

Genome sequence and origin analyses of the recombinant novel IBV virulent isolate SAIBK2

Xuan Wu¹ · Xin Yang¹ · Pengwei Xu² · Long Zhou¹ · Zhikun Zhang¹ · Hongning Wang¹

Received: 24 December 2015 / Accepted: 5 April 2016 / Published online: 25 April 2016
© Springer Science+Business Media New York 2016

Abstract Recombination between infectious bronchitis viruses (IBVs), together with point mutations, insertions, and deletions, is thought to be responsible for the emergence of new IBV variants. SAIBK2 is a nephropathogenic strain isolated from layer flocks vaccinated with live attenuated H120 vaccine in Sichuan province, China in 2011. SAIBK2 causes severe kidney lesions and results in 50 % mortality in 30-day-old specific-pathogen-free chickens (with a dose of 10^5 EID₅₀/0.1 mL SAIBK2 per chicken). The complete genome of SAIBK2 consists of 27669 nucleotides, excluding the poly-A tail at the 3' end. SAIBK2 has the highest identity to YX10 in terms of complete genome. Phylogenetic analysis of complete sequence showed that SAIBK2 belongs to the most dominant genotype in China. Comparison and recombination analyses with other IBV strains revealed that SAIBK2 may originate from recombination events among a YX10-, a YN-, and a Mass-like strain. Furthermore, whole gene 5 and parts of nsp 3, nsp 4, nsp 16, and N genes are involved in the recombination events, and the uptake of these regions from YN and Mass strains by SAIBK2 may

increase its replication efficiency and be responsible for its increased virulence in specific-pathogen-free chickens.

Keywords Infectious bronchitis virus · Complete genome · Virulent strain · Recombination event

Introduction

Avian infectious bronchitis (IB) is a common, highly contagious, and acute disease of chickens; its pathogen infectious bronchitis virus (IBV) is a member of *Gamma-coronavirus* of *Coronaviridae* family. IBV can cause pathological lesions in different organs varying from respiratory tract to kidney and gonads, resulting in heavy economic losses.

IBV is an enveloped, non-segmented, single-stranded, positive-sense RNA virus, with a genome approximate 27.6 kb in length. In the genome, two-thirds of the 5' region encodes two polyproteins, polyprotein 1a (pp 1a) and polyprotein 1ab (pp 1ab), in which pp 1ab is an extension of pp 1a by a minus 1 frameshift translation mechanism [1]. These two polyproteins are post-translationally processed into fifteen non-structural proteins (nsp 2–16, IBV does not have nsp 1 that is found in some other coronaviruses). The processing of pp 1a/1ab is performed by two virus-encoded proteases: papain-like protease (PL^{pro}) contained in the nsp 3 and main protease (M^{pro}) contained in the nsp 5 [2]. Then these fifteen mature non-structural proteins take part in the process of viral RNA replication and other aspects of viral pathogenesis. Recently, some researches in reverse genetics had found that these nsps may play vital roles in pathogenesis [3, 4]. The remaining one-third of IBV genome encodes four structural proteins: spike glycoprotein (S), small membrane

Edited by Juergen A Richt.

Xuan Wu and Xin Yang have contributed equally to the research.

✉ Hongning Wang
whongning@163.com

¹ School of Life Science, Sichuan University, Animal Disease Prevention and Food Safety Key Laboratory of Sichuan Province, “985 Project” Science Innovative Platform for Resource and Environment Protection of Southwestern, Key Laboratory of Bio-resources and Eco-environment, Ministry of Education, 29# Wangjiang Road, Chengdu, China

² School of Public Health, Xinxiang Medical University, Xinxiang, Henan, China

protein (E), membrane glycoprotein (M), phosphorylated nucleocapsid protein (N), and four accessory proteins: 3a, 3b, 5a, and 5b. The S glycoprotein has been demonstrated to be the determinant of cell tropism [5], and the S1 subunit of S glycoprotein can induce neutralizing and serotype-specific antibodies [6, 7]. The M and E proteins are two membrane-binding proteins. The N protein has the activity to bind viral RNA forming ribonucleoprotein (RNP). All these four structural proteins together with the viral RNA constitute the enveloped virion. The four accessory proteins have been demonstrated to be dispensable for virus replication; their actual functions are still undefined [8, 9].

IBV was firstly detected in the USA as the Massachusetts (Mass)-type in early 1940s. To date, numerous IBV strains have been defined worldwide. IBV has high mutation frequency. Point mutations, insertions, deletions, and recombination are of common occurrence in the IBV genome, which are responsible for the emergence of new IBV strains [10–14]. In this present study, a novel virulent strain was isolated and designated as SAIBK2 (abbreviation of Sichuan Avian Infectious Bronchitis Kidney 2). To investigate the origin of SAIBK2 and the reason for its increased virulence, the complete genome was sequenced and the complete sequence analyses were performed. The results revealed that SAIBK2 originated from multiple recombination events among a YN-, a YX10-, and a Mass-like strains and these recombination events may account for its increased virulence.

Materials and methods

Virus isolation and amplification

Throughout 2011, 435 layer chickens suspected to be infected with IBV were collected from 15 chicken farms from 8 areas in Sichuan province, China. All these flocks were vaccinated with live attenuated H120 vaccine. 53 of the 435 chickens were confirmed to be infected with IBV. Of the 53 disease chickens, 12 were infected by a novel virulence IBV strain which was designated as SAIBK2. 12 tissue samples of swollen kidneys from diseased chickens, infected with SAIBK2, were prepared as 10 % w/v tissue suspensions in 0.1 % phosphate-buffered saline (PBS). Then the tissue suspensions were clarified by centrifugation at 1500 g at 4 °C for 10 min, and filtered by 0.22- μ m membrane filters. The supernatant 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 fluid was collected at 72 h post-inoculation and stored at –80 °C.

Pathogenicity studies

Twenty-one-day-old SPF chicks were used for assessing the pathogenicity of SAIBK2, ten in experiment group and ten in control group, and kept in two isolators with negative pressure. The experiment group was inoculated by ocular administration at one-day old with a dose of 10^5 EID₅₀/0.1 mL SAIBK2 per chick, and chicks of the control group were mock-inoculated with 0.1 mL sterile allantoic fluid. All chicks were observed daily for signs of disease and death for 30 days after infection. Dead chicks were dissected for the observations of organs.

Viral RNA extraction, RACE, and RT-PCR

Primers were designed on the base of IBV complete genome sequences from GenBank (Table 1). 5 primers were used for RACE PCR to amplify 5' and 3' terminal sequences, and 17 pairs of primers were used to amplify the remaining parts of SAIBK2 genome (Table 2).

Viral RNA was extracted from virus-infected allantoic fluid with TRIzol (Invitrogen, USA) following the manufacturer's instructions. Rapid amplification of cDNA ends (RACE) and reverse transcription were performed using 3'/5' RACE kit (TaKaRa, Japan) and SuperScript III First-Strand Synthesis System (Invitrogen, USA) following the manufacturer's instructions.

