Discovery of a novel nidovirus in cattle with respiratory disease

Rafal Tokarz,1 Stephen Sameroff,1 Richard A. Hesse,2 Ben M. Hause,2 Aaloki Desai,1 Komal Jain1 and W. Ian Lipkin1

1Center for Infection and Immunity, Mailman School of Public Health, Columbia University, NY 10032, USA
2Kansas State Veterinary Diagnostic Laboratory and Department of Diagnostic Medicine and Pathobiology, Kansas State University, Manhattan, KS 66508, USA

The family Coronaviridae represents a diverse group of vertebrate RNA viruses, all with genomes greater than 26 000 nt. Here, we report the discovery and genetic characterization of a novel virus present in cattle with respiratory disease. Phylogenetic characterization of this virus revealed that it clusters within the subfamily Torovirinae, in the family Coronaviridae. The complete genome consists of only 20 261 nt and represents the smallest reported coronavirus genome. We identified seven ORFs, including the canonical nidovirus ORF1a and ORF1b. Analysis of polyprotein 1ab revealed that this virus, tentatively named bovine nidovirus (BoNV), shares the highest homology with the recently described python-borne nidoviruses and contains several conserved nidovirus motifs, but does not encode the NendoU or O-MT domains that are present in other viruses within the family Coronaviridae. In concert with its reduced genome, the atypical domain architecture indicates that this virus represents a unique lineage within the order Nidovirales.

The order Nidovirales comprises enveloped, positive-sense, ssRNA viruses that share similar genome organization, homology in the replicase polyprotein, and a unique replication strategy (deGroot et al., 2011a; Gorbalenya et al., 2006). The genomic architecture of nidoviruses is characterized by two large, partially overlapping ORFs, ORF1a and ORF1b, that encode components of the viral replicase and occupy approximately two-thirds of the genome. ORFs encoding the structural components are located downstream of ORF1a/b and are expressed from 3' co-terminal subgenomic mRNAs (Brian & Baric, 2005; Sawicki et al., 2007).

Nidoviruses are divided taxonomically into four families, Arteriviridae, Coronaviridae, Mesoniviridae and Roniviridae (deGroot et al., 2011a; Lauber et al., 2012). In addition to phylogenetic classification, genome size is a distinctive feature of these viral families. Arteriviruses (genome size of 12.7–15.7 kb) have been referred to as ‘small nidoviruses’, mesoniviruses (20 kb) as ‘intermediate’, and both roniviruses (~26 kb) and coronaviruses (26–33 kb) as ‘large’ (deGroot et al., 2011a; Gorbalenya et al., 2006; Nga et al., 2011). On the basis of genetic and phenotypic differences, the family Coronaviridae is further divided into the subfamilies Coronavirinae and Torovirinae (deGroot et al., 2011b). The subfamily Torovirinae consists of two genera, Torovirus and Bafinivirus. The genus Torovirus includes four viruses originating from mammals (human torovirus, bovine torovirus, equine torovirus and porcine torovirus) (Beards et al., 1984; Draker et al., 2006; Sun et al., 2014; Weiss et al., 1983). The genus Bafinivirus includes two viruses isolated from fish (White bream virus and White minnow nidovirus) (Batts et al., 2012; Schütze et al., 2006). Within the past year, nidoviruses have been described in pythons with pneumonia [python nidovirus (PNV) and ball python nidovirus (BPNV)] and these viruses are proposed to represent a novel genus within the subfamily Torovirinae (Bodewes et al., 2014; Stenglein et al., 2014; Uccellini et al., 2014). Here, we report on the discovery and genomic characterization of a novel virus that represents a distinctive lineage within this subfamily.

