Genome sequencing and characterization analysis of a Beijing isolate of chicken coronavirus infectious bronchitis virus

JIN Weiwu, CHEN Chen, ZHANG Ying, ZHAO Yiqiang, FENG Jidong, CHEN Fuyong, WU Qingming, YANG Hanchun, WANG Ming, YU Jialin, LI Ning, GONG Yuanshi, SUN Qixin & CHEN Zhangliang

Animal Coronavirus Research Group, China Agricultural University, Beijing 100094, China
Correspondence should be addressed to Li Ning (e-mail: ninglbau@public3.bta.net.cn) or Chen Zhangliang (e-mail: zhangliangchen@pku.edu.cn)

Abstract Avian infectious bronchitis virus (AIBV) is classified as a member of the genus coronavirus in the family coronavirusidae. The enveloped virus has a positive-sense, single-stranded RNA genome of approximately 28 kilo-bases, which has a 5′ cap structure and 3′ polyadenylation tract. The S protein is cleaved to release S1 and S2 proteins after translation. The S1 protein related to SARS coronavirus. Phylogenetic analyses based on the whole genome sequence of infectious bronchitis virus (IBV), Beijing isolate, was determined by cloning sequencing and primer walking. The whole genome is 27733 nucleotides in length, has ten open reading frames: 5′-orf1a-orf1ab-3a-3b-3c-3b-c-m-6a-6b-n-3′. Alignments of the genome sequence of IBV Beijing isolate with those of two AIBV strains and one SARS coronavirus were performed respectively. The genome sequence of IBV Beijing isolate compared with that of the IBV strain LX4 (uncompleted, 19440 bp in size) was 91.2% similarity. However, the full-length genome sequence of IBV Beijing isolate was 85.2% identity to that of IBV Strain Beaudette, and was only 50.8% homology to that of SARS coronavirus. The results showed that the genome of IBV has remarkable variation. And IBV Beijing isolate is not closely related to SARS coronavirus. Phylogenetic analyses based on the whole genome sequence, S protein, M protein and N protein, also showed that AIBV Beijing isolate is lone virus in group III and is distant from SARS coronavirus. In conclusion, this study will contribute to the studies of diagnosis and diseases control on IBV in China.

Keywords: avian infectious bronchitis virus (AIBV), severe acute respiratory syndrome coronavirus (SARS-CoV), primer walking, sequence analysis.

DOI: 10.1360/03we0347

Avian infectious bronchitis virus (AIBV) is a highly infectious, contagious pathogen of chicken and causes disease of the respiratory and urogenital tracts, resulting in significant economic losses to poultry industries throughout the world[1]. In China, IBV leads to more than RMB one billion yuan losses every year. IBV, as a member of the family coronavirusidae in the new order Nidovirales, has a single-stranded positive-sense RNA genome, which is 27.6 kb in length[2]. The genome of IBV has more than 10 open reading frames, which encode two polyproteins and three major structural proteins and other small proteins. Three major structural proteins are the spike glycoprotein (S), the membrane glycoprotein (M), and the phosphorylated nucleocapsid protein (N). The S protein induces protective immunity[3–5]. And the most antigenic of the virus neutralization antibody-inducing epitopes is formed by a few amino acids that occur in the first and third quarters of the S1 subunit, especially in the hyper-variable regions[6,7].

IBV is the lone virus in the antigenic group III of coronavirusidae[8], however, more than 20 serotypes within IBV have been recognized worldwide on the basis of virus-neutralization test[9]. Therefore, it is very difficult to prevent from the infections of virus. The whole genome of IBV was determined in 1987[10]. And the genome sequences of IBV are very important and useful for disease diagnosis and vaccination. However, remarkable variation was observed in various IBV isolates: the genome of IBV isolate in this region may be very different with that in the other regions. IBV Beijing isolate, which caused great economic losses, was obtained from a chicken house in Yukou, Beijing[10].

In addition, an outbreak of atypical pneumonia, named severe acute respiratory syndrome (SARS) by the WHO, was identified in Guangdong Province, China, and spread to several countries. This disease is great threatening to the health of people worldwide. Many reports showed that a novel coronavirus was the pathogen of the SARS[11–13]. And scientists considered that SARS coronavirus could be from some animals. After the complete genome of IBV Beijing isolate was sequenced, we determined the relations between IBV Beijing isolate and SARS coronavirus.

