Arch Virol (2009) 154:495–499
DOI 10.1007/s00705-009-0325-9

BRIEF REPORT

Identification of sequence changes responsible for the attenuation of avian infectious bronchitis virus strain Arkansas DPI

Arun Ammayappan · Chitra Upadhyay · Jack Gelb Jr · Vikram N. Vakharia

Received: 1 November 2008 / Accepted: 14 January 2009 / Published online: 14 February 2009

Abstract Infectious bronchitis virus (IBV) is the causal agent of infectious bronchitis, which still remains one of the most important poultry diseases worldwide because of numerous serotypes and variants. A virulent strain of IBV, isolated from Arkansas (Ark), was propagated in embryonated eggs (Ark DPI 11). Following 101 serial passages in embryonated eggs, an attenuated strain of IBV was established (Ark DPI 101) that does not induce histopathological lesions in the tracheae of infected chicks. To identify sequence changes responsible for the attenuation of IBV, complete genome sequences of both virulent and attenuated Ark DPI viruses were obtained. Comparison of the genome sequences of the virulent and attenuated Ark DPI viruses reveals that these viruses are similar and differ only by 21 nucleotides, resulting in 17 amino acids changes. Most of those substitutions are located in the replicase 1a and spike genes. No differences were observed in gene 3, M or 5a, and only one nucleotide substitution each was present in 5b, N and 3′UTR. By comparing the deduced amino acid sequences of virulent and attenuated viruses, we identified sequence changes responsible for the adaptation and attenuation of the IBV-Ark DPI strain.

A. Ammayappan · C. Upadhyay · V. N. Vakharia (✉)
Center of Marine Biotechnology,
University of Maryland Biotechnology Institute,
701 E. Pratt St., Baltimore, MD 21202-3101, USA
e-mail: vakharia@umbi.umd.edu

A. Ammayappan
Department of Veterinary Medicine,
University of Maryland, College Park, MD 20742, USA

J. Gelb Jr
Department of Animal and Food Sciences,
College of Agriculture and Natural Resources,
University of Delaware, 531 South College Avenue,
Newark, DE 19716, USA

Avian infectious bronchitis virus (IBV), a member of the family *Coronaviridae*, order *Nidovirales* [6], is a highly infectious pathogen of domestic fowl. IBV is an enveloped virus that replicates in the cell cytoplasm and has a single-stranded, positive-sense RNA genome of 27.6 kb in size [2]. The IBV genome comprises ten open reading frames (ORFs). The ORF1 or replicase gene contains two overlapping open reading frames, ORF 1a and 1b [2]. The ORF 1b is produced as a fusion protein of 1a and 1b by -1 frameshift translation [3]. The IBV genome encodes four major structural proteins: the spike (S) glycoprotein, the small envelope (E) protein, the membrane (M) glycoprotein, and the nucleocapsid (N) protein [22, 23]. The spike protein is cleaved into S1 and S2, of which S1 produces neutralizing and serotype-specific antibodies [8, 19]. Because of the error-prone nature of RNA polymerase, coronavirus genomic RNA accumulates several point mutations during it replication, which leads to the emergence of new serotypes and variants [13]. In the case of IBV, most mutations occur in the spike glycoprotein, which is necessary for viral attachment and entry into host cells [11, 27, 28].

While it is possible to study the evolution of viruses and its impact on viral pathogenicity by comparing genomic sequences of heterologous strains, the analysis of homologous strains provides a unique opportunity to understand specific genes that are likely to be involved in viral pathogenicity. To identify specific sequence changes responsible for adaptation of the field virus to chick embryonic tissue and subsequent attenuation, we carried out comparative sequence analysis of the virulent Arkansas (Ark) DPI 11 (passage 11 in chick embryo) strain and its egg-adapted attenuated vaccine virus, Ark DPI 101 (passage 101 in chick embryo). Here, we have chosen Ark DPI as a model to determine the molecular basis of attenuation of IBV. This is the first report of comparative and complete genome sequence analysis of two homologous infectious bronchitis viruses to identify sequence changes responsible for adaptation to chick embryo and subsequent attenuation of IBV.

