Evidence of genetic diversity generated by recombination among avian coronavirus IBV

C.-W. Lee and M. W. Jackwood

Department of Avian Medicine, College of Veterinary Medicine, The University of Georgia, Athens, Georgia, U.S.A.

Accepted March 24, 2000

Summary. Previously, we demonstrated that the DE072 strain of IBV is a recombinant which has an IBV strain D1466-like sequence in the S gene. Herein, we analyzed the remaining 3.8 kb 3′ end of the genome, which includes Gene 3, Gene 4, Gene 5, Gene 6, and the 3′ non-coding region of the DE072 and D1466 strains. Those two viruses had high nucleotide similarity in Gene 4. However, the other individual genes had a much different level of sequence similarity with the same gene of the other IBV strains. The genome of five IBV strains, of which the complete sequence of the 3′ end of the genome has been determined, were divided at an intergenic (IG) consensus sequence (CTGAACAA or CTTAACAA) and compared phylogenetically. Phylogenetic trees of different topology indicated that the consensus IG sequences and the highly conserved sequence around this regions may serve as recombination ‘hot spots’. Phylogenetic analysis of selected regions of the genome of the DE072 serotype field isolates further support those results and indicate that isolates within the same serotype may have different amounts of nucleotide sequence similarity with each other in individual genes other than the S gene. Presumably this occurs because the consensus IG sequence serves as the template switching site for the viral encoded polymerase.

Introduction

Infectious bronchitis virus causes a highly contagious upper-respiratory disease in chickens. The disease is characterized by increased ocular and nasal secretions, excess mucus in the trachea, decreased weight gain and feed efficiency in broilers, and declines in egg production and egg quality in layers. Although live attenuated vaccines are available, IBV continues to be a severe economic problem in commercial chickens because many different serotypes of the virus exist and do not cross protect [3].
Infectious bronchitis virus (IBV) is a coronavirus in the new order Nidovirales [4]. Members of the Nidovirales order have a single stranded positive sense RNA genome and produce a 3’ nested set of subgenomic mRNAs when they replicate [4]. Coronaviruses are divided into three antigenic groups based primarily on their structural proteins. Infectious bronchitis virus is the type strain of coronaviruses and is the only virus placed in antigenic group three. Characteristics of this group are a cleaved spike (S) glycoprotein, an N-glycosylated membrane (M) protein, and no hemagglutinin/esterase protein [19]. The genome of IBV is approximately 27 kilobases in length [1]. It is organized into six regions, each containing one or more open reading frames (ORF’s), which are separated by intergenic sequences (IG) that contain the signal for transcription of subgenomic mRNAs [1, 17]. The viral RNA-dependent RNA polymerase is encoded in the 5’ two thirds of the viral genome by two overlapping open reading frames (ORF1a and ORF1b) [1]. The structural protein genes are located 3’ to the viral polymerase gene and are in order from 5’ to 3’, the S glycoprotein gene (gene 2), the small envelope (E) gene (gene3), the M glycoprotein gene (gene4), and the nucleocapsid (N) gene (gene6) [19, 20].

Evolution in IBV has been observed through the occurrence of variant viruses and analysis of known serotypes. More than twenty serotypes within IBV have been recognized worldwide and are thought to be generated by insertions, deletions, point mutations and RNA recombination [2, 3, 6, 14]. Evidence of natural recombination for several IBV strains has been reported [10, 15, 22]. However, because of the limited sequence information, recombination has only been described for a small part of the genome. So far, the complete sequence of the 3’ end of the genome (from the 3’ end of the polymerase gene to the poly A tail) of only three strains, Beaudette, KB8523 and CU-T2 have been determined [1, 11, 20].

The DE072 strain was first isolated in 1992 in the Delmarva peninsula region of the USA and initial characterization of this virus indicated this virus was serologically distinct from any other IBV serotypes in North America [7]. Previously, we demonstrated that the DE072 strain is a recombinant which has a D1466-like sequence in the S1 and S2 genes [18]. D1466 is an IBV vaccine strain of the D212 serotype from the Netherlands [7, 13, 14]. Herein, we describe the sequences of the remaining genes of the DE072 and D1466 strains with the exception of gene 1 (the polymerase gene). We conducted phylogenetic analysis by dividing the genome in the IG sequence to elucidate possible role of this sequence in the homologous recombination in IBV. Further, we conducted sequence analysis of six isolates of the DE072 serotype in order to determine if recombination is frequently occurring in this region in field isolates of IBV.