1 Materials and methods

(1) Virus. IBV Beijing isolate[10] was propagated in 10-d-old specific-pathogen-free (SPF) embryonated chicken eggs. The allantoic fluid was collected 24–36 h after inoculation, clarified by centrifugation at 5000 r/min for 30 min, 4°C. The pellet was discarded and the supernatant was concentrated by more than 2 h high-speed centrifugation (30000 r/min) at 4°C. The supernatant was discarded and the pellet was resuspended in PBS (pH 7.2). The virus suspension was ultra-centrifuged through a 30%/40%/50% saccharose step gradient at 32000 r/min for 3 h at 4°C. Mixture between 40% and 50% saccharose was transferred and PBS (pH 7.2) was added. Samples were mixed and concentrated by centrifugation at 30000 r/min for 2 h at 4°C. The supernatant was discarded and
the virus pellet was resuspended in PBS. The virus sus-
pension was divided into 250 μL each tube and was stored at −80°C until use.

(ii) Viral RNA extraction and RT-PCR. Viral genomic RNA was extracted with Trizol LS reagent (Invitrogen, USA). Each 250 μL virus suspension was added with 750 μL Trizol LS reagent and mixed. The homoge-
nized samples was incubated for 5 min at 15°C—30°C and then added with 200 μL chloroform. The capped sam-
ple tubes were shaken vigorously by hand for 15 s and incubated at 15°C for 30 min. The supernatant was discarded and RNA pellet was air-dried (about 15 min) until use.

The random hexameric primer (Promega, USA) or a specific reverse primer was used in cDNA first strand synthesis. After cDNA first-strand was synthesized, each specific primer pair was used in the reaction of PCR. A 20-μL reaction volume was used for the first strand cDNA synthesis. Total RNA (10 μL) and random hexameric primer (0.5 μg/μL, Promega, USA) or a specific reverse primer 1 μL were added to a nuclease-free microcentrifuge tube. Mixture was heated to 70°C for 10 min and was quickly chilled on ice for 1—2 min. After brief centrifuga-
tion, the following components: 4 μL 5 × first-strand buffer, 2 μL 0.1 mmol/L DTT, 1 μL RNaseOut (40 U/μL) and 1 μL M-MLV reverse transcriptase (200 U/μL) were added. After being mixed gently, the mixture was incu-
bated at 37°C for 1 h. Samples were inactivated at 70°C for 15 min and subsequently treated with 2 U of Ribonu-


clease H at 37°C for 30 min. The cDNA was used as tem-
plate for amplification in PCR. A 30 μL reaction volume
was used in PCR. The components were 1.5 μL cDNA, 3 μL 10 × buffer, 2 μL 2.5 mmol/L dNTP each, 1 μL of each primer (25 μmol/L), 0.5 μL Taq DNA polymerase (5 U/μL) and 21 μL ddH2O. PCR was carried out with the following thermocycling parameters: denaturation at 94°C for 5 min, 35 or 40 cycles of denaturation at 94°C for 40 s, annealing at 49°C—53°C for 1 min, and extension at 72°C for 1—2 min or 68°C for 4—5 min, a final extension at 72°C for 10 min and store at 4°C. Annealing and extension con-
ditions were dependent on primers and the length of amplification fragments. Some primers used in PCR are listed in Table 1.

Rapid amplification of cDNA ends (RACE) was performed to capture the 3′ and 5′ ends of the viral ge-

