

## Accepted Manuscript

Title: Origin and evolution of LX4 genotype infectious bronchitis coronavirus in China

Author: Wenjun Zhao Mengying Gao Qianqian Xu Yang Xu Yan Zhao Yuqiu Chen Tingting Zhang Qiuling Wang Zongxi Han Huixin Li Lingfeng Chen Shuling Liang Yuhao Shao Shengwang Liu



PII: S0378-1135(16)30657-5  
DOI: <http://dx.doi.org/doi:10.1016/j.vetmic.2016.11.014>  
Reference: VETMIC 7442

To appear in: *VETMIC*

Received date: 16-8-2016  
Revised date: 14-11-2016  
Accepted date: 15-11-2016

Please cite this article as: Zhao, Wenjun, Gao, Mengying, Xu, Qianqian, Xu, Yang, Zhao, Yan, Chen, Yuqiu, Zhang, Tingting, Wang, Qiuling, Han, Zongxi, Li, Huixin, Chen, Lingfeng, Liang, Shuling, Shao, Yuhao, Liu, Shengwang, Origin and evolution of LX4 genotype infectious bronchitis coronavirus in China. *Veterinary Microbiology* <http://dx.doi.org/10.1016/j.vetmic.2016.11.014>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

**Origin and evolution of LX4 genotype infectious bronchitis  
coronavirus in China**

**Origin and evolution of LX4 genotype IBV in China**

Wenjun Zhao<sup>#</sup>, Mengying Gao<sup>#</sup>, Qianqian Xu, Yang Xu, Yan Zhao, Yuqiu Chen,  
Tingting Zhang, Qiuling Wang, Zongxi Han, Huixin Li, Lingfeng Chen, Shuling  
Liang, Yuhao Shao, Shengwang Liu\*

*Division of Avian Infectious Diseases, State Key Laboratory of Veterinary  
Biotechnology, Harbin Veterinary Research Institute, the Chinese Academy of  
Agricultural Sciences, Harbin 150001, the People's Republic of China*

*<sup>#</sup>These authors contributed equally to this work.*

\*Corresponding Author.

Tel: (86) 451-51051698.

Fax: (86) 51051698.

*E-mail: [swliu@hvri.ac.cn](mailto:swliu@hvri.ac.cn)*

Shengwang Liu

Division of Avian Infectious Diseases,  
State Key Laboratory of Veterinary Biotechnology,  
Harbin Veterinary Research Institute,  
the Chinese Academy of Agricultural Sciences,  
Harbin 150001, the People's Republic of China

swliu@hvri.ac.cn (S. Liu)

## Highlights

- ▶ LX4 is one the most important genotypes of infectious bronchitis coronavirus worldwide.
- ▶ Two LX4 genotype viruses have novel genomic organizations which lacked 3a and 5b gene, respectively.
- ▶ Recombination events may be responsible for the emergence of the LX4 genotype.
- ▶ Most of these viruses disappeared likely because they were not “fit” to adaptation in chickens.
- ▶ The “fit” viruses continued to evolve and have become widespread and predominant in commercial poultry.

## ABSTRACT

We investigated the genomic characteristics of 110 LX4 genotype strains of infectious bronchitis viruses (IBVs) isolated between 1995 and 2005 in China. The genome of these IBVs varies in size from 27596 bp to 27790 bp. Most IBV strains have the typical genomic organization of other group III coronaviruses, however, two strains lacked *3a* and *5b* genes as a result of a nucleotide change within the start codon in both the *3a* and *5b* genes. Analysis of our 110 viruses revealed that recombination events may be responsible for the emergence of the LX4 genotype with different topologies. Most of these viruses disappeared (before mid-2005) because they were not “fit” to adaptation in chickens. Finally, those of the “fit”

viruses (after mid-2005) continued to evolve and have become widespread and predominant in commercial poultry. In addition, few of these viruses experienced recombination with those of the vaccine strains at the 3' end of the genome.

**Keywords:** Infectious bronchitis virus (IBV); LX4 genotype (QX-like); Evolution; Topology; Clade I; Clade II

## 1. Introduction

Avian infectious bronchitis (IB) is a highly contagious viral respiratory disease in birds caused by infectious bronchitis coronavirus (IBV) and considered to be one of the major causes of economic losses to the poultry industry worldwide. Although nearly all IBV strains primarily result in respiratory disease; some strains can also cause lesions in the enteric, urinary and reproductive tracts, which results in nephritis, reduced egg production and quality in layers, decreased feed conversion efficiency and significant mortality in commercial broilers (Cavanagh, 2005).

IBV is the prototype avian coronavirus, belonging to the genus *Gammacoronaviridae*. IBV is an enveloped virus, with single-stranded, positive sense, 5' capped and 3' polyadenylated RNA genome that is approximately 27 Kb (Bourisnell et al., 1987). The 3' end of the genome encodes four structural proteins, including spike (S), envelope (E), membrane (M) and nucleocapsid (N), and four accessory proteins. Genetic diversity in IBV is the result of recombination events and/or mutation, including substitutions, deletions and insertions that occur in the

genome. The S1 subunit of the spike protein is particularly variable especially during viral replication. The 5' end of the genome encodes the replication genes, which are translated into two large polyproteins, pp1a and pp1ab, which are processed into 15 non-structural proteins (nsp) via proteolytic cleavage (Thiel et al., 2003). For IBV, as for other coronaviruses, the recombination events are thought to result from a unique template switching copy choice mechanism during RNA replication, while the high mutation rates are attributed to the minimal proof reading capabilities of the viral RNA-dependent RNA-polymerase (Simon-Loriere and Holmes, 2011).

The S1 gene of IBV is highly variable among different viral strains, which results in the diversity of IBV serotypes/genotypes, this is because the S1 subunit of the spike glycoprotein is responsible for inducing neutralizing and serotype-specific antibodies in chickens (Cavanagh, 2007). Since IBV was first described in 1936, many IBV genotypes/serotypes and variants have been identified (Jackwood, 2012). It is believed that only a small proportion of these have become widespread and predominant in countries with significant poultry industries, it is believed that the majority of these strains have either disappeared or become endemic in certain geographical areas (Khataby et al., 2016). In the last few years, one of the most predominant IBV genotypes circulating in the chicken flocks worldwide is thought to be the LX4 strain (also known as QX-like) (Liu and Kong, 2004; de Wit et al., 2011; Jackwood, 2012). The LX4 genotype is thought to have originated in the mid-1990s in China (Liu and Kong, 2004). Subsequently, it has been shown to be the predominant genotype circulating in chicken flocks in China (Han et al., 2011).