In 2013, an outbreak of severe respiratory disease occurred in a cattle feedlot located in the south-western USA. The animals suffered from severe tracheitis and pneumonia, with a high mortality rate. Although a viral agent was suspected, initial culture efforts were unsuccessful. In an effort to examine the breadth of viral agents present, postmortem tissue samples were collected from four animals and analysed by high-throughput sequencing (HTS) using the Illumina HiSeq 2500 platform (Table 1). The GenBank/EMBL/DBJ accession number for the sequence of bovine nidovirus determined in this study is KM589359.
Analysis of HTS data identified the presence of a virus with <35 % amino acid homology to viruses within the subfamily Torovirinae in samples from three animals (Table 1). The complete genome of this virus, provisionally called bovine nidovirus (BoNV), was assembled from the lung sample of one of the animals. The authenticity of the draft genome was confirmed by PCR using specific primers designed to generate overlapping fragments along the length of the genome and was followed by dideoxy sequencing of the resulting amplification products. The genome termini were obtained by 5’ and 3’ RACE. Using this genome sequence as a reference, we assembled >80 % of the BoNV genomes present in the other two virus-positive animals. These sequences were 99 % identical to the reference sequence, suggesting that all three animals were infected with the same strain of BoNV.

The complete genome of BoNV comprises 20,261 nt and contains seven ORFs (Fig. 1a). Flanking the ORFs are 486 nt 5’ and 453 nt 3’ UTRs. The initial two ORFs are homologous to the canonical ORF1a and ORF1b of mem-

### Table 1. List of samples analysed by HTS

<table>
<thead>
<tr>
<th>Animal</th>
<th>Sample type</th>
<th>Bovine nidovirus*</th>
<th>Other viruses identified by HTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Lung homogenate</td>
<td>–</td>
<td>Bovine coronavirus</td>
</tr>
<tr>
<td></td>
<td>Lymph node homogenate</td>
<td>+</td>
<td>Parvovirus, bovine herpesvirus</td>
</tr>
<tr>
<td>2</td>
<td>Trachea homogenate</td>
<td>–</td>
<td>Bovine herpesvirus</td>
</tr>
<tr>
<td>3</td>
<td>Lung homogenate</td>
<td>+</td>
<td>Bovine herpesvirus</td>
</tr>
<tr>
<td></td>
<td>Abomasum homogenate</td>
<td>–</td>
<td>6 &gt;</td>
</tr>
<tr>
<td>4</td>
<td>Trachea homogenate</td>
<td>+</td>
<td>Bovine rhinovirus B, bovine herpesvirus</td>
</tr>
</tbody>
</table>

*Nidovirus sequences were identified by HTS and were confirmed by PCR.

Fig. 1. (a) Schematic representation of the BoNV genome. Each ORF is designated by an arrow corresponding to its approximate size. The ribosomal frameshift sequence UUUAAAC that results in translation of ORF1a/b is indicated by a black arrow. The ORF designations S, M and N correspond to the glycoprotein spike, membrane and nucleocapsid proteins, respectively. The nucleotide coordinates are displayed on top. Genome assemblies were done in Geneious v.6.1.5. (b) Comparison of the length and domain organization of polyprotein 1ab of BoNV and its closest homologue, the BPNV. The location and approximate length of each domain is shown. The ORF1a portion of the polyprotein is indicated in grey, ORF1b in white. The amino acid coordinates are shown on top. The schematic of the BPNV was modified from Stenglein et al. (2014).
bers of the order Nidovirales. During viral replication, the expression of ORF1α yields polyprotein 1a (pp1a) while translation of ORF1b is initiated by a −1 ribosomal frameshift sequence upstream of the ORF1a stop codon and results in polyprotein 1ab. In BoNV, the putative frameshift sequence of UUUAAAC was identical to the sequence found in toroviruses and banifiviruses but differs from the sequence identified in the python-borne viruses (AAA-AAC). BLASTP homology searches using the amino acid sequences of both pp1a and pp1b revealed that BoNV is most similar to viruses within the subfamily Torovirinae, although limited overall sequence identity was observed throughout either polyprotein (<30 %). The highest identity to BoNV within pp1a or pp1b was observed within the python-borne nidoviruses PNV and BPNV (Table S1, available in the online Supplementary Material). Notably, the combined size of BoNV pp1ab was the smallest among all members of the subfamily Coronaviridae. Of all nidoviruses, only arteriviruses were found to contain a smaller polyprotein.

The pp1a and pp1ab contain several conserved domains encoding enzymes that are processed into their individual components following post-translational cleavage of the polyprotein (deGroot et al., 2011a). We used protein prediction tools including INTERPRO, HMMER3 and BLASTP to identify such signature motifs within the polyproteins of BoNV. Analysis of pp1a revealed the presence of a serine/threonine protein kinase, three hydrophobic regions (TM1–TM3) and the main protease (Mpro) involved in post-translational processing of pp1a/b (Fig. 1b). The Mpro domain was encoded between TM2 and TM3, a location conserved in all nidoviruses. We also identified a homologue of coronavirus non-structural protein 10 (nsp10), a zinc-binding protein that is involved in RNA synthesis (Bouvet et al., 2014).