tabelle 1  Primers for PCR

| Primer | Position/nt | Size/nt | Forward primer | Reverse primer |
|--------|-------------|---------|----------------|----------------|
| gp5t   | 1—355       | 355     | ACT TAA GAT AGA TAT TAA TAT A | GGT GAC CCC GTA/G TTT C |
| gp1    | 1042—1928   | 887     | CAT GGA TGC TT/CT GGG CTC AAA T/C | TCC/T TCA AGA/G TTT/G CGC ACT TTT CC |
| gp3    | 2817—6014   | 3198    | G/A/GC GTA AAG AC/TT TGG ATT GC/TA T | GCA/G TGT TTA/G TAC AAA TGA AGC |
| gp3-2  | 2793—3540   | 748     | AGA ATG TC/CT GAC CAT TTG TTG ATA A | TT/CT TAG CC/TT TCT/C TCT CAC GTT |
| gp3-3  | 4050—6014   | 1965    | TGA G/ACC AAG AA/GC AC/TA/G TCG ATT AT T | AA/GG CG/AT GTT GT/AT ACA AAT GAA G |
| gp4    | 6708—10454  | 3474    | TTC TTA AGG AT/AG CAC A/TA GA AT A/G | TAA CA/CA TG/AA GA/CT TTA GA/TA C |
| gp4-2  | 6960—9601   | 2642    | ATT ATG AA GA/GG GCA CAC TTC G | ATT G/AT C/CC/G CA/TT CCC ACT TAT TA |
| gp5    | 10498—14242 | 3745    | TTG TTA CTT TGG A/GTT C/TTG TTT G | CTA TAC GCA ACA AAT TT/AG GCA T |
| gp5-2  | 12341—12956 | 616     | TCT GTT CAA TCA GTT GCT GGA/T | CCA ACA TAA CCT/C TTT TCA ACC A |
| gp5a   | 14594—17221 | 2627    | AAA AGT TTT ATT CTT A/CA/TG/AT G | CTT CTA CAT CAA AAC CNA C |
| gp5a-2 | 14896—16710 | 1815    | GTA GAT GAC/A GTG GAT AAG ACA | CTG G/ATT CAT AGC ATG GTA AGG T |
| gp6    | 17612—18097 | 486     | TTA ATT CTC CTA CTC AG/AG A/CTT A | ACA TTG ACAA TAA C TCA C |
| gp8    | 18442—21319 | 2878    | TAT/C AAT GCA GCT GTT/C ACA | TAG/AAAT TAT AAT A AC CAC TCT |
| gp9    | 21706—23182 | 1477    | TTT AG/AA TAA/C G/TTG TGG GTA G/TTA T | GTA TGGA TTT AA/GG AA CAT GTC TGT |
| gp9-2  | 22193—22681 | 489     | TCT/T AAC/C TTA/G ACT GTT ACA GAT G | CCA TAG AA/GT CTA CTA G/AAG AA |
| gp3c-2 | 25184—25901 | 718     | GAT ACT GGC GAG CTA GAA AG | ACG CCG GAC AAG GCT CTG |
| gp3c-3 | 26209—27445 | 1237    | GGT TTA AGG CAG GCA AAG GT | CTG GCG ATAG CAT AG |

nome. A 20-μL reaction volume was used for the first strand cDNA synthesis in 3′ RACE. The procedure was done as follows: AP (5′-GGCCACGCGTCGACTAGTA-

CTTTTTTTTTTTTTTTTTTTTTTTT-3′) 1 μL, mixture contained 10 μL total RNA, were heated to 70°C for 10 min and was quickly chilled on ice for 1—2 min. After brief cen-
trifugation, 2 μL 10× buffer, 2 μL 0.1 mol/L DTT, 2 μL 25 mmol/L MgCl₂, 1 μL 10mmol/L dNTP Mix, 1 μL RNaseOut (40 U/μL) and 1 μL M-MLV reverse transcriptase (200 U/μL) were added. Mixture was mixed gently and incubated at 37°C for 1 h. Samples were inactivated at 70°C for 15 min and subsequently treated with 1 μL Ribonuclease H (2 U/μL) at 37°C for 30 min. 1.5 μL cDNA was used as template in a 25-μL PCR reaction volume. 0.5 μL AP, 0.5 μL GSP (5’-ATTACCTACATGTCTATCGCCA-3’), 2.5 μL 10× buffer, 25 mmol/L MgCl₂ 1.5 μL, 0.5 μL 10 mmol/L dNTP Mix, Taq DNA polymerase (5 U/μL) 0.5 μL, 17.5 μL ddH₂O were mixed gently and heated to 94°C for 3 min to denature. 35 cycles were carried out for 30 s at 94°C, 1 min at 62°C, 1 min at 72°C, with a final extension at 72°C for 10 min, then stored at 4°C.

A 25-μL reaction volume was used in 5’RACE for the first strand cDNA synthesis. Mixture consisted of 15.2 μL total RNA, 0.3μL GSP1 (5’-ACTTAAGATAGATATTAATATA-3’), 2.5 μL 10× buffer, 25 mmol/L MgCl₂ 1.5 μL, 0.5 μL 10 mmol/L dNTP Mix, Taq DNA polymerase (5 U/μL) 0.5 μL, 17.5 μL ddH₂O were mixed gently and heated to 94°C for 3 min to denature. 35 cycles were carried out for 30 s at 94°C, 1 min at 62°C, 1 min at 72°C, with a final extension at 72°C for 10 min, then stored at 4°C.

