Research paper

Evolutionary genetics of canine respiratory coronavirus and recent introduction into Swedish dogs

Michelle Wille[a,1], Jonas Johansson Wensman[b], Simon Larsson[c,2], Renaud van Damme[d], Anna-Karin Theelke[e], Juliette Hayer[d], Maja Malmberg[c,d,*]

[a] *Zoonosis Science Center, Department of Medical Biochemistry and Microbiology, Uppsala University, 751 23 Uppsala, Sweden*
[b] *Department of Clinical Sciences, Swedish University of Agricultural Sciences, Box 7054, 750 07 Uppsala, Sweden*
[c] *Section of Virology, Department of Biomedical Sciences and Veterinary Public Health, Swedish University of Agricultural Sciences, Box 7028, 750 07 Uppsala, Sweden*
[d] *SLU Global Bioinformatics Centre, Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, Box 7023, 750 07 Uppsala, Sweden*
[e] *Department of Microbiology, National Veterinary Institute, 751 89 Uppsala, Sweden*

ABSTRACT

Canine respiratory coronavirus (CRCoV) has been identified as a causative agent of canine infectious respiratory disease, an upper respiratory infection affecting dogs. The epidemiology is currently opaque, with an unclear understanding of global prevalence, pathology, and genetic characteristics. In this study, Swedish privately-owned dogs with characteristic signs of canine infectious respiratory disease (n = 88) were screened for CRCoV and 13 positive samples (14.7%, 8.4–23.7% [95% confidence interval (CI)]) were further sequenced. Sequenced Swedish CRCoV isolates were highly similar despite being detected in dogs living in geographically distant locations and sampled across 3 years (2013–2015). This is due to a single introduction into Swedish dogs in approximately 2010, as inferred by time structured phylogeny. Unlike other CRCoVs, there was no evidence of recombination in Swedish CRCoV viruses, further supporting a single introduction. Finally, there were low levels of polymorphisms, in the spike genes. Overall, we demonstrate that there is little diversity of CRCoV which is endemic in Swedish dogs.

1. Introduction

Canine infectious respiratory disease (CIRD) complex, colloquially referred to as kennel cough, is a contagious disease in dogs, particularly prolific in rehoming centers and kennels. Dogs suffer from a dry hacking cough, which is usually cleared in 1–3 weeks, however, severe bronchopneumonia can develop (Appel, 1987). CIRD is a multifactorial disease with identified disease agents including canine parainfluenza virus (CPIV) (Appel and Percy, 1970), canine adenovirus type 2 (CAV-2) (Ditchfield et al., 1962), canine pneumovirus (Mitchell et al., 2013), and the bacteria *Bordetella bronchiseptica* (Bemis, 1992) and *Mycobacterium cynos* (Mitchell et al., 2017). More recently, canine respiratory coronavirus (CRCoV) was also identified as a causative agent

of CIRD. This virus was first identified in the UK in a rehoming center with a high incidence of CIRD (Erles et al., 2003), however, a retrospective study has suggested that it may have circulated as early as 1996 in Canada (Ellis et al., 2005). Additional surveys have identified antibodies in dogs in the UK, Ireland, Italy, Japan, and the US, with antibody prevalence as high as 87.7% in the state of Kentucky, USA (An et al., 2010b; Decaro et al., 2007; Erles and Brownlie, 2005; Knesl et al., 2009; Priestnall et al., 2006, 2007; Schulz et al., 2014). A small number of isolates have also been sequenced and characterized (An et al., 2010a; Erles et al., 2003, 2007; Jeoung et al., 2014; Yachi and Mochizuki, 2006).

Coronaviruses cause numerous diseases, including upper and lower respiratory diseases, gastroenteritis, and central nervous system

infections in a number of avian and mammalian species. In humans, coronavirus infections range from the common cold (e.g. HCoV-OC83) to more severe zoonotic diseases such as Severe Acute Respiratory Syndrome (SARS) (Peiris et al., 2003; van der Hoek et al., 2004) and Middle East Respiratory Syndrome (MERS) (Cunha and Opal, 2014; de Groot et al., 2013). Coronaviruses have large zoonotic potential, and the ability to cross species boundaries lies in not only mutation, but also the propensity for these viruses to recombine, with important breakpoints identified around the spike (S) gene in both birds and mammals (Vijaykrishna et al., 2007; Woo et al., 2009). Extensive homologous and heterologous recombination events have been documented in both human and animal group 2 coronaviruses leading to the generation of various genotypes and strains (Woo et al., 2009). CRCoV is a group 2 coronavirus, or betacoronavirus, which are comprised of mammalian coronaviruses. The most closely related species to CRCoV is bovine coronavirus (BCoV), with many closely related species or strains such as Sambar deer coronavirus, Waterbuck coronavirus and human enteric coronavirus (HEC 4408), illustrating the proliferation of these BCoV-like viruses.

In this study, we screened nasopharyngeal swabs from privately-owned dogs in Sweden with and without CIRD for CRCoV. We identified the first CRCoV positive dogs from Sweden, and we used these to assess the evolutionary genetics of CRCoV, both locally and globally. Specifically, we wanted to determine genetic variation and quasispecies of CRCoV in Swedish dogs to clarify diversity within Sweden and the relationship of Swedish viruses to other CRCoV isolates in Europe and elsewhere. Furthermore, we aimed to understand the introduction of these viruses into Sweden by using time-structured phylogeny and recombination analysis. Finally, given the large number of CRCoV isolates sequenced in this study, we were able to better elucidate the emergence of CRCoV through analysis of recombination dynamics of CRCoV and BCoV.

2. Materials and methods

2.1. Ethics statement

This research was reviewed, approved and conducted in accordance with the regulations provided by the Swedish Board of Agriculture and approved by the Swedish Animal Research Ethics Board (Uppsala djurförsöksetiska nämnd, reference numbers C227/11 and C127/14).

2.2. Sample collection

Between April 2013 and December 2015, privately owned dogs with characteristic upper respiratory signs of CIRD (dry cough) for up to 7 days were enrolled in a study investigating the cause of CIRD in Sweden. Maximum two dogs per households were sampled, and sampled dogs were not treated by antibiotics at the sampling. Samples were taken from seven veterinary clinics across Sweden, with dogs residing in 12 Swedish counties (Fig. 1). Nasopharyngeal swabs (E-swabs with Amies medium and regular nylon flocked applicator, Copan Italia, Brescia, Italy) were collected and stored at −80 °C within 24-48 h of collection. A total of 88 dogs with CIRD were swabbed. As controls, we also swabbed 20 healthy dogs that had not suffered from respiratory signs for the last six months.

2.3. Statistics

We used generalized linear models (glm, family = binomial) to evaluate the effect of age, breed, location, and sex on CRCoV prevalence. Explanatory variables were entered in isolation or in combination, where the resulting model improvements were χ^2 tested for significance. Statistics were done in R 3.5.1 (R Development Core Team, 2008) integrated into R Studio 1.0.143.
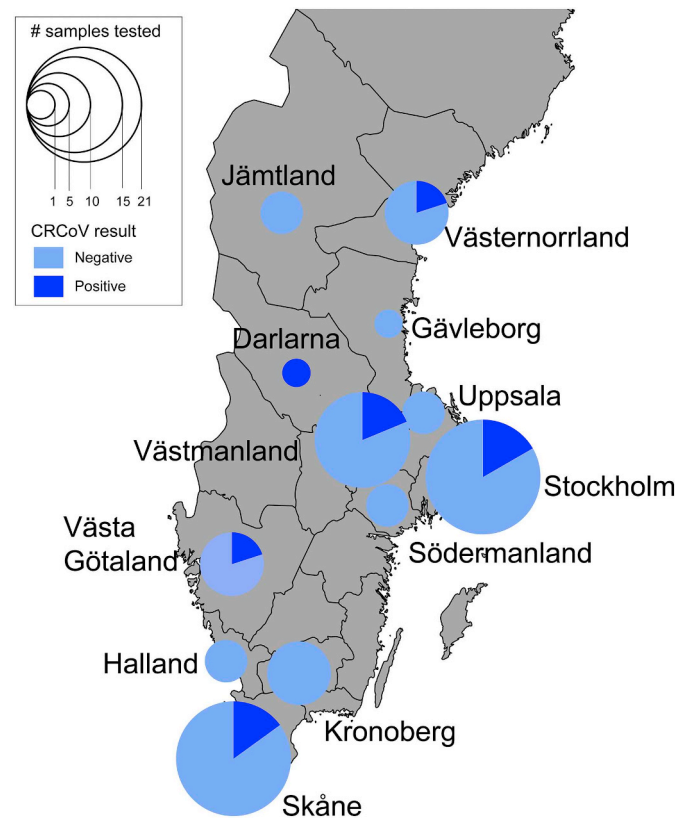