ORF1b encodes the more highly conserved portion of pp1ab and is characterized by a set of domains encoding replicative enzymes. The order of these domains is conserved and in coronaviruses encompasses the RNA-dependent RNA polymerase (RdRp), superfamily 1 helicase (Hel1), 3’ to 5’ exoribonuclease (ExoN), N7-methyltransferase (NMT), uridylyl-specific endonuclease (NendoU) and 2’-O-methyltransferase (O-MT). Although all nidoviruses contain the RdRp and Hel, the presence of the remainder of these domains varies across nidoviruses. Analysis of the pp1b of BoNV revealed that it encodes domains homologous to RdRp, Hel1 and ExoN in the canonical order. We did not identify an NMT domain, consistent with its absence in all viruses within the subfamily Torovirinae. Surprisingly, we also did not identify a homologue of NendoU and O-MT domains. All nidoviruses reported to date encode at least one of these enzymes.

For phylogeny inference, the amino acid sequences of the Hel and RdRp domains of viruses within the family Coronaviridae were aligned and compared in MEGA 6.0 (Tamura et al., 2013). For both domains, BoNV forms a monophyletic branch within the subfamily Torovirinae, indicative of a novel taxonomic group within this subfamily (Figs 2 and S1).

BoNV encodes five structural ORFs downstream of ORF1a/b (Fig. 1a). ORF2 is located within the region of the genome typically occupied by the gene encoding the S (‘spike’) protein, a type I glycoprotein involved in receptor binding and membrane fusion. This ORF is 1689 nt in length, and encodes a putative 562 aa protein containing a signal peptidase cleavage sequence between aa 18 and 19, a transmembrane domain between aa 537 and 559 and multiple potential glycosylation sites. Homology searches revealed similarity only to the S of the python-borne viruses, although with only 17 % amino acid sequence identity within the complete protein. In addition, although the BPNV and PNV S proteins are among the shortest within the family Coronaviridae, the BoNV S is only 58 % of their size and represents the shortest coronavirus spike protein reported to date. ORF3 of BoNV encodes a putative 231 aa protein that contains multiple transmembrane (TM) spanning regions similar to coronavirus membrane proteins (M). ORF4 encodes a 178 aa protein predicted by genomic location as a nucleocapsid (N) protein. Similar to other N proteins of members of the subfamily Torovirinae, it is short (approx. half the size of a typical coronavirus N) and highly basic with a predicted isoelectric point of 12.0. ORF5 encodes a 435 aa protein that contains multiple glycosylation sites and a predicted TM region between aa 429 and 448 and likely represents another glycoprotein (designated G2). We also identified a very short, putative 267 nt ORF that would encode an 88 aa protein near the 3’ end of the genome. Unlike ORF1a/b and ORF2, ORF3–6 do not demonstrate significant similarity to any viral or non-viral proteins.

To examine the seroprevalence of BoNV, we generated a luciferase immunoprecipitation assay (LIPS) to measure antibody levels against BoNV (Burbelo et al., 2009). We implemented this assay to examine 127 bovine sera originating from the central and western USA (Kansas, Idaho, Missouri and Texas). All samples were previously submitted for diagnostic testing for other bovine infectious agents. Eleven sera (9 %) were positive for BoNV (Fig. S2). The specificity of the BoNV LIPS assay was demonstrated by the lack of reactivity with bovine-coronavirus-positive sera.

