Infection, Genetics and Evolution 14 (2013) 29–38

Contents lists available at SciVerse ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

# Comparative analysis of four Massachusetts type infectious bronchitis coronavirus genomes reveals a novel Massachusetts type strain and evidence of natural recombination in the genome

Xiaoli Liu, Yuhao Shao, Huijie Ma, Chuyang Sun, Xiaonan Zhang, Chengren Li, Zongxi Han, Baolong Yan, Xiangang Kong, Shengwang Liu *

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

ARTICLE INFO

*Article history:*
Received 18 July 2012
Received in revised form 23 September 2012
Accepted 25 September 2012
Available online 22 November 2012

*Keywords:*
Avian infectious bronchitis coronavirus
Mass-type
Mutation
Recombination event
Genome

ABSTRACT

Four Massachusetts-type (Mass-type) strains of infectious bronchitis coronavirus (IBV) were compared genetically with the pathogenic M41 and H120 vaccine strains using the complete genomic sequences. The results revealed that strains ck/CH/LNM/091017 and ck/CH/LDL/101212 were closely related to the H120 vaccine, which suggests that they might represent re-isolations of vaccine strains or variants of vaccine strains that have resulted from the accumulated point mutations after several passages in chickens. In contrast, strains ck/CH/LHLJ/07VII and ck/CH/LHLJ/100902 had a close genetic relationship with the pathogenic M41 strain. In addition, molecular markers have been identified that distinguish between field and vaccine (or vaccine-like) Mass-type viruses, which may be able to differentiate between field and vaccine strains for diagnostic purposes. Phylogenetic analysis, and pairwise comparison of full-length genomes and the nine genes, identified the occurrence of recombination events in the genome of strain CK/VH/LHLJ/07VII, which suggests that this virus originated from recombination events between M41- and H120-like strains at the switch site located at the 3′ end of the nucleocapsid (N) genes. To our knowledge, this is the first time that evidence for the evolution and natural recombination under field conditions between Mass-type pathogenic and vaccinal IBV strains has been documented. These findings provide insights into the emergence and evolution of the Mass-type IB coronaviruses and may help to explain the emergence of Mass-type IBV in chicken flocks all over the world.



## 1. Introduction

In 1931, Schalk and Hawn described "an apparently new respiratory disease of chicks" in North Dakota in the United States, which was considered to be infectious bronchitis (IB) by later researchers of avian respiratory diseases (Schalk and Hawn, 1931). Currently, IB still occurs in nearly all poultry-producing countries; it is a highly contagious, acute, and economically important viral disease of chickens. The etiology of IB, which was first demonstrated by Beach and Schalm (1936), is infectious bronchitis virus (IBV).

IBV is grouped in the genus *Gammacoronavirus* of the family *Coronaviridae* in the order *Nidovirales* (de Groot et al., 2011). The coronavirus genomes are the largest among the known RNA viruses and are polycistronic, generating a nested set of subgenomic RNAs with common 5′ and 3′ sequences (Masters, 2006). Like those of all other coronaviruses, the 5′ two-thirds of the IBV genome consists of two large replicase open reading frames (ORFs), ORF1a and ORF1b. The ORF1a polyprotein (pp1a) can be extended with ORF1b-encoded sequences via a −1 ribosomal frameshift at a conserved slippery site (Brierley et al., 1987), which generates the polyprotein pp1ab, comprising more than 7000 amino acids, which includes the putative RNA-dependent RNA polymerase (RdRp) and RNA helicase (HEL) activity (Ziebuhr et al., 2001). The pp1a and pp1ab of IBV are processed autocatalytically by two different viral proteases, encoded by a papain-like protease (PLP) and a 3C-loke protease ($3CL^{pro}$) (Lee et al., 1991; Ziebuhr et al., 2000, 2001). Other putative domains, presumably associated with a 3′-to-5′ exonuclease (ExoN) activity, a poly(U)-specific endo-RNase (XendoU) activity, and a 2′-*O*-methyltransferase (2′-*O*-MT) activity, have been predicted in pp1ab (Ivanov et al., 2004; Snijder et al., 2003). The 3′ end of a coronavirus genome includes the viral structural and accessory protein genes: a spike (S) glycoprotein gene; an envelope (E) protein gene; a membrane (M) glycoprotein gene; a nucleocapsid (N) phosphoprotein gene; and several ORFs that encode putative non-structural accessory proteins (Masters, 2006). Of the virus-encoded proteins, the S1 subunit of the S protein carries

