ELSEVIER

# Comparisons of envelope through 5B sequences of infectious bronchitis coronaviruses indicates recombination occurs in the envelope and membrane genes

Judy Elaine Brooks [a], Aaron Cameron Rainer [a], Rebecca Lynn Parr [b], Peter Woolcock [c], Fred Hoerr [d], Ellen Whited Collisson [a,*]

[a] *Department of Veterinary Pathobiology, College of Veterinary Medicine, Texas A&M University, College Station, TX 77843, USA*
[b] *Department of Biochemistry and Biophysics, College of Agriculture and Life Sciences, Texas A&M University, College Station, TX 77843, USA*
[c] *California Veterinary Diagnostic Laboratories System, University of California, Davis, CA 95616, USA*
[d] *Department of Pathobiology, College of Veterinary Medicine, Auburn University, Auburn, AL 36849, USA*

## Abstract

A 1.78 kb sequence, including the E, M, 5a and 5b genes, and the intergenic region between the M and 5a genes, of six US strains of infectious bronchitis (corona)virus (IBV) were sequenced and compared to the published sequences for two additional strains. The overall identities as determined through pairwise analyses of nucleotide sequences of the entire 1.78 kb region ranged from 90 to 99%, with the 5b open reading frame (ORF) having the greatest identity (94–99%) while the identities of the E, 5a and M ORFs ranged from 87 to 100%. Nucleotide sequencing of recent field isolates from Alabama (Ala1) and California (Cal3) revealed distinct shifts in homology in the M gene, indicating two apparent recombination events between the Holland 52/Mass41-like strain and an Ark-like strain, both origins of commonly used vaccine strains. Putative sites of recombination could also be identified in both the E and M ORFs of laboratory strains of IBV.
© 2003 Elsevier B.V. All rights reserved.

*Keywords:* Infectious bronchitis virus; Membrane ORF; Recombination; Nucleotide sequence comparisons

## 1. Introduction

Infectious bronchitis virus (IBV) is a linear, positive-sense RNA coronavirus that is responsible for recurring economic losses to the poultry industry. First described by Schalk and Hawn (1931) as a "gasping disease" of baby chicks, the clinical signs of IBV infection include respiratory manifestations, such as tracheal rales and coughing, diarrhea, egg loss and nephritis (Beaudette and Hudson, 1937; Winterfield and Hitchner, 1962; Crinion and Hofstad, 1972; Gilchrist, 1963; Cumming, 1963; Julian and Willis, 1969). Different strains of the virus may cause distinct illness. The Gray and Holte strains, for example, have been shown to cause both nephrogenic and respiratory disease (Winterfield and Hitchner, 1962), whereas the Mass41 and Ark99 strains cause primarily respiratory disease (Beaudette and Hudson,

1937; Winterfield and Hitchner, 1962; Cumming, 1963; Darbyshire et al., 1978; Johnson and Marquardt, 1976; Sutou et al., 1988). The pathogenesis and genome arrangement of IBV are similar to the recently described severe acute respiratory syndrome coronavirus or SCoV (Peiris et al., 2003; Drosten et al., 2003; Ksiazek et al., 2003; Marra et al., 2003; Rota et al., 2003).

The RNA genome of IBV is about 27.6 kilobases (kb) in length (Boursnell et al., 1987). IBV, like other coronaviruses, has a unique discontinuous transcription which forms a nested set of genomic and subgenomic mRNAs with common 3′ termini with the smaller mRNAs. The 5′ end of each RNA contains the leader sequence found only in the 5′ unique untranslated region of the genome (Spaan et al., 1983; Lai et al., 1984; Stern and Kennedy, 1990a,b; Gorbalenya, 2001). Due to the economic importance of IBV for much of the last 72 years, an accumulated wealth of genetic IBV variants with distinct serotypes and pathotypes are available.

Four major structural proteins make up the virion of IBV and SCoV; the glycosylated spike (S) protein,

---

* Corresponding author. Tel.: +1-409-845-4122;
fax: +1-409-862-1088.
*E-mail address:* ecollisson@cvm.tamu.edu (E.W. Collisson).

post-translationally cleaved in IBV but not SCoV, the hydrophobic envelope or small membrane (E) protein, the membrane (M) protein, glycosylated at the amino termini, and the phosphorylated nucleocapsid (N) protein, associated with the viral RNA (Collisson et al., 1992; Cavanagh and Naqi, 1997). A lipid bilayer membrane formed by budding from intracellular membranes surrounds the nucleocapsid.

The genome of the Beaudette (Beau) strain of IBV was the first coronavirus genome completely sequenced (Binns et al., 1985; Boursnell et al., 1984, 1985a,b; Boursnell and Brown, 1984; Brown et al., 1984). The conserved and variable regions of the S and N genes and 3′ untranslated region from a number of US strains of IBV have been identified, as well as potential recombination sites within the S protein (Binns et al., 1986a,b; Niesters et al., 1986a; Boursnell et al., 1985b; Williams et al., 1992, 1993; Wang et al., 1993, 1994; Kwon and Jackwood, 1995; Jia et al., 1995; Gelb et al., 1997; Keeler et al., 1998; Callison et al., 1999; Nix et al., 2000; Kingham et al., 2000). However, comparisons of genomic sequences between the S and N genes on the IBV genome have not been examined with most US strains. Only the nucleotide sequences of the Mass41 (Niesters et al., 1986b; Cavanagh and Davis, 1988; Liu et al., 1991) and Holl52 (Cavanagh and Davis, 1988) E and M genes have been reported.

