Virus Genes 31:1, 5–11, 2005

Using DNA Shuffling to Create Novel Infectious Bronchitis Virus S1 Genes: Implications for S1 Gene Recombination

SCOTT CALLISON, DEBORAH HILT & MARK JACKWOOD*

Department of Avian Medicine, Poultry Diagnostic and Research Center, The University of Georgia, Athens, GA, 30605, USA

Received November 24, 2004; Accepted December 15, 2004

Abstract. We employed the staggered extension process (StEP) to shuffle the S1 genes from four infectious bronchitis virus (IBV) strains representing four unique serotypes. Upon creating a shuffled S1 gene library, we randomly selected 25 clones and analyzed them by DNA sequencing. In total, eleven clones contained novel S1 gene recombinants. Based on sequence data, each recombinant was unique and contained a full-length open reading frame. The average number of crossovers per recombinant was 5 and the average number of point mutations was 1.3, leading mostly to non-synonymous amino acid changes. No recombinant contained sequences from all four parental genes and no recombinant contained any sequence from the distantly related Delaware 072 strain. Our data suggests that recombination between distantly related IBV strains within the S1 gene probably does not readily occur. This finding is extremely important in light of the common industry vaccination practice of mixing different live-attenuated IBV strains.

Key words: DNA shuffling, infectious bronchitis virus (IBV), staggered extension process (StEP), S1 gene

Introduction

Infectious bronchitis virus (IBV) continues to cause disease in poultry flocks around the world. Although commercial poultry flocks are routinely vaccinated for IBV, outbreaks of infectious bronchitis still happen due to naturally occuring variant viruses that continue to arise. Currently, modified live and killed viruses that do not cross protect are used to vaccinate chickens. Better vaccines that are low in virulence, highly immunogenic and most importantly, cross protective between different IBV strains/serotypes are needed. To create better vaccines, a strategy that keeps pace with the rate of viral change in nature is needed.

Infectious bronchitis virus is a Coronavirus that contains a large, single stranded, positive sense RNA genome that is 27.5 kb long. The genome encodes four major structural proteins: the membrane glycoprotein, the envelope protein, the nucleocapsid protein, and the spike glycoprotein, which contains conformationaly dependent virus-neutralizing and serotype specific epitopes [1]. Unfortunately, the three-dimensional structure of the spike is not known and development of an efficacious recombinant vaccine has not been successful. The ability of IBV to rapidly change in nature is attributable to its method of replication. Infectious bronchitis virus undergoes genetic drift due to point mutations that occur during replication because the viral polymerase lacks proof-reading ability. In addition, genetic shift occurs due to the recombination of two or more strains during replication. Both genetic drift and shift have been documented for IBV in the laboratory and in birds [2–7].

DNA shuffling is the process of purposeful, but random recombination of parental genes into novel recombinant genes. Basically, genes (parental) with some level of similarity are recombined in a test tube by any number of techniques, such that the newly created genes

*Author for all correspondence:
E-mail: mjackwoo@uga.edu

contain genetic information from some or all of the parental genes. The new recombinant genes can be expressed as protein and then selected or screened for a desired property. Within the last 10 years, researchers have used DNA shuffling to create new genes that encode novel or improved proteins [8–11]. As an example, Zhao et al. developed the Staggered Extension Process or StEP to shuffle different sequences encoding the subtilisin E protein and improved its half-life at 65°C by 50-fold [12]. Briefly, they promoted recombination by using a modified polymerase chain reaction in which the annealing and extension steps were combined and allowed only brief periods of time for polymerization.

In this study, we use DNA shuffling to create unique spike glycoprotein genes from four different strains of IBV that represent four different serotypes using the StEP method. Examining S1 sequence recombinations following DNA shuffling ought to provide insights into coronaviral antigenic shift and thereby lead to predictions of future variant IBVs. In addition, it might be possible to create spike glycoproteins with multiple virus-neutralizing epitopes that could be used to develop a broadly-protective IBV vaccine.

