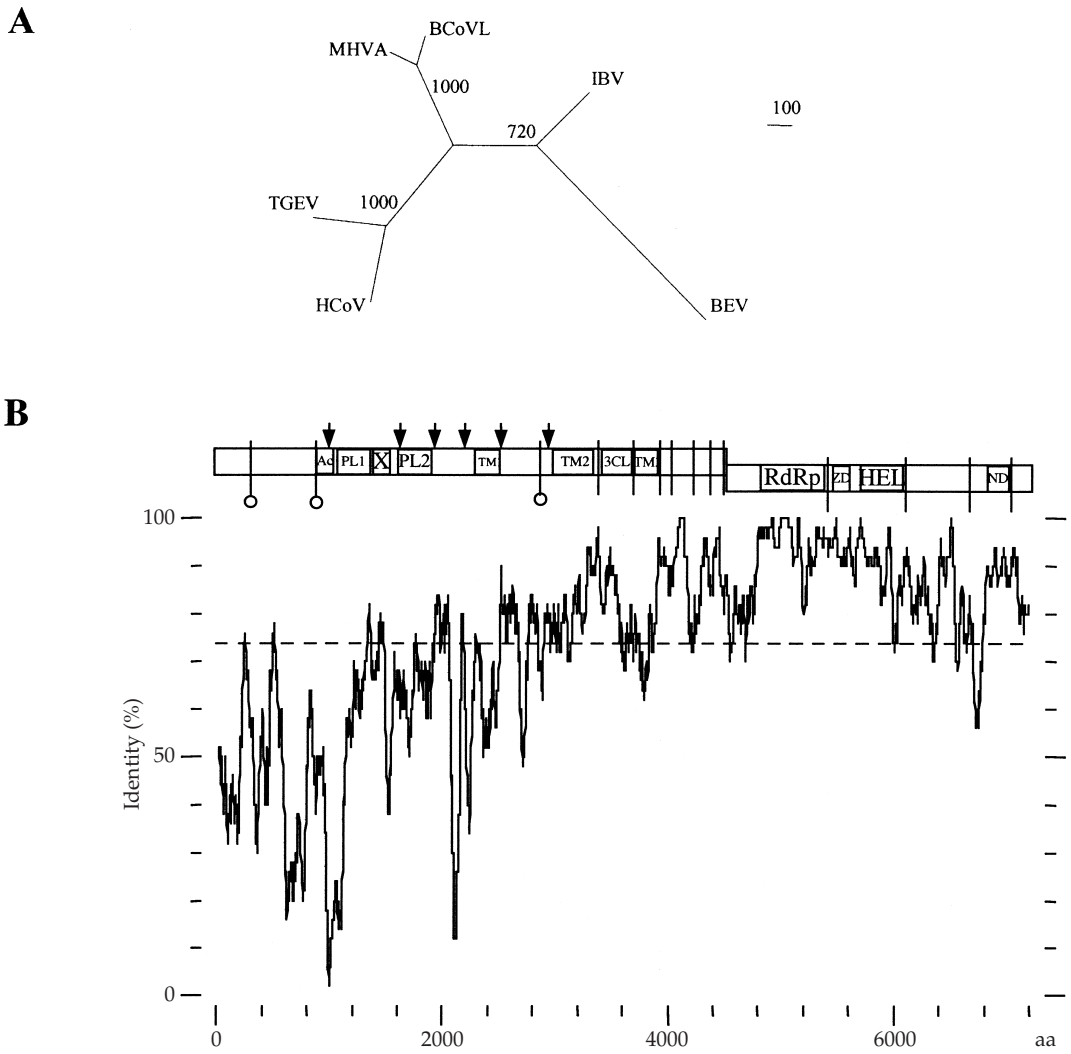

Fig. 2. Comparison of replicases of BCoV and other coronaviruses. A tree (A) and a similarity plot (B) were generated using a multiple alignment between replicases of the LUN and ENT isolates of BCoV and other coronaviruses, including MHV (two strains) (Bonilla *et al.*, 1994; Bredenbeek *et al.*, 1990; Lee *et al.*, 1991), transmissible gastroenteritis virus (TGEV) (Eleouet *et al.*, 1995), human coronavirus (HCoV) (Herold *et al.*, 1993), infectious bronchitis virus (IBV) (Boursnell *et al.*, 1987) and Berne torovirus (BEV) (Snijder *et al.*, 1991). For accession numbers of these sequences and alignments of selected replicase domains, see Ziebuhr *et al.* (2000, 2001). (A) Unrooted tree for the pp1b portion of coronavirus replicases. Because only the pp1b portion is available for BEV, this part of the alignment was used to assess the phylogenetic relationship between replicases of BCoV and other coronaviruses. The number of trees in which a particular bifurcation was sustained in the course of 1000 bootstrap simulations is given at each node. Bar, 100 amino acid replacements. (B) Similarity plot between pp1ab of BCoV and MHV A. The plot was produced by recording percentage sequence identity within a 50 aa window, which was moved along the alignment between MHV A and BCoV, using the Plotsimilarity routine of the GCG Wisconsin package. The broken line depicts the mean similarity between the two polyproteins. The domain organization of pp1ab conserved in BCoV and positions of amino acid differences between LUN and ENT strains (arrows) are shown on the top of the plot for comparison (for details of strain-specific nucleotides and amino acids, see Fig. 1 A and Supplementary Table 2, available at http://vir.sgmjournals.org/).

