



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.

# Complete sequences of 3' end coding region for structural protein genes of turkey coronavirus

Tsang Long Lin\*, Chien Chang Loa, Ching Ching Wu

*Department of Veterinary Pathobiology, Purdue University, ADDL, 406 South University Street, West Lafayette, Indiana 47907-2065, USA*

Received 6 October 2003; received in revised form 27 May 2004; accepted 12 June 2004

Available online 3 August 2004

## Abstract

Overlapping fragments of genomic RNA spanning 6963 nucleotides from 5' end of spike (S) protein gene to 3' end of nucleocapsid (N) protein gene of turkey coronavirus (TCoV) were amplified by reverse-transcription-polymerase chain reaction (RT-PCR). The primers were derived from the corresponding sequences of infectious bronchitis virus (IBV). The PCR products were cloned and sequenced and their nucleic acid structure and similarity to published sequences of other coronaviruses were analyzed. Sequencing and subsequent analysis revealed 9 open reading frames (ORFs) representing the entire S protein gene, tricistronic gene 3, membrane (M) protein gene, bicistronic gene 5, and N protein gene in the order of 5'–3'. The overall nucleic acid structures of these encoding regions of TCoV were very similar to the homologous regions of IBV. The consensus transcription-regulating sequence (TRS) of IBV, CT(T/G)AACAA, was highly conserved in TCoV genome at the levels of nucleotide sequence and location in regarding to the initiation codon of individual genes. Pair-wise comparison of gene 3, M gene, gene 5, or N gene sequences with their counterparts of IBV revealed high levels (82.1–92.0%) of similarity. Phylogenetic analysis based on the deduced amino acid sequences of S, M, or N protein demonstrated that TCoV was clustered within the same genomic lineage as the IBV strains while all the other mammalian coronaviruses were grouped into separate clusters corresponding to antigenic groups I or II. There were substantial differences of S protein sequence between TCoV and IBV with only 33.8–33.9% of similarity.

© 2004 Elsevier B.V. All rights reserved.

**Keywords:** Turkey coronavirus; Infectious bronchitis virus; Coronavirus; Genomic relationship

## 1. Introduction

Turkey coronavirus (TCoV) was identified in the early 1970s as the major causative agent of the most costly disease of turkey encountered in Minnesota between 1951 and 1971 (Nagaraja and Pomeroy, 1997). Outbreaks of turkey poult enteritis associated with TCoV have caused severe economical losses in the turkey industry in Indiana, North Carolina, and other states for the last several years. Although the economical importance of this disease has been recognized for decades, the organization of genomic structure of TCoV is poorly understood and reports regarding the relationships of TCoV with other coronaviruses remained controversial (Van Regenmortel et al., 2000; Gonzalez et al., 2003).

Coronaviruses are pleomorphic, enveloped spherical particles surrounded by a fringe of 20 nm long club-shaped spikes. The diameter of coronaviral particles are around 140–150 nm. The coronavirus genome is a positive single-stranded capped RNA with a polyadenylated 3' end. Complete genomic RNA sequences of coronaviruses has been determined for infectious bronchitis virus (IBV; 27,569 nucleotides; Bournsnel et al., 1987), murine hepatitis virus (MHV; 31,092 nucleotides; Lee et al., 1991), human coronavirus (HCoV) strain 229E (27,277 nucleotides; Herold et al., 1993), and transmissible gastroenteritis virus (TGEV; 28,579 nucleotides; Eleouet et al., 1995; Penzes et al., 2001). The 5' two-thirds of the coronavirus genome, approximately 20 kb, consists of two overlapping open reading frames (ORFs) that encode non-structural proteins including the viral RNA-dependent RNA polymerase and proteases. Another one-third nucleotide sequences from 3' end contain ORFs encode the major structural proteins:

\* Corresponding author. Fax: +1 765 494 9181.  
E-mail address: [tlin@purdue.edu](mailto:tlin@purdue.edu) (T.L. Lin).

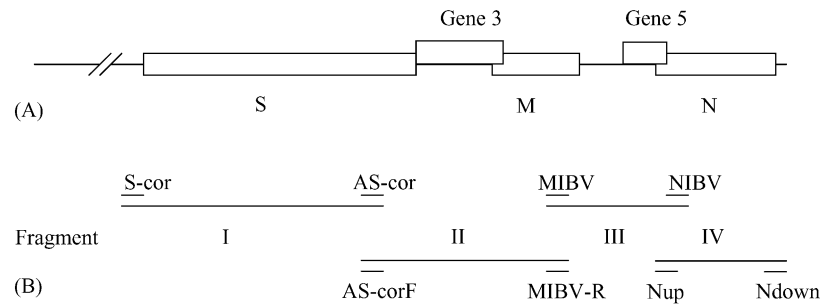


Fig. 1. (A) Schematic representation of turkey coronavirus genomic RNA, showing locations of coding regions for the spike (S) protein, gene 3, membrane (M) protein, gene 5, and nucleocapsid (N) protein. (B) Schematic representation of the location of primers used in the polymerase chain reaction (PCR), along with the intervening sequences (fragments I–III, or IV) amplified by the PCR. The sequences of primers: S-cor, tgaaaactgaacaaaagacagact; AS-cor, ccaaacatccaaggc-cactt; AS-corF, aagtgccctgtatgtttgg; MIBVR, gtcacacttagcaagcactg; MIBV, taagcttcagtggcttctaagtgtgaacc; NIBV, tggatcaccgctactctcaaaacttgggcgc; Nup, tcctttgccatggcaag c; Ndown, tactcaaaagtcatctc.

spike (S), membrane (M), and nucleocapsid (N) proteins in the order of 5′–3′ along the genome, respectively.

Turkey coronavirus was initially determined to be antigenically distinct from all other coronaviruses based on antigenic differences revealed by immunoelectron microscopy (Ritchie et al., 1973) and hemagglutination-inhibition (Dea et al., 1986). This unique antigenicity was questioned when the close relationship between TCoV and bovine coronavirus (BCoV) was demonstrated in a series of antigenic studies (Dea et al., 1990) and by sequence analysis of TCoV M and N genes (Verbeek and Tijssen, 1991). In contrast, recent antigenic (Guy et al., 1997; Loa et al., 2000) and genomic (Breslin et al., 1999a,b; Akin et al., 2001; Cavanagh et al., 2001, 2002; Lin et al., 2002) analysis of TCoV, however, demonstrated that TCoV and IBV, two avian coronaviruses, are closely related. The causes for these discrepant results regarding the relationships of TCoV with BCoV or IBV remained unclear. Further analysis of genomic structure of TCoV is important to clarify this enigma. Thus, the purpose of the present study was to determine the sequences of the 3′ end coding region for structural protein genes of TCoV.

