Original Paper

Intervirology

Intervirology 2006;49:230–238
DOI: 10.1159/000091470

Received: January 17, 2005
Accepted after revision: June 8, 2005
Published online: February 16, 2006

Comparison of 3′-End Encoding Regions of Turkey Coronavirus Isolates from Indiana, North Carolina, and Minnesota with Chicken Infectious Bronchitis Coronavirus Strains

Chien Chang Loa Ching Ching Wu Tsang Long Lin

Department of Veterinary Pathobiology, Purdue University, West Lafayette, Ind., USA

Key Words
Turkey coronavirus · Infectious bronchitis virus · Coronavirus · Genomic relationship

Abstract
Objective: To analyze the 3′-end structural protein-encoding region of turkey coronavirus (TCoV) isolates associated with outbreaks of acute enteritis in Indiana, North Carolina, or Minnesota. ***Methods:*** Four isolates of TCoV were sequenced over the entire 3′-end structural protein-encoding region and compared phylogenetically along with the corresponding sequences of infectious bronchitis virus (IBV) strains. ***Results:*** The sequence similarity between TCoV and IBV was lower than that among TCoV isolates or that among IBV strains. The variation of sequences between TCoV and IBV was mainly contributed by the S protein gene. The sequence similarity of S gene between TCoV and IBV was lower than that among TCoV isolates or that among IBV strains. The phylogenetic tree based on the S protein region was similar to that based on the entire 3′-end structural protein-encoding region with TCoV isolates and IBV strains grouped in two separate clusters. The phylogenetic tree based on other genes had a very different topology with TCoV isolates randomly forming groups with different IBV strains. ***Conclusions:*** These results suggested that TCoV probably shared the same origin with that of IBV and acquired sequences of S gene for turkey intestine tropism during the process of evolution in a separate environment.

Introduction

Turkey coronavirus (TCoV) causes an acute and highly contagious enteric disease. Turkey coronaviral enteritis is the most costly disease of turkey encountered in Minnesota between 1951 and 1971 and is characterized by acute enteritis, diarrhea, decreased body weight gain, inappetence, ruffled feathers, and uneven flock growth. The disease accounted for 23% of all turkey mortality and over a half million dollars in lost income in Minnesota in 1966 [1]. Outbreaks of turkey poult enteritis associated with TCoV have contributed to significant economic losses encountered by a large number of turkey producers in Indiana, North Carolina, and other states for the last several years. The disease is not easily eliminated and is frequently encountered in areas with high concentrations of turkeys on a year-round basis [1]. There are currently no effective vaccines or medications to prevent the disease and treatments of the disease are often unsuccessful. In order to develop strategies for the diagnosis, control, and prevention of turkey coronaviral infection, in-depth understanding of the molecular biology of TCoV is essential.

KARGER

Fax +41 61 306 12 34
E-Mail karger@karger.ch
www.karger.com

0300–5526/06/0494–0230$23.50/0

Accessible online at:
www.karger.com/int

Tsang Long Lin
Purdue University, Department of Veterinary Pathobiology, ADDL
West Lafayette, IN 47907-2065 (USA)
Tel. +1 765 494 7927, Fax +1 765 494 9181
E-Mail tllin@purdue.edu

Coronaviruses are pleiomorphic, enveloped spherical particles surrounded by a fringe of 20-nm-long club-shaped spikes. The diameter of coronaviral particle varies from 50 to 150 nm. The coronavirus genome is a positive and single-stranded capped RNA with a polyadenylated 3′-end. The 5′ two thirds of the genome, roughly 20 kb, consist of two overlapping open reading frames (ORFs) that encode nonstructural proteins including the viral RNA-dependent RNA polymerase and proteases. Another one third nucleotide sequences from 3′-end contain ORFs encoding the major structural proteins, spike (S), membrane (M), and nucleocapsid (N) proteins, in the order of 5′ to 3′ along the genome, respectively. Spike protein contains neutralizing and/or group-specific epitopes and is highly variable among different coronaviruses. In contrast, M and N proteins are more conserved among coronaviruses between different antigenic groups. There are several ORFs encoding other structural or nonstructural proteins, which varied in number, size, and order of genes among different coronaviruses. Like many other RNA viruses, the genome of coronavirus is very dynamic due to the lack of exonuclease editing the function of the viral RNA-dependent RNA polymerase and a discontinuous transcription mechanism for coronavirus replication. Mutation rates during viral replication have been estimated in the range of 10^{-3} to 10^{-5} substitutions per nucleotide. Given the size of coronavirus genome at approximately 30,000 nucleotides, an average of 3 mutations per template copied is anticipated.