Recently, the prevalence of this genotype has been reported in many European and Asian countries (de Wit et al., 2011; Jackwood, 2012; Promkuntod, 2016). LX4 genotype is becoming one of the most important genotypes of IBV resulting in major economic problems in IB-vaccinated flocks in many countries of the world (de Wit et al., 2011). It appears that this genotype is still able to spread rapidly among susceptible flocks in other countries of the world.

It has been more than 20 years since the LX4 genotype was first described in China. However, the dynamics of this genotype's circulation in commercial birds has not been extensively investigated. Therefore, the aim of this study was to investigate and genetically characterize the LX4 genotype in China between 1995 and 2015. We sequenced the complete genomes of 110 IBV strains isolated in China and compared the sequences with each other and with the other IBV sequences available in GenBank. We performed phylogenetic, molecular and recombination analyses, and reported our findings here.

## **2. Materials and methods**

### *2.1. Virus*

Of the 110 IBV strains, 50 were isolated previously (Liu and Kong, 2004; Liu et al., 2006; Liu et al., 2008; Liu et al., 2009; Han et al., 2011; Sun et al., 2011; Ma et al., 2012) and 60 were isolated in this study and purified as previously described (Chen et al., 2015). All the viruses were isolated from the chicken flocks suspected

to be infected by IB. Information about the regions, years and the organs from which the isolates were obtained are listed in Supplemental Table 1. Viruses were isolated by inoculating and blind passaging in the allantoic cavity of 9-day-old specific pathogen-free (SPF) embryonated chicken eggs (Harbin Veterinary Research Institute, China) until characteristic IBV lesions were observed (Liu and Kong, 2004). Each of the virus stocks was prepared by propagating in 9-day-old SPF chicken eggs, as described previously (Sun et al., 2011). After 48h incubation, the eggs were chilled for 12–18h at 4°C and the allantoic fluid collected and stored at -80°C until RNA extraction for genome sequencing.

Of the 110 IB viruses, the *SI* gene of the 50 viruses isolated between 1995 and 2010 were sequenced previously (Liu and Kong, 2004; Liu et al., 2006; Liu et al., 2008; Liu et al., 2009; Han et al., 2011; Sun et al., 2011; Ma et al., 2012). In addition, the sequences from *S2* to *N* genes of eight strains, including strains ck/CH/LHLJ/95I, ck/CH/LLN/98I, ck/CH/LHLJ/99I, ck/CH/LHLJ/02I, LX4, ck/CH/LJL/04I, ck/CH/LSD/03I and ck/CH/LXJ/02, were also sequenced previously (Liu et al., 2008b). In this study, the complete genome of all 50 of the previously isolated viruses, together with those of the 60 IBVs isolated in this study, were sequenced.

## 2.2. RNA extraction

An aliquot of each of the virus stocks was clarified by centrifuging at 2500×*g* for 10 min. Two-hundred microliters of the supernatant was then used for RNA extraction using the RNAiso Plus kit (TaKaRa, Shiga, Japan), following the

manufacturer's protocol and the RNA template was used immediately for RT-PCR or stored at -80 °C until its use.

### 2.3. RT-PCR amplification

Overlapping fragments of the genome of the 110 IBV strains were obtained through RT-PCR using primer sets based on the conserved regions in the genome among most of the IBV strains (Liu et al., 2013). A one-step method was adopted using PrimeScript™ One Step RT-PCR kit Ver.2 (TaKaRa) and the following 25  $\mu$ L mixture: 12.5  $\mu$ L of 2 $\times$  1 step Buffer, 7.5  $\mu$ L of PrimeScript I step Enzyme Mix, 15 nmol each of downstream and upstream primers and 3  $\mu$ L of template RNA. The reaction was conducted at 95 °C for 5 min, and 30 cycles of 94 °C for 1 min; 50 °C for 1 min; 72 °C for 2 min, and a final extension step of 72 °C for 10 min. All gaps and ambiguous sequences were corrected by additional RT-PCR assays and subsequent sequencing attempts using primers designed on the alignment of the sequenced viruses in this study.

The far 5' and 3' ends were amplified using 5' and 3' RACE for Rapid Amplification of cDNA Ends (Invitrogen, Grand Island, USA), respectively, following the manufacturer's instructions. The PCR products were detected by electrophoresis of a 1% agarose gel and visualization under UV light after ethidium bromide staining.



#### 2.4. Sequence comparison and analysis

RT-PCR products were subjected to direct sequencing and/or cloned into a pMD 18-T vector (Takara Bio Inc.) and three to five clones were sequenced. The genomic fragments of each virus were sequenced at least three times to determine a consensus sequence for any given genomic region. The nucleotide sequences from all the sequenced IBV strains were manually edited and analyzed using the ClustalW method (available in the Bioedit software package, <http://www.mbio.ncsu.edu/bioedit>) and the NCBI's (<http://www.ncbi.nlm.nih.gov>) tools. Nucleotide sequences of the different ORFs and comparative sequence analysis with five reference IBV sequences was carried out on the complete genomic sequence. The nucleotide sequences of the spike genes of our 110 viruses were converted into amino acid sequences and compared with those of the reference strains.

Multiple sequences alignments, including the spike genes, the sequences from the *N* gene to the 3' UTR, and the complete genomes, were performed with our 110 IBV strains using the Muscle algorithm, implemented in MEGA software, version 6.06 (<http://www.megasoftware.net/>). Five IBV reference strains including Baudette (NC\_001451), H120 (GU393335), M41 (DQ834384), 4/91 (KF377577) and ck/CH/LDL/97I (JX195177), available in GenBank database, were added to the alignments, respectively. Phylogenetic analyses were elaborated on the spike genes, the sequences from the *N* gene to 3' UTR, and the complete genomes using the

neighbor-joining method with 1000 bootstrap replicates (MEGA software version 5.0; available at <http://www.megasoftware.net/>).