## 2. Materials and methods

### 2.1. Turkey coronavirus

The TCoV isolate (isolate 540) used in the present study were recovered from fecal contents and intestines of turkey poult with acute coronaviral enteritis in Indiana, US in 1994. The viruses were passaged 5 times in 22-day-old embryonating turkey eggs. The presence of TCoV in the intestines of embryos were confirmed by TCoV-specific immunofluorescence antibody assays and electron microscopy at the Indiana State Animal Disease Diagnostic Laboratory in West Lafayette, Indiana, US.

### 2.2. RNA isolation and reverse transcription

Total RNA was extracted from the intestines and intestinal content of turkey embryo infected with TCoV by a modified

method using guanidinium thiocyanate and acid-phenol (Chomczynski and Sacchi, 1987; Akin et al., 1999). Conversion of total RNA to cDNA was essentially performed according to a protocol supplied by the manufacturer of the reverse transcriptase (Superscript II system, Life Technologies, Gaithersburg, MD).

### 2.3. PCR amplification

Three microliters of cDNA were used in PCR amplifications with the primers designed from IBV genomic sequences. The locations and sequences of primers for the amplification of 4 fragments I–IV for 3′ end coding region of TCoV structural protein genes are outlined in Fig. 1. PCR was performed with a mixture (64:1, v:v) of Taq (Promega Corp., Madison, WI) and Pfu polymerases (Stratagene, La Jolla, CA) in a 96-well thermal cycler (GeneAmp, Perkin–Elmer Cetus Corp., Norwalk, CT) (Barnes, 1994; Akin et al., 1999). The cyclic parameters of the PCR was as follows: 94 °C for 1 min for denaturation, 37 °C for 2 min for annealing, and 72 °C for 5 min for extension for 40 cycles followed by 72 °C for 10 min for final extension.

### 2.4. Molecular cloning and sequencing

One microliter of the amplification product was used to ligate with pCR-II plasmid vector according to the manufacturer's instructions (Invitrogen, San Diego, CA). Determination of the nucleotide sequences of the selected clone with amplified sequences was performed by dideoxycycle sequencing method with the corresponding sequencing primers for both strands (DAVIS Sequencing, Davis, CA).

### 2.5. Sequence analysis

The nucleotide and deduced amino acid sequences between the TCoV and other coronaviruses were analyzed by DNASTar program (Lasergene Corp, Madison, WI), respectively. Percent similarities were calculated to find nucleic