Materials and methods

Viruses

Viruses used in this study are listed in Table 1. The viruses were propagated in 9-day-old embryonated specific-pathogen-free (SPF) chicken eggs (SELECT Laboratories, Gainesville,
Recombination in IBV

Table 1. Viruses used in this study

| Strain/isolates | Serotype | Origin          | Source                                      |
|-----------------|----------|-----------------|---------------------------------------------|
| DE072           | DE072    | Delmarva, USA   | J. Gelb Jr.a                                |
| D1466           | D212     | Netherlands     | Y. Weismanb                                 |
| 97-6370         | DE072    | Minnesota, USA  | PDRCc                                       |
| 97-6386         | DE072    | Arizona, USA    | PDRC                                        |
| 98-2831         | DE072    | Illinois, USA   | PDRC                                        |
| 99-5381         | DE072    | Georgia, USA    | PDRC                                        |
| 99-5425         | DE072    | Kansas, USA     | PDRC                                        |
| 99-5658         | DE072    | Georgia, USA    | PDRC                                        |

aUniversity of Delaware, Newark, DE USA
bKimron Veterinary Institute, Israel
cPoultry Diagnostic and Research Center, Athens, GA, USA

GA, USA). The D1466 strain of IBV was obtained as phenol-inactivated allantoic fluid using USDA import permit #42290.

Viral RNA extraction and RT-PCR

Viral RNA from IBV grown in embryonating eggs was extracted using the High Pure PCR Template Preparation Kit (Boehringer Mannheim, Indianapolis, IN, USA) according to the manufacturers recommendation. RNA from the phenol-inactivated allantoic fluid of D1466 was extracted with a modification in first several step of the High Pure PCR Template Preparation Kit. Briefly, 1.5 ml of the infectious allantoic fluid was placed into a microcentrifuge tube and centrifuged at 13,000 g for 5 min. The aqueous top layer, approximately 200 μl, was transferred to new tube. Binding buffer (200 μl) and 40 μl of proteinase K (18 mg/ml) was added and incubated for 10 min at 70°C. Then 150 μl of chloroform/isoamyl alcohol (49:1) was added, vortexed gently for 5–10 sec and then placed on ice for 15 min. The mixture was centrifuged at 13,000 × g for 10 min. The upper phase was transferred to a clean 1.5 ml tube and 100 μl of chloroform/isoamyl alcohol (49:1) was added. The mixture was vortexed gently for 5–10 sec. This was centrifuged for 2 min at 13,000 × g, and the upper phase was transferred to a clean 1.5 ml tube. Remaining steps were followed sequentially as described by the manufacturer.

Gene 3, Gene 4, Gene 5, Gene 6, and a 421 bp hypervariable region (HVR) of the S1 gene were amplified separately using the Titan One Tube RT-PCR System (Boehringer Mannheim). Primer sets used to amplify Gene 3, Gene 4, and the HVR in S1 are listed in Table 2. The primers utilized for amplification of Gene 5 and Gene 6 have been reported [8, 23]. The reaction conditions for RT-PCR were previously described [16, 23].

Sequencing and analysis

PCR products were cut from 1% agarose gels and purified using the QIA quick Gel Extraction Kit (Qiagen, Santa Clarita, CA, USA). Purified PCR products were either sequenced directly or cloned into the TA cloning vector (Invitrogen, Carlsbed, CA, USA), and automated sequencing with the Prism DyeDeoxy terminator cycle sequencing kit (Perkin Elmer, Foster City, CA, USA) was conducted at the Molecular Genetics Instrumentation Facility, University of Georgia. Sequencing primers to various regions of the gene for DE072 and
Table 2. The oligonucleotide sequences of primers used in this study

