

# Full-length sequence analysis of four IBDV strains with different pathogenicities

Daniel Petkov · Erich Linnemann ·  
Darrell R. Kapczynski · Holly S. Sellers

Received: 2 December 2005 / Accepted: 7 February 2006 / Published online: 22 August 2006  
© Springer Science+Business Media, LLC 2006

**Abstract** Characterization of field isolate 9109, Lukert, Edgar cell culture-adapted (CCA), and Edgar chicken embryo-adapted (CEA) serotype 1 IBDV strains using full-length genomic sequences is reported. IBDV genomic segments A and B were sequenced and the nucleotide and deduced amino acid (aa) sequences were compared with previously reported full-length sequenced IBDV strains. We found that the viral protein VPX and amino acid sequences between aa 202–451 and 210–473 of VP2 but not the entire VP2 protein are the best representatives of the entire IBDV genome. The greatest variability was found in the VP2 and 5′ non-coding region of segment B among IBDV strains. The deduced amino acid sequences of the VP1 protein varies in length among the strains analyzed. The RNA-

dependent, RNA-polymerase motifs within VP1 and the VP5 protein were highly conserved among isolates. Although within the VP2 processing site, amino acid sequence of Lukert was similar to the classical while the Edgar CCA, and CEA were more similar to the very virulent strains, it was determined that these strains have sequence characteristics of the classical strains. In addition, close relatedness between Lukert, Edgar CCA and CEA was observed. Although phylogenetic analysis of the VP1, VP3, and VP4 proteins indicated that 9109 is a classical type virus, this isolate shares unique amino acid changes with very virulent strains within the same proteins. Phylogenetic analysis of the 3′ and 5′ non-coding regions of segment A revealed that 9109 is more similar to the very virulent strains compared to the classical strains. In the VP2 protein, several amino acids were conserved between variant E and 9109 strains. Thus, it appears that 9109 isolate has characteristics of classical, very virulent, and variant strains. Our analysis indicates that although VPX amino acid comparison may be initially useful for molecular typing, full-length genomic sequence analysis is essential for thorough molecular characterization as partial sequences may not designate a particular strain as very virulent, classical, or variant.

---

D. Petkov  
Department of Infectious Diseases, College of Veterinary  
Medicine, University of Georgia,  
Athens, GA 30602, USA

*Present Address:*  
D. Petkov  
Southeast Poultry Research Laboratory, Agricultural  
Research Services, USDA, 934 College Station Road,  
Athens, GA 30605, USA

E. Linnemann · H. S. Sellers (✉)  
Department of Population Health, Poultry Diagnostic and  
Research Center, College of Veterinary Medicine,  
University of Georgia, 953 College Station Road,  
Athens, GA 30602, USA  
e-mail: hsellars@uga.edu

D. R. Kapczynski  
Southeast Poultry Research Laboratory, Agricultural  
Research Services, USDA, 934 College Station Road,  
Athens, GA 30605, USA

**Keywords** IBDV · Sequencing ·  
Phylogenetic analysis

## Introduction

Infectious bursal disease has a great economic impact on the poultry industry worldwide [1]. Infectious bursal disease virus (IBDV) causes lymphoid depletion in the bursa of Fabricius, a primary lymphoid organ of

chickens. Infected chickens are immunosuppressed, can be predisposed to secondary infections, and respond poorly to immunization [2].

Two serotypes of IBDV have been identified [3], however, only serotype 1 viruses cause disease in chickens. Strains within serotype 1 differ in their pathogenicity and antigenicity, and are generally classified as very virulent, classical, or antigenic variants [4, 5].

IBDV is a non-enveloped, bi-segmented, double-stranded RNA virus belonging to the *Birnaviridae* family [6]. Segment A is approximately 3,260 nucleotides (nts) and contains two open reading frames (ORFs). The ORF1 (438 nts) partially overlaps ORF2 (3,039 nts) and encodes the viral VP5 protein [7]. The monocistronic ORF2 encodes a precursor polyprotein VPX–VP4–VP3 and is auto-catalytically cleaved at amino acid (aa) position 512–513 for VPX–VP4 and at aa position 755–756 for VP4–VP3. The VPX protein (aa 1–512) is processed to VP2 (aa 1–441) protein [8]. Most of the research has been focused on the VP2 protein as it is the primary determinant for cell tropism [9] and induces neutralizing antibodies [10]. It has been demonstrated that changes within the VP2 antibody binding region, at aa 206–350, can lead to antigenic variants [5, 11]. The group-specific VP3 protein induces non-neutralizing antibodies [12] and is involved in virus replication and genome packaging [13]. VP4 is a *cis*-acting viral protease that processes the polyprotein [14].

Segment B is approximately 2827 nts and encodes for VP1, the RNA-dependent, RNA-polymerase (RdRp). The protein plays a role in viral encapsidation [15] and primes the viral RNA synthesis [16]. In addition, VP1 initiates protein synthesis by resolving 5′ unique terminal repeats and recruits the 40S ribosomal subunit [17]. The RdRp activity is 3′ non-coding region (NCR) dependent on segment A and B. The 3′ NCR has a signal recognition function for replication, transcription, and translation [18]. The replication efficiency and virulence factors have been linked to segment B [19, 20]. Although, the VP1 protein of the very virulent (vv) IBDV strains form a distinct cluster separation from the classical strains [21], chimeric IBDV viruses containing segment A and B from classical and very virulent strains, respectively do not show increase of the viral virulence or pathogenicity [22].

The level of attenuation and the system used for viral propagation affects the pathogenicity of related viruses. The Lukert strain is a further attenuated Edgar cell-culture-adapted strain [23]. The Edgar, bursa-derived strain is a pathogenic field isolate [24] and compared with both Edgar CEA and Lukert strains is the most invasive and pathogenic [25].

Most of the current molecular techniques used for IBDV typing utilize amplification of a representative genomic sequence [26] of the VP2 fragment by reverse transcriptase (RT) polymerase chain reaction (PCR) followed by restriction fragment length polymorphism (RFLP) analysis [27]. The most widely used typing method is the amplification of fragment of VP2 gene between nucleotides with position 701–1444 of ORF2 using RT-PCR followed by digestion with restriction enzymes. Different strains can be grouped into genotypes based on restriction fragment length polymorphism (RFLP) patterns [28]. Based on segment-reassortment and other approaches it has been previously suggested that VP2 is not the only virulent determinant and that viral pathogenicity can be determined by genomic regions beyond this protein [4, 19, 20, 22, 29]. In addition the hypervariable region of the VP2 protein is highly genetically unstable during viral propagation *in vivo* [30].

We report the molecular characterization of four IBDV strains with different pathogenicities: field isolate 9109, Lukert, Edgar CCA, and Edgar CEA strains. The full-length genomes of these strains were sequenced and analyzed. Analysis of the nucleotide and deduced amino acid sequences of segments A and B were compared with previously reported full-length sequences from serotype 1 very virulent, classical, antigenic variants, and serotype 2 IBDV strains. Our analysis revealed that although amino acid sequence of VPX may be useful, complete genomic sequence information for strain characterization is essential for thorough molecular typing as partial viral sequences may not designate a particular strain as very virulent, classical, or variant.