[ORF1b and ORF2 (N_s 32 kDa, 21504–22340); ORF2 and ORF3 (HE, 22352–23626); ORF3 and ORF4; ORF5 and ORF6 (N_s 4·8 kDa, 27889–28026); ORF6 and ORF7; ORF8 and ORF9 (M, 28691–29383) and ORF9 and ORF10]. The complete ORF1a and ORF1b sequences of BCoV contained 13 152 and 8163 nt, respectively, and are reported for the first time. These ORFs overlap by 31 nt in a region including the 'slippery' sequence UUUAAAC that is believed to direct a −1 frameshift and is located immediately upstream of a predicted stem–loop stabilized through a complex RNA pseudoknot (not shown) (reviewed in Brierley, 1995).

A multiple sequence alignment of replicase polyproteins pp1ab of BCoV and other coronaviruses (not shown) was produced using the programs CLUSTAL X (Thompson *et al.*, 1997), Dialign2 (Morgenstern, 1999) and Macaw (Schuler *et al.*, 1991) and used for phylogenetic analysis of pp1b. By

Table 1. Predicted end-products of proteolytic processing of the BCoV replicase polyproteins pp1a and pp1ab mediated by putative PL1pro, PL2pro and 3CL^{pro}

Proteins are numbered in their order in polyproteins pp1a/pp1ab. Domains not defined elsewhere are: Ac, domain enriched in acidic Asp/Glu residues; PL1pro, papain-like proteinase 1; PL2pro, papain-like proteinase 2; 3CL^{pro}, 3C-like proteinase; GFL, Cys-rich domain originally thought to be growth factor-like.