PCR was conducted with 2 μ L cDNA as template in a total of 25 μ L reaction volume containing 2.5 μ L of $10 \times$ LA buffer, 5 μ L of 2.5 mM dNTPs, 2 μ L of 25 mM MgCl₂, 10 pmol of each primer, and 0.3 μ L of LA Taq polymerase (TaKaRa, Japan). The PCR parameters included an initial denaturation for 5 min at 94 °C followed by 30 cycles of denaturation at 94 °C for 50 s, annealing at 53 °C for 50 s, and extension at 72 °C for 1–3 min depending on the sizes of the products and a final extension at 72 °C for 10 min.

Cloning and sequencing of the genome sequence

17 overlapped PCR fragments spanning the entire genome, 5' and 3' terminal sequences (19 fragments altogether) were purified from 0.8 % agarose gel with QIAquick PCR Purification Kit (QIAGEN Inc., Valencia, CA). The purified products were cloned into the pMD19-T Vector (TaKaRa, Japan). Nucleotide sequences of each fragment were determined by consensus sequence of three, at least, positive clones. The nucleotide sequences of the positive clones were determined by Sangon Biological Engineering Technology and Services Co., Ltd. Sequences were assembled into complete genome sequence using SeqMan II program of DNASTar software package (DNASTar, Madison, WI).

Table 1 Sequences used in this study

Strains	GenBank accession no.	Strains	GenBank accession no.
California_99 ^{a,b}	AY514485.1	CK/CH/LSD/05I ^a	EU637854.1
Arkansas_DPI ^{a,b}	GQ504720.1	ck/CH/LHLJ/140906 ^a	KP036502.1
ck/CH/LJL/111054 ^{a,c}	KC506155.1	Connecticut_vaccine ^{a,b}	KF696629.1
Delaware_072 ^{a,b}	GU393332.1	IBVUkr27-11 ^a	KJ135013.1
Georgia_1998 ^{a,b}	GQ504722.1	4/91_vaccine ^{a,b}	KF377577.1
H120 ^{a,b,c}	FJ888351.1	YN ^{a,b,c}	JF893452.1
ck/CH/LHB/121040 ^a	KJ425495.1	ck/CH/LDL/97I ^a	JX195177.1
H52 ^{a,b}	EU817497.1	ck/CH/LHB/100801 ^a	JF330898.1
Peafowl/GD/KQ6/2003 ^a	AY641576.1	JMK ^a	GU393338.1
M41 ^a	DQ834384.1	Holte ^{a,b}	GU393336.1
ck/CH/LHLJ/100902 ^a	JF828980.1	Gray ^{a,b}	GU393334.1
CK/CH/XDC-_2/2013 ^a	KM213963.1	BJ ^a	AY319651.1
SDZB0808 ^a	KF853202.1	CQ04-1 ^a	HM245924.1
Beaudette ^b	AJ311317.1	DY07 ^a	HM245923.1
SDIB821/2012 ^a	KF574761.1	ITA/90254/2005 ^a	FN430414.1
GX-NN09032 ^a	JX897900.1	A2 ^a	EU526388.1
YX10 ^{a,b,c}	JX840411.1	SAIBK ^{a,b}	DQ288927.1
GX-YL9 ^a	HQ850618.1	Mass41_1985 ^b	FJ904723.1
Sczy3 ^a	JF732903.1	ck/ZA/3665/11 ^a	KP662631.1
CK/SWE/0658946/10 ^a	JQ088078.1	Mass41_2006 ^{b,c}	FJ904713.1
TW2575/98 ^a	DQ646405.2	HKU11-934 ^d	FJ376619.2
SC021202 ^a	EU714029.1	HKU13-3514 ^e	NC_011550.1

^a IBV sequences involved in phylogenetic analysis

^b IBV sequences involved in sequence identity analysis

^c IBV sequences involved in recombination analysis

^d Bulbul coronavirus sequence

^e Munia coronavirus sequence

Genome sequence comparison, phylogenetic, and recombination analysis

42 IBV sequences and 2 bird coronavirus strains (as out-group) downloaded from GenBank were used in this study. Sequence comparisons and nucleotide identities of complete or partial genome were computed using MegAlign clustal V method. Phylogenetic analysis was conducted using MEGA version 6 with the neighbor-joining method. The bootstrap values were determined from 1000 replicates of the original data. The IBV sequences were used in this study and their accession numbers are listed in Table 1.

To identify recombination events, complete genomic sequence of SAIBK2 were compared with those of YX10, YN, and three Mass-type strains (H120, Mass41 2006, and ck/CHLJL/111054), and the results were used for Similarity Plotting analysis using the SimPlot program version 3.5.1 with a window size of 200 bp and a step size of 20 bp. Pairwise comparisons between SAIBK2 and those above-mentioned five viruses were performed to confirm the precise recombination breakpoints. Phylogenetic trees

based on every recombinant fragments were constructed to avoid phylogenetic biases derived from ignoring recombination events.

Results

Pathogenicity studies

Obvious clinical signs were observed in all SAIBK2-infected chicks from 3–19 days postinfection (dpi), including ruffled feathers and huddled together. Five of the ten chicks died: one on 4 dpi, two on 6 dpi, and two on 7 dpi, in the SAIBK2-infected group. Gross lesions in the dead chicks were confined mainly to the kidneys. No overt disease was observed and none of the chicks died in control group.

Genome organization of isolate SAIBK2

To obtain the genome sequence of SAIBK2 strain, 17 overlapping large fragments, ranging from 1309 to

Table 2 Primers used in this study

Primers	Location	Forward primer (5'–3')	Reverse primer (5'–3')
1	218-1526	CATCCGTTGCTTGGGCTACCTAG	GCAGACCAACCTTCTGGTTCC
2	1439-3082	CACAAGTTGTTGTTCTGTTGGC	GAGCGTCCTTATGAACTATAC
3	2545-3801	GCGGGTGGCAAGACTGTCAC	AATGCCTTTGACATAAGAAGG
4	3738-5606	GGACTATGTTAAGAAACATGG	CCATCTCTACTTTTATCGGC
5	5434-7340	GTTGATGGGTAACTATGGGC	GTACCACTAACGATACCACC
6	7219-9220	TCAGCGACTGTCAAGTCAGG	TACCACCATAAGAGCAAGC
7	9171-10949	GAAAGCAAATTGTGGTGATAG	GAGGGAATGTGTGAAAACCTC
8	10815-12586	GGACAATTTGTTGGGTATGC	GTTTATAATTACTAGGAGTGG
9	12412-14028	CCTGATGTTGTAGAGCGAGCC	AAGAATGGACACACCTGCCAC
10	13780-15652	TTTGAATGTTATGAAGGTGG	TAGTCTTACCAGGTTCCACG
11	15520-16868	TCATTAAGACGCTTTGCTG	ATGACTACAAGTATACCACGC
12	16738-19004	GTAGACTCTTCAAGGTTTC	TAAACATACAGATTCGCTCC
13	18402-20306	GGCTTCTTATAATGCAGCTG	AAATTCACAACCTGGTGTTC
14	20070-22284	CAAGATTGTGCATGGTGGAC	TATGTTATCACAACAGGACC
15	21262-22963	ACACAAACAGCTCAGAGTGG	CTTGAATTGCATCAAGTTGC
16	22788-24887	CATTGGTCATATGCAGGAAGG	GGAGTATTGAACCTACGGCATT
17	24627-27112	TGGATACGCAACTAGGAGTGC	CATCATCAAATTCAGCTGTGC
5' region ^a	1-449	CATGGCTACATGCTGACAGCCTA	TACAAACGTCGTATAGCCGACC
3' region ^b	26942-27669	GTAAGACCAAAATCACGCCAAG	AAGCAGTGGTATCAACGCAGAGT
3' RT ^c	poly-A	AAGCAGTGGTATCAACGCAGAGTAC(T) ₃₀	

