Virus Genes (2009) 38:56–65
DOI 10.1007/s11262-008-0282-5

Complete genome sequence analysis of a predominant infectious bronchitis virus (IBV) strain in China

Xing-li Liu · Jing-liang Su · Ji-xun Zhao · Guo-zhong Zhang

Received: 5 February 2008 / Accepted: 19 August 2008 / Published online: 3 September 2008

Abstract Infectious bronchitis (IB) is one of the major diseases in poultry flocks all over the world caused by infectious bronchitis virus (IBV). In the study, the complete genome sequence of strain A2 was sequenced and analyzed, which was a predominant IBV strain in China. The results indicated that there were mutations, insertions, and deletions distributed in the whole genome. The A2 virus had the highest identity to S14 and BJ in terms of full genome, whereas had a further distance to Massachusetts strains. Phylogenetic analysis showed that A2 isolate clustered together with most Chinese strains. The results of this study suggest that strain A2 may play an important role in IBV's evolution and A2-like IBVs are predominant strains in China.

Keywords Infectious bronchitis virus · Isolate · Complete genome · Phylogenetic analysis

Introduction

Infectious bronchitis virus (IBV) causes tremendous economic losses to the poultry industry in many countries. Domestic chicken is the most important but not the only host for IBV [1]. IB has been reported in peafowl, teal, partridge, turkey, pheasant, racing pigeon, and guinea fowl [2–5]. IBV is the prototype of *Coronaviridae*, which is a diverse family consists of three groups based on genetic and serologic properties [6, 7]. IBV belongs to *Coronaviridae* group 3, which differs extensively from group 1 and group 2 that are consisting of mammalian coronaviruses.

Infectious bronchitis virus is an enveloped positive-sense, single-strand RNA virus, with a large genome RNA about 27 kb in size [8]. The whole IBV genome has at least 10 open reading frames (ORF), from 5′ to 3′ are as follows: 5′-1a-1b-S(S1,S2)-3a,b,c(E)-M-5a,b-N-Poly(A)-3′, consisting of four main structural proteins (glycosylation spike glycoprotein (S), small envelope protein (E), membrane glycoprotein (M), and nucleocapsid protein (N)) and numerous non-structural proteins of known and unknown functions [1, 9–12]. S protein was cleavaged into N-terminal S1 and C-terminal S2 glycopolypeptides. The S1 spike glycoprotein carries virus-neutralizing and serotype-specific antigenic determinants, which can evolve rapidly especially within the three hypervariable regions (HVR). S2 protein anchors to viron and can affect S1 protein fusion to membrane. N protein plays important role in virus replication, assembly, and cell immunity. E and M proteins are required in virus assembly and budding. Gene 1 codes two polymerase proteins 1a and 1b, which are incised into 15 non-structural proteins (nsp) and are associated with RNA replication and transcription. NS proteins, encoded by gene 3 and gene 5, are not essential in virus replication, but they may play a role in antagonism of innate immune responses. They can also serve as targets for rational attenuation of IBV pathogenicity [1, 8, 13–16].

Protection for IBV infection is not satisfactory because of the numerous existed and newly emerged serotypes, poor cross-protection and the inefficient application of vaccine [17–24]. Currently, more than 60 serotypes have been reported as a consequence of mutation in its large genome. Choosing appropriate vaccine to control IB infection is very

X.-l. Liu · J.-l. Su · J.-x. Zhao · G.-z. Zhang (✉)
Department of Preventive Veterinary Medicine, College of Veterinary Medicine, China Agricultural University, Beijing 100193, People's Republic of China
e-mail: zhanggz@cau.edu.cn

important due to the poor cross-protection among the different serotypes. In North America, the mainly used vaccines are Massachusetts, Connecticut, and Arkansas serotypes. California strains and other important region serotypes are also used. In Europe, various "Holland variants" are available. In Australia, native B and C serotypes are used to control IB.

In China, IB was first reported in 1972 and the virus was found nearly all over the country in the following years. The disease has occurred frequently in chicken flocks and caused severe economic losses in recent years, although vaccines based on Massachusetts strains, such as H120 and H52, have been broadly used for prevention and control of the infection. Our previous study revealed that there are at least four groups of IBVs circulating in China and the disease outbreaks might have been caused by infection of multiple strains of IBV, but the predominant strain was correlative tightly with 4/91 serotype [21].

In this study, we sequenced the complete genome of IBV strain A2, which was an early isolate and closely related to strain 4/91 in neutralization test and the immune protection test. Sequence and phylogenetic analysis were carried out to compare A2 and other IBVs, which included the reference strains, the vaccine strains or the isolates from China reported in recent years. The results from this study will be useful to better understand the evolution of IBV in China and improve the efficacy of the vaccines for IBV infection in poultry industry.

Materials and methods

Virus amplification

Strain A2 was inoculated into the allantoic cavity of 10-day-old embryonated specific pathogen free (SPF) eggs. The embryos were incubated at 37°C and examined twice daily for their viability. The allantoic fluids were harvested from three eggs after incubation for 40 h and three other eggs were further incubated for 144 h to observe the lesions of embryos. Several blind passages were performed until the dwarfing and death of embryos were observed between 48 and 144 h after inoculation.

Primer design

Based on the cDNA sequences of IBVs from GenBank (M41, AY851295 and BJ, AY319651) and the sequenced A2 fragments (AY043312), 22 pairs of primers were designed and used to amplify the complete genome sequence of strain A2 except 5′ terminal. 5 primers were used for RACE PCR to amplify 5′ terminal sequence. Primer sequences were listed in Table 1.

Reverse transcription-polymerase chain reaction (RT-PCR)

Reverse transcription was carried out at 37°C for 1 h using 5 μl of total RNA, 1 μl random primer (Promega, USA) and 1 μl M-MLV reverse transcriptase (Promega, USA). PCR was performed in PCR Machine (Biometra, Germany) with 3 μl cDNA as template in a total of 25 μl reaction volume containing 0.5 μl of each primer and 12.5 μl PCR TaqMix (Meilaibo, China). Reactions were carried out at 94°C for 5 min followed by 35 cycles of 94°C for 45 s, 45–60°C (depending on the special PCR reaction) for 45 s, 72°C for 2 min, and finally, 72°C for 10 min.