Protein	Position in pp1a/pp1ab (aa)*	Length (aa)	Length of MHV orthologue (aa) (percentage identity)	Associated putative functional domain(s)†	Putative protease(s) predicted to release protein from polyproteins
1	Met^{1}–Gly^{246}	246	247 (47)	?	PL1pro
2	Val^{247}–Ala^{851}	605	585 (48)	?	PL2pro/PL1pro
3	Gly^{852}–Gly^{2750}	1899	2008 (63)	Ac, PL1pro, X, PL2pro, TM1	PL2pro/PL1pro
4	Ala^{2751}–Gln^{3246}	496	493 (79)	TM2	PL2pro + 3CL^{pro}
5	Ser^{3247}–Gln^{3549}	303	302 (84)	3CL^{pro}	3CL^{pro}
6	Ser^{3550}–Gln^{3836}	287	286 (76)	TM3	3CL^{pro}
7	Ser^{3837}–Gln^{3925}	89	89 (89)	?	3CL^{pro}
8	Ala^{3926}–Gln^{4122}	197	194 (89)	?	3CL^{pro}
9	Asn^{4123}–Gln^{4232}	110	110 (86)	?	3CL^{pro}
10	Ala^{4233}–Gln^{4369}	137	137 (89)	GFL	3CL^{pro}
11	Ser^{4370}–Val^{4383} of pp1a	14	14 (93)	?	3CL^{pro}
12	Ser^{4370}–Gln^{5297}	928	928 (90)	*RdRp*	3CL^{pro}
13	Ser^{5298}–Gln^{5900}	603	600 (92)	*ZD, NTPase, HEL*	3CL^{pro}
14	Cys^{5901}–Gln^{6421}	521	521 (86)	?	3CL^{pro}
15	Ser^{6422}–Gln^{6795}	374	374 (77)	ND	3CL^{pro}
16	Ala^{6796}–Ile^{7094}	299	299 (86)	?	3CL^{pro}

* Amino acid numbering of the ORF1a and ORF1a/1b gene products, pp1a and pp1ab, is based upon predictions of the coronavirus −1 frameshift site that is believed to occur at the 'slippery sequence', UUUAAAC, to form a peptide bond Asn^{4377}–Arg^{4378}. Protein junctions were tentatively identified through an alignment of replicases of BCoV and other well-characterized coronaviruses (Ziebuhr *et al.*, 2000, 2001).
† Putative domains are in italics. ?, Function unknown.

conducting a parsimonious exhaustive tree search with a UNIX version of the program PAUP* 4.0.0d55 (Swofford, 2000), which is part of the GCG Wisconsin package, we found that the most closely related replicases were those of BCoV and MHV (Fig. 2A). A similar observation was reported recently for the most conserved fragment of pp1b using a larger set of coronaviruses (Stephensen *et al.*, 1999). This replicase clustering parallels that of the structural proteins of MHV and BCoV, which previously led to the placement of these viruses in coronavirus group 2 (Siddell, 1995). The pp1ab polyprotein amino acid sequences of BCoV and MHV share approximately 75% identity, distributed unevenly (Fig. 2B). Identities of 47–93% were found between BCoV and MHV for 16 replicase protein sequences (Table 1). The domain organization and end-products of proteolytic processing of pp1a/pp1ab were predicted from comparisons of BCoV and other, better-characterized coronaviruses (reviewed in Ziebuhr *et al.*, 2000; see also Kanjanahaluethai & Baker, 2000; Ziebuhr *et al.*, 2001). The viral proteinases and the X, Cys-rich (GFL) and predicted (Rost *et al.*, 1995) membrane-spanning domains (TM1, TM2 and TM3) are derived from the pp1a part of the pp1ab polyprotein. The major replication enzymes, including the putative RNA-dependent RNA polymerase (RdRp) and RNA helicase (HEL), as well as Cys/His-rich putative Zn-binding domain (ZD) and nidovirus-specific conserved domain (ND), are derived from pp1b (Table 1; Fig. 1A).

The genomes of the two isolates differed in 107 out of 31028 positions scattered throughout the genome (available as Supplementary Table 2 at JGV Online; http://vir.sgmjournals.org/). Differences in 25 positions were non-synonymous and resulted in 24 amino acid changes, as discussed below. The remaining 82 nucleotide differences did not cause amino acid changes. However, these mutations might modulate phenotypic properties by affecting the RNA structure per se and/or RNA interaction(s) with partners. ORF1b represented the most conserved part of the BCoV genome and contained 16 synonymous differences between the LUN and ENT isolates.

