



## Spike Gene Analysis of the DE072 Strain of Infectious Bronchitis Virus: Origin and Evolution

CHANG-WON LEE & MARK W. JACKWOOD\*

Department of Avian Medicine, College of Veterinary Medicine, University of Georgia, Athens, Georgia 30602-4875

Received April 20, 2000; Accepted June 10, 2000

**Abstract.** The entire S2 gene of the DE072 strain of infectious bronchitis virus (IBV) was sequenced. The nucleotide and amino acid sequence was most similar to the D1466 strain and was 84.8% and 89.9% identity, respectively. The nucleotide and amino acid sequence similarity among the DE072 strain and other IBV strains was less than 71.9% and 76.6%, respectively. Phylogenetic analysis, based on both nucleotide and amino acid sequence, showed that IBV isolates were divided into two distinct groups. The DE072 strain clustered only with the D1466 strain, and all of the other strains were distinct from those two viruses. Further the nucleotide sequence analysis of the entire spike glycoprotein gene of the DE072 strain demonstrated that most of the gene contained a D1466-like sequence, and five putative cross-over sites were identified. Based on cross-over site, phylogenetic trees were constructed for different regions of the spike gene, and a difference in topology between these trees was observed. Considering the difference in S2 gene sequence identity and tree topology, we assume that DE072 and D1466 viruses share a different origin from other isolates of IBV. Furthermore, entire spike gene analysis indicates that the DE072 strain has undergone recombination event as well as extensive antigenic variation.

**Key words:** infectious bronchitis virus, DE072, recombinant

### Introduction

Infectious bronchitis (IB) is an acute, highly contagious upper-respiratory disease in chickens and is characterized by tracheal rales, sneezing, and coughing. Avian infectious bronchitis virus (IBV), the etiologic agent of IB, causes considerable economic losses in the poultry industry. Signs of the disease include a decline in weight gain and feed efficiency as well as decreased egg production and egg shell quality. When IBV is a component of mixed infections with *Escherichia coli*, airsacculitis usually occurs and results in condemnations at the processing plant (7).

IBV is the prototype strain of the *Coronaviridae*, and the single stranded positive sense RNA genome contains the information for four structural proteins

(7,22). The spike (S) protein, which is involved with virus neutralization, serotype specificity, and cell attachment, is cleaved post-translationally into N-terminal S1 and C-terminal S2 subunits (6,8). Because virus neutralizing epitopes have been associated with the S1 subunit, it has been studied more extensively than any other gene component (12,20). However, it has been shown that the S2 subunit can also induce neutralizing antibodies, and IBV serotypes can be grouped based on S2 gene sequence (3,16).

Diversity in the genome of IBV is thought to be generated by insertions, deletions, point mutations, and RNA recombination (2,5,13). Several isolates have been identified that have major segment of their genome derived from two different isolates. Dual infections of a field strain and a vaccine could lead to such recombination events (4,14). Furthermore, genetic drift and shift occur naturally in coronavirus (15,24).

The DE072 strain of IBV is a relatively new variant, which was isolated in the northeastern region

\*Author for all correspondence: Department of Avian Medicine, University of Georgia, Athens, GA 30602-4875.  
E-mail: mjackwoo@arches.uga.edu

of the United States in 1992 (9). Large differences between the S1 gene sequence of the DE072 strain and other IBV isolates hindered molecular epidemiology studies of this virus. We chose to sequence the entire S2 gene because that sequence is more conserved than the sequence of the S1 gene yet still shows sequence variability among different serotypes of IBV (3, 16). That sequence was used to elucidate the evolution of the DE072 virus.

## Materials And Methods

### *Viruses*

The DE072 strain was obtained from Dr. Jack Gelb, Jr. (University of Delaware, Newark DE). The virus was propagated in nine-day-old embryonated specific-pathogen-free (SPF) chicken eggs (SELECT Laboratories, Gainesville, GA) to prepare stock as previously described (19).