5′ RACE of genome RNA

5′ terminal sequence of genome RNA was amplified using 5′-full race kit (TaKaRa, Japan). 5′ terminal cDNA was synthesized using 1 μl RT primer, 5 μl total RNA, and 1 μl AMV reverse transcriptase. After reverse transcription, the hybrided RNA was degenerated using RNase H and the cDNA were ligated with T4 RNA ligase at 16°C for 15–18 h. First PCR was carried out using 10-fold diluted ligated liquid as cDNA template. The first PCR product was 100-fold diluted and serve as template for the second PCR. Two PCR were carried out at 94°C for 5 min followed by 25 or 30 cycles of 94°C for 45 s, 50°C for 45 s, 72°C for 1 min, and finally, 72°C for 10 min.

Cloning and sequencing of the genome sequence

Polymerase chain reaction products were inserted into pGEM®-T Easy Vector (Promega, USA) and transformed into DH5α competent cell according to the manufacturer's instructions. Positive clones were screened from three different PCRs through blue and white cloney. The nucleotide (nt) sequences of the positive clones were determined using T7 and SP6 sequence primers on a commercial service.

Sequence analysis of the genome sequence

The sequences of IBVs were downloaded from NCBI in order to analyze the relationship between A2 and other strains. Nucleotide and amino acid (aa) identity of complete and partial genome were computed using DNAstar computer software. Phylogenetic and molecular evolutionary analyses were conducted using MEGA version 3.1 (Kumar, Tamura, Nei 2004).

Accession numbers of IBV sequences

The complete genome sequences used for comparisons are as follows: A2 (EU526388); Beaudette (NC_001451);

Table 1 Primers used for A2 complete genome sequence amplification

Primer	Location[a]	Upstream primer	Downstream primer	Length/bp
PCR				
366	366–1643	TCACCTCCCCCCACATACCTC	GACAGACA/TACACGCTCACAA/GCAAT	1278
1466	1466–2616	AATGACTTAACACTGCTTGA	TAATAA/GA/GCACAACATCAGGTG	1151
2172	2172–3045	TTGCACAGCATTGTTTTCA	TAATCATCATCAAGAGGCAGA	874
2846	2846–4380	TGCCATCTCATTTACCGC	TTTCCATCACGCCACTCA	1535
4162	4162–6039	GGGTCAATTTGGACAGGTT	TCTACGCTATAAGCATGTTTG/ATA	1878
5672	5672–7407	TTTGGGTTTACTAGGGAGC	CACTAAGAGCCATTTACAACAA	1736
7279	7279–9125	ATAACAAAGTCTGGTGCTAAACAA	CTG/ACTGACAACATTCAAAGTAACA	1847
8981	8981–10833	TGTGGTTGGGTGACTCTATC	ATTAACATCATCTGAGGCAAG	1853
10607	10607–11638	CACCTAAAACTGTGTGGGA	TGTGCTTGATCTACACCTG	1033
11583	11583–13011	CTTATGCCGCACGGTGTT	TTGCCATTGAGGTCTTGGTTAT	1429
12784	12784–14076	AGAAATACTTGTCACTTATGGTTGT	AACTGCCTATTAGTCATAGTAGAAAG	1293
13807	13807–15164	TGAATGTTATGAAGGTGGTT	AAGATTGTAGGGTTGTTGG	1358
14920	14920–16226	TGTGGCTGTTATGGAGCGTTA	TGTACTTTTTGCCTGTGTCATT	1307
16009	16009–18032	TCCTGGCAGTGGTAAGTCA	ATTGGGTACGATTGGGTT	2024
17810	17810–19606	ATTTTGTC/AAAGATGTCAACTGGGA	GCAAGGTTCCATAACACAATTCTG	1797
19525	19525–20602	ACCCACAACTTCAATCAGC	AGAACACCCACAGAACACG	1078
20336	20336–22180	ACTGAACAAAAGACMGACTT	CTGYTAGGTATGAGCACATT	1845
22050	22050–23896	TGGGGTTATGTTAGTTATGG	GACTTTTTAGGTCTGTTTTGTT	1847
23616	23616–25056	GGCGTTATACAGGGTCTT/CAATGACT	TATCTTTAGGCAAGTGGTCTGGTTC	1441
24773	24773–26252	CACAGGAGGTCTTGTCGCA	AAATACCAAGCATCTGGGACTG	1480
25975	25975–27496	ATGGCAAGCGGTAAGGCAACTG	TACCGTTCGTTTCCAGGCTACTAAG	1522
27305	27305–27715	TTAGTTGGTTTCCCTTGTTGG	TTTTTTTTTTTTGCTCTAACTCTAT	411
5′RACE PCR				
RT primer[b]		CTCCCAGATTACGGTCAAAC		
S1[c]		GTGATTTGTGGTGGTCTTGGAC		
A1[c]		CGGTTTCTGTAAGGGCTAGTTGA		
S2[d]		AGTGGAGTCCCCAACAAACC		
A2[d]		GCGACTACGAAAGCGAAAA		

[a] Primer locations were listed according to strain A2 (EU526388)

[b] 5′-phosphated RT primer used for 5′RACE

[c] The primers used for the first PCR in 5′RACE

[d] The primers used for the second PCR in 5′RACE

BJ (AY319651); California 99 (AY514485); M41 (AY85 1295); S14 (AY646283); KQ6 (AY641576); LX4 (AY33 8732); SAIBK (DQ288927). Partial gene sequences used for comparisons have the following GenBank accession numbers: LKQ3 (AY702085); LDT3 (AY702975); Jilin (AY84 6833, AY839144); HK (AY761141); Vic (DQ490221); S (DQ490213); Tw1171/92 (DQ646406). The accession numbers of all sequences used for phylogenetic analysis were included in Table 2.