Fig. 1. Swedish counties from which samples were submitted. Size of pie chart is roughly proportional to the number of samples collected. Dark blue indicates proportion of samples positive for CRCoV, and light blue indicates the proportion of samples negative for CRCoV. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2.4. Virus sequencing

Viral RNA was extracted from nasopharyngeal swabs with the Magnatrix 8000+ extraction robot (Magnetic Biosolutions, Sweden) and Vet Viral NA kit (NorDiag ASA, Oslo, Norway), as described in (Jinnerot et al., 2015). cDNA was subsequently synthesized using SuperScript™ III Reverse Transcriptase (Invitrogen, Life Technologies, Carlsbad, CA, USA), and the second strand was synthesized using Klenow Fragment™ 3′->5′ exo- (New England Biolabs, M0212S, Ipswich, MA, USA).

Four approaches were utilized to sequence viruses. First, using traditional PCR and Sanger sequencing of the PCR products, and second, using Illumina MiSeq to sequence PCR products of partial S gene. Third, a viral metagenomics approach was done on two samples (CRCoV4 and CRCoV6) with the aim to get close to complete genomes. Lastly, a probe-based capture method was used on one sample (CRCoV1) (Table A1).

2.4.1. Sanger sequencing

For the first approach, PCR reactions targeting the S, Membrane (M) and Hemagglutinin-esterase (HE) genes were carried out using previously published primers (An et al., 2010a; Erles et al., 2007) (Table A2) using KAPA2G Robust HotStart ReadyMix PCR Kit (KAPA Biosystems, Roche Sequencing, Pleasanton, CA, USA). Thermocycling conditions were 95 °C for 3 min and then 40 cycles of 95 °C for 15 s, a reaction specific annealing temperature for 15 s, a reaction specific elongation time at 72 °C, and finally 72 °C for 1 min (Table A2). The PCR products were run on a 1.5% agarose gel stained with GelRed, visualized under UV transillumination (GelDoc, Bio-Rad Laboratories,

Inc., Richmond, CA, USA), purified using GeneJet Gel Extraction Kit (Life Technologies, Carlsbad, CA, USA) and sequenced at Macrogen Europe (Amsterdam, NL).

2.4.2. Illumina MiSeq sequencing of PCR products of partial spike gene

A 2480 bp fragment was amplified using LongAmp Taq DNA Polymeranse (2.5 U), dNTP mix (0.33 mM), 1 × LongAmp Taq Reaction buffer (New England BioLabs, M0323S, Ipswich, MA, USA), 0.4 μM of respective primer Sp1F and Sp10R (Table A2), and 3 μl of template. Thermocycling conditions were 94 °C for 30 s and then 40 cycles of 94 °C for 30 s, 64 °C for 45 s, elongation time at 65 °C for 220 s, and finally 65 °C for 10 min. The PCR products were purified using the GenJET PCR Purification kit (Thermo Fisher Scientific, Waltham, MA, USA). Sequencing libraries were made using Nextera XT Library Preparation Kit (Illumina, San Diego, CA, USA), normalization and pooling of 2 nM sequencing libraries was done based on concentration measurements from Agilent High Sensitivity DNA Kit (2100 Bioanalyzer, Agilent Technologies, Palo Alto, CA, USA). The pool of sequencing libraries was denatured with NaOH and further diluted with hybridization buffer to a final concentration of 10 pM and spiked with 5% PhiX for diversity. Paired-end sequencing was performed with MiSeq Reagent Kit v3 600 cycles on the MiSeq instrument (Illumina, San Diego, CA, USA) at the National Veterinary Institute, Uppsala, Sweden.

2.4.3. Viral metagenomics

For the viral metagenomics approach, 600 μl of the swab media was freeze-thawed thawed twice, centrifuged at 627g at 4 °C for 3 min, the supernatant transferred to a filtrate 0.65 μm (Millipore, Burlington, MA, USA), and centrifuged for 4 min at 10000g. The filtrated sample was treated with 16 U TurboDNase in TurboDNase Buffer (Invitrogen, Thermo Fisher Scientific, Waltham, MA, USA) and DNaseI (2 μl) (Invitrogen, Life Technologies, Carlsbad, CA, USA) for 30 min at 37 °C, followed by RNase cocktail™ Enzyme Mix (Invitrogen, Thermo Fisher Scientific, Waltham, MA, USA) treatment for 5 min at room temperature. Thereafter, RNA was extracted using Trizol, chloroform and RNeasy kit (Qiagen, Hilden, Germany). The kit Oviation® RNA-Seq System V2 (NuGEN, Redwood City, CA, USA) was used to amplify the RNA in accordance with the provided protocol, and the GenJET PCR Purification kit (Thermo Fisher Scientific, Waltham, MA, USA) was used for purification.

Sequencing libraries were constructed at the National Sequencing Infrastructure in Uppsala, Sweden, using the AB Library Builder System (Ion Xpress™ Plus and Ion Plus Library Preparation for the AB Library Builder™ System protocol, Thermo Fisher Scientific, Waltham, MA, USA) and size selected on the Blue PippinTM (Sage Science, Beverly, MA, USA). Library size and concentration were assessed by a Bioanalyzer High Sensitivity Chip (Agilent Technologies, Santa Clara, CA, USA) and by the Fragment Analyzer system (Advanced Analytical; Agilent Technologies, Santa Clara, CA, USA). Template preparation was performed on the Ion Chef™ System using the Ion 520 & Ion 530 Kit-Chef (Thermo Fisher Scientific, Waltham, MA, USA). Samples were sequenced on 530 chips using the Ion S5™ XL System (Thermo Fisher, Waltham, MA, USA).

2.4.4. Probe-based-capture

Probes were designed for feline coronavirus 30 kb - NC_002306.3, canine respiratory coronavirus 30 kb - JX860640.1, bovine coronavirus 30 kb - U00735.2, kobuvirus 8 kb - KM977675.1, porcine rubulavirus 15 kb - NC_009640.1, African swine fever virus 180 kb - Benin97/1 AM712239.1, canine parainfluenza virus 15 kb - KC237064.1, by Agilent Technologies SureSelect DNA Advanced Design Wizard. This resulted in a 297.965 kbp capture based on 2772 probes.

One sample that was positive for CRCoV with a Cq-value of 33.15 as determined by qPCR was selected (CRCoV1). RNA extracted from nasopharyngeal swabs as described above was used. In total 8 μl of RNA

was converted into cDNA using SuperScript™ III Reverse Transcriptase (Invitrogen, Life Technologies, Carlsbad, CA, USA). Thereafter, treated with RNAse H for 20 min at 37 °C, and made double-stranded using, Klenow Fragment™ 3′->5′ exo- (Thermo Fisher Scientific, Waltham, MA, USA).

