Available online at www.sciencedirect.com
SCIENCE DIRECT
Virus Research 115 (2006) 56–68
Virus Research
www.elsevier.com/locate/virusres

The complete sequence of the bovine torovirus genome

Ryan Draker [a,1], Rachel L. Roper [b], Martin Petric [c], Raymond Tellier [a,d,*]

[a] *Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, Ont., Canada*
[b] *Department of Microbiology and Immunology, Brody School of Medicine, East Carolina University, NC, USA*
[c] *British Columbia Center for Disease Control, Vancouver, BC, Canada*
[d] *Division of Microbiology, The Hospital for Sick Children, 555 University Avenue, Toronto, Ont., Canada M5G 1X8*

Received 16 March 2005; received in revised form 5 July 2005; accepted 12 July 2005
Available online 30 August 2005

Abstract

Viruses in the family *Coronaviridae* have elicited new interest, with the outbreaks caused by SARS-HCoV in 2003 and the recent discovery of a new human coronavirus, HCoV-NL63. The genus *Torovirus*, within the family *Coronaviridae*, is less well characterized, in part because toroviruses cannot yet be grown in cell culture (except for the Berne virus). In this study, we determined the sequence of the complete genome of Breda-1 (BoTV-1), a bovine torovirus. This is the first complete torovirus genome sequence to be reported. BoTV-1 RNA was amplified using long RT-PCR and the amplicons sequenced. The genome has a length of 28.475 kb and consisted mainly of the replicase gene (~20.2 kb) which contains two large overlapping ORFs, ORF1a and ORF1b, encoding polyproteins pp1a and pp1b, respectively. Sequence analysis identified conserved domains within the predicted sequences of pp1a and pp1b. Sequence alignments and protein secondary structure prediction data suggest the presence of a 3C-like serine protease domain with similarity to the arterivirus 3C-like serine protease and a single papain-like cysteine protease domain with similarity to the picornavirus leader protease. The ADRP (APPR-1″) domain – unique to the *Coronaviridae* – was also located in BoTV pp1a. In addition, several hydrophobic domains were identified that are typical of a nidovirus replicase. Within the pp1b sequence the polymerase and helicase domains were identified, as well as sequences predicted to be involved in ribosomal frameshifting, including the conserved slippery sequence UUUAAAC and two potential pseudoknot structures.

Keywords: Bovine torovirus; Complete sequence; Proteases; Nidovirales; Long PCR

1. Introduction

Toroviruses are enveloped, single-stranded, positive-sense RNA viruses with a pleomorphic virion morphology. The first report on toroviruses described the identification of an unclassified virus from diarrheic calves (Woode et al., 1982), now designated as the Breda virus (BRV), or bovine torovirus (BoTV), two serotypes of which have been identified (BoTV-1 and BoTV-2). The virus is endemic in cattle herds, with asymptomatic cows possibly acting as reservoirs. Newborn calves generally develop symptoms of diarrhea, typically lasting 2–13 days (Koopmans et al., 1990, 1991) and BoTV has been commonly recovered from calves with diarrhea (Duckmanton et al., 1998a,b; Hoet et al., 2003a,b). Infection with BoTV has also been conclusively associated with diarrhea in veal calves (Hoet et al., 2003a,b). Interestingly, BoTV has been reported as well in nasal secretions, and may therefore be able to infect both the respiratory and gastrointestinal tract (Hoet et al., 2002), as is the case with some coronaviruses (Lai and Holmes, 2001), including the SARS coronavirus (Leung et al., 2003). The equine torovirus (EqTV), or Berne virus (BEV), has been designated as the genus prototype; it can be grown in cell culture (unlike BoTV) and consequently is better characterized. Despite the presence of anti-EqTV antibodies in horses, no disease has been firmly associated with EqTV infection.

* Corresponding author. Tel.: +1 416 813 6592; fax: +1 416 813 6257.
E-mail address: raymond.tellier@sickkids.ca (R. Tellier).
[1] Present address: Ontario Cancer Institute, Princess Margaret Hospital, and Department of Medical Biophysics, University of Toronto, Toronto, Ont., Canada.

0168-1702/$ – see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.virusres.2005.07.005

Toroviruses have also been reported in humans with gastroenteritis (Beards et al., 1984; Duckmanton et al., 1997; Jamieson et al., 1998) and pigs (Kroneman et al., 1998; Smits et al., 2003). Antibodies reacting against BoTV and EqTV antigens have been detected in various other mammals, suggesting toroviruses may be widespread (Brown et al., 1987; Weiss et al., 1984).