There are more than 20 different serotypes of infectious bronchitis coronavirus (IBV) recognized worldwide. Most of these serotypes do not cross-protect and cause outbreaks of infectious bronchitis in chicken flocks vaccinated with attenuated virus of different serotype. The diversity of IBV is generated by insertions, deletions, point mutations, and RNA recombinations. A new serotype of IBV is still emerging [2]. A mutation rate of 1.5% nucleotide changes and the evolutionary rate of 2.5% amino acid changes per year in a hypervariable region of the S protein gene were estimated for the arising of a new serotype GA98 from serotype DE072 [2]. In a previous study, we demonstrated that the genomic structure of 3′-end structural protein-encoding region of an Indiana isolate of TCoV is very similar to that of IBV [3]. It is speculated that a similar phenomenon of genetic diversity in IBV genome may exist among TCoV isolates. In addition, the evolutionary interactions between these two avian viruses with distinct species and tissue tropism are not clear and merit further investigation. Analysis of genomic structure of different TCoV isolates from different geographic areas is necessary to delineate the relationships between TCoV and IBV. The objective of the present study is to examine the sequences of the 3′-end structural protein-encoding region of TCoV isolates associated with recent outbreaks of acute enteritis of turkeys in Indiana, North Carolina, and Minnesota as well as an isolate recovered from affected turkeys in Minnesota in the early 1970s. The TCoV sequences were compared with the corresponding published sequences of IBV strains Beaudette [4], KB8523 [5], CU-T2 [6], DE072 [7, 8], and D1466 [7, 8].

Materials and Methods

Viruses

The TCoV isolates used in the present study were recovered from fecal contents and intestines of turkey poults with recent outbreaks of acute coronaviral enteritis in Indiana in 1994 (isolate 540), Minnesota in 1996 (isolate 310), or North Carolina in 1999 (isolate 1440). The prototype isolate of TCoV recovered from turkey poults with outbreaks of acute coronaviral enteritis in Minnesota in the early 1970s was obtained from American Type Culture Collection (ATCC, Rockville, Md., USA). The viruses were passaged 5 times in 22-day-old embryonating turkey eggs. The presence of TCoV in the intestines of embryos was confirmed by TCoV-specific immunofluorescence antibody assays and electron microscopy at the Indiana State Animal Disease Diagnostic Laboratory in West Lafayette, Ind., USA [9].

The published sequences of the IBV strains Beaudette isolated in 1937, KB8523 isolated in the 1980s, CU-T2 isolated in between 1991 and 1992, DE072 isolated in 1992, and D1466 isolated in 1979 were retrieved from GenBank. The accession numbers for Beaudette, KB8523, and CU-T2 are AJ311317, M21515, and U49858, respectively. The sequences for DE072 are compiled from GenBank accession numbers: U77298, AF201930, AF202998, AF202999, AF203000, and AF203001. The sequences for D1466 are compiled from GenBank accession numbers M21971, X58001, AF203003, AF203004, AF203005, AF203006, and AF203007.

RNA Isolation

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 [10, 11]. The RNA was dissolved in 150 μl of diethyl pyrocarbonate-treated sterile double-distilled water and a portion of it was quantified by spectrophotometry at 260 nm wavelength.

Reverse Transcription

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., USA). Briefly, the total RNA (2 μg) was heat denatured at 100° for 3 min and slowly cooled to 22° in 15 min in reverse transcription buffer (Life Technologies) containing random hexamers (40 ng). Then, the reverse transcription was carried out at 42° for 60 min.

Table 1. Comparison of sizes of ORF in the 3′-end encoding regions between TCoV isolates and IBV strains

	ORF							
	spike	3a	3b	3c	membrane	5a	5b	nucleocapsid
TCoV								
540	3,612	174	195	318	669	198	243	1,230
ATCC	3,612	174	195	312	672	198	249	1,224
310	3,612	174	195	312	672	198	249	1,230
1440	3,624	NA	195	324	672	198	249	1,230
IBV								
Beau	3,489	174	195	330	678	198	249	1,230
CU	3,507	174	195	282	687	198	249	1,230
KB	3,489	NA	195	330	678	198	249	1,230
DE	3,483	174	195	330	678	198	249	1,230
D1466	3,456	174	195	330	678	198	249	1,230

NA = Not applicable.