Chicken embryo passage numbers 11, 26, 51, and 101 of the Ark DPI strain were performed in Dr. Gelb's laboratory at the University of Delaware. Seed stocks of each passage number were prepared by inoculating 9-day-old specific-pathogen-free (SPF) embryonated chicken eggs and collecting allantoic fluid 72 h post-inoculation. Forty-one-day-old SPF leghorn chickens (SPAFAS, Inc., Norwich, CT) were assigned to 5 treatment groups of 8 birds each (Table 1). Chicks in groups 1 through 4 were inoculated intratracheally with $10^{4.5}$ embryo infectious dose$_{50}$ (EID$_{50}$) per chick of virus from each of the different passage numbers. Experimental inoculation of one-day-old chicks was carried out to evaluate the virulence of Ark DPI 11 and Ark DPI 101. The results of pathogenicity studies, summarized in Table 1, clearly demonstrate that the virulent IBV-Ark DPI strain is gradually attenuated after passage in chicken embryos.

Viral RNA was extracted from allantoic fluid seed stocks and stored at −20°C using the Qiagen RNAeasy kit according to the manufacturer's instructions. The RT-PCR and the cloning were carried out as described earlier [1]. DNAs from three independent clones were sequenced for each amplicon to exclude errors that can occur from RT and PCR reactions. The assembly of contiguous sequences and multiple sequence alignments were performed with the GeneDoc software [17]. The complete sequences of Ark DPI strain embryo passage numbers 101 and 11 have been submitted to GenBank with the accession numbers EU418975 and EU418976, respectively.

Table 1 Results of experimental inoculation of day-old chicks with IBV-Ark DPI strain obtained after different numbers of passages in embryos

IBV embryo passage number	Mortality (%)	Microscopic tracheal lesions[A] (%)	Mean body weight[B] (g)
11	8^a	100	242 ± 64^a
26	8^a	100	267 ± 24^a
51	0^b	20	277 ± 25^b
101	0^b	20	305 ± 22^b
Negative control	0^b	0	347 ± 12^b

[A] Presence of microscopic lesion in the trachea of chicks 6 days post-inoculation, when the most severe lesions are detected. Criteria for IBV infection are the observation of moderate-to-severe lesions in one or more of the following categories: deciliation, edema, epithelial hyperplasia, mucous gland hyperplasia, mucous gland exhaustion and mononuclear cell infiltration

[B] Body weights ± standard deviation on day 14 post-inoculation when differences between treatments are greatest

Within a column, values followed by different superscript letters are significantly ($p < 0.05$) different

The genomes of both viruses of Ark DPI consist of 27,620 nucleotides (nts) excluding the poly (A) tail and include ten ORFs flanked by 5′ (529 nts) and 3′ (507 nts) untranslated regions (UTRs). The genome organization of Ark DPI is 5′-Rep1a-Rep1b-S-3-M-5-N-3′, as shown in Fig. 1. In this comparative study, we found only 21 nucleotides differences between virulent and avirulent ArkDPI strains, which result in 17 amino acids changes (see Table 2). A single amino acid substitution was found in the p87 protein at nucleotide position 1,107. A similar kind of single amino acid change was reported previously in this coding region (nt 529–1,263) when challenge and vaccine viruses of the M41 strain were compared [14]. The role of the p87 protein is not clearly defined, but it may have a negative effect on PL^{pro}-mediated proteolytic cleavage at the p87/p195 site [30]. The amino acid substitution found in the region of PL1^{pro} at position 945 is unpredictable, because PL1^{pro}, is inactive in IBV [30]. Two nucleotide differences were observed in viral proteinase PL2^{pro} and one of them is silent. The amino acid substitution from acidic Asp to neutral Gly found in PL2^{pro} region could be considered an important one. This amino acid substitution is very close to the active catalytic site (nucleophile cysteine) of PL2^{pro} and could possibly interfere with proteolytic processing. Therefore, we speculate that this amino acid substitution may restrict viral maturation or replication. The amino acid substitution found in domain Y and p9 is difficult to predict because of role of both is unknown. A substitution of Ser to Pro was found in the virulent strain at nucleotide position 10,036 in the HD3 domain. Earlier, it was shown that some of the ORF1a-encoded hydrophobic domains are involved in membrane association of the replication complex of members of the *Nidovirales* [21, 24, 25]. Hence, this substitution might be critical for adaptation of virus by controlling the replication rate of the virus in different hosts. One amino acid substitution was found in the growth-factor-like (GFL) protein, which is involved in the growth factor signaling pathway [16]. The amino acid change from polar Thr to non-polar Ile may interrupt membrane association of this protein and thereby affect viral replication.