Materials and Methods

Virus Strains

The virus strains used in this study were Massachusetts 41 [13], Arkansas DPI [14], Connecticut 46 [13], and Delaware 072 [15]. All viral strains were propagated by allantoic sac inoculation of 9–11 day-old embryonating chicken eggs. After 48 h of incubation, the allantoic fluid was harvested and stored at −70°C [16].

RNA extraction, RT-PCR, and cloning of parental genes

Viral RNA was extracted from allantoic fluid using the High Pure RNA extraction kit (Roche Diagnostics Corporation, Indianapolis, IN). The S1 gene was amplified by RT-PCR as previously described [17,18] using the Titan One Step RT-PCR kit (Roche Diagnostics Corporation, Indianapolis, IN). The 1700 bp amplicon from each virus strain was gel purified on a 1% agarose gel using GenElute spin columns (Sigma-Aldrich Co, St. Louis, MO) and Microcon 30 columns (Millipore, Bedford, MA) per the manufacturers directions. The resultant DNA was cloned into TOPO XL per the manufacturers directions (Invitrogen Corp., Carlsbad, CA). Plasmid DNA was extracted from colonies using a plasmid miniprep kit (Qiagen, Inc., Valencia, CA). Plasmids containing inserts of the appropriate size were screened by DNA sequencing of the 5′ and 3′ ends. Clones containing S1 genes with appropriate 5′ and 3′ end sequences were fully sequenced (SeqWright, Houston, TX). Plasmid DNA from each parental S1 gene was purified, linearized by digestion with MIu I (New England Biolabs, Inc., Beverly, MA) and gel purified on a 1% agarose gel as stated earlier, for use as StEP template.

StEP reaction

The staggered extension process was performed as previously descried [12]. Briefly, a 100 μl PCR was assembled containing 0.05 picomoles of each linear parental template, 40 picomoles of each S1 primer, 50 μl of a commercial PCR buffer mixture, Premix 7 (Epicentre, Madison, WI), and water up to 100 μl. The reaction mixture was subjected to a thermocycle program consisting of 95°C, 5 min; 100 cycles of 95°C, 30 s, 55°C, 2 s; and 4°C hold until the products could be analyzed.

Cloning of StEP recombinants

StEP products were electrophoresed on a 1% agarose gel. The diffuse band around 1700 bp was excised from the gel and purified as previously stated. The purified DNA was reamplified by PCR using 1 μl as template in the same reaction mixture as for the StEP. The thermocycle program was 95°C, 2 min; 25 cycles of 95°C, 30 s, 45°C, 30 s, 72°C, 90 s; and 4°C, hold. After the reaction, the 1700 bp amplicon was gel purified and cloned into TOPO XL according to the manufacturer's directions.

Characterization of StEP recombinant clones

Twenty-five colonies were randomly selected from the transformation reaction. Each clone was grown

overnight in LB broth in the presence of kanamycin. Plasmid DNA was prepared from each clone as stated earlier and the inserts were sequenced (Seq-Wright, Houston, TX). Sequence data was analyzed using MacDNASIS Pro V3.5 (Hitachi Software Engineering Corp., San Bruno, CA) and Lasergene V3.12 (DNASTAR, Inc., Madison, WI).

Results

Creating a Novel IBV S1 Gene Library Using the StEP and Initial Screening of 25 Clones

We used the S1 gene of the Massachusetts 41, Arkansas DPI, Connedticut 46, and Delaware 072 IBV strains, which represent four different serotypes, in the StEP and examined the reaction products after 20, 40, 60, 80, and 100 cycles on an agarose gel (Fig. 1). In total, 11 of the 25 clones examined contained recombinant S1 genes. Each of the recombinant genes had a unique sequence (Fig. 2) that maintained a full-length ORF from the ATG start site to the cleavage site. On average, each recombinant gene contained 5 crossovers. Seven of the recombinants (DS-6, DS-8, DS-9a, DS-9, DS-12, DS-21, DS-22) contained sequence from two of the parental genes, while four of the recombinants (DS-1, DS-2, DS-13, DS-16) contained sequences from three of the parental genes. None of the recombinants contained sequences from all four of the parental genes. There was very little Arkansas DPI parental sequence and no Delaware 072 parental sequence in any of the recombinants.

