Virus Research 211 (2016) 145–150

Contents lists available at ScienceDirect

Virus Research

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

Molecular detection and Smoothing spline clustering of the IBV strains detected in China during 2011–2012

Zhikun Zhang[1], Yingshun Zhou[1], Hongning Wang*, Fanya Zeng, Xin Yang, Yi Zhang, Anyun Zhang

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

ARTICLE INFO

Article history:
Received 21 April 2015
Received in revised form
22 September 2015
Accepted 13 October 2015
Available online 19 October 2015

Keywords:
Infectious bronchitis virus (IBV)
S1 gene
Smoothing spline clustering (SSC)
Molecular characteristics
Evolutionary analysis

ABSTRACT

Infectious bronchitis virus (IBV) is a highly variable virus with a large number of genotypes. During 2011–2012, nineteen wild IBV strains were isolated in China. Sequence analysis showed that these isolates were divided into five sub-clusters: A2-like, CKCHLDL08I-like, SAIBK-like, KM91-like and TW97/4-like. Phylogenetic analysis based on the 1118 sequences available on line suggested that all IBVs were classified into six clusters. The prevalent strains including all the isolates were in cluster VI with a 0.194–0.259 genetic distance to Mass type vaccines. In addition, we introduced the smoothing spline clustering (SSC) method to estimate the highly variable sites for some sub-clusters. The results showed that highly variable sites range from sub-clusters, the N-terminal sequences of 4/91-like, TW97/4-like and Arkansas-like are more variable than other sub-clusters. This is the first time that the SSC method has been used for the evolution study of IBV.

1. Introduction

Avian infectious bronchitis (IB) is an endemic viral respiratory disease. It affects chickens of different ages and results in economic loss for the poultry industry. The causative agent is infectious bronchitis virus (IBV) (Cavanagh, 2007). The virus is enveloped with a positive-sense, single-stranded RNA genome about 27.6 kb. The genome encodes four structural proteins: spike, envelope, membrane, nuclear protein (S, E, M, N) (Zhou et al., 2013). Spike structural protein (S), located at the surface of the virus, is cleaved into matured S1 and S2 subunits during infection (Cavanagh et al., 1986). The S1 subunit is highly variable, it contains virus-neutralizing epitopes and serotype-specific epitopes (Cavanagh et al., 1992; Ignjatovic and Sapats, 2005). It is also responsible for tissue tropism (Casais et al., 2003; Wickramasinghe et al., 2011). The analyses of the S1 gene sequences are usually effective for differentiating the genotypes of IBV (Lee et al., 2003).

IBV was firstly identified at early 1940s in US as the Massachusetts (Mass). Vaccines of Mass type have been developed subsequently and used worldwide until now. Nowadays, researchers have done many epidemiology studies of IBV. Many genotypes have been definite (Cook et al., 2012). Variants were found worldwide, and vaccines developed from Mass showed poor cross-protection for some variants (Bourogaa et al., 2014; Liu et al., 2009). In China, IBV has been isolated since the late 1970s. Many kinds of variants have been isolated in China, such as QX-IBV, 4/91, KM91 and field strains SAIBK in Sichuan, TW97/4 in Taiwan (Han et al., 2011).

In order to analyze the new 19 isolates in Southwestern China, we compared these S1 genes with previously published records in the NCBI to evaluate genotypes of these isolates and made a general table of all these sequences. Furthermore, the smoothing spline clustering (SSC) method is used to postulate the molecular characterisation of IBV and to analyze the highly variable regions in the S1 sequence. SSC is a program based on the R project for statistical computing, it can count up the consistency rate of each site and draw a smoothing curve. We can visualize the consistency and variable regions of different sub-clusters in the result. These variable regions may be responsible for the antigenic domains and associated with the evolution of IBV.

* Corresponding author. Fax: +86 02885471599.
E-mail address: whongning@163.com (H. Wang).
[1] These authors contributed equally to the research.

http://dx.doi.org/10.1016/j.virusres.2015.10.015
0168-1702/

2. Materials and methods

2.1. Sample collection and virus amplification

66 field samples were collected from laying birds of different ages with respiratory and nephropathogenic symptoms in Southwestern China. The collected lung, kidney and tracheal tissues were ground in 0.1% phosphate-buffered saline (PBS), frozen and thawed three times, treated with 200 μg/ml gentamicin and 200 U/ml penicillin overnight, followed by centrifuged at 4000 × *g* for 5 min. RNA of these samples was extracted and determined by RT-PCR with the M gene and 3′-UTR primers (MF: 5′-CCTAAGAACGGTTGGAAT-3′, MR: 5′-TACTCTCTACACACACA C-3′, 3′F: 5′-GGAAGATAGGCATGTAGCTT-3′, 3′R: 5′-CTAACTCTATACTAG CCTAT-3′) (Ji et al., 2011; Luo et al., 2012). The isolation of IB viruses was performed as previously described (Sun et al., 2011). In order to obtain the full length sequence of S1, the homogenised extract of IBV positive tissue was inoculated in SPF chicken embryos for virus amplification. Blind passages were carried out until the occurrence of characteristic embryo changes, such as dwarfing, stunting, or curling of the embryos.