To obtain more information, the SimPlot analysis was performed with the 110 complete genomic sequences using the SimPlot program (Lole et al., 1999). The whole sequence of H120 was used as a query. Finally, SimPlot was also used to detect the recombination events in the sequence from the *N* gene to the 3' UTR of strains ck/CH/LHLJ/130744 and ck/CH/LJL/140734, respectively. The IBV strains 4/91 and H120 were used as a query, respectively, and the reference strain, ck/CH/LHLJ/130822 was isolated in this study. To confirm the precise recombination breakpoints, pairwise comparison of the sequences from the *N* gene to the 3' UTR of ck/CH/LHLJ/130744 and ck/CH/LJL/140734 were performed using strain ck/CH/LHLJ/130822, and the 4/91 and H120 strains, respectively.

### *2.5. Nucleotide sequence accession number*

All 110 complete genomic sequences reported here have been deposited in the GenBank database, and the accession numbers are list in Supplemental Table 1.

## **3. Results**

### *3.1. Molecular characteristics of the spike gene*

The exploratory phylogenetic tree based on the *S* gene showed a distinct group formed by the LX4 strain, when compared with those of the Massachusetts, 793/B and ck/CH/LDL/97I, genotypes (Fig. 1). Within this group, two distinct clades could

be seen with strains clustered according to the years when these strains were isolated, although some of the strains showed variability. Clade I contained all virus strains isolated between 1995 and 2003, in contrast, Clade II contained strains isolated between 2004 and 1995. The Beaudette, H120 and M41 strains clustered with Massachusetts types, 793/B, 4/91 and ck/CH/LDL/97I.

In this study, 29 of the 110 viruses were selected and used for homology analysis of the S protein according to the results of the phylogenetic analysis (Fig. 1). The year and region of isolation were considered when making the selection. The percentage of similarities for the viruses in Clade I ranged from 94.0% to 97.4% at the amino acid level, whereas, viruses in Clade II ranged from 96.3% to 99.7% at the amino acid level (Supplemental Table 2). Generally, viruses in the same Clade showed higher similarity than those in different Clades, although some of the strains showed diversity, which was in line with the results obtained from the phylogenetic analysis. These results indicated that our 110 viruses belonged to a same genotype (designated LX4). In contrast, the LX4 genotypes shared less than 83.6% similarity with H120 strain, indicating the distinct genetic relationship.

The spike cleavage recognition site sequences of IBV correlates with geographic distribution of the viruses although it does not appear to correlate with serotype and pathogenicity (Jackwood et al., 2001). Our 110 LX4 genotypes showed five different cleavage site sequences. The most common cleavage recognition site was His-Arg-Arg-Arg-Arg which was shared by 105 viruses. A second cleavage recognition site, His-Arg-His-Arg-Arg, was observed for viruses ck/CH/LSHH/03II

and ck/CH/LLN/090312. In addition, three IBV strains, ck/CH/LDL/05III, ck/CH/LLN/06I and ck/CH/LXJ/111265, had the Arg-Arg-Tyr-Arg-Arg, His-Arg-Pro-Arg-Arg and Arg-Arg-Pro-Arg-Arg cleavage recognition site sequences, respectively.

### *3.2. Genomic organization of IBV LX4 genotype isolated in China*

The genomes of the 110 viruses isolated in this study have genomes of varying size, from 27,596 bp to 27790 bp, excluding the 3' poly (A) tail (Supplemental Table 3). Of the 110 viruses, 37 strains were 27673 bp in size, 14 had a genome of 27663 bp, 14 were 27670 bp in size, 4 had a genome of 27676 bp, 4 were 27664 bp, 3 27660 bp, 3 27669 bp, 3 27671 bp, 3 27666 bp, 2 27672 bp and 2 27655 bp in size. In addition, the genomic sizes of the 21 viruses selected were 27790 bp, 27742 bp, 27654 bp, 27620 bp, 27702 bp, 27665 bp, 27675 bp, 27642 bp, 27635 bp, 27644 bp, 27662 bp, 27633 bp, 27630 bp, 27685 bp, 27628 bp, 27682 bp, 27596 bp, 27674 bp, 27697 bp, 27667 bp and 27689 bp, respectively. The varying genome size is the result of deletions and/or insertions scattered in different regions of the genome, especially in the 3' UTR (Supplemental Table 3).

The overall size and position of the genomes and individual genes of our 110 LX4 genotype IBVs are summarized in Supplemental Table 3. One-hundred and eight out of 110 viruses showed the typical IBV genome organization, 5'-UTR-Gene 1 (ORF) 1a, 1b)-S-Gene 3 (3a, 3b and 3c)-M-Gene 5 (5a and 5b)-N-UTR-3' (Supplemental Fig. 1), this contrasted with the genomic organization of two novel

IBV strains: ck/CH/LLN/98I and ck/CH/LJL/08-1. Approximately 7 kb of the 3' region of ck/CH/LLN/98I was sequenced previously (Liu et al., 2008b) and the complete genome was sequenced in this study, confirming the occurrence of a nucleotide change at the corresponding position of the CK/CH/LLN/98I start codon in the 3a gene leading to the absence of ORF 3a in this virus. This nucleotide change resulted in a novel genomic organization 5'-UTR-Gene 1 (ORF1a, 1b)-S-Gene 3 (3b and 3c)-M-Gene 5 (5a and 5b)-N-UTR-3' (Supplemental Fig. 1). In addition, we found in this study that a nucleotide change at the start codon in the 5b gene of ck/CH/LJL/08-1 led to the absence of ORF 5b (Supplemental Fig. 2), resulting in the genomic organization 5'-UTR-Gene 1 (ORF1a, 1b)-S-Gene 3 (3a, 3b and 3c)-M-Gene 5 (5a)-N-UTR-3' (Supplemental Fig. 2).

### *3.3. Analysis of genetic diversity in the genomes of LX4 genotype IBVs*

In general, phylogenetic analysis of the complete genome of our 110 LX4 genotype strains, together with those of the Massachusetts, 793/B and ck/CH/LDL/97I genotypes, divided IBV strains into four distinct groups in which our 110 strains showed a distinct group formed by the LX4 strain, similar to that of the phylogenetic tree constructed using the S protein (Fig. 2). In this group, they also formed two clusters, with virus strains isolated between 1995 and 2005 and a strain isolated in 2006 clustered with the LX4 type strain in Clade I, and another virus isolated in 2006 and all viruses isolated between 2007 and 2015 in Clade II. The

Massachusetts type Beaudette, H120 and M41 strains, the 793/B type 4/91 strain and ck/CH/LDL/97I and ck/CH/LDL/97I strains clustered separately.

