Sequence Analysis of the Matrix/Nucleocapsid Gene Region of Turkey Coronavirus

Jamie J. Breslin  Lynda G. Smith  Frederick J. Fuller  James S. Guy

Department of Microbiology, Pathology, and Parasitology, North Carolina State University, Raleigh, N.C., USA

Key Words
Coronavirus  Infectious bronchitis virus  Turkey coronavirus  Polymerase chain reaction

Summary
A reverse transcriptase, polymerase chain reaction (RT-PCR) procedure was used to amplify a segment of the genome of turkey coronavirus (TCV) spanning portions of the matrix and nucleocapsid (MN) protein genes (approximately 1.1 kb). The MN gene region of three epidemiologically distinct TCV strains (Minnesota, NC95, Indiana) was amplified, cloned into pUC19, and sequenced. TCV MN gene sequences were compared with published sequences of other avian and mammalian coronaviruses. A high degree of similarity (>90%) was observed between the nucleotide, matrix protein, and nucleocapsid protein sequences of TCV strains and published sequences of infectious bronchitis virus (IBV). The matrix and nucleocapsid protein sequences of TCV had limited homology (<30%) with MN sequences of mammalian coronaviruses. These results demonstrate a close genetic relationship between the avian coronaviruses, IBV and TCV.

Introduction
The Coronaviridae are a large group of RNA-containing viruses that infect a wide variety of avian and mammalian species [1, 2]. The family is comprised of two genera, Coronavirus and Torovirus, that share similarities in morphology, genome organization and genome expression [3]. The coronavirus genome consists of a positive sense, single-stranded RNA molecule that is 20–30 kb in size [3, 4]. Virions are enveloped, pleomorphic, 80–220 nm in diameter, and have club-shaped surface projections approximately 20 nm in length [3, 4]. Three major structural proteins are known, the surface (S) glycoprotein (90–180 kD), an intermembrane matrix (M) protein (20–35 kD), and a nucleocapsid (N) protein (50–60 kD) [3, 4].

Coronaviruses initially were subdivided into four antigenic groups based on differences determined primarily by immunofluorescent (FA), enzyme-linked immunosorbent assay (ELISA), and immunoelectron microscopic (IEM) studies [1, 2, 5, 6]. Mammalian coronaviruses were shown to comprise antigenic groups 1 and 2, whereas avian coronaviruses, infectious bronchitis virus (IBV) and turkey coronavirus (TCV), comprised groups 3 and 4, respectively [1, 2, 5, 6]. IBV and TCV were determined to be antigenically distinct from each other and mammalian coronaviruses based on IEM, hemagglutination inhibition (HI) and virus-neutralization (VN) studies [7, 8]; how-
ever, subsequent antigenic and genomic analyses have questioned these taxonomic groupings. Studies done by Dea et al. [9] and later by Verbeek and Tijssen [10] indicated that TCV and bovine coronavirus (BCV), a mammalian group 2 coronavirus, were closely related based on antigenic and nucleotide sequence analyses. In contrast to the findings of these studies, antigenic analyses by Guy et al. [11] indicated that the avian coronaviruses, IBV and TCV, are closely related and comprise a single antigenic group within the coronavirus genus [11].

TCV is the cause of an acute highly contagious enteric disease of turkeys referred to as bluecomb disease [12]. Bluecomb disease was first identified in turkeys in 1951 and a coronavirus was identified as the cause of the disease in 1973 [12]; however, antigenic and molecular characterization of TCV has been hampered by difficulties associated with in vitro culture of the virus. Minnesota strain, a reference TCV strain, and field isolates have been successfully propagated in embryonated turkey and embryonated chicken eggs by inoculation of the amniotic cavity [12]. Recently, Dea et al. [7] reported the adaptation and serial propagation of TCV in a human rectal adenocarcinoma (HRT) cell line.

The present study was undertaken to examine relationships between TCV and other avian and mammalian coronaviruses based on nucleotide sequence analysis of TCV RNA. A well-characterized region of the coronavirus genome [13–15] comprised of nucleotides from both the matrix and nucleocapsid (MN) genes was used as a basis for comparison of three epidemiologically distinct TCV strains and selected coronaviruses.