## Materials and methods

### Viruses and virus propagation

The field isolate 9109 and the classical attenuated Lukert strain were kindly provided by Dr. Pedro Villegas (University of Georgia, Athens, GA). The 9109 isolate was isolated at the Poultry Diagnostic and Research Center (PDRC), University of Georgia from immunosuppressed broiler flocks showing subclinical signs of IBDV such as bursal atrophy, and minimal mortality. Previous typing of 9109 identified a unique RFLP pattern when compared to the reference strains used for genotyping at PDRC. Edgar cell-culture-adapted (CCA) (cat. No. 120-ADV-2001) and Edgar chicken-embryo-adapted (CEA) (cat. No. 124-ADV-9501) challenge strains were purchased from National

Veterinary Services Laboratory (Ames, IA). The 9109 isolate was propagated in 21 day of age (d.a.) specific pathogen free (SPF) chickens. At 72 hours post infection, bursas were aseptically collected, snap frozen in liquid nitrogen and stored at  $-80^{\circ}\text{C}$  until future use. Lukert and Edgar CCA were propagated in primary chicken embryo fibroblasts (CEF) derived from 9–11 d.a. SPF embryos as previously described [31].

#### RNA extraction

Total RNA was extracted directly from the Edgar CEA strain from the purchased virus without propagation using the RNeasy mini kit (Qiagen, Valencia, CA). Total RNA was extracted from Lukert or Edgar CCA infected chicken embryonic fibroblasts using the same kit. The cell culture supernatant was removed 48–72 hours post infection, the cell monolayer overlaid with 300  $\mu\text{l}$  RLT buffer containing 0.143 M  $\beta$ -mercaptoethanol (Bio-Rad Laboratories, Hercules, CA) and scraped, as recommended by the manufacturer. Total RNA was extracted from homogenized bursas harvested from chickens infected with the 9109 isolate. All samples of extracted RNA were incubated with Proteinase K (Qiagen) at  $55^{\circ}\text{C}$  for 30 min prior to RNA column purification.

#### Reverse transcription-polymerase chain reaction (RT-PCR) of segment A and B

The first strand cDNA was synthesized using modified, previously reported segment-specific primers: segACEF5 or segACEF3 for segment A, and segBCEF3 for segment B [32]. Additional primers for PCR and sequencing were designed as needed using PrimerSelect software (DNASTAR, Lasergene, v. 5, Inc., Madison, WI) and previously published sequences obtained from GenBank. Initially 4–5  $\mu\text{l}$  total RNA and the segment-specific primer were denatured at  $98^{\circ}\text{C}$  for 5 min and cooled at  $4^{\circ}\text{C}$  for 10 min. The first-strand cDNA was synthesized at  $50^{\circ}\text{C}$  for 30 min in a 20  $\mu\text{l}$  reaction containing 1x first-strand buffer, 0.01 M DTT, 0.5  $\mu\text{M}$  segment-specific primer, 40 units RNase-out (Invitrogen, Life Tech., Carlsbad, CA), 200 units SuperScript II, and 0.5 mM deoxynucleoside triphosphates mix (dNTP) (Invitrogen). The RT mixture was then incubated at  $70^{\circ}\text{C}$  for 15 min. Two units of RNase H<sup>-</sup> was added and incubated at  $37^{\circ}\text{C}$  for 20 min.

Segment A was amplified in two overlapping fragments with PCR primer pair segAVP25/segACEF3 and segAVP23/segACEF5 with an expected product size of 2691 and 1417 nts, respectively. Segment A positive RT-PCR controls were generated using the

primer pair segAVP25/segAVP23 with an expected size of 848 nts. The last 129 nts at the 5' end of segment A, Edgar CEA was amplified using the product from 5' RACE and primer pair segAVP23/segACEF5. Segment B was amplified with a single PCR reaction using the primer pair segBCEF3/segBCEF5 annealing to the extreme 3' and 5' ends, respectively, and yielding an expected product of 2827 nts. The PCR reaction was synthesized in a 100  $\mu\text{l}$  mixture containing 5  $\mu\text{l}$  of the RT reaction, 1x PCR buffer, 0.2 mM dNTP, 1.5 mM  $\text{MgCl}_2$ , primers at 0.5  $\mu\text{M}$  each and 2.5 units of recombinant or Platinum *Taq* DNA polymerase (Invitrogen). The PCR cycle parameters were as follows:  $94^{\circ}\text{C}$  for 3 min followed by 35 cycles at  $94^{\circ}\text{C}$  for 15 s,  $55^{\circ}\text{C}$  for 15 s,  $72^{\circ}\text{C}$  for 5 min, and a final extension at  $72^{\circ}\text{C}$  for 30 min. The RT-PCR amplified products were electrophoresed on a 0.8% agarose gel and visualized with ethidium bromide as per manufacturer's recommendations.

#### 5' RACE (rapid amplification of cDNA ends)

The sequences of the 5' termini of the viral genome were amplified using the 5' RACE system Version 2.0 (Invitrogen) as per the manufacturer's recommendations. Segment A and B specific primers were constructed with the PrimerSelect software (DNASTAR, Lasergene, v. 5) using the previously reported nucleotide sequence of CEF94 (GenBank accession AF194428 and AF194429, respectively) (The sequence of the primers is available upon request). First strand cDNA for segment A was synthesized using 4–5  $\mu\text{l}$  total RNA using SuperScript II reverse transcriptase, segment A specific reverse primer segA676. The reaction was carried out at  $98^{\circ}\text{C}$  for 10 min,  $4^{\circ}\text{C}$  for 5 min,  $50^{\circ}\text{C}$  for 50 min (reverse transcription),  $70^{\circ}\text{C}$  for 15 min, and a final incubation with an RNase mix at  $37^{\circ}\text{C}$  for 30 min. The first strand was S.N.A.P. column purified and an oligo-dC tail was added to the 3' end of the cDNA using terminal transferase and dCTP. The tailed cDNA was used as a template for polymerase chain reaction (PCR1) with abridged anchor (AAP) and segment A specific primer segA507. The second round of PCR (PCR2) was performed using 1  $\mu\text{l}$  of PCR1 product with abridged universal amplification (AUAP) and segment A specific primer segA323. Similarly the first strand cDNA for segment B was synthesized using segment B specific reverse primer segB671. The oligo-dC tailed segment B cDNA was then used as a template for PCR1 with AAP and the segment B specific primer segB460. One microliter of the PCR1 product was used in PCR2 with AUAP and segment B specific primer segB344. All RACE-PCR

reactions were carried out at 94°C for 3 min followed by 35 cycles at 94°C for 15 s, 55°C for 15 s, 72°C for 1 min, and final extension at 72°C for 30 min. The 5' RACE products were electrophoresed on a 2.0% agarose gel. Gel purification was performed with the QIAEXII-gel extraction kit (Qiagen).