	S gene	S	
TCoV	TGAAAA <u>TTGAACAA</u> AAGACAGACTTAGTTTTTGGTTTAGTTAAGAGTGGCAAGTTATTAGTTAGAGATGTTGGCGCCGTT		80
Beau	..... C. . . . . AA. . . . . A. . . . . T. . . . . T. . . . . C. G. . . . . A. . . . . <u>ATGTTGGCGCCGTT</u>		80
TCoV	GTGCTGCTTGACGATAATCTTGTCACTTGCATGTAATG- TGAGTTCTATTCCCACAATTTGTAACAATGCTACTGTAGGT		159
Beau	TCTT. TAC. AGT. . CTC. T. . . . . GTGCACT. . . . . G. . . . . C. . . . . TT. . . . . G. . . . . GA. AGT. G. . . . . CT. . . . . CGTT. A. . . . . ACCAA-		159
TCoV	AGTGGTGAGTGTGTACCTTGCA- CAGACAT- - ACCAGAGTATGCACATGGTAAAAATGGGGGTGAAGCTCTAGATTTTTTA		236
Beau	. . . . . CCTTCA. ACC. . . . . A. TGGTT. G. . . . . TT. . . . . A. . . . . G. GG. . . . . TT. . . . . CGGT. G. TAACA. TT- . . . . . AGC. . . . . A. . . . . A.		236
TCoV	TAGTCCAGATGTT- - ATGAGACCGC- CAGATGGTGCATATATACAATCCGGTTATTATGAACCACCTTTTCACAGGTTGCT		313
Beau	. . . . . A. G. . . . . GCTC. TC. . . . . C. . . . . GGT. TA. T. T. . . . . AT- . . . . . TC. TGG. GGTCG. G. . . . . G. T. . . . . T- G. . . . . C. TCT. - . . . . A. . . . .		312
TCoV	TTAATCAAACCTAATCAAACCTGATAGCACCTGTGCAAAATGGACTTTATAGAGGGTCCGACAGGTAATTTTACTATACAAGGT		393
Beau	A. G. CGGC. . CG. . . . . T- . . . . . A. G. . . . . TGG. T. . . . . GT. T. GCA. T. AG. T. T. TACTG. A. ACT. . . . . - TC. G. T. CTAC		389
TCoV	GATTTTCTTGAAAAATTATGATGCTATTGGCATAATGTTT- TGGTGGGGTCTAGGTAATAAGTGGGTGTCATCACGATCCCA		472
Beau	AG. G. . . . . TG. T. . . . . - C. CAT. . . . . T. . . . . AAA. . . . . GG. . . . . GG. G. CC. TTAAGTGGCA. G. . . . . TC. ACA. AATC. T. TA- . . . . . GTG		463
TCoV	CTTTTAATTTAACATGGGGCAACTTTTTCTTAATTTCTAAGAATTTACAAGTTTCCCTAAAGTTAAAAGTGTTATATTC		552
Beau	T. . . . . C. GC. A. G. A. AAT. . . . . C. GC. . . . . CTA. . . . . . . . . . . C. G. . . . . - . . . . . G. AG. . . . . GTACCCT. C. T. . . . . G. . . . .		536
TCoV	ATTGCCAC- TGGAGATATTTTCGTAATGGCGTTTTAATGGGTGTTTATAATCTAAATTTTACGCCAACCTTAACAATTT		631
Beau	. . . . . T. AGTG. . . . . TTA. . . . . A. . . . . AAC. TCC. TATA. . . . . AT. . . . . A. CT. GT. TAC. CC. C. . . . . ATG. G. . . . . A. . . . . GATG. . . . . A		616
TCoV	GGTTAGCACAGTGTGTTGGCACAATGAAGGTTGTTATTTTACGTAATAGTAATGCTCTAGTTCAGTTTTCAGCTGGCAAC		711
Beau	CA. C- - TG. . . . . - . . . . . T. TTT. A. . . . . C. G. . . . . GGACC. . . . . TA- . . . . . C. TA. . . . . A. T. A. GAGAG. AG. . . . . AA. . . . . CC- TGG.		689
TCoV	GTAGTTGATTTTGAACCCTGTACAGGAGACACTACTAT- TAATAAGTTACGTTGTG- CTTATCAGCAATTTAATTTTTTCT		789
Beau	T. . . . . T. . . . . TG. . . . . AATGGTA. . . . . C. . . . . A. . . . . TGTTA. TT. G. G. G. . . . . GGA. C. . . . . C. A. A. G. . . . . G. T. . . . . TGCC. G. A. AATA		769
TCoV	ACGGGATTTTATGACATAGAT- - ACTTTTGTACCTGTGACA- TCTAAT- - ATTACATATTTACC- - TTACCCAGATTTA		861
Beau	CT. . . . . C. A. . . . . T. C. G. . . . . G. C. TTT. TCC. . . . . T. . . . . TAA. AGT. G. T. . . . . G. TA. GC. G. AG. . . . . TTGTC. . . . . T. GT. . . . . AAAT		849
TCoV	AAAGATAAATACTGGTCAAGAGGTATATGATTTTTATGTAGCTCTTAAAGGAGATTCTGTTAGTTATAATCAAAGTTGTGT		941
Beau	. . . . . GT. T. . . . . . . . . . AC. TGTAC. T. . . . . C. CA. . . . . C- . . . . . T- . . . . . T. A. . . . . T. AGAC. GGC. CC. ACCC. . . . . CT. . . . . G. . . . .		926
TCoV	. . . AGA- - CTCTAACTACCCATTCTT- - TAAATTAAGTG- - TA- - ATAATACTTATCTATGGGATCAGGATAAGTTGT		1009
Beau	TC. . . . . ATAT. . . . . A. . . . . TA. . . . . AA. AAAAAAC. GC. C. G. . . . . GT. . . . . TT. . . . . TT. A. . . . . T. T. CCTT. . . . . T. AG. . . . . T.		1005
TCoV	GTATATTAGGGTCTGAAATATATGTGCCGGGGTATAAGGTTTATACGCAACATTCATATGCTGGCACTATACCGAATTAC		1089
Beau	. . . . . T. . . . . A. . . . . AGTCT. . . . . T. T. . . . . AT- . . . . . ATCT. . . . . TCACCC. AGTTGT. A. . . . . T. G. C- TA. AA. . . . . TAAT. . . . . GG.		1083
TCoV	CCA- - GCTTGTCTTCCAAAGGTCTTAGTT- TGGAAAAATTTTATAAGAATATAGGTTTCAATAGATTTTGTGTAACATA		1166
Beau	TTGTG. T. . . . . AA. TCA. TTTCA. . . . . T. C. A. . . . . GCTT. CGG. CC. CT. C. AGG. - . . . . . G. . . . . GCA. . . . . C- . . . . . CTT. . . . .		1157
TCoV	ATTGTTGGAGACTCGT- GATATA- TCATCTATAACCCAATACACTTGTGTTTAT- GTCGTTAGTGTAGAAGGTTTTGATA		1243
Beau	. . . . . GGTAGA. CA. . . . . T. . . . . T. T. . . . . GC. T. . . . . TC. . . . . TGGAGG. C. TTCGG. . . . . G. . . . . AA. GT. . . . . TA. TC. . . . . GT. AG. . . . . A. . . . . C		1237
TCoV	GTTATAATTGGGGTATTATATACAGGTATGTTGGTGTAACTTTGATTTTGGTACAGCTATGTATAGTGTGAAAAGTGCA		1323
Beau	A. A. . . . . TT. GAAT. . . . . GGAC. G. TAGTT. . . . . AC. AAG. G. GG. . . . . GC. C. C. . . . . T. CAA. CAGCC. C. . . . . AACCGCCA. TT		1317
TCoV	CCTATGCAGTTCATCTATA- CAGAGCAATGTCA- - TAATTATAATATTTACAACATAAGAGGTG- - TTGGTCAAATATT		1397
Beau	ATA. CT. . . . . AAA. . . . . AT. . . . . AT. ATATT. CT. T. AA. . . . . C. . . . . G. GT. GA. . . . . T. . . . . T. . . . . TAT. . . . . CAGAAC. . . . . C. . . . . GGT. . . . .		1397
TCoV	AAATGTCACTGGTAAAGATAAATACTACACTTGTAGACGGTGGTCTAGCAATAACTAGTGGCAGTGGTTTATTGGCCCTTC		1477
Beau	- . . . . A. T. . . . . AA. GTGACCG. CT. AG. TG. . . . . AGTT. TAA. TA. . . . . . . . . . GACG. AG. . . . . TTGGC. AT. . . . . GATA. A. C. G		1474
TCoV	GTAATAATGGAAGTCTTTATAGTGTACAACCGTGCCAGACAGTTTCAACT- - CAAGCTGTAATTGTTAACGGGAGTTTTG		1555
Beau	. . . . . TCC. - . . . . A. . . . . CA. . . . . G. - . . . . . G- . . . . . A- . . . . . T. TG. . . . . C. T. . . . . T. ATT. TAAG. . . . . T. ACCC. TG. . . . . AAGA. G. CA		1549
TCoV	TAGGCTTGTTTCTTCTGCATCTTGTAGGTTAGCCGGACAATTTGGTCTAGGTAACCACACAGAATATGTAGATGGTGGT		1635
Beau	ACCAGCA. . . . . G. AGT. T. TGG. G. . . . . AA. . . . . TA. . . . . TATTC. . . . . AC. TC. C. . . . . TG. G. . . . . T. GT. C. CAGCT. CT. . . . . AG		1629
TCoV	TGTCTTATTAGTAATAAGACTAGTACTAGAGAGCGTAGGTCTACTAACGCTGTTTATACTGGTGAATGC- ACAGGTT- TA		1713
Beau	AAC. AGT. . . . . TAC. TC. . . . . A. TC. C. . . . . A. G. . . . . ACA. . . . . C. T. T. . . . . G. . . . . T. C. A. . . . . - . . . . . A- A. . . . . TTG. . . . . AA. . . . . GCC		1705
TCoV	TTTGC- - - AGCTATGGGTACATCTTGTATTTATAGTAATGGCACAGTTCTTAATAGAACATTACCACAACCATCTGTTCA		1790
Beau	C. . . . . ATGTT. . . . . T. . . . . . . . . . AG. T. . . . . AA. ACC. G. . . . . T. . . . . A- . . . . . GCC. C. . . . . T. G. . . . . A. . . . . A. . . . . TG. AA. . . . .		1781
TCoV	AACTAGTGTGCAACCATTAAGTGTGTTGTTGGCAAAATGTGTCTATTCAAAAGAATTGACACTTGTGTAACCCACAGAAT		1870
Beau	GTT. - . . . . GC. . . . . T. . . . . T. TAA. . . . . ACT. A. . . . . . . . . . CTC. . . . . A. . . . . T. . . . . CAGT. . . . . C. ACT. AA. . . . . T. . . . . AGAT. . . . . G.		1858
TCoV	ACTTACAAACTCGTTATCAGAAAGTTACTATTGATTGTGCTAAGTATGTCTGTGGTGATTCTATAAGATGCCGCACTCTG		1950
Beau	. . . . . A. . . . . . . . . . G. . . . . ATGG. T. . . . . G. . . . . CCAA. . . . . A. . . . . CCTGC. . . . . T. . . . . CAG. . . . . C. GGAT. . . . . TA. A. AGT. . . . .		1938

Fig. 2. Nucleotide sequence of the amplified fragments containing entire spike (S) protein gene, gene 3, membrane (M) protein gene, gene 5, and nucleocapsid protein gene region of turkey coronavirus (TCoV) and their similarity to those of infectious bronchitis (IBV) strain Beaudette (GenBank accession number AJ311317). The positions where nucleotide bases are missing are indicated as (-) and identical nucleotides as (.). Heavy underlines below the sequence of TCoV indicate the putative start codons. Light lines above the sequence of TCoV indicate the stop codons. The conserved transcription-regulating nucleotide sequence (A/C)T(T/G)AACA, which is located upstream from the start codons of individual genes, is boxed. The start codon of IBV M protein gene is also underlined because it is at different position from that of TCoV M protein gene.