PCR Amplification

Three microliters of cDNA were used in PCR amplifications with the primers designed for four overlapping fragments covering the entire 3′-end structural protein-encoding regions as described previously [3]. Eight microliters of the PCR products were electrophoresed on 1% agarose gels in TBE buffer, stained with ethidium bromide, and observed under ultraviolet light.

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, Calif., USA). The ligation reactions were transformed into the *Escherichia coli* strain TOP10F′ and the recombinants were selected by blue-white screening. Determination of the nucleotide sequences of the amplified products was performed by dideoxy cycle sequencing method with the corresponding sequencing primers for both strands (DAVIS Sequencing, Davis, Calif., USA).

Sequence Analysis

The nucleotide and deduced amino acid sequence similarities between the TCoV isolates and IBV strains were analyzed by the Clustal W method in the MegAlign module of the DNAstar program (Lasergene, Madison, Wisc., USA). Percent similarities were calculated to find nucleic acid and amino acid pair distances. Based on the obtained sequences of TCoV isolates and previously published sequences of IBV strains, phylogenetic trees were constructed using the neighbor-joining method [12].

Results

The primary structures of the 3′-end structural protein-encoding regions of TCoV isolates, containing the entire S protein gene, gene 3, M protein gene, gene 5, and N protein gene in the order from 5′ to 3′, were very similar to those found in the corresponding genomic areas of IBV strains. As shown in table 1, the size of most of the individual ORF was similar between TCoV isolates and IBV strains except the ORF for the S gene. The size of ORF for the S gene of TCoV isolates was longer than that of IBV strains. The S protein gene consisted of 3,612–3,624 nucleotides for TCoV isolates and 3,456–3,507 nucleotides for IBV strains. The ORF 3a for TCoV isolate 1440 was not found in two separate trials of cloning and sequencing of the same region. The ORF 3a for IBV strain KB8523 was also not found in the published sequence [5]. The canonical consensus transcription-regulating sequence (TRS) of IBV, CT(T/G) AACAA, was also found in TCoV. 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 isolates and IBV strains (table 2).

The pairwise comparison of nucleotide sequence distance for the entire 3′-end structural protein-encoding region between TCoV isolates and IBV strains is summarized in table 3. The similarity score among TCoV isolates and IBV strains ranged from 69.4 to 99.4%. The similarity scores among the TCoV isolates ranged from 92.7 to 99.4% and the similarity scores among the IBV strains ranged from 79.1 to 93.2%.

The pairwise comparison of nucleotide and deduced amino acid sequence distance of the S protein gene between the TCoV isolates and the IBV strains is summarized in table 4. The similarity score among TCoV isolates and IBV strains ranged from 49.1 to 99.7% at the nucleo-

Table 2. Comparison of TRS of genes in the 3′-end encoding regions between TCoV isolates and IBV strains

	Genes				
	spike	gene 3	membrane	gene 5	nucleocapsid
TCoV					
540	ctgaacaa (52)	atgaacaa (23)	cttaacaa (74)	cttaacaa (9)	cttaacaa (93)
ATCC	ctgaacaa (52)	ctgaacaa (23)	cttaacaa (77)	cttaacaa (9)	cttaacaa (93)
310	ctgaacaa (52)	ctgaacaa (23)	cttaacaa (77)	cttaacaa (9)	cttaacaa (93)
1440	ctgaacaa (52)	atgaacaa (187)[a]	cttaacaa (74)	cttacaacaa (9)	cttaacag (93)
IBV					
Beau	ctgaacaa (52)	ctgaacaa (23)	cttaacaa (77)	cttaacaa (9)	cttaacaa (93)
CU	ctgaacaa (52)	ctgaacaa (23)	cttaacaa (58)	cttaacaa (9)	cttaacaa (93)
KB	ctgaacaa (52)	ctgaacaa (198)[a]	cttaacaa (77)	cttaacaa (9)	cttaacaa (93)
DE	NA	ctgaacaa (23)	cttaacaa (77)	cttaacaa (9)	cttagcaa (93)
D1466	NA	ctgaacaa (23)	cttaacaa (77)	cttaacaa (9)	cttaacaa (93)

The numbers in parentheses indicate the distance calculated as nucleotides between 3′-end of TRS and the ATG start codon of the corresponding first downstream ORF. NA = Not applicable because the sequences were not available.

[a] The TCV isolate 1440 and IBV strain KB do not have ORF 3a and the first downstream ORF is 3b.