#### Cloning and sequencing of RT-PCR products

The purified DNA was cloned into the pCR4-TOPO or pCRII-TOPO vectors according to the manufacturer's recommendations with maximum incubation times and transformed into competent *E. coli*, DH5 $\alpha$ -T1R cells (Invitrogen). Transformants were selected on Luria-Bertani (LB) agar media (Q-BIOgene, CA, USA) containing 50  $\mu$ g/ml kanamycin and screened by PCR with segment specific primers. Six clones from each PCR reaction were stored and three used for sequencing. Clones were expanded overnight at 37°C in 10 ml LB media containing 50  $\mu$ g/ml kanamycin, centrifuged at 4°C for 10 min at 800 g, resuspended in 600  $\mu$ l fresh LB media with 10% glycerol, and stored at -80°C until further use.

Plasmid DNA was purified using the Qiagen mini-prep kit (Qiagen) according to manufacturer's recommendations. Double-stranded DNA was sequenced by the dideoxy chain termination procedure using the automated ABI Prism 310, Genetic Analyzer (Applied Biosystems, ABI, Foster City, CA). Plasmid specific (M13 reverse and forward), as well as, IBDV segment specific primers were used for sequencing with primer walking. Clones obtained from 5' RACE were sequenced with primers segA323 and segB344 for segment A and B, respectively. A consensus sequence was obtained using three clones which were sequenced at least three times with forward and reverse primers.

#### Multiple alignment and phylogenetic analysis

The nucleotide sequences obtained from the four IBDV strains in this report were deposited in GenBank with accession numbers for 9109 (AY462027, AY459321), Lukert (AY918948, AY918947), Edgar CCA (AY462026, AY459320), and Edgar CEA (AY918950, AY918949).

Previously published full-length nucleotide sequences of segment A and B used in our analysis have the following GenBank accession numbers: serotype 1 very virulent isolates D6948 (AF240686, AF240687), UPM97/61 (AF247006, AF527040), HK46 (AF092943, AF092944), UK661 (X92760, X92761), OKYM (D49706, D49707), serotype 1 classical isolates

CEF94 (AF194428, AF194429), CEF (AF493979), P2 (X84034, X84035), wild type Cu1(wt) (AF362747, AF362748), Cu1 (D00867, AF362775), segment B Winterfield-2512 (AF083092) and 002/73 (M19336), serotype 1 Delaware, variant E isolate (AF133904, AF133905) and serotype 2 isolates OH (U30818, U20950) and 23/82 (AF362773, AF362774).

The nucleotide and deduced amino acid sequences were aligned with the previously published full-length IBDV sequences using ClustalW multiple sequence alignment program and the pair distance calculated using Lasergene, v. 5 (DNASTar). Unrooted cladograms were generated using maximum parsimony analysis with 1000 bootstrap replicates using the Phylogenetic Analysis Using Parsimony software (PAUP, v.10). The 5' and 3' NCR of segment A and B were analyzed using the Mfold software (v. 3) [33].

## Results

There was a high nucleotide and amino acid similarity between IBDV strains used in this study. The amino acid sequences had higher similarities than nucleotide sequences. The similarity between both nucleotide and amino acid sequence of all strains analyzed for the VPX-VP4-VP3 polyprotein was representative for VP1, VPX, VP2, VP3, VP4, and VP5 genes (Table 1). The VPX-VP4-VP3 amino acid sequence similarity for the Edgar CCA and CEA was 96.6% and for Lukert and Edgar CEA and CCA was 96.4% and 96%, respectively. Lukert, 9109, and Edgar CEA were more similar to Cu1 strains 97.3%, 97.1%, and 96.5%, respectively than to Edgar CCA strain.

The similarity between amino acid and nucleotide sequences of Lukert and the classical strains Cu1, Cu1wt, and P2 was higher within VP4, VP5, and the VP2 processing site at 100% and 98.6%, respectively. The VP2 processing site in the 9109 isolate had 100% amino acid similarity to the classical strains Cu1, Cu1wt, CEF94, and P2. Within the same site Edgar CEA was more similar to Lukert and very virulent strains at 90.4% compared to the other IBDV strains.

#### Segment A analysis

Segment A in all four strains was 3260 nts in length. ORF1, encoding the VP5 protein, contains 438 nts and encodes 145 amino acids in the four sequenced strains with the start codon at nucleotide position 97 and stop codon at 543 nt. The amino acid substitution at position 16 (D  $\rightarrow$  A) is conserved between Lukert, Edgar



**Table 1** Comparison of deduced amino acid and nucleotide sequences of the VPX–VP4–VP3 polyprotein

% Nucleotide identity	Strain	% Amino acid identity										
		1	2	3	4	5	6	7	8	9	10	11
1	9109	<b>100</b>	94.4	94.6	95.7	96.0	96.7	97.1	97.1	96.8	96.7	88.6
2	Edgar CCA	95.0	<b>100</b>	96.6	96.4	94.9	95.4	95.9	96.1	95.7	94.5	87.6
3	Edgar CEA	95.3	97.9	<b>100</b>	96.0	95.5	95.9	96.2	96.5	96.5	94.7	87.8
4	Lukert	95.8	97.9	97.5	<b>100</b>	96.0	96.5	97.1	97.3	97.1	95.7	88.8
5	vvUPM9761	94.8	93.8	94.2	94.6	<b>100</b>	98.9	97.2	97.3	97.5	95.8	89.0
6	vvOKYM	95.1	94.0	94.5	94.8	98.4	<b>100</b>	97.7	97.8	98.0	96.4	89.6
7	P2	96.8	96.2	96.3	96.9	95.4	95.7	<b>100</b>	99.4	98.7	97.1	90.1
8	Cu1	96.7	96.1	96.3	96.9	95.3	95.6	99.7	<b>100</b>	98.6	97.1	90.1
9	Cu1wt	96.5	96.1	96.4	97.0	95.4	95.7	97.9	97.7	<b>100</b>	96.9	89.7
10	Variant E	97.5	95.3	95.5	96.2	95.0	95.3	97.0	96.9	96.8	<b>100</b>	88.4
12	OH	83.0	82.7	82.5	83.3	82.9	82.9	83.7	83.6	83.6	83.0	<b>100</b>

CCA, CEA and serotype 2 OH and 23/82 strains. The second substitution in Edgar CEA is at amino acids at position 135 (H → R). VP5 in very virulent strains contains four additional amino acids at the start (MLSL) compared to the rest of the strains analyzed (Table 2, Fig. 1).

The ORF2, encoding for the polyprotein, contains 3,039 nts and encodes 1012 amino acids. The polyprotein start codon is at nucleotide position 131 and stop codon at 3169 nt. Comparison of the polyprotein nucleotide sequences revealed that isolate 9109 has an insertion of C at position 2178 and deletion at 2209 nt. The VPX protein consists of 512 amino acids, VP4 consists of 243 amino acids, and VP3 consists of 257 amino acids in the four sequenced strains.