| Primer       | 5’->3’ sequence                  | Position |
|--------------|----------------------------------|----------|
| Gene 3 U     | catgacttgtgtgttgtgtgtg           | −141−−121|
| Gene 3 L     | cctttctat.instructions          | 1222−1242|
| Gene 4 U     | ttttctttgtgtattttg              | 920−940  |
| Gene 4 L     | gttatttttttgtgtgttgtg           | 1677−1697|
| HVR in S1    | agtacaggcctcctaatgg             | 95−113   |
| Ag072 5’     | agtacaggcctcctaatgg             | 95−113   |
| Ag072 3’     | cacygctcttaactaat              | 535−553  |

The relative primer positions were calculated using the ATG start site of Gene 3 as 1 for primers gene 3 and 4, and ATG start site of S1 gene as 1 for primers HVR in S1

D1466 were designed using OLIGO version 4.0 software (National Bioscience, Plymouth, MN, USA) and are available upon request.

Assembly of sequencing contigs, translation of nucleotide sequence into protein sequence, and initial multiple sequence alignments were performed with the Clustal V method in MegAlign software versin 1.03 (DNASTAR Inc., Madison, WI, USA). Phylogenetic trees for each gene were generated using the maximum parsimony method with 100 bootstrap replicates in a heuristic search using the PAUP 3.1 software program [21].

Nucleotide sequence accession numbers

The nucleotide sequences reported here have been deposited with the GenBank. The accession numbers are as follows: DE072 (Gene 3), AF202998; DE072 (Gene 4), AF202999; DE072 (Gene 5), AF203000; DE072 (Gene 6), AF203001; DE072 (3’ end non-coding region), AF203002; D1466 (Gene 3), AF203003; D1466 (Gene 4), AF203004; D1466 (Gene 5), AF203005; D1466 (Gene 6), AF203006; D1466 (3’ end non-coding region), AF203007; 98-2831 (HVR in S1), AF206254; 99-5831 (HVR in S1), AF206255; 99-5425 (HVR in S1), AF206256; 99-5658 (HVR in S1), AF206257; 97-6370 (HVR in S1), AF206258; 97-6386 (HVR in S1), AF206259; 98-2831 (Gene 3), AF206260; 99-5381 (Gene 3), AF206261; 99-5425 (Gene 3), AF206262; 99-5658 (Gene 3), AF206263; 97-6370 (Gene 3), AF206264; 97-6386 (Gene 3), AF206265; 98-2831 (Gene 4), AF206266; 97-6386 (Gene 4), AF206267; 99-5425 (Gene 4), AF206268; 99-5658 (Gene 4), AF206269; 97-6370 (Gene 4), AF206270; 97-6386 (Gene 4), AF206266.

The complete sequence of the 3’ end of the genome of three strains, Beaudette, KB8523 and CU-T2 and Gene 6 of Holl52 strain have been previously reported [1, 11, 20, 23].

Results

Sequence analysis of DE072 and D1466

A total of 3839 nucleotide and 3861 nucleotide were found, respectively, in a region beginning from the 5’ end of gene 3 to the 3’ end of DE072 and D1466 genome. The intergenic sequence CTGAACAA or CTTAACAA was found immediately upstream of the start site for each gene of both strains. The sequences were identical to those found in the corresponding genomic areas of the Beaudette, KB8523, and CU-T2 strains (Fig. 1).
Fig. 1 (continued)
Fig. 1 (continued)
Fig. 1 (continued)
Fig. 1 (continued)
Fig. 1. The nucleotide sequence alignment of gene 3, gene 4, gene 5, and gene 6, and 3’ non-coding region. Dots indicates nucleotide identical to that of DE072 strain. The conserved nucleotide sequences ctgaacaa or cttaaacaa, which is located at the starting site of each gene, is in bold character. Heavy underlines indicate the putative start codons, asterisks above the sequence indicate the stop codons.