Table 3. Sequence pair distances for nucleic acid sequence of the entire 3′-end structural protein gene-encoding region between TCoV isolates and IBV strains

	Nucleotide identity, %								
	1	2	3	4	5	6	7	8	9
1 540	100	92.7	92.9	94.8	69.8	70.1	70.3	72.2	70.0
2 ATCC		100	99.4	93.4	70.0	70.6	71.1	71.5	69.4
3 310			100	93.6	70.1	70.8	71.3	71.6	69.5
4 1440				100	69.7	70.5	70.5	72.4	70.0
5 Beau					100	87.8	93.2	79.7	79.8
6 CU						100	88.7	80.0	79.1
7 KB							100	80.4	79.3
8 DE								100	85.3
9 D1466									100

Table 4. Sequence pair distances for nucleic acid and deduced amino acid sequence of the entire spike protein gene region between TCoV isolates and IBV strains

	Nucleotide identity, %								
	1	2	3	4	5	6	7	8	9
1 540	100	93.0	93.3	95.7	50.1	50.0	49.9	52.4	49.8
2 ATCC	92.5	100	99.7	93.7	50.0	50.4	50.1	52.3	50.0
3 310	93.0	99.3	100	94.0	50.0	50.5	50.1	52.2	49.9
4 1440	95.0	93.4	93.9	100	49.9	50.1	49.8	52.1	49.1
5 Beau	38.7	38.3	38.4	38.7	100	85.7	94.4	68.2	67.4
6 CU	38.8	38.9	39.0	39.0	83.5	100	86.0	68.2	68.3
7 KB	39.0	38.7	38.8	39.0	94.0	85.1	100	68.3	67.9
8 DE	41.5	41.4	41.2	41.3	64.5	65.0	65.3	100	77.0
9 D1466	38.9	39.0	38.8	39.0	64.3	64.2	65.3	76.4	100
	Amino acid identity, %								

Table 5. Sequence pair distances for nucleic acid and deduced amino acid sequence of the entire membrane protein gene region between TCoV isolates and IBV strains

		Nucleotide identity, %								
		1	2	3	4	5	6	7	8	9
1	540	100	93.4	93.4	93.9	91.3	88.7	91.5	90.7	91.9
2	ATCC	93.3	100	100.0	93.6	90.0	87.8	92.9	89.2	90.2
3	310	93.3	100.0	100	93.6	90.0	87.8	92.9	89.2	90.2
4	1440	95.5	95.1	95.1	100	92.4	89.9	90.9	91.7	92.6
5	Beau	93.8	93.3	93.3	95.1	100	89.0	94.1	96.8	95.7
6	CU	86.2	86.2	86.2	87.9	84.6	100	86.0	89.1	86.9
7	KB	92.4	94.2	94.2	95.1	95.6	85.0	100	94.6	95.0
8	DE	92.9	92.9	92.9	96.0	96.9	85.9	96.9	100	97.2
9	D1466	93.3	93.3	93.3	96.0	97.4	85.9	97.4	98.7	100
		Amino acid identity, %								

Table 6. Sequence pair distances for nucleic acid and deduced amino acid sequence of the entire nucleocapsid protein gene region between TCoV isolates and IBV strains

		Nucleotide identity, %								
		1	2	3	4	5	6	7	8	9
1	540	100	92.1	92.9	93.1	91.6	92.5	92.3	93.7	91.0
2	ATCC	88.2	100	97.5	93.2	91.4	91.9	92.8	93.2	90.4
3	310	92.2	93.1	100	94.1	92.5	93.0	93.9	94.3	91.6
4	1440	92.0	90.2	94.2	100	91.6	93.8	94.0	95.0	92.5
5	Beau	90.0	88.7	92.9	91.5	100	91.5	91.5	91.6	90.6
6	CU	91.0	88.7	93.2	93.7	91.7	100	93.0	94.9	95.9
7	KB	92.5	90.7	95.4	93.9	93.4	94.6	100	93.9	91.9
8	DE	93.2	90.4	95.1	94.2	92.2	95.1	95.6	100	92.9
9	D1466	90.3	88.0	92.7	92.9	91.5	95.9	94.2	94.4	100
		Amino acid identity, %								

tide level or from 38.3 to 99.3% at the amino acid level. The similarity scores among the TCoV isolates within the S protein gene region ranged from 93.0 to 99.7% at the nucleotide level or from 92.5 to 99.3% at the amino acid level. The similarity scores among the examined IBV strains within the S protein gene region ranged from 67.4 to 94.4% at the nucleotide level or from 64.2 to 94.0% at the amino acid level.