As early as 1988, Cavanagh and Davis postulated that IBV recombination could have occurred between two Dutch strains. Sequence analysis of a portion of the S1 and M genes indicated that one recombinant could have the M gene of one parent strain and the S1 gene of the other (Cavanagh and Davis, 1988; Niesters et al., 1986a). Kusters et al. (1989, 1990) and Zwaagstra et al. (1992) used phylogenetic trees to illustrate that some strains may have evolved by recombination. Cavanagh et al. (1990) demonstrated that a Portuguese isolate was a recombinant, and that part of the S1 gene of this isolate was very similar to that of the vaccine strain H120. Wang et al. (1993) found evidence of natural recombination in the S1 gene of field isolates of IBV from Texas, concluding that vaccine strains may be involved in recombination of IBV and that recombination occurred in nature. Later studies by Wang et al. (1994) concluded that shifts in homology, as well as point mutations within the S1 gene were common and contribute to the evolution of IBV. Jia et al. (1995) sequenced the isolate CU-T2 and observed recombination events between three different IBV strains within the S2 gene and the N gene with evidence that the Holl52 vaccine strain N contributed to half that gene in CU-T2. Recombination continues to be identified in recent isolates from the United States and Southeast Asia (Lee and Jackwood, 2000a,b, 2001; Yu et al., 2001). However, to our knowledge, recombination within the M gene has not been previously reported from a field isolate.

This report compares the genomic sequences from the 5′ end of the E gene through the 3′ end of the 5b gene of eight strains of IBV. Six of these were completely sequenced in this region for the first time (Mass41, Iowa, Conn, Arkansas Delmarva Peninsula (Ark DPI), Gray, and Holl52). In addition, the E, M, 5a, and 5b genes of one field isolate from Alabama (Ala1) and most of the M gene of two additional field isolates from California (Cal2 and Cal3) were sequenced and analyzed. Apparent recombinations in the genome of two recent IBV field isolates are described.

## 2. Materials and methods

### 2.1. Viral stocks

The Gray strain of IBV was obtained from Dr. R.W. Winterfield at Purdue University, and the Arkansas DPI strain of IBV was obtained from Dr. J.K. Rosenberger at the University of Delaware. Dr. Fred Hoerr kindly provided a field isolate (99 06426 P-3, referred to in this manuscript as Ala1); Dr. Peter Woolcock provided two field isolates used in this study (F9601491, referred to as Cal2, and F9700100, as Cal3).

### 2.2. RT-PCR amplification

Virus was propagated in embryonating chicken eggs (Williams et al., 1992). The protocol described by Burleson et al. (1992) was used to extract RNA. Reverse transcriptase-polymerase chain reaction (RT-PCR) was performed, using the protocol for the SuperScript Preamplification System for First Strand cDNA Synthesis from GibcoBRL (Rockville, MD). Primers were made in the DNA Core Facility, Department of Pathobiology, College of Veterinary Medicine, Texas A&M University (Dr. James Derr, Director, College Station, TX). Table 1 shows the sequences of the primers that were used for PCR amplification and sequencing. In addition, primer D1, complementary to the sequence of the 3′ end of the genome, was used for the RT reactions. A 1–5 μg RNA sample was mixed with 2 mM primer in DEPC-treated water and incubated for 10 min at 70 °C. After a 1 min incubation on ice, the RNA-primer mixtures were mixed with 10× PCR buffer, 25 mM MgCl₂, 10 mM dNTP mix, and 0.1 M dithiothreitol (DTT). This mixture was incubated for 5 min at 42 °C. SuperScript II reverse transcriptase (RT) was added to each tube, and this mixture was incubated at 42 °C for 50 min. The reaction was terminated first by incubating at 70 °C for 15 min and then by chilling on ice. RNase H was added, and the reactions were incubated for 20 min at 37 °C. The first strand complementary DNA (cDNA) was either used immediately for PCR amplification or stored at −20 °C.

Target cDNA was amplified by PCR, using the SuperScript Preamplification System for First Strand cDNA Synthesis from GibcoBRL (Rockville, MD). The following reaction was mixed in PCR tubes: 10× PCR buffer, 25 mM MgCl₂, 10 mM dNTP mix, 10 mM forward and reverse amplification primers, Taq DNA polymerase, cDNA, and autoclaved, distilled water. PCR amplification was performed at 94 °C for 3 min, then 30 cycles of 95 °C for 1 min, 50 °C

Table 1
Primers used for PCR amplification of fragments of infectious bronchitis virus to sequence the region from the E gene through the 5b gene

| Primer | Sequence | Sense | Position |
| --- | --- | --- | --- |
| E1 | ACTATTGTTATTACAGAGGA | Positive | 24,141–24,160 |
| E2 | TGCATAGCCATACTG | Negative | 24,613–24,627 |
| M1 | CTTAACAATCCGGAATTAGAA | Positive | 24,420–24,440 |
| M2 | TTTCCTAAGAACGGTTGGAATA | Positive | 24,459–24,480 |
| M3 | ACTCTTTAAGCGGTGTAGGTC | Positive | 24,810–24,830 |
| M4 | GGCACAAAAGAGTGGAAAAACTG | Negative | 6,479–6,499 |
| M5 | CTGTCTTTCCAGTTGCCTTACC | Negative | 25,882–25,903 |
| M6 | TCCCGCGTGTACCTCTCTAGTA | Positive | 25,801–25,822 |
| D1 | GGATCCGCTCTAACTCTATACTAGCCTA | Negative | 27,580–27,608 |

Primers are written 5'–3'. All primer positions are in relationship to primer locations in the Beaudette genome except primer M4, which is within a deleted region of the Beau genome. The position of primer M4 is relative to the start codon of the S gene of strain KB8523.

for 30 s, and 72 °C for 1 min, followed by a 3 min period at 72 °C. Samples of the PCR products were run on a 1% agarose gel to identify cDNA products.

### 2.3. Nucleotide sequencing

The cDNA samples were purified with the QIAquick PCR Purification Kit from Qiagen (Santa Clarita, CA). PCR products were directly sequenced or were sequenced after cloning into the pCR 2.1 plasmid from the TOPO TA Cloning Kit (Invitrogen, Carlsbad, CA). Sequencing was performed using the ABI PRISM dRhodamine Terminator Cycle Sequencing Kit (Applied Biosystems, Foster City, CA). Sequencing PCR reactions were performed for 25 cycles, with 10 s at 96 °C, 5 s at 50 °C, and 4 min at 60 °C. Residual primers were removed from these samples by ethanol precipitation or Bio-Rad Micro Bio-Spin chromatography columns (Hercules, CA). The products of these reactions were sent to the DNA Core Facility, Department of Pathobiology, College of Veterinary Medicine, Texas A&M University (College Station, TX) for dideoxy sequence analysis. Each region was sequenced at least twice from at least two PCR products from different RT reactions.