Amino acid differences between the two BCoV isolates are listed in Fig. 1. Within ORF1a, the isolates differed in 32 positions, six of which were non-synonymous and were located in the third (five mutations) and fourth (one) putative proteins (Fig. 1; Table 1). The third protein is the largest (1899 aa) replicase subunit in BCoV and is also known as p195 or p210 (p240) in other coronaviruses (Denison & Perlman, 1987; Lim *et al.*, 2000; Schiller *et al.*, 1998). The replicase

mutations were found within or proximal to the interdomain junctions involving the PL1pro, PL2pro, TM1 and TM2 domains (Fig. 1A). The BCoV and MHV p195/p210 proteins are known to tolerate substantial changes including large deletions and insertions (compare the sizes of proteins in Table 1; see also Fig. 2B). Therefore, the predominant clustering of mutations within p195/p210 may be due to the rapid evolution of this protein. However, it is evident that only two of the five mutations, those at positions 910 and 2091, are located in the least conserved regions ('hot' spots) in p195/p210. Furthermore, the first and second proteins, which have not accepted mutations after the divergence of the LUN and ENT strains, have been shown to be the fastest evolving proteins in coronaviruses and specifically in the BCoV/MHV lineage (compare the percentages of identical residues in the first three proteins in Table 1; see also Fig. 2B; Lee *et al.*, 1991; A. E. Gorbalenya, unpublished). These observations suggest that certain amino acid differences observed between LUN and ENT p195/p210 may be meaningful because they have evolved under selective pressure. Mutation-induced changes in PL1pro/PL2pro functions might affect the expression pattern of the poorly characterized p195/p210 and flanking proteins, which are under the control of these proteases. Also, these mutations could affect expression of subgenomic mRNAs, which may be stimulated by a non-proteolytic activity associated with these proteases (Herold *et al.*, 1999; Tijms *et al.*, 2001). Mutations in the vicinity of transmembrane domains, which are thought to form a scaffold for the RNA-synthesizing machinery (Gorbalenya *et al.*, 1989; Shi *et al.*, 1999; van der Meer *et al.*, 1999), might affect the biogenesis and/or functions of this multisubunit structure.

Downstream of the replicase genes, the LUN and ENT isolates differed in 59 positions, 18 of which were non-synonymous and were scattered over all proteins (Fig. 1B). Comparison of all available BCoV nucleotide sequences revealed that only the amino acid change at position 179 (Arg for LUN and Gln for ENT) of the S protein may be specific for respiratory versus enteric isolates, as has been demonstrated for transmissible gastroenteritis virus (Almazán *et al.*, 2000; Ballesteros *et al.*, 1995, 1997; Krempl *et al.*, 1997; Rasschaert *et al.*, 1990; Sanchez *et al.*, 1999).

Previous work with BCoV and other coronaviruses has suggested that S also plays an important role in virus-induced membrane fusion and plaque size (Chambers *et al.*, 1990; Gallagher *et al.*, 1991; Kubo *et al.*, 1993; Routledge *et al.*, 1991; Sturman *et al.*, 1990; Yoo *et al.*, 1991). Interestingly, when the LUN and ENT strains were assayed on HRT-18G cells at 48 h after infection, the LUN strain caused extensive cell-to-cell fusion and formed two- to threefold larger plaques than did the ENT strain (data not shown). Single amino acid changes within S are prime candidates for causing changes in virus-induced cell fusion and plaque size. However, other strain-specific mutations may also contribute to these phenotypes.

The genome structures of BCoV reported here extend the list of known coronavirus genomes (Boursnell *et al.*, 1987; Eleouet *et al.*, 1995; Bonilla *et al.*, 1994; Bredenbeek *et al.*, 1990; Herold *et al.*, 1993; Lee *et al.*, 1991). A theoretical analysis of amino acid differences between the LUN and ENT isolates indicated that the replicative protein p195/p210 and structural protein S might control strain-specific properties. Current work is focusing on the evaluation of the relative importance of the isolate-specific mutations of BCoV in cell culture and in the natural host through the construction and use of specific LUN/ENT recombinant viruses.