The pairwise comparison of nucleotide and deduced amino acid sequence distance of M protein gene between the TCoV isolates and the IBV strains is summarized in table 5. The similarity score among TCoV isolates and IBV strains ranged from 86.0 to 100.0% at the nucleotide level or from 84.6 to 100.0% at the amino acid level. The similarity scores among the TCoV isolates within the M protein gene region ranged from 93.4 to 100.0% at the nucleotide level or from 93.3 to 100.0% at the amino acid level. The similarity scores among the examined IBV strains within the M protein gene region ranged from 86.0 to 97.2% at the nucleotide level or from 84.6 to 98.7% at the amino acid level.

The pairwise comparison of nucleotide and deduced amino acid sequence distance of N protein gene between the TCoV isolates and the IBV strains is summarized in table 6. The similarity score among TCoV isolates and IBV strains ranged from 90.4 to 97.5% at the nucleotide level or from 88.0 to 95.9% at the amino acid level. The similarity scores among the TCoV isolates within the N protein gene region ranged from 92.1 to 97.5% at the nucleotide level or from 88.2 to 94.2% at the amino acid level. The similarity scores among the examined IBV strains within the N protein gene region ranged from 90.6 to 95.9% at the nucleotide level or from 91.5 to 95.9% at the amino acid level.