TCoV	CTGAACTATTTGACCCCTTTGAAGTTTGTGTTTACAGAGGAGGTAATTATTTGGGAGTTAGAGTCAGCTGACGAGTTTTCA	4021
Beau	. A . . . G . . . . . T . . . . . CTA . . . TA . . . . . T . . . . . AA . . . . . A . . . T . G . . . .	3898
TCoV	GGTGATGACGAATATATTGAGTAAATCGCTAGAGGAGAACGGAAGTTTCCTAACAGCAGTTTACATATTTATTGGATTTT	4101
Beau	. . . . . T . . . . . T . . . . . A . . . . . G . . . . . T . . . . . T . . . . . GC . . . . . T . . . . . A . . . G . A . . . . .	3978
TCoV	TAGCACTTTACCTATTAGGTAGAGCACTTCAAACATTTGTACAGGCTGCTGATGCTTGTGTGTTTGTGTTTGGTATACATGG	4181
Beau	. . . . . T . . . TC . . . . . G . . . . . A . . . . . T . . . . . T . . . . . T . . . . . T . . . . . T . . . . .	4058
TCoV	GTAGTAGTTCCTGGAGCTAAGGGTGCAGCCTTTGTATACAATTATACATATGGTAAAAA	4261
Beau	. . . . . A . . . . . A . . . . . A . . . . . A . . . . . G . . . . . G . . . . . G . . . . . T . . . . . T . . . . . A . . . . .	4138
TCoV	AGCGGTTATTGTTAACGAGTTCCTAAGAACGGTTGGAATAGT- - - - - CCAGCAAATTTCCGACGTG- - - - - ATGGAA- -	4329
Beau	. . . . . A . . . . . T . . . . . T . . . . . A . . . . . A . . . . . A . . . . . T . . . . . T . . . . . A . . . . . GA . . . . . C . . . . . C . . . . . AG . . . . . CA	4218
TCoV	AATTGCACACTTAACTCAGAGCAGGCAGTTCTGCTTTTTAAAGAATATAATTTATTTATAACCGCATTCTATTGTTTTT	4409
Beau	. . . . . T . . . . . T . . . . . G . . . . . TT . . . . . T . . . . . A . . . . . T . . . . . G . . . . . T . . . . . T . . . . . T . . . . . G . . . . . C . . . . .	4298
TCoV	AACCATACTACTTCAGTATGGGTACGCAACTAGGAGTCCGGTTATTTACATACTGAAAATGATAGTGTATGGTGCTTTT	4489
Beau	. . . . . A . . . . . A . . . . . C . . . . . T . . . . . A . . . . . A . . . . . AA . . . . . T . . . . . C . . . . . T . . . . . C . . . . .	4378
TCoV	GGCCCATTAACATTGCAGTAGGTGTAATTTTCATGTGTATATCCACCAAATACAGGAGGTCTTGTGCGCAGCGATAATACTT	4569
Beau	. . . . . C . . . . . C . . . . . C . . . . . C . . . . . C . . . . . C . . . . . C . . . . . C . . . . . C . . . . . C . . . . .	4458
TCoV	ACTGTGTTTGCTTGTCTTTCTTTTATAGGTTATTGGATCCAGAGTATCAGACTCTTTAAAAGGTGTAAGTCTTGGTGGTC	4649
Beau	. . . . . A . . . . . G . . . . . G . . . . . G . . . . . G . . . . . T . . . . . GC . . . . . G . . . . . A . . . . .	4538
TCoV	ATTTAACCCCGAATCTAATTCCTAGGTTCAACTCTCTAACTAATGGTCAACAATGCAATTTTGTCTATAGAGAGTGTGC	4729
Beau	. . . . . T . . . . . A . . . . . G . . . . . T . . . . . T . . . . . T . . . . . T . . . . . T . . . . . T . . . . . T . . . . .	4618
TCoV	CTATGGTGCTTTCCCTATTATAAAGAATGGTGTCTTTATTGTGAGGGTCAGTGGCTTGCTAAGTGTGAACCAGACCAC	4809
Beau	. . . . . A . . . . . T . . . . . A . . . . . T . . . . . T . . . . . T . . . . . T . . . . . T . . . . . T . . . . . T . . . . .	4698
TCoV	TTGCCTAAAGATATATTTGTGTGCACACCGGATAGACGTAATATCTATCGTATGGTGCAGAAATACACTGGTGATCAAAG	4889
Beau	. . . . . T . . . . . T . . . . . C . . . . . T . . . . . T . . . . . T . . . . . T . . . . . T . . . . . T . . . . .	4778
TCoV	CGGAAATAAGAAAAGGTTTGTACATTTGTCTATGCAAAGCAGTCAGTAGACACCGGCGAGCTAGAAAAGTGTAGCAACGT	4969
Beau	. . . . . G . . . . . T . . . . . T . . . . . T . . . . . T . . . . . T . . . . . T . . . . . T . . . . . T . . . . .	4858
TCoV	CTGGAAGTAGCCTTTACACATAAATGTGTGTGTGTAGAGAGTATTTAAAATTATTCTTCAATAGTGCCTCTATTTTAAAG	5049
Beau	GA . . . . . T . . . . . T . . . . . T . . . . . T . . . . . T . . . . . T . . . . . T . . . . . T . . . . . T . . . . .	4938
TCoV	GCGCGAAGAGTATTTGTTTTGAGGATATTAATATAAATCCTTTTTGTTTTGCACTCTCTTTCGCAAGAGTTATTATTTAA	5129
Beau	. . . . . AT . . . . . A . . . . . C . . . . . C . . . . . C . . . . . AT . . . . . C . . . . . T . . . . . C . . . . .	5012
TCoV	GCAACAGTTTTTCTTTCTTTTGTGTTGGAAGAAAGTTGTTGTTAATGGTGTAGAATTCGAAGTAGAAAATGGAAAAGTCC	5209
Beau	. . . . . G . . . . . G . . . . . A . . . . . T . . . . . T . . . . . T . . . . . T . . . . . T . . . . . T . . . . . T . . . . .	5049
TCoV	ACTACGAAGGAAATCCAATCTTCCAAAAAGGTTGTTGTAGTTTGTGGTCCGATTATAAGAAAGATTAGAATAATTAAGCC	5289
Beau	. . . . . CA . . . . . G . . . . . T . . . . . GA . . . . . A . . . . . G . . . . . A . . . . . G . . . . . A . . . . .	5124
TCoV	ACCAACTACACTTATTTTTAAAAGGGCGTTTTAGCTTATAAAT	5369
Beau	. . . . . T . . . . . T . . . . . GA . . . . . T . . . . . CT . . . . . A . . . . . C . . . . . AA . . . . .	5204
TCoV	TTTGGAAGAGCGGTTGTTTTCATGTTACAAAGCACTACTATTAACCTCAACTTAGAGTGTAGATAGGTTAATTTTATAGATCA	5449
Beau	. . . . . A . . . . . A . . . . . T . . . . . T . . . . . T . . . . . C . . . . . T . . . . . T . . . . . T . . . . . T . . . . .	5284
TCoV	CGGACCAAAAACGCGTTTTAACGTGTAGTAGGCGAGTGCTTTTTGTTTCAGTTAGATTTAGTTTATAGGTTGGCGTTTTACGC	5529
Beau	. . . . . T . . . . . CT . . . . . C . . . . . C . . . . . AG . . . . . A . . . . . A . . . . . A . . . . . A . . . . .	5364
TCoV	CCACCCAACCGCTGGT	5603
Beau	. . . . . T . . . . . C . . . . . TAGTAA . . . . . G . . . . . T . . . . . T . . . . . T . . . . . T . . . . .	5444
TCoV	GAGAAGGATTAGATTGTGTTACTT	5683
Beau	. . . . . G . . . . . T . . . . . C . . . . . G . . . . . T . . . . . C . . . . . T . . . . . T . . . . . T . . . . .	5524
TCoV	CGAGGGAAAAATTTGTGAGGAACACATAAATAATAATAATCTTTTGTCA	5763
Beau	. . . . . A . . . . . C . . . . . C . . . . . C . . . . . C . . . . . C . . . . . C . . . . . C . . . . . C . . . . .	5604
TCoV	GCTCCGGCGCCAGTCATCAAAC	5843
Beau	. . . . . C . . . . . A . . . . . T . . . . . T . . . . . A . . . . . C . . . . . C . . . . . A . . . . . CA . . . . .	5684
TCoV	AATAAAAAGCCAAGAAGCTAAATTCACCTCAGCCTAAGTTTGAAGGTAGCGGTGTTCTGATAATGAAAATCTTAAAACAA	5923
Beau	. . . . . T . . . . . A . . . . . C . . . . . C . . . . . C . . . . . C . . . . . C . . . . . CA . . . . . GC . . . . .	5764