^a PCR Primers used in 5' RACE

^b PCR Primers used in 3' RACE

^c RT Primer used in 5' RACE

2486 bp, were sequenced and assembled using DNASTar software.

The complete genome sequence of SAIBK2 obtained in this study was submitted to Genbank database under the accession number of KU317090. The complete genome of SAIBK2 strain consists of 27669 nt, excluding the poly-A tail at the 3' end, which encodes 6 genes. Gene 1 has a length of 19892 nt, consisting of 2 overlapping ORFs, ORF1a, and ORF1b which are separately 11859 and 8105 nt in length. There are 50 nt overlap between Gene 1 and Gene 2. Gene 2 with a length of 3501 nt has a single ORF, encoding S glycoprotein which is cleaved into 2 subunits S1 and S2 with 1620 and 1881 nt in length, respectively (encoding 540 and 626 aa). Between Gene 2 and Gene 3, there are 4 nt overlap. Gene 3 has 675 nt consisting of 3 overlapping ORFs, encoding 2 accessory proteins (3a and 3b), and structural protein E with 174, 192, and 330 nt in length, respectively (encoding 57, 62, and 109 aa). Between 3a and 3b, 3b and E, there are 1 and 20 nt overlap, respectively. There are 29 nt overlap between Gene 3 and Gene 4. Gene 4 contains 678 nt, encoding M protein of 225 aa. A 361 nt non-encoding region is detected between Gene 4 and Gene 5. Gene 5 containing 443 nt

consists of 2 overlapping ORFs, ORF5a and ORF5b, which are 198 and 249 nt in length, respectively (encoding 65 and 82 aa). There are 58 nt overlap between Gene 5 and Gene 6. Gene 6, encoding N protein, contains 1230 nt (encoding 409 aa). Untranslated regions (UTRs) at 5'-end and 3' end of the genome are 528 and 505 nt in length, respectively (Fig. 1).

Genome sequence analyses

Genome sequence comparison

BLAST searches with the complete genome sequence of SAIBK2 showed that it is most similar to that of YX10, sharing 96.1 % nucleotide identity. Compared with YX10, there are a 3-nt insertion at 341-343 nt, a 3-nt deletion at 3269-3298 nt, and a 4-nt deletion at 25554-25557 nt in SAIBK2 genome. To further evaluate the nucleotide sequence identities between isolate SAIBK2 and other IBV strains, sequence alignments between the complete or partial genomic sequences of SAIBK2 and 16 IBV strains from GenBank (Table 1) were performed. As shown in Table 3, SAIBK2 shows the highest similarity with YX10

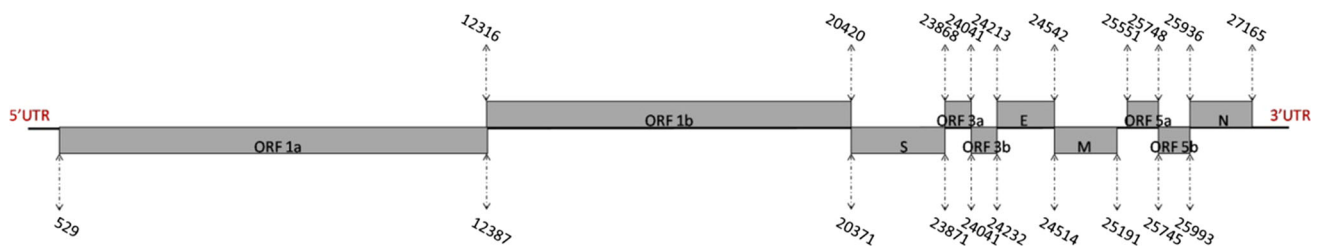


Fig. 1 Expression of ten ORFs of the SAIBK2 genome. The gray boxes represent ORFs, and the numbers above or under them show the positions in SAIBK2 genome

Table 3 Percentage of sequence identities of different regions of SAIBK2 compared with other IBV strains

Strain	Complete genome	Gene 1						
		Gene 1	Nsp2	Nsp3	Nsp4	Nsp5	Nsp6	Nsp7
Beaudette	86.7	87.3	85.2	85.8	85.6	86.8	85.8	92.4
Gray	87.0	87.5	84.3	86.1	87.5	87.2	85.9	92.0
Holte	85.7	87.6	86.0	85.9	86.4	86.6	85.4	92.4
YN	90.8	91.9	92.8	93.9	92.5	85.3	86.7	94.4
Mass41_1985	86.5	87.3	84.9	86.0	85.2	87.8	84.8	92.0
4/91_vaccine	86.3	86.7	85.3	84.3	86.0	85.6	84.0	92.4
Connecticut_vaccine	87.1	87.3	84.5	85.6	86.2	86.9	85.3	92.4
Mass41_2006	87.2	87.3	84.3	85.6	86.7	86.9	85.3	92.4
California_99	86.9	87.2	85.4	85.2	87.0	86.2	85.1	92.0
Arkansas_DPI	87.1	87.4	85.3	85.1	86.8	86.9	86.1	92.0
Delaware	84.6	86.8	85.5	84.1	86.8	85.6	84.2	92.4
Georgia_1998_Vaccine	85.0	87.2	85.0	85.4	86.7	83.4	85.9	92.0
H52	86.4	87.0	85.5	84.6	86.9	85.6	84.5	92.0
H120	86.5	87.1	85.3	84.7	86.6	85.6	84.2	92.0
YX10	96.1	96.1	97.2	93.6	91.8	99.7	98.2	96.8
SAIBK	89.1	90.2	93.0	92.7	86.5	86.6	86.4	92.0

