Infection, Genetics and Evolution 12 (2012) 1870–1878

Contents lists available at SciVerse ScienceDirect

Infection, Genetics and Evolution

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

Molecular and phylogenetic analysis of bovine coronavirus based on the spike glycoprotein gene

Nadia Martínez [a], Paulo E. Brandão [b], Sibele Pinheiro de Souza [b], Maritza Barrera [a], Nelson Santana [b,c], Heidy Díaz de Arce [a], Lester J. Pérez [a,*]

[a] Centro Nacional de Sanidad Agropecuaria (CENSA), La Habana, Cuba
[b] Department of Preventive Veterinary Medicine and Animal Health, College of Veterinary Medicine, University of São Paulo, São Paulo, Brazil
[c] Universidad Nacional de Colombia, Bogotá, Colombia

ARTICLE INFO

Article history:
Received 15 February 2012
Received in revised form 15 May 2012
Accepted 17 May 2012
Available online 24 May 2012

Keywords:
Bovine coronavirus
Winter dysentery
Phylogenetic analysis
Recombination analysis
Heptad repeat

ABSTRACT

Bovine coronavirus has been associated with diarrhoea in newborn calves, winter dysentery in adult cattle and respiratory tract infections in calves and feedlot cattle. In Cuba, the presence of BCoV was first reported in 2006. Since then, sporadic outbreaks have continued to occur. This study was aimed at deepening the knowledge of the evolution, molecular markers of virulence and epidemiology of BCoV in Cuba. A total of 30 samples collected between 2009 and 2011 were used for PCR amplification and direct sequencing of partial or full S gene. Sequence comparison and phylogenetic studies were conducted using partial or complete S gene sequences as phylogenetic markers. All Cuban bovine coronavirus sequences were located in a single cluster supported by 100% bootstrap and 1.00 posterior probability values. The Cuban bovine coronavirus sequences were also clustered with the USA BCoV strains corresponding to the GenBank accession numbers EF424621 and EF424623, suggesting a common origin for these viruses. This phylogenetic cluster was also the only group of sequences in which no recombination events were detected. Of the 45 amino acid changes found in the Cuban strains, four were unique.

1. Introduction

Bovine coronavirus (BCoV) was first identified in association with diarrhoea in newborn calves (Mebus et al., 1973) and later associated with winter dysentery (WD) in adult cattle (Saif et al., 1991) and respiratory tract infections in calves and feedlot cattle (Storz et al., 2000). Although the affected animals rarely die, coronavirus infection causes dramatic reductions in milk production in dairy herds and loss of body condition in both calves and adults (Saif et al., 1998), resulting in severe economic losses. Thus, BCoV is currently considered an important pathogen that causes enteric disease, often in combination with clinical respiratory signs.

BCoV is included in the genus *Betacoronavirus* of the family Coronaviridae, which, together with the families *Arteriviridae* and *Roniviridae*, constitute the order *Nidovirales* (International Committee for Taxonomy of Viruses (ICTV): http://talk.ictvonline.org/cfs-filesystemfile.ashx/key/CommunityServer.Components.PostAttachments/00.00.00.06.26/2008.085_2D00_122V.01.Coronaviridae.pdf). The BCoV genome consists of a single molecule of linear, positive-sense, single-stranded RNA of 31 kb in length that is transcribed into a nested set of several 3′-coterminal subgenomic mRNAs that produce both non-structural and structural proteins (Chouljenko et al., 2001). The virion contains five structural proteins: the nucleocapsid (N) protein, the transmembrane (M) protein, the small envelope (E) protein, the haemagglutinin–esterase (HE) protein and the spike (S) protein (Lai and Cavanagh, 1997).

The S glycoprotein is important for viral entry and pathogenesis, forms large petal-shaped spikes on the surface of the virion and is cleaved into S1 (N-terminus) and S2 (C-terminus) subunits (Abraham et al., 1990). The S1 is the globular subunit and is responsible for virus binding to host–cell receptors (Kubo et al., 1994), the induction of neutralising antibody expression (Yoo and Deregt, 2001) and haemagglutinin activity (Schultze et al., 1991). It's sequences are variable, and mutations in this region have been associated with changes in antigenicity and viral pathogenicity (Ballesteros et al., 1997). On the other hand, the S2 is the transmembrane subunit and is required to mediate the fusion of viral and cellular membranes (Luo and Weiss, 1998).

Variations in the host range and tissue tropism of the coronaviruses have been largely attributable to variations in the S glycoprotein (Gallagher and Buchmeier, 2001). Therefore, to identify biological, antigenic and genetic characteristics that are distinct between respiratory and enteric BCoV strains, several studies based on partial or complete S gene sequences have been conducted (Brandão et al., 2006; Decaro et al., 2008; Hasoksuz et al.,

* Corresponding author. Tel.: +53 47863206; fax: +53 47861104.
E-mail address: lesterjosue@censa.edu.cu (L.J. Pérez).

1567-1348/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.meegid.2012.05.007

2002; Kanno et al., 2007; Liu et al., 2006). Although no clear markers have been established, comparative nucleotide sequence analyses have been useful for investigating the molecular phylogeny of BCoV (Brandão et al., 2006; Decaro et al., 2008; Kanno et al., 2007; Liu et al., 2006).

In Cuba, the presence of BCoV was first reported in 2006 (Barrera et al., 2006), and sporadic outbreaks have continued to occur since that time (Martínez et al., 2010).