Multiple alignment of the predicted amino acid sequence of the VPX protein revealed several amino acid substitutions and potential antigenic regions (manuscript in preparation). Analysis of the VP3 protein revealed two unique substitutions 785 (I) and 993 (P) in Lukert, Edgar CCA and CEA (Table 2, Fig. 2). The 9109 isolate has unique substitution 1005 (A) also present in OKYM and vvUPM9761. Several other substitutions, 922 (Q) in Lukert and 858 (F), 922 (Q), and 948 (A) in Edgar CCA were also revealed.

Analysis of the VP4 protein revealed that the following unique amino acid substitutions: 14 in 9109, 7 in Edgar CEA, and 1 in Edgar CCA (Table 2, Fig. 3). The 9109 isolate has substitution at 541 (I), present in OKYM, variant E, and OH strains, 547 (D) within motif I, 680 (Y) present in vvUPM9761 and OKYM, and 686 (I) present in Edgar CCA and variant E, and a unique extensive (SQSTRLGQA) sequence with residue position 726–734). Edgar CEA has amino acid substitutions between serine-protease motifs I (aa 539–550) and II (aa 583–597) at position 576 (R), RIRPF with residue position 578–582 and in motif II at position 584 (E → G). No unique substitutions were found within motif III (aa 644–661), substrate-binding motif IV (aa 697–705), catalytic triad with aa positions 546 (H), 589 (D), 652 (S), and catalytic dyad with aa positions 652 (S) and 692 (K) [14, 33].

Segment B analysis

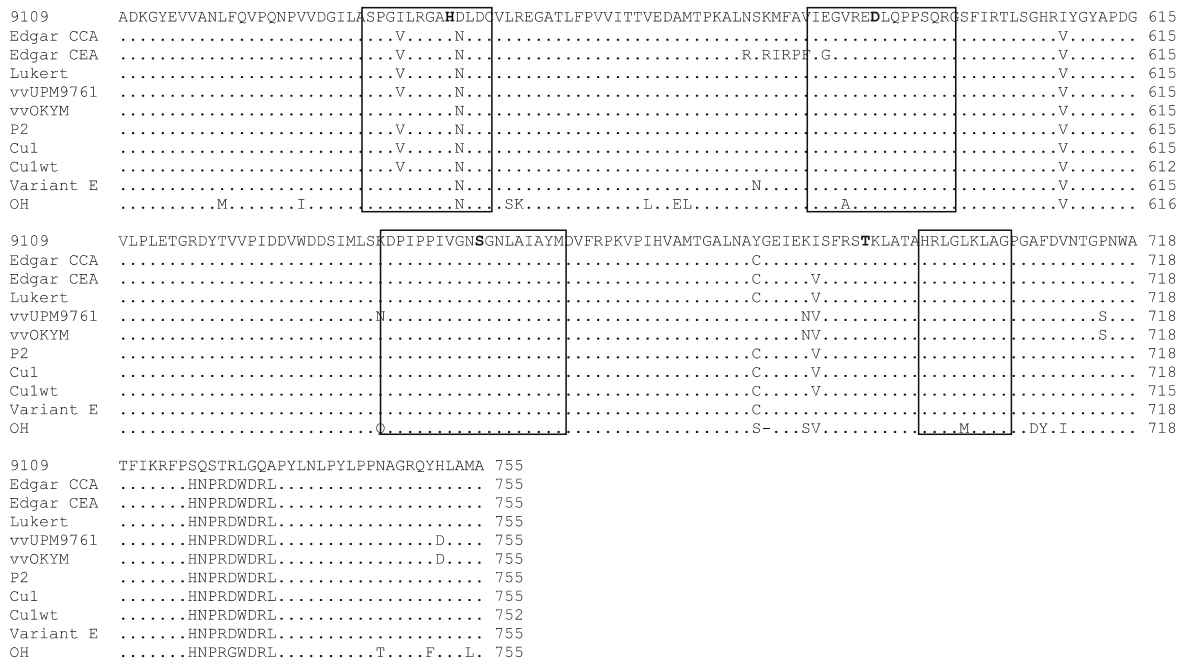
Segment B in Lukert, 9109, Edgar CCA and CEA strains was 2827 nucleotides (nts) in length. Edgar CCA has a coding region of 2646 nts (112–2757) and encodes for VP1 protein containing 881 amino acids (aa). The 9109 isolate and Edgar CEA have coding

**Table 2** Total amino acid substitutions in the four sequenced IBDV strains within segment A and B coding regions and comparison with the other serotype 1 and 2 strains used in this study

Strain	Segment A						Segment B	
	ORF1		ORF2				Substitutions	Additions
	VP5	VP2 (1–441) <sup>a</sup>	VP2 (442–512)	VP4 (513–755)	VP3 (756–1012)	Total		
9109	0	10	0	14	2	26	5	–
Edgar CCA	1	16	15	1	5	38	1	2
Edgar CEA	2	16	8	7	2	35	1	–
Lukert	1	18	1	0	3	23	5	2

<sup>a</sup> The start site of the VP2 is at amino acid (aa) 1 from ORF2, VPX–VP4–VP3 polyprotein of segment A. The VPX–VP4 is at aa 512–513 and the VP2 processing site is at aa 442–512. The amino acid positions are based upon the cleavage site for isolate 9109





**Fig. 3** Multiple alignment of the predicted amino acid sequence of VP4 protein of IBDV Lukert, 9109, Edgar CCA and CEA strains with corresponding sequences of published IBDV serotype one very virulent (vv) UPM9761 and OKYM, classical P2, Cu1, Cu1wt, variant E, and serotype two, OH strain. The start site of the VP4 is at amino acid (aa) 513 from ORF2 of

segment A. VP4 motifs I (aa 539–550), II (aa 583–597), III (aa 644–661), and IV (aa 697–705) are indicated with boxes. The amino acids sequence identical to the 9109 field isolate are indicated with dots and non-identical amino acids are presented with letters. The amino acid positions are based upon the 9109 isolate

The 9109 isolate has one unique substitution at 809 (M) and amino acids 13 (T) and 546 (L) are conserved in P2 and Cu1wt, 682 (R) is conserved in Winterfield and variant E. In addition, at amino acid position 147 field isolate 9109 has (D) and very virulent strains have (N) compared to the other strains.

The 5' and 3' non-coding regions (NCR) sequences

The nucleotide sequences of the 5' and 3' NCR of the four sequenced IBDV strains were compared to the published NCR sequences of vvD6948, CEF94, and P2, segment A and vvD6948, CEF94, P2, Cu1, and CEF, segment B IBDV strains. The 5' NCR of segment A in all strains is 97 and 131 nts for ORF1 and ORF2, respectively, while segment B has 112 nts. The 3' NCR of segment A has 91 nts in all strains, however the length of segment B differs: 78 nts in P2, 75 nts in vvD6948, 9109, Edgar CEA, and CEF94, and 70 nts in Edgar CCA and Lukert. Segment A in Lukert has mutation at position 75 (U → C) and Edgar CEA at nucleotide position 3257 (C → A). The 18S rRNA binding motif in segment A [34] and the 5' NCR of segment A ORF2 promoter region – 131 to –100 nts [35] were highly conserved among all strains analyzed.