Results and discussion

After four passages in 10-day-old embryonated SPF eggs, pathological changes in A2-infected embryos including dwarfing and death of embryos between 48 and 144 h post-inoculation were clearly observed (Figure not shown). The allantoic fluids were collected and used for RNA extraction.

The A2 strain was tested in Reverse transcription-polymerase chain reaction (RT-PCR) and 5′RACE PCR using the gene specific primers. 23 overlapping fragments were produced ranging from 411 to 2024 bp in size (Figure not shown).

To obtain the complete genome sequence of A2 strain, 23 sequenced large fragments were assembled using DNAstar software. The complete genome of strain A2 consists of 27,715 nucleotides (nt), which encodes six different genes. Gene 1 has a length of 20,446 nt, consisting of two overlapping ORFs: 1a and 1b. Gene 2, encoding S protein,

Table 2 The accession numbers of IBVs used for phylogenetic analysis

S1		**S2 (Continued)**		**N (Continued)**	
4/91 attenuated	AF093793	TW2575/98	DQ646405	Gx2-98	AY277721
A2	EU526388	Vic	DQ490221	H120	AY028296
Ark99	M99482	W93	EF602443	H52	AY044185
ArkDPI	AF006624	ZJ971	AF352314	HK	AY761141
Beaudette	NC_001451	**E**		IBN	AY856349
BJ	AY319651	A2	EU526388	IBVQ	DQ472165
BJ03-1	DQ459475	Ark99	AF317209	J	DQ084440
Cal 99	AY514485	Beaudette	NC_001451	Jilin	AY839145
Conn	L18990	BJ	AY319651	KQ6	AY641576
CQ04-1	DQ459476	Cal 99	AY514485	LAH03I	DQ494213
De072	AF274435	Conn	AF317471	LDL04II	DQ352156
Florida 18288	AF027512	De072	AF202998	LDL97I	EF213582
GAV 92	U16157	H52	AF317210	LDL98I	DQ287910
Gray	L14069	HK	AY761141	LDT3	AY702975
GX1-98	AY319302	IBN	EF602439	LGD04II	DQ352154
H120	M21970	JAAS	EF602441	LGD96I	DQ287911
H52	AF352315	Jilin	AY846833	LHB96I	DQ287912
HK	AY761141	KQ6	AY641576	LHLJ/04XI	DQ352152
Hotle	L18988	LDT3	AY702975	LHLJ02I	DQ352149
Jilin	AY839144	LGD04II	EF602444	LHLJ04 V	DQ352153
KQ6	AY641576	LHB96I	EF602460	LHLJ95I	DQ287913
LDT3	AY702975	LHLJ04XI	EF602461	LHLJ99I	AY842864
LGX06I	EF213580	LHLJ95I	EF602454	LHN00I	DQ352147
LKQ3	AY702085	LKQ3	AY702085	LKQ3	AY702085
LNM05I	EF213567	LSC99I	EF602450	LLN98I	DQ287914
LX4	AY338732	LSD03I	EF602457	LSC95I	DQ287915
LXJ02I	DQ167152	LTJ95I	EF602448	LSC99I	DQ287916
M41	AY851295	LX4	AY338732	LSD03I	DQ352151
PSH050513	DQ16004	LXJ02I	EF602458	LTJ95I	DQ287917
QXIBV	AF193423	M41	AY851295	LX4	AY338732
S	DQ490213	S	DQ490213	LXJ02I	DQ352150
S14	AY646283	S14	AY646283	M41	AY851295
SAIBK	DQ288927	SAIBK	DQ288927	PSH050513	DQ144422
SC021202	AY237817	TW1171/92	DQ646406	QXIBV	AF199412
SC03-1	DQ459472	TW2296/95	DQ646404	S	DQ490213
T	AY775779	TW2575/98	DQ646405	S14	AY646283
TA03	AY837465	Vic	DQ490221	SAIBK	DQ288927
TW1171/92	DQ646406	W93	EF602443	SH	DQ472166
TW2296/95	DQ646404	**M**		TW1171/92	DQ646406
TW2575/98	DQ646405	A2	EU526388	TW2296/95	DQ646404
UK793	Z83979	Ark DPI	AY942739	TW2575/98	DQ646405
Vic	DQ490221	Beaudette	NC_001451	Vic	DQ490221
YN05-01	DQ459474	BJ	AY319651	W93	AY842861
ZJ971	AF352313	Cal 99	AY514485	WF	DQ472167
S2		Conn	AY942741	ZJ971	AF352308
A2	EU526388	De072	AF202999	**Papain-like protease**	
Ark DPI	AF335555	H52	AF286185	A2	EU526388
Ark 99	AF094814	HK	AY761141	Ark99	AY392061
Beaudette	NC_001451	IBN	AY856347	Ark DPI	AY392062

Table 2 continued

BJ	AY319651	JAAS	EF602441	Beaudette	NC_001451
Cal 99	AY514485	Jilin	AY846833	Beaudette p65	DQ001339
California	AF335553	KQ6	AY641576	Beaudette Vero	AY692454
Conn	AF094818	LDT3	AY702975	BJ	AY692454
De072	AY024337	LGD04II	EF602444	Cal 99	DQ001339
Florida 18288	AF094819	LHB96I	EF602460	Conn	AY392059
GAV 92	AF094817	LHLJ04XI	EF602461	CU-T2	AY561734
Gray	AF394180	LHLJ95I	EF602454	De072	AY392064
H52	AF352315	LKQ3	AY702085	Florida	AY392060
HK	AY761141	LSC99I	EF602450	GA98	AY392063
Holte	AF334685	LSD03I	EF602457	Gray	AY392066
IBN	EF602439	LTJ95I	EF602448	H52	AY392058
Iowa	AF334684	LX4	AY338732	KQ6	AY338732
JAAS	EF602441	LXJ02I	EF602458	LX4	AY514485
JMK	AF239982	M41	AY851295	M41	AY646283
KQ6	AY641576	S	DQ490213	Ma5	AY561730
LDT3	AY702975	S14	AY646283	S14	AY319651
LGD04II	EF602444	SAIBK	DQ288927	SAIBK	AY851295
LHB96I	EF602460	TW1171/92	DQ646406	**3C-LP**	
LHLJ04XI	EF602461	TW2296/95	DQ646404	Beaudette	NC_001451
LHLJ95I	EF602454	TW2575/98	DQ646405	BJ	AY319651
LKQ3	AY702085	Vic	DQ490221	Cal 99	AY514485
LSC99I	EF602450	W93	AY846835	S14	AY646283
LSD03I	EF602457	**N**		LX4	AY338732
LTJ95I	EF602448	A2	EU526388	M41	AY851295
LX4	AY338732	Ark DPI	AY942745	KQ6	AY641576
LXJ02I	EF602458	Beaudette	NC_001451	SAIBK	DQ288927
M41	AY851295	BJ	AY319651	A2	EU526388
S	DQ490213	Cal 99	AY514485	Beaudette Vero	AY692454
S14	AY646283	Conn	AY942746	Beaudette p65	DQ001339
SAIBK	DQ288927	De072	AF203001		
TW1171/92	DQ646406	GE	DQ377140		
TW2296/95	DQ646404	Gx1-98	AY278110		