Strain	Gene 1								
	Nsp8	Nsp9	Nsp10	Nsp11	Nsp12	Nsp13	Nsp14	Nsp15	Nsp16
Beaudette	86.5	89.5	90.8	90.9	88.7	89.6	91.6	86.7	85.8
Gray	87.8	88.9	90.1	92.4	88.8	89.4	91.2	87.8	86.6
Holte	88.3	87.4	89.4	92.4	89.6	89.6	90.3	87.3	88.6
YN	87.6	91.3	91.5	97.0	95.3	89.9	89.1	88.8	94.7
Mass41_1985	87.0	88.9	90.1	95.5	89.7	89.3	90.4	86.5	86.2
4/91_vaccine	87.1	88.6	90.1	89.4	89.2	89.3	88.8	86.2	86.7
Connecticut_vaccine	88.1	89.8	88.7	93.9	89.3	89.9	90.4	87.6	85.8
Mass41_2006	87.0	89.2	91.3	92.4	89.1	89.8	90.0	88.1	86.6
California_99	87.6	88.3	88.3	93.9	89.3	89.6	90.2	87.4	85.3
Arkansas_DPI	88.1	89.8	89.9	92.4	88.8	89.9	90.3	87.5	88.2
Delaware	87.1	88.3	91.3	89.4	89.1	88.8	89.3	87.2	86.8
Georgia_1998_Vaccine	87.8	89.8	91.3	93.9	88.9	89.5	90.5	87.6	85.9
H52	86.7	88.6	90.1	95.5	89.7	89.2	90.2	86.6	86.4
H120	87.3	89.5	91.5	89.4	89.1	89.6	90.2	88.0	86.6
YX10	94.9	97.3	98.6	95.5	97.7	98.0	99.0	97.1	93.7
SAIBK	86.7	88.9	90.8	89.4	89.4	89.9	88.8	88.7	93.1

Table 3 continued

Strain	Gene 2		Gene 3			Gene 4	Gene 5		Gene 6
	S1	S2	3a	3b	E	M	5a	5b	N
Beaudette	77.4	85.4	83.3	76.9	87.9	90.1	92.9	97.6	89.1
Gray	75.3	86.6	85.6	74.9	86.1	89.2	100.0	98.0	89.5
Holte	76.1	86.0	79.9	74.4	87.9	88.2	93.4	95.2	89.0
YN	79.8	93.2	93.1	75.9	87.6	91.3	83.8	96.0	87.6
Mass41_1985	77.4	85.8	83.3	77.4	88.2	90.1	89.9	97.2	90.4
4/91_vaccine	78.1	85.1	84.5	72.3	81.2	90.6	89.9	97.2	88.5
Connecticut vaccine	76.6	86.4	82.2	73.3	86.1	90.9	100.0	100.0	90.9
Mass41_2006	77.4	86.1	82.2	73.3	85.8	91.0	100.0	99.2	90.9
California_99	75.5	85.4	82.2	73.3	88.8	91.6	100.0	98.8	89.9
Arkansas_DPI	75.6	86.4	86.2	73.8	85.8	90.7	100.0	99.6	90.7
Delaware	57.9	75.0	82.2	73.3	87.0	89.4	98.0	99.2	89.5
Georgia_1998_Vaccine	58.6	74.8	82.8	74.4	88.8	90.4	99.5	100.0	90.8
H52	77.7	85.4	82.8	77.4	87.6	90.3	93.9	96.8	88.7
H120	77.4	85.6	83.3	77.4	86.7	90.0	94.4	97.2	88.3
YX10	98.7	98.0	99.4	93.8	97.3	97.5	83.8	92.8	90.1
SAIBK	80.3	93.1	91.4	76.9	86.7	90.1	84.3	96.4	88.2

The highest nucleotide identities of different regions are indicated in bold values

in most regions (nsp 2, nsp 5–10, nsp 12–15, Gene 2, Gene 3, and Gene 4). Nsp 3, nsp 4, nsp 11, and nsp 16 of SAIBK2 were most similar to those of YN. ORF5a of SAIBK2 is completely the same as that of Gray, Connecticut vaccine, Mass41_2006, California_99, and Arkansas_DPI. ORF5b of SAIBK2 is completely the same as that of Connecticut vaccine and Georgia_1998_Vaccine. Gene 6 of SAIBK2 is most similar to that of Connecticut vaccine and Mass41_2006.

Phylogenetic analysis

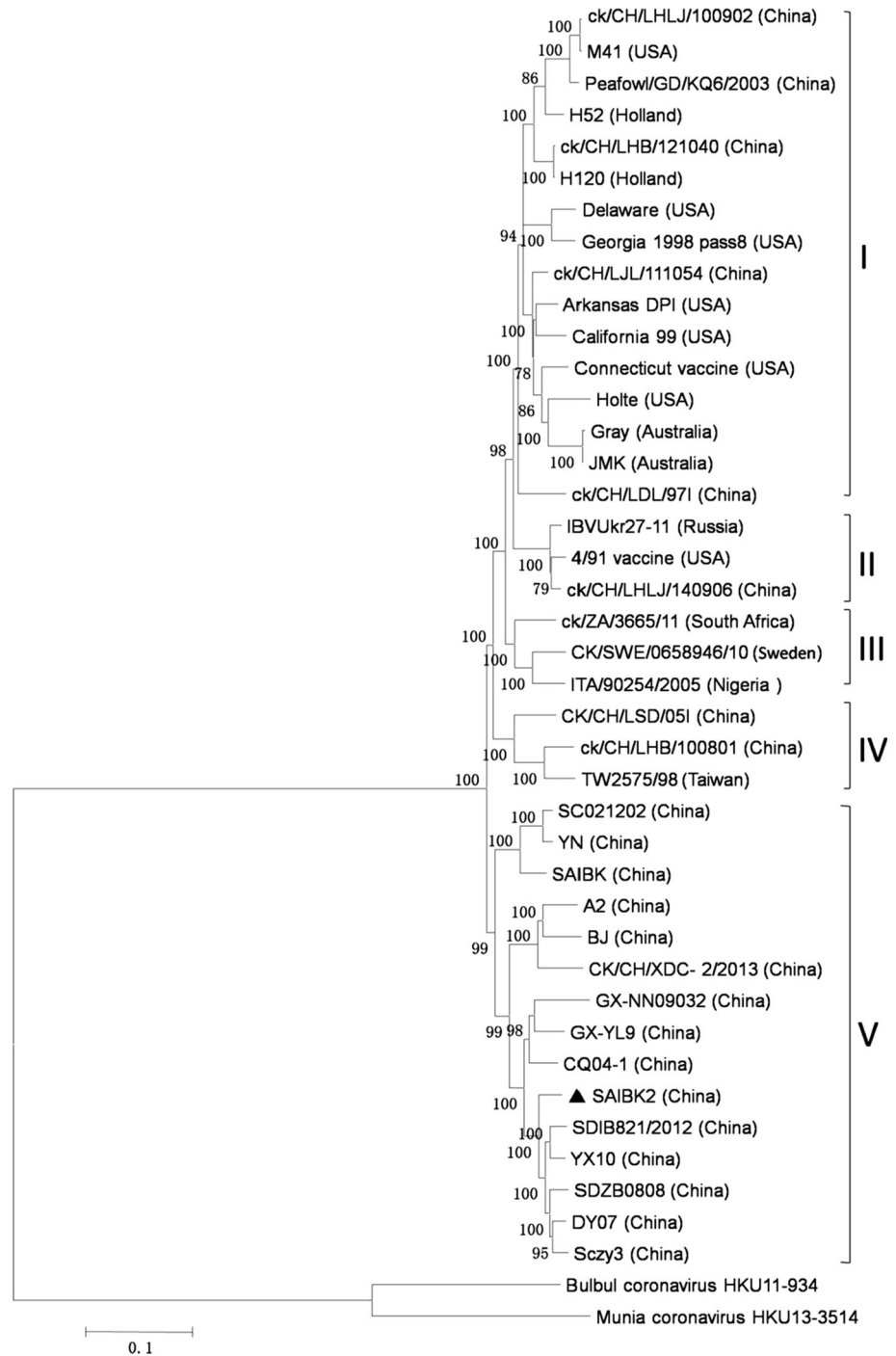
In order to study the genetic relatedness between SAIBK2 and other IBV strains, the sequence of SAIBK2 was compared with complete genomic sequences of 39 IBV and 2 bird coronavirus from GenBank (Table 1) to construct phylogenetic tree. Based on the Phylogenetic tree, we grouped IBV strains into 5 clades (Fig. 1). Group I was constituted by Mass-genotype strains. Group II was 4/91 genotype. The third group is constituted of two Africa strains and one Europe strain. The fourth group was formed by two China strains and one Taiwan strain. Isolate SAIBK2 was clustered into group V, which was exclusively formed by China strains. This result revealed that SAIBK2 belongs to the most predominant genotype in China (Fig. 2).