**Materials and Methods**

**Viruses**

TCV (Minnesota) was obtained from the American Type Culture Collection (Rockville, Md., USA). TCV (NC95) was isolated from turkeys in North Carolina [11]. TCV (Indiana) was isolated from turkeys in Indiana and obtained from Tom Hooper, Purdue University, Dubois, Ind., USA. TCV strains were propagated by amniotic inoculation of embryonated turkey eggs [11]. IBV (Massachusetts) was obtained from SPAFAS, Inc. (Norwich, Conn., USA) and propagated in 9- to 11-day-old embryonated chicken eggs [16].

**Virus Purification**

Template RNA for reverse transcriptase, polymerase chain reaction (RT-PCR) was obtained from partially purified preparations of TCV as follows. TCV-infected turkey embryo intestines were prepared as 10% (w/v) suspensions in TNE buffer (0.01 M Tris-hydrochloride (pH 7.4), 0.1 M NaCl, 1 mM EDTA), homogenized, and clarified by centrifugation at 8,000 g for 20 min. Polyethylene glycol (Sigma) was added to the supernatant fluid to a final concentration of 10%, then incubated overnight at 4 °C with gentle stirring. The precipitate was collected by centrifugation at 8,000 g for 20 min. The pellet was resuspended in 1/10 of the original volume in TNE buffer pH 7.4, layered onto a 20–60% (w/w) sucrose step gradient, and centrifuged at 80,000 g for 1 h; the pellet was resuspended in RNase free water. IBV was similarly purified from virus-laden allantoic fluids.

Template RNA also was obtained from TCV (NC95) which was purified by banding to density in sucrose gradients. TCV (NC95) was propagated and partially purified as described above. Material at the interface of the 20–60% (w/w) step gradient was resuspended in an equal volume TNE buffer pH 7.4 and layered onto a 20% to 60% (w/w) continuous sucrose gradient and centrifuged at 80,000 g overnight. The gradient was fractionated in 1-ml aliquots by upward displacement with a density gradient fractionator. Fractions were diluted in TNE buffer pH 7.4 and centrifuged at 80,000 g for 1.5 h; pellets were resuspended in TNE buffer pH 7.4.

**Nucleic Acid Purification**

Nucleic acid was harvested from purified virus by incubation in 0.5% SDS (Gibco BRL) for 5 min at room temperature followed by a phenol-chloroform extraction was described [17]. Pellets were resuspended in RNase free water.

**RT-PCR**

Viral cDNA was prepared using a commercially available RT reaction kit (Promega). The RT reaction consisted of RT buffer [100 mM Tris-HCl (pH 8.8), 500 mM KCl, 1% Triton® X-100], 5 mM MgCl₂, 1 mM each deoxynucleoside triphosphate (dNTP), 20 units rRNasin, 12.5 units AMV RT, 0.5 µg random primers, 1 µg template RNA, and nuclease free water to 20 µl. The reaction was completed in a programmable thermal cycler at 42 °C for 30 min, 99 °C for 5 min and 4 °C for 5 min per the directions in the kit.

PCR was conducted in the same tube on the resultant cDNA in a final volume of 100 µl. The PCR consisted of 2.5 units PWO DNA polymerase (Boehringer Mannheim), PCR buffer [100 mM Tris-HCl (pH 8.85), 250 mM KCl, 50 mM (NH₄)₂SO₄, 2.5 mM MgSO₄, 100 ng of each custom primer EcoM and EcoN (Gibco BRL), and nuclease free water to volume. Samples were placed in a hot (94 °C) thermal cycler and amplified as previously described [14] with the annealing temperature at 52 °C for 2 min. PCR products were analyzed on a 1% Ultra Pure Agarose (Gibco BRL) gel with 0.1 µg/ml ethidium bromide (Sigma).