A similarity plot (Simplot) analysis comparing the 110 full genomes showed that most of the genomes of these viruses are highly similar throughout the genome, except for nine strains isolated between 1995 and 2005 in the Clade I (Fig. 3). Strains LH1 and LD3 have the same topology and show obvious diversity between the 3' end of nsp 2 and the 5' end of nsp3. Strains ck/CH/LDL/05II and ck/CH/LDL/05III show the same topology and are obviously different from other strains between the 3' end of nsp 6 and the 5' end of nsp 8. Strains LX4 and ck/CH/LHLJ/95I, show similar topology and are clearly different for most of the nsp 3 region. Strain ck/CH/LLN/98I showed variability from the 3' end of the *M* gene to the 3' UTR. Strain ck/CH/LSD/03I showed the most diversity at different regions, including from the 5' UTR to the 3' end of nsp 2, the 3' end of nsp 2 and the 5' end of nsp 8 and the 3' end of ORF3 to the 3' UTR.

#### 3.4. Recombination analysis

As a result of the SimPlot analysis using the complete genome we were able to demonstrate that the sequences at the 3' end of the genomes showed different topologies. So we analyzed the genetic diversities by constructing the phylogeny first using the sequences from the *N* gene to the 3' UTR. In the neighbor-joining tree, all the viruses clustered separately into seven groups (Groups I to VII) (Fig. 4). Of the 110 viruses, 59 were clustered together in Group I which included the viruses

isolated between 2007 and 2015. Thirty-four viruses formed Group II which mainly included the viruses isolated between 2005 and 2015. Viruses in Groups III to VII showed higher diversities and some viruses showed close genetic relationship with 4/91 or H120 strains, indicating possible recombination events during the origin of these viruses.

In order to evaluate the possible recombination events at the 3' end of the genome, two viruses, ck/CH/LJL/140734 and ck/CH/LHLJ/130744 which clustered in Groups IV and VI, were selected for further investigation. Strains 4/91 and ck/CH/LHLJ/130822 were selected as potential parental viruses of ck/CH/LHLJ/130744 for SimPlot analysis because they were closely related ck/CH/LHLJ/130744 between the 5' UTR and the *M* gene (Fig. 3) and the *N* gene and the 3' UTR (Fig. 4), respectively. SimPlot analysis confirmed the aforementioned results, and it clearly showed that ck/CH/LHLJ/130744 arose from a homologous RNA recombination event from a template switch (Fig. 5A). A crossover point (nt 26144–26161) was found located at the 5' end of the *N* gene (Fig. 5B) in strain ck/CH/LHLJ/130744. Similarly, strains H120 and ck/CH/LHLJ/130822 were selected as potential parental viruses of ck/CH/LJL/140734 for SimPlot analysis. It was also shown that ck/CH/LJL/140734 arose from a homologous RNA recombination event (Fig. 5C) and a crossover point (nt 26682–26721) was found to be located at the 3' end of the *N* gene of strain ck/CH/LHLJ/130744 (Fig. 5D). These results suggested that recombination events might account for the genetic diversities at the 3' end of the genomes of the viruses in Groups IV and VI.

#### 4. Discussion

The complete genomes of 110 LX4 genotype IBV strains were analyzed and compared to each other and to that of three reference types including Beaudette, H120 and M41 as Massachusetts type, 4/91 as a 793/B type and ck/CH/LDL/97I as ck/CH/LDL/97I type. The latter is another very important genotype that first emerged in China and spread to other regions of the world (Ababneh et al., 2012; Marandino et al., 2015). This comparison revealed that most of the viruses (108 out of 110) had a previously known genomic organization (5'-UTR-Gene 1 (ORF1a, 1b)-S-Gene 3 (3a, 3b and 3c)-M-Gene 5 (5a and 5b)-N-UTR-3'), although various lengths of genomes have been found which may be the result of insertions/deletions. Interestingly, different genomic organizations were found in two of the IBV strains, ck/CH/LLN/98I and ck/CH/LJL/08-1, respectively. The 3a gene of ck/CH/LLN/98I was previously found to be absent because of a nucleotide change at the start codon in the 3a gene (Liu et al., 2008b). We sequenced the complete genome in this study and confirmed the previous result. An identical genomic organization lacking the 3a gene was also observed in another IBV strain, ck/CH/LHLJ/111043 (Xu et al., 2016). In this study, we found another novel IBV strain, ck/CH/LJL/08-1, lacked the 5b gene which was also the result of a nucleotide change at the start codon. Novel IBV strains which lacked either all or most of the genes coding for accessory proteins at the 3' end of the genome have been also isolated in Australia (Mardani et al., 2008).



These novel IBVs grew at a slower rate and reached lower titers *in vitro* and *in vivo* and were markedly less immunogenic in chicks. Further, although the novel IBVs induced histopathological lesions in the tracheas of infected chicks that were comparable to those induced by classical strains, they did not induce lesions in the kidneys (Mardani et al., 2008). However, the nucleotide sequences of all the essential genes of the novel IBV strains that lacked the accessory genes differed significantly from those of their counterpart IBV strains with a full accessory gene complement, suggesting that these sequences were not derived by mutation or recombination with the more commonly isolated classical strain. In contrast, lack of *5b* in ck/CH/LJL/08-1 was the result of a nucleotide change, and therefore had a very similar backbone and high nucleotide sequence similarity to most of the classical IBV strains. It has been demonstrated that Gene 5 (including *5a* and *5b*) of IBV is not essential for replication using the reverse genetics system in the non-pathogenic strain Beaudette (Casais et al., 2005). Recently, it has been suggested that IBV accessory protein *5b* is a functional equivalent of *nsp1* of Alpha- and Beta coronavirus and is indispensable for limiting interferon production by inducing a host shutoff which plays an important role in antagonizing the host's innate immune response, thereby indicating that *5b* may be an important virulence factor of IBV (Kint et al., 2016). It is not possible to determine if the *5a* gene was ever present in strain ck/CH/LJL/08-1 or whether it was deleted as a result of a mutation during their circulation in commercial poultry. The biological characterization, including growth properties, virulence and immunogenicity, of strain ck/CH/LJL/08-1 requires

further investigation.