Recombination analysis

To detect possible recombination events within SAIBK2 genomic sequence, SimPlot analysis was used to show the

consecutive nucleotide identity between SAIBK2 and other IBV complete sequences from GenBank. The results indicated that SAIBK2 may be a mosaic formed by a YX10-, a YN-, and a Mass-like strain (Fig. 3). To find out the regions where the template switches (breakpoint) have taken place, pairwise comparison between the complete genomic sequences of SAIBK2, YX10, YN, and Mass41_2006 was performed, and six breakpoints were found as shown in Fig. 4. The first recombination breakpoint is located in the nsp 3 (5874–5887 nt), the second is located in the nsp 4 (8103–8116), the third and fourth breakpoints are located in nsp 16 (19843–19868 and 20326–20345 nt, respectively), the fifth breakpoint is located in the non-coding region between ORF4 and ORF5 (25321–25328 nt), the last is located in N gene (26225–26247 nt). These results revealed that the YX10-like strain is the major parent of SAIBK2 and the first two recombinant regions (5888–8102 and 19869–20325 nt) are from the YN-like strain, the third region (25329–26224 nt) is from the Mass-like strain. To further verify these three above-mentioned recombinant events, the sequence identity (Table 4) and phylogenetic analyses (Fig. 3b) among corresponding regions of SAIBK2, YN, YX10, and three Mass-like strains were performed. The results showed that SAIBK2 has higher similarity with YX10 than YN and Mass-like viruses in the regions of 1–5873, 8111–19842, 20346–25320, and 26248–27713 nt. In contrast, SAIBK2 is more similar to YN than YX10 and Mass-type strains in the regions of 5888–8102 and 19869–20325 nt, and is more similar to Mass-like strains in the region of

Fig. 2 Phylogenetic tree by neighbor-joining method (bootstrapping for 1000 replicates with its value >70 %) based on complete genome. Sequence of SAIBK2 was labeled by filled triangle



25329-26224 nt (Table 4). In line with the results obtained from the Simplot analysis and identity analysis (Table 4), phylogenetic trees constructed using the corresponding gene fragments also showed the same results (Fig. 3).

These results suggested that SAIBK2 arose from a series of homologous RNA recombination events from multiple template switches involving a YX10-, a YN-, and a Mass-

like strain, and that nsp3, nsp4, nsp12, ORF5, and N gene are involved in the three recombinant events.

Discussion

According to the result of sequence identity analysis based on complete genome, SAIBK2 was most similar to YX10.