The replicase gene is usually not subjected to host immunity and is quite conserved in coronaviruses [15]. The main replicase proteins, RNA-dependent RNA polymerase (RdRp) and 3C-like cysteine protease (3CL^{pro}) or main proteinase (M^{pro}), were highly conserved, and not a single amino acid difference was noted. Virulent and attenuated

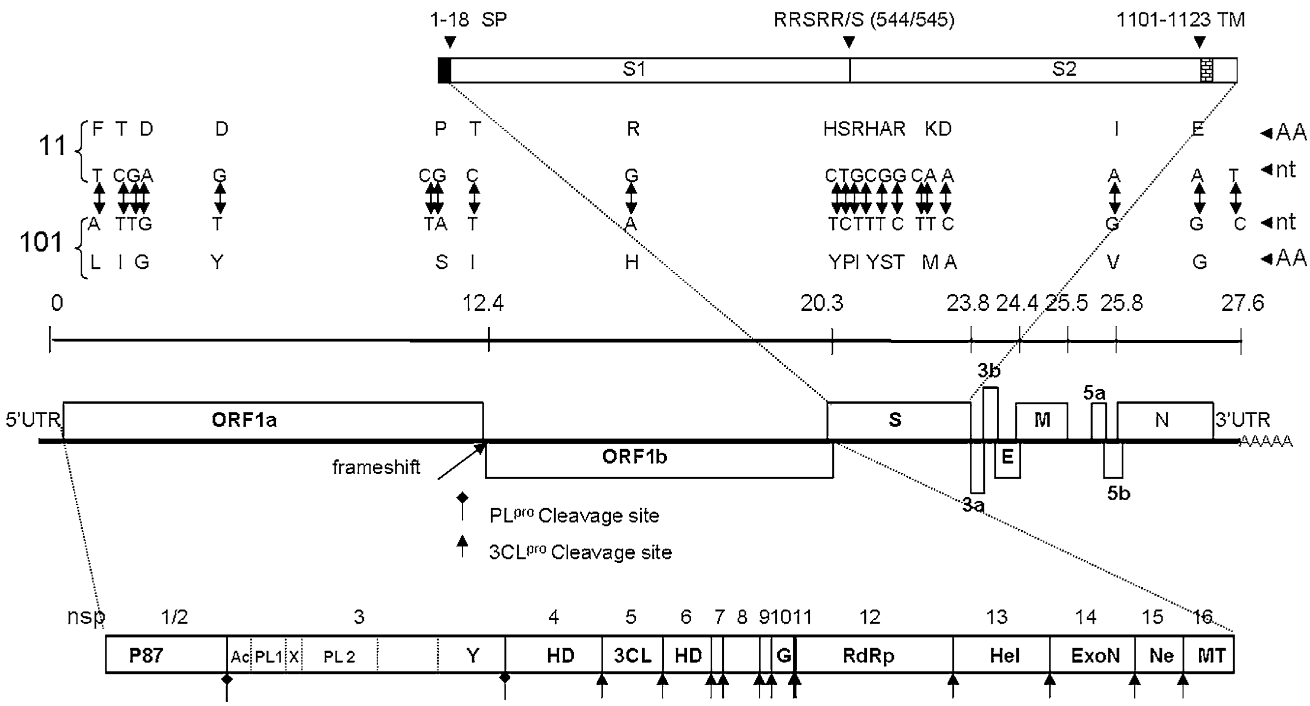

Fig. 1 Organization of the infectious bronchitis virus genome. The genome of Ark DPI is 27,620 nt long, excluding the poly (A) tract. *Middle* ten genes and their ORFs. The *scale* indicates the approximate positions and sizes of genes in the Ark DPI genome. *Bottom* putative domains of ORF1a/1b polyprotein: *nsp* non-structural protein; *Ac* acidic domain; *X* unknown domain X; *PL1* papain-like proteinase1; *PL2* papain-like proteinase 2; *Y* unknown domain Y; *HD* hydrophobic domain; *3CL* 3C-like proteinase; *G* growth-factor-like protein (GFL); *RdRp* RNA-dependent RNA polymerase; *Hel* helicase; *ExoN* exoribonuclease; *Ne* nidoviral uridylate-specific endoribonuclease; *MT* 2′-O-ribose methyltransferase. *Top* details of spike protein. *SP* signal peptide; *RRSRR/S* spike protein cleavage site between 544 and 545aa; *TM* transmembrane domain of spike protein. Nucleotide *nt* and amino acid *AA* differences between ArkDPI 11 and 101 and their approximate positions are depicted

strains differed by one amino acid in the helicase domain at nucleotide position 15,763. The amino acid change of Arg for His in the attenuated strain might significantly alter viral replication.

Among the structural genes, most of the nucleotide differences were located in the spike gene. Out of eight amino acid differences in the S protein, six were in the S1 region, located between amino acid positions 42 and 324. It has been shown that the S protein of coronaviruses is responsible for cell tropism [4, 12, 20]. Earlier workers predicted three hypervariable regions (HVR) in S1 of the spike protein [7, 15, 18] depending upon clustering of amino acid differences. In this study, we found a single amino acid change in each one of the HVRs. The changes were positive His to neutral Tyr in HVR I, polar Ser to nonpolar Pro in HVR II, and positive polar Arg to neutral nonpolar Ile in HVR III. Apart from HVRs, the region between residues 162 and 214 had three amino acid substitutions. Previous