EcoM and EcoN primers were synthesized on previously described primer sequences (MIBV, NIBV) for PCR amplification of IBV RNA [13, 14]. EcoM and EcoN primers were identical to MIBV and NIBV primers with the exception that restriction sites in these primers were replaced with EcoRI restriction sites. EcoM recognizes the 3’ end of the matrix protein gene and EcoN recognizes the 5’ end of the nucleocapsid protein gene. Primer sequences are as follows:

EcoM 5’-TGAATTTCTAGTGTTGTCAATTTGGACCT-3’
EcoN 5’-TGAATTCTACGGCTACCTTAAAATTTGCGG-3’

**Cloning**

RT-PCR products and cloning vector pUC19 were digested with restriction endonuclease EcoRI (Gibco BRL) according to manufacturer’s recommendations. RT-PCR products were inserted into pUC19 and transformed into a competent Escherichia coli strain DH5α (Gibco BRL) as described [17].

---

Sequence Analysis of Turkey Coronavirus

Intervirology 1999;42:22–29
Fig. 1. Agarose gel electrophoresis of RT-PCR products obtained from IBV (Massachusetts) and three TCV strains. Lane 1 and 6 are molecular-weight markers λDNA/HindIII and φX174/HaeIII respectively; numbers on the vertical axis indicate size in kilobase pairs of molecular-weight markers. Lane 2, IBV (Massachusetts); lane 3, TCV (NC95); lane 4, TCV (Minnesota); lane 5, TCV (Indiana).

Sequencing/Phylogenetic Tree Analysis

DNA was sequenced at the University of North Carolina (Chapel Hill) Automated DNA Sequencing Facility, on a Model 373A DNA Sequencer (Applied Biosystems) using the Taq DyeDeoxy™ Terminator Cycle Sequencing Kit (Applied Biosystems). All sequences were confirmed by sequencing both strands. GenBank accession numbers are: TCV (Minnesota) AF072911; TCV (NC95) AF072912; TCV (Indiana) AF072913.

Comparative analyses of nucleotide and protein sequences were performed using the MegAlign application of the Lasergene software package (DNASTAR, Madison, Wisc., USA). Phylogenetic tree construction was based on the neighbor-joining method of Thompson et al. [18] using an unrooted tree analysis in the program CLUSTAL X with 1,000 bootstrap trials. Bootstrap values ≥950 are considered significant at the 95% confidence level (p ≤ 0.05). Protein comparison consisted of 70 amino acids at the carboxy-terminus of the matrix gene and 55 amino acids at the amino-terminus of the nucleocapsid gene. Coronavirus sequences were obtained from the GenBank database. These included IBV (Beaudette, Gray, KB8523 strains), BCV (Mebus, F15 strains), mouse hepatitis virus (MHV) (A59, JHM strains), transmissible gastroenteritis virus (TGEV) (Purdu, FS772 strains), porcine respiratory coronavirus (PRCV) (86-137004, RM4 strains), canine coronavirus (CCV) (Insavc-1 strain), human coronavirus (229E, OC43 strains), feline infectious peritonitis virus (79-1146), and feline enteric coronavirus (79-1683) [15, 19–34]. The sequence reported by Verbeek and Tijsen for TCV, and herein referred to as TCV (Verbeek), was also included in the study [10].

Results

RT-PCR

Preliminary studies indicated that RNA obtained from turkey embryo-propagated TCV (NC95) could be amplified in an RT-PCR using IBV-specific synthetic primers
Fig. 2. Comparison of MN gene nucleotide sequence of infectious bronchitis virus (IBV) Beaudette strain [20] and TCV strains NC95, Minnesota, and Indiana. Nucleotide sequence differences are shown for TCV strains. The positions where nucleotide bases are missing are indicated as (–) and similar nucleotides as (.).

MIBV and NIBV (data not shown). These primers were previously described by Andreason et al. [13] for RT-PCR amplification of the MN gene region of IBV RNA. Subsequent RT-PCR of TCV strains utilized synthetic DNA primers EcoM and EcoN; these primers were identical to MIBV and NIBV with the exception that EcoRI restriction sites were substituted for restriction sites in these primers.

RNA obtained from partially purified TCV (NC95), TCV (Indiana) and TCV (Minnesota) was used as template in RT-PCR using synthetic primers EcoM and EcoN. The PCR products were approximately 1.1 kb in size and slightly larger than the product obtained using IBV RNA as template (fig. 1). IBV RNA was used as a positive control in the RT-PCR procedure and produced a product of approximately 1.0 kb, consistent with that described by Andreason et al. [13]. No PCR product was observed when RNA was harvested from uninfected embryos, and amplified by RT-PCR, or when RT-PCR was run without RT (data not shown).