In the mid-1990s, a disease characterized by “proventriculitis” emerged and spread in chicken flocks in China, especially in commercial broilers. In most cases, IBV could be isolated from the proventriculus of the diseased chickens; hence, many researchers in China suggested that a novel pathotype of IBV was the etiological cause of the disease although infection of both SPF and commercial chickens with the IBVs isolated from the proventriculus had not resulted in similar disease progression as observed in field conditions. QXIBV was isolated in 1997 and was among the strains considered to be an etiology causing proventriculitis (Yudong et al., 1998). However, it was not recognized as a novel genotype of IBV at that time. During the dynamic surveillance of IBVs in China, we found that QX-like IBVs were prevalent in many regions in China and they represented a new genotype (LX4 strain as representative) which were nephropathogenic (Liu and Kong, 2004; Liu et al., 2006; Liu et al., 2008a; Liu et al., 2009). Subsequently, this genotype of IBV has spread in chicken flocks in many parts of the world (de Wit et al., 2011; Jackwood, 2012; Promkuntod, 2016). In this retrospective study, we looked at ck/CH/LHLJ/95I isolated in 1995 in China and, to our knowledge, this is the first isolate of the LX4 genotype. The strains of the genotype appeared in different locations in China in the 10 following years, e.g. ck/CH/LHLJ/95I in Heilongjiang in 1995, QXIBV in Shandong in 1996 (Yudong et al., 1998), ck/CH/LLN/98I in Liaoning in 1998, LX4 in Xinjiang in 1999 and ck/CH/LSHH/03II in Shanghai in 2003. It could be argued that either there were more independent introductions of the genotype from an

as-yet-unidentified source into commercial poultry at regions geographically distinct from China or the introduction of a single variant occurred at a particular time. It is difficult to answer this question by analysis of only the *S* gene although generally, the *S* gene of LX4 genotypes isolated before 2006 are genetically different from those of strains isolated after 2006. Phylogenetic and SimPlot analysis of the complete genomic sequences of the 110 LX4 genotypes suggested that IBV strains isolated before 2006 were genetically different not only at the *S* gene but at other part(s) of genome as well. These differences occurred between isolates from similar time periods and from those isolated later, possibly indicating that there were independent introductions of the LX4 genotype into commercial poultry in various regions of China and that these viruses evolved in parallel. The topology of some regions in the genomes of most of the IBV strains isolated before 2006 were clearly different from each other and from those IBVs isolated after 2006 (Fig. 3). Most of these regions have no obvious sequence similarity to their counterparts in other IBV or other coronaviruses (data not shown), indicating that the origin of these strains was not mutation, but might be from recombination events between strains which have not yet been determined, similar to that of Turkey coronavirus (Jackwood et al., 2010).

Interestingly, nearly all the strains of the LX4 genotype isolated after 2006 had similar topology as assessed by phylogenetic and SimPlot analysis which was different from those of most viruses isolated before 2006. Only few strains isolated after 2006 showed diversity at the 3' end of the genome. Those variations were the

result of a recombination event with the vaccine strains which were commonly used in China. Comparatively, two strains (ck/CH/LXJ/02I and ck/CH/LJL/05I) isolated before 2006 showed similar topology by SimPlot analysis, suggesting that strains isolated after 2006 might be derivatives of these strains after evolution. The disappearance after 2006 of most of the LX4 genotype with various topologies in commercial poultry is unexpected. In most cases, IBV variants persist only transiently, particularly if they are targeted by vaccination (Mardani et al., 2008; de Wit et al., 2011). This may be the case for some of our strains and not for others, such as LD3 and LH1, which were isolated in Heilongjiang and Shandong provinces in 2001 and 2003, respectively, and ck/CH/LHLJ/95I and LX4, which were isolated in Heilongjiang and Xinjiang provinces in 1995 and 1999, respectively. No novel vaccines were introduced into the commercial poultry industry in China in the mid-2000s. Hence, as-yet-unknown factor(s) contributing to the disappearance of some of the virus strains identified here could have been a result of their reduced capacity for replication in chickens. For example, mutations in these strains resulted in the attenuation of these viruses which meant they could not compete with “fitter” viruses (Mardani et al., 2008). In contrast, the LX4 genotype viruses isolated after 2006 may be fitter for replication in chickens and consequently adapted in poultry.

Taken together, a hypothesis can be formed to explain the scenario of the origin and evolution of Chinese LX4 genotype viruses, in which independent recombination events of as-yet-unidentified parental viruses may firstly be responsible for the emergence of IBV LX4 strains with different topologies. Most of

these viruses disappeared because they were not “fit” to adapt in chickens and lost the competition to others. Finally, those of the “fit” viruses were still evolving and have become widespread and predominant in commercial poultry in China after 2006. In addition, few of these viruses experienced recombination with those of the vaccine strains at the 3' end of the genome.

### **Competing interests**

The authors declare that they have no competing interests.

### **Acknowledgement**

This work was supported by grants from the China Agriculture Research System (No. CARS-41-K12), National "Twelfth Five-Year" Plan for Science and Technology Support (2015BAD12B03), and Special Fund for Agro-scientific Research in the Public Interest (No. 201303033).

### **References**

Ababneh, M., Dalab, A.E., Alsaad, S., Al-Zghoul, M., 2012. Presence of Infectious Bronchitis Virus Strain CK/CH/LDL/97I in the Middle East. *ISRN Veterinary Science* 11.