### 2.4. Computer analysis

Computer analysis of the sequences was performed using the CLUSTLW, SIXFRAME, and DRAWGRAM programs, all from Biology Workbench 5.2 (San Diego Supercomputer Center, University of California, San Diego, CA, 2000). GenBank submissions were made with the BankIt software program through the National Center for Biotechnology Information (NCBI), US National Library of Medicine (Bethesda, MD).

### 2.5. GenBank accession numbers

GenBank accession numbers for the genes submitted through this study included: Gray E, AF318282; Ark DPI E, AF317209; Holl52 E, AF317210; Conn E, AF317471; Iowa E, AF467921; Gray M, AF286180; Mass41 M, AF286184; Ark DPI M, AF286181; Holl52 M, AF286185; Conn M, AF286182; Iowa M, AF286183; Gray intergenic through 5b, AF469011; Mass41 intergenic through 5b, AF469015; Ark DPI intergenic through 5b, AF469012; Holl52 intergenic through 5b, AF469016; Conn intergenic through 5b, AF469013; Iowa intergenic through 5b, AF469014; Alabama 1 E, AF470627; Alabama 1 M, AF470628; Alabama 1 5a and 5b, AF470626; California 2 M, AF470629; California 3 M, AF470630.

## 3. Results

The nucleotide sequences of a region extending from the E gene through the 5b gene of six US strains of IBV (Gray, Mass41, Ark DPI, Holl52, Conn, and Iowa) and putative amino acid sequences were determined and compared with each other and the published sequences of the Beaudette (Boursnell et al., 1987) and KB8523 (Sutou et al., 1988) strains. The E, M, 5a, and 5b genes were sequenced for one new field isolate (Ala1). In addition, the M genes of two new field isolates from California were sequenced. This region represented a 1780 nucleotide sequence from the KB8523 strain genome (Sutou et al., 1988), but was shorter in the Beaudette strain (Boursnell et al., 1987) and in the Iowa strain due to lengthy deletions. The remaining five strains had about the same number of nucleotides as the KB8523 strain. The longest sequence of complete conservation was a stretch of 38 identical bases near the 5′ end of the M gene. For each strain examined, the intergenic (IG) consensus sequence CTTAACAA, found at positions 78 through 85 upstream from the start codon of the M gene, positions 9 through 16 upstream from the start codon of the 5a gene, and positions 96 through 103 upstream from the N gene, was identical and in the same positions in the previously published strains.

### 3.1. Pairwise sequence homologies

Pairwise nucleotide sequence homologies in the E gene through 5b region of the various strains ranged from 90 to
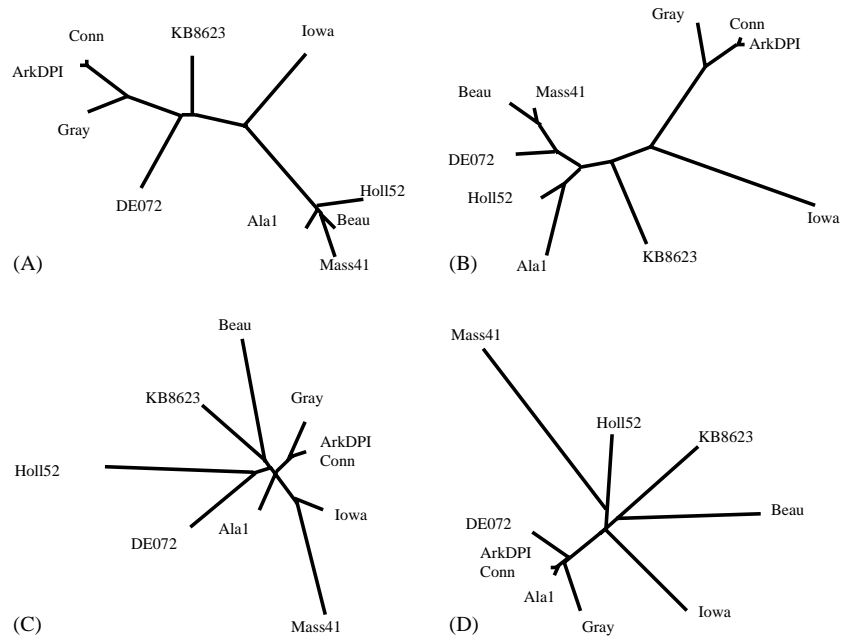
Fig. 1. Phylogenetic trees comparing Ala1, DE072, and eight other strains of infectious bronchitis virus. (A) E gene; (B) M gene; (C) 5a gene; (D) 5b gene.

99% (data not shown). Overall, the 5b ORF, with 94–99% identity among strains, was the most conserved region and the least conserved E ORF, had 87–99% nucleotide sequence homology. In the 5a gene, the nucleotide sequences of the Ark DPI and Conn strains were 100% identical and 97% identical to those of the Gray strain. The intergenic region between the M gene and the 5a gene was much more variable with an overall identity of only 74%. In addition, the Beaudette strain had a large deletion in this region, and the nucleotide sequences for the Ark DPI, Conn, and Gray strains were 97% identical.

Pairwise putative amino acid identities for each individual putative protein ranged from a low of 79% identity for the E, between the Gray and Holl52 strains and between the Conn and Holl52 strains, to 100% identity for M, between the Mass41 and Holl52 strains, and for the 5a gene, between the Conn and Ark DPI strains.