RT-PCR Analysis of Sucrose-Gradient-Purified TCV

As an additional means of validating the RT-PCR, template RNA was obtained from sucrose-gradient-purified TCV (NC95). RT-PCR was performed on RNA extracted from individual gradient fractions after isopycnic centrifugation of TCV (NC95). Positive PCR amplification was evident in fractions with densities of 1.14–1.20 g/cm³, consistent with the density of coronaviruses [3]. Positive PCR amplification was seen in only one other fraction, at a density of 1.113 g/cm³.
Fig. 3. Comparison of matrix protein (A) and nucleocapsid protein amino acid sequences (B) of IBV Beaudette strain [20] and TCV strains NC95, Minnesota, and Indiana. Amino acid sequence differences are shown for TCV strains. The positions where amino acids are similar are indicated as (.).

Fig. 4. Phylogenetic relationship of avian and mammalian coronaviruses based on comparison of matrix protein (A) and nucleocapsid protein sequences (B) in the 1.1-kb MN gene region. Amino acid sequences were aligned using the CLUSTAL X method [18] and unrooted phylogenetic trees were constructed using the neighbor-joining method with 1,000 bootstrap trials. The bootstrap values are indicated at the 17 internal branch points.

Sequencing/Phylogenetic Trees

EcoRI digestion of the TCV RT-PCR product, in preparation for cloning, revealed an interior EcoRI site in both TCV (NC95) and TCV (Indiana), but not in TCV (Minnesota). Therefore, TCV (NC95) and TCV (Indiana) were cloned and sequenced in two pieces, one approximately 400 bases and the other approximately 700 bases. TCV (Minnesota) was cloned and sequenced as one 1.1 kb piece. Figure 2 compares the MN gene nucleotide sequence (1.1 kb) of three strains of TCV with the published sequence of IBV (Beaudette) [20]. Figure 3 compares the matrix and nucleocapsid amino acid sequence of three strains of TCV with the published sequence of IBV (Beaudette) [20].

MN nucleotide sequences of TCV strains are compared in table 1 with published sequences of IBV strains (Beaudette, KB8523) and two representative mammalian coronaviruses, BCV (Mebus) and TGEV (Purdue); results are shown as the percentage of sequence homology between viruses. Nucleotide sequence homology among the three TCV ranged from 87.2 to 89.6%. Nucleotide sequence homology between TCV strains and IBV strains ranged from 85.8 to 92.6%. Homology between TCV strains and mammalian coronaviruses was less than 30%.

Comparisons between matrix and nucleocapsid protein sequences are shown in table 2. Matrix and nucleocapsid protein homologies between TCV and IBV ranged from 95.6 to 100% and 90.9 to 98.2%, respectively. In contrast, homology between TCV matrix and nucleocapsid protein sequences and those of mammalian coronaviruses was less than 30%. Similarly, homology between matrix and nucleocapsid proteins of TCV strains examined in this study (Minnesota, NC95, Indiana) and TCV (Verbeek) was less than 30%. These findings demonstrate that TCV (Minnesota, NC95, Indiana) and IBV are highly...
Table 1. Percent nucleotide sequence homology in MN gene region between TCV (Minnesota, NC95, Indiana), IBV (Beaudette, KB8523) [20, 21], and representative mammalian coronaviruses, TGEV (Purdue) [27, 28] and BCV (Mebus) [15]

<table>
<thead>
<tr>
<th></th>
<th>IBV(Beaudette)</th>
<th>IBV(KB8523)</th>
<th>TCV(NC95)</th>
<th>TCV(Minnesota)</th>
<th>TCV(Indiana)</th>
<th>BCV(Mebus)</th>
</tr>
</thead>
<tbody>
<tr>
<td>IBV(KB8523)</td>
<td>90.3</td>
<td>87.0</td>
<td>90.4</td>
<td>88.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TCV(NC95)</td>
<td>85.8</td>
<td>88.4</td>
<td>89.6</td>
<td>87.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TCV(Indiana)</td>
<td>85.8</td>
<td>88.4</td>
<td>89.6</td>
<td>87.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BCV(Mebus)</td>
<td>21.6</td>
<td>21.5</td>
<td>21.0</td>
<td>21.8</td>
<td>19.6</td>
<td></td>
</tr>
<tr>
<td>TGEV(Purdue)</td>
<td>21.1</td>
<td>22.7</td>
<td>20.9</td>
<td>20.5</td>
<td>21.2</td>
<td>27.5</td>
</tr>
</tbody>
</table>