### *Viral RNA Extraction and RT-PCR*

RNA was isolated from virus containing allantoic fluids using the polymerase chain reaction (PCR) Template Preparation Kit (Boehringer Mannheim, Indianapolis, IN) according to the manufacturer's recommendations. The 5' half and the 3' half of the S-2 gene of the DE072 strain were amplified separately by reverse transcriptase-polymerase chain reaction (RT-PCR) using the Titan™ One Tube RT-PCR System (Boehringer Mannheim). The reaction conditions for RT-PCR were previously described (17). Two sets of primers for the amplification of the S-2 gene were designed using OLIGO™ version 4.0 software (National Bioscience, Plymouth, MN). The primer sequences are as follows: primers for the 5' half amplification are 5'ccgtagaattagaaggtccac3' and 5'cagaagaatagcaccacaaat3', and primers for the 3' half amplification are 5'ttcagaagaaccaggagaaa3' and 5'agtacaagtttacaccaaaga3'.

### *Cloning and Sequencing the S-2 Gene*

PCR products were cut from 1% agarose gels and purified using the QIAquick Gel Extraction Kit (Qiagen). Purified PCR products were cloned into the TA cloning vector (Invitrogen, Carlsbad, CA), and ligated areas of the plasmids were sequenced with the

M13 universal forward and reverse primers (Molecular Genetics Instrumentation Facility, University of Georgia). The nucleotide sequence of the S2 gene reported here has been deposited with the GenBank. (accession number: AF201930)

### *Sequence Analysis*

For the S2 gene analysis, 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 version 1.03 (DNASStar Inc., Madison, WI). 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 (23). Selected sequences from GenBank were used to construct a phylogenetic tree.

For the entire S gene analysis, alignments and phylogenetic analysis were conducted using Clustal V method in MegAlign software version 1.03 (DNASStar Inc., Madison, WI). These methods use a multiple alignment algorithm, and the unweighted pair group method with arithmetic mean algorithm (UPGMA) (21).