Fig. 2. (Continued)

TCoV	GCCAGCAACATGGGTATTGGAGACGCCAAGCTAGGTTTAAAGTCAAGTAGAGGAGGAAGAAAACCAGTCCCAGATGCCTGG	6003
Beau	..... A . C . . . . . C . . . . . C . . . . . G . C . A . . . T . . . . . T . . . . .	5844
TCoV	TACTTCTATTATACTGGAACAGGACCAGCCGCTGACCTGCAATGGGGTGATAGCCAAGCGGGTATAGTGTGGGTTGCTGC	6083
Beau	..... T . C . . . . . G . . . . . T . . . . . A . C . . . . . C T . . . . . A T . . . . .	5924
TCoV	AAAGGGTGCTGATGTTAAATCTAAATCTAACCAGGGTACAAGGGACCCTGACAAGTTTGACCAATATCCACTACGTTTCT	6163
Beau	T . . . . . A C . . . . . G . . . . . C . . . . . T . . . . . A . . . . . T . . . . . T . . . . . C . . . . . A . . . . .	6004
TCoV	CTGATGGTGGACCTGATGGTAATTTCCGTTGGGATTTTCATTCTTTGAATCGTGGTAGGAGTGAAGATCAACAGCAGTT	6243
Beau	. G . . . . . C . . . . . C . . . . . G . . . . . C . . . . . C C . . . . . C . . . . . C . . . . .	6084
TCoV	TCATCAGCAGCATCTATTAGAGCACCATCACGTGAAGGTTTCGGTGGTTCGTAGAAGTGGTGCTGAAGATGATCTTATTGC	6323
Beau	..... G . . . . . G . . . . . G . . . . . A . T . . . . . G . . . . . C . . . . .	6164
TCoV	TCGCGCAGCAAAGATAATTCAGGATCAGCAGAAGAGGGGTTCTCGCATTACCAAGGCTAAGGCTGAAGAAATGGCCCATC	6403
Beau	..... T . . . . . C . . . . . C . . . . . A . A . . . . . C . . . . . A . . . . . T . . . . . T . . . . .	6244
TCoV	GCCGGTACTGCAAGCGCACTGTCCCACCTGGTTATAAGGTTGAGCAAGTTTTTGGTCCCCGTACTAAAGGTAAGGAGGGA	6483
Beau	..... T . . . . . A . . . . . A A . . . . . G . . . . . T . . . . . G . . . . . G . . . . .	6324
TCoV	AATTTTGGTGATGACAAGATGAATGAGGAAGGTATTAAGGATGGGCGTGTTACAGCAATGCTCAACCTAGTCCCTAGCAG	6563
Beau	.....	6404
TCoV	CCATGCTTGTCTTTTGAAGTAAATTGACGCCCAAACCTCAACCAGATGGGCTGCACTTGAAATTTGAATTTACTACTG	6643
Beau	..... G . . . . . G . G . . . . . A . . . . . T . . . . . T . . . . . G . . . . .	6484
TCoV	TGGTCCCACGTGATGATCCGCAGTTTGATAATTATGTTAGTATTTGTGATCAGTGTGTGATGGTGTAGGAACACGTCCA	6723
Beau	..... T . . . . . C . . . . . G . A A . . . . . G . . . . .	6564
TCoV	AAAGATGACGAACCGAGACCAAAGTCACGCGCAAGTTCAAGACCTGCTACAAGAGGAAATTCCTCAGCTCCAAGACAACA	6803
Beau	..... A . A . . . . . T . . . . . G . . . . .	6644
TCoV	GCGTCTAAAGAAGGAGAAAAAGCCAAAGAAGCAGGATGATGAAGTAGATAAAGCATTGACCTCAGATGAGGAGAGGAACA	6883
Beau	..... C . C . . . . . T . . . . . C . . . . .	6724
TCoV	ATGCACAGCTGGAATTTGATGATGAACCAAGGTGATTAAGTGGGGGGATTTCAGCTCTAGGGGAGAATGAACCTTTGAGTA	6963
Beau	..... T . . . . . G . . . . . A . . . . . G . . . . .	6804

Fig. 2. (Continued).

acid and amino acid pair distances. Based on the obtained sequences of TCoV and previously published sequences of different coronaviruses, phylogenetic trees were constructed according to the coding sequences for S, M, and N genes.