The bold letters indicated the used genes in phylogenetic analysis

contains 3567 nt and has a single ORF with 3507 nt. S gene is cleaved into two subunit S1 and S2 with 1614 and 1890 nt long, respectively (potentially encoding 538 and 621 aa). There are 699 nt in gene 3, encoding non-structural protein 3a, 3b and structural protein E with 174nt, 192nt and 327 nt, respectively. Between gene 2 and gene 3, there are 22 nt overlap. Between 3a and 3b, 3b and 3c, there are 1 and 18 nt overlap. Gene 4 contains 757 nt, encoding M protein of 225 aa. Gene 3 has 115 nt of overlap with gene 4. Gene 5 has two ORFs: 5a (198 nt) and 5b (249 nt). Between gene 4 and ORF 5a, there is a 363 nt fragment which does not encode protein. Gene 6 contains 1332 nt, encoding N protein of 409 aa. After the N gene, the 3′UTR and polyA tail were observed.

The full genome sequence of strain A2 was compared with all the available complete genomic sequences from GenBank. The results revealed that the nucleotide identities between A2 and these strains ranged from 84.43 to 94.27%. A2 strain had the highest similarity to S14 at 94.27% identity and least like M41 with 84.43% identity. In our current study, point mutations, insertions, and deletions were distributed throughout the whole genome RNA. Nsp1, S1, gene 3, and ORF 5a were hypervariable, whereas nsp4 to nsp13, M and ORF 5b were relatively conserved.

Nucleotide identities between different fragments of gene 1 were compared using DNAstar software. All gene 1 fragments of A2 had highest identity with S14, except for nsp9 and nsp11-nsp13 (Table 3). For nsp9 and nsp11-nsp13, A2 was 97.1% and 93.2% identical to BJ, whereas was 97.0% and 92.2% identical to S14, respectively. Nsp2

Table 3 Pairwise sequence comparisons of the A2 gene 1 fragments with other IBV strains

Strains	Nsp1	Nsp2	Nsp3	Nsp4–6	Nsp7	Nsp9	Nsp10	Nsp11–13
LX4	91.0	88.4	87.5	93.1	90.6	95.4	94.9	91.9
SAIBK	82.4	85.4	86.1	88.2	88.8	89.4	89.7	89.2
BJ	95.7	92.4	95.0	97.4	89.4	**97.1**	94.8	**93.2**
S14	**95.9**	**95.4**	**99.1**	**98.4**	**95.9**	97.0	**96.9**	92.2
M41	82.2	85.7	84.4	87.5	90.1	89.6	88.6	88.3
Beaudette	81.8	85.5	85.3	88.3	89.9	89.1	90.0	88.4
California 99	81.5	84.8	85.5	88.3	89.2	89.0	89.1	88.0
KQ6	82.3	85.6	84.2	87.2	89.9	89.1	89.0	88.4

Highest identity are indicated in bold letters

is the putative 3C-like proteinase (3C-LP), also known as the main proteinase (M^{pro}), which plays a pivotal role in viral gene expression and replication as well as in proteolytic processing of gene 1-encoded polyproteins in coronaviruses. Between different IBVs 3C-LP is highly conserved [25–27]. But compared with other gene 1 fragments in this study, nucleotide identity between A2 and other strains for 3C-LP was relatively low.

To interpret relationship between A2 and other IBV strains in 3′-terminal 7.3 kb genome, 15 other IBV sequences were obtained from GenBank. Parewise comparisons were performed using DNAstar computer software, as shown in Table 4.

In the S1 gene, the A2 virus and other strains were 58.7–91.9% identical and the SAIBK strain was the most similar at the nt level (91.9% identical). There was an eight aa insertion between 75–82 in deduced aa sequences, in which a putative N-glycosylation site was found. The BJ S1 gene had the same eight aa insert identical of the A2 virus. The insertion was also found in SAIBK with one aa difference. The insertion position was located in or near HVR1, which had high relationship with neutralization antibody [21, 28]. So insertion of eight aa and change of glycosylation pattern may potentially affect the immunoreaction of the S1 protein. In addition, there were two aa deletions (aa 22 and 124) and 95 aa mutations in the A2 S1gene compared with M41. 70 out of 95 mutations were located in the N-terminal 300 aa which contained HVR1, HVR2 and HVR3 [18, 29, 30].

Compared to S1 sequence, S2 was less variable. Between A2 and the other IBVs, nucleotide and aa identity were 74.6–94.8% and 74.2–94.9%, respectively. The A2 S2 gene was 94.5% identical to LX4 (94.3% identical to SAIBK and LDT3 at aa level). At position of aa 15, 293, 377, 435, 493, 565, 621 of S2 protein, the aa were S, I, R, L, N, R, and A for A2, while all other strains were P, V, Q, V, D, W, and T.