## Results

The nucleotide and deduced amino acid sequence of the S2 subunit of the DE072 strain was compared with previously published sequence from other IBV strains (Table 1). The S2 nucleotide and amino acid sequence of D1466 was most similar to the DE072 strain with 84.8% and 89.9% identity, respectively. The nucleotide and amino acid sequence similarity between the DE072 strain and other IBV strains was less than 71.9% and 76.6%, respectively. The S2 gene nucleotide sequence of selected isolates were aligned and individual bases were compared (Fig. 1). There were homology shift junctions at bases 984, 1091, and 1594 counting from the first base of the S2 gene (Fig. 1). The sequence of the DE072 strain was most similar to the D274 strain in the region between nucleotides 984–1027 (Table 2). In the 30 bp region between nucleotides 1091–1120, there was only one nucleotide difference between the DE072 strain and Mass41 strain (Table 2). The sequence of the DE072 strain at the 3' end of the S2 gene had the most identity with the

```

DE072 TCCACTAGTGA AAAATGTTACATCTCTGCCATATGTAACCTTATGGAAAATPTTGTATTA AACCCAGATGGTGATATTC AAMATATGTCGCCAGAAGAATTTA
D1466 ..T.....G.....T.....C.....T.....CT.G.CAT.T.A.....
D274 ..T.T.C.GG.....AA.T...T...T...C.G...A...T...TCA...C.CA.A.A...A...T.GG
M41 ..T.T.C.....G..AA.T.C.T...T.G.....T.G.....A...T...TCA...G.C.CA.A.A...A.C.T.GG

101 AAGATTATACATCTCTTTTCCPTAATAGGACTGATTAATGACTTATAACCCACAGTPTTAACTGTTACAGATGAGTPTTATACAGACTCAATPTCA
.....T.....C.....C.G.C.A.....AC.....C.A.G..
..C..T.GTG.A.C.A...GTA...AA...G.C...TG.....A.C.T...AC...A.G.GTA.GG
..C.G.T.GTG.A.C..A...GTT...AA...G.C...T.....AC...A.G.GTA.GG

201 AAAGATCAAAATTAATTCATTC AATATGTGTGTGGAAATGCTATTCATTCATGCAAGCAATGTTTCAACAGTATGGTTCGTGPTTGTGATAACATACTTCT
..G.....A...T...T...A.....G.....C.G..A...
T...G.C.....C...G...T.C.C.A...T.GG.G.T.GAA.G...A...GC...C...T.G...
T...G.C.....TC.G.G...T...C.A...C.G.G.T.T.GAG.T...A...GC...C...T.G...

301 ATTGTGAATGGTATAGCATTTC AAGATAATGCTGAAATGCTCCACTPTTATCTCTACTAAGCCCTAGAGGTTTGGATACAAACTCAITTTGTAATTTTA
...T...T.G.GCA.....G..A.T.G.....A.....C..AGC.C...
G.A.A...A..G.T.GTCA.A...AG..ATG...C.TT.A...C.C.....A..RTCT...A...CCGTTT.C..AG...G...
G.A.A...A...T.GTCA.A...AG..ATG...C.TT.GA.T.C.....A..GCCT...A...CCA.TTC..AG...G...

401 CAGCAGGTGAGTPTTAATATCTCTCTTGTATTACTAAGAATGGACACCCACAGGGGGCTCTTTTATAGAGGATTTGTTGTTGATAAAGTGAATCACT
GT..T.G.T...T...T.AA...C...C.....T...C.T...G...
GTA.T...T.G...A.ACTCC.A.TAGTG.T.GT...T...T.A...C.T.A.CACA.GT..T...TG-
GCA.T...T...C.G...A.A.CTCC.A.TAGT..T.GAA...T...T.A.CC.TC.A..ACA.GC..T...TG-

501 TGGTTTAACTGGTGAACCTGCTTACCAAAAATGTACATCTGGAACCTTAGGTTTGGTTAAGACTTAGTGTGTGCACAGAAATATATGGGTTACTTGTGA
A..A...T...T...G.C...C...T...T...TC...T.C.T...A...
A...AACCA..TGAC..A..A.G.G.C.TG.A...T...C.TC...G..C...T.GTG.A...T.G...G
A...AACCA..TGAC..A..A...T.C.TG.A...T...TC...G..C..GC...T.GTG.A...T.G...G

601 CTACCCCAATTTATAACTGCGGAAAATGCAAGCAATGATATACTAGTCTCTTTGTTGTTCTCAATGGCTTTTGGTGGAAATTACATCTGCGGGAGCCATACCTT
T...G.T...G.A...T...C...C...A..G...A...
T.G.T.T...A.A...A.CT...C...CT..T...T...T.A.TT..T.T...
T.G.T.C...A.A...A.TT...C...CT..T...T...TG.A.T.T.T...

701 TTCAACTCAAAATTCAGCGGGCATAAAACACTCTGGTATTACCCAGAGGGTTTTCAGAGAGAACCGGAGAAAATTTGCTGCATGTTCAATAAAGCTAT
...G..C...A.T...A...T.A...T.A...T...
C.C.A..T.G.G.TA.A.T.T..T.G...C...T.AC.C..TT...T.A.A...T.C.T...G.C..
...C.A..C.G.G..TA.A..T.T..CT.G...T.AC..TT...T.A.A...T.C.T...G.C..

801 GAAACATATGCAAGGTGGCTTTTTCAGCTACATCTCTTGCATTCACCAAGTGCAGATGTTGTTAATGAACAGGGTGCANPTCTCCAGCAGACTATGCAT
...A.G...A.T...G.A...T.A...G...T.A.T...A...
GGCC...A.T..AG.AG..G...A..T...A.T...A.G..A.T.T.C...ACTG...C...GGC
TGCT.G...A.T..AG.AG...A..T...A.T...A.G..A...T...TACTG...GCA

901 TCACCTTAATAAAAATTTTGGTGCATATCACACTGTGCTATCAGATATTTATTAACAACTAGATGCATTAGAAGCAAAATGCACAGTTGATGCCATATTA
.....G.....T.....T...C...C.....A.A..A.AC.G...
.....T.TTC...AC.T...CC.G...T..T.TA.TC...G...T...T...
.....T.T.TTC...G.T...A.C..CC.G...T.C.CA.C...T.A.G...TC.T.A...

1001 CTGGCAGGCTTTGCTTTTATCAGTTTTTACGATCTGCTTAAACAATTTAGAGTACACTTAAAGTTTCCAGCAACGTGAATTTGGCTAAAGCAAAAATTTAATGA
.....A.C...C...T.AC...GC...G.C...CA...G...C.T...G...G...
T.A...A...T...T...A...GGC...TAC.G...A.A.G...G.A.C.CTC.G...
T..AT.G..A..AC.T..T...G..GGC...C.T.T..G..G..A.A.G...G.A...CTC.G...

1101 GGTGTAAAGTTCACAGTTTAAATAGACATGGTTTTTGTGGAGAGGCCATGCATATCATGTCCTTCCCGAGAACCCACTAATGGAATGCTCTTTTGCAT
A...A...C.C...T...C...GT...T...C.A...G...G.A...C.A...
T..A..T.A..T...GT..TCA...TA.T.A.GA..G.TC.AKOC..A.A.A.T.T...C.A.G..A.A.C
T...T...T...GT.CTCC...TA.T..ACGA..G.TC.AA.C..A.G..A.T...T.A.G..A.A.C

1201 TTCAOCTACGCCOCTGAAACATATGCAAAATGTGACAGCAGTTGTAGGTTTTTGTGTRAAACCGGCCAATGGTACTGAAATATGGALTTGTGCTGTAGTAG
.....A.C.GT...T.....G.....C..T.T.AT...A...G...ACT...
T.A...G.A.G.GT.T.TC...G..A.A...G...C...CT...C.GCC.G...C.A.A...CTAAT...
TT.T.T.T.A.T.GT.T.TT...T.T..A.A.G...G...CT...C.G.C.G...C.A.A..A.C.TAAT...

1301 CCAGAGGGGTTTTTATAGAAGTAAATGGTACTTACTATATCACAGGAGCTGATATGATAGCCACGAGCCATTACAGCAGGAGATGTAGTTAAACTGAC
.....TA...G.....A.....TCT...C.A...A...T...GG...T...
.....TA...C...T...G...C...T.C.A.A...TG..A..AT...T...A...OG..T...
T.G.TA...C...T...G...C...C..A...TG..A..T...T...A...OG..T...