In this study, sequence comparisons and phylogenetic studies based on S gene sequences were performed to deepen the knowledge of the evolution, potential molecular markers of virulence and epidemiology of BCoV in Cuba.

2. Materials and methods

2.1. Samples collection

A total of 30 faecal samples from dairy and beef cows that were affected with enteric manifestations resembling WD were selected from a total of 136 samples that were sent to the Animal Virology Group of the CENSA for BCoV diagnosis. The samples for this study were selected based on the geographic region of origin and the year of the collection (Fig. 1).

2.2. Laboratory procedure

To inactivate potential RT-PCR inhibitors contained in the faecal samples, the faeces were diluted in nuclease-free water (Promega, Madison, WI, USA) at a ratio of 3:1 (v/v). The final suspensions were centrifuged at 5000*g* for 10 min at 4 °C. RNA was extracted from 250 μL of supernatant recovered using the TRIzol reagent (Invitrogen™/Life Technologies, Carlsbad, CA, USA) according to the manufacturer's instructions. Finally, the RNA pellet was diluted in 30 μL of nuclease-free water (Promega, Madison, WI, USA).

First-strand complementary DNA (cDNA) was synthesised using Moloney-Murine leukaemia virus reverse transcriptase (M-MLV RT) (Invitrogen) and random primers (50 ng/μL) (Invitrogen) in a 20 μL final reaction volume. The cDNA of each sample was screened for the BCoV genome using the PCR method described by Tsunemitsu et al. (1999).

2.3. Sequencing

In the first approach, the hypervariable region of the S1 gene (approx. 488 bp) was amplified from BCoV-positive samples using the primer pairs reported by Souza et al. (2010) and the reaction conditions described by Brandão et al. (2006). The full S gene was then amplified from four positive samples (see samples marked with asterisk in Fig. 1) using several primer pairs (Table 1). Each fragment of the S gene was amplified in a reaction volume of 50 μL containing 10 μL cDNA, 2.5 U Platinum® Taq DNA polymerase (Invitrogen), 200 μM of each dNTP, 2.5 mM $MgCl_2$ and 0.5 μM of each primer.

The resulting amplicons were purified from agarose gels using a GFX PCR DNA and GEL BAND Purification Kit® (GE Healthcare) and submitted to bi-directional DNA sequencing using a BigDye Terminator v3.1 cycle sequencing kit following the manufacturer's directions (Applied Biosystems). Sequencing products were read on an ABI PRISM-3100 Genetic Analyzer (Applied Biosystems). The sense and antisense sequences obtained from each amplicon were assembled, and a consensus sequence for each gene was generated using the ChromasPro V1.5 program (Technelysium Pvt. Ltd., 2009). Nucleotide BLAST analysis (http://www.ncbi.nlm.nih.gov/blast/Blast.cgi) was initially used to verify the identity of each fragment sequence obtained. The sequences were submitted to the GenBank database under accession numbers HE616734–HE616737 for the hypervariable region and HE616738–HE616741 for the full S gene.

Total samples selected	Year of collection	Category	Geographic region	Positive RT-PCR results
10	2009	dairy cow	Mayabeque (A)	3
5	2010	dairy cow	Matanzas (B)	1
5	2010	dairy cow	Cienfuegos (C)	2
5	2011	beef cow	Ciego de Avila (D)	1
5	2011	beef cow	Camagüey (E)	1

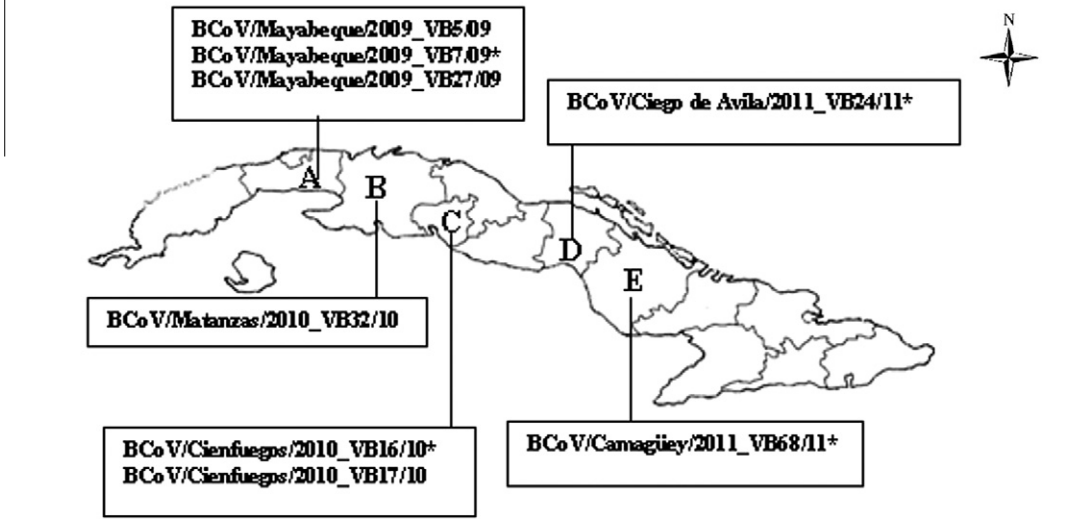

Fig. 1. Map of Cuba showing the geographic distribution of the sample collection sites. The associated table indicates the quantity of samples analysed from each area and the positive RT-PCR results obtained from each of them (*complete S gene sequenced samples).