The S glycoprotein of IBV is translated as a precursor protein (S_0), then cleaved into two subunits S1 and S2. Previous studies reveal that the cleavage recognition site sequence, which consists of five basic aa residues, does not appear to correlate with increased cleavability, host cell range and increased virulence as it does with envelope glycoproteins in orthomyxoviruses and paramyxoviruses, but correlates with IBVs in different geographic regions [30, 31]. In the present study, strain A2 shared the same

Table 4 Pairwise sequence comparisons of the A2 3′-terminal 7.3 kb genome with other IBVs

Strain	S1	S2	3a	3b	E	M	5a	5b	N	3′UTR
LX4	80.3(79.3)	**94.8**(94.9)	84.6(76.3)	89.1(80.0)	94.8(92.7)	**93.3**(93.0)	**90.5**(86.6)	95.2(94.0)	92.9(94.6)	38.5
LDT3	87.7(85.1)	93.5(94.3)	**97.7**(98.3)	**95.3**(92.2)	**98.2**(95.5)	91.9(93.0)	85.4(71.6)	**96.4**(92.9)	94.2(95.1)	NA[a]
SAIBK	**91.9**(90.1)	92.9(94.3)	87.4(81.4)	80.5(76.2)	86.6(88.2)	91.2(93.0)	84.9(80.6)	92.0(89.3)	87.4(90.8)	41.1
BJ	85.9(84.0)	85.1(88.0)	97.1(96.6)	**95.3**(92.2)	93.6(90.0)	90.0(90.7)	86.4(89.6)	88.0(84.5)	92.5(92.2)	97.2
S14	88.2(86.3)	93.4(94.1)	**97.7**(98.3)	**95.3**(92.2)	97.9(95.5)	91.9(93.0)	85.4(82.1)	**96.4**(92.9)	**94.3**(95.1)	**97.8**
M41	82.7(81.3)	85.9(87.0)	83.4(81.4)	77.4(69.2)	87.8(85.5)	90.8(93.4)	77.9(71.6)	90.4(85.7)	87.2(90.8)	40.8
Beaudette	82.7(81.5)	85.4(86.1)	84.6(86.4)	77.9(69.2)	88.4(84.5)	90.8(92.1)	82.4(77.6)	91.2(90.5)	86.9(89.3)	87.3
California 99	76.9(74.9)	85.7(86.2)	85.7(91.5)	71.6(63.1)	86.0(81.8)	91.6(94.7)	80.9(76.1)	91.2(89.3)	87.0(91.0)	87.3
LKQ3	82.8(81.9)	85.2(86.2)	81.7(83.1)	77.4(69.2)	88.4(86.4)	90.6(92.5)	77.9(82.1)	90.4(88.1)	87.2(90.5)	NA
KQ6	82.7(81.1)	85.6(86.2)	83.4(81.4)	77.4(69.2)	88.4(86.4)	90.8(93.0)	82.4(82.1)	92.4(91.7)	87.9(91.2)	40.8
Jilin	78.7(75.3)	NA	86.9(86.4)	72.1(61.5)	86.3(81.5)	91.7(94.2)	NA	NA	87.4(91.7)	NA
HK	82.7(81.5)	85.7(87.7)	86.3(84.7)	72.1(61.5)	86.3(81.5)	91.5(93.8)	80.9(76.1)	91.6(88.1)	87.2(91.2)	NA
Vic	80.2(77.2)	85.1(86.9)	82.9(83.1)	72.6(61.5)	83.2(80.6)	86.8(90.7)	81.9(70.1)	92.4(89.3)	86.8(90.8)	NA
S	80.2(77.0)	85.1(86.7)	82.9(83.1)	72.6(61.5)	83.2(80.6)	87.1(91.1)	81.4(70.1)	92.0(88.1)	86.8(90.8)	NA
Tw1171/92	83.7(83.6)	89.7(92.7)	84.6(84.7)	77.4(69.2)	87.8(87.3)	89.8(91.6)	83.4(79.1)	92.4(89.3)	85.6(87.8)	90.6

Sequences with the highest identity to A2 at nucleotide level are indicated in bold letters

[a] NA: identity of nucleotide and amino acid not analyzed

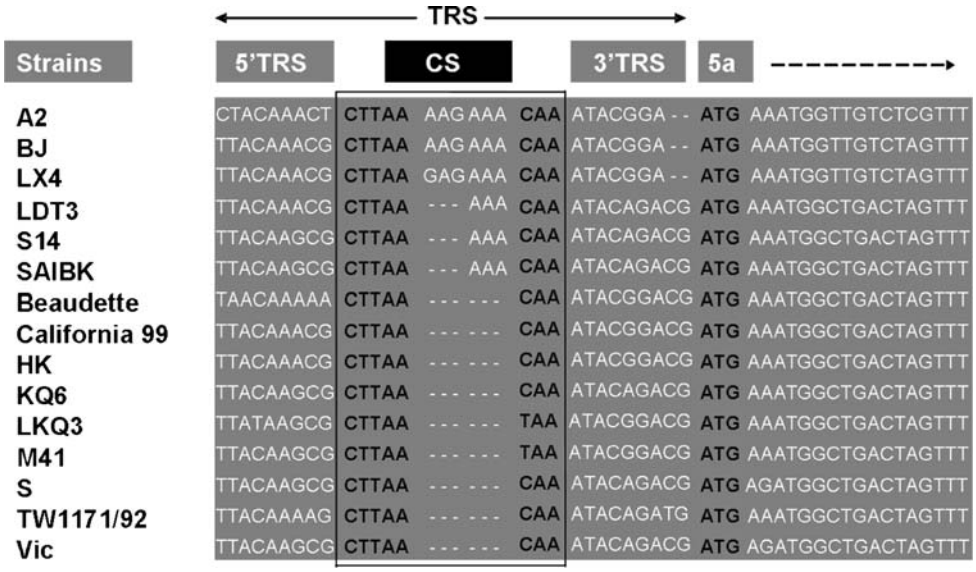

Fig. 1 Sequence comparisons of strain A2 with other strains in TRS of gene 5. Strain A2 and BJ have six nucleotides (AAGAAA) insertion and two nucleotides deletion compared to most strains

cleavage recognition site sequence with a majority of other strains (10 of 15 viruses): Arg-Arg- Phe -Arg-Arg, which was the most common cleavage recognition site observed [31, 32].