The classification of the order Nidovirales is strictly based on sequence similarity. Based on these criteria, the phylogeny of BoNV suggests that it is a member of a novel genus within the subfamily Torovirinae. However, BoNV has several genetic characteristics that are atypical of viruses in this subfamily. First, at approximately 20 000 nt, the BoNV genome is much shorter than the eight viruses currently classified as members of the subfamily Torovirinae that have genomes that range from 26 000 to 33 000 nt. The ‘intermediate’ size of the BoNV genome is more characteristic of mesoniviruses although it shares little sequence homology with...
Fig. 2. Maximum-likelihood phylogeny of members of the family Coronaviridae based on the amino acid sequence of the RdRp (a) and the helicase (b). Trees were generated using the JTT matrix-based model. Each genus within the subfamilies Coronavirinae and Torovirinae is highlighted by a grey box. For the subfamily Coronavirinae, only representative International Committee on Taxonomy of Viruses-approved species with available complete sequences were included. For the subfamily Torovirinae, newly reported viruses were also included. The GenBank accession numbers for sequences used in the tree are: feline infectious peritonitis virus (NC002306), human coronavirus 229E (NC002645), human coronavirus NL63 (NC005831), porcine epidemic diarrhea virus (NC003436), bat coronavirus HKU2 (NC009988), Scotophilus bat coronavirus 812 (NC009657), bovine coronavirus (NC003045), human coronavirus HKU1 (NC006577), murine hepatitis virus (NC001846), Pipistrellus bat coronavirus HKU5 (NC009020), Rousettus bat coronavirus HKU9 (NC009021), severe acute respiratory syndrome (SARS) coronavirus (NC004718), Tylonycteris bat coronavirus HKU4 (NC009019), munia coronavirus HKU13 (NC011550), thrush coronavirus HKU12 (NC011549), bulbul coronavirus HKU11 (FJ376619), avian coronavirus (NC001451), beluga whale coronavirus SW1 (NC010646), fathead minnow nidovirus (GU002364), porcine torovirus (NC022787), Breda virus (NC007447), Berne virus (X52374), PNV (KJ935003) BPNV (NC024709). Bars, 0.2 substitutions per amino acid position.
viruses in this family. Secondly, BoNV would currently be the only virus within the subfamily Torovirinae lacking an O-MT and NendoU homologue in its genome. Interestingly, BoNV shares the highest amino acid similarity with PNV and BPNV, viruses with the largest RNA genomes discovered to date. For contrast, the genome of BoNV measures only 60 % of the size of these python-borne viruses.

We speculate that this virus has undergone an atypical route of genome expansion. The expansion of the nidovirus genome has been correlated with an evolutionary procurement of enzymes controlling RNA replication fidelity. The acquisition of 3′ to 5′ exoribonuclease (ExoN), a homologue of a DNA proofreading enzyme, has been proposed to be of particular importance in the transition from 'small' genomes (arteriviruses) to 'intermediate' and 'large' (genomes of <20 kb) (Gorbalenya et al., 2006; Lauber et al., 2013; Nga et al., 2011). Consistent with this hypothesis, we identified an ExoN homologue within the 'intermediate'-sized genome of BoNV. The gain of additional enzymes such as O-MT or NendoU, and their role in expansion of nidovirus genome size is less understood. If, as proposed by Lauber et al. (2013), NendoU was present in an ancestral lineage that culminated in the four known families of nidoviruses, it would suggest that this domain was lost at some point in the BoNV evolutionary history. Such domain loss would not be unprecedented as NendoU has been lost in both mesoniviruses and roniviruses (Lauber et al., 2013; Nga et al., 2011). In contrast, the O-MT domain is present in all nidoviruses with 'intermediate' and 'large' genomes. It was suggested that this enzyme may have been acquired following the split of the larger nidoviruses from arteriviruses and if so, its absence in BoNV suggests it too would have been lost similar to NendoU. Alternatively, it is also possible that the ancestral line that gave rise to BoNV never acquired it.

The clinical significance of BoNV is unknown. Although coronaviruses are frequently implicated in disease in mammals, we do not have conclusive data to associate BoNV with illness. BoNV was not the only viral agent detected by HTS (Table 1). Furthermore, samples from healthy cows from the same feedlot were not available for analysis.

In summary, we propose that BoNV may be a descendant of the ancestral nidovirus lineage, one that acquired ExoN, facilitating the expansion of its genome beyond the size of 'small' nidoviruses. Based on their sequence similarity, BoNV and the python-associated viruses may have shared a common ancestor and probably split prior to a major expansion of members of the subfamily the Torovirinae.

Acknowledgements

This work was supported by grants from the National Institutes of Health (IU19AI109761; Center for Research in Diagnostics and Discovery) and by Zoetis Inc.

References