The complete genome sequence of IBV Beijing isolate was determined using PCR and cloning before sequencing as well as primer walking approaches. After clones (recombinant plasmids) were sequenced, the assembly of 140 sequences were totally 11 kb in length. These sequences comprised 12 contigs, covering these regions more than 4-fold. 101 primers were used for direct sequencing, among which 36 primer pairs were used for PCR and 3 primers for RACE. 136 sequences were obtained by primer walking. The sequences assembled were about 17 kb in size. Totally, 276 sequences resulted in the complete 27733 bp sequence of IBV Beijing isolate genome. The average useful length of the component sequences was 350 bases. The IBV Beijing isolate genomic sequence has been deposited into GenBank (Accession: AJ319651).

(i) Genome organization. The 3’ and 5’ ends of the viral genome were captured after rapid amplification of cDNA ends (RACE) was performed. A fragment about 300 bp in size from 3’RACE and a band about 350-bp-length from 5’RACE were purified and sequenced directly or cloned into the PMD 18-T vector before sequencing. The sequences of RACE were almost 300 bp in size, with a 17polyA tract in 3’RACE, and 13polyG tail in 5’RACE (Fig. 1).

The IBV genome has five major genes, with a typical gene order: rep, spike, envelope, membrane, nucleocapsid. The rep gene, which comprises approximately two-thirds of the genome, is encoded two polyproteins (encoded by orf1a and orf1b). The four genes downstream of rep encoded the structural proteins S, E, M and N. There are also
Table 2  Ten ORF of IBV Beijing isolate

| ORF      | Position/nt | Peptide length/a.a. |
|----------|-------------|---------------------|
| RNA polymerase | 530—12391   | 6641                |
| Orf1a    | 12391—20454 | 808                 |
| S protein | 20405—23914 | 1169                |
| 3a protein | 23914—24087 | 75                  |
| 3b protein | 24087—24275 | 62                  |
| E protein | 24259—24588 | 122                 |
| M protein | 24557—25237 | 226                 |
| 6a protein | 25601—25798 | 65                  |
| 6b protein | 25795—26040 | 81                  |
| N protein | 25983—27212 | 409                 |

Sequences encoding five proteins of IBV Beijing isolate were also compared with those of IBV strain Beaudette and SARS coronavirus, at the levels of nucleotide and amino acid, respectively (Tables 3 and 4). S protein and E protein had low identity at the levels of nucleotide and amino acid between two strains of IBV. While RNA polymerase was of relatively higher conversation at the amino acid level between two strains. However, four structural proteins of IBV Beijing isolate had fairly low homologies (<50%) to those of SARS coronavirus. On the contrary, the homologies of four structural proteins at the amino acid level were lower than those at the nucleotide level, respectively, between IBV Beijing isolate and SARS coronavirus.
Table 3  nt identity of five proteins (%)

| Gene         | AIBV% | SARS-CoV% |
|--------------|-------|-----------|
| orf1a        | 83.5  | 48.9      |
| orf1b        | 88.7  | 59.2      |
| spike        | 81.5  | 49.8      |
| envelope     | 85.5  | 48.0      |
| membrane     | 87.5  | 49.0      |
| nucleocapsid | 88.0  | 48.6      |

a) AIBV strain Beaudette (Accession: NC_001451.1); b) SARS coronavirus (Accession: NC_004718.1/TOR2).

Table 4  a.a identity of five proteins (%)

| Protein     | AIBV% | SARS-CoV% |
|-------------|-------|-----------|
| Orf1a       | 91.1  | 31.5      |
| Orf1b       | 96.8  | 67.2      |
| Spike       | 86.5  | 32.4      |
| Envelope    | 86.2  | 29.1      |
| Membrane    | 92.1  | 43.8      |
| Nucleocapsid| 94.0  | 33.3      |

a) AIBV strain Beaudette (Accession: NC_001451.1); b) SARS coronavirus (Accession: NC_004718.1/TOR2).

3 Discussion

There is very remarkable variation in the genome sequence of IBV. The full-length genome of IBV Beijing isolate, which was 27733 bp in length, was 125 bp longer than that of IBV strain Beaudette (Accession: NC_001451.1), sharing only 85.2% identity of the nucleotide. And the nucleotide homology (91.2%) between IBV Beijing isolate and IBV LX4 (Accession: AY223860) will decline after the complete genome of IBV LX4 finished. For there is a gap of more than 6 kb at the 3’end of the genome, which contains three proteins: S, M and E. These sequences have low conservation. The number of amino acid of RNA polymerase and four structural proteins was
also observed. The orf1a of IBV Beijing isolate was 36 nt longer than that of IBV strain Beaudette (Accession: NC_001451.1), which resulted in 12 amino acids more for RNA polymerase than that of IBV strain Beaudette.