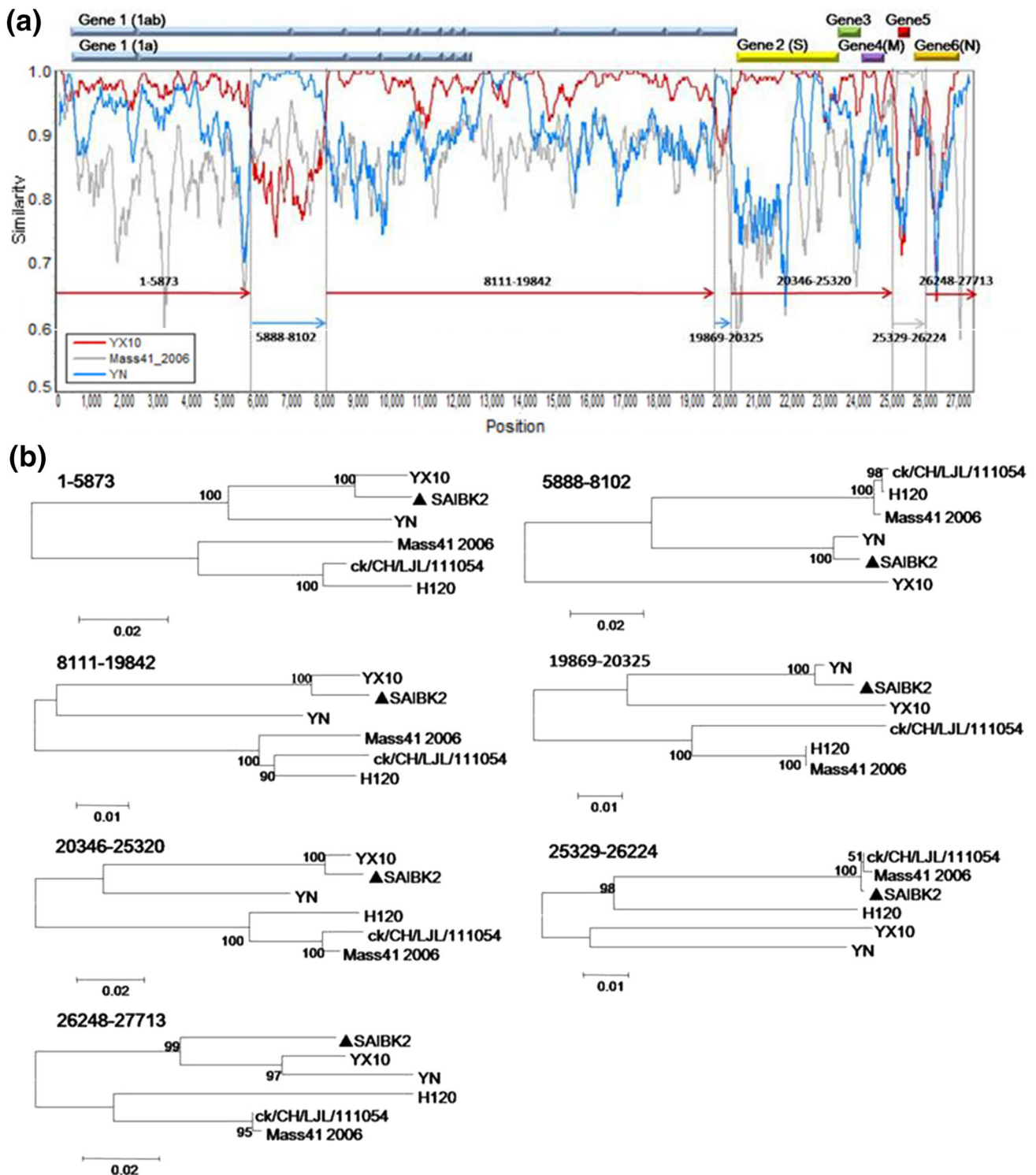


Fig. 3 Recombination analysis of isolate SAIBK2. SimPlot analysis was performed using SAIBK2 as the query sequence (**a**). Bars at the top represent relative positions of Genes of IBV. The dummy lines show the deduced breakpoints. The arrows show different regions and their colors are the same as those of the parent viruses. The numbers above those arrows show corresponding nucleotide positions in the alignment. The y axis shows the percentage similarity within a sliding window of 200 bp centered on the position plotted, with a step size

between plots of 20 bp. Phylogenetic trees of genome positions 1-5873, 5888-8102, 8111-19842, 19869-20325, 20346-25320, 25329-26224, and 26248-27713 among SAIBK2, YX10, YN, and three Mass-type viruses (H120, ck/CH/LJL/111054, and Mass41_2006) (**b**). Phylogenetic trees were constructed by neighbor-joining method (bootstrapping for 1000 replicates). Sequence of SAIBK2 was labeled by filled triangle

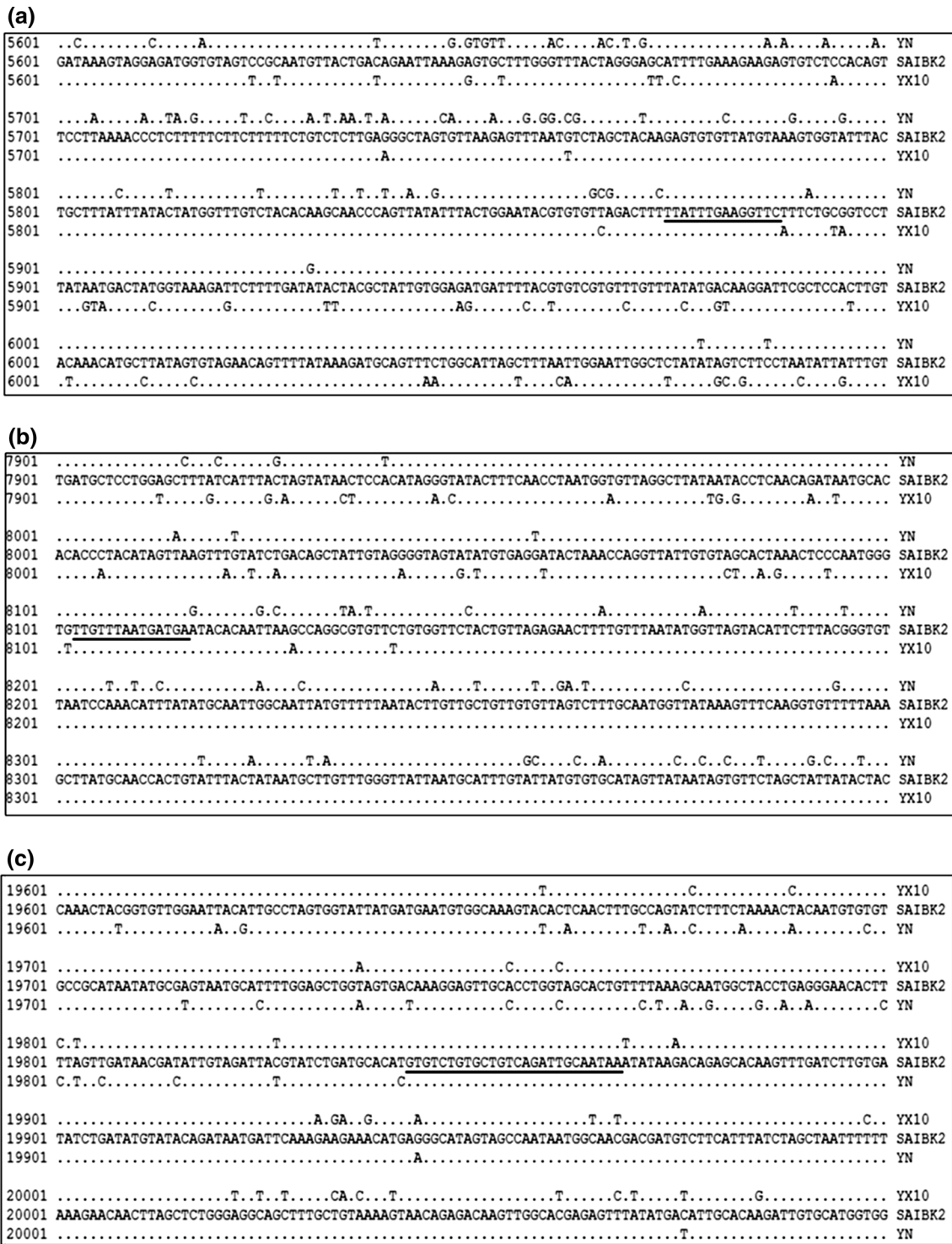


Fig. 4 Analysis of breakpoint. **a–d** are breakpoint analysis between SAIBK2, YX10, and YN. **e, f** are breakpoint analysis between SAIBK2, YX10, and Mass41 2006. Numbers in the left show the

nucleotide position in the alignment. Only the nucleotides which differ from those of SAIBK2 are listed. The breakpoints are *underlined*. The deleted nucleotides are indicated by *prime*

Subsequently, sequence identity analyses based on partial genomic regions showed that SAIBK2 shared higher similarities with other reference strains than YX10 in some partial genomic regions, as shown in Table 3. These results

implied that recombination events might happen during the evolution of SAIBK2.