### 3.2. Phylogenetic trees

The sequence relationships of these eight strains of IBV, strain DE072, and the field isolate Ala1 are illustrated in phylogenetic trees for the E, M, 5a, and 5b genes (Fig. 1). The trees for all four genes indicated that Gray, Ark DPI, and Conn cluster together. The E gene of the recent Ala1 isolate clustered with Beau, Holl52, and Mass41 (Fig. 1A) and the M gene of Ala1 was also grouped closely with Holl52 (Fig. 1B). However, in the 5a and 5b trees (Fig. 1C and D), Ala1 sequences were placed in a cluster with Conn, Ark DPI, and Gray. The inconsistencies in the clustering observed among the individual trees of each gene suggested shifts in sequencing homology.

### 3.3. Shifts in homology

Analysis of the region from the E gene through the 3′ end of the 5b region located the shifts in homology or potential recombination sites in the E and M genes (Fig. 2), while the remainder of the 1.78 kb sequence did not appear to have any
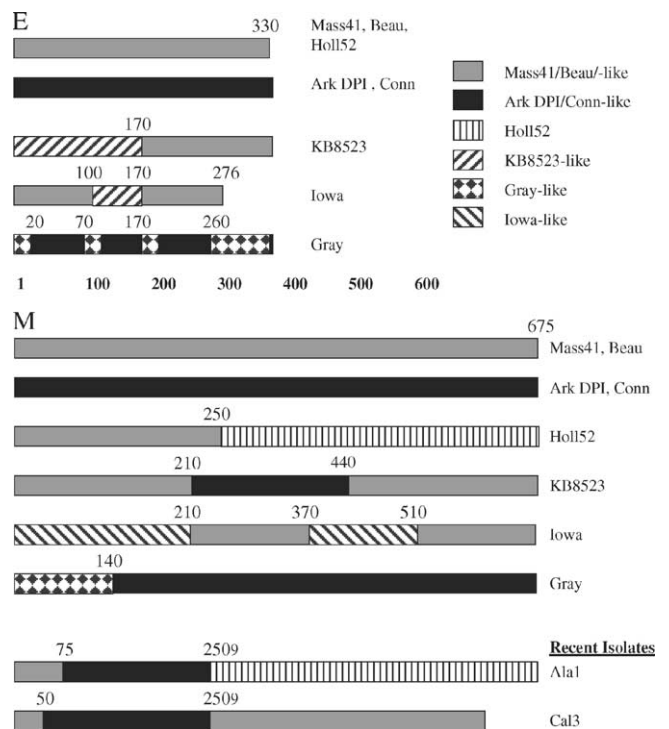


Fig. 2. Shifts in homology in the M gene of infectious bronchitis virus.

Table 2
Percentage identities of sequences within the M genes of Ala1 and Cal3

| Nucleotide location | ArkDPI | Mass41 | Holl52 |
|---|---|---|---|
| Ala1 | | | |
| 1–60 | 64 | 100 | 98 |
| 61–215 | 99 | 90 | 89 |
| 216–678 | 94 | 95 | 99 |
| Cal3 | | | |
| 1–59 | 69 | 93 | 91 |
| 60–225 | 96 | 90 | 89 |
| 226–600 | 91 | 93 | 93 |

obvious changes in homology. The E genes of the Mass41, Holl52, and Beau strains were very similar in nucleotide sequence (Fig. 2). The Iowa strain was similar to Mass41, Holl52, and Beau for most of the 5′ end and the 3′ end of the E gene, but the middle region was closer to that of the KB8523 strain. The Iowa E gene contained a deletion from bases 276 through 301. The entire E gene of the Ark DPI and Conn strains were very close. The Gray strain E gene was unlike that of the other strains except for three small regions of approximately 50, 80, and 70 nucleotides that were similar to Ark DPI and Conn.

The nucleotide sequences of the M gene for Mass41 and Beaudette were clustered together (Fig. 2), whereas the Ark DPI and Conn nucleotide sequences were nearly identical to each other. Although the 5′ end of the Gray M gene differed from the other strains, the 3′ end was quite similar to that of Ark DPI and Conn and in contrast, the 5′ end of Holl52 was similar to that of Mass41, KB8523, and Beau, while the 3′ end differed from all the other strains. The middle regions of the M represented shifts in homology or potential recombination in the Gray and KB8523. Putative sites of recombination were also identified at the 3′ end of the Iowa strain M gene. Only 46% of the 5′ 50 bases of the M gene were conserved. This region encodes the amino acids of M that lie external to the virion envelope.

### 3.4. Analysis of field isolates

When the E, M, 5a, and 5b genes for the 1999 Ala1 isolate were sequenced and compared with the other genomes (Table 2), shifts in homology were obvious. The nucleotide sequence of the E gene of this isolate was similar to those of Mass41, Holl52, and Beau. The M gene appeared to have originated from at least two distinct parental strains (Fig. 2). The first approximately 75 base sequence of the Ala1 M gene was similar to Mass41, Holl52, KB8523, and Beau, but from nucleotide 80 through 140 was identical to the M sequences of Ark DPI and Conn. A region of approximately 600 nucleotides from nucleotide 150 to the 3′ end and into the intergenic region downstream of the M, was nearly identical with Holl52. However, the 5a gene and most of the 5b gene for Ala1 were again identical to Conn and Ark DPI. These comparisons with the Ala1 sequences identified two appar-

ent recombination events within the M gene. The first recombination occurred between nucleotide 60 and nucleotide 70 of the M gene, shifting from a Mass41, Holl52, KB8523, or Beau-like sequence to a sequence perfectly identical to Ark DPI and Conn. Another apparent recombination event occurred between nucleotide 150 and nucleotide 215 of the M gene and thereafter, the Ala1 nucleotide sequence most closely resembled the Holl52 strain.

Wang et al. (1993) noted that the sequence CTT(A/T) (A/T)G was located near every putative junction where homology shifts were found. The sequence CTTTTG was identified between bases 171 and 176 of the M gene (Fig. 3). This site is in the middle of a 60-nucleotide region of 100% identity, including the Ala1 isolate. In addition, the second apparent recombination event in the Ala1 isolate occurred within this region of total homology. The nucleotide sequences of the 5a and 5b genes for Ala1 were 99% identical to those of Conn and Ark DPI and only 93 and 96% identical to the respective genes for Holl52, indicating that another recombination event may have occurred within the intergenic region between the M gene and the 5a gene.