The unrooted phylogenetic analysis

Unrooted phylogenetic analysis using maximum parsimony analysis with 1000 bootstrap replicates was performed with the aligned segment A and B deduced amino acid sequences. Unrooted analysis allows identification of related descendants without explicitly designating a common ancestor. The analysis of the less divergent trees from VP1 (Fig. 4A) and RdRp motif (data not shown), VP4 (Fig. 5C), I–IV motifs (data not shown), and VP3 (Fig. 5D) demonstrated formation of two branches, very virulent in one branch and classical strains in the second branch. In the same trees, 9109 clustered with the classical strains. The phylogenetic analysis of the VP1, VP3, and VP4 identified a closer relationship between Lukert, Edgar CCA and CEA and classical strains with higher bootstrap probability. vvD6948 in VP3 clustered with classical strains (Fig. 3D). In the VP2 processing site, Lukert was similar to the classical strains and Edgar CCA and CEA were more similar to the very virulent strains (manuscript in preparation). In VP2, Edgar CCA and Lukert clustered together, Edgar CEA did not group with any of the clusters and formed a branch wherein it was sole the sole taxa and 9109 clustered with variant E (Fig. 4C). The 5' and 3' NCR of segment

**Table 3** Unique amino acid substitutions in VP1 protein

Strain	Amino acid position <sup>a</sup>																						
	4	12	13	46	61	69	75	90	92	104	121	125	136	145	146	147	150	160	163	170	172	174	192
9109	I	S	T	S	V	E	T	G	V	Q	P	P	Y	N	E	D	D	R	A	M	Q	T	P
Edgar CCA	.	.	K	.	.	.	.	.	.	.	T	.	.	.	.	G	.	.	.	.	.	.	.
Edgar CEA	.	.	K	G	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.
Lukert	.	.	K	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	E	.	.
vvUPM9761 <sup>b</sup>	V	.	K	.	I	.	.	.	.	.	.	.	T	D	N	.	K	P	.	.	.	.	S
vvOKYM <sup>b</sup>	V	Q	K	.	I	D	.	.	.	.	S	.	T	D	N	.	.	.	.	.	.	.	.
Winterfield	.	.	K	.	.	.	.	R	.	H	.	.	.	.	.	G	.	.	.	I	.	.	.
P2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	N
Cu1	.	T	.	.	.	.	.	.	A	.	.	.	.	.	.	G	.	.	.	.	.	.	.
Cu1wt	.	R	K	.	.	.	A	.	.	.	.	F	.	.	.	G	E	.	.	.	.	.	.
Variant E	.	.	K	.	.	.	.	R	.	H	.	.	.	.	.	G	.	.	.	I	.	.	.
OH <sup>c</sup>	V	T	K	.	.	.	.	.	.	.	A	.	.	.	.	G	.	.	.	.	.	.	.
	204	205	212	213	242	287	289	310	311	319	356	370	380	390	393	425	426	431	469	479	484	508	
9109	Q	L	V	G	D	T	L	Y	W	F	M	P	G	L	E	T	R	A	S	T	S	R	
Edgar CCA	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Edgar CEA	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Lukert	.	.	.	.	.	.	.	.	.	V	.	.	.	.	.	.	.	.	.	P	.	.	.
vvUPM9761	.	.	.	.	E	A	.	.	.	.	.	.	M	D	.	.	.	.	.	.	.	.	K
vvOKYM	.	.	.	E	A	.	.	.	.	T	V	M	V	I	S	V	.	.	.	.	.	.	K
Winterfield	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	M	.
P2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Cu1	H	V	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Cu1wt	H	V	.	.	.	.	F	G	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Variant E	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	M	.
OH	.	.	G	S	.	.	Q	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	K
	511	515	527	546	562	617	640	646	650	667	675	677	682	685	687	695	697	718	720	722	788	789	
9109	R	E	N	L	S	A	A	G	R	S	A	E	R	V	S	K	V	N	L	T	K	V	
Edgar CCA	.	.	.	P	.	.	.	.	.	.	.	.	K	.	.	.	.	.	.	.	.	.	.
Edgar CEA	.	.	.	P	.	.	.	.	.	.	.	.	K	.	.	.	.	.	.	.	.	.	.
Lukert	.	.	S	P	.	.	.	.	.	.	.	.	K	.	.	.	.	F	.	.	.	.	.
vvUPM9761	S	.	.	P	P	.	.	S	.	.	.	.	K	.	P	R	.	.	N	.	.	.	.
vvOKYM	S	.	.	P	P	.	.	S	.	S	V	K	.	P	R	A	.	.	.	.	.	.	.
Winterfield	.	D	.	P	.	G	.	.	.	.	.	.	I	.	.	.	S	.	.	.	.	.	.
P2	.	.	.	.	.	.	.	.	.	.	.	.	K	.	.	.	.	.	.	.	.	.	.
Cu1	.	.	.	.	.	.	.	.	.	.	.	.	K	.	.	.	.	.	.	.	.	.	.
Cu1wt	.	.	.	P	.	.	.	.	.	.	.	.	K	.	.	.	.	.	.	.	.	.	.
Variant E	.	D	.	P	.	G	.	.	.	.	.	.	I	.	.	.	S	.	.	.	.	.	.
OH	S	.	.	P	.	G	.	S	-	A	.	.	K	.	.	.	.	.	.	.	N	.	-
	793	809	831	871	879	880	881																
9109	A	M	H	R	Q	-	-																
Edgar CCA	.	V	.	.	.	Q	P																
Edgar CEA	.	V	.	.	.	-	-																
Lukert	.	V	.	.	.	Q	P																
vvUPM9761	.	V	.	.	.	-	-																
vvOKYM	.	V	R	.	.	-	-																
Winterfield	.	V	.	.	.	-	-																
P2	.	V	.	.	.	-	-																
Cu1	.	V	.	.	.	-	-																
Cu1wt	.	V	.	.	.	-	-																
Variant E	.	V	.	.	.	-	-																
OH	G	V	.	A	.	-	-																

