Short communication

Molecular characterization of a canine respiratory coronavirus strain detected in Italy

Alessio Lorusso, Costantina Desario, Viviana Mari, Marco Campolo, Eleonora Lorusso, Gabriella Elia, Vito Martella, Canio Buonavoglia, Nicola Decaro

Department of Public Health and Animal Sciences, Faculty of Veterinary Medicine of Bari, Strada per Casamassima km 3, 70010 Valenzano (Bari), Italy

Corresponding author. Tel.: +39 0804679832; fax: +39 0804679843.
E-mail address: n.decaro@vet.uniba.it (N. Decaro).

Coronaviruses (CoVs) are positive-stranded, non-segmented RNA viruses generally responsible for the emergence of respiratory and enteric disease in humans, companion animals and livestock. Their aptitude to evolve by genetic recombination and/or point mutation is recognized, thus giving rise to new viral genotypes and mutants with different tissues or host tropism. In particular, a probable origin from the strictly related bovine coronavirus (BCoV) or, alternatively, from a common ancestor has been suggested for some group 2a CoVs, including canine respiratory coronavirus (CRCoV). In this study, we report the sequence analysis of the viral RNA 3'-end of an Italian CRCoV, strain 240/05, together with the sequence comparison with extant bovine-like viruses including the sole CRCoV strain 4182 previously described. Interestingly, although the structural proteins show the same features of CRCoV 4182, the genomic region between the spike and the envelope protein genes of CRCoV 240/05 encodes for three distinct products, including the equivalent bovine 4.9 kDa non-structural protein and a truncated form of the 4.8 kDa protein, whereas CRCoV 4182 has a unique 8.8 kDa protein.

© 2008 Elsevier B.V. All rights reserved.
using the QIAamp® RNeasy Mini Kit (Qiagen S.p.A.), following the RNAlater RNA Stabilization Reagent (Qiagen S.p.A., Milan, Italy) by extracted from a fragment of the original lung sample stored in fidelity DNA polymerase (Platinum® Taq Hi Fi). The RNA was (Life Technologies, Invitrogen, Milan, Italy) which contains a high-

Schematic comparison of the CRCoV 240/05 and 4182 genomes. Below the diagram, the length in amino acids is reported for the encoded structural and the putative non-structural proteins.

A 9.6-kb region encompassing the complete 3′-end of the viral genome of CRCoV 240/05 was determined through RT-PCR amplifications of overlapping fragments by using primers previously published (Decaro et al., 2008; Decaro and Buonavoglia, 2008) and the kit SuperScriptTM One-Step RT-PCR for Long Templates (Life Technologies, Invitrogen, Milan, Italy) which contains a high-fidelity DNA polymerase (Platinum® Taq Hi Fi). The RNA was extracted from a fragment of the original lung sample stored in RNAlater RNA Stabilization Reagent (Qiagen S.p.A., Milan, Italy) by using the QIAamp® RNeasy Mini Kit (Qiagen S.p.A.), following the manufacturer’s instructions, with the template RNA being eluted in 50 µl of RNase-free water. DNA samples generated from two different RT-PCR runs were sequenced in both directions by Cogenics Europe (Meylan, France). Additional RT-PCR assays and sequencing attempts were performed to close gaps between assembled contigs using strain-specific primers. The analysis of the 240/05 3′-end genomic sequence by means of the NCBI graphical analysis ORF Finder tool (http://www.ncbi.nlm.nih.gov/gorf/gorf.html) identified nine ORFs that can be deduced to encode the non-structural and structural proteins of the virus (Fig. 1). All the predicted ORFs but the 2.7 kDa gene were preceded by a TRS sequence (CUAAAC or CCAAAAC) which is assumed to interact with the viral polymerase during the discontinuous transcription of the negative sg RNAs (for a review, see Britton and Cavanagh, 2008). The ORFs encoding structural proteins are ORF 2b (nt 1017–2291), ORF 3 (nt 2306–6397), ORF 5 (nt 7087–7341), ORF 6 (nt 7356–8048), ORF 7 (nt 8058–9404). The predicted translation products are the HE protein (47.7 kDa), S (151.1 kDa), E (9.2 kDa), M (26.4 kDa) and N (49.3 kDa) proteins, respectively. Analogously to other bovine-like CoVs, the group 2a-specific accessory structural HE protein was 424 amino acids long and identical, in terms of amino acid sequence, to that of CRCoV 4182. Only one synonymous nucleotide change differentiated the E protein encoding gene of strains 240/05 and 4182. Two residues were found to be different from the same protein of BCoV-Mebus, only one if compared with the extant BCoVs. The TRS sequence was located 123 nucleotides upstream the AUG codon as already shown in 4182 strain. SignalP assigned a signal anchor at position 1–34 and no glycosylation sites were found.