The 3a amino acid sequence of the A2 virus was 100% identical to those of the S14 and LDT3, except a substitution of Gln by an His at aa 54. For the 3b protein, strain A2 was 92.2% identical to those of S14, LDT3, and BJ at aa level. The third protein encoded by gene 3, E protein was also the most similar to S14 and LDT3 with 95.5% identity. In addition, all other strains were Leu and Trp at aa 12 and 52, except for A2 which were replaced by Gln and Arg, respectively.

Comparisons of the A2 M protein with other IBV strains showed that the aa identity ranged from 90.7 to 94.7%, with California 99 strain being the closest matches, and with BJ and Vic strains the least similar. The identity was obviously higher than any other structural proteins. The N-terminal 18 amino acids were highly variable in M protein. A2 together with most other strains had a putative N-glycosylation site in this 18 aa sequence. At aa 197, all other strains had a basic aa Lys, whereas A2 changed to an acidic aa Glu.

In gene 5, the 5a gene of the A2 strain was 90.5% identical to BJ strain at nt level. As for the 5b gene, A2 was 96.4% identity to those of LX4 and S14 strains.

The N protein is another conserved structural protein. The aa identity of the A2 N protein with other IBV strains ranged from 87.8 to 95.1% and a maximum 95.1% aa identity with S14 and LDT3 viruses. Point mutations were randomly distributed in the whole sequence. But no deletion or insertion occurred in the N protein of the A2 strain.

Transcription-regulating sequences (TRS) precede each gene and include a conserved core sequence (CS) surrounded by relatively variable sequences (5′TRS and 3′TRS). The consensus TRS of IBV, CTT/GAACAA, was highly conserved in IBV genome at the levels of nucleotide sequence and location in regarding to the initiation codon of individual genes [33, 34]. The canonical TRS of IBV was also found in strain A2 in the present study as shown in Fig. 1, except for gene 5. In gene 5, A2 had a six nt insertion (AAGAAA) in TRS compared to those corresponding regions of IBVs. The BJ virus had the same insert observed. The TRSs maybe serve as binding site of transcript complex and the distance between TRS and initiation codon may affect the subgenomic RNA synthesis and translation and in turn change the virus biologic activity [35–37]. The results in the present study also suggest that the TRS sequence is not essential for virus replication and multiplication.

Phylogenetic trees were derived from the nucleotide sequences of specific gene of the A2 virus, as shown in Fig. 2. Phylogenetic tree constructed from the S1 gene sequences revealed six genetic groups. The analysis showed that A2 and most Chinese strains formed into the same groupI [21, 30]. Strains of groupIranged from 1996 to 2006, most of them were isolated between 2002 and 2006. The results revealed that A2 related IBVs were prevalent in China in recent years (Fig. 2a). A2 strain had a more close evolutionary distance to other group compared with other groupIstrains. We infer that A2 may play an important role in groupIIBVs evolution.

In S2, E, M, and N genes, the same results were obtained and strain A2 grouped with most Chinese isolates into one subcluster (Fig. 2b–e).

For 3C-LP and the papain-like protease, the analysis was unclear because the limited strains in GenBank (Fig. 2f, g). In all phylogenetic trees, A2-like viruses and Massachusetts strains, such as H120 and H52, were distributed into different groups. Athough different genotypes of infectious