**3. Results**

*3.1. Complete nucleotide sequences of 3' end coding region for structural protein genes of turkey coronavirus*

Cloning and sequencing of the 4 overlapping fragments revealed a total of 6963 nucleotides in a region containing the entire S protein gene, tricistronic gene 3, M protein gene, bicistronic gene 5, and N protein gene of TCoV in the present study. The primary structures of the coding sequences for these genes of TCoV in the present study were very similar to those found in the corresponding genomic regions of IBV strain Beaudette as shown in Fig. 2 and Table 1. The canonical consensus transcription-regulating sequence (TRS) of IBV, CT(T/G)AACAA, was also found in TCoV in the present study. Both the nucleotide sequence of the TRS and the distance between the 3' end of the TRS and the initiation codon of the downstream adjacent ORF were highly conserved between TCoV in the present study and IBV (Table 1).

Table 1

Comparison of the 3' end encoding regions between turkey coronavirus (TCoV) and infectious bronchitis virus (IBV) strain Beaudette (GenBank accession number AJ311317)

Virus	Gene	ORF <sup>a</sup> size (nucleotides)	TRS <sup>b</sup> sequence	RS distance <sup>c</sup> (nucleotides)
TCoV	Spike	3612	ctgaacaa	52
	Gene 3		atgaacaa	23
	3a	174		
	3b	195		
	3c	318		
	Membrane	669	cttaacaa	74
	Gene 5		cttaacaa	9
	5a	198		
	5b	243		
	Nucleocapsid	1230	cttaacaa	93
IBV	Spike	3489	ctgaacaa	52
	Gene 3		ctgaacaa	23
	3a	174		
	3b	195		
	3c	330		
	Membrane	678	cttaacaa	77
	Gene 5		cttaacaa	9
	5a	198		
	5b	249		
	Nucleocapsid	1230	cttaacaa	93

<sup>a</sup> ORF: open reading frame.

<sup>b</sup> TRS: transcription-regulating sequence.

<sup>c</sup> The distance is calculated as nucleotides between 3' end of TRS and the ATG start codon of the corresponding first downstream ORF.

Table 2

Sequence pair distances for nucleic acid and deduced amino acid sequence of the entire spike protein gene region of turkey coronavirus (TCoV) with other coronaviruses

		Nucleotide identity (%)															
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	TCoV <sup>a</sup>	100	95.1	95.4	51.7	52.5	52.1	42.9	41.8	41.8	38.3	42.6	42.8	41.1	38.8	41.3	39.3
2	TCoV-Gh <sup>b</sup>	90.2	100	96.8	50.6	51.3	51.0	42.2	41.4	41.3	38.1	41.9	42.2	40.2	38.1	41.0	38.4
3	TCoV-GI <sup>c</sup>	91.1	95.1	100	50.3	51.0	50.7	42.0	41.3	41.3	37.9	41.8	42.1	40.1	38.3	41.0	38.2
4	IBV-CU <sup>d</sup>	33.9	32.8	32.7	100	85.3	85.6	41.7	39.9	39.7	38.6	41.5	41.8	40.6	38.6	40.3	39.5
5	IBV-KB <sup>e</sup>	33.8	33.1	32.7	83.2	100	94.4	42.6	40.7	40.6	38.1	42.2	42.7	41.1	39.3	41.2	40.6
6	IBV-Beau <sup>f</sup>	33.9	33.1	32.7	84.6	94.0	100	42.8	40.7	40.7	38.1	42.3	42.9	41.6	39.3	41.1	40.4
7	BCoV <sup>g</sup>	22.3	21.8	21.8	20.4	21.3	20.7	100	38.2	38.6	33.2	93.9	98.7	67.1	36.1	38.7	38.6
8	CCoV <sup>h</sup>	25.4	24.0	24.5	23.8	23.2	23.3	21.3	100	90.4	44.5	35.5	35.9	35.6	50.3	82.7	33.7
9	FECoV <sup>i</sup>	25.2	23.9	24.4	24.0	23.3	23.5	21.3	93.1	100	44.1	36.3	36.2	36.1	50.9	84.3	33.7
10	HCoV229E <sup>j</sup>	25.9	25.2	25.6	24.7	24.8	24.5	18.4	38.8	39.1	100	38.0	38.5	37.8	57.5	55.3	36.5
11	HCoVOC43 <sup>k</sup>	22.3	21.8	21.7	20.3	21.2	20.7	91.9	20.2	20.1	21.0	100	94.1	67.0	36.2	38.6	38.2
12	HECoV <sup>l</sup>	22.3	21.6	21.7	20.3	21.2	20.6	97.9	20.0	19.8	21.4	92.2	100	67.0	36.2	38.6	38.5
13	MHV <sup>m</sup>	22.3	22.1	21.7	21.1	21.8	21.2	65.7	19.5	19.5	22.1	65.4	65.9	100	35.8	38.4	37.2
14	PEDV <sup>n</sup>	26.4	25.4	25.9	25.1	25.4	25.5	19.8	42.6	43.3	49.7	20.2	20.2	20.9	100	52.6	32.8
15	TGEV <sup>o</sup>	25.0	23.6	24.2	24.1	23.5	23.6	21.3	80.5	81.5	48.0	21.6	21.4	21.2	44.8	100	33.8
16	SARS <sup>p</sup>	19.5	18.6	18.8	21.6	21.1	21.4	22.6	18.2	18.3	22.1	22.4	22.7	22.3	19.7	17.9	100
		Amino acid identity (%)															

<sup>a</sup> TCoV: a US, Indiana, isolate of TCoV.<sup>b</sup> TCoV-Gh: an isolate of TCoV. GenBank accession number AY342356.<sup>c</sup> TCoV-GI: an isolate of TCoV. GenBank accession number AY342357.<sup>d</sup> IBV-CU: a German strain, CU-T2, of infectious bronchitis virus (IBV). GenBank accession number U49858.<sup>e</sup> IBV-KB: a Japanese strain, KB8523, of IBV. GenBank accession number M21515.<sup>f</sup> IBV-Beau: a US strain, Beaudette, of IBV. GenBank accession number AJ311317.<sup>g</sup> BCoV: bovine coronavirus. GenBank accession number M64668.<sup>h</sup> CCoV: canine coronavirus. GenBank accession number X77047.<sup>i</sup> FECoV: feline enteric coronavirus. GenBank accession number X80799.<sup>j</sup> HCoV229E: human coronavirus strain 229E. GenBank accession number AF344186.<sup>k</sup> HCoVOC43: human coronavirus strain OC43. GenBank accession number L14643.<sup>l</sup> HECoV: human enteric coronavirus. GenBank accession number L07748.<sup>m</sup> MHV: murine hepatitis coronavirus. GenBank accession number U72635.<sup>n</sup> PEDV: porcine epidemic diarrhea coronavirus. GenBank accession number Z25483.<sup>o</sup> TGEV: porcine transmissible gastroenteritis coronavirus. GenBank accession number AJ271965.<sup>p</sup> SARS: severe and acute respiratory syndrome coronavirus. GenBank accession number NC\_004718.