We acknowledge the technical assistance of Mamie Burrell. We are grateful to John Ziebuhr for permission to use unpublished data in this study. USDA grants 98-35204-6585 of the National Research Initiative Program to K.G.K. and 98-34362-6071 from the USDA to J.S. and K.G.K. supported this work. We are indebted to the LSU School of Veterinary Medicine for support. This publication is identified as GeneLab publication #GL1203. A.E.G. was supported with funds from the National Cancer Institute, National Institutes of Health, under contract no. NO1-CO-56000. The content of this publication does not necessarily reflect the views or policies of the US Department of Health and Human Services.

References

Almazán, F., González, J. M., Pénzes, Z., Izeta, A., Calvo, E., Plana-Durán, J. & Enjuanes, L. (2000). Engineering the largest RNA virus genome as an infectious bacterial artificial chromosome. *Proceedings of the National Academy of Sciences, USA* **97**, 5516–5521.

Ballesteros, M. L., Sanchez, C. M., Martin-Caballero, J. & Enjuanes, L. (1995). Molecular bases of tropism in the PUR46 cluster of transmissible gastroenteritis coronaviruses. *Advances in Experimental Medicine and Biology* **380**, 557–562.

Ballesteros, M. L., Sanchez, C. M. & Enjuanes, L. (1997). Two amino acid changes at the N-terminus of transmissible gastroenteritis coronavirus spike protein result in the loss of enteric tropism. *Virology* **227**, 378–388.

Bonilla, P. J., Gorbalenya, A. E. & Weiss, S. R. (1994). Mouse hepatitis virus strain A59 RNA polymerase gene ORF 1a: heterogeneity among MHV strains. *Virology* **198**, 736–740.

Boursnell, M. E. G., Brown, T. D. K., Foulds, I. J., Green, P. F., Tomley, F. M. & Binns, M. M. (1987). Completion of the sequence of the genome of the coronavirus avian infectious bronchitis virus. *Journal of General Virology* **68**, 57–77.

Bredenbeek, P. J., Pachuk, C. J., Noten, A. F., Charite, J., Luytjes, W., Weiss, S. R. & Spaan, W. J. (1990). The primary structure and expression of the second open reading frame of the polymerase gene of the coronavirus MHV-A59; a highly conserved polymerase is expressed by an efficient ribosomal frameshifting mechanism. *Nucleic Acids Research* **18**, 1825–1832.

Brierley, I. (1995). Ribosomal frameshifting on viral RNAs. *Journal of General Virology* **76**, 1885–1892.

Chambers, P., Pringle, C. R. & Easton, A. J. (1990). Heptad repeat sequences are located adjacent to hydrophobic regions in several types of virus fusion glycoproteins. *Journal of General Virology* **71**, 3075–3080.

Chouljenko, V. N., Kousoulas, K. G., Lin, X. & Storz, J. (1998). Nucleotide and predicted amino acid sequences of all genes encoded by

the 3′ genomic portion (9·5 kb) of respiratory bovine coronaviruses and comparisons among respiratory and enteric coronaviruses. *Virus Genes* **17**, 33–42.

Denison, M. & Perlman, S. (1987). Identification of putative polymerase gene product in cells infected with murine coronavirus A59. *Virology* **157**, 565–568.

Eleouet, J. F., Rasschaert, D., Lambert, P., Levy, L., Vende, P. & Laude, H. (1995). Complete sequence (20 kilobases) of the polyprotein-encoding gene 1 of transmissible gastroenteritis virus. *Virology* **206**, 817–822.

Gallagher, T. M., Escarmis, C. & Buchmeier, M. J. (1991). Alteration of the pH dependence of coronavirus-induced cell fusion: effect of mutations in the spike glycoprotein. *Journal of Virology* **65**, 1916–1928.

Gorbalenya, A. E., Koonin, E. V., Donchenko, A. P. & Blinov, V. M. (1989). Coronavirus genome: prediction of putative functional domains in the non-structural polyprotein by comparative amino acid sequence analysis. *Nucleic Acids Research* **17**, 4847–4861.

Herold, J., Raabe, T., Schelle-Prinz, B. & Siddell, S. G. (1993). Nucleotide sequence of the human coronavirus 229E RNA polymerase locus. *Virology* **195**, 680–691.

