Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Molecular characterization of a canine respiratory coronavirus strain detected in Italy

Alessio Lorusso, Costentina 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.
between the S and E protein genes that are also expressed from sg mRNAs (Narayanan et al., 2008). These sg mRNAs each contain a short leader sequence at their 5′-end, followed by a transcription regulatory sequence (TRS). TRS sequences are also present in the genome upstream of each gene and they serve as signals for the generation of the sg mRNAs (for a review, see Britton and Cavanagh, 2008). Despite the serological evidence that BCoV-like particles are circulating in dogs, few partial genomic sequences of CRCoVs are publicly available, mostly regarding the S and HE proteins, except the English isolate CRCoV 4182 whose entire genomic 3′-end has been determined and analyzed (Erles et al., 2007).

In the present study, we report the sequence and phylogenetic analyses of the entire genomic 3′-end of the CRCoV 240/05 strain previously detected in the lungs of a dog in Italy (Decaro et al., 2007). 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 SuperScript™ 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 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 (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 showed 2.7 kDa gene were preceded by a TRS sequence (CUAAAC) located nine nucleotides upstream of the AUG codon, indicating the presence of a putative 5′ non-structural RNA 5′UTR. The AUG codon was preceded by the TRS sequence. Amino acid identity between the S protein of CRCoV 240/05 and BCoV/bovine-like viruses S protein ranged from 91.2% to 95.7% (Table 1). A potential N-terminal signal peptide was identified at amino acids 1–17 and 21 potential glycosylation sites were detected throughout the protein. No deletions or insertions were observed in strain 240/05 compared with bovine-like CoVs. Moreover, the amino acid stretch KRRSRR, responsible for the proteolytic cleavage of the S protein at residue 768 into subunits S1 and S2, was conserved.

The S protein was long 1363 amino acids and its AUG codon was directly preceded by the TRS sequence. Amino acid identity in comparison with the S protein of CRCoV 4182 was 98.4% (Table 1) showing 22 amino acid substitutions, of which nine were found to be unique to 240/05 strain whereas 10 were present in other BCoV and derivatives. Twenty-one synonymous nucleotide variations were encountered in the S gene. Amino acid identity between 240/05 and BCoV/bovine-like viruses S protein ranged from 91.2% to 95.7% (Table 1). A potential N-terminal signal peptide was identified at amino acids 1–17 and 21 potential glycosylation sites were detected throughout the protein. No deletions or insertions were observed in strain 240/05 compared with bovine-like CoVs. Moreover, the amino acid stretch KRRSRR, responsible for the proteolytic cleavage of the S protein at residue 768 into subunits S1 and S2, was conserved.

The E protein was 84 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 of 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 M protein was 230 amino acids in length and the AUG translation initiation codon was separated from the TRS stretch by three nucleotides. Five synonymous and five non-synonymous nucleotide variations were found with respect to the M gene of CRCoV 4182, three of which were unique to 240/05 strain while the remaining

![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.](image-url)
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. The 32 kDa HE S 4.9 kDa 2.7 kDa a 12.8 kDa E M N I.
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 (DQ082406); HCoV-OC43 (NC_005147); PHEV-VW572 (DQ011855); MHV-A59 (AY780211); SDAV (AF207551); 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; Haijema 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 descend 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 or both viruses descended from a common ancestor. Nevertheless, 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 (CIT, Department of Biology, Cork, Ireland) for the English revision of the manuscript.

References

Abraham, S., Kienzle, T.E., Lapps, W.E., Brian, D.A., 1990. Sequence and expression analysis of potential nonstructural proteins of 4.9, 4.8, 12.7, and 9.5 kDa encoded between the spike and membrane protein genes of the bovine coronavirus. Virology 177, 488–495.

Banner, L.R., Lai, M.M., 1991. Random nature of coronavirus RNA recombination in the absence of selection pressure. Virology 185, 441–445.

Britton, P., Cavanagh, D., 2008. Nidovirus genome organization and expression mechanisms. In: Perlman, S., Gallagher, T., Snijder, E.J. (Eds.), Nidoviruses. ASM Press, Washington, DC, pp. 29–46.

Buonavoglia, C., Martella, V., 2007. Canine respiratory viruses. Vet. Res. 38, 355–373.

Curtis, K.M., Yount, B., Baric, R.S., 2002. Heterologous gene expression from transgenic mice. Virology 308, 13–22.

Decaro, N., Disario, C., Elia, G., Mari, V., Lucente, M.S., Cordioli, P., Colaianni, M.L., Martella, V., Buonavoglia, C., 2007. Serological and molecular evidence that canine respiratory coronavirus is circulating in Italy. Vet. Microbiol. 121, 225–230.

Decaro, N., Martella, V., Elia, G., Campolo, M., Mari, V., Disario, C., Lucente, M.S., Lorusso, A., Greco, G., Corrente, M., Tempesta, M., Buonavoglia, C., 2008. Biological and genetic analysis of a bovine-like coronavirus isolated from water buffaloes (Bubalus bubalis) calves. Virology 370, 213–222.

Decaro, N., Buonavoglia, C., 2008. An update on canine coronavirus: viral evolution and pathobiology. Vet. Microbiol. 132, 221–234.

Dijkman, R., Jebbink, M.F., Wilbrink, B., Pryck, K., Zaaier, H.L., Minor, P.D., Franklin, S., Berkhourt, B., Thié, V., van der Hoek, L., 2006. Human coronavirus 229E encodes a single ORF4 protein between the spike and the envelope genes. Virol. J. 3, 106.

Ellis, J.A., McLean, N., Hupaelo, R., Haines, D.M., 2005. Detection of coronavirus in cases of tracheobronchitis in dogs: a retrospective study from 1971 to 2003. Can. Vet. J. 46, 447–448.