studies have shown that HVRs encode the serotype- and neutralization-specific epitopes, and the amino acid substitutions observed between Ark DPI 11 and Ark DPI 101 in the S1 region of the spike protein may have a similar function [7, 15]. Our findings are supported by a recent study, which revealed that a single passage of Ark DPI vaccine in a chicken led to selection of virus populations with an S1 gene that is similar to that of the virulent parental strain [26]. It is evident that the markers of virulence and adaptation reside mostly in the S1 protein. Of the two amino acid substitutions in the S2 region, one is located downstream and in the vicinity of the fusion peptide, and the other one is located in heptad repeat region 2. These mutations in S2 may alter the fusogenic properties of the S protein. The S1 undergoes more nucleotide changes than S2, which is quite conserved. But minimal changes in S2 are enough to alter the membrane fusion ability of the spike protein and thereby infectivity [9]. Interestingly, out of eight amino acid substitutions in S, six of the charged residues in Ark DPI 11 were mutated to neutral residues in Ark DPI 101. The two charged amino acids of Ark DPI 11 S2 were changed to membrane-interacting (hydrophobic) residues in Ark DPI 101. These residue changes in the S protein may contribute to adaptation of field virus to chick embryonic tissue and subsequent attenuation of the virus.

There was one amino acid difference between the attenuated and the virulent strain found in the 5b protein. The role of gene 5 in pathogenesis and replication is not

Table 2 Nucleotide and deduced amino acid differences between virulent and avirulent IBVs of Ark DPI strain

Substitution number	Nucleotide position	Ark DPI[a] 11	Ark DPI[a] 101	Amino acid substitutions[b]	Gene/protein involved[c]
1	1,107	T	A	F → L	P87
2	3,362	C	T	T → I	$PL1^{pro}$
3	4,188	G	T	Silent	$PL2^{pro}$
4	4,256	A	G	D → G	$PL2^{pro}$
5	5,967	G	T	D → Y	Unknown domain Y
6	10,036	C	T	P → S	HD3
7	10,650	G	A	Silent	P9
8	12,008	C	T	T → I	GFL
9	15,763	G	A	R → H	Helicase
10	20,437	C	T	H → Y	S1
11	20,665	T	C	S → P	S1
12	20,798	G	T	R → I	S1
13	20,821	C	T	H → Y	S1
14	20,947	G	T	A → S	S1
15	21,278	G	C	R → T	S1
16	22,344	C	T	Silent	S2
17	22,472	A	T	K → M	S2
18	23,420	A	C	D → A	S2
19	25,748	A	G	I → V	5b
20	27,101	A	G	E → G	N
21	27,580	T	C	Non-coding	3′UTR