KAPA HyperPlus Library Prep kit (KAPA Biosystems, Roche Sequencing, Pleasanton, CA, USA) was used in combination with an unofficial protocol for using Agilent's SureSelectXT Target Enrichment System for Illumina Paired-End multiplexed sequencing libraries Version B5 (June 2016), 200 ng DNA samples. A 3× bead clean-up was used on 20 μl of double-stranded cDNA prior to fragmentation. The samples were fragmented at 37 °C for 25 min. At the adapter ligation step the SureSelect Adapter Oligo Mix was used and the sample was incubated at 20 °C for 45 min. The Pre-Capture libraries were vacuum centrifuged to a final concentration of 221 ng/μl. Hybridization was done according to the SureSelect protocol and the samples were incubated at 65 °C for 24 h in a ProFlex PCR machine (Applied Biosystems, Foster City, CA, USA). Thereafter, Dynabeads MyOne Streptavidin T1 (Thermo Fisher Scientific, Waltham, MA, USA) were used to capture the SureSelectXT enriched libraries. An on-bead PCR with KAPA HiFi HotStart amplified the libraries. For this, 16 PCR cycles were used. After a final clean up with Agencourt AMPure beads (Beckman Coulter Indianapolis, IN, USA), the libraries were quality assured using Bioanalyzer HS DNA assay (Agilent Technologies, Santa Clara, CA, USA). Prior to sequencing a 4 nM pool was made, denatured with NaOH and further diluted with hybridization buffer to a final concentration of 8 pM and spiked with 2.4% PhiX for diversity. Paired-end, 151 cycles sequencing was performed with MiSeq Reagent Kit v2 300 cycles on the MiSeq instrument (Illumina, San Diego, CA, USA) at the National Veterinary Institute in Uppsala, Sweden.

2.5. Bioinformatics

2.5.1. Sanger sequencing

Reads were cleaned, forward and reverse reads aligned and a consensus per sample was made using DNASTAR (Madison, WI, USA).

2.5.2. Illumina MiSeq sequencing of PCR products of partial spike gene

The sequenced reads were demultiplexed using bcl2fastq (https://github.com/brwnj/bcl2fastq.) and the adapters were removed using fastp version 0.19.5 (Chen et al., 2018) (https://github.com/OpenGene/fastp). The reads were then assembled using MEGAHIT version 1.1.4 (Li et al., 2016), with default settings. For each sample, the only contig produced corresponding to the size of the amplicon was extracted (longest contig). The reads were trimmed for quality using *fastp* with the a PHRED score threshold of 30 and a minimum length of 30 bp, and aligned to the longest contig obtained by MEGAHIT using BWA-MEM version 0.7.12 (Li and Durbin, 2009) with default parameters. The result of the alignment was converted and sorted in a bam file using SAMtools version 1.3.1 (Li et al., 2009). From these alignments, Single Nucleotide Variants were then analyzed using both ShoRAH version 1.0 (Zagordi et al., 2011) and QuasiRecomb version 2.1 (Topfer et al., 2013).

2.5.3. High-throughput sequencing of viral metagenomics prepared samples

The reads were assembled using MEGAHIT version 1.1.4 (Li et al., 2016), with default settings. A taxonomic classification of the contigs using Diamond version 0.9.24 (Buchfink et al., 2015) against the non-redundant protein database from NCBI (*nr*, release February 2019) was then performed. The produced output files (daa format) were uploaded into MEGAN (version 6.12.6) (Huson et al., 2007) and all the contigs classified as Betacoronavirus 1 were extracted and inspected. The longest contigs were selected for further structural and functional annotation.

2.5.4. Probe-based-capture

Reads were assembled using SPAdes (v 3.7.0) (Bankevich et al., 2012) but as the coverage was too high to get a correct assembly, the dataset was randomly reduced down to 75,000 reads. Those reads were then assembled using SPAdes. The obtained contigs were then taxonomically assigned using Diamond. The outputs from Diamond were visualized in MEGAN and contigs classified as Betacoronavirus1 were retrieved.

2.5.5. Annotation

The CRCoV1, CRCoV4 and CRCoV6 genomes were annotated using an annotation pipeline for prokaryotic and viral genomes, PROKKA (Seemann, 2014). We had previously extracted all the protein sequences belonging to the Coronaviridae family from UniProtKB (release April 2019). This dataset was provided to PROKKA for annotating the newly assembled coronavirus genomes.

2.6. Sequence availability

All generated sequences from this study have been deposited to the European Nucleotide Archive at EBI, under the BioProject PRJEB34079. High Throughput Sequencing (HTS) reads have been deposited in the EBI Short Read Archive (accession numbers: spike amplicons datasets: ERR3489845-ERR3489847, ERR3489112-ERR3489114, ERR3489104-ERR3489106. CRCoV1 probe-based capture dataset: ERR3486809, metagenomics datasets: CRCoV4: ERR3489103 and CRCoV6: ERR3489111). Final full and partial annotated genomes generated from HTS have been deposited in European Nucleotide Archive under the accession numbers: ERZ1079470 (CRCoV1), ERZ1079468 (CRCoV4), ERZ1079469 (CRCoV6). Full length genes generated through sanger sequencing have been deposited in European Nucleotide Archive, Accession numbers ERZ1080291 to ERZ1080310.

2.7. Virus genetic analysis

Resulting sequences were aligned using the MAFFT algorithm (Katoh et al., 2009) within Geneious R11 (Biomatters, New Zealand). Maximum likelihood phylogenetic trees were constructed for each gene (ORF1ab, hemagglutinin-esterase HE, S, and matrix M) using PhyML 3.0 (Guindon et al., 2010) implementing the best substitution model for each gene.

For the full length S gene, we utilized BEAST 1.8 (Drummond and Rambaut, 2007) to better infer the evolutionary relationship within CRCoVs. Shortly, we used Maximum Likelihood trees constructed in MEGA 6 to test for clock-like behavior in each data set by performing linear regressions of root-to-tip distances across years of sampling in TempEst (Rambaut et al., 2016). Using BEAST, time-stamped data were analyzed using both the uncorrelated lognormal relaxed and strict

molecular clock, the SRD06 codon position model - a HKY85 substitution model and a different rate of nucleotide substitution for the 1 + 2 codon position and the 3rd codon position (Bahl et al., 2013; Shapiro et al., 2006). We implemented the Bayesian Skyline coalescent tree prior. Three independent analyses of 50 million generations were performed and convergence assessed using Tracer 1.6. Independent runs were combined in LogCombiner v1.8 following a burnin of 10%. Maximum credibility clade trees were generated using TreeAnnotator v1.8. The maximum credibility clade trees were visualized in FigTree v1.4.3.

To assess recombination, a concatenated sequence was generated including the viruses from which there were sequences from the partial Non-structural protein 2a (NS), and full length HE, S, envelope (E), and M, resulting in 6541 bp for analysis and the genes were placed in genomic order. To assess evolutionary patterns of all segments a split-stree network was constructed with CRCoV as well as BCoV outgroups using SplitsTree4 (Huson and Bryant, 2006). Splitstree builds a network which takes recombination into account. To better understand the recombination process, the concatenated alignment was used in RPD4 to estimate break points (Martin et al., 2010). Specifically, we used the algorithms RDP (Martin and Rybicki, 2000) and BOOTSCAN (Martin et al., 2005) to detect the recombination window, and used additional algorithms within RPD4 to cross reference support for the detected window.

To understand genetic variation within each sample, we utilized sequences of the partial spike gene generated using Illumina MiSeq of nine samples. The amplicon sequenced ranged from position 1630 to the 3′ end (4092) of the spike gene, so variants were only studied in this 3′ region. The genetic variant population was estimated using a combination of both ShoRAH version 1.0 (Zagordi et al., 2011) and QuasiRecomb version 2.1 (Topfer et al., 2013) with default parameters. We used a combination of both to ensure the accuracy of the SNPs found.

3. Results

3.1. Patterns of detection of CRCoV in Swedish dogs

A total of 108 dogs were sampled as part of this study, comprising 88 dogs with signs of disease and 20 healthy dogs. Thirteen samples collected from dogs with signs of disease were positive for CRCoV (14.7%, 8.4–23.7% [95% confidence interval (CI)]), and were collected from dogs living in the counties of Stockholm (4/22), Skåne (3/17), Västmanland (3/14), Västernorrland (2/8) and Dalarna (1/1) (Fig. 1). None of the healthy dogs tested positive for CRCoV. Overall prevalence of CRCoV did not vary significantly by dog clinic (X^2 = 11.95, df = 7, p = 0.1022) or county from which dogs originated (X^2 = 13.194, df = 11, p = 0.2809). There was further no statistical difference across dog breed (X^2 = 3.037, df = 1, p = 0.08137), sex (X^2 = 3.8, df = 1,

Table 1
Canine Respiratory Coronavirus (CRCoV) positive samples included in this study.