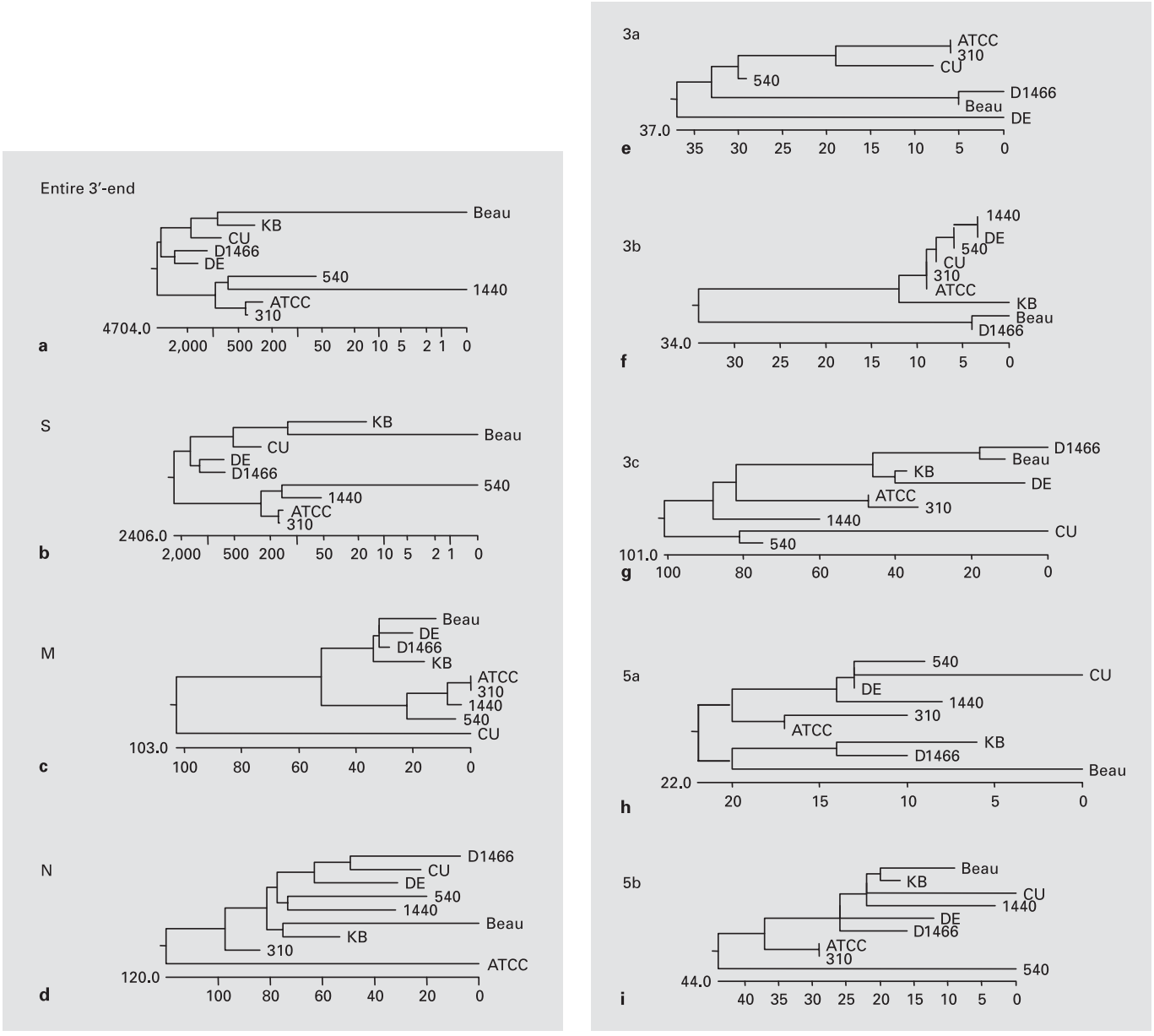

Fig. 1. Phylogenetic analyses within the entire 3′-end structural protein-encoding region (**a**) and the deduced amino acid sequence of spike (S) protein gene (**b**), membrane (M) protein gene (**c**), nucleocapsid (N) protein gene (**d**), ORF 3a (**e**), 3b (**f**), 3c (**g**), 5a (**h**), and 5b (**i**) demonstrating relationships among TCoV isolates and IBV strains. The similarity scores were calculated by the Lasergene program using the unweighted pair group method with arithmetic averages. The horizontal distances are proportional to the number of nucleotide differences. The GenBank accession numbers for each coronavirus were described in the Materials and Methods.

The similarity score among TCoV isolates and IBV strains within the ORF 3a region ranged from 82.3 to 100.0% at the nucleotide level or from 76.3 to 100.0% at the amino acid level. The similarity score among TCoV isolates and IBV strains within the ORF 3b region ranged from 83.7 to 99.5% at the nucleotide level or from 72.7 to 100.0% at the amino acid level. The similarity score among TCoV isolates and IBV strains within the ORF 3c region ranged from 86.6 to 99.4% at the nucleotide level or from 77.9 to 98.1% at the amino acid level.

The similarity score among TCoV isolates and IBV strains within the ORF 5a region ranged from 89.9 to 99.5% at the nucleotide level or from 88.1 to 98.5% at the amino acid level. The similarity score among TCoV isolates and IBV strains within the ORF 5b region ranged from 91.8 to 100.0% at the nucleotide level or from 81.7 to 100.0% at the amino acid level.

Phylogenetic analysis according to the entire 3′-end structural protein-encoding region (fig. 1a) or the deduced amino acid sequence of S protein gene (fig. 1b) indicated that TCoV isolates were clustered within the same genomic lineage while all the IBV strains were grouped into another separate cluster. However, the phylogenetic trees for the deduced amino acid sequences of M protein gene, N protein gene, or ORF 3a, 3b, 3c, 5a, and 5b (fig. 1c–i, respectively) had a very different topology. The TCoV isolates randomly formed groups with the IBV strains in the phylogenetic trees for different genes. In M protein gene, no TCoV isolate or IBV strain clustered with IBV strain CU (fig. 1c). In contrast, no TCoV isolate or IBV strain clustered with TCoV isolate ATCC in the phylogenetic tree for N protein gene (fig. 1d).

Discussion

Since its initial recognition in the early 1970s, the relationship of TCoV to other coronaviruses is still under debate. It was speculated that the discrepant results regarding the relationship between TCoV and other coronaviruses might be caused by different isolates of TCoV from various geographical areas at different times. In order to test this possibility, the antigenicity of 18 isolates of TCoV from various geographical areas at different times including the prototype isolate TCoV-ATCC was examined in one of our previous studies [13]. It was found that all the different isolates of TCoV had the same pattern of antigenic reactivity with the same panel of antibodies and all of them cross-reacted with polyclonal antibody to IBV or a monoclonal antibody to M protein of IBV. In contrast, none of the examined TCoV isolates reacted with a monoclonal antibody to S protein of IBV. In line with these previous results of antigenic study, the sequence analysis in the present study demonstrated that isolates of TCoV from various geographical areas at different times including the prototype isolate TCoV-ATCC had a high similarity in the 3′-end structural protein-encoding region. In addition, the M protein region of TCoV isolates shared a high similarity (86.2–96.0%) with that of IBV strains, while the S protein region of TCoV isolates had a relatively low similarity (38.3–41.5%) with that of IBV strains.