Among the four structural proteins, S protein, M protein, and N protein of IBV Beijing isolate were longer than those of IBV strain Beaudette (7a.a, 1a.a, 1a.a, respectively). The genome sequence of IBV Beijing isolate was very different from that of IBV strain Beaudette, and it would be very important for the studies of diagnosis and diseases control on IBV in China.

Moreover, the full-length genome of IBV Beijing isolate showed 50.8% homology to that of SARS coronavirus (Accession: NC_004718.1). The identities of five proteins of IBV Beijing isolate were fairly low at the amino acid level, compared to those of SARS coronavirus, respectively. Phylogenetic analyses also showed that Beijing isolate was not resembled closely to SARS coronavirus, which was consistent with other reports.[11–13]

Acknowledgements This work was supported by SARS Coronavirus Program of the Ministry of Education of China and the State High-Tech Research and Development Program (Grant No. 2002AA206111).

References
1. Cavanagh, D., Naqi, S. A., Infectious bronchitis, Diseases of Poultry, 10th ed. (ed. Calnek, B. W.), Ames: Iowa State University Press, 1997, 511–526.
2. Boursell, M. E., Brown, T. D. K., Foulds, I. J. et al., Completion of the sequence of the genome of the Coronavirus avian infectious bronchitis virus, J. Gen. Virol., 1987, 68: 55—77.
3. Cavanagh, D., Coronavirus IBV: structural characterization of the spike protein, J. Gen. Virol., 1983, 64: 2577—2583.
4. Cavanagh, D., Coronavirus IBV glycopolypeptides: size of their polypeptides and nature of their oligosaccharides, J. Gen. Virol., 1983, 64: 1187—1191.
5. Cavanagh, D., Davis, P. J., Coronavirus IBV: removal of spike glycopolypeptide S1 by urea abolishes infectivity and haemagglutination but not attachment to cells, J. Gen. Virol., 1986, 67: 1443—1448.
6. Cavanagh, D., Davis, P. J., Cook J. K. A. et al., Location of the amino acid differences in the S1 spike glycoprotein subunit of closely related serotypes of infectious bronchitis virus, Avian Pathol., 1992, 21: 33—43.
7. Wang, C. H., Huang, Y. C., Relationship between serotypes and genotypes based on the hypervariable region of the S1 gene of infectious bronchitis virus, Arch. Virol., 2000, 145: 291—300.
8. Siddell, S. G., The coronaviridae, The Viruses (eds. Fraenkel-Conrat, H., Wanger, R. R.), New York: Plenum Press, 1995, 1—49.
9. Gelb, J. J., Wolff, J. B., Moran, C. A., Variant serotypes of infectious bronchitis virus isolated from commercial layer and broiler chickens, Avian Dis., 1991, 35: 82—87.
10. Chen, F. Y., Studies on etiology of infectious bronchitis, Chin. J. Vet. Med. (in Chinese), 1998, 11: 8—10.
11. Rota, P. A., Oberste, S. M., Monroe, S. S. et al., Characterization of a novel coronavirus associated with severe acute respiratory syndrome, Science, 2003, 300: 1394—1399.
12. Marra, M. A., Jones, S. J., Astell, C. R. et al., The genome sequence of the SARS-associated coronavirus, Science, 2003, 300: 1399—1404.
13. Qin, E. D., Zhu, Q. Y., Yu, M. et al., A complete sequence and comparative analysis of a SARS-associated virus (Isolate BJ01), Chinese Science Bulletin, 2003, 48(10): 941—948.
14. Sterky, F., Lundeberg, J., Sequence analysis of genes and genomes, J. Biotech., 2000, 76: 1—31.
15. Enjuanes, L., Brian, D., Cavanagh, D. et al., Coronavirus, Virus Taxonomy: Classification and Nomenclature of Viruses (eds. van Regenmortel, M. H. V., Fauquet, C. M., Bishop, D. H. L. et al.), New York: Academic Press, 2000, 835—849.

(Received July 4, 2003; revised October 14, 2003; accepted January 13, 2004)