Nucleotide and deduced amino acid similarity of recombinants to parental genes

The nucleotide similarity for each of the 11 recombinants, when compared with the Massachusetts 41 strain, ranged from 91.3% to 99.6%. The nucleotide similarity for each recombinant when compared with the Connecticut 46 strain, ranged from 90.4% to 98.8%. The nucleotide similarity for each recombinant when compared with the Arkansas DPI strain, ranged from 75.7% to 82.8% (Table 1). The amino acid similarity for each of the 11 recombinants, when compared with the Massachusetts 41 strain, ranged from 89.9% to 99.3%. The amino acid similarity for each recombinant when compared with the Connecticut 46 strain, ranged from 87.6% to 97.4%. The amino acid similarity for each recombinant when compared with the Arkansas DPI strain, ranged from 75.1% to 81.4% (Table 1).

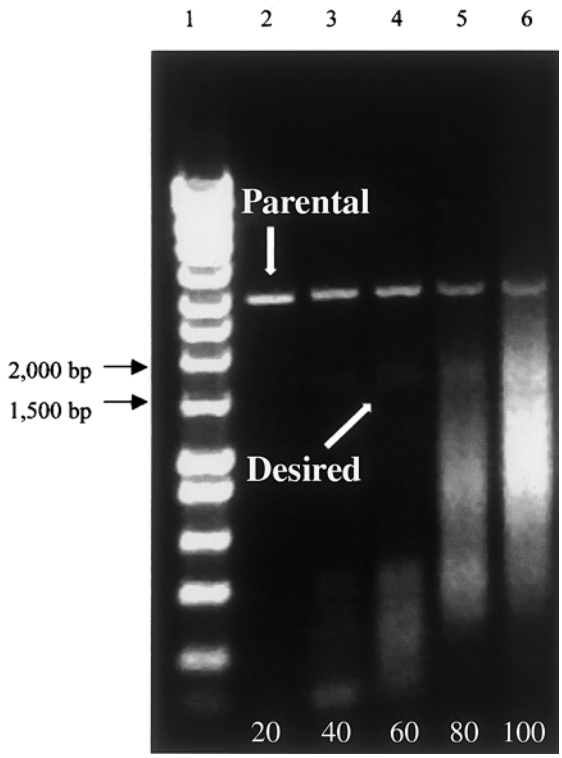

Fig. 1. Representative gel analysis of Staggered Extension Process products. Lane 1 = DNA ladder 1 (Gene Choice, PGC Scientifics Corp., Frederick, MD), from top to bottom: 10000, 8000, 6000, 5000, 4000, 3000, 2500, 2000, 1500, 1000, 800, 600, 400, 200; Lane 2 = Aliquot of StEP reaction after 20 cycles; Lane 3 = Aliquot of StEP reaction after 40 cycles; Lane 4 = Aliquot of StEP reaction after 60 cycles; Lane 5 = Aliquot of StEP reaction after 80 cycles; Lane 6 = Aliquot of StEP reaction after 100 cycles. The desired product is ~1700 bp. The bright band around 3000 bp is parental template. The smear below the desired product is incomplete amplification products. The numbers at the bottom of the gel indicate the number of reaction cycles in that lane.

Recombinant S1 deduced amino acid sequence analysis

The cleavage site sequence and the number of potential glycosylation sites were analyzed for each recombinant (Table 2). Each of the recombinants contained a cleavage sequence of the type X_1RX_2RR, where X_1 represents Arg, His, or Tyr and X_2 represents Ser, Phe, Ile, Arg, or His [22,23].