- Boursnell, M.E., Brown, T.D., Foulds, I.J., Green, P.F., Tomley, F.M. & Binns, M.M., 1987. Completion of the sequence of the genome of the coronavirus avian infectious bronchitis virus. *Journal of General Virology* 68, 57-77.
- Casais, R., Davies, M., Cavanagh, D., Britton, P., 2005. Gene 5 of the avian coronavirus infectious bronchitis virus is not essential for replication. *Journal of Virology* 79, 8065-7808.
- Cavanagh, D., 2005. Coronaviruses in poultry and other birds. *Avian Pathology* 34, 439-448.
- Cavanagh, D., 2007. Coronavirus avian infectious bronchitis virus. *Veterinary Research* 38, 281-297.
- Chen, L., Zhang, T., Han, Z., Liang, S., Xu, Y., Xu, Q., Chen, Y., Zhao, Y., Shao, Y., Li, H., Wang, K., Kong, X., Liu, S., 2015. Molecular and antigenic characteristics of Massachusetts genotype infectious bronchitis coronavirus in China. *Veterinary Microbiology* 181, 241-251.
- de Wit, J.J., Cook, J.K., van der Heijden, H.M., 2011. Infectious bronchitis virus variants: a review of the history, current situation and control measures. *Avian Pathology* 40, 223-235.
- Han, Z., Sun, C., Yan, B., Zhang, X., Wang, Y., Li, C., Zhang, Q., Ma, Y., Shao, Y., Liu, Q., Kong, X., Liu, S., 2011. A 15-year analysis of molecular epidemiology of avian infectious bronchitis coronavirus in China. *Infection, Genetics and Evolution* 11, 190-200.
- Jackwood, M.W., 2012. Review of infectious bronchitis virus around the world.

- Avian Diseases 56, 634-641.
- Jackwood, M.W., Boynton, T.O., Hilt, D.A., McKinley, E.T., Kissinger, J.C., Paterson, A.H., Robertson, J., Lemke, C., McCall, A.W., Williams, S.M., Jackwood, J.W., Byrd, L.A., 2010. Emergence of a group 3 coronavirus through recombination. *Virology* 398, 98-108.
- Jackwood, M.W., Hilt, D.A., Callison, S.A., Lee, C.W., Plaza, H., Wade, E., 2001. Spike glycoprotein cleavage recognition site analysis of infectious bronchitis virus. *Avian Diseases* 45, 366-372.
- Khataby, K., Fellahi, S., Loutfi, C., Mustapha, E.M., 2016. Avian infectious bronchitis virus in Africa: a review. *The Veterinary quarterly* 36, 71-75.
- Kint, J., Langereis, M.A., Maier, H.J., Britton, P., van Kuppeveld, F.J., Koumans, J., Wiegertjes, G.F., Forlenza, M., 2016. Infectious bronchitis coronavirus limits interferon production by inducing a host shutoff that requires accessory protein 5b. *Journal of Virology* pii: JVI.00627-16.
- Liu, S., Kong, X., 2004. A new genotype of nephropathogenic infectious bronchitis virus circulating in vaccinated and non-vaccinated flocks in China. *Avian Pathology* 33, 321-327.
- Liu, X., Ma, H., Xu, Q., Sun, N., Han, Z., Sun, C., Guo, H., Shao, Y., Kong, X., Liu, S., 2013. Characterization of a recombinant coronavirus infectious bronchitis virus with distinct S1 subunits of spike and nucleocapsid genes and a 3' untranslated region. *Veterinary Microbiology* 162, 429-436.
- Liu, S., Wang, Y., Ma, Y., Han, Z., Zhang, Q., Shao, Y., Chen, J., Kong, X., 2008a.

- Identification of a newly isolated avian infectious bronchitis coronavirus variant in China exhibiting affinity for the respiratory tract. *Avian Diseases* 52, 306-314.
- Liu, S., Zhang, Q., Chen, J., Han, Z., Liu, X., Feng, L., Shao, Y., Rong, J., Kong, X., Tong, G., 2006. Genetic diversity of avian infectious bronchitis coronavirus strains isolated in China between 1995 and 2004. *Archives of Virology* 151, 1133-1148.
- Liu, S., Zhang, Q., Chen, J., Han, Z., Shao, Y., Kong, X., Tong, G., 2008b. Identification of the avian infectious bronchitis coronaviruses with mutations in gene 3. *Gene* 412, 12-25.
- Liu, S., Zhang, X., Wang, Y., Li, C., Han, Z., Shao, Y., Li, H., Kong, X., 2009. Molecular characterization and pathogenicity of infectious bronchitis coronaviruses: complicated evolution and epidemiology in china caused by cocirculation of multiple types of infectious bronchitis coronaviruses. *Intervirology* 52, 223-234.
- Lole, K.S., Bollinger, R.C., Paranjape, R.S., Gadkari, D., Kulkarni, S.S., Novak, N.G., Ingersoll, R., Sheppard, H.W., Ray, S.C., 1999. Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. *Journal of virology* 73, 152-160.
- Ma, H., Shao, Y., Sun, C., Han, Z., Liu, X., Guo, H., Liu, X., Kong, X., Liu, S., 2012. Genetic diversity of avian infectious bronchitis coronavirus in recent years in



- China. *Avian Diseases* 56, 15-28.
- Marandino, A., Pereda, A., Tomás, G., Hernández, M., Iraola, G., Craig, M.I.,  
Hernández, D., Banda, A., Villegas, P., Panzera, Y., Pérez, R., 2015.  
Phylogenetic analysis of avian infectious bronchitis virus in South America.  
*Journal of General Virology* 96, 1340-1346.
- Mardani, K., Noormohammadi, A.H., Hooper, P., Ignjatovic, J., Browning, G.F.,  
2008. Infectious bronchitis viruses with a novel genomic organization. *Journal  
of Virology* 82, 2013-2024.
- Promkuntod, N., 2016. Dynamics of avian coronavirus circulation in commercial and  
non-commercial birds in Asia--a review. *The Veterinary quarterly* 36, 30-44.
- Simon-Loriere, E., Holmes, E.C., 2011. Why do RNA viruses recombine? *Nature  
Reviews Microbiology* 9, 617-626.
- Sun, C., Han, Z., Ma, H., Zhang, Q., Yan, B., Shao, Y., Xu, J., Kong, X., Liu, S.,  
2011. Phylogenetic analysis of infectious bronchitis coronaviruses newly  
isolated in China, and pathogenicity and evaluation of protection induced by  
Massachusetts serotype H120 vaccine against QX-like strains. *Avian  
Pathology* 40, 43-54.
- Thiel, V., Ivanov, K.A., Putics, A., Hertzog, T., Schelle, B., Bayer, S., Weissbrich, B.,  
Snijder, E.J., Rabenau, H., Doerr, H.W., Gorbalenya, A.E., Ziebuhr, J., 2003.  
Mechanisms and enzymes involved in SARS coronavirus genome expression.  
*Journal of General Virology* 84, 2305-2315.
- Xu, Q., Han, Z., Wang, Q., Zhang, T., Gao, M., Zhao, Y., Shao, Y., Li, H., Kong, X.,

Liu, S., 2016. Emergence of novel nephropathogenic infectious bronchitis viruses currently circulating in Chinese chicken flocks. *Avian Pathology* 45, 54-65.