```

Fig. 1. The nucleotide sequence alignment of the S2 gene. Dots indicates nucleotides identical to that of the DE072 strain. (I) indicates possible junction sites.



Mass41 strain (Table 2). Phylogenetic analysis based on both nucleotide and amino acid sequences showed that the DE072 strain clustered only with the D1466 strain (Fig. 2). All of the other strains were distinct from DE072 and D1466.

The entire S gene was divided into six sections based on homology shift junctions and phylogenetic trees were constructed, respectively (Fig. 3). In section I and III, which covers most of the S1 gene and S2 gene, respectively, the DE072 strain clustered with D1466. In the regions, designated as sections II and IV, the DE072 strain shifted towards the D274 branch. The DE072 strain clustered with Mass41 and was distant from D1466 strain in section V, which contains 30 bp in the middle of the S2 gene and section VI, which is about 300 bases at the 3' end of the S2 gene.

## Discussion

The DE072 strain was first isolated in 1992 in the Delmarva Peninsula region of the United States (9). Initial characterization of the DE072 strain indicated that this virus was serologically distinct from other IBV serotypes. The origin of the DE072 strain was uncertain. Differences between the S1 gene sequence of the DE072 strain and other isolates of IBV indicated that the DE072 virus did not originate directly from a live vaccine virus or other common IBV strain in the United States (9).

Callison et al. (3) showed that the S2 gene sequence was more conserved than the sequence of the S1 gene but still showed sequence variability among different strains of IBV. Indeed, serotypes were grouped based on S2 gene sequence. The