Gene 3 of both strains contained three ORFs, 3a, 3b, and 3c. Gene 4 consisted of the M protein gene with a single ORF and a non-coding region between the 3’ end of the M protein gene and gene 5. Gene 5 contained two ORFs (5a and 5b). Gene 6 consisted of the N protein gene with a single ORF and a 3’ non-coding region. Downstream from the stop codon of the N gene, a 15 base insertion was found in the D1466 genome which also occurs in the genome of the Holl52 (Fig. 1).

Sequence comparison and phylogenetic analysis

The 3’-terminal 3.8 kb of the genome of five strains and gene 6 of the Holl52 strain were compared. The nucleotide sequence similarities among coding regions of gene 3, M, gene 5, and the N protein gene of DE072 and other strain were between 83.3–97.6%. Those of D1466 and other strains were between 78.7–98.2% identical. D1466 showed only 1.8% nucleotide difference with Holl52 in Gene 6. Gene 3c and gene 5b were relatively more conserved than the other genes (Table 3).

Genes were divided by IG sequences (CTGAACAA/CTTAACAA) and phylogenetic analysis was conducted. The DE072 strain clustered with the CU-T2
Table 3. Percentage nucleotide homologies between coding regions of gene 3, M protein gene, gene 5, and N protein gene of IBV strains

| IBV        | Percent homology with DE072 | Percent homology with D1466 |
|------------|----------------------------|----------------------------|
|            | Gene 3  | Gene 5  | Gene 3  | Gene 5  |
|            | 3a      | 3b      | 3c      | M       | 5a      | 5b      | N       | 3a      | 3b      | 3c      | M       | 5a      | 5b      | N       |
| D1466/DE072 | 83.3    | 83.6    | 87.8    | 96.9    | 92.4    | 95.2    | 92.4    | 83.3    | 83.6    | 87.8    | 96.9    | 92.4    | 95.2    | 92.4    |
| Beaudette  | 85.6    | 83.6    | 88.4    | 96.5    | 90.9    | 96.8    | 91.1    | 95.4    | 98.5    | 97.6    | 95.6    | 91.4    | 94.4    | 90.1    |
| KB8523     | 85.6    | 89.7    | 93.6    | 94.1    | 92.4    | 96.8    | 93.7    | 80.5    | 86.2    | 91.2    | 94.8    | 92.9    | 95.2    | 91.3    |
| CU-T2      | 85.6    | 97.4    | 90.8    | 84.8    | 97.0    | 97.6    | 94.7    | 78.7    | 82.6    | 83.7    | 82.3    | 91.4    | 95.6    | 95.5    |
| Holl52     | N/A     | N/A     | N/A     | N/A     | N/A     | N/A     | 92.6    | N/A     | N/A     | N/A     | N/A     | N/A     | 98.2    |

N/A Not available

Fig. 2. Phylogenetic analysis of DE072 and D1466 with other IBV strains in Genes 3, 4, 5, and 6. A The linear structure of IBV genomic RNA. Genes are divided by intergenic (IG) sequences which is a stretch of consensus sequences (CTGAACAA or CTTAACAA). B Phylogenetic analysis using parsimony for five IBV genes based on nucleotide sequence. All trees were constructed by general bootstrap analysis using 100 replicates and midpoint rooted. Branch lengths are provided in each tree.
Recombination in IBV 2145

strain in all genes except gene 4, where it clustered with D1466 and separated far from CU-T2. On the other hand, D1466 clustered with Beaudette in gene 3 and gene 5, and clustered with Holl52 in gene 6 (Fig. 2). KB8523, which is only the nephropathogenic strain, was solely placed in all genes compared.

**Phylogenetic analysis of field isolates of DE072 serotype**

In order to demonstrate the genetic heterogeneity of the same serotype isolates of IBV, we conducted phylogenetic analysis using six DE072 serotype field isolates. Phylogenetic analysis of the hypervariable region (HVR) in S1, clustered all the DE072 serotype isolates in one group with the prototype strain of the DE072

---

**Fig. 3.** Phylogenetic analysis of field isolates of DE072 serotype IBV. A Schematic representation of the genome of IBV. I–III indicate the regions used to construct the phylogenetic tree. B Phylogenetic trees for the regions I–III as indicated in A. I Trees of HVR in S1, which is 421 bp; II trees of gene 3 sequences; III trees of partial gene 4 sequences upstream residue 670 bp
serotype of IBV. This group was far from other serotypes of IBV strains in tree length. However, phylogenetic tree of gene 3 and gene 4 showed differences in tree topology among six isolates. In Gene 3, only one isolate, 98-2831, clustered with DE072. In gene 4, no isolates clustered with DE072 and formed groups randomly with other serotypes of IBV (Fig. 3).