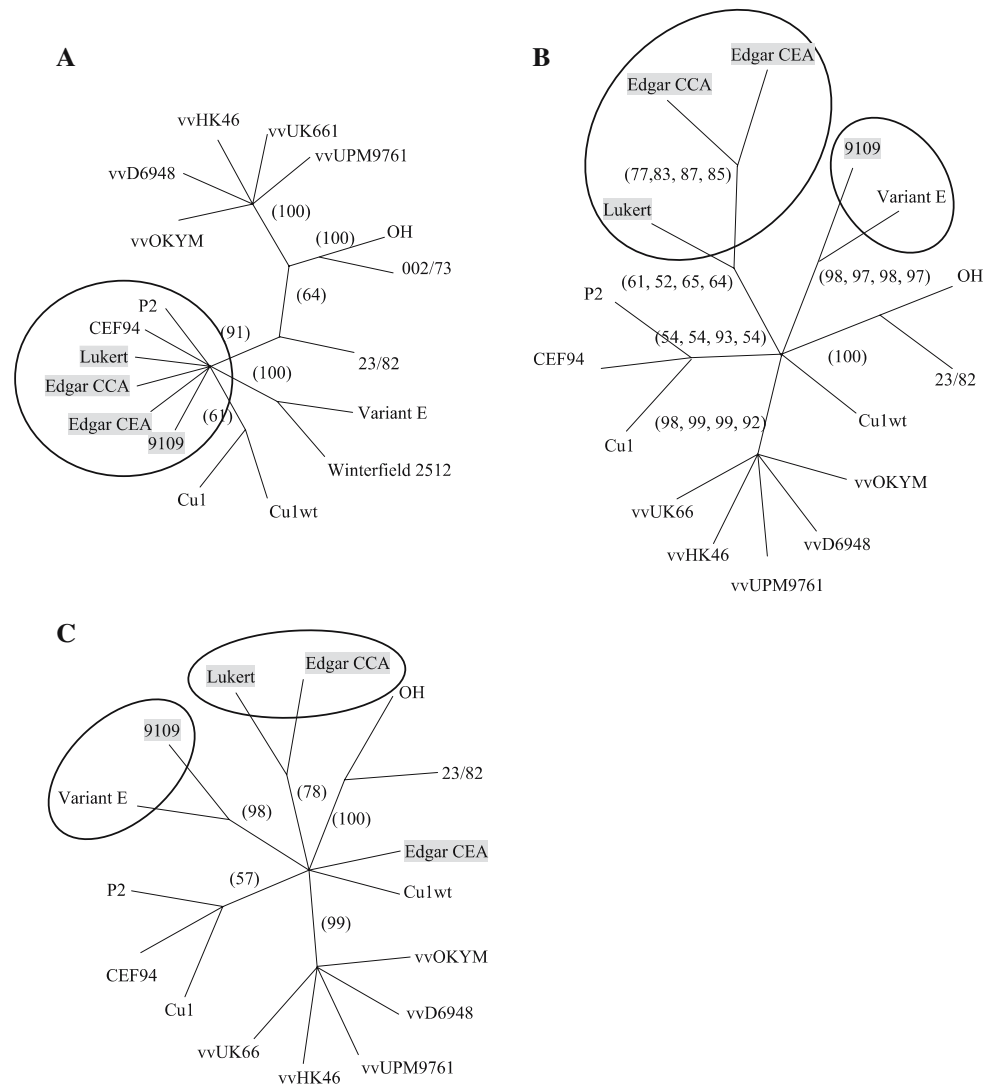
<sup>a</sup> The amino acid positions are based upon the 9109 isolate. The amino acids sequence identical to the 9109 field isolate are indicated with dots and non-identical amino acids are presented with letters. The dashes indicate amino acid deletion

<sup>b</sup> vvUPM9761 and vvOKYM represent all very virulent strains used in this analysis

<sup>c</sup> Not all amino acid substitutions between positions 100–135 in OH strain are included in the table



**Fig. 4** Phylogenetic relationship based on the amino acid (aa) sequences. The tree presented is representative of the independent phylogenetic analyses of the following segments: **(A)** VP1, **(B)** VP2 aa 202–451, VP2 aa 210–473, VPX aa 1–512, VP5–VPX–VP3–VP4–VP1, and **(C)** VP2 for Edgar chicken embryo adapted (CEA), Edgar cell culture adapted (CCA), Lukert, and 9109 (shaded boxes) with other IBDV strains. The cladograms were generated using maximum parsimony analysis with 1000 bootstrap replicates (confidence levels listed in parentheses) using the Phylogenetic Analysis Using Parsimony software (PAUP, v.10)



A 9109 isolate was found to be similar to vvD6948 (data not presented).

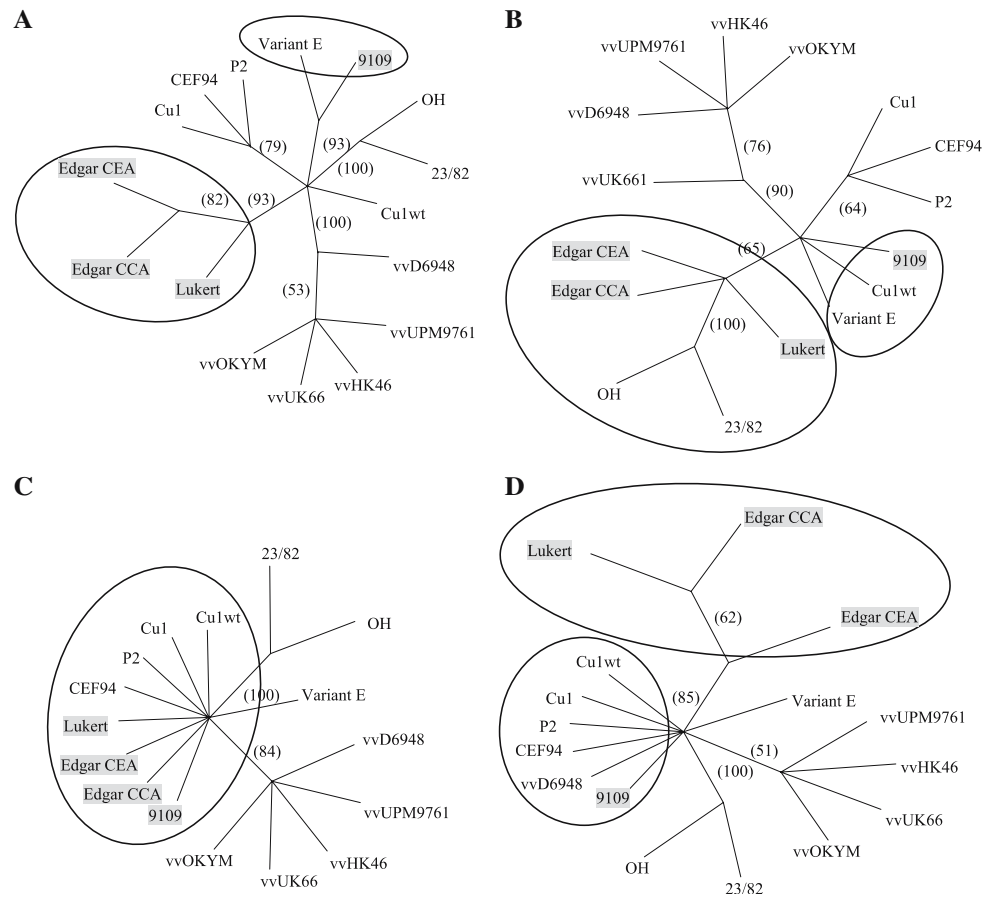
In the cladograms obtained from multiple alignments of the VPX–VP4–VP3, VP5 (Fig. 5A, B) and VP2, VP5–VPX–VP4–VP3–VP1, VPX the very virulent and serotype 2 strains (Fig. 4B, C) were represented in two distinct branches. In the same trees, 9109 and variant E clustered together with very high probability as predicted by a high bootstrap confidence level. In the same cladograms, Lukert, Edgar CCA, and CEA, and classical strains formed individual clusters. Edgar CEA and CCA were more similar to Lukert in VPX, VPX–VP4–VP3, VP5, and VP3 (Figs. 4B, 5A, B, D). The clustering pattern obtained from the VPX and amino acid sequences between amino acids at positions 202–451 and 210–473 cladograms represented the VP5–VPX–VP4–VP3–VP1 and VPX–VP4–VP3 (Figs. 4B, 5A).