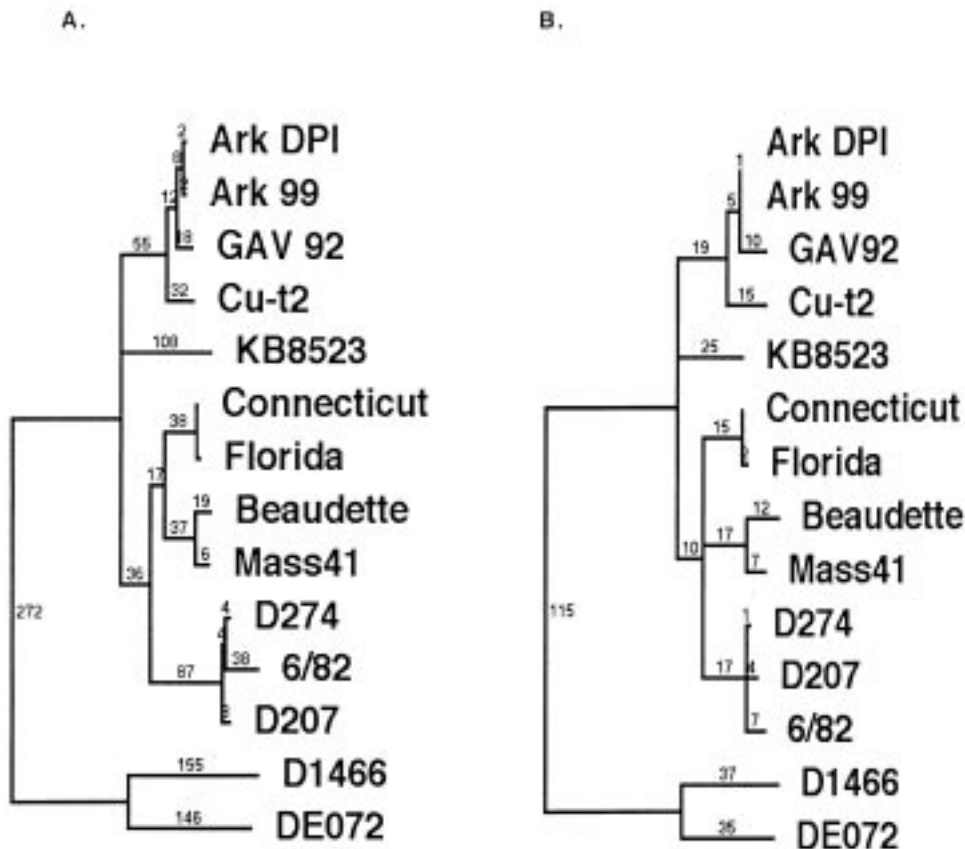


Fig. 2. Phylogenetic tree based on the nucleotide (A) and amino acid (B) sequence of the S-2 gene. Both trees are midpoint rooted and branch lengths are provided in each tree.