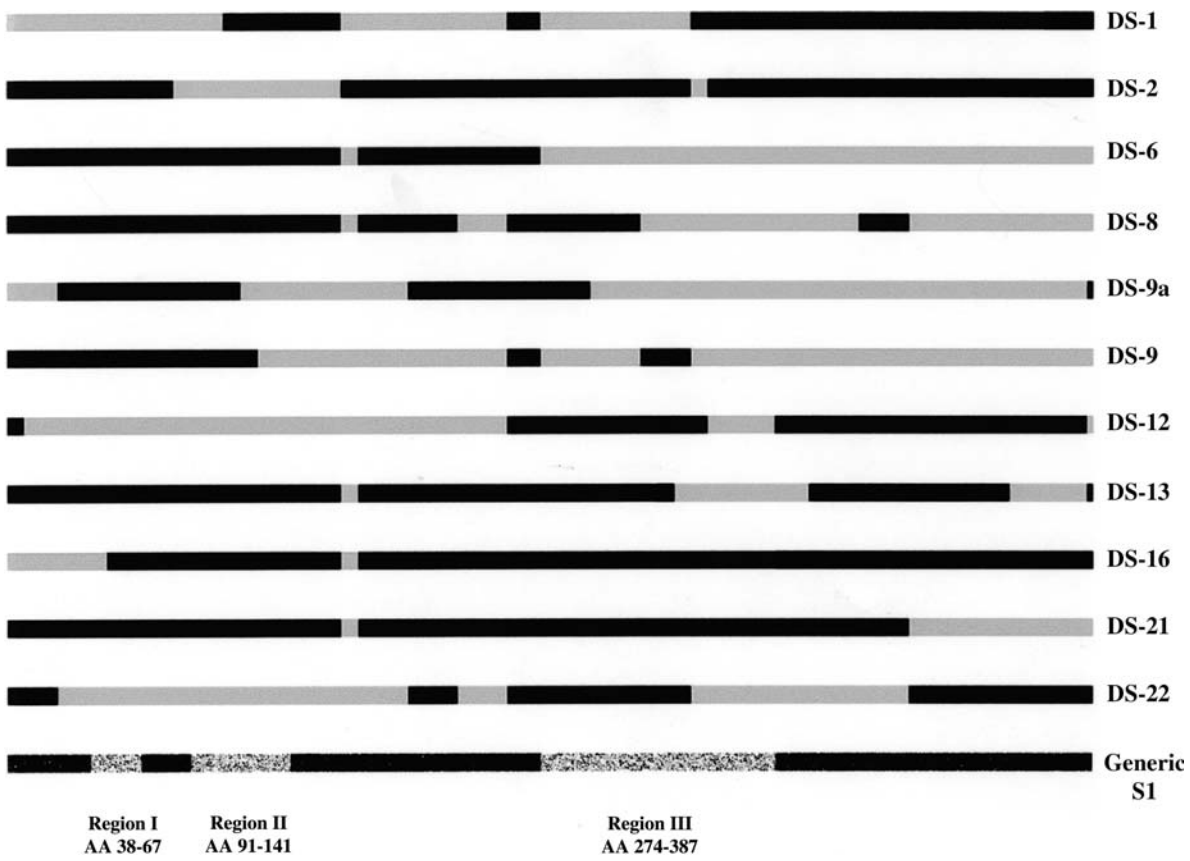

Fig. 2. Comparison of recombinant S1 gene sequences. Dark grey = Massachusetts 41 parental sequence; Light grey = Connecticut 46 parental sequence; Black = Arkansas DPI parental sequence. Shown at the bottom is a generic S1 gene with regions and their amino acid positions known to be involved in the formation of neutralizing and serotype specific epitopes (light grain) [19–21].