## Discussion

We describe the full-length sequence characterization of four IBDV strains with different pathogenicities. The nucleotide and deduced amino acid sequences of these strains were compared with previously reported full-length sequenced serotype 1 and 2 IBDV strains.

Substitutions at amino acid position 785 (L → I) and 993 (Q → P) in Lukert, Edgar CCA, and CEA are in the VP3- and double-stranded (ds)RNA-binding domains, respectively [36]. The 9109 isolate has (Q) at amino acid position 981 in the dsRNA- and (A) at position 1005 in VP1-binding domain. In addition, it was determined that 1005 (A) is a vvIBDV characteristic and may be significant for the pathogenicity [37], however, 9109 is not a very virulent strain. Previously, the VP3 carboxyl terminus, aa 857–1012, was found to be important for vvIBDV pathogenicity [38] and

**Fig. 5** Phylogenetic relationship based on the amino acid (aa) sequences. The tree presented is representative of the independent phylogenetic analyses of the following segments: **(A)** VPX–VP4–VP3, **(B)** VP5, **(C)** VP4, I–IV motifs, and **(D)** VP3 proteins for Edgar chicken embryo adapted (CEA), Edgar cell culture adapted (CCA), Lukert, and 9109 (shaded boxes) with other IBDV strains. The cladograms were generated using maximum parsimony analysis with 1000 bootstrap replicates (confidence levels listed in parentheses) using the Phylogenetic Analysis Using Parsimony software (PAUP, v.10)



formation of the VP3–VP2 complex, which has a role in IBDV replication and genome packaging [13, 15].

Mutations in the VP4 protease may affect the poly-protein cleavage and may have a role in viral adaptation and pathogenicity. Substitutions in the VP4 protein at amino acid positions 541 (V → I) and 547 (N → D) in 9109 are within motif I and several substitutions in Edgar CEA are between serine-protease motifs I and II. The amino acid 680 (Y) observed in 9109 was previously proposed as very virulent marker [37].

Edgar CCA, CEA, and Lukert strains have substitutions within the VP5 protein at amino acid position 16 (D → A) indicating that this amino acid may not be important for viral pathogenicity since non-pathogenic serotype 2 strains have the same substitution. In addition Edgar CEA contains a second substitution at amino acid position 135 (H → R). The cytotoxic VP5 protein is an important for the viral release [39], induction of apoptosis [29], and pore formation in the cells [24]. Although previous analysis of VP5 did not reveal amino acids unique to the very virulent pathotype [40], amino acid substitutions could affect VP5 function. An IBDV VP5 knock-out mutant has been

shown to replicate in the bursa without bursal lesions and does not induce apoptosis [41] as does the wild type.

The stem-loop structure, not the sequence of the 3' NCR, is an important functional determinant [38] and plays a role in the protein-primed RNA synthesis [42] and RNA packaging [43]. The mutation at nucleotide position 3257, segment A in Edgar CEA is within the inverted terminal repeat 3255–3260 nts at the 3' NCR and part of the stem-loop structure with energy level –3.1 vs. –11.7 kcal/mole in the other sequenced strains which changes the stability of the stem-loop. The mutation at position 75 (U → C) in Lukert within the 5' NCR of segment A is in the putative 18S rRNA binding domain [34] and may affect the transcription and translation efficiency.

The viral protein VPX and amino acid sequences between aa at positions 202–451 and 210–473 of VP2 but not the entire VP2 protein are the best representatives of the entire IBDV genome. The greatest variability was found in the VP2 and 5' NCR of segment B among IBDV strains. The RdRp within VP1 and the VP5 protein were highly conserved among isolates.

Close nucleotide and amino acid relatedness between Lukert, Edgar CCA and CEA was established. Although Lukert, Edgar CCA and CEA strains have as expected sequence characteristics of the classical strains when, using the VP2 processing site, Edgar CCA and CEA could be classified as very virulent strains as shown in this analysis.

Phylogenetic and sequence analysis of field isolate 9109 revealed unique features. The isolate has characteristics of classical, variant, and very virulent strains depending on the region analyzed. The 9109 isolate is more similar to the classical strains in the VP2 processing site, VP3, and VP4 proteins. Although we did not find the unique *SspI* restriction site within the VP2 of 9109 isolate [44], (manuscript in preparation) previously suggested as a very virulent marker [45], we have identified two amino acid substitutions 680 (Y) in VP4 and 1005 (A) in VP3 which were also reported as vvIBDV markers. In addition, the amino acid at position 147 in the VP1 protein was substituted only in 9109 and very virulent strains. Phylogenetic analysis of the 5' and 3' NCR of segment A revealed similarity between 9109 and the vvD6948 strain. There are also several similarities between variant E and 9109 in VPX, VP2, and VPX–VP4–VP3 polyprotein. The predicted antigenic regions for 9109 isolate demonstrated similarities between 9109 and variant E (manuscript in preparation).

The four sequenced strains were determined by the phylogenetic analysis of VP1 protein to be classical strains. The segment B was highly conserved among analyzed strains. The VP1 protein in Lukert and Edgar CCA contains two more amino acids compared to 9109 and Edgar CEA strains. Similar variations have been previously described for P2, Cu1, Cu1M, and CEF94 strains [21] but the significance is still unknown. Edgar CEA has a substitution at amino acid position 46 (S → G) within the Ser-phosphorylation region [46]. At amino acid position 147 only 9109 and virulent strains have substitutions. Amino acids previously suggested to play a role in cell-specific replication [47], virulence [20], phosphorylation, glycosylation, NTP-binding motifs [48], and the RNA-dependent, RNA-polymerase (RdRp) motifs [46, 49] were highly conserved between all strains analyzed.

While the molecular basis for IBDV pathogenicity has not been fully defined, full length sequence analysis of viruses with different pathogenicities is an important factor in identifying potential predictors of viral pathogenicity. Finally, genomic analysis of IBDV at both the nucleotide and amino acid level is necessary for typing new field isolates to provide complete

pathogen data which can be correlated with virulence and pathogenicity.

**Acknowledgments** This project was funded by the Veterinary Medical Agricultural Research fund at The University of Georgia. The authors would like to thank Dr. Mark Jackwood for his review and comments.