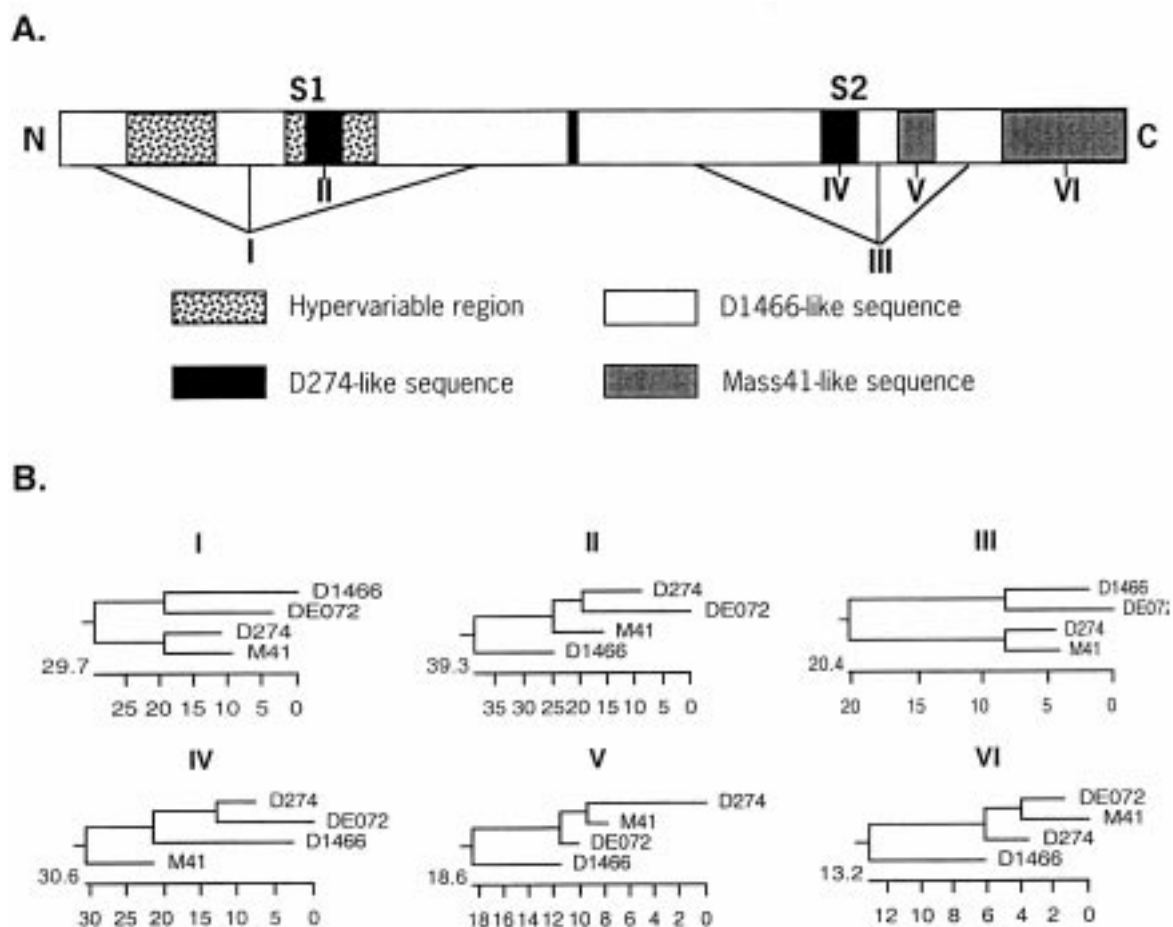


Fig. 3. (A) Schematic representation of the putative recombination site in DE072 strain. (B) Phylogenetic trees for the regions I–VI as indicated in (A). I: trees of S1 sequences between 1–120, 270–725, and 1171–1631, II: trees of S1 sequences between 1009–1109, III: trees of S2 sequences between 1–983, 1028–1090, and 1121–1593, IV: trees of S2 sequences between 984–1027, V: trees of S2 sequences between 1091–1120, VI: trees of S2 sequences between 1594–1876.

nucleotide sequence of the S2 genes available to date, in general, differ by less than 10% (14). However, the S2 gene sequence of the DE072 strain showed less than 71.9% nucleotide similarity with other IBVs except for the European vaccine strain D1466 (Table 1). Phylogenetic analysis of the S2 gene revealed that the DE072 and D1466 strains were distinct from other IBV strains. Isolates other than the DE072 and the D1466 strain formed a cluster, containing subgroups that were the same as groupings observed when S1 sequence data were analyzed.

Following alignment of the DE072 strain S2 sequence with those previously published from other IBV strains, we observed regions of sequence variation between residues 984–1027, 1091–1120,

and 1594–1876 (Figs. 1 and 3). Sequence variation in the first two regions was subtle and was probably caused by either antigenic drift or recombination. However sequence variation in the region between residues 1594–1876, which is at the 3' end of the S2 gene, probably occurred due to recombination because of a drastic homology shift. Banner et al. (1) showed that RNA recombination occurred preferentially at certain sites in the genome of murine coronavirus. This hot spot was also observed in the CU-T2 strain of IBV (10). That work supports our observation that the 3' end sequence variance in the S2 gene of the DE072 strain was caused by a recombination event. Recombination is thought to occur by template switching, which depends in part on

the presence of homologous nucleotide sequence similarities between the strains (2,10). Because the 3' end of the S2 gene is conserved between IBV strains, it is logical that that region would favor recombination. Furthermore, secondary structure reported at the 5' end of gene 3 may influence recombination at the 3' end of the S2 gene (18). However, it is also possible that this recombination hot spot may be the result of selection pressure on the virus (1,2).