[a] Nucleotide that differed between Ark DPI 11 and Ark DPI 101 at the corresponding position

[b] Single-letter codes of corresponding amino acids. → Amino acid change from Ark DPI 11 to Ark DPI 101

[c] *PL1*^{pro} papin-like proteinase 1; *PL2*^{pro} papin-like proteinase 2; *HD* hydrophobic domain; *GFL* growth-factor-like protein

clear, and it is considered non-essential for replication of virus [5]. Therefore, this amino acid difference in the 5b protein could be regarded as non-significant for viral attenuation. There was a notable nucleotide substitution found in the N protein gene, and one in the 3′UTR. In the attenuated strain, at nucleotide position 27,101 in the N gene, G was changed to A, and at the same time, at nucleotide position 21,580 in the 3′UTR, C was changed to T. This corresponding nucleotide change seems to be significant, because earlier studies demonstrated that the N protein binds very strongly to the extreme 3′ end of UTR [29]. The binding of the N protein to the 3′UTR is essential for synthesis of negative-strand viral RNA. It has been shown that the N protein interacts with the 3′UTR, but the sequence–specific interaction between the N gene and the 3′UTR is not clear [29]. The nucleotide substitutions found in N and the 3′UTR suggest that it may have an impact on viral replication and thereby on viral pathogenesis.

The role of the replicase gene of IBV in pathogenicity is not well understood. However, the amino acid changes in the ORF1a/1b proteins give an insight into putative residues that may be involved in the adaptation to chick embryonic tissue and subsequent attenuation of the virus. Although Ark DPI 11 and Ark DPI 101 are 99.92% similar in their nucleotide sequences, the pathogenicity of these viruses is entirely different. The spike protein is the major determinant of cell tropism in IBV, and the majority of nucleotide differences observed in the S1 gene in this study support and extend earlier observations [4]. The substitutions in the replicase proteins should be considered critical for their role in replication, and thus the pathogenicity of the virus. Even though only structural genes of IBV are known for affecting pathogenicity [4, 10], this study also suggests the involvement of the replicase gene.

Acknowledgments This project was supported by the National Research Initiative of the United States Department of Agriculture (USDA) Cooperative State Research, Education and Extension Service; Grant No. 2004-35204-14814.

References

1. Ammayappan A, Upadhyay C, Gelb J, Vakharia VN (2008) Complete genomic sequence analysis of infectious bronchitis virus Ark DPI strain and its evolution by recombination. Virol J 5:157
2. Boursnell MEG, Brown TDK, Foulds IJ, Green PF, Tomely FM, Binns MM (1987) Completion of the sequence of the genome of the coronavirus avian infectious bronchitis virus. J Gen Virol 68:57–77
3. Brierley I, Digard P, Inglis SC (1989) Characterization of an efficient coronavirus ribosomal frameshifting signal: Requirement for an RNA pseudo knot. Cell 57:537–547
4. Casais R, Dove B, Cavanagh D, Britton P (2003) Recombinant avian infectious bronchitis virus expressing a heterologous spike gene demonstrates that the spike protein is a determinant of cell tropism. J Virol 77:9084–9089