Herold, J., Siddell, S. G. & Gorbalenya, A. E. (1999). A human RNA viral cysteine proteinase that depends upon a unique Zn^{2+}-binding finger connecting the two domains of a papain-like fold. *Journal of Biological Chemistry* **274**, 14918–14925.

Hofmann, M. A., Chang, R. Y., Ku, S. & Brian, D. A. (1993). Leader–mRNA junction sequences are unique for each subgenomic mRNA species in the bovine coronavirus and remain so throughout persistent infection. *Virology* **196**, 163–171.

Kanjanahaluethai, A. & Baker, S. C. (2000). Identification of mouse hepatitis virus papain-like proteinase 2 activity. *Journal of Virology* **74**, 7911–7921.

Krempl, C., Schultze, B., Laude, H. & Herrler, G. (1997). Point mutations in the S protein connect the sialic acid binding activity with the enteropathogenicity of transmissible gastroenteritis coronavirus. *Journal of Virology* **71**, 3285–3287.

Kubo, H., Takase-Yoden, S. & Taguchi, F. (1993). Neutralization and fusion inhibition activities of monoclonal antibodies specific for the S1 subunit of the spike protein of neurovirulent murine coronavirus JHMV c1-2 variant. *Journal of General Virology* **74**, 1421–1425.

Lee, H. J., Shieh, C. K., Gorbalenya, A. E., Koonin, E. V., La Monica, N., Tuler, J., Bagdzhadzhyan, A. & Lai, M. M. C. (1991). The complete sequence (22 kilobases) of murine coronavirus gene 1 encoding the putative proteases and RNA polymerase. *Virology* **180**, 567–582.

Lim, K. P., Ng, L. F. & Liu, D. X. (2000). Identification of a novel cleavage activity of the first papain-like proteinase domain encoded by open reading frame 1a of the coronavirus avian infectious bronchitis virus and characterization of the cleavage products. *Journal of Virology* **74**, 1674–1685.

Morgenstern, B. (1999). DIALIGN 2: improvement of the segment-to-segment approach to multiple sequence alignment. *Bioinformatics* **15**, 211–218.

Rasschaert, D., Duarte, M. & Laude, H. (1990). Porcine respiratory coronavirus differs from transmissible gastroenteritis virus by a few genomic deletions. *Journal of General Virology* **71**, 2599–2607.

Rost, B., Casadio, R., Fariselli, P. & Sander, C. (1995). Transmembrane helices predicted at 95% accuracy. *Protein Science* **4**, 521–533.

Routledge, E., Stauber, R., Pfleiderer, M. & Siddell, S. G. (1991). Analysis of murine coronavirus surface glycoprotein functions by using monoclonal antibodies. *Journal of Virology* **65**, 254–262.

Sanchez, C. M., Izeta, A., Sanchez-Morgado, J. M., Alonso, S., Sola, I., Balasch, M., Plana-Duran, J. & Enjuanes, L. (1999). Targeted recombination demonstrates that the spike gene of transmissible gastroenteritis coronavirus is a determinant of its enteric tropism and virulence. *Journal of Virology* **73**, 7607–7618.

Schiller, J. J., Kanjanahaluethai, A. & Baker, S. C. (1998). Processing of the coronavirus MHV-JHM polymerase polyprotein: identification of precursors and proteolytic products spanning 400 kilodaltons of ORF1a. *Virology* **242**, 288–302.

Schuler, G. D., Altschul, S. F. & Lipman, D. J. (1991). A workbench for multiple alignment construction and analysis. *Proteins* **9**, 180–190.

Shi, S. T., Schiller, J. J., Kanjanahaluethai, A., Baker, S. C., Oh, J. W. & Lai, M. M. (1999). Colocalization and membrane association of murine hepatitis virus gene 1 products and de novo-synthesized viral RNA in infected cells. *Journal of Virology* **73**, 5957–5969.