Based on our data, we conclude that the DE072 virus probably shares the same origin with the D1466 strain, has undergone extensive antigenic variation, and has acquired sequences similar to the Mass41 strain in the 3' region of the S2 gene most likely by recombination. Phylogenetic analysis indicated that the D274 (11), a Dutch field isolate, originated from the Mass41 vaccine strain. As more sequence data for the S2 subunit and other genes are accumulated, the genetic basis for the diversity of IBV will be more clearly understood.

Because the DE072 strain continues to persist in the field and has a unique sequence compared with other American strains of IBV, it is an excellent choice for studying genetic drift and recombination. Furthermore, analyzing field isolates from birds with different vaccine histories can be used to elucidate the role of vaccine strains in the evolution of IBV.

## References

- Banner L.R., Keck J.G. and Lai M.M.C., *Virology* 175, 548–555, 1990.
- Banner L.R. and Lai M.M.C., *Virology* 185, 441–445, 1991.
- Callison S.A., Jackwood M.W. and Hilt D.A., *Virus Genes* 19, 1–8, 1999.
- Cavanagh D., Davis P.J. and Cook J., *Avian Pathol* 21, 401–408, 1992.
- Cavanagh D., Davis P.J., Cook J., Li D., Kant A., and Koch G., *Avian Pathol* 21, 33–43, 1992.
- Cavanagh D., Davis P.J., Pappin D.J.C., Binns M.M., Bournnell M.E.G., and Brown T.D.K., *Virus Res* 4, 133–143, 1986.
- Calnek B.W., Barnes H.J., Beard C.W., Reid W.M., and Yoder H.W. (eds), *Disease of Poultry*, 10th ed., Iowa State University Press, Ames, Iowa, 1997, pp. 511–526.
- Cavanagh D., Davis P.J. and Mockett A., *Virus Res* 11, 141–150, 1988.
- Gelb Jr. J., Keeler Jr. C.L., Nix W.A., Rosenberger J.K., and Cloud S.S., *Avian Dis* 41, 661–669, 1997.
- Jia W., Karaca K., Parrish D.R., and Naqi S.A., *Arch Virol* 140, 259–271, 1995.
- Jordi B.J.A.M., Kremers D.A.W.M., Kusters H.G., and van der Zeijst B.A.M., *Nucleic Acids Res* 17, 6726, 1989.
- Koch G., Hartog L., Kant A., and van Roozelaar D.J., *J Gen Virol* 71, 1929–1935, 1990.
- Kottier S.A., Cavanagh D. and Britton P., *Virology* 213, 569–580, 1995.
- Kusters J.G., Jager E.J., Niesters H.G.M., and van der Zeijst B.A.M., *Vaccine* 8, 605–608, 1990.
- Kusters J.G., Niesters H., Lenstra J.A., Horzinek M.C., and van der Zeijst B.A.M., *Virology* 169, 217–221, 1989.
- Kusters J.G., Jager E.J., Lenstra J.A., Koch G., Posthumus W.P., Meloen R.H., and van der Zeijst B.A., *J Immunol* 143, 2692–2698, 1989.
- Kwon H.M., Jackwood M.W. and Gelb Jr. J., *Avian Dis* 37, 194–202, 1993.
- Liu D.X. and Inglis S.C., *J Virol* 66, 6143–6154, 1992.
- Hitchner S.B., Domermuth C.H., Purchase H.G., and Williams J.E. (eds), *Isolation and Identification of Avian Pathogens*, Creative Printing Co. Inc., New York, 1980, pp. 70–72.
- Niesters H., Bleumink P.N., Osterhaus A., Horzinek M.C., and van der Zeijst B.A.M., *Virology* 161, 511–519, 1989.
- Sokal R.R. and Sneath P.H.A. (eds), *Principles of Numerical Taxonomy*. Freeman Press, San Francisco, CA, 1963, p. 359.
- Stern D.F. and Sefton B.M., *J Virol* 44, 794–803, 1982.
- Swofford D.L., *PAUP: Phylogenetic analysis using parsimony. Version 3*. Illinois Natural History Survey, Champaign, 1989.
- Wang L., Junker D. and Collison E.W., *Virology* 192, 710–716, 1993.