Sample number	Swedish county	Date sample received	Breed	Age (years)	Disease status[a]	Duration of disease (days)
CRCoV 1	Skåne	12 April 2013	Border collie	1.25	Mild	1
CRCoV 2	Skåne	25 October 2013	Mixed breed	2.25	Mild	1
CRCoV 3	Stockholm	4 December 2013	Boxer	0.8	Severe	4
CRCoV 4	Stockholm	30 December 2013	Dobermann	0.7	Severe	2
CRCoV 5	Stockholm	30 January 2014	English Bulldog	0.5	Mild	No information
CRCoV 6	Dalarna	30 January 2014	Mixed breed	0.33	Mild	1
CRCoV 7	Västmanland	11 February 2014	Mixed breed	4.9	Mild	4
CRCoV 8	Västmanland	20 February 2014	Golden retriever	0.6	Mild	3
CRCoV 9	Västmanland	20 February 2014	Mixed breed	5.4	Mild	1
CRCoV 10	Skåne	7 March 2014	West highland white terrier	12.5	Mild	1.5
CRCoV 11	Västenorrland	19 March 2014	Mixed breed	0.8	Mild	1
CRCoV 12	Stockholm	19 December 2014	Mixed breed	1.1	Mild	2.5
CRCoV 13	Västernorrland	19 March 2015	Danish–Swedish Farm dog	7	Mild	0.5

[a] Dogs with mild disease signs displayed mild intermittent coughing only, and dogs with severe disease signs had affected general condition and deep coughing.

p = 0.0504), or age category (X^2 = 6.778, df = 3, p = 0.07932) (Fig. A1). The largest number of positive samples were detected in winter months (n = 8) compared to the autumn (n = 1) and spring (n = 5); no positive samples were detected in the summer (Table 1). Despite this detection difference, prevalence of CRCoV was not significantly different across season (X^2 = 6.83, df = 3, p = 0.075) (Fig. A1).

3.2. Full and partial genomes of CRCoV

Using probe-based capture one complete genome was generated (CRCoV1). The complete genome of CRCoV1 was 31,190 bp and following annotation 9 coding regions were predicted: Orf1a, Orf1ab, Non-structural protein 2a, HE, S, 12.8 kDa non-structural protein, E, M and Nucleoprotein (NP) (Fig. A2). This full genome now brings the number of complete CRCoV genomes to 3 [CRCoV-K37 (An et al., 2010a) and CRCoV-BJ232 (Lu et al., 2017)]. Viral metagenomics resulted in partial genomes of CRCoV4 and CRCoV6 covering 7933 bp and 16,085 bp, respectively. Following annotation of the CRCoV4 genome, 6 coding regions were predicted: a partial gene coding for the HE, and all full genes coding for: S, 12.8 kDa non-structural protein, E, M and the NP. Annotation of the CRCoV6 genome predicted a partial ORF1ab gene and all the remaining genes located between ORF1ab and the 3′ end of the genome. Of the remaining viruses, 1–3 genes were sequenced through a combination of Sanger sequencing and Illumina high-throughput sequencing (Table A1).

3.3. Swedish CRCoV are the result of a single introduction

Assessment of the S, E, M, NS2 and HE genes generated using Sanger and high-throughput sequencing (Table A1) demonstrated high genetic similarity of viruses detected in Swedish dogs, despite sampling from numerous clinics across Sweden and an over a period of 3 years. Indeed, all samples for all genes were within 99% identity (Fig. A3). Phylogenetic reconstruction of the larger genes (S, HE and M) indicates that Swedish viruses are most closely related to a CRCoV isolated in Italy in 2005 (EU999954), and in some trees CRCoV isolated in the UK in 2003 (DQ682406) (Fig. A4). Given discordance between trees, we constructed a consensus tree from a concatenated alignment using Splitstree, suggesting that all Swedish viruses are likely the result of a single introduction and despite recombination being a common feature in CRCoV's, are most similar to the two aforementioned viruses (Figs. 4, A4). Time-structured phylogenetic analysis of the S gene suggests that the most recent common ancestor of all Swedish CRCoVs was introduced to Sweden at the beginning of 2010 [2010.152 with a confidence interval ranging from 2004.035 to 2012.697], with subsequence proliferation within Sweden (Fig. 2). As suggested with previous analysis, the viruses isolated in Sweden are most closely related to an isolate from Italy, but not necessarily the UK, and that CRCoVs found in Europe may have emerged and been introduced from Asia. However, due to the sparseness of sequences, we have not undertaken phylogeographic analysis, and these observations should be interpreted with caution.

3.4. Recombination between CRCoV and BCoV

Largely, CRCoV and BCoV form independent lineages in phylogenetic analysis, and large number of Swedish sequences in each tree help to better define the CRCoV clade. This lineage separation is most evident in the S tree (Fig. A4C), where all CRCoVs form a clade that is different from BCoV. However, CRCoV does not always fall into a distinctive clade; some genes like the HE and M of some viruses are more closely related to BCoV sequences than CRCoV. This was not the case with the viruses from Sweden, which are very similar to each other, and more distantly related to BCoV than other CRCoV sequences (Fig. S2A, Fig. 3). When assessing pairwise identity, it is clear that the percentage identity of all global CRCoVs only (Fig. A3B) and a combination of all

CRCoV and BCoV sequences (Fig. A3C) for the partial NS, HE and the complete M gene is rather similar. This is in contrast to the S gene, where percentage identity of all CRCoVs ranges from 99 to 100% whereas when including BCoV outgroups the range is from 97.5 to 99%. This is because for many genes, viruses from Korea (K9, K37, K39) and a recent sequence from China which were mined from GenBank are more similar to BCoV than to CRCoVs detected in Europe (Fig. 4A). Recombination detection algorithms detect that there is indeed a recombination event in these previously described Korean and Chinese viruses, where the NS, HE, E and M sequence is more similar to BCoV than to other CRCoV sequences (Region probability with MC correction: 0.002691556) (Table A3). The breakpoint, represented by the end of the HE gene and start of the S gene, has a high probability (0.008164); however, the certainty of the breakpoint at the end of the S gene is undetermined due to a number of potential breakpoints (Fig. A5), which is revealed in the BOOTSCAN analysis. This is because all CRCoVs, but especially the Asian CRCoVs, have higher signal for BCoV at the end of the S gene (Fig. 4B). Whether this is due to recombination within the S gene, or due to evolutionary processes of conservation in this section of the S gene, is unclear given some similarity between Swedish CRCoV and BCoV in this region as well. Regardless, these recombinant isolates have been maintained in Asia, demonstrated by the isolation of a virus in China in 2014 sharing the same mosaic CRCoV-BCoV pattern as viruses isolated in Korea in 2008. Overall, viruses sequenced in Sweden as part of this study have a different recombination profile as compared to previously described viruses from China and Korea.

3.5. Single nucleotide variants analysis

To better understand the variation within each CRCoV virus detected in Sweden, partial spike amplicon high-throughput sequencing of all CRCoV isolates except CRCoV1, 4 and 8 were analyzed for variations with 2 different methods: ShoRAH and QuasiRecomb. Overall, ShoRAH identified more Single Nucleotide Variants (SNVs) compared to QuasiRecomb (Fig. 5), however approximately half of the SNVs predicted by ShoRAH were also predicted by QuasiRecomb (except for the insertions/deletions that are only predicted by ShoRAH). Furthermore, regardless of method, most of the SNVs observed were at a frequency lower than 0.1. There was variation in SNV's across samples. There were no variants identified in CRCoV5 with either method. In contrast, CRCoV10 contained 17 different SNV's. CRCoV7 and CRCoV12 had marked different pattern depending on method used, whereby SNV's were identified with ShoRAH but not with QuasiRecomb due to differences in detection of indels. There are no positions of SNVs that are common to all samples. Only one SNV was common to CRCoV3 and CRCoV9 (position 2807). This variation analysis suggests that the studied region of the S gene is not variable among the CRCoV isolates collected from Swedish dogs.