* Corresponding author. Tel.: +86 451 85935000; fax: +86 451 82733132.
*E-mail address:* swliu@hvri.ac.cn (S. Liu).

1567-1348/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.meegid.2012.09.016

virus-neutralizing activity, determines the serotype of IBV and is responsible for viral attachment to cells. It is also a major determinant of cell tropism in culture (Casais et al., 2003). The accumulation of point mutations, deletions, insertions and recombination events that have been observed in multiple structural genes, especially the S1 gene, of IBV recovered from naturally occurring infections have been considered to contribute to the genetic diversity and evolution of IBV, and consequently, to a number of IBV serotypes (Cavanagh, 2007).

The occurrence and emergence of multiple serotypes of the virus have complicated control by vaccination because many serotypes and variants do not confer complete cross-protection against each other (Cavanagh and Gelb, 2008). The originally discovered Massachusetts (Mass) type of IBV was identified in the United States, beginning in the 1950s (Fabricant, 2000; Johnson and Marquardt, 1975; Mondal et al., 2001). Mass-type strains have been isolated in Europe and Asia since the 1950s and up to the present day (Cavanagh and Gelb, 2008), together with dozens of other serotypes that have been isolated in Africa, Asia, India, Australia, Europe, and South America (Cavanagh, 2001, 2003, 2005). The first Mass-type "H" vaccines were developed in about 1960. They include H120 and H52 (Bijlenga et al., 2004), and are used very commonly and widely around the world. However, virus of this type is occasionally isolated from Massachusetts-vaccinated and -unvaccinated flocks with respiratory clinical signs. Some of the viruses have shown close genetic relationships with pathogenic Mass-type, rather than vaccine, strains by S1 gene analysis. However, conclusions based on the genetic analysis of a single gene sequence, and sometimes even a partial gene sequence, require caution because the true phylogeny can only be demonstrated by analyzing complete genomic sequences. Herein, we sequenced the complete genome of four IBV Mass-type strains that showed S1 gene diversity (Liu et al., 2009; Ma et al., 2012; Sun et al., 2011), and we present evidence for in-field recombination between pathogenic and vaccinal strains. Furthermore, we characterized the molecular variability of the four Mass-type strains to gain insight into the emergence and evolution of these viruses.

## 2. Materials and methods

### 2.1. Virus strains

Four Mass-type IBV strains were used for complete genomic sequence comparison and analysis in this study. Strain ck/CH/LHLJ/07VII was isolated in 2007 from the kidney of a layer hen vaccinated with H120 and 4/91 in Heilongjiang province, China (Liu et al., 2009). Strain ck/CH/LNM/091017 was isolated in 2009 from the swollen proventricular tissues of a broiler vaccinated with H120 in Neimenggu province, China (Sun et al., 2011). Strains ck/CH/LDL/101212 and ck/CH/LHLJ/100902, both of which were isolated in 2010, were obtained from laying hens in Dalian and Heilongjiang provinces, respectively, in China; the birds were suffering from nephropathogenic lesions and respiratory signs, respectively. In addition, the diseased birds in both flocks were suffering from proventriculitis (Ma et al., 2012). All of the IBV strains have been associated with various IB outbreaks in recent years in China and were assigned to the Mass-type strains by S1 sequence analysis. To avoid the possible mutation in the viral genome after serial passages in specific-pathogen-free (SPF) embryonated chicken eggs, the first passage of each original virus stock was used and purified once by propagating in 9- to 11-day-old SPF chicken eggs with a dose of $10^4$-fold dilutions per egg, and the presence of viral particles in the allantoic fluids of inoculated eggs was confirmed with a negative contrast electronic microscope (JEM-1200, EX) and reverse transcriptase-polymerase chain reaction (RT-PCR) as described previously (Han et al., 2011). In addition, since these viruses were isolated from chickens vaccinated with H120, it is possible that mixed IBV infections are present in one chicken flock. To exclude this, nine clones of S1 gene of each virus obtained from three independent PCR reactions were sequenced and compared. Sequences of each virus identical to the previously results were obtained (Liu et al., 2009; Ma et al., 2012; Sun et al., 2011).