Overall, the nucleotide sequence of the M for the recent isolate Cal2 was similar to that of the Gray strain (data not shown). However, the Cal3 isolate had putative sites of recombination in the M gene. From nucleotide 170 of the E gene through nucleotide 50 of the M gene, the Cal3 nucleotide sequence was similar to that of the Mass41, Holl52, KB8523, and Beau group. However, from approximately nucleotide 60 of the M gene through 250 of the M gene, the Cal3 sequence was similar to that of Ark DPI. From about 250 the Cal3 then resembled the M gene of Mass41 and Beau. Thus, Cal3 had two apparent recombination events within the M gene.

### 4. Discussion

This is the first report of nucleotide sequence comparisons, representing the E through the 5b region of the IBV genome. The comparisons of sequences from eight strains of IBV showed abrupt shifts in homology that suggested recombination had occurred in several locations within this region. Williams et al. (1992, 1993) analyzed the nucleocapsid gene sequences and the 3′ non-coding regions and found that the five strains that they analyzed (Mass41, Beaudette, KB8523, Ark99, and Gray strains) had a percent nucleotide identity ranging from 93.2 (between the Ark99 and KB8523 strains) to 99.1 (between the Beaudette and Mass41 strains) for the 3′ non-coding region. The similarities of the amino acid sequences of the nucleocapsid proteins of those five strains ranged from 90.7 to 96.3% and the Gray strain showed the least overall similarity to the other strains.

Wang et al. (1994) produced a phylogenetic tree that showed the relationships of 25 different IBV strains and field isolates based on their S1 amino acid sequences. These were divided into five genotypic groups based on the homologies

```
Ark DPI    ATGATG---G AAA---ATTG CACACTTAAC TTAGAGCAGG CAACTCTGCT
Holl52     ...TCCAAC. ...CAA.... T..T...G.. ..T..A...T ..GT.GA...
Ala1       ...TCCAAC. .G.CAA.... T..T...G.. ..T..A...T ..GT.GA...
Cal3       ...TCCAAC. CGGCAA.... ...T...G.. .GT..A...T ..GT.GA...

Ark DPI    TTTTAAAGAA TATAATTTAT TTATAACCGC ATTCCTATTG TTTCTTACTA
Holl52     .........G .......... .......T.. ....T.G... ..CT.A..C.
Ala1       .........G .......... .......... .......... ..........
Cal3       .........G .......... .......... .......... ..........

Ark DPI    TACTACTTCA GTATGGGTAC GCAACTAGGA GTCGGTTTAT TTATATACTG
Holl52     ..A....... ......T..T .....A..A. ..A....... .........T
Ala1       .......... .......... .......... .......... ......A...
Cal3       .......... ......A..T .....A.... .....N.... ...C...A..

Ark DPI    AAAATGATAG TGTTATGGTG CTTTTGGCCC CTTAACATTG CAGTAGGTGT
Holl52     .......... .......... .......... .......... ..........
Ala1       .......... .......... .......... .......... ..........
Cal3       .......... .......... .......... .......... ..........

Ark DPI    AATTTCATGT ATATATCCAC CAAATACAGG AGGTCTTGTC GCAGCGATAA
Holl52     .......... .....C.... ....C..... .......... ..........
Ala1       .......... .....C.... ....C..... .......... ..........
Cal3       .......... .......... .......... .......... .......A..

Ark DPI    TACTTACTGT GTTTGCTTGT CTTTCTTTTG TAGGTTATTG GATTCAGAGT
Holl52     .......... ......G... .......... .......... ...C......
Ala1       .......... ......G... .........T .......... ...C......
Cal3       .......A.. .G....G... ..G....... A...A..... ..........

Ark DPI    TGTAGACTCT TTAAAAGGTG TAGGTCTTGG TGGTCTTTTA ACCCCGAGTC
Holl52     AT........ ....GC.... ...A...... ....A.... ....A..A..
Ala1       AT........ ....GC.... ...A...... ....A.... ...A..A..
Cal3       NTC....... ....GC..G. .G..CAA... ...G.A.... ....T.....

Ark DPI    TAATGCCGTA GGTTCAATAC TCCTCACAAA TGGTCAACAA TGTAATTTTG
Holl52     ...C...... .......... ....A..T.. .......... ..........
Ala1       ...C...... .......... ....A..T.. .......... ..........
Cal3       .......... .......... ....AT.T.. .......... ..........

Ark DPI    CTATAGAGAG TGTGCCTATG GTGCTTGCTC CAATTATAAA GAACGGTGTC
Holl52     .......... ......G... ......T... .T........ ...T.....T
Ala1       .......... ......G... .......... .T........ ...T.....T
Cal3       .......... ......A... ......T... .......... ...T..A..T

Ark DPI    CTTTATTGTG AGGGTCAGTG GCTTGCTAAA TGTGAACCAG ACCACTTGCC
Holl52     .......... .......... .......... .......... ..........
Ala1       .......... .......... .......... .......... ..........
Cal3       .......... .......... .......... .......... ..........

Ark DPI    TAAAGACATA TTTGTATGCA CACCGGATAG ACGTAATATC TATCGTATGG
Holl52     .......... .......... ....A..... .......... ..........
Ala1       .......... .......... ....A..... .......... ..........
Cal3       .......... .....T...C ....A..... .......... ..C.......

Ark DPI    TGCAGAAATA CACTGGTGAC CAAAGCGGAA ATAAGAAAAG GGTTGCTACA
Holl52     .......... .......... .......... .......... .T........
Ala1       .......... .......... .......... .......... .T........
Cal3       .......... .......... .......... .......... .T........

Ark DPI    TTTGTCTATG CAAAGCAGTC AGTAGACACT GGCGAGCTAG AAAGTGTACC
Holl52     .......... .......... A......... ..C....... ........G.
Ala1       .......... .......... A......... ..C....... ........G.

Ark DPI    AACAGGAGGA AGTAGTCTTT ACACATAA
Holl52     ......T... .....C.... ........
Ala1       ......T... .....C.... ........
```