YuDong, W., YongLin, W., ZiChun, Z., GenChe, F., YiHai, J., XiangE, L., 1998. Isolation and identification of glandular stomach type IBV (QX IBV) in chickens. *Chinese Journal of Animal Quarantine* 15, 1-3.

**Figure legends****Fig. 1.**

Phylogenetic tree inferred from the S amino acid sequences of 110 LX4 genotype IBV strains and five reference strains. The trees were constructed using the neighbor-joining method and bootstrap values calculated from 1000 trees. The IBV strains used for comparison of amino acid similarities are in bold.

**Fig. 2.**

Phylogenetic tree inferred from the complete genomic sequences of 110 LX4 genotype IBV strains and five reference strains. The trees were constructed using the neighbor-joining method and bootstrap values calculated from 1000 trees.

**Fig. 3.**

SimPlot analysis of our 110 LX4 IBV strains. The complete sequence of the H120 strain was used as the query sequence. 101 out of 110 LX4 strains showed similar topology and are indicated in grey. Those with different topologies are indicated in different colors with the name of the virus in the box. The position from the *N* gene to the 3' UTR is also indicated.

**Fig. 4.**

Phylogenetic tree inferred from the nucleotide sequence from the *N* gene to the 3' UTR of 110 LX4 IBV strains and five reference strains. The trees were constructed using

the neighbor-joining method and bootstrap values calculated from 1000 trees.

**Fig. 5.**

Recombination analysis of the IBV strains ck/CH/LHLJ/130744 (A and B) and ck/CH/LJL/140734 (C and D). SimPlot using H120 as the query sequence. The dotted lines show the deduced recombination breakpoints (A and C). The hollow arrows show the different fragments similar to those of the parental viruses. Multiple sequence alignments of the predicted breakpoints and flanking sequences among 4/91, ck/CH/LHLJ/130822 and ck/CH/LHLJ/130744 (B), and H120, ck/CH/LHLJ/130822 and ck/CH/LJL/140734 (D). The numbers on the right of each alignment show the nucleotide positions in the genome of each virus. The sequences of ck/CH/LHLJ/130744 (B) and ck/CH/LJL/140734 (D) are listed, respectively, and only the nucleotides differing from those of these two strains are depicted. The regions where the template switches (breakpoint) have taken place are in bold.

Fig. 1.

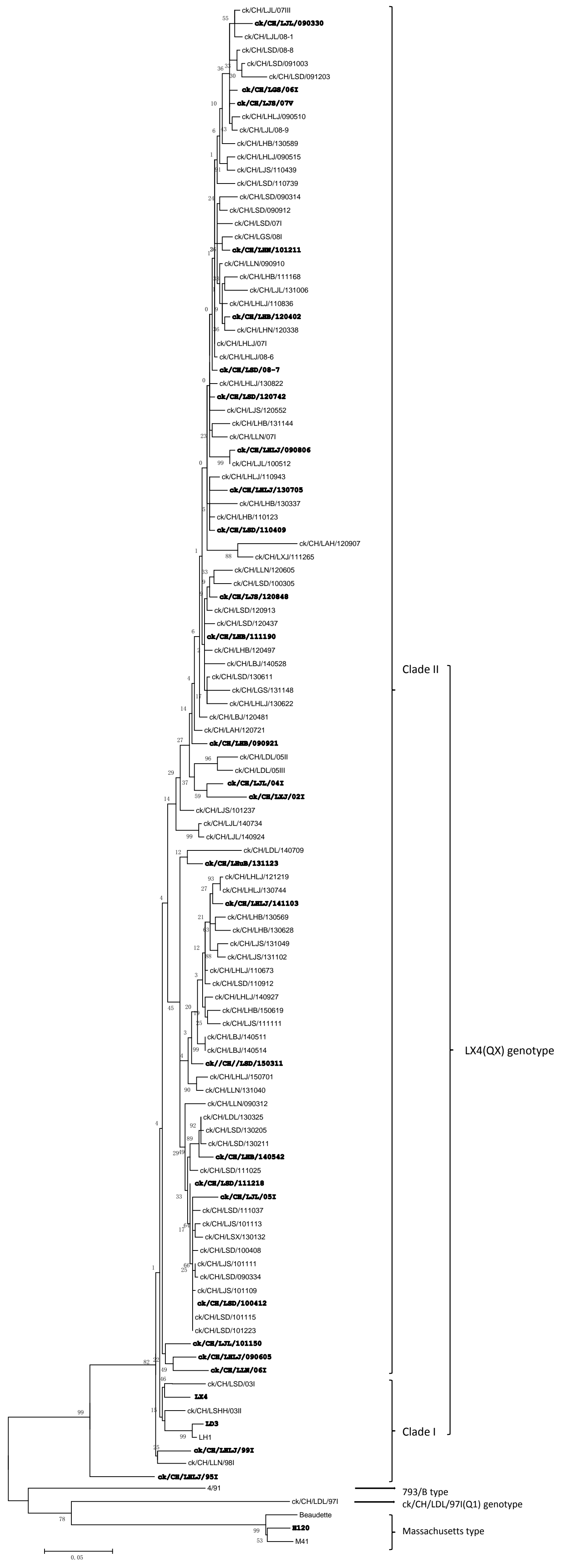


Fig. 2.