The number of potential N-linked glycosylation sites ranged from 17 to 18. When secondary structure predictions of the amino acid sequence for each recombinant were performed using the Chou and Fasman algorithm [24], and compared to the parental amino acid sequences, each recombinant was found to have a different secondary structure prediction (data not shown).

Recombinants containing sequences from three parental genes

The DS-1, DS-2, DS-13, and DS-16 recombinants contained parental sequence from Massachusetts 41, Connecticut 46, and Arkansas DPI (Fig. 3). The recombinant DS-1 had 6 crossovers and no point mutations. The recombinant DS-2 had four crossovers and four point mutations, of which three were non-synonymous changes. The recombinant DS-13 had 7 crossovers and 1-point mutation that was synonymous. The recombinant DS-16 had four crossovers and two point mutations, of which one was a non-synonymous change (Table 2).

Discussion

The StEP was used to produce recombinant S1 genes from four different strains of IBV, representing four different serotypes. We cloned the putative recombinant S1 genes and randomly selected 25 clones from the library for sequence analysis. Eleven of the clones contained recombinant S1 genes. All eleven recombinant clones (44% of total clones) contained a unique S1 sequence. This percentage might be improved by performing the StEP at a lower temperature

Table 1. S1 nucleotide and amino acid alignment pair distances

	Nucleotide–percent similarity															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14		
1		94.9	76.3	96.9	91.3	98.1	98.4	97.2	97.1	96.5	99.1	93.6	99.6	95.9	1	M41 S1
2	90.1		75.5	97.3	90.8	96.5	96.3	97.7	97.9	98.1	95.2	90.4	95.3	98.8	2	Conn46 S1
3	75.4	75.7		76.0	82.8	75.9	76.1	76.5	76.1	75.7	76.5	82.4	76.2	76.0	3	Ark DPI S1
4	94.0	95.3	75.3		90.1	96.3	96.2	95.3	95.7	97.8	96.3	92.1	96.9	96.6	4	DS-1 S1
5	89.9	89.3	81.4	87.3		91.1	91.2	92.2	91.9	91.3	91.1	96.9	91.3	91.8	5	DS-2 S1
6	96.3	93.1	75.2	93.6	89.2		99.3	98.3	98.6	95.3	98.4	93.3	98.6	96.3	6	DS-6 S1
Deduced amino acid–percent similarity																
7	97.6	92.1	75.2	93.1	90.1	98.3		98.3	98.4	95.7	98.8	93.4	98.9	96.6	7	DS-8 S1
8	94.6	95.5	76.7	91.6	91.6	96.1	6.3		99.1	96.7	97.5	92.6	97.5	97.6	8	DS-9a S1
9	94.4	95.7	76.4	92.5	91.1	97.0	96.1	98.0		96.4	97.4	92.2	97.5	97.3	9	DS-9 S1
10	92.7	96.6	75.1	95.3	90.1	90.4	91.8	94.0	93.1		96.1	90.8	96.6	98.6	10	DS-12 S1
11	98.7	90.6	75.8	93.4	89.9	96.8	98.1	95.2	95.0	92.5		93.3	99.1	96.1	11	DS-13 S1
12	93.9	87.6	81.0	90.8	94.6	92.7	93.7	91.4	90.9	88.8	93.5		93.6	90.9	12	DS-16 S1
13	99.3	90.8	75.4	94.4	90.3	97.0	98.3	95.0	95.2	93.1	98.7	94.2		95.8	13	DS-21 S1
14	92.3	97.4	75.7	93.4	91.4	92.3	93.3	95.5	94.6	97.4	92.9	88.8	92.3		14	DS-22 S1
	1	2	3	4	5	6	7	8	9	10	11	12	13	14		

and/or by decreasing the amount of time for annealing/extension. Lowering the temperature should increase the chance for recombination by slowing down the DNA polymerase and allowing for less similar sequences to hybridize, while decreasing the annealing/extension time would produce shorter fragments during each PCR cycle. Those changes should increase the number of unique recombinants, and might produce recombinants that contain sequences from all of the parental genes.