4. Discussion

In this study, we aimed to reveal the epidemiology and evolutionary genetics of CRCoV, one of the causative agents of canine infectious respiratory disease (CIRD), in Sweden. CIRD is a major cause of morbidity in dogs and an important animal welfare condition. This disease is multi-factorial, and a recent European survey reported that CRCoV, canine pneumovirus and the bacteria *Mycoplasma cynos*, besides the previously established main causes (mainly CPIV and *B. bronchiseptica*), all played a role in the disease (Mitchell et al., 2017). To date, the highest prevalence of CRCoV remains in rehoming centers (Erles and Brownlie, 2005; Erles et al., 2003; Mitchell et al., 2017), or other instances where dogs may be cohoused, such as in racing greyhounds (Sowman et al., 2018). Epidemiology is unclear in pet dogs, such as those sampled in our study, but we predict that occasions where many dogs meet (e.g. dog day-care centers, dog shows and other activities),
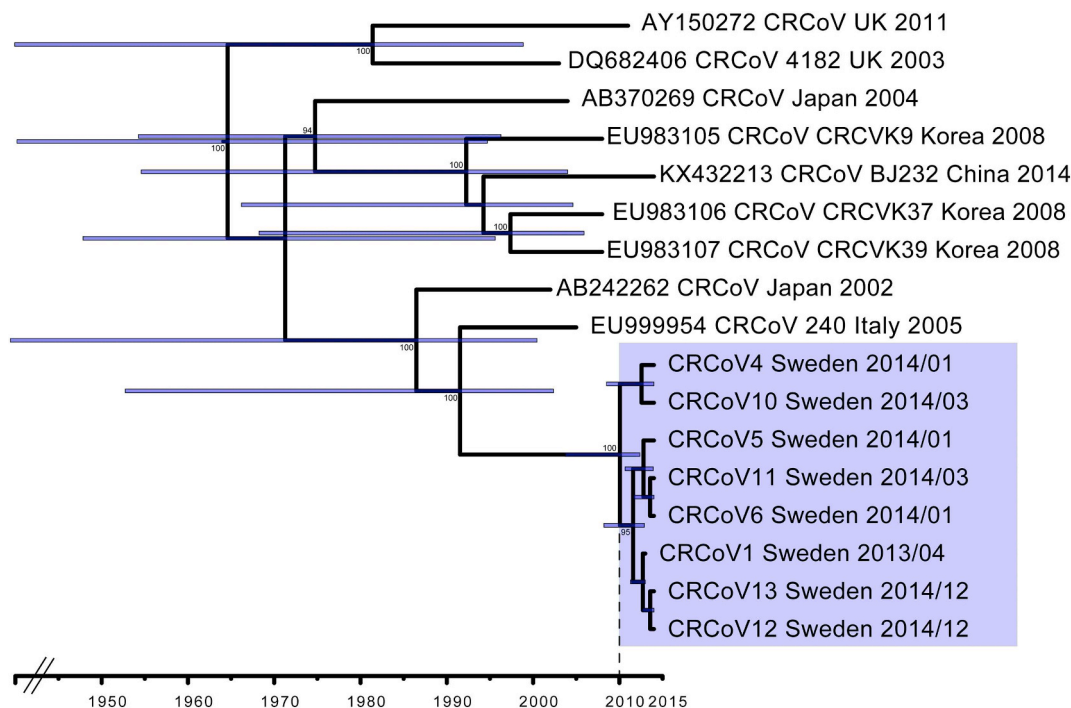
Fig. 2. Time structures phylogeny of full-length spike (S) gene of CRCoV illustrating all viruses from Sweden comprise of a single clade representing a single introduction into Swedish dogs. Node bars correspond to the 95% highest posterior density interval (HDP) on the root height and posterior probabilities are indicated for each node. Scale bar represents time in years.

are likely contributing to spread, similarly to other CIRD-causing pathogens. Overall, we found that 13% of pet dogs with signs of CIRD in Sweden, ranging from mild (cough only) to severe (affected general condition and deep coughing) disease signs, were positive for CRCoV, demonstrating the role that this virus plays in the complex aetiology of CIRD. Importantly, we found CRCoV in dogs of all breeds, ages, and sexes demonstrating that this virus affects all dogs, equally. CRCoV was isolated between December and March; no viruses were detected in the summer months. This lack of significance of season was likely due to

few tests being undertaken in the summer, resulting in a large uncertainty.

The evolutionary genetics of coronaviruses is complex, largely driven by high rates of substitution, and more importantly recombination. This has allowed for interspecies variability, interspecies host jumps and novel coronaviruses to emerge under the "right" conditions (Decaro et al., 2009; Pyrc et al., 2006; Ren et al., 2015; Su et al., 2016; Woo et al., 2009; Zhang et al., 2015). Lineage A of betacoronaviruses is comprised of a number of closely related viruses including
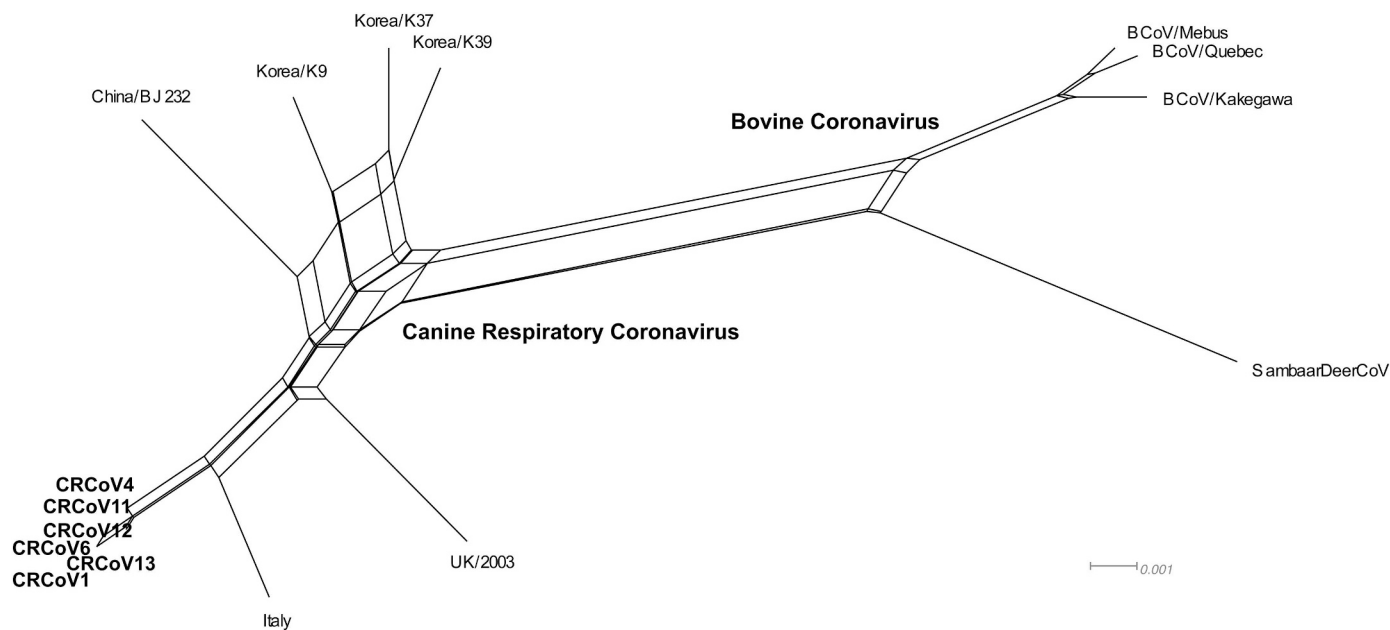