In spite of the similar primary genomic structure to the S gene, tricistronic gene 3, M gene, bicistronic gene 5, and N gene from 5′ to 3′, the sequence similarity of the entire 3′-end structural protein-encoding region between TCoV and IBV was low (69.4–72.4%) in comparison with that among TCoV isolates (92.7–99.4%) or that among IBV strains (79.1–93.2%). The differences were mainly contributed by the S protein gene. The size of TCoV S gene (3,612–3,624 nucleotides) was larger than that of IBV S gene (3,456–3,507 nucleotides). The sequence similarity of the S gene nucleotide sequences between TCoV and IBV was lower (49.1–52.4%) than that among TCoV isolates (93.0–99.7%) or that among IBV strains (67.4–94.4%). In other genes, the sequence differences between TCoV and IBV were comparable with those among TCoV isolates or among IBV strains. The phylogenetic tree based on the S protein region was the same as that based on the entire 3′-end structural protein-encoding region with TCoV isolates and IBV strains grouped in two separate clusters. These results suggested that TCoV probably shared the same origin with that of IBV, evolved independently in a separate environment a long time ago, and acquired sequences of S gene for turkey intestine tropism. Accumulation of S gene sequence of more TCoV isolates will further clarify the relationship between TCoV and IBV and may lead to the understanding of origin and evolution of TCoV.

The replication of coronavirus is characterized by the synthesis of a 3′-coterminal nested set of polycistronic subgenomic mRNAs with a discontinuous transcription mechanism. Only the first ORF from 5′-end of each polycistronic subgenomic mRNA is translated. There is a common 5′-terminal leader sequence derived from the 5′-end of the coronavirus genome on the body sequence of each subgenomic mRNA. The acquisition of the leader sequence to each subgenomic mRNAs involves the conserved TRS sequence present along the genomic RNA and 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 [14, 15]. 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 an important role in the transcription of mRNAs. As shown in the present

study, the TRS sequences and the distance between the TRS 3′-end and the initiation codon of the first ORF of TCoV isolates were highly conserved with those of the corresponding genes of IBV strains. The TRS of gene 5 for the IBV strain Beaudette has 2 consensus sequences, CTTAACAA, in a tandem repeat (CTTAACAAAAACTTAACAA). According to analysis with reported genes, it was demonstrated that only one of this TRS is required for the expression of gene 5 and either sequence can function as an acceptor site for acquisition of the leader sequence [15]. In contrast to this result, no similar structure of 2 TRS sequences in a tandem repeat for gene 5 was found in any other IBV strains or TCoV isolates. However, it is interesting to note that there are deletions of 8 nucleotides in the corresponding region of TCoV isolate 1440. These deletions resulted in a TRS sequence of CTTACAACAA for gene 5 of TCoV isolate 1440.

The predicted proteins of ORF 3a, 3b, 3c, 5a, and 5b were small (about or less than 10 kDa). The protein encoded by ORF 3c was shown to be associated with viral envelope [16], while the functions of all the other gene products are not clear. In general, ORFs of gene 5 are more conserved than those of gene 3 among TCoV isolates and IBV strains. The ORF 3a was not found in the genome of TCoV isolate 1440 in two different trials of cloning and sequencing analysis in the present study. In comparison, the IBV strain KB8523 does not have ORF 3a either. These results suggested that the product of ORF 3a is not important for pathogenesis and replication of TCoV and IBV.

The highly conserved TRS sequences of IBV have been shown to be a recombination 'hot spot' and may serve as the template switching sites for the viral encoded RNA-dependent RNA polymerase [7]. These recombination events play an important role in the emergence of new IBV variants responsible for continuous outbreaks in the chicken flocks vaccinated with live attenuated viruses due to a failure of cross-protection. For example, the DE072 strain is a recombinant with a D1466-like sequence in the S gene, while the CU-T2 strain is a recombinant between Arkansas and Massachusetts strains [7, 8]. It is possible that the similar recombination events of IBV in chicken may contribute to the origin and evolution of TCoV in turkey. Except for the S protein gene, the sequences of other individual genes among TCoV isolates and IBV strain had different degrees of similarity. The phylogenetic tree based on these individual genes had a very different topology. These results implied that a series of complex events including recombination occurred between TCoV and IBV in the process of their evolution.