### 2.2. Eggs

Fertile White Leghorn SPF chicken eggs were obtained from the Laboratory Animal Center, Harbin Veterinary Research Institute, the Chinese Academy of Agricultural Sciences, China.

### 2.3. RNA extraction and RT-PCR

To determine the full-length genomic sequences of the four viruses, 15 pairs of overlapping primers encompassing the entire genome were used. The primers were designed in regions that are conserved among most of the IBV strains available in the GenBank database. The sequences and locations of the primers used in RT and PCR in this study are presented in Table 1.

Viral RNA was extracted from 200 μl of infectious allantoic fluid using TRIzol reagents (Invitrogen, Grand Island, USA), following the manufacturer's protocol. Complementary DNA (cDNA) was synthesized using 80 μl of the first strand mixture (Invitrogen) containing 20 μM of primers N (−), 0.5 mM each of dNTP (TaKaRa, Dalian, China) and 40 μl of total RNA. The mixture was incubated at 70 °C for 5 min and then quick-chilled on ice for 2 min. The RT master mix was composed of 16 μl 5× RT buffer (Invitrogen), 8 μl 10 mM DTT, 200 U of M-MLV Reverse Transcriptase (Invitrogen), and 20 U RNAse inhibitors (Invitrogen). This RT master mix was incubated at 37 °C for 3 h. The reaction was terminated by heating at 70 °C for 10 min then chilling on ice for 5 min.

The PCR was performed in a 25 μl reaction containing 2 μl first strand cDNA; 15 nmol each of downstream and upstream primers; 5 μl of 10× PCR buffer ($Mg^{2+}$ Plus, TaKaRa); 4 μl of 2.5 mmol dNTPs; 2 U *Taq* polymerase (TaKaRa); and 18 μl of water. 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. A product, detectable by ethidium bromide staining, of the expected size was generated.

### 2.4. The 5′- and 3′-ends of the genome

A cDNA clone representing the 5′ and 3′ ends of the genome of the four IBV strains was synthesized according to the 5′ RACE and 3′ RACE System for rapid amplification of cDNA ends (TaKaRa). PCR was performed according to the instructions accompanying the kits. The sense and antisense primers used to amplifying the 5′- and 3′-ends of the genome had been designed on the basis of the sequences obtained above that were constant in the four IBV strains, respectively. The outer and inner primers used to amplify the 5′-end of the four IBV strains were 5′-CAGCTATGGCAATGCG CAG-3′ and 5′-CATCTTTGGTGTCTCA/TCC-3′, respectively. The primer used to amplify the 3′-end was 5′-GAGGAGAGGAACAATGC ACA-3′.

### 2.5. DNA cloning and sequence determination

The DNA generated by PCR amplification was cloned using a T-tailed vector, pMD18-T (TaKaRa), and transformed using JM109 competent cells (TaKaRa) according to the manufacturer's instructions. At least five clones of each fragment in each strain were sequenced and the consensus sequence was determined. The sequences were analyzed using the Sequencher 4.5 sequence