Fig. 3. Splits tree of concatenated NS, HE, S, E and M genes. Each band of parallel edges indicates a split, which is a collection of sequence differences that separates one group of sequences from another. Non-tree like evolution is expected to be from recombination. The distance between taxa represents the sum of weights of all splits that separate taxa.
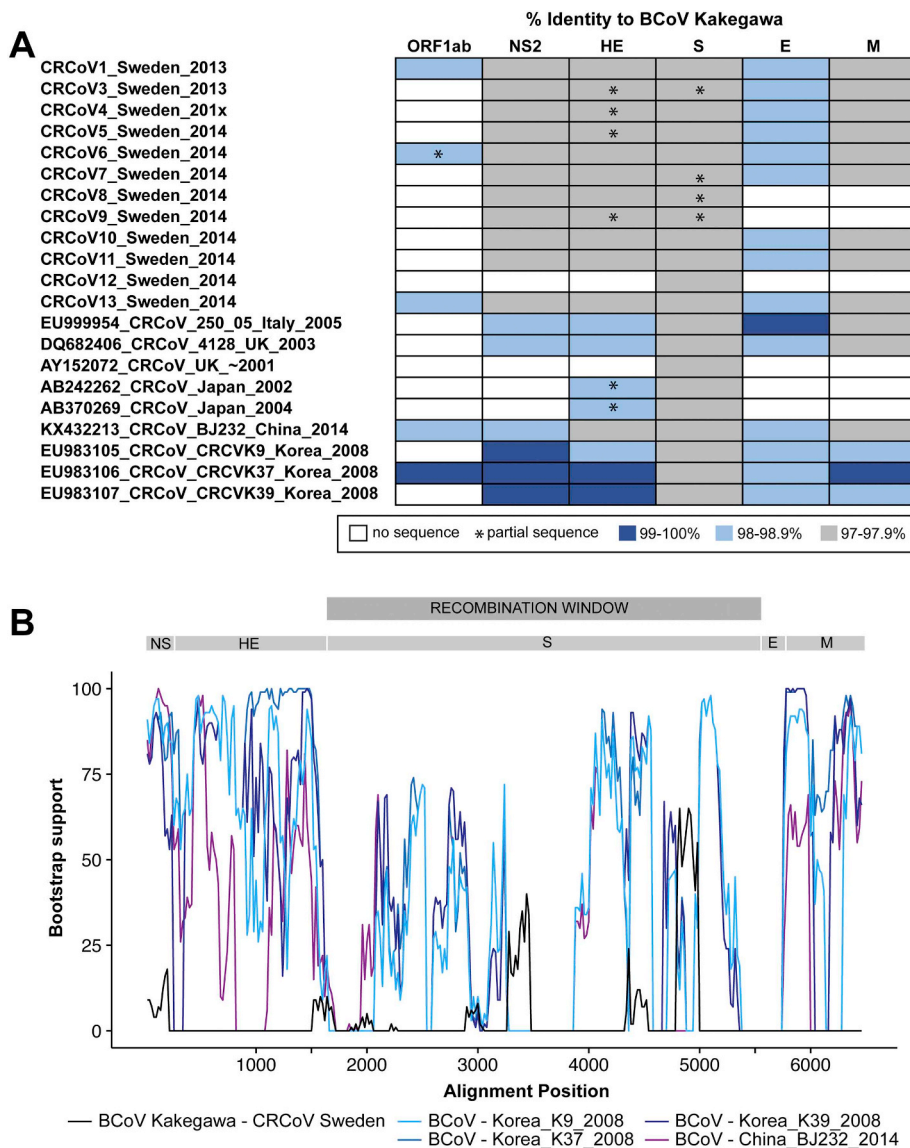
Fig. 4. Recombination between CRCoV and BCoV. (A) Percentage identity between CRCoV and BCoV strain Kakegawa. White indicates no sequence from the gene was available, and grey/blue shading corresponds to the percentage identity to BCoV. Asterisk indicates only a partial gene sequence was available. (B) Bootscan analysis of concatenated sequence illustrating a recombination where CRCoV from Korea and China are more similar to BCoV in the NS, HE and M genes. Genes are identified above the plot, and the statistically supported recombination window is further illustrated above the plot. Lines plotted are percentage bootstrap support include only potentially recombined viruses, in addition to one of the Swedish CRCoV viruses for reference (black). Statistical support for recombination window is presented in Table A3 and Chi-squared breakpoint in Fig. A5. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

CRCoV, BCoV, BCoV-like, and HCoV-OC43. A study by Kin et al., 2016 demonstrated that BCoV and HCoV-OC43 have >96% global nucleotide identity and they reported in an earlier study that HCoV-OC43 may be the result of a zoonotic transmission between bovines and humans (Kin et al., 2016; Kin et al., 2015). Furthermore, there is a 97.3% global nucleotide similarity between CRCoV and BCoV (Erles et al., 2003), but more importantly, cross-reactivity between BCoV antigen with CRCoV antibodies (Erles et al., 2003), and infectivity of puppies with BCoV (Kaneshima et al., 2007; Priestnall et al., 2006). Finally, BCoV and CRCoV (and HCoV-OC43) use sialic acids or heparan sulphate to attach to the cell surface, and use the same molecule as entry receptors (HLA-I, Szczepanski et al., 2019). Overall, this provides a lot of evidence for a very putative host switch event. Our splitstree and recombination analyses of global viruses support the results in Lu et al., which demonstrate high levels of recombination between BCoV and previously described CRCoV strains, including those from UK, Italy, Korea and China (Lu et al., 2017).

The viruses sequenced in this study, at a more local level, add an interesting piece to the puzzle. Viruses from Sweden were most closely related to a sequenced strain from Italy in 2005, and therefore the most parsimonious conclusion is that these viruses were introduced to Sweden from elsewhere in Europe. Despite being detected in a large number of dog species, large geographic area and across 3 years, there was very little genetic variation in CRCoV from Sweden – very little genetic drift, few SNVs and no evidence of recombination, suggesting a single, recent introduction, and subsequent spread. Potentially the most interesting finding, is that the CRCoV circulating in Sweden is the most genetically distant virus to BCoV as compared to CRCoV previously described (particularly those from Asia), with the majority of genes clustering into a "CRCoV clade" rather than BCoV, and most genes being <98% similar to BCoV. Whether this suggests better adaptation to dogs with time, or whether it represents geographical differences, is unclear, and will be revealed with further sequencing of these viruses, globally.

Recent descriptions of novel aetiological causes of CIRD, such as CRCoV, urge for further studies on epidemiology and pathogen genetics to facilitate future vaccine development and diagnostics. To date, commercial vaccines are only available for the well-established CIRD pathogens CPIV, CAV-2 and *B. bronchiseptica*. Despite frequent vaccinations in dogs, CIRD remains a burden for dogs and dog-owners, and as such, it is imperative to better understand the epidemiology and evolution to prevent virus introductions and spread, and to provide informed advice on quarantine decisions related to dog housing facilities, ranging from dog day-care centers to show dogs and racing kennels.
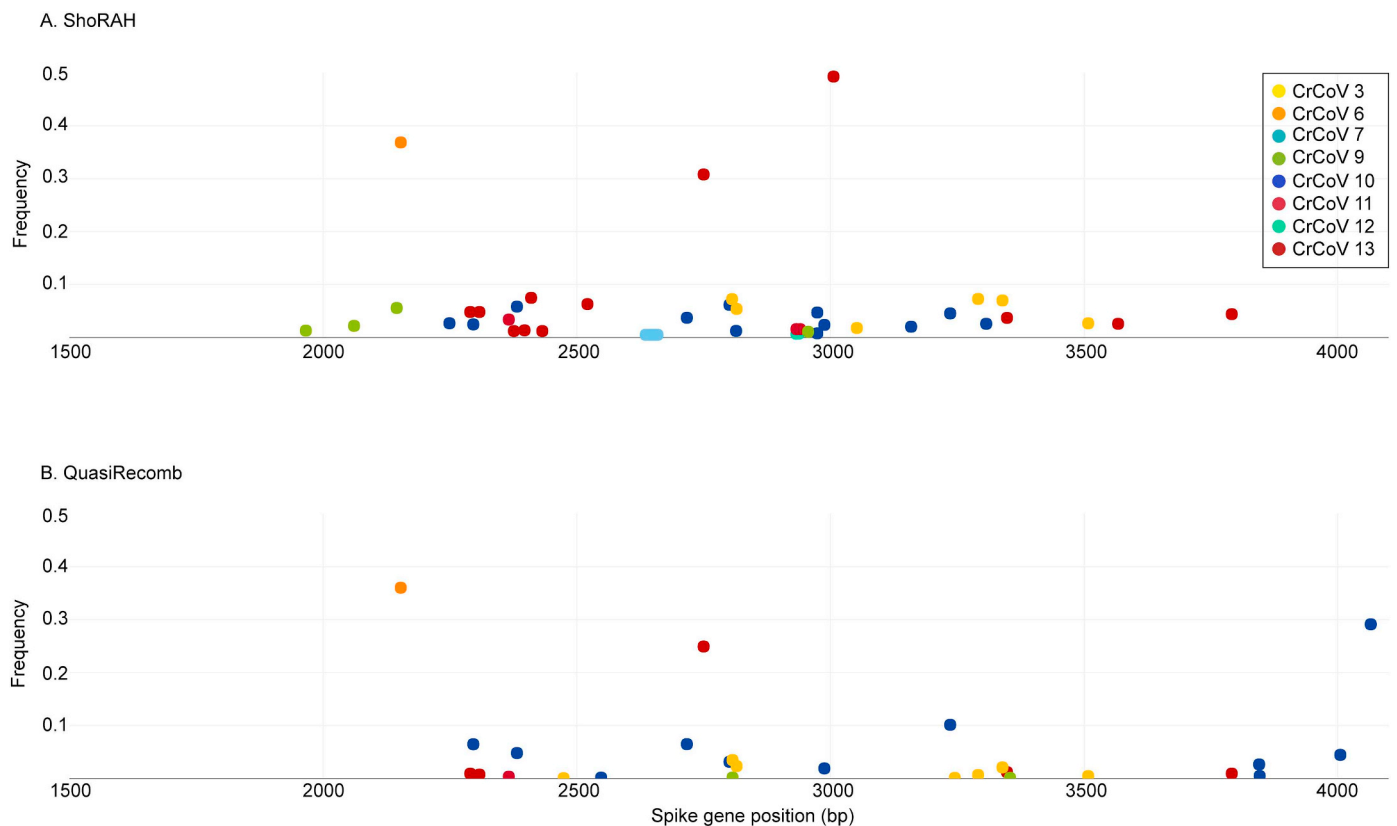
A. ShoRAH