Fig. 3. Nucleotide sequences of the M gene of Ark DPI, Holl52, Alabama 1 (Ala1), and California 3 (Cal3).

among the strains. The Iowa 609 and Gray strains were in one group; the KB8523, H52, Beaudette, Mass41, and Conn strains were in another; Ark99 strain was in still another group. Like the Williams et al. (1992, 1993) studies of the nucleocapsid and 3′ UTR sequences, the phylogenetic trees of the genes of the 1.78 kb region clustered Gray and Ark at one end of the tree, while Beau, Holl52, and Mass41 were placed at the other end. Wang et al. (1994), however, found that the S1 genes of the Iowa and Gray strains were clus-

tered together and the Ark99 strain was in another group. As with the present study and the Williams et al. (1992, 1993) studies, Wang et al. (1994) noted that the S1 gene of Beau, Holl52, and Mass41 formed a common group. Combined with the present work, these results indicated a homology shift for the Ark, Gray, and Iowa strains somewhere in the region between the S1 gene and the E gene. Moreover, the phylogenetic trees in this study suggested that putative recombination occurred to generate the recently isolated Ala1

strain. The tree for the E gene placed Ala1 in a cluster with Beau and Mass41. In the M gene, Ala1 was grouped with Holl52, whereas in the 5a and 5b genes, Ala1 was close to Conn, Ark DPI, and Gray.

It has been postulated that the intergenic (IG) consensus sequences CTTAACAA or CTGAACAA and the conserved region around these sequences may be "hot spots" for recombination (Jia et al., 1995; Lee and Jackwood, 2000b). In the present study, no recombination "hot spots" were found at the IG consensus sequences (CTTAACAA). The sequence CTTTTG, that Wang et al. (1993) described adjacent to every putative recombination junction, was found twice within the M gene and once within the 5b gene for every strain studied in this paper. In addition, this sequence was found once within the E gene, once within the intergenic region between the M and 5a genes, and once within the 5a gene for several of the strains. Data from the present study demonstrated one putative recombination site in the M gene of the Ala1 isolate that was 40 nucleotides downstream from the first CTTTTG sequence in the M gene. These so-called "hot spots" could possibly promote recombination because they are A-T-rich regions. However, recombination events among the eight strains would seem to occur randomly rather than at specific "hot spots", because only one apparent recombination occurred near the CTTTTG sequence.

Whereas there were no detectable recombination events in the 5a or 5b genes, the nucleotide sequence of the 5b gene is well conserved. Sequencing of field isolates demonstrated recombination within the M gene. A recombination sequence could be identified as related to a portion of the Holl52 vaccine. Such natural recombination is apparently a widespread event and occurs throughout the genome, providing a natural process for the construction of genetic variants that produce more virulent strains of IBV.

## Acknowledgements

## References

Beaudette, F.R., Hudson, C.B., 1937. Cultivation of the virus of infectious bronchitis. J. Am. Vet. Med. Assoc. 90, 51–60.

Binns, M.M., Boursnell, M.E.G., Cavanagh, D., Pappin, D.J.C., Brown, T.D.K., 1985. Cloning and sequencing of the gene encoding the spike protein of the coronavirus IBV. J. Gen. Virol. 66, 719–726.

Binns, M.M., Boursnell, M.E., Tomley, F.M., Brown, T.D.K., 1986a. Comparison of the spike precursor sequences of coronavirus IBV strains M41 and 6/82 with that of IBV Beaudette. J. Gen. Virol. 67, 2825–2831.

Binns, M.M., Boursnell, M.E., Tomley, F.M., Brown, T.D., 1986b. Nucleotide sequence encoding the membrane protein of the IBV strain 6/82. Nucl. Acids Res. 14, 55–58.

Boursnell, M.E.G., Binns, M.M., Brown, T.D.K., 1985a. Sequencing of coronavirus IBV genomic RNA: three open reading frames in the 5′ 'unique' region of mRNA D. J. Gen. Virol. 66, 2253–2258.

Boursnell, M.E.G., Binns, M.M., Foulds, I.J., Brown, T.D.K., 1985b. Sequences of the nucleocapsid genes from two strains of avian infectious bronchitis virus. J. Gen. Virol. 66, 573–580.

Boursnell, M.E.G., Brown, T.D.K., 1984. Sequencing of coronavirus IBV genomic RNA: a 195-base open reading frame encoded by mRNA B. Gene 29, 87–92.

Boursnell, M.E.G., Brown, T.D.K., Binns, M.M., 1984. Sequence of the membrane protein gene from avian coronavirus IBV. Virus Res. 1, 303–313.

Boursnell, M.E.G., Brown, T.D.K., Foulds, I.J., Green, P.F., Tomley, F.M., Binns, M.M., 1987. Completion of the sequence of the genome of the coronavirus avian infectious bronchitis virus. J. Gen. Virol. 68, 57–77.

Brown, T.D.K., Boursnell, M.E.G., Binns, M.M., 1984. A leader sequence is present on mRNA A of avian infectious bronchitis virus. J. Gen. Virol. 65, 1437–1442.

Burleson, F.G., Chambers, T.M., Wiedbrauk, D.L., 1992. Virology—A Laboratory Manual. Academic Press, San Diego.

Callison, S.A., Jackwood, M.W., Hilt, D.A., 1999. Infectious bronchitis virus S2 gene sequence variability may affect S1 subunit specific antibody binding. Virus Genes 19, 143–151.

Cavanagh, D., Davis, P.J., 1988. Evolution of avian coronavirus IBV: sequence of the matrix glycoprotein gene and intergenic region of several serotypes. J. Gen. Virol. 69, 621–629.