In coronaviruses, including IBV, recombinant events efficiently happen and it has been well documented to be

(d)

```

20101 .....T..T.....T..A.....T..A.....G.....T..T.....C..... YX10
20101 ACAATGTTTTGTACAGCAGTAAATGCCTCATCTTCAGAGGCATTCTTCTGGGTATTAAATATTGGGTGCAAGTGAAAAGGTGAAAGTTAGTGGAAAA SAIBK2
20101 .....G..... YN

20201 ..A.....C.....C.....T.....C.....A.....C.....A..C..... YX10
20201 CGTGCATGCAAATTATATATTTGGAGGAATTGTAAGTATTACAACTCTGCTTATAGTGCAATTTGATGTTGCTAAGTTTGATTGAAATTAAGC SAIBK2
20201 .....C.....T..... YN

20301 ...G.....A.....A..CT.....G.....G..... YX10
20301 AACACCAGTTGTGAATTTGAGAAAGAACAAAGACCGACTTAGTTGTTAATCTACTAAAGAACGGCAAGTTATTGATTAGAGATGTTGGGAATTCAC SAIBK2
20301 .....A.C.....T.....G..... YN

20401 .....C.....T.....A.. YX10
20401 GTTTTTAGTACCATTGTTGTGCACTATGTAGTGCAAAATTTGTTTATTCTGGCAATAATATGTGTACTACTACCAAGTGCCTATAGGCCTCCAGAT SAIBK2
20401 ..AA.....TG.....T.....C.C.C.''''.A.ACTC.CG.....T.....A..A..A.. YN

20501 .....C.....G..G.....'''' YX10
20501 GGATGGCATCTGCAAGGGGGTCTTATGCACTAGTGAATCTACTAATATAATAATAATGCCGGTCTGCAATGAGTGCACCTGTTGGTGTATT'''' SAIBK2
20501 .....T.....A.A.C.....A..CG.C..T.GC....TT...A.....C..TTTA YN

```

(e)

```

25101 ..... YX10
25101 GGTGCAGAAATATACCTGGTACCAAAGCGGAATAAGAAAAGTTGCTACATTTGCTCTAGCAAAGCAGTCAGTAGACACTGGCCAGCTAGAAAGTGA SAIBK2
25101 .....C..... Mass41_2006

25201 .....T.....T..... YX10
25201 GCAACAGGAGGAAGTAGTCTTTACACATAAATGTGTGTGTAGAGAGTATTTAAAATTACTCTTTGATAGTGCCTCTATTCTAAGAGCCATAAGAGTA SAIBK2
25201 .....G.C..T...A.....T.....C..... Mass41_2006

25301 .....A.....C.....T..... YX10
25301 TTCATTTGAGGATTTCAACATAAATCCTCTTTGTTTATACTCTCTTTTCAAGAGCTATTATTTAAAAACAGTTTTCCACTCTTTTGCCAAAAAC SAIBK2
25301 ..T.....A..T..T..... Mass41_2006

25401 ..G.....T....AG...C...A.....T.....AC..CA...T..G.....C..GT..A.....C..G YX10
25401 TATTGTTGTTAAGCGTGTACCTTTCAAGTGATAAATGGAAGGCTACTACGAGGAACACCAGTTTCCAAAAGGTTGTTGATAGATGTTGTTCAAT SAIBK2
25401 ..... Mass41_2006

25501 .....GG....A...G.C..A.....A..A..TC..G...T..A.....GGAT.....G.....G..G..A..AA.TACGGA..... YX10
25501 TATAAGAAAGATTAGAATAATTAAGCCCACTACACTTATTCTATAAGAGGCGTTTATCTTACAAACGCTTAACAAATACGGACG''''ATGAAAT SAIBK2
25501 .....T.....''''..... Mass41_2006

```

(f)

```

26001 .....T..T.....A.....C.....A..G.....A.....G.....T..G..... YX10
26001 GAAAGACAGACGCCCCAGCGCAGTCATCAAACCTAGGAGGACCAAGCCACTAAAGTTGGTCTCTCTGGAATGCATCTTGGTTTCAAGCAATAAAGC SAIBK2
26001 ..... Mass41_2006

26101 .....G..G...G.....T.....G.....G..... YX10
26101 CAAGAAGCTAAATTCACCTCCACTAAGTTGAAAGTAGCGGTGTCCTGATAAGGAAATCTTAAAACAGCCAGCAACATGGATAGTGGAGCGCCAA SAIBK2
26101 ..... Mass41_2006

26201 .....A.....T.....T.....C..... YX10
26201 GCTAGGTTTAAAGCCAGGTAAGGCGGAGAAAACAGTCCAGATGCGTGGTATTCTTATTACATGGAACAGGACCAGCGCTGACCTGAATGGGGTG SAIBK2
26201 .....C.....T.....T.....C..... Mass41_2006

26301 ..T.....T.....T.....C.. YX10
26301 ATCCTCAGGATGGTATAGTGTGGTTGCTGCTAAGGGTGTGATGTAATAATCCAGATCCAATCAGGGTACAAGAGATCCTGACAAGTTTGACCAATATCC SAIBK2
26301 ..AGC..A.....A.....T.....T.....T..C.....G..C..... Mass41_2006

26401 G.....C..A.....T..A.....C.....G..T..A.....T..G.....A YX10
26401 ACTACGATTTTACAGATGGGGACCTGACGTAATTTCCGTTGGGATTTCAITCCTCTGAATCGTGGTAGGAGTGAAGATCAACCGCAGCTCATCAGCG SAIBK2
26401 .....C..G..C..A.....T.....C.....A.....T.....A Mass41_2006

```

Fig. 4 continued

responsible for the emergence of new variants [11, 14–16]. To investigate whether recombination events took place during the evolution of SAIBK2. SimPlot analyses were conducted and three recombination events were illustrated in the results (Fig. 3a). The study of sequence identities (Table 4) and phylogenetic analyses (Fig. 3b) based on

corresponding regions both further confirmed the results of SimPlot analyses. As the results of sequence identity and SimPlot analyses exhibited, YX10 is the major parent of SAIBK2. YX10 was isolated in Zhejiang province, China in 2010, and caused 25 % mortality in SPF chicken [17, 31]. Compared with YX10, mortality in SPF chicken

Table 4 Percentage of sequence identity of corresponding regions as indicated among SAIBK2, YX10, YN, and three Mass-type viruses (H120, ck/CH/LJL/111054, and Mass41_2006)

Virus	Position in alignment						
	1-5873	5888-8102	8111-19842	19869-20325	20346-25320	25329-26224	26248-27713
YX10	97.5	84.2	98.0	89.9	98.1	86.8	93.2
YN	92.0	98.6	90.3	98.9	87.2	86.3	90.1
H120	84.6	88.7	88.4	88.0	83.1	90.4	83.8
ck/CH/LJL/111054	85.5	88.6	88.5	86.7	83.0	99.9	86.6
Mass41_2006	85.0	88.8	88.6	88.0	83.3	99.7	87.7

caused by SAIBK2 was higher (50 %). Considering the recombination events that happened in the procedure of formation of SAIBK2 genome, it is with reason to speculate that the acquisition of the regions from the YN- and Mass-like strains should account for the increased virulence of SAIBK2 compared with YX10. YN was isolated in Yunnan province, China in 2005. It can result in 65 % mortality in SPF chickens [18]. According to the result of SimPlot analyses, SAIBK2 genomic regions of 5888-8102 and 19869-20325 nt were from YN-like strain, and these two regions encode several highly hydrophobic stretches and S-adenosyl-L-methionine (SAM)-dependent methyltransferases (MTase), respectively [19–23]. Meanwhile, Mass-based live attenuated IBV vaccines, including H120, are used extensively in China, and SAIBK2 genomic region of 25329-26224 nt was from Mass-like strain; this region contains Gene 5 and N-terminal domain of N protein. All the three regions above-mentioned involved in recombination events have been documented to act important roles in virus replication [24–30]. Nevertheless, IBV pathogenicity is regulated by multiple genes; these above-mentioned recombination events, together with mutations in the SAIBK2 genome, altered the replication efficiency and resulted in the increased virulence, as observed. However, further investigations by reverse genetic system should provide more profound insight into this issue and increase the understanding of IBV pathogenesis.