B. QuasiRecomb

Fig. 5. SNVs frequencies in the partial spike genome of CRCoV using (A) ShoRAH and (B) QuasiRecomb. The position refers to the position of the spike gene, and only the region from 1500 to 4000 bp is shown given this was the region amplicon sequenced. Each sample is represented by a unique colour code. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

CRediT authorship contribution statement

Michelle Wille: Writing - original draft, Visualization, Data curation, Formal analysis, Methodology, Software, Conceptualization. **Jonas Johansson Wensman:** Conceptualization, Methodology, Supervision, Writing - review & editing, Funding acquisition, Resources. **Simon Larsson:** Investigation, Formal analysis, Writing - review & editing. **Renaud van Damme:** Visualization, Data curation, Formal analysis, Software, Writing - review & editing. **Anna-Karin Theelke:** Investigation, Writing - review & editing. **Juliette Hayer:** Methodology, Data curation, Formal analysis, Software, Writing - review & editing. **Maja Malmberg:** Conceptualization, Methodology, Supervision, Writing - review & editing, Project administration, Investigation, Validation.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: The study was in part financed by MSD Animal Health. This sponsor did not have any influence on the design of the study or the decision to publish the results.

Acknowledgements

Funding

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.meegid.2020.104290.

References

An, D.-J., Jeong, W., Yoon, S.H., Jeoung, H.-Y., Kim, H.-J., Park, B.-K., 2010a. Genetic analysis of canine group 2 coronavirus in Korean dogs. Vet. Microbiol. 141, 46–52.

An, D.J., Jeoung, H.Y., Jeong, W., Chae, S., Song, D.S., Oh, J.S., Park, B.K., 2010b. A serological survey of canine respiratory coronavirus and canine influenza virus in Korean dogs. J. Vet. Med. Sci. 72, 1217–1219.

Appel, M.J.G., 1987. Canine infectious tracheobronchitis short review: kennel cough. In: Appel, M.J.G. (Ed.), Virus Infections of Carnivores. Elsevier Science Publishing, Amsterdam, pp. 201–211.

Appel, M.J., Percy, D.H., 1970. SV-5-like parainfluenza virus in dogs. J. Am. Vet. Med. Assoc. 156, 1778–1781.

Bahl, J., Krauss, S., Kuhnert, D., Fourment, M., Raven, G., Pryor, S.P., Niles, L.J., Danner, A., Walker, D., Mendenhall, I.H., Su, Y.C., Dugan, V.G., Halpin, R.A., Stockwell, T.B., Webby, R.J., Wentworth, D.E., Drummond, A.J., Smith, G.J., Webster, R.G., 2013. Influenza A virus migration and persistence in North American wild birds. PLoS Pathog. 9, e1003570. https://doi.org/10.1001371/journal.ppat.1003570.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., Pyshkin, A.V., Sirotkin, A.V., Vyahhi, N., Tesler, G., Alekseyev, M.A., Pevzner, P.A., 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J. Comput. Biol. 19, 455–477.

Bemis, D.A., 1992. Bordetella and mycoplasma respiratory infections in dogs and cats. Vet. Clin. North Am. Small Anim. Pract. 22, 1173–1186.

Buchfink, B., Xie, C., Huson, D.H., 2015. Fast and sensitive protein alignment using DIAMOND. Nat. Methods 12, 59–60.

Chen, S., Zhou, Y., Chen, Y., Gu, J., 2018. fastp: an ultra-fast all-in-one FASTQ pre-processor. Bioinformatics 34, i884–i890.

Cunha, C.B., Opal, S.M., 2014. Middle East respiratory syndrome (MERS): a new zoonotic viral pneumonia. Virulence 5, 650–654.

de Groot, R.J., Baker, S.C., Baric, R.S., Brown, C.S., Drosten, C., Enjuanes, L., Fouchier, R.A., Galiano, M., Gorbalenya, A.E., Memish, Z.A., Perlman, S., Poon, L.L., Snijder, E.J., Stephens, G.M., Woo, P.C., Zaki, A.M., Zambon, M., Ziebuhr, J., 2013. Middle East respiratory syndrome coronavirus (MERS-CoV): announcement of the Coronavirus Study Group. J. Virol. 87, 7790–7792.

Decaro, N., Desario, 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., Mari, V., Campolo, M., Lorusso, A., Camero, M., Elia, G., Martella, V., Cordioli, P., Enjuanes, L., Buonavoglia, C., 2009. Recombinant canine coronaviruses related to transmissible gastroenteritis virus of swine are circulating in dogs. J. Virol. 83, 1532–1537.

Ditchfield, J., Macpherson, L.W., Zbitnew, A., 1962. Association of canine adenovirus (Toronto A 26/61) with an outbreak of laryngotracheitis ("kennel cough"): a pre-liminary report. Can Vet J 3, 238–247.

Drummond, A.J., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 7, 214–222.

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.

Erles, K., Brownlie, J., 2005. Investigation into the causes of canine infectious respiratory disease: antibody responses to canine respiratory coronavirus and canine herpesvirus in two kennelled dog populations. Arch. Virol. 150, 1493–1504.

Erles, K., Toomey, C., Brooks, H.W., Brownlie, J., 2003. Detection of a group 2 cor-onavirus in dogs with canine infectious respiratory disease. Virology 310, 216–223.

Erles, K., Shiu, K.-B., Brownlie, J., 2007. Isolation and sequence analysis of canine re-spiratory coronavirus. Virus Res. 124, 78–87.

Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W., Gascuel, O., 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst. Biol. 59, 307–321.

Huson, D.H., Bryant, D., 2006. Application of phylogenetic networks in evolutionary studies. Mol. Biol. Evol. 23, 254–267.

Huson, D.H., Auch, A.F., Qi, J., Schuster, S.C., 2007. MEGAN analysis of metagenomic data. Genome Res. 17, 377–386.

Jeoung, S.-Y., Ann, S.-Y., Kim, H.-T., Kim, D., 2014. M gene analysis of canine coronavirus strains detected in Korea. J. Vet. Sci. 15, 495–502.

Jinnerot, T., Malm, K., Eriksson, E., Wensman, J.J., 2015. Development of a Taqman real-time PCR assay for detection of *Bordetella bronchiseptica*. Veterinary Science: Research and Reviews 1, 14–20.