Cavanagh, D., Davis, P., Cook, J., Li, D., 1990. Molecular basis of the variation exhibited by avian infectious bronchitis coronavirus (IBV). Adv. Exp. Med. Biol. 276, 369–372.

Cavanagh, D., Naqi, S., 1997. Infectious bronchitis. In: Calnek, B.W., et al. (Ed.), Diseases of Poultry, 10th ed. Iowa State University Press, Ames, IA, pp. 511–526.

Collisson, E., Williams, A., Parr, R., Wang, L., 1992. The molecular biology of avian infectious bronchitis virus. Poult. Sci. Rev. 4, 41–55.

Crinion, R.A.P., Hofstad, M.S., 1972. Pathogenicity of four serotypes of avian infectious bronchitis virus for the oviduct of young chickens of various ages. Avian Dis. 16, 351–363.

Cumming, R.B., 1963. Infectious avian nephrosis (uraemia) in Australia. Aust. Vet. J. 39, 145–147.

Darbyshire, J.H., Cook, J.K.A., Peters, R.W., 1978. Growth comparisons of avian infectious bronchitis virus strains in organ cultures of chicken tissues. Arch. Virol. 56, 317–325.

Drosten, C., Gunther, S., Preiser, W., van der Werf, S., Brodt, H.R., Becker, S., Rabenau, H., Panning, M., Kolesnikova, L., Fouchier, R.A., Berger, A., Burguiere, A.M., Cinatl, J., Eickmann, M., Escriou, N., Grywna, K., Kramme, S., Manuguerra, J.C., Muller, S., Rickerts, V., Sturmer, M., Vieth, S., Klenk, H.D., Osterhaus, A.D., Schmitz, H., Doerr, H.W., 2003. Identification of a novel coronavirus in patients with severe acute respiratory syndrome. N. Engl. J. Med. 348, 1967–1976.

Gelb Jr., J., Keeler Jr., C.L., Nix, W.A., Rosenberger, J.K., Cloud, S.S., 1997. Antigenic and S-1 genomic characterization of the Delaware variant serotype of infectious bronchitis virus. Avian Dis. 41, 661–669.

Gilchrist, P., 1963. A survey of avian respiratory diseases. Aust. Vet. J. 39, 140–144.

Gorbalenya, A.E., 2001. Big Nidovirus genome. When count and order of domains matter. Adv. Exp. Med. Biol. 494, 1–17.

Jia, W., Karaca, K., Parrish, C.R., Naqi, S.A., 1995. A novel variant of avian infectious bronchitis virus resulting from recombination among three different strains. Arch. Virol. 140, 259–271.

Johnson, R.B., Marquardt, W.W., 1976. Strains of infectious bronchitis virus on the Delmarva Peninsula and in Arkansas. Avian Dis. 20, 382–386.

Julian, R.J., Willis, N.G., 1969. The nephrosis–nephritis syndrome in chickens caused by a Holte strain of infectious bronchitis virus. Can. Vet. J. 10, 18–19.

Keeler Jr., C.L., Reed, K.L., Nix, W.A., Gelb Jr., J., 1998. Serotype identification of avian infectious bronchitis virus (IBV) by RT-PCR of the peplomer (S-1) gene. Avian Dis. 42, 275–284.

Kingham, B.E., Keeler Jr., C.L., Nix, W.A., Ladman, B.S., Gelb Jr., J., 2000. Identification of avian infectious bronchitis virus by direct automated cycle sequencing of the S-1 gene. Avian Dis. 44, 325–335.

Ksiazek, T.G., Erdman, D., Goldsmith, C.S., Zaki, S.R., Peret, T., Emery, S., Tong, S., Urbani, C., Comer, J.A., Lim, W., Rollin, P.E., Dowell, S.F., Ling, A.E., Humphrey, C.D., Shieh, W.J., Guarner, J., Paddock, C.D., Rota, P., Fields, B., DeRisi, J., Yang, J.Y., Cox, N., Hughes, J.M., LeDuc, J.W., Bellini, W.J., Anderson, L.J., SARS Working Group, 2003. A novel coronavirus associated with severe acute respiratory syndrome. N. Engl. J. Med. 348, 1953–1966.

Kusters, J.G., Jager, E.J., Niesters, H.G., van der Zeijst, B.A., 1990. Sequence evidence of RNA recombination in field isolates of avian coronavirus infectious bronchitis virus. Vaccine 8, 605–608.

Kusters, J.G., Niesters, H.G., Lenstra, J.A., Horzinek, M.C., van der Zeijst, B.A., 1989. Phylogeny of antigenic variants of avian coronavirus IBV. Virology 169, 217–221.

Kwon, H.M., Jackwood, M.W., 1995. Molecular cloning and sequence comparison of the S1 glycoprotein of the Gray and JMK strains of avian infectious bronchitis virus. Virus Genes 9, 219–229.

Lai, M.M.C., Baric, R.S., Brayton, P.R., Stohlman, S.A., 1984. Characterization of RNA leader sequences on the virion and mRNAs of mouse hepatitis virus, a cytoplasmic RNA virus. Proc. Natl. Acad. Sci. U.S.A. 81, 3626–3630.

Lee, C.-W., Jackwood, M.W., 2000a. Redesign of primer and application of the reverse transcriptase polymerase chain reaction and restriction fragment length polymorphism test to the DE072 strain of infectious bronchitis virus. Avian Dis. 44, 650–654.

Lee, C.-W., Jackwood, M.W., 2000b. Evidence of genetic diversity generated by recombination among avian coronavirus IBV. Arch. Virol. 145, 2135–2148.

Lee, C.-W., Jackwood, M.W., 2001. Spike gene analysis of the DE072 strain of infectious bronchitis virus: origin and evolution. Virus Genes 22, 85–91.

Liu, D.X., Cavanagh, D., Green, P., Inglis, S.C., 1991. A polycistronic mRNA specified by the coronavirus infectious bronchitis virus. Virology 184, 531–544.