5. Casais R, Davies M, Cavanagh D, Britton P (2005) Gene 5 of the avian coronavirus infectious bronchitis virus is not essential for replication. J Virol 79:8065–8078
6. Cavanagh D (1997) Nidovirales: a new order comprising Coronaviridae and Arteriviridae. Arch Virol 142:629–633
7. Cavanagh D, Davis PJ, Mockett APA (1988) Amino acids within hypervariable region 1 of avian coronavirus IBV (Massachusetts serotype) spike glycoprotein are associated with neutralization epitopes. Virus Res 11:141–150
8. Cavanagh D, Davis PJ, Cook JKA, 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
9. Fang SG, Shen S, Tay FPL, Liu DX (2005) Selection of and recombination between minor variants lead to the adaptation of an avian coronavirus to primate cells. Biochem Biophy Res Comm 336:417–423
10. Gallagher TM, Buchmeier MJ (2001) Coronavirus spike proteins in viral entry and pathogenesis. Virology 279:371–374
11. Jia W, Karaca K, Parrish CR, Naqi SA (1995) A novel variant of avian infectious bronchitis virus resulting from recombination among three different strains. Arch Virol 140:259–271
12. Kuo L, Godeke GJ, Raamsman MJB, Masters PS, Rottier PJM (2000) Retargeting of coronavirus by substitution of the spike glycoprotein ectodomain: crossing the host cell species barrier. J Virol 74:1393–1406
13. Lai MM, Cavanagh D (1997) The molecular biology of coronaviruses. Adv Virus Res 48:1–100
14. Mondal SP, Cardona CJ (2004) Comparison of four regions in the replicase gene of heterologous infectious bronchitis virus strains. Virology 324:238–248
15. Moore KM, Jackwood MW, Hilt DA (1997) Identification of amino acids involved in a serotype and neutralization specific epitope within the S1 subunit of avian infectious bronchitis virus. Arch Virol 142:2249–2256
16. Ng LFP, Liu DX (2002) Membrane association and dimerization of a cysteine-rich, 16-kilodalton polypeptide released from the C-terminal region of the coronavirus infectious bronchitis virus 1a polyprotein. J Virol 76:6257–6267
17. Nicholas KB, Nicholas HBJ, Deerfield DW (1997) GeneDoc: Analysis and visualization of genetic variation. EMBNEW NEWS 4:14
18. Niesters HGM, Lenstra JA, Spaan WJM, Zijderveld AJ, Bleumink-Pluym NMC, Hong F, van Scharrenburg GJM, Horzlnek MC, van der Zeijst BAM (1986) The peplomer protein sequence of the M41 strain of coronavirus IBV and its comparison with Beaudette strains. Virus Res 5:253–263
19. Parr RL, Collisson EW (1993) Epitopes on the spike protein of a nephropathogenic strain of infectious bronchitis virus. Arch Virol 133:369–383
20. Sanchez CM, Izeta A, Sanchez-Morgado JM, Alonso S, Sola I, Balasch M, Plana-Duran J, Enjuanes L (1999) Targeted recombination demonstrates that the spike gene of transmissible gastroenteritis coronavirus is a determinant of its enteric tropism and virulence. J Virol 73:7607–7618
21. Snijder EJ, van Tol H, Roos N, Pedersen KW (2001) Nonstructural proteins 2 and 3 interact to modify host cell membranes during the formation of the arterivirus replication complex. J Gen Virol 82:985–994
22. Spaan W, Cavanagh D, Horzinek MC (1988) Coronaviruses: structure and genome expression. J Gen Virol 69:2939–2952
23. 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
24. van der Meer Y, van Tol H, Locker KJ, Snijder EJ (1998) ORF1a-encoded replicase subunits are involved in the membrane association of the arterivirus replication complex. J Virol 72:6689–6698
25. van der Meer Y, Snijder EJ, Dobbe JC, Schleich S, Denison MR, Spaan WJ, Locker KJ (1999) Localization of mouse hepatitis virus nonstructural proteins and RNA synthesis indicates a role for late endosomes in viral replication. J Virol 73:7641–7657
26. van Santen VL, Toro H (2008) Rapid selection in chickens of subpopulations within ArkDPI-derived infectious bronchitis virus vaccines. Avian Pathol 37:293–306
27. Wang L, Junker D, Collisson EW (1993) Evidence of natural recombination within the S1 gene of infectious bronchitis virus. Virology 192:710–716
28. Wang L, Junker D, Hock L, Ebiary E, Collisson EW (1994) Evolutionary implications of genetic variations in the S1 gene of infectious bronchitis virus. Virus Res 34:327–338
29. Zhou M, Williams AK, Chung S, Wang L, Collisson EW (1996) The infectious bronchitis virus nucleocapsid protein binds RNA sequences in the 3′ terminus of the genome. Virology 217:191–199
30. Ziebuhr J, Thiel V, Gorbalenya AE (2001) The autocatalytic release of a putative RNA virus transcription factor from its polyprotein precursor involves two paralogous papain-like proteases that cleave the same peptide bond. J Biol Chem 276:33220–33232