Nucleotide sequence information was obtained from the GenBank database. Comparative analyses of nucleotide sequences were performed using the MegAlign application of the Lasergene software package (DNASTAR, Madison, Wisc.).

Table 2. Amino acid sequence comparison of TCV matrix and nucleocapsid proteins with corresponding regions of other avian and mammalian coronaviruses

<table>
<thead>
<tr>
<th></th>
<th>IBV (Beaudette)</th>
<th>IBV(KB8523)</th>
<th>TCV(NC95)</th>
<th>TCV(Minnesota)</th>
<th>TCV(Indiana)</th>
<th>BCV(Mebus)</th>
<th>MHV(A-59)</th>
<th>CCV(Insavc-1)</th>
<th>TGEV (Purdue)</th>
</tr>
</thead>
<tbody>
<tr>
<td>IBV (Beaudette)</td>
<td>93.4</td>
<td>90.9</td>
<td>96.4</td>
<td>90.9</td>
<td>20.3</td>
<td>20.0</td>
<td>19.7</td>
<td>20.4</td>
<td></td>
</tr>
<tr>
<td>IBV (KB8523)</td>
<td>95.6</td>
<td>92.7</td>
<td>98.2</td>
<td>92.7</td>
<td>20.3</td>
<td>18.3</td>
<td>19.7</td>
<td>20.2</td>
<td></td>
</tr>
<tr>
<td>TCV (NC95)</td>
<td>100.0</td>
<td>98.6</td>
<td>94.5</td>
<td>92.7</td>
<td>20.0</td>
<td>20.0</td>
<td>18.2</td>
<td>20.0</td>
<td></td>
</tr>
<tr>
<td>TCV (Minnesota)</td>
<td>98.6</td>
<td>97.1</td>
<td>98.6</td>
<td>94.5</td>
<td>20.0</td>
<td>20.0</td>
<td>18.2</td>
<td>20.0</td>
<td></td>
</tr>
<tr>
<td>TCV (Indiana)</td>
<td>98.6</td>
<td>97.1</td>
<td>98.6</td>
<td>97.1</td>
<td>20.0</td>
<td>20.0</td>
<td>18.2</td>
<td>21.8</td>
<td></td>
</tr>
<tr>
<td>TCV (Verbeek)</td>
<td>28.4</td>
<td>27.6</td>
<td>19.6</td>
<td>17.6</td>
<td>19.6</td>
<td>99.6</td>
<td>67.4</td>
<td>17.1</td>
<td>19.9</td>
</tr>
<tr>
<td>BCV (Mebus)</td>
<td>28.4</td>
<td>28.0</td>
<td>19.6</td>
<td>17.6</td>
<td>19.6</td>
<td>99.6</td>
<td>67.2</td>
<td>17.1</td>
<td>19.9</td>
</tr>
<tr>
<td>MHV (A-59)</td>
<td>28.9</td>
<td>28.4</td>
<td>17.6</td>
<td>17.6</td>
<td>17.6</td>
<td>83.8</td>
<td>84.2</td>
<td>21.5</td>
<td>23.0</td>
</tr>
<tr>
<td>CCV (Insavc-1)</td>
<td>19.6</td>
<td>19.1</td>
<td>23.5</td>
<td>23.5</td>
<td>23.5</td>
<td>34.3</td>
<td>34.8</td>
<td>32.0</td>
<td>89.0</td>
</tr>
<tr>
<td>TGEV (Purdue)</td>
<td>20.4</td>
<td>19.6</td>
<td>23.5</td>
<td>23.5</td>
<td>36.5</td>
<td>37.0</td>
<td>35.1</td>
<td>86.3</td>
<td></td>
</tr>
</tbody>
</table>

Matrix percent similarity.

related (>85% at the nucleotide level; >90% at the level of the matrix and nucleocapsid proteins).