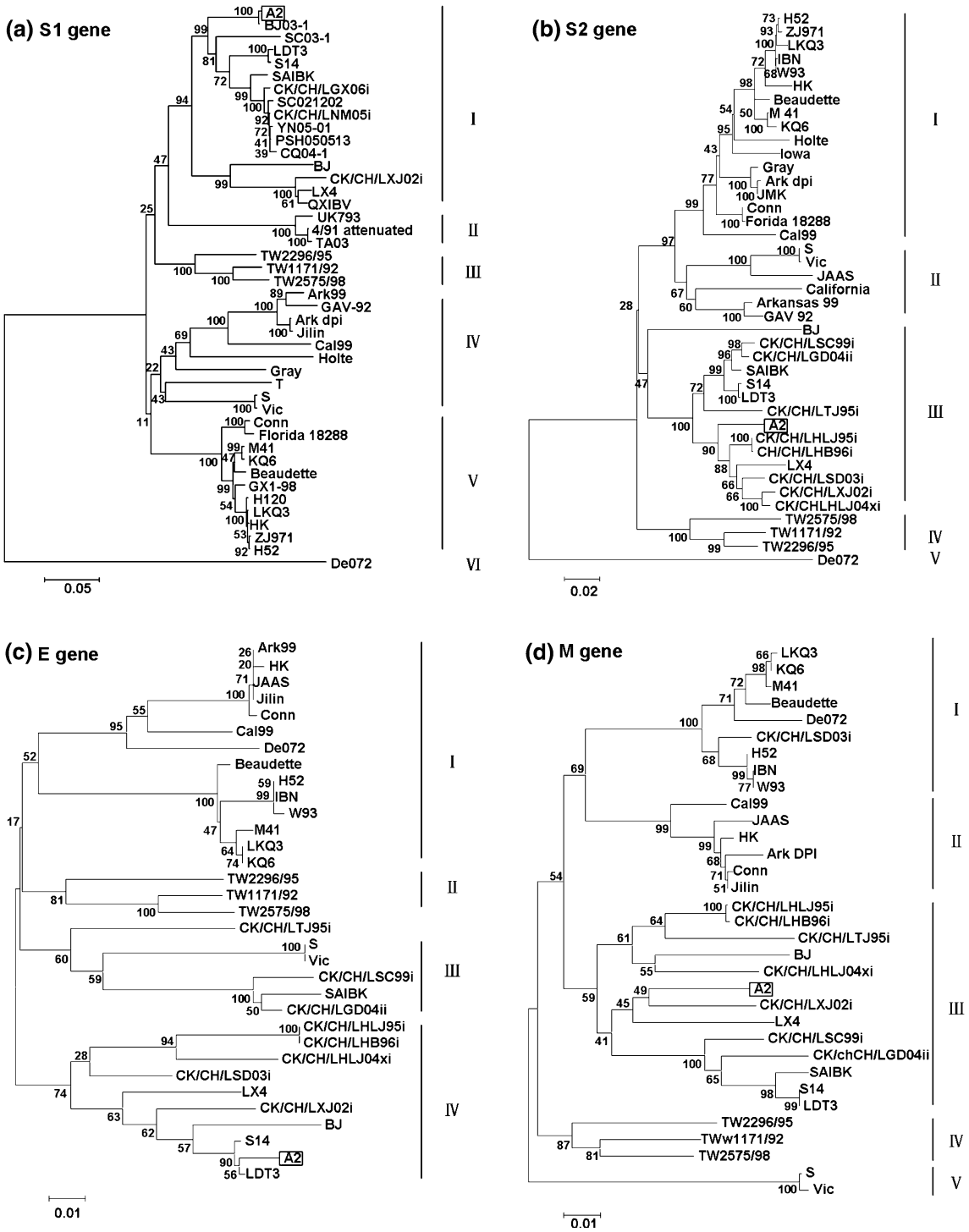

Fig. 2 Phylogenetic trees of each fragment, generated by the neighbor-joining method with 1,000 bootstrap replicates. The horizontal bar indicates the nucleotide substitutions per site, while vertical distances are for clarity. Phylogenetic trees derived from all fragments revealed that most Chinese isolates were related to A2 strain and A2-like strains had a far distance to Mass and Conn serotype strains

bronchitis viruses co-circulated in China [21, 38], the results from our study revealed that A2-like viruses were the dominant IBV strains in China and had a far distance to Massachusetts strains, which were the most common vaccine strain in China.

Infectious bronchitis has occurred frequently in China in recent years. The results of this study indicate that the A2 virus, which is the first strain of 4/91 like serotype isolated in China, plays an important role in China IBV evolution. Due to vaccine of A2-like or 4/91 related serotype are not

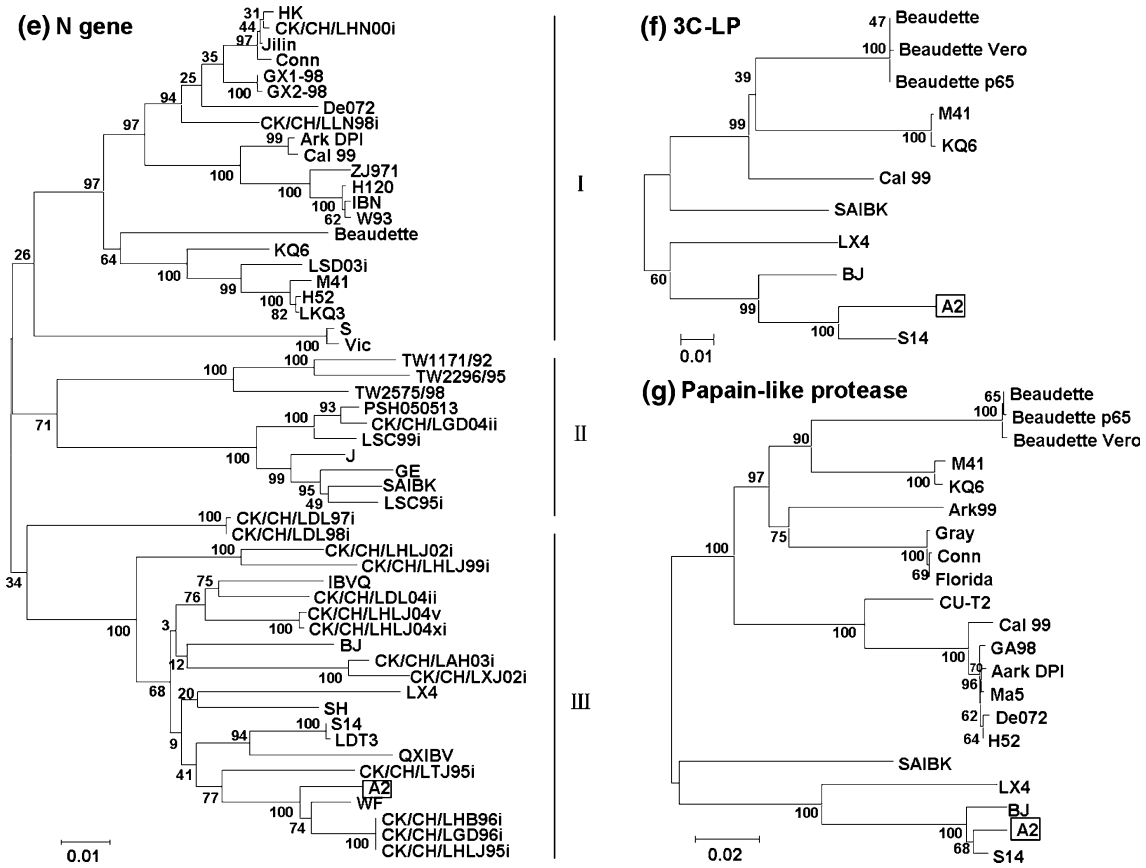

Fig. 2 continued

available, the A2-like IBV are still circulating in China. To prevent A2-like IBV outbreak frequently, attenuated or inactived vaccine are urgently needed, though the use of vaccine may contribute to the mutation and recombination of IBV.