The E protein was 84 amino acids long and identical, in terms of amino acid sequence, to that of CRCoV 4182. The HE protein gene displayed an accompanying inter-

Fig. 1. Schematic comparison of the CRCoV 240/05 and 4182 genomes. Below the diagram, the length in amino acids is reported for the encoded structural and the putative non-structural proteins.
two changes were shared with extant BCoV strains. Moreover, four potential O-glycosylation sites (http://www.cbs.dtu.dk/services/NetOGlyc/) and one N-glycosylation site were identified at the N-terminus of the protein as previously described for all bovine-like CoVs.

The N protein of strain 240/05 had a length of 448 amino acids and was highly conserved among the bovine and bovine-like viruses. Six synonymous and two non-synonymous changes were present with respect to strain 4182. The same amino acid changes were present in extant BCoV-like viruses.

The putative 32 kDa nsp showed the same length (278 amino acids) as most other bovine-like CoVs, including GiCoV, ApCoV and sable antelope coronavirus (SACoV), as well as HCoV-OC43 and CRCoV 4182. Three amino acid changes were found in comparison with 4182 strain, of which one was unique to 240/05 whereas the remaining two substitutions were shared with other bovine-like strains. Five were the synonymous nucleotide substitutions. The corresponding TRS sequence was located seven nucleotides upstream of the AUG codon confirming the genome organization of CRCoV 4182.

In the majority of their genome, CRCoV 240/05 and 4182 possessed the same genomic arrangement, but they displayed a diverse genome structure in the accessory genes encoding for nsp located between the S and E genes. In this region, most BCoV-like CoVs display three accessory genes, namely the 4.9 kDa, 4.8 kDa and 12.8 kDa protein genes. In CRCoV 4182 only two accessory genes were detected (Erles et al., 2007), whereas the genomic arrangement of other bovine-like CoVs was partially conserved in strain 240/05. Indeed, CRCoV 4182 showed a unique 8.8 kDa protein fashioned by the fusion of the 4.9 kDa protein and a truncated form of the 4.8 kDa protein. This was the result of a two nucleotide deletion that inactivated the stop codon of the 4.9 kDa protein encoding sequence and introduced additional 12 amino acids not found in any BCoV strains (Erles et al., 2007). In contrast, in the 240/05 genome the terminating codon of the 4.9 kDa encoding sequence is conserved exactly as in most group 2a CoVs, including BCoVs–Mebus, Quebec and DB2, HECV-4408 and BuCoV. The corresponding TRS sequence is located 317 nucleotides upstream of the initiation codon according to previous observations for other ruminant CoVs. However, the 4.8 kDa protein of most BCoV and bovine-like CoVs is replaced in strain 240/05 by a 2.7 kDa protein (25 amino acids in length) due to the presence of an early stop codon. An identical truncated protein has been already described for CRCoV strain G9142 (Erles et al., 2007), whereas a slightly longer truncated protein is present in GiCoV and SACoV (Hasoksuz et al., 1999). The corresponding TRS sequence was not detected in any bovine-like viruses with the exception of CRCoV G9142 (Erles et al., 2007). The 12.8 kDa protein was 109 amino acids in length according to the extant BCoV and bovine-like CoVs except BCoV Quebec that exhibits a truncated form of the protein. The related TRS was located 73 nucleotides upstream of the AUG codon. Two synonymous and three non-synonymous nucleotide substitutions with respect to strain 4182 were detected in the coding sequence, with one change being unique if compared to other bovine-like CoVs previously described. The I protein was 207 amino acids in length according to CRCoV 4182. Only one amino acid change and two synonymous nucleotide variations differentiated the two CRCoV strains and strain 240/05 showed almost 95% average amino acid identity with most of the bovine-derivative viruses.

The rooted phylogenetic analysis (Fig. 2) performed on the S(a) and M/N (b) proteins confirmed the high genomic relatedness of CRCoV with BCoV and its derivatives, thus supporting data obtained previously (Decaro et al., 2007, 2008; Erles et al., 2007). Noteworthy, in the S and M/N phylogenetic trees, CRCoV 240/05 and 4182 formed a separate bunch among the group 2a CoVs, most likely as a consequence of the speciation of BCoV in the canine host.