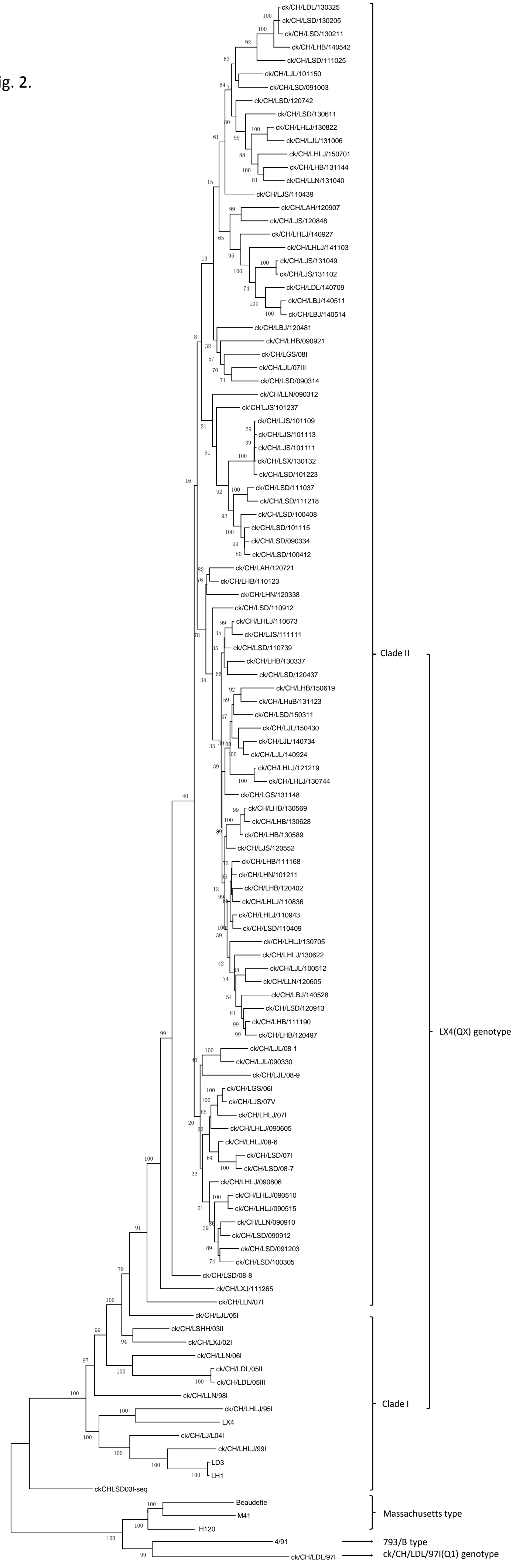


Fig. 3.

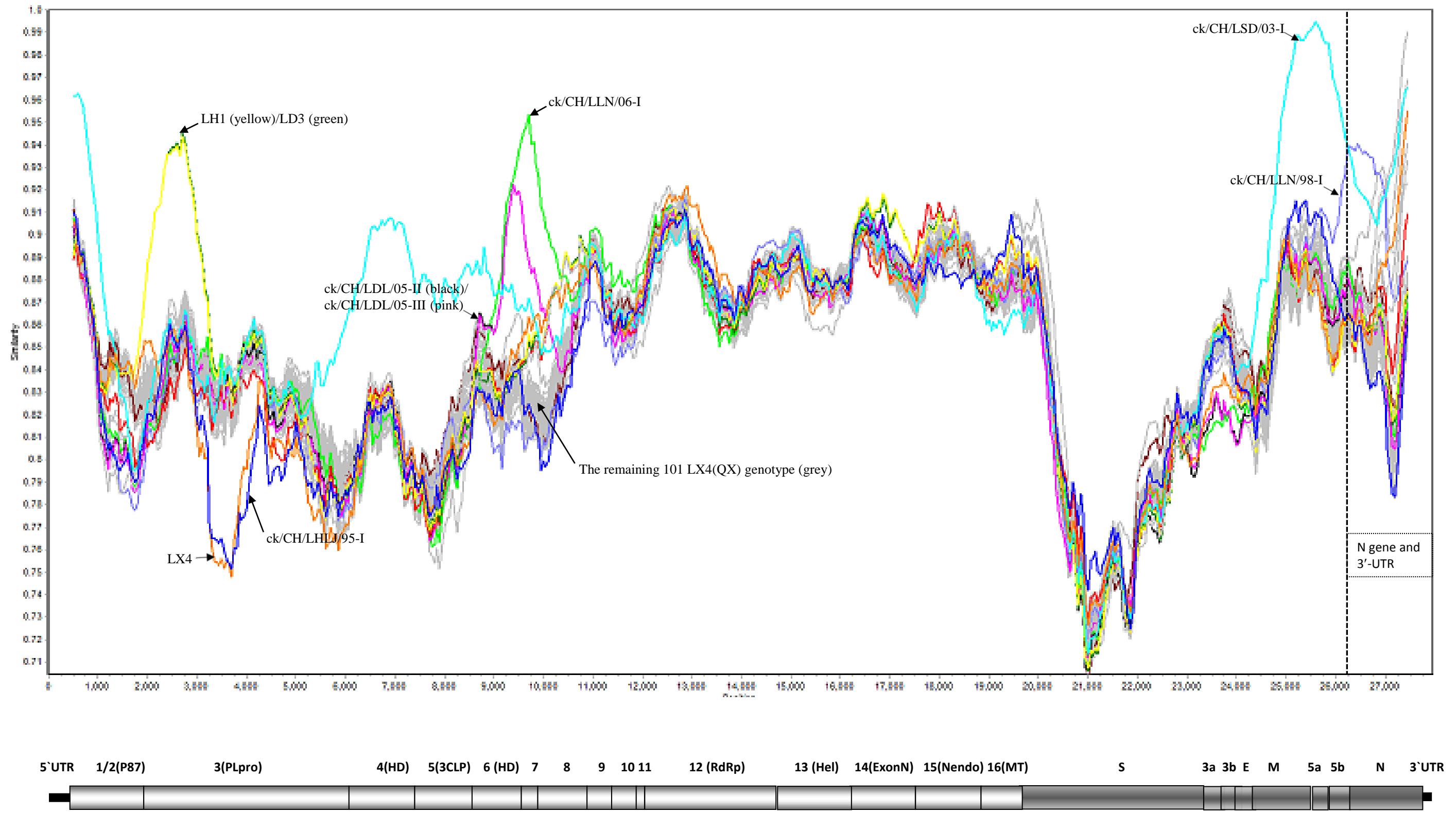


Fig. 4.