## References

1. N. Etteradossi, (OIE, Paris, 1995) pp. 75–82
2. J.M. Sharma, I.J. Kim, S. Rautenschlein, H.Y. Yeh, *Dev. Comp. Immunol.* **24**, 223–235 (2000)
3. J. McFerran, M.S. McNulty, E.R. McKillop, T.J. Connor, R.M. McCracken, D.S. Collins, G.M. Allan, *Avian Pathol* **9**, 395–404 (1980)
4. T.P. Van den Berg, D. Morales, N. Etteradossi, G. Rivallan, D. Toquin, R. Raue, K. Zierenberg, M.F. Zhang, Y.P. Zhu, C.Q. Wang, H.J. Zheng, X. Wang, G.C. Chen, B.L. Lim, H. Muller, *Avian Pathol.* **33**, 470–476 (2004)
5. H.G. Heine, M. Haritou, P. Failla, K. Fahey, A. Azad, *J. Gen. Virol.* **72**, 1835–1843 (1991)
6. P. Dobos, J.B. Hill, R. Hallett, T.D. Kells, H. Becht, D. Teninges, *J. Virol.* **32**, 593–605 (1979)
7. E. Mundt, J. Beyer, H. Muller, *J. Gen. Virol.* **76**, 437–443 (1995)
8. N. Lejal, B. Da Costa, C.J. Huet, B. Delmas, *J. Gen. Virol.* **81**, 983–992 (2000)
9. E. Mundt, *J. Gen. Virol.* **80**, 2067–2076 (1999)
10. K.J. Fahey, K. Erny, J. Crooks, *J. Gen. Virol.* **70**, 1473–1481 (1989)
11. V.N. Vakharia, J. He, B. Ahamed, D.B. Snyder, *Virus Res.* **31**, 265–273 (1994)
12. H. Becht, H. Muller, H.K. Muller, *J. Gen. Virol.* **69**, 631–640 (1988)
13. M.G. Tacken, J.P. Rottier, L.A. Gielkens, B.P. Peeters, *J. Gen. Virol.* **81**, 209–218 (2000)
14. C. Birghan, E. Mundt, A.E. Gorbalenya, *EMBO J.* **19**, 114–123 (2000)
15. E. Lombardo, A. Maraver, R.J. Castn, J. Rivera, A. Fernandez-Arias, A. Serrano, L.J. Carrascosa, J.F. Rodriguez, *J. Virol.* **73**, 6973–6983 (1999)
16. U. Spies, H. Muller, H. Becht, *Virus Res.* **8**, 127–140 (1987)
17. A.C. Gingras, B. Raught, N. Sonenberg, *Annu. Rev. Biochem.* **68**, 913–963 (1999)
18. E.G. Strauss, J.H. Strauss, *Curr. Top. Microbiol. Immunol.* **105**, 1–98 (1983)
19. H.J. Boot, J.A. Hoekman, A.L. Gielkens, *Arch. Virol.* **150**, 137–144 (2005)
20. M. Liu, V.N. Vakharia, *Virology* **330**, 62–73 (2004)
21. M.R. Islam, K. Zierenberg, H. Muller, *Arch. Virol.* **146**, 2481–2492 (2001)
22. H.J. Boot, A.A. ter Huurne, J.A. Hoekman, P.B. Peeters, A.L. Gielkens, *J. Virol.* **74**, 6701–6711 (2000)
23. P.D. Lukert, J. Leonard, R.B. Davis, *Am. J. Vet. Res.* **36**, 539–540 (1975)
24. S.A. Edgar, Y. Cho, *Poult. Sci.* **52**, 492–497 (1973)
25. I.R. Rodriguez-Chavez, K.J. Rosenberger, S.S. Cloud, C.R. Pope, *Avian Pathol.* **31**, 485–492 (2002)
26. C. Le Nouen, G. Rivallan, D. Toquin, N. Etteradossi, *Arch. Virol.* **150**, 313–325 (2005)
27. N. Majo, J. El-Attrache, A. Banda, P. Villegas, A. Ramis, A. Pages, N. Ikuta, *Avian Dis.* **46**, 859–868 (2002)

28. D. Jackwood, S.E. Sommer, *Avian Dis.* **41**, 627–637 (1997)
29. K. Yao, V.N. Vakharia, *Virology* **285**, 50–58 (2001)
30. J. Smiley, D. Jackwood, *Avian Dis.* **45**, 1–8 (2001)
31. P. Lukert, R.B. Davis, *Avian Dis.* **18**, 243–250 (1974)
32. H.J. Boot, H.A. ter Huurne, B.P. Peeters, *J. Virol. Methods* **84**, 49–58 (2000)
33. D.H. Mathews, J. Sabina, M. Zuker, D.H. Turner, *J. Mol. Biol.* **288**, 911–940 (1999)
34. E. Mundt, H. Muller, *Virology* **209**, 10–18 (1995)
35. M.M. Nagarajan, F.S. Kibenge, *Arch. Virol.* **142**, 2499–2514 (1997)
36. M.G. Tacken, A.P. Van Den Beuken, P.B. Peeters, A.A. Thomas, J.P. Rottier, H.J. Boot, *Virology* **312**, 306–319 (2003)
37. L.L. Kong, R.A. Omar, M. Hair- Bejo, I. Aini, H.F. Seow, *Arch. Virol.* **149**, 425–434 (2004)
38. H.J. Boot, S.B. Pritz-Verschuren, *Nucleic Acids Res.* **32**, 211–222 (2004)
39. E. Lombardo, A. Maraver, I. Espinosa, A. Fernandez-Arias, J.F. Rodriguez, *Virology* **277**, 345–357 (2000)
40. M.F. Rudd, G.H. Heine, I.S. Sapats, L. Parede, J. Ignjatovic, *Arch. Virol.* **147**, 1303–1322 (2002)
41. K. Yao, A.M. Goodwin, V.N. Vakharia, *J. Virol.* **72**, 2647–2654 (1998)
42. P. Dobos, *Virology* **193**, 403–413 (1993)
43. W. Yao, K. Muqtadir, J.A. Bruenn, *J. Virol.* **69**, 1917–1919 (1995)
44. A. Banda, P. Villegas, J. El-Attrache, C. Estevez, *Avian Dis.* **45**, 620–630 (2001)
45. Z. Lin, A. Kato, Y. Otaki, T. Nakamura, E. Sasmaz, S. Ueda, *Avian Dis.* **37**, 315–323 (1993)
46. P.S. Shwed, P. Dobos, A.L. Cameron, N.V. Vakharia, R. Duncan, *Virology* **296**, 241–250 (2002)
47. M. Brandt, K. Yao, M. Liu, A.R. Heckert, V.N. Vakharia, *J. Virol.* **75**, 11974–11982 (2001)
48. R. Duncan, L.C. Mason, E. Nagy, A.J. Leong, P. Dobos, *Virology* **181**, 541–552 (1991)
49. A.E. Gorbalenya, M.F. Pringle, L.J. Zeddarn, T.B. Luke, E.C. Cameron, J. Kalmakoff, N.T. Hanzlik, H.K. Gordon, V.K. Ward, *J. Mol. Biol.* **324**, 47–62 (2002)