### 3.2. Sequence comparison and phylogenetic analysis

Pair-wise comparison of nucleotide and deduced amino acid sequence distance between TCoV S protein gene in the present study and the homologous gene sequences of other known coronaviruses is summarized in Table 2. The similarity score between TCoV in the present study and other non-TCoV coronaviruses within the S protein gene region ranged from 38.3% to 52.5% at the nucleotide level or from 19.5% to 33.9% at the amino acid level.

The similarity scores between TCoV in the present study and IBV strains within the M or N protein gene region were high (>80%). In contrast, the similarity score between TCoV in the present study and other mammalian coronaviruses within the M or N protein gene region ranged from 24.8% to 30.8% for nucleotide sequence and from 16.9% to 29.1% for deduced amino acid sequence. The tricistronic gene 3 with 3 overlapping ORFs, 3a–c, in between S and M genes as well as the dicistronic gene 5 with 2 overlapping ORFs, 5a and 5b, in between M and N genes of TCoV in the present study all shared high similarity with the corresponding genomic sequences of IBV strains.

Phylogenetic analysis according to the deduced amino acid sequence of S, M, or N proteins indicated that TCoV in the present study was clustered within the same genomic lineage as the IBV strains while all the other mammalian coronaviruses were grouped into separate clusters corresponding to antigenic groups I and II (data not shown).

## 4. Discussion

Total 6963 nucleotides of TCoV genome were cloned and sequenced in the present study. This region is likely to include all of the viral genes excluding the polymerase gene and, thus, provides substantial genetic information of the virus for comparison with other coronaviruses. The genomic structures of ORFs for S protein, 3a–c, M protein, 5a–b, and N protein were very similar to those of IBV. The phylogenetic analysis based on the deduced amino acid sequences of S, M, or N protein all showed that the TCoV in the present study was classified within the same genomic lineage with IBV strains while all the other mammalian coronaviruses including BCoV were grouped to separate clusters. The nucleotide sequences of ORFs for 3a–c, M protein, 5a–b, and N protein of TCoV shared high similarity (82.1–92.0%) with the corresponding sequences of IBV. These results clearly demonstrated the close relationship of TCoV in the present study to avian IBV.

The presence of tricistronic gene 3 in between 3' end of S gene and 5' end of M gene as well as the presence of dicistronic gene 5 in between 3' end of M gene and 5' end of N gene are particular features of avian coronaviruses, TCoV and IBV. These particular genomic structures are not found in any other mammalian coronaviruses as determined to date. These distinct features of genome structure implied that TCoV shares a relatively close evolutionary relationship with IBV.

The predicted proteins of ORF 3a–c, 5a and 5b were small (about or less than 10 kd). The functions of these gene products are not known. Several ORFs encoding non-structural proteins have been recognized in coronavirus genomes (Bournnell et al., 1987; Lee et al., 1991; Herold et al., 1993; Eleouet et al., 1995). The number, nucleotide sequence, and gene order of these ORFs varied remarkably among different coronaviruses. It is speculated that these genes were inserted into different sites in the coronavirus genomes due to the RNA recombination-prone discontinuous transcription mechanism and were not essential for virus replication and pathogenesis. However, sequence analysis in the present study indicated that both nucleotide sequences and locations of these ORFs and their consensus TRS of TCoV are highly conserved with those of IBV. Given such a highly conserved sequences and structures within avian coronaviruses, genes 3 and 5 may play important roles in coronavirus pathogenesis to avian species.

One of the characteristic features for coronavirus replication is the synthesis of a 3' coterminal nested set of polycistronic subgenomic mRNAs by a discontinuous transcription mechanism. Several conserved TRS have been identified for different coronavirus proximal to the initiation codon of the first ORF for each particular subgenomic mRNA. The consensus sequences of the TRS sites are CT(T/G)AACAA for IBV, ATC(T/C)AAAC for BCoV, AACTAAAC for TGEV, AATC(T/C)A(A/T)AC for MHV, and AACTAAAC for FIPV (Spaan et al., 1988; Stirrups et al., 2000). The distance between the TRS and the first ORF is different for each subgenomic mRNA of different coronaviruses. Both the nucleotide sequence of TRS and the distance between the TRS 3' end and the initiation codon of first ORF are suggested to play important role in the transcription of mRNAs. As shown in the present study, the TRS sequences of TCoV were found highly conserved with those of the corresponding genes of IBV except one nucleotide substitution in that of gene 3. The TRS of gene 3 is ATGAACAA for TCoV and CTGAACAA for IBV. The distances between TRS and initiation codon of S gene, gene 3, gene 5, and N gene of TCoV were all the same as those of IBV while the distances for TCoV or IBV M gene are 74 or 77, respectively. The highly conserved sequences and structures of TRS between TCoV and IBV provide further evidence that these two avian coronaviruses share close evolutionary relationship. These highly conserved TRS sequences of IBV has been shown to be recombination "hot spot" and may serve as the template switching sites for the viral encoded RNA dependent RNA polymerase (Lee and Jackwood, 2000). These recombination events play important role to the emergence of new IBV variants responsible for continuous outbreaks in the chicken flocks vaccinated with live attenuated viruses due to failure of cross protection. It is possible that the similar recombination events of IBV in chicken may contribute to the origin and evolution of TCoV in turkey and merit further investigation.

Even though the close genetic relationship between TCoV and IBV was clearly demonstrated as discussed above,

these two avian coronaviruses are dramatically different at the S protein gene level. The similarity of S protein sequences between TCoV in the present study and IBV strains (33.8–33.9%) is much lower than that among IBV strains (83.2–94.0%). The difference of nucleotide between TCoV in the present study and IBV seems to be randomly distributed throughout the entire S gene except a stretch of 225 nucleotides from the 3' end that shared high similarity (88.9%) with the corresponding sequences of IBV. These observations suggested that cross-over homologous recombination, very likely by a template switching mechanism, occurred around the consensus TRS site of S gene and within the 3' end 225 nucleotides region (involving the TRS site of gene 3) and resulted in a whole new codon reading frame for S protein of TCoV with conserved TRS and other genomic structure features of IBV. Spike protein of coronaviruses has been well known as the major structural protein responsible for attachment, fusion, and penetration of virions to the target cells. The substantial difference of S protein gene between TCoV and IBV well explains the different host tropism and different tissue pathogenicity of these two avian coronaviruses. Turkey coronavirus is associated with enteric disease of turkey while IBV is usually associated with respiratory disease in chicken.

Two group-specific monoclonal antibodies, which reacted with a broad spectrum of homologous and heterologous IBV serotypes, were tested for reactivity with TCoV in a previous study (Loa et al., 2000). The antibody specific to M protein (Mab 919) of IBV had strong cross reactivity with TCoV but the antibody specific to S protein (Mab 94) of IBV did not react with TCoV. In line with these previous observations of antigenicity, the sequence analysis in the present study revealed a high homology of M protein gene between TCoV in the present study and IBV. On the other hand, the difference of S protein gene between TCoV in the present study and IBV is substantial. Therefore, molecular diagnostic assay or antigenic analysis using antibody specific to S protein or gene will be useful tools to differentiate TCoV from IBV.