China has a large number of chickens and most of them are kept at a rather high density; this situation increased the odds of chicken multiple-infection by several IBV strains. In this case of SAIBK2, YX10- and YN-like strains are very likely to spread in chicken flocks simultaneously; meanwhile, Mass-type live attenuated vaccine was used in the flocks from which we got the disease chickens. Overall, current results provide convincing evidence that SAIBK2 originated from recombination events, in which a YX10-like strain is the major parent and another two strains, a YN- and a Mass-like strains, are the minor parents.

In conclusion, this work implicates that recombination between IBV strains may account for the increasing of IBV

virulence. In addition, multiple IBV infection may be a requirement for IBV recombination. Consequently, it continues to be important to put emphasis on surveillance for better control of IBV.

Acknowledgments The study was supported by the earmarked fund for Modern Agro-industry Technology Research System (CARS-41-K09), Project for Science and Technology Support Program of Sichuan province (2014NZ0020), and Natural Science Foundation of China (31302094).

References

1. I. Brierley, M.E. Bournsnel, M.M. Binns, B. Bilimoria, V.C. Blok, T.D. Brown, S.C. Inglis, *EMBO J.* **6**, 3779 (1987)
2. J. Ziebuhr, V. Thiel, A.E. Gorbalenya, *J. Biol. Chem.* **276**, 33220–33232 (2001)
3. M. Armesto, D. Cavanagh, P. Britton, *PLoS ONE* **4**, e7384 (2009)
4. T. Hodgson, R. Casais, B. Dove, P. Britton, D. Cavanagh, *J. Virol.* **78**, 13804–13811 (2004)
5. R. Casais, B. Dove, D. Cavanagh, P. Britton, *J. Virol.* **77**, 9084–9089 (2003)
6. D. Cavanagh, P.J. Davis, J.K. Cook, D. Li, A. Kant, G. Koch, *Avian Pathol.* **21**, 33–43 (1992)
7. R.L. Parr, E.W. Collisior, *Arch. Virol.* **133**, 369–383 (1993)
8. T. Hodgson, P. Britton, D. Cavanagh, *J. Virol.* **80**, 296–305 (2006)
9. R. Casais, M. Davies, D. Cavanagh, P. Britton, *J. Virol.* **79**, 8065–8078 (2005)
10. Y. Zhang, H. Wang, T. Wang, W. Fan, A. Zhang, K. Wei, G. Tian, X. Yang, *Virus Genes* **41**, 377–388 (2010)
11. T. Zhang, Z. Han, Q. Xu, Q. Wang, M. Gao, W. Wu, Y. Shao, H. Li, X. Kong, S. Liu, *Infect. Genet. Evol.* **32**, 377–387 (2015)
12. P.S. Masters, *Adv. Virus Res.* **66**, 193–292 (2006)
13. X. Liu, H. Ma, Q. Xu, N. Sun, Z. Han, C. Sun, H. Guo, Y. Shao, X. Kong, S. Liu, *Vet. Microbiol.* **162**, 429–436 (2013)
14. S. Liu, Q. Xu, Z. Han, X. Liu, H. Li, H. Guo, N. Sun, Y. Shao, X. Kong, *Infect. Genet. Evol.* **23**, 189–195 (2014)
15. S.W. Thor, D.A. Hilt, J.C. Kissinger, A.H. Paterson, M.W. Jackwood, *Viruses* **3**, 1777–1799 (2011)
16. M.W. Jackwood, T.O. Boynton, D.A. Hilt, E.T. McKinley, J.C. Kissinger, A.H. Paterson, J. Robertson, C. Lemke, A.W. McCall, S.M. Williams, J.W. Jackwood, L.A. Byrd, *Virology* **398**, 98–108 (2010)
17. Y. Xue, Q. Xie, Z. Yan, J. Ji, F. Chen, J. Qin, B. Sun, J. Ma, Y. Bi, *J. Virol.* **86**, 13812–13813 (2012)

18. J. Feng, Y. Hu, Z. Ma, Q. Yu, J. Zhao, X. Liu, G. Zhang, *Emerg. Infect. Dis.* **18**, 1994–2001 (2012)
19. M. Oostra, M.C. Hagemeijer, M. van Gent, C.P.J. Bekker, E.G. Te Lintelo, P.J.M. Rottier, C.A.M. de Haan, *J. Virol.* **82**, 12392–12405 (2008)
20. A. Kanjanahaluethai, Z. Chen, D. Jukneliene, S.C. Baker, *Virology* **361**, 391–401 (2007)
21. S.H. van den Worm, K. Knoops, J.C. Zevenhoven-Dobbe, C. Beugeling, Y. van der Meer, A.M. Mommaas, E.J. Snijder, *J. Virol.* **85**, 5669–5673 (2011)
22. M.M. Lai, R.S. Baric, P.R. Brayton, S.A. Stohlman, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 3626–3630 (1984)
23. Y. Chen, C. Su, M. Ke, X. Jin, L. Xu, Z. Zhang, A. Wu, Y. Sun, Z. Yang, P. Tien, T. Ahola, Y. Liang, X. Liu, D. Guo, *PLoS Pathog.* **7**, e1002294 (2011)
24. J.D. Lewis, E. Izaurrealde, *Eur. J. Biochem.* **247**, 461–469 (1997)
25. B. Schwer, X. Mao, S. Shuman, *Nucleic Acids Res.* **26**, 2050–2057 (1998)
26. S. Kuo, H. Kao, M. Hou, C. Wang, S. Lin, H. Su, *Vet. Microbiol.* **162**, 408–418 (2013)
27. H. Jayaram, H. Fan, B.R. Bowman, A. Ooi, J. Jayaram, E.W. Collisson, J. Lescar, B.V.V. Prasad, *J. Virol.* **80**, 6612–6620 (2006)
28. K.R. Hurst, R. Ye, S.J. Goebel, P. Jayaraman, P.S. Masters, *J. Virol.* **84**, 10276–10288 (2010)
29. N.E. Grosseohme, L. Li, S.C. Keane, P. Liu, C.E. Dann, J.L. Leibowitz, D.P. Giedroc, *J. Mol. Biol.* **394**, 544–557 (2009)
30. S.A. Kopecky-Bromberg, L. Martinez-Sobrido, M. Frieman, R.A. Baric, P. Palese, *J. Virol.* **81**, 548–557 (2006)
31. K. Feng, Y. Xue, J. Wang, W. Chen, F. Chen, Y. Bi, Q. Xie, *Vaccine* **33**, 1113–1120 (2015)