This study provides useful data for the molecular comprehension of bovine-like CoVs in the canine host. The CoV ecology is intricate due to the number of apparently frequent cross-species jumps that entail its evolution. Indeed, the presence of bovine-like CoVs in dogs was clearly demonstrated in previous studies by serological methods, sequence comparison and phylogenetic analysis, but molecular data available to date is incomplete and limited in number. Sequence analysis of the 3′-end of the viral genome showed that strain 240/05 has a genomic organization similar to BCoV, including the presence of the ORFs encoding for the nsp between the S and E protein genes. Although the structural genome are highly conserved among CRCoV strains 240/05 and 4182, the location and number of the accessory genes differ between the two viruses. Moreover, their function in the viral life cycle has not been established yet. In general, as already demonstrated for other CoVs, accessory genes are not essential for replication and their expression could even decrease viral fitness in vitro (Schwarz et al., 1990; Yokomori et al., 1991; Curtis et al., 2002; Ortego et al., 2002; Youn et
Fig. 2. Rooted neighbor-joining tree inferred from multiple amino acid sequence alignment of the S (a) and M/N (b) protein, illustrating the relationship of CRCoV in the group 2a. For the analysis, CCov-II CB/05 (DQ112226) served as outgroup and the following CoVs strain were used: BCoV-Mebus (U00735), Quebec (AF220295), DB 2 (DQ811784), ENT (AF391541), LUN (AF391542), E-AH65 (EF424615), R-AH65 (EF424617), E-AH65-TC (EF424616), E-AH65-TC (EF424618), R-AH65 (EF424619), R-AH65 (EF424620); GiCoV (EF424623); ApCoV (DQ915164); SACoV (EF424621); CRCoV-4182 (DQ882406); HCoV-OC43 (NC_005147); PHEV-VW572 (DQ901585); MHV-A59 (AY700211); SDAV (AF20755); HCoV-HKU1 (NC_006577); BuCoV-179/07-11 (EU019216). A statistical support was provided by bootstrapping over 1000 replicates and bootstrap values >70 are indicated at the correspondent node. The scale bars indicate the estimated numbers of amino acid substitutions per site.

al., 2005; Yount et al., 2005). Nevertheless, in field conditions those genes are constantly maintained (Herrewegh et al., 1995; Smits et al., 2005; Dijkman et al., 2006) and their loss is often accompanied by the decline of virulence in the natural host (de Haan et al., 2002; Ortego et al., 2003; Hajjema et al., 2004). CRCoV 240/05 was directly sequenced from the lung sample of a dog allowing the molecular characterization of viral RNA coming straight from field conditions with no adaptation to cell culture. Apparently, in natural conditions the 4.8 kDa protein, strictly maintained in other bovine-like CoVs, was truncated in CRCoV 240/05, and this may be potentially associated to the cross-species transmission and subsequent adaptation of the ancestor BCoV to a different host. However, directly downstream of the S protein gene, CRCoV 4182 possessed a unique 8.8 kDa protein gene, whereas CRCoV G9142 displayed the canonical set of BCoV accessory genes but with the equal truncated form of the 4.8 kDa protein gene as in strain 240/05. Noteworthy, the identical truncated 4.8 kDa terminating codon is present in 240/05, G9142 and in the unique 8.8 kDa corresponding nucleotide sequence of CRCoV strain 4182 (Erles et al., 2007). Accordingly, more than one BCoV strain or ancestor virus was likely involved in the origin of CRCoV, thus leading to the emergence of different canine strains with a different organizations of the accessory genes. It has been hypothesized that the 4.9 kDa and 4.8 kDa proteins of BCoV may have arisen through mutation from a bovine 11 kDa protein (Abraham et al., 1990). According to this scenario, CRCoV 4182 may descent from a mutation of an ancestral bovine-like strain that exhibited the full-length 11 kDa protein, whereas CRCoV 240/05 and G9142 presumably descended from a different ancestor that showed the two distinct non-structural proteins. Furthermore, nsp 4.9 and 4.8 are not present in the bovine-derivative HCoV-OC43 (Vijgen et al., 2005) and their function is yet unknown in BCoV itself where the 4.9 kDa protein could not been expressed due to
the absence of a start codon in its mRNA (Hofmann et al., 1993). Obviously, it cannot be ruled out that the elevated level of genomic differences among the accessory genes is due to the high frequency of mutations/deletions that occur during CoV evolution, rather than to RNA recombination events that characterize CoV ecology (Lai et al., 1985; Makino et al., 1986; Banner and Lai, 1991). Accordingly, all CRCoV strains analyzed so far constantly cluster at the phylogenetic level with extant BCoV strains (Fig. 2), reinforcing previous suggestions that CRCoV apparently originated as a host variant of BCoV or both viruses descended from a common ancestor. Nevertheless, the elevated level of amino acid similarity with extant bovine-like CoVs characterizing the 12.8 kDa protein suggests a more important role of this protein in viral pathobiology (Erles et al., 2007). Unfortunately, CRCoV 240/05 could not be propagated in tissue culture probably as a consequence of the long storage of the tissue sample, thus preventing further analysis of viral mRNAs in the context of the infected cells. However, several studies are warranted in order to investigate the effective functionality of the accessory genes in BCoV and CRCoV, thus elucidating the relationship and variations in their expression and the interconnected evolution of highly similar viruses in different hosts.

Acknowledgements

This work was supported by grants from the Italian Ministry of Health, Ricerca finalizzata 2008, project "Mammalian coronaviruses: molecular epidemiology, vaccine development and implications for animal and human health". The authors are grateful to P.J. Collins (CT, Department of Biology, Cork, Ireland) for the English revision of the manuscript.

References