The results of sequence analysis in the present study stress the close relationship of TCoV to IBV. Coronavirus genomes are dynamic with high frequency of recombination, insertion, and deletion, subsequently, may result in significant genetic differences. Further cloning and sequencing analysis of full-length genomic sequences of more TCoV isolates are under way for revealing a faithful picture of the TCoV genome.

## Acknowledgements

The authors thank the financial support provided by US Department of Agriculture.

## References

- Akin, A., Lin, T.L., Wu, C.C., Bryan, T.A., Hooper, T., Schrader, D., 2001. Nucleocapsid protein gene sequence analysis reveals close genomic relationship between turkey coronavirus and avian infectious bronchitis virus. *Acta Virol.* 45 (1), 31–38.
- Akin, A., Wu, C.C., Lin, T.L., 1999. Amplification and cloning of complete infectious bursal disease virus genomic RNA segments by a long and accurate PCR. *J. Virol. Methods* 82, 55–61.
- Barnes, W.M., 1994. PCR amplification of up to 35 kb DNA with high fidelity and high yield from bacteriophage templates. *Proc. Natl. Acad. Sci. U.S.A.* 91, 2216–2220.
- Bournsell, M.E., Brown, T.D., 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. *J. Gen. Virol.* 68, 57–77.
- Breslin, J.J., Smith, L.G., Fuller, F.J., Guy, J.S., 1999a. Sequence analysis of the matrix/nucleocapsid gene region of turkey coronavirus. *Intervirology* 42, 22–29.
- Breslin, J.J., Smith, L.G., Fuller, F.J., Guy, J.S., 1999b. Sequence analysis of the turkey coronavirus nucleocapsid protein gene and 3' untranslated region identifies the virus as a close relative of infectious bronchitis virus. *Virus Res.* 65, 187–193.
- Cavanagh, D., Mawditt, K., Sharma, M., Drury, S.E., Ainsworth, H.L., Britton, P., Gough, R.E., 2001. Detection of a coronavirus from turkey poult in Europe genetically related to infectious bronchitis virus of chickens. *Avian Pathol.* 30, 355–368.
- Cavanagh, D., Mawditt, K., Welchman, D.B., Britton, P., Gough, R.E., 2002. Coronaviruses from pheasants (*Phasianus colchicus*) are genetically closely related to coronaviruses of domestic fowl (infectious bronchitis virus) and turkeys. *Avian Pathol.* 31, 81–93.
- Chomczynski, P., Sacchi, N., 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate–phenol–chloroform extraction. *Anal. Biochem.* 162, 156–159.
- Dea, S., Marsolais, G., Beaubien, J., Ruppanner, R., 1986. Coronaviruses associated with outbreaks of transmissible enteritis of turkeys in Quebec: hemagglutination properties and cell cultivation. *Avian Dis.* 30, 319–326.
- Dea, S., Verbeek, A.J., Tijssen, P., 1990. Antigenic and genomic relationships among turkey and bovine enteric coronaviruses. *J. Virol.* 64, 3112–3118.
- Eleouet, J.F., Rasschaert, D., Lambert, P., Levy, L., Vende, P., Laude, H., 1995. Complete genomic sequence of the transmissible gastroenteritis virus. *Adv. Exp. Med. Biol.* 380, 459–461.
- Gonzalez, J.M., Gomez-Puertas, P., Cavanagh, D., Gorbalenya, A.E., Enjuanes, L., 2003. A comparative sequence analysis to revise the current taxonomy of the family Coronaviridae. *Arch. Virol.* 148, 2207–2235.
- Guy, J.S., Barnes, H.J., Smith, L.G., Breslin, J., 1997. Antigenic characterization of a turkey coronavirus identified in poult enteritis and mortality syndrome-affected turkeys. *Avian Dis.* 41, 583–590.
- 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.
- Lee, H.J., Shieh, C.K., Gorbalenya, A.E., Koonin, E.V., La Monica, N., Tuler, J., Bagdzhadzhyan, A., Lai, M.M., 1991. The complete sequence (22 kb) of murine coronavirus gene 1 encoding the putative proteases and RNA polymerase. *Virology* 180, 567–582.
- Lee, C.W., Jackwood, M.W., 2000. Evidence of genetic diversity generated by recombination among avian coronavirus IBV. *Arch. Virol.* 145, 2135–2148.
- Lin, T.L., Loa, C.C., Wu, C.C., 2002. Existence of gene 5 indicates close genomic relationship of turkey coronavirus to infectious bronchitis virus. *Acta Virol.* 46, 107–116.
- Loa, C.C., Lin, T.L., Wu, C.C., Bryan, T.A., Thacker, H.L., Hooper, T., Schrader, D., 2000. Detection of antibody to turkey coronavirus by antibody-capture enzyme-linked immunosorbent assay utilizing infectious bronchitis virus antigen. *Avian Dis.* 44, 498–506.
- Nagaraja, K.V., Pomeroy, B.S., 1997. Coronaviral enteritis of turkeys (bluecomb disease). In: Calnek, B., Barnes, H.J., Beard, C.W., McDougald, L.R., Saif, Y.M. (Eds.), *Diseases of Poultry*, 10th ed. Iowa State University Press, Ames, Iowa, pp. 686–692.

- Penzes, Z., Gonzalez, J.M., Calvo, E., Izeta, A., Smerdou, C., Mendez, A., Sanchez, C.M., Sola, I., Almazan, F., Enjuanes, L., 2001. Complete genome sequence of transmissible gastroenteritis coronavirus PUR46-MAD clone and evolution of the purdue virus cluster. *Virus Genes* 23, 105–118.
- Ritchie, A.E., Desmukh, D.R., Larsen, C.T., Pomeroy, B.S., 1973. Electron microscopy of coronavirus-like particles characteristic of turkey bluecomb disease. *Avian Dis.* 17, 546–558.
- Spaan, W., Cavanagh, D., Horzinek, M.C., 1988. Coronaviruses: structure and genome expression. *J. Gen. Virol.* 69, 2939–2952.
- Stirrup, K., Shaw, K., Evans, S., Dalton, K., Casais, R., Cavanagh, D., Britton, P., 2000. Expression of reporter genes from the defective RNA CD-61 of the coronavirus infectious bronchitis virus. *J. Gen. Virol.* 81, 1687–1698.
- Verbeek, A., Tijssen, P., 1991. Sequence analysis of the turkey enteric coronavirus nucleocapsid and membrane protein genes: a close genomic relationship with bovine coronavirus. *J. Gen. Virol.* 72, 1659–1666.
- Van Regenmortel, M.H.V., Fauquet, C.M., Bishop, D.H.L., Carstens, E.B., Estes, M.K., Lemon, S.M., Maniloff, J., Mayo, M.A., McGeoch, D.J., Pringle, C.R., Wickner, R.B., 2000. Virus taxonomy: the classification and nomenclature of viruses. The Seventh Report of the International Committee on Taxonomy of Viruses. Academic Press, San Diego, California.