It was interesting to note that there was very little Arkansas DPI sequence and no Delaware 072 sequence in the recombinant genes. This may be due to the lower level of similarity between these two parental genes as compared to the Massachusetts 41 and Connecticut 46 S1 genes (Table 1). In fact, it may be impossible to recombine the Delaware 072 gene by the StEP because its S1 gene nucleotide similarity to the Massachusetts 41, Conneticut 46, and Arkansas DPI strains is extremely low (52.3%, 50.7%, and 48.7%, respectively). It is difficult to shuffle genes by the StEP with similarities below 70% (F.H. Arnold, California Institute of Technology, personal communication).

Table 2. Properties of parental and novel S1 gene recombinants

Clone name	Nucleotides/ Amino acids	Crossovers	Point mutations (non-synonmous changes)	Glycosylation sites	Cleavage sequence
Massachusetts 41	1611/537	NA	NA	17	RRFRR
Conneticut 46	1602/534	NA	NA	18	RRSRR
Arkansas DPI	1632/544	NA	NA	17	HRSRR
DS-1	1602/534	6	0	18	HRSRR
DS-2	1611/537	4	4(3)	17	HRSRR
DS-6	1611/537	3	2(2)	18	RRSRR
DS-8	1611/537	7	l(l)	17	RRSRR
DS-9a	1611/537	5	0	18	RRFRR
DS-9	1611/537	5	0	18	RRSRR
DS-12	1602/534	5	2(2)	17	RRSRR
DS-13	1611/537	7	1(0)	17	RRFRR
DS-16	1611/537	4	2(1)	17	HRSRR
DS-21	1611/537	3	0	17	RRSRR
DS-22	1602/534	6	2(1)	17	RRFRR

Previous research indicates that most natural recombination of infectious bronchitis viruses occurs in hot spots termed intergenic consensus sequences [5]. These areas are regions of similarity outside of the coding sequence for a gene. This normally leads to recombination of whole genes, not sections of genes, although some exceptions have been reported [2,3]. Our research has shown that even under near ideal conditions for recombination, recombinants from very different sequences within the S1 gene could not be produced. The Delaware 072 virus is unique in that it is less than 60% similar in the S1 gene to all other IBV isolates. Thus, it would seem logical to assume that natural Delaware 072 recombinants containing crossovers within the S1 gene coding sequence with other IBV types would be rare. Indeed, a recent report details a new serotype of IBV, designated GA98, which appears to have evolved from the Delaware 072 strain by antigenic drift and not by antigenic shift [6]. It appears that recombination at intergenic consensus regions, which replaces whole genes in IBV, is more likely to occur when the S1 genes of two strains are less than 70–75%, whereas antigenic shift could occur to produce variants within the S1 gene from the recombination of IBV strains with a S1 gene similarity above 70–75%. This finding is extremely important in light of the common practice of simultaneously vaccinating with different live attenuated strains of IBV to achieve broad protection in commercial chicken flocks. However, these percentages have been determined from our data and should be considered an estimation only, because *in vivo* RNA recombination for IBV may differ from the *in vitro* DNA recombination reported here.

Although we did not produce any recombinants with sequences from all four parental S1 genes, we did produce several that contained sequences from three of the parental S1 genes. In particular, recombinants DS-2 and DS-16 contained large S1 gene segments from parental strains, Massachusetts 41, Connecticut 46, and Arkansas DPI. The relatively low S1 amino acid sequence similarity for the DS-2 and DS-16 clones to any of the parental clones suggests that they may induce unique antibodies and makes them good candidates for further serologic characterization. Interestingly, the crossover sites for DS-2 ensure that different parental sequences reside within regions known to play key roles in the formation of virus neutralizing epitopes [19–21,25].