Marra, M.A., Jones, S.J., Astell, C.R., Holt, R.A., Brooks-Wilson, A., Butterfield, Y.S., Khattra, J., Asano, J.K., Barber, S.A., Chan, S.Y., Cloutier, A., Coughlin, S.M., Freeman, D., Girn, N., Griffith, O.L., Leach, S.R., Mayo, M., McDonald, H., Montgomery, S.B., Pandoh, P.K., Petrescu, A.S., Robertson, A.G., Schein, J.E., Siddiqui, A., Smailus, D.E., Stott, J.M., Yang, G.S., Plummer, F., Andonov, A., Artsob, H., Bastien, N., Bernard, K., Booth, T.F., Bowness, D., Czub, M., Drebot, M., Fernando, L., Flick, R., Garbutt, M., Gray, M., Grolla, A., Jones, S., Feldmann, H., Meyers, A., Kabani, A., Li, Y., Normand, S., Stroher, U., Tipples, G.A., Tyler, S., Vogrig, R., Ward, D., Watson, B., Brunham, R.C., Krajden, M., Petric, M., Skowronski, D.M., Upton, C., Roper, R.L., 2003. The Genome sequence of the SARS-associated coronavirus. Science 300 (5624), 1399–1404.

Niesters, H.G., Lenstra, J.A., Spaan, W.J., Zijderveld, A.J., Bleumink-Pluym, N.M., Hong, F., van Scharrenburg, G.J., Horzinek, M.C., van der Zeijst, B.A., 1986a. The peplomer protein sequence of the M41 strain of coronavirus IBV and its comparison with Beaudette strains. Virus Res. 5, 253–263.

Niesters, H.G., Zijderveld, A.J., Seifert, W.F., Lenstra, J.A., Bleumink-Pluym, N.M., Horzinek, M.C., van der Zeijst, B.A., 1986b.

Infectious bronchitis virus RNA D encodes three potential translation products. Nucl. Acids Res. 14, 31–44.

Nix, W.A., Troeber, D.A., Kingham, B.F., Keeler Jr., C.L., Gelb Jr., J., 2000. Emergence of subtype strains of the Arkansas serotype of infectious bronchitis virus in Delmarva broiler chickens. Avian Dis. 44, 568–581.

Peiris, J.S.M., Chu, C.M., Cheng, V.C.C., Chan, K.S., Hung, I.F.N., Poon, L.L.M., Law, K.I., Tang, B.S.F., Hon, T.Y.W., Chan, C.S., Chan, K.H., Ng, J.C.S., Zheng, B.J., Ng, W.L., Lai, R.W.M., Guan, Y., Yuen, K.Y., Members of the HKU/UCH SARS Study Group, 2003. Clinical progression and viral load in a community outbreak of coronavirus-associated SARS pneumonia: a prospective study. Lancet 361, 1767–1772.

Rota, P.A., Oberste, M.S., Monroe, S.S., Nix, W.A., Campagnoli, R., Icenogle, J.P., Penaranda, S., Bankamp, B., Maher, K., Chen, M.H., Tong, S., Tamin, A., Lowe, L., Frace, M., DeRisi, J.L., Chen, Q., Wang, D., Erdman, D.D., Peret, T.C., Burns, C., Ksiazek, T.G., Rollin, P.E., Sanchez, A., Liffick, S., Holloway, B., Limor, J., McCaustland, K., Olsen-Rasmussen, M., Fouchier, R., Gunther, S., Osterhaus, A.D., Drosten, C., Pallansch, M.A., Anderson, L.J., Bellini, W.J., 2003. Characterization of a novel coronavirus associated with severe acute respiratory syndrome. Science 300 (5624), 1394–1399.

Schalk, A.F., Hawn, M.C., 1931. An apparently new respiratory disease of baby chicks. J. Am. Vet. Med. Assoc. 78, 413–422.

Spaan, W., Delius, H., Skinner, M., Armstrong, J., Rottier, P., Smeekens, S., van der Zeijst, B.A.M., Siddell, S.G., 1983. Coronavirus mRNA synthesis involves fusion of non-contiguous sequences. EMBO J. 2, 1839–1844.

Stern, D.F., Kennedy, S.I.T., 1990a. Coronavirus multiplication strategy. I. Identification and characterization of virus-specific RNA. J. Virol. 34, 665–674.

Stern, D.F., Kennedy, S.I.T., 1990b. Coronavirus multiplication strategy. II. Mapping the avian infectious bronchitis virus intracellular RNA species to the genome. J. Virol. 36, 440–449.

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

Wang, L., Junker, D., Collisson, E.W., 1993. Evidence of natural recombination within the S1 gene of infectious bronchitis virus. Virology 192, 710–716.

Wang, L., Junker, D., Hock, L., Ebiary, E., Collisson, E.W., 1994. Evolutionary implications of genetic variations in the S1 gene of infectious bronchitis virus. Virus Res. 34, 327–338.

Williams, A.K., Wang, L., Sneed, L.W., Collisson, E.W., 1992. Comparative analyses of the nucleocapsid genes of several strains of infectious bronchitis virus and other coronaviruses. Virus Res. 25, 213–222.

Williams, A.K., Wang, L., Sneed, L.W., Collisson, E.W., 1993. Analysis of a hypervariable region in the 3′ non-coding end of the infectious bronchitis virus genome. Virus Res. 28, 19–27.

Winterfield, R.W., Hitchner, S.B., 1962. Etiology of an infectious nephritis–nephrosis syndrome of chickens. Am. J. Vet. Res. 23, 1273–1279.

Yu, L., Wang, Z., Jiang, Y., Low, S., Kwang, J., 2001. Molecular epidemiology of infectious bronchitis virus isolates from China and Southeast Asia. Avian Dis. 45, 201–209.

Zwaagstra, K.A., van der Zeijst, B.A., Kusters, J.G., 1992. Rapid detection and identification of avian infectious bronchitis virus. J. Clin. Microbiol. 30, 79–84.