Phylogenetic trees were prepared to examine relationships between TCV and other coronaviruses based on comparison of amino acid sequences within the matrix and nucleocapsid proteins (fig. 4). Phylogenetic trees were prepared based on matrix and nucleocapsid protein sequences for TCV strains and published sequence data for selected avian and mammalian coronaviruses. These phylogenetic trees divided the coronaviruses into three genotypic groups with the avian coronaviruses, IBV and TCV, representing one distinct group.
Discussion

Previous studies in our laboratory indicated a close antigenic relationship between IBV and TCV [11]. TCV (NC95) and TCV (Minnesota) were shown to be closely related to IBV based on FA procedures using polyclonal and monoclonal antibodies [11]. These studies suggested that these avian coronaviruses comprise a single antigenic group within the Coronavirus genus. The present study provides additional support for those findings. Sequencing data demonstrated a close genetic relationship between IBV and TCV, and phylogenetic analyses divided the avian and mammalian coronaviruses into three distinct genotypic groups with the avian coronaviruses, IBV and TCV, comprising a distinct genotype.

Synthetic primers that previously were described for RT-PCR amplification of IBV RNA within the MN gene region were shown in this study to amplify TCV RNA [13]. The RT-PCR product was very similar in size (1,100 bp) to the expected product; RT-PCR amplification of IBV RNA resulted in a 1,000-bp DNA product [13]. Subsequent nucleotide sequence analyses of the resultant DNA products demonstrated that the sequence of the MN gene region of TCV was closely related to IBV, and only distantly related to mammalian coronaviruses. A high degree of similarity (>90%) was observed between the matrix protein and nucleocapsid protein sequences of TCV strains and sequences of IBV strains. In contrast, a significantly lower degree of homology (<30) was evident between matrix and nucleocapsid protein sequences of TCV and mammalian coronaviruses.

Several control procedures were used to validate the RT-PCR procedure for amplification of TCV RNA. No RT-PCR product was obtained when RT-PCR was performed with material obtained from uninfected embryo intestines, and no product was obtained when the RT-PCR was performed without RT. The source of template RNA was examined by isopycnic centrifugation of TCV (NC95) in a sucrose gradient, followed by gradient fractionation, nucleic acid extraction, and RT-PCR amplification of individual gradient fractions. Gradient fractions which yielded positive RT-PCR products had densities of 1.113 g/cm³ and 1.14 to 1.20 g/cm³. Coronaviruses have a buoyant density in sucrose of approximately 1.15–1.19 g/cm³ [3]. Thus, the material used for RNA extraction and RT-PCR banded at densities consistent with that of coronaviruses. The nature of the material that banded at a density of 1.113 g/cm³ and produced a positive RT-PCR product has not been determined; however, this material likely resulted from virion degrada-

dation and release of viral RNA during sedimentation in sucrose.

Studies conducted by Dea et al. [9] and Verbeek and Tijsen [10] indicated a close antigenic and genomic relationship between TCV and BCV; however, our studies contradict those findings. In the studies by Dea et al. [9], TCV was shown to replicate in HRT cells and to be closely related to BCV based on immunoblotting studies, HI, VN, and hybridization of BCV cDNA to TCV RNA. Verbeek and Tijsen [10] also used HRT-propagated TCV strains and their studies demonstrated nucleotide sequence homology of 99% between TCV and BCV in the MN gene region. In contrast to the findings of Dea et al. [9] and Verbeek and Tijsen [10], our previous studies failed to detect antigenic similarity between BCV and TCV using cross-immunofluorescence, and our attempts to propagate TCV in HRT cells were unsuccessful [11]. In the present studies, sequence analyses supported our previous antigenic analyses and indicated that TCV strains NC95, Minnesota, and Indiana were closely related to IBV.

The results of the present studies demonstrate a strong genetic relationship between the avian coronaviruses, IBV and TCV, and support the contention that IBV and TCV comprise a single antigenic group/genotype within the Coronavirus genus. Further studies examining the genetic relationship between the surface glycoprotein (S) genes of TCV and other coronaviruses are warranted.
References