2.2. Hemagglutination assay (HA) and S1 sequencing

Hemagglutination assay (HA) was used to detect IBV in the allantoic fluid, according to the protocol described by the Office International des Epizooties, with the positive controls of SAIBK, M41 and H120 virus and the negative controls of SPF allantoic fluid and PBS. Viral RNA was extracted from inoculated allantoic fluid using the RNA isolation kit (RNeasy. Qiagen) and reverse transcribed with SuperScript II reverse transcriptase (Invitrogen). To amplify the whole S1 cDNA from various field samples, PCR primers (IBVS1-F: 5′-AACTCTGCACGCAAATTA-3′, IBVS1-R: 5′-TGTTGTCGCAAACAGGACC-3′) were designed according to the sequences of M41, H120, SAIBK, QX, LX4, Ma5, Beaudette and some other IBV sequences from NCBI. The cycling conditions were: initial at 94 °C for 5 min, followed by 29 cycles of 94 °C 30 s, 53 °C for 30 s and 68 °C for 3 min, followed by a single final extension step at 68 °C for 10 min. The amplified S1 fragments were cloned for subsequent sequencing reactions (Sangon Biological Engineering Technology & Services Co., Ltd.).

2.3. Gene comparison, phylogenetic analysis and smoothing spline clustering (SSC)

The amino acid alterations in S1 were confirmed by software MegAlign (www.dnastar.com) and MEGA 5.0 (Tamura et al., 2011). All these sequences have been submitted to GenBank. The total 1118 full-length S1 sequences were utilised to construct a phylogenetic tree of IBV (Table 1). The full-length S1 was defined as the sequence with a starting codon ATG and the proteolytic cleavage site (RRXRR) of the S protein. Combined with the 19 S1 isolated in this study, we align a total of 1137 S1 sequences at nucleotide level by using MEGA 5.0 by ClustalW method. We constructed the phylogenetic tree using the Neighbor-Joining method with 2000 Bootstrap replicates. Recombinant events of the isolates were detected with recombination detection program (RDP) and confirmed by Simplot. Sequence variations of S1 were estimated by smoothing spline clustering (SSC) analysis. This method is a program based on statistical software R: sequences of different groups were aligned with MEGA 5.0 by ClustalW method, the alignment results were input to R software by “seqinr” package (Charif and Lobry, 2007), then count up the consistency rate of each site, draw curves with locally weighted scatter plot smoothing LOESS (LOWESS) method and show sites separately with consistencies below 0.6. All this steps were programmed into the SSC script. We can easily get our result by inputting the alignment results of each sub-cluster into our script and running the software R i386 3.0.1, as we previously described (Zhou et al., 2012).

Table 1
The detailed information of the 1137 sequences (starting at the ATG translation initiation codon and ending at the cleavage recognition motifs): The table was constructed following the shape of the phylogenetic tree (Supplementary figures). The phylogenetic tree was constructed using the MEGA version 5.0 by the neighbor-joining method with 2000 bootstrap replicates. Represented isolates, the number of sequences, distribution, and mean genetic distances within each sub-cluster are listed below. The isolates are marked with ” +”. Reported nephropathogenic represented strains in cluster VI are shown in bold words.

Cluster	Subcluster	Represent isolates	Number of sequences	Distribution	Mean genetic distance
VI	6.10	**A2**	349 + 5	China, Thailand	0.028
	6.9	**QX**	119	China, Europe, Korea	0.073
	6.8	4/91	99	USA, Europe, Japan, IR, China	0.123
	6.7	**CKCHLDL08I**	29 + 2	China	0.012
	6.6	**KM91**	54 + 2	China, Korea	0.039
	6.5	**D971**	21	China, Thailand	0.151
	6.4	SAIBwj	1	China	–
	6.3	**SAIBK**	76 + 3	China	0.070
	6.2	SAIB4	1	China	–
	6.1	**TW97/4**	42 + 7	China, Thailand, USA, Taiwan	0.128
V	5.8	RF-19-99	1	Russia	–
	5.7	NGA/SOK/4032007	1	Nigeria	–
	5.6	NGA/324/2006	2	Nigeria and Belgian	0.252
	5.5	T3	19	China, Nigeria, Egypt	0.138
	5.4	UFGM	13	Brazil	0.214
	5.3	Mass	117	Asia, USA, Europe	0.026
	5.2	JAAS	20	China, Europe, Australia	0.163
	5.1	Arkansas	110	Korea, China, USA	0.147
IV	4	CKCHLAH08I	3	China	0.179
III	3	N4-02	3	Australia	0.002
II	2.1	Q388	3	Australia	0.177
	2.2	TC07-2	7	China	0.171
I	1	DE072	28	USA	0.066