**Discussion**

DE072 is a recent isolate made in 1992 [6]. In a previous study of the S gene, we demonstrated that this virus was closely related to D1466 which is an IBV vaccine strain of the D212 serotype from the Netherlands [7, 13, 18]. Analysis of gene 4 also reveals a high sequence relatedness between DE072 and D1466 (Table 3). However, in the other genes analyzed in this study, DE072 shares high sequence similarity with the CU-T2 strain which has also been reported to be a recombinant between Arkansas and Massachusetts strains [10]. Considering the fact that both strains were isolated in the northeastern USA, it is possible that they had undergone similar selection pressure. On the other hand, D1466 shows high similarity with Beaudette and Holl52 strains in genes other than the S gene. The percent similarity in the N gene and a 15 base insertion in the 3’ non-coding region suggests that both D1466 and Holl52 are closely related (Table 3, Fig. 1). The Holl52 strain has been extensively used as a live vaccine in Europe [5]. This finding provides more convincing evidence that vaccine strains are contributing to the emergence of variants in the field. Based on these results, we suggest that DE072 and D1466 had the same origin, but diverged a long time ago and evolved independently in different geographical locations.

Since recombination in coronaviruses is thought to occur by a template switching mechanism [8, 19], we speculate that IG sequences may serve as ‘hot spots’ for homologous recombination. So far, recombinations suggested in IBV have been used on a small part of the genome [10, 15, 22]. Examining only a small part of the genome may result in misleading conclusions because of point mutations or conserved regions of the gene. We conducted phylogenetic analysis by dividing 3.8 kb of the 3’ end of the genome among five IBV strains at the IG sequences. Phylogenetic trees of this sequence data had very different topology (Fig. 2), which indicates that recombination had occurred. It has been reported that RNA recombination in IBV can occur randomly in non-localized sites in vitro [12]. However, considering the selection pressure in vivo recombination in the IG sequences should be advantageous to virus in two aspects. First, since crossovers occur at the site of consensus IG sequences, there would be no shift in the codon reading frame. Second, since whole genes are substituted, there would be no drastic change in the conformation of proteins encoded by individual genes. Further, crossovers at each of the five IG sequences would generate tremendous genetic diversity. This amount of diversity may contribute to persistence and to the continuing emergence of new variants of IBV despite vaccination efforts.

Finally, we conducted sequence analysis of 6 isolates of the DE072 serotype to demonstrate how random recombination occurs within the same serotype.
Recombination in IBV

Phylogenetic analysis of the HVR in S1 shows that these 6 isolates cluster together because they are the same serotype. However, these 6 isolates had a much different level of nucleotide sequence similarity with each other in gene 3 and gene 4, and clustered randomly with other serotypes of IBV (Fig. 3). Based on this result, it is clear that isolates of the same serotype can differ substantially in individual genes. Thus, every field isolate of IBV could be unique in each gene sequence because of recombination.

Acknowledgements

We express appreciation to Dr. Yoram Weisman for providing D1466 virus and Deborah Hilt for technical assistance. Thanks are also extended to Drs. Bruce Seal, Maricarmen Garcia, and Holly Sellers for the review of this manuscript.