It is impossible to predict the antigenic outcome of a mixture of parental sequences from serologically distinct viruses within the S1 gene. Virus neutralizing epitopes for IBV are conformationally dependent [19–21] and it is not known whether hypervariable regions form neutralizing epitopes within themselves or with other hypervariable or constant regions. Furthermore, due to the known association of non-contiguous amino acids in the functional spike glycoprotein and the non-conservative point mutations within our recombinant S1 genes, the biological functionality and serological character of these genes remains to be determined.

References

1. King D.J. and Cavanagh D., in Calnek B.W. (ed.), *Diseases of Poultry,* Iowa State University Press, Ames, 1991, pp. 471–484.
2. Wang L., Junker D., and Collisson E.W., Virology *192*(2), 710–716, 1993.
3. Jia W., Karaca K., Parrish C.R., and Naqi S.A., Arch Virol *140*(2), 259–271, 1995.
4. Kottier S.A., Cavanagh D., and Britton P., Virology *213*(2), 569–580, 1995.
5. Lee C.W. and Jackwood M.W., Arch Virol *145*(10), 2135–2148, 2000.
6. Lee C.W. and Jackwood M.W., Virus Res *80*(1–2), 33–39, 2001.
7. Lee C.W. and Jackwood M.W., Virus Genes *22*(1), 85–91, 2001.
8. Stemmer W.P., Proc Natl Acad Sci USA *91*(22), 10747–10751, 1994.
9. Stemmer W.P., Nature *370*(6488), 389–391, 1994.
10. Soong N.W., Nomura L., Pekrun K., Reed M., Sheppard L., Dawes G., and Stemmer-W.P. , Nat Genet *25*(4), 436–439, 2000.
11. Miyazaki K., Nucleic Acids Res *30*(24), e139, 2002
12. Zhao H., Giver L., Shao Z., Affholter J.A., and Arnold F.H., Nat Biotechnol *16*(3), 258–261, 1998.
13. Jungherr, in Proc 60th Ann Mtg U.S.LS.A, 1956.
14. Gelb J. Jr., Perkins B.E., Rosenberger J.K., and Allen P.H., Avian Dis *25*(3), 655–666, 1981.
15. Gelb J. Jr., Keeler C.L., Nix W.A., Rosenberger J.K., and Cloud S.S., Avian Dis *41*(3), 661–669, 1997.
16. Senne D.A., in Swayne, D.E. (ed.), *A Laboratory Manual for the Isolation and Identification of Avian Pathogens,* American Association of Avian Pathologists: Kennett Square, 1998, pp. 235–240.
17. Kwon H.M., Jackwood M.W., and Gelb J. Jr., Avian Dis *37*(1), 194–202, 1993.

18. Lee C.W., Hilt D.A., and Jackwood M.W., Avian Dis *44*(3), 650–654, 2000.
19. Niesters H.G., Bleumink-Pluym N.M., Osterhaus A.D., Horzinek M.C., and van der Zeijst B.A., Virology *161*(2), 511–519, 1987.
20. Cavanagh D., Davis P.J., and Mockett A.P., Virus Res *11*(2), 141–150, 1988.
21. Moore K.M., Jackwood M.W., and Hilt D.A., Arch Virol *142*(11), 2249–2256, 1997.
22. Moore K.M., Bennett J.D., Seal B.S., and Jackwood M.W., Virus Genes *17*(1), 63–83, 1998.
23. Jackwood M.W., Hilt D.A., Callison S.A., Lee C.W., Plaza H., and Wade E., Avian Dis *45*(2), 366–372, 2001.
24. Chou P.Y. and Fasman G.D., Adv Enzymol Relat Areas Mol Biol *47*, 45–148, 1978.
25. Kant A., Koch G., van Roozelaar D.J., Kusters J.G., Poelwijk F.A., and van der Zeijst B.A., J Gen Virol *73*(Pt3), 591–596, 1992.