References

1. D. Cavanagh, Vet. Res. **38**, 281–297 (2007). doi:10.1051/vetres:2006055
2. S.W. Liu, J.F. Chen, J.D. Chen, X.G. Kong, Y.H. Shao, Z.X. Han et al., J. Gen. Virol. **86**, 719–725 (2005). doi:10.1099/vir.0.80546-0
3. G.H. Zhang, J.D. Fu, T. Ren, W.S. Cao, K.J. Luo, C.G. Xu et al., Prog. Nat. Sci. **16**, 1275–1280 (2006). doi:10.1080/10020070612330141
4. D. Cavanagh, K. Mawditt, M. Sharma, S.E. Drury, H.L. Ainsworth, P. Britton et al., Avian Pathol. **30**, 355–368 (2001). doi:10.1080/03079450120066368
5. D. Cavanagh, K. Mawditt, D.B. Welchman, P. Britton, R.E. Gough, Avian Pathol. **31**, 81–93 (2002). doi:10.1080/03079450120106651
6. S. Dea, A.J. Verbeek, P. Tijssen, J. Virol. **64**, 3112–3118 (1990)
7. E.J. Snijder, P.J. Bredenbeek, J.C. Dobbe, V. Thiel, J. Ziebuhr, L.L. Poon et al., J. Mol. Biol. **331**, 991–1004 (2003). doi:10.1016/S0022-2836(03)00865-9
8. M.M. Binns, J. Gen. Virol. **68**, 57–77 (1987). doi:10.1099/0022-1317-68-1-57
9. D.F. Stern, B.M. Sefton, J. Virol. **50**, 22–29 (1984)
10. A.O. Pasternak, W.J.M. Spaan, E.J. Snijder, J. Gen. Virol. **87**, 1403–1421 (2006). doi:10.1099/vir.0.81611-0
11. W. Spaan, D. Cavanagh, M.C. Horzinek, J. Gen. Virol. **69**, 2939–2952 (1988). doi:10.1099/0022-1317-69-12-2939
12. S. Sutou, S. Sato, T. Okabe, M. Nakai, Virology **165**, 589–595 (1988). doi:10.1016/0042-6822(88)90603-4
13. S. Youn, J.L. Leibowitz, E.W. Collisson, Virology **332**, 206–215 (2005). doi:10.1016/j.virol.2004.10.045
14. T. Hodgson, P. Britton, D. Cavanagh, J. Virol. **80**, 296–305 (2006). doi:10.1128/JVI.80.1.296-305.2006
15. R. Casais, M. Davies, D. Cavanagh, P. Britton, J. Virol. **79**, 8065–8078 (2005). doi:10.1128/JVI.79.13.8065-8078.2005
16. S. Shen, Z.L. Wen, D.X. Liub, Virology **311**, 16–27 (2003). doi:10.1016/S0042-6822(03)00117-X
17. K.J. Worthington, R.C. Jones, Vet. Rec. **159**, 291–292 (2006)
18. R. Smati, A. Silim, C. Guertin, M. Henrichon, M. Marandi, M. Arella et al., Virus Genes **25**, 85–93 (2002). doi:10.1023/A:1020178326531
19. S.W. Liu, X.G. Kong, Avian Pathol. **33**, 321–327 (2004). doi:10.1080/0307945042000220697
20. W.A. Nix, D.S. Troeber, B.F. Kingham, C.L. Keeler Jr., J. Gelb Jr., Avian Dis. **44**, 568–581 (2000). doi:10.2307/1593096
21. C.P. Xu, J.X. Zhao, X.D. Hu, G.Z. Zhang, Vet. Microbiol. **122**, 61–71 (2007). doi:10.1016/j.vetmic.2007.01.006
22. L. Yu, Y. Jiang, S. Low, Z. Wang, S.J. Nam, W. Liu et al., Avian Dis. **45**, 416–424 (2001). doi:10.2307/1592981

23. J. Gelb Jr., Y. Weisman, B.S. Ladman, R. Meir, Avian Pathol. **34**, 194–203 (2005). doi:10.1080/03079450500096539
24. D. Cavanagh, R. Casais, M. Armesto, T. Hodgson, S. Izadkhasti, M. Davies et al., Vaccine **25**, 5558–5562 (2007). doi:10.1016/j.vaccine.2007.02.046
25. S.P. Monda, C.J. Cardona, Virology **324**, 238–248 (2004). doi:10.1016/j.virol.2004.03.032
26. L.F. Ng, D.X. Liu, Virology **272**, 27–39 (2000). doi:10.1006/viro.2000.0330
27. H.Y. Xu, K.P. Lim, S. Shen, D.X. Liu, Virology **288**, 212–222 (2001). doi:10.1006/viro.2001.1098
28. D. Cavanagh, P.J. Davis, A.P. Mockett, Virus Res. **11**, 141–150 (1988). doi:0.1016/0168-1702(88)90039-1
29. C.H. Wang, Y.C. Huang, Arch. Virol. **145**, 291–300 (2000). doi:10.1007/s007050050024
30. S.W. Liu, Q.X. Zhang, J.D. Chen, Arch. Virol. **151**, 1133–1148 (2006). doi:10.1007/s00705-005-0695-6
31. M.W. Jackwood, D.A. Hilt, S.A. Callison, C.W. Lee, H. Plaza, E. Wade, Avian Dis. **45**, 366–372 (2001). doi:10.2307/1592976
32. D. Cavanagh, P.J. Davis, D.J. Pappin, M.M. Binns, M.E. Boursnell, T.D. Brown, Virus Res. **4**, 133–143 (1986). doi:10.1016/0168-1702(86)90037-7
33. T.L. Lin, C.C. Loa, C.C. Wu, Virus Res. **106**, 61–70 (2004). doi:10.1016/j.virusres.2004.06.003
34. I. Sola, J.L. Moreno, S. Zúñiga, S. Alonso, L. Enjuanes, J. Virol. **79**, 2506–2516 (2005). doi:10.1128/JVI.79.4.2506-2516.2005
35. A.O. Pasternak, E. van den Born, W.J. Spaan, E.J. Snijder, EMBO J. **20**, 7220–7228 (2001). doi:10.1093/emboj/20.24.7220
36. M.A. Ijms, D.D. Nedialkova, J.C. Zevenhoven-Dobbe, A.E. Gorbalenya, E.J. Snijder, J. Virol. **81**, 10496–10505 (2007). doi:10.1128/JVI.00683-07
37. P.S. Masters, Adv. Virus Res. **66**, 193–292 (2006). doi:10.1016/S0065-3527(06)66005-3
38. G.X. Bing, X. Liu, J. Pu, Q.F. Liu, Q.M. Wu, J.H. Liu, Virus Genes **35**, 333–337 (2007). doi:10.1007/s11262-007-0100-5