References

1. Boursnell MEG, Brown TDK, Foulds IJ, Green PF, Tomley FM, Binns MM (1987) Completion of the sequence of the genome of the coronavirus avian infectious bronchitis virus. J Gen Virol 68: 57–77
2. Cavanagh D, Davis PJ, Cook J, Li D, Kant A, Koch G (1992) Location of the amino acid differences in the S1 spike glycoprotein subunit of closely related serotypes of infectious bronchitis virus. Avian Pathol 21: 33–43
3. Cavanagh D, Naqi SA (1997) Infectious bronchitis. In: Calnek BW, Barnes HJ, Beard CW, Reid WM, Yoder HW (eds) Disease of poultry, 10th ed. Iowa State University Press, Ames, pp 511–526
4. Cavanagh D (1997) Nidovirales: a new order comprising Coronaviridae and Arteriviridae. Arch Virol 142: 629–633
5. Davelaar FG, Kouwenhoven B, Burger AG (1984) Occurrence and significance of infectious bronchitis virus variant strains in egg and broiler production in the Netherlands. Vet Q 6: 114–120
6. Gelb Jr J, Wolff JB, Moran CA (1991) Variant serotypes of infectious bronchitis virus isolated from commercial layer and broiler chickens. Avian Dis 35: 82–87
7. Gelb Jr J, Keeler Jr CL, Nix WA, Rosenberger JK, Cloud SS (1997) Antigenic and S-1 genomic characterization of the Delaware variant serotypes of infectious bronchitis virus. Avian Dis 41: 661–669
8. Jackwood MW, Kwon HM, Hilt DA (1992) Infectious bronchitis virus detection in allantoic fluid using the polymerase chain reaction and a DNA probe. Avian Dis 36: 403–409
9. Jarvis TC, Kirkegaard K (1992) Poliovirus RNA recombination: mechanistic studies in the absence of selection. EMBO J 11: 3135–3145
10. Jia W, Karaca K, Parrish DR, Naqi SA (1995) A novel variant of avian infectious bronchitis virus resulting from recombination among three different strains. Arch Virol 140: 259–271
11. Jia W, Naqi SA (1997) Sequence analysis of gene 3, gene 4 and gene 5 of avian infectious bronchitis virus strain CU-T2. Gene 189: 189–193
12. Kottier SA, Cavanagh D, Britton P (1995) Experimental evidence of recombination in coronavirus infectious bronchitis virus. Virology 213: 569–580
13. Kuster JG, Niesters HM, Bleumink-Pluym NMC, Davelaar FG, Horzinek MC, Van Der Zeijst BAM (1987) Molecular epidemiology of infectious bronchitis virus in the Netherlands. J Gen Virol 68: 343–352
Kusters JG, Niesters H, Lenstra JA, Horzinek MC, Van Der Zeijst BAM (1989) Phylogeny of antigenic variants of avian coronavirus IBV. Virology 169: 217–221

Kusters JG, Jager EJ, Niesters HGM, Van Der Zeijst BAM (1990) Sequence evidence for RNA recombination in field isolates of avian coronavirus infectious bronchitis virus. Vaccine 8: 605–608

Kwon HM, Jackwood MW, Gelb Jr J (1993) Differentiation of infectious bronchitis virus serotypes using polymerase chain reaction and restriction-fragment-length-polymorphism analysis. Avian Dis 37: 194–202

Lai MMC, Liao C-L, Lin Y-J, Zhang X (1994) Coronavirus: How a large RNA viral genome is replicated and transcribed. Infect Agent Dis 3: 98–105

Lee C-W, Jackwood MW (2000) Spike gene analysis of the DE072 strain of infectious bronchitis virus: origin and evolution. Virus Genes (in press)

Siddell SG (1995) The coronaviridae. In: Fraenkel-Conrat H, Wagner RR (eds) The viruses. Plenum Press, New York, pp 1–49

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

Swofford DL (1989) PAUP: Phylogenetic analysis using parsimony. Version 3. Illinois Natural History Survey, Champaign

Wang L, Junker D, Collison EW (1993) Evidence of natural recombination within the S1 gene of infectious bronchitis virus. Virology 192: 710–716

Williams AK, Wang L, Sneed LW, Collisson EW (1992) Comparative analyses of the nucleocapsid genes of several strains of infectious bronchitis viruses and other coronaviruses. Virus Res 25: 213–222

Authors’ address: Dr. M. W. Jackwood, Department of Avian Medicine, The University of Georgia, 953 College Station Road, Athens, GA 30602, U.S.A.

Received December 1, 1999