In vitro studies on EqTV grown in cell culture demonstrated the generation of sub-genomic RNAs (sgRNAs) for translation of the ORFs coding for the structural proteins. Approximately 16 kb of sequence from the 3′ end of the genome, and approximately 1.5 kb from the 5′ end of genome have been determined. These sequence data have revealed the basic torovirus genome organization and the presence of conserved replicase domains(Snijder and Horzinek, 1993). These findings were instrumental in the revision of the *Coronaviridae* taxon to include the genus *Torovirus*, as well as the creation of the order *Nidovirales* which currently includes the families *Coronaviridae* (comprising two genera, *Coronavirus* and *Torovirus*), *Arteriviridae* and *Roniviridae* (Cavanagh, 1997; Cavanagh et al., 1994; Cavanagh and Horzinek, 1993; González et al., 2003). The complete genome sequence of EqTV has yet to be published, although Smits et al. (2003) reported the completion of the EqTV genome sequence in a recent study of torovirus field variants. The published sequence of BoTV is limited to that of the structural genes, which account for approximately 7.5 kb of the 3′ end of the genome (Cornelissen et al., 1997; Duckmanton et al., 1998a,b). Consequently, the genus *Torovirus* remains the only genus within the order *Nidovirales* without a representative complete genome sequence.

In this study, we sequenced the entire genome of the bovine torovirus, BoTV-1, using long RT-PCR followed by the sequencing of amplicons. Assembly of the sequences resulted in a genome of 28,475 nucleotides. Sequence analysis revealed the presence of a nidovirus-like replicase gene with conserved domains.

2. Materials and methods

2.1. Source of BoTV-1

BoTV-1 was obtained from Dr. Gerald Woode, as described (Duckmanton et al., 1998a,b). The virus, consisting of stool specimens from gnotobiotic calves infected with the original Breda-1 virus, was aliquoted and stored at −80 °C.

2.2. Extraction of viral RNA

BoTV-1 RNA was extracted from fecal specimens as described (Duckmanton et al., 1998a,b) using TRIzol Reagent (Invitrogen, Burlington, Ont., Canada). The RNA was resuspended in 10 μl of ddH_2O containing 10% 100 mM dithiothreitol (Invitrogen) and 5% 20–40 U/μl RNasin (Promega, Mississauga, Ont.), and stored at −80 °C.

2.3. Primers

Primers used in RT-PCR (synthesized by Invitrogen) were designed using Gene Runner 3.05 (Hastings Software, Inc.), based on the available BoTV-1 sequence, and EqTV sequence (GenBank accession no., X52374, and X56016). Primer names, sequences, and positions within the BoTV-1 genome are given in Table 1. One primer, APPR-RS, was designed within a conserved region of the coronavirus open reading frame (ORF) 1a of the replicase gene. This region was first identified by ORF1a amino acid alignment, followed by the corresponding nucleotide alignment of the infectious bronchitis virus (IBV), GenBank accession no. NC001451, nucleotides (nts) 3646–3699; human coronavirus 229E (HCoV 229E), NC002645, nts 4193–4246; murine hepatitis virus (MHV) A59, NC001846, nts 4239–4292; and transmissible gastroenteritis virus (TGEV), AJ271965, nts 4365–4418. Alignments were done using ClustalX 1.8 (Thompson et al., 1997). The sequence of the APPR-RS primer was based on the consensus sequence of the nucleotide alignment described above. The EEAT7 primer used in the amplification of the 3′ non-coding region (3′ NC) is a coxsackie virus B6 specific primer (Martino et al., 1999). The complementary sequence of the first 21 nucleotides (3′–5′) of this primer were incorporated into the BRE-3′NC-CVB6 primer used in the reverse transcription reaction. Lastly, the AAP-CVB6 primer used in the 5′ RACE PCR reaction was designed to bind within the poly(A) of the cDNA transcripts, and contained sequence complementary to the 3′ end of the EEAT7 primer, allowing for the option of using EEAT7 in a reamplification of 1st round PCR products.

2.4. Long RT-PCR

The long RT-PCR was done essentially as described (Tellier et al., 1996a,b, 2003). Briefly, purified viral RNA was thawed on ice, incubated at 65 °C for 2 min then placed back on ice. To the RNA was added 10 μl of a master mix composed of 4 μl of 5× 1st Strand Synthesis Buffer (Invitrogen), 0.5 μl of RNasin (20–40 U/μl) (Promega), 1 μl of 100 mM dithiothreitol (Invitrogen), 1 μl of a 10 mM solution of deoxynucleotide triphosphates (dNTPs) (Pharmacia, Piscataway, NJ), 2.5 μl of a 10 μM solution of antisense primer, and 1 μl of Superscript II reverse transcriptase (Invitrogen). The reaction mixture was incubated at 42 °C for 1 h, after which 1 μl each of RNase H (1–4 U/μl) and RNase T1 (900–3000 U/μl) (Invitrogen) were added and the reaction mixture was incubated at 37 °C for 20 min. Ten-fold dilutions of the newly synthesized cDNA were then prepared using ddH_2O and the neat and diluted cDNA samples were kept on ice for use in PCR or stored at −80 °C.

The PCR master mix was composed of 5 μl of 10× Advantage 2 PCR Reaction Buffer (Clontech, Palo Alto, CA), 1.25 μl of a 10 mM solution of dNTP mixture (Amersham