Kaneshima, T., Hohdatsu, T., Hagino, R., Hosoya, S., Nojiri, Y., Murata, M., Takano, T., Tanabe, M., Tsunemitsu, H., Koyama, H., 2007. The infectivity and pathogenicity of a group 2 bovine coronavirus in pups. J. Vet. Med. Sci. 69, 301–303.

Katoh, K., Asimenos, G., Toh, H., 2009. Multiple alignment of DNA sequences with MAFFT. Methods Mol. Biol. 537, 39–64.

Kin, N., Miszczak, F., Lin, W., Gouilh, M.A., Vabret, A., Consortium, E., 2015. Genomic analysis of 15 human coronaviruses OC43 (HCoV-OC43s) circulating in France from 2001 to 2013 reveals a high intra-specific diversity with new recombinant genotypes. Viruses 7, 2358–2377.

Kin, N., Miszczak, F., Diancourt, L., Caro, V., Moutou, F., Vabret, A., Gouilh, M.A., 2016. Comparative molecular epidemiology of two closely related coronaviruses, bovine coronavirus (BCoV) and human coronavirus OC43 (HCoV-OC43), reveals a different evolutionary pattern. Infect. Genet. Evol. 40, 186–191.

Knesl, O., Allan, F.J., Shields, S., 2009. The seroprevalence of canine respiratory cor-onavirus and canine influenza virus in dogs in New Zealand. N. Z. Vet. J. 57, 295–298.

Li, H., Durbin, R., 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25, 1754–1760.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., Genome Project Data Processing, S, 2009. The sequence alignment/map format and SAMtools. Bioinformatics 25, 2078–2079.

Li, D., Luo, R., Liu, C.M., Leung, C.M., Ting, H.F., Sadakane, K., Yamashita, H., Lam, T.W., 2016. MEGAHIT v1.0: a fast and scalable metagenome assembler driven by advanced methodologies and community practices. Methods 102, 3–11.

Lu, S., Wang, Y.Q., Chen, Y.Z., Wu, B.J., Qin, K., Zhao, J.C., Lou, Y.L., Tan, W.J., 2017. Discovery of a novel canine respiratory coronavirus support genetic recombination among betacoronavirus1. Virus Res. 237, 7–13.

Martin, D., Rybicki, E., 2000. RDP: detection of recombination amongst aligned se-quences. Bioinformatics 16, 562–563.

Martin, D.P., Posada, D., Crandall, K.A., Williamson, C., 2005. A modified bootscan al-gorithm for automated identification of recombinant sequences and recombination breakpoints. Aids Res Hum Retrov 21, 98–102.

Martin, D.P., Lemey, P., Lott, M., Moulton, V., Posada, D., Lefeuvre, P., 2010. RDP3: a flexible and fast computer program for analyzing recombination. Bioinformatics 26, 2462–2463.

Mitchell, J.A., Cardwell, J.M., Renshaw, R.W., Dubovi, E.J., Brownlie, J., 2013. Detection of canine pneumovirus in dogs with canine infectious respiratory disease. J. Clin. Microbiol. 51, 4112–4119.

Mitchell, J.A., Cardwell, J.M., Leach, H., Walker, C.A., Le Poder, S., Decaro, N., Rusvai, M., Egberink, H., Rottier, P., Fernandez, M., Fragkiadaki, E., Shields, S., Brownlie, J., 2017. European surveillance of emerging pathogens associated with canine infectious respiratory disease. Vet. Microbiol. 212, 31–38.

Peiris, J.S., Lai, S.T., Poon, L.L., Guan, Y., Yam, L.Y., Lim, W., Nicholls, J., Yee, W.K., Yan, W.W., Cheung, M.T., Cheng, V.C., Chan, K.H., Tsang, D.N., Yung, R.W., Ng, T.K., Yuen, K.Y., group, S.s, 2003. Coronavirus as a possible cause of severe acute re-spiratory syndrome. Lancet 361, 1319–1325.

Priestnall, S.L., Brownlie, J., Dubovi, E.J., Erles, K., 2006. Serological prevalence of ca-nine respiratory coronavirus. Vet. Microbiol. 115, 43–53.

Priestnall, S.L., Pratelli, A., Brownlie, J., Erles, K., 2007. Serological prevalence of canine respiratory coronavirus in southern Italy and epidemiological relationship with ca-nine enteric coronavirus. J. Vet. Diagn. Investig. 19, 176–180.

Pyrc, K., Dijkman, R., Deng, L., Jebbink, M.F., Ross, H.A., Berkhout, B., van der Hoek, L., 2006. Mosaic structure of human coronavirus NL63, one thousand years of evolution. J. Mol. Biol. 364, 964–973.

R Development Core Team, 2008. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Rambaut, A., Lam, T.T., Carvalho, L.M., Pybus, O.G., 2016. Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). Virus Evolution 2. https://doi.org/10.1093/ve/vew1007.

Ren, L., Zhang, Y., Li, J., Xiao, Y., Zhang, J., Wang, Y., Chen, L., Paranhos-Baccala, G., Wang, J., 2015. Genetic drift of human coronavirus OC43 spike gene during adaptive evolution. Sci. Rep. 5, 11451.

Schulz, B.S., Kurz, S., Weber, K., Balzer, H.J., Hartmann, K., 2014. Detection of re-spiratory viruses and Bordetella bronchiseptica in dogs with acute respiratory tract infections. Vet. J. 201, 365–369.

Seemann, T., 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30, 2068–2069.

Shapiro, B., Rambaut, A., Drummond, A.J., 2006. Choosing appropriate substitution models for the phylogenetic analysis of protein-coding sequences. Mol. Biol. Evol. 23, 7–9.

Sowman, H.R., Cave, N.J., Dunowska, M., 2018. A survey of canine respiratory pathogens in New Zealand dogs. N. Z. Vet. J. 66, 236–242.

Su, S., Wong, G., Shi, W., Liu, J., Lai, A.C.K., Zhou, J., Liu, W., Bi, Y., Gao, G.F., 2016. Epidemiology, genetic recombination, and pathogenesis of coronaviruses. Trends Microbiol. 24, 490–502.

Szczepanski, A., Owczarek, K., Bzowska, M., Gula, K., Drebot, I., Ochman, M., Maksym, B., Rajfur, Z., Mitchell, J.A., Pyrc, K., 2019. Canine respiratory coronavirus, bovine coronavirus, and human coronavirus OC43: receptors and attachment factors. Viruses 11.

Topfer, A., Zagordi, O., Prabhakaran, S., Roth, V., Halperin, E., Beerenwinkel, N., 2013. Probabilistic inference of viral quasispecies subject to recombination. J. Comput. Biol. 20, 113–123.

van der Hoek, L., Pyrc, K., Jebbink, M.F., Vermeulen-Oost, W., Berkhout, R.J., Wolthers, K.C., Wertheim-van Dillen, P.M., Kaandorp, J., Spaargaren, J., Berkhout, B., 2004. Identification of a new human coronavirus. Nat. Med. 10, 368–373.

Vijaykrishna, D., Smith, G.J., Zhang, J.X., Peiris, J.S., Chen, H., Guan, Y., 2007. Evolutionary insights into the ecology of coronaviruses. J. Virol. 81, 4012–4020.

Woo, P.C., Lau, S.K., Huang, Y., Yuen, K.Y., 2009. Coronavirus diversity, phylogeny and interspecies jumping. Exp. Biol. Med. 234, 1117–1127.

Yachi, A., Mochizuki, M., 2006. Survey of dogs for group 2 canine coronavirus infection. J. Clin. Microbiol. 44, 2615–2618.

Zagordi, O., Bhattacharya, A., Eriksson, N., Beerenwinkel, N., 2011. ShoRAH: estimating the genetic diversity of a mixed sample from next-generation sequencing data. BMC Bioinformatics 12, 119.

Zhang, Y., Li, J., Xiao, Y., Zhang, J., Wang, Y., Chen, L., Paranhos-Baccala, G., Ren, L., Wang, J., 2015. Genotype shift in human coronavirus OC43 and emergence of a novel genotype by natural recombination. J. Inf. Secur. 70, 641–650.