Recently, two small fragments of sequences for a TCoV UK isolate were reported [17]. Sequence comparisons of these sequences with the corresponding regions of the TCoV isolates from the US as examined in the present study showed a high similarity. However, the similarity scores between the UK isolate and the US isolates were apparently lower than that among the US isolates. For the 844-bp fragment of the UK isolate spanning from the 3′-end of S gene to the 5′-end of M gene, the similarity scores between the UK isolate and the US isolates were 90.8–91.8%, while the similarity scores among the US isolates were 92.6–99.8%. For the other 866-bp fragment of UK isolate spanning from gene 5 to the 5′-end of N gene, the similarity scores between the UK isolate and the US isolates were 86.9–89.3% while the similarity scores among the US isolates were 91.1–99.9%. These results indicate that the UK isolate and the US isolates of TCoV had the same origin and evolved independently in separate geographical areas, which resulted in a genetic drift.

Acknowledgements

The authors thank Pfizer Animal Health, North Carolina Turkey Spiking Mortality Task Force, and the US Department of Agriculture for providing financial support. The authors also thank Tom Hooper and Donna Schrader for virus isolation and Jing Dai for PCR amplification.

References

1 Nagaraja KV, Pomeroy BS: Coronaviral enteritis of turkeys (bluecomb disease); in Calnek B, Barnes HJ, Beard CW, McDougald LR, Saif YM (eds): Diseases of Poultry, ed 10. Ames, Iowa State University Press, 1997, pp 686–692.

2 Lee CW, Jackwood MW: Origin and evolution of Georgia 98 (GA98), a new serotype of avian infectious bronchitis virus. Virus Res 2001;80:33–39.

3 Lin TL, Loa CC, Wu CC: Complete sequence of 3′ end encoding regions reveals close genomic relationship between turkey coronavirus and avian infectious bronchitis virus. Virus Res 2004;106:61–70.

4 Boursnell ME, Brown TD, Foulds IJ, Green PF, Tomley FM, Binns MM: Completion of the sequence of the genome of the coronavirus avian infectious bronchitis virus. J Gen Virol 1987;68:57–77.

5 Sutou S, Sato S, Okabe T, Nakai M, Sasaki N: Cloning and sequencing of genes encoding structural proteins of avian infectious bronchitis virus. Virology 1988;165:589–595.

6 Jia W, Naqi SA: Sequence analysis of gene 3, gene 4 and gene 5 of avian infectious bronchitis virus strain CU-T2. Gene 1997;189:189–193.

7 Lee CW, Jackwood MW: Evidence of genetic diversity generated by recombination among avian coronavirus IBV. Arch Virol 2000;145:2135–2148.

8 Lee CW, Jackwood MW: Spike gene analysis of the DE072 strain of infectious bronchitis virus: origin and evolution. Virus Genes 2001;22:85–91.

9 Loa CC, Lin TL, Wu CC, Bryan TA, Thacker HL, Hooper T, Schrader D: Detection of antibody to turkey coronavirus by antibody-capture enzyme-linked immunosorbent assay utilizing infectious bronchitis virus antigen. Avian Dis 2000;44:498–506.

10 Akin A, Wu CC, Lin TL: Amplification and cloning of complete infectious bursal disease virus genomic RNA segments by a long and accurate PCR. J Virol Methods 1999;82:55–61.

11 Chomczynski P, Sacchi N: Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal Biochem 1987;162:156–159.

12 Saitou N, Nei M: The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 1987;4:406–425.

13 Lin TL, Loa CC, Wu CC, Bryan TA, Hooper T, Schrader D: Antigenic relationship of turkey coronavirus isolates from different geographic locations in the United States. Avian Dis 2002;46:466–472.

14 Spaan W, Cavanagh D, Horzinek MC: Coronaviruses: structure and genome expression. J Gen Virol 1988;69:2939–2952.

15 Stirrups K, Shaw K, Evans S, Dalton K, Casais R, Cavanagh D, Britton P: Expression of reporter genes from the defective RNA CD-61 of the coronavirus infectious bronchitis virus. J Gen Virol 2000;81:1687–1698.

16 Liu DX, Inglis SC: Association of the infectious bronchitis virus 3c protein with the virion envelope. Virology 1991;185:911–917.

17 Cavanagh D, Mawditt K, Sharma M, Drury SE, Ainsworth HL, Britton P, Gough RE: Detection of a coronavirus from turkey poults in Europe genetically related to infectious bronchitis virus of chickens. Avian Pathol 2001;30:355–368.