Siddell, S. (1995). The *Coronaviridae*: an introduction. In *The Coronaviridae*, pp. 1–10. Edited by S. G. Siddell. New York: Plenum Press.

Snijder, E. J., den Boon, J. A., Horzinek, M. C. & Spaan, W. J. (1991). Comparison of the genome organization of toro- and coronaviruses: evidence for two nonhomologous RNA recombination events during Berne virus evolution. *Virology* **180**, 448–452.

Spaan, W., Cavanagh, D. & Horzinek, M. C. (1988). Coronaviruses: structure and genome expression. *Journal of General Virology* **69**, 2939–2952.

Stephensen, C. B., Casebolt, D. B. & Gangopadhyay, N. N. (1999). Phylogenetic analysis of a highly conserved region of the polymerase gene from 11 coronaviruses and development of a consensus polymerase chain reaction assay. *Virus Research* **60**, 181–189.

Storz, J., Stine, L., Liem, A. & Anderson, G. A. (1996). Coronavirus isolation from nasal swab samples in cattle with signs of respiratory tract disease after shipping. *Journal of the American Veterinary Medical Association* **208**, 1452–1455.

Storz, J., Lin, X., Purdy, C. W., Chouljenko, V. N., Kousoulas, K. G., Enright, F. M., Gilmore, W. C., Briggs, R. E. & Loan, R. W. (2000*a*). Coronavirus and *Pasteurella* infections in bovine shipping fever pneumonia and Evans′ criteria for causation. *Journal of Clinical Microbiology* **38**, 3291–3298.

Storz, J., Purdy, C. W., Lin, X., Burrell, M., Truax, R. E., Briggs, R. E., Frank, G. H. & Loan, R. W. (2000*b*). Isolation of respiratory bovine coronavirus, other cytocidal viruses, and *Pasteurella* spp. from cattle involved in two natural outbreaks of shipping fever. *Journal of the American Veterinary Medical Association* **216**, 1599–1604.

Sturman, L. S., Ricard, C. S. & Holmes, K. V. (1990). Conformational change of the coronavirus peplomer glycoprotein at pH 8·0 and 37 °C correlates with virus aggregation and virus-induced cell fusion. *Journal of Virology* **64**, 3042–3050.

Swofford, D. L. (2000). PAUP*: phylogenetic analysis using parsimony (*and other methods), version 4. Sunderland, MA: Sinauer Associates.

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

Tijms, M. A., van Dinten, L. C., Gorbalenya, A. E. & Snijder, E. J. (2001). A zinc finger-containing papain-like protease couples subgenomic mRNA synthesis to genome translation in a positive-stranded RNA virus. *Proceedings of the National Academy of Sciences, USA* **98**, 1889–1894.

van der Meer, Y., Snijder, E. J., Dobbe, J. C., Schleich, S., Denison, M. R., Spaan, W. J. & Locker, J. K. (1999). Localization of mouse

hepatitis virus nonstructural proteins and RNA synthesis indicates a role for late endosomes in viral replication. *Journal of Virology* **73**, 7641–7657.

Wege, H., Siddell, S. & ter Meulen, V. (1982). The biology and pathogenesis of coronaviruses. *Current Topics in Microbiology and Immunology* **99**, 165–200.

Yoo, D. W., Parker, M. D. & Babiuk, L. A. (1991). The S2 subunit of the spike glycoprotein of bovine coronavirus mediates membrane fusion in insect cells. *Virology* **180**, 395–399.

Ziebuhr, J., Snijder, E. J. & Gorbalenya, A. E. (2000). Virus-encoded proteinases and proteolytic processing in the *Nidovirales*. *Journal of General Virology* **81**, 853–879.

Ziebuhr, J., Thiel, V. & Gorbalenya, A. E. (2001). The autocatalytic release of a putative RNA virus transcription factor from its polyprotein precursor involves two paralogous papain-like proteases that cleave the same peptide bond. *Journal of Biological Chemistry* **276**, 33220–33232.

Received 11 May 2001; Accepted 30 August 2001