HAL
open science

First full length sequences of the S gene of European isolates reveal further diversity among turkey coronaviruses.

Stéphan Maurel, Didier Toquin, François-Xavier Briand, Maryline Queguiner, Chantal Allee, Joel Bertin, Charlotte Retaux, Vincent Turblin, Hervé Morvan, Nicolas Eterradossi

To cite this version:

Stéphan Maurel, Didier Toquin, François-Xavier Briand, Maryline Queguiner, Chantal Allee, et al.. First full length sequences of the S gene of European isolates reveal further diversity among turkey coronaviruses.. Avian Pathology, 2011, 40 (02), pp.179-189. 10.1080/03079457.2011.551936 . hal-00687805

HAL Id: hal-00687805

https://hal.science/hal-00687805

Submitted on 15 Apr 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

First full length sequences of the S gene of European isolates reveal further diversity among turkey coronaviruses.

Journal:	*Avian Pathology*
Manuscript ID:	CAVP-2010-0118.R2
Manuscript Type:	Original Research Paper
Date Submitted by the Author:	24-Nov-2010
Complete List of Authors:	MAUREL, Stéphan; French Agency for Food Environmental and Occupational Health Safety (Anses) Ploufragan-PLouzane Laboratory, Avian and Rabbit Virology Immmunology and Parasitology Reserach Unit (VIPAC) TOQUIN, Didier; French Agency for Food Environmental and Occupational Health Safety (Anses) Ploufragan-PLouzane Laboratory, Avian and Rabbit Virology Immmunology and Parasitology Reserach Unit (VIPAC) BRIAND, François-Xavier; French Agency for Food Environmental and Occupational Health Safety (Anses) Ploufragan-PLouzane Laboratory, Avian and Rabbit Virology Immmunology and Parasitology Reserach Unit (VIPAC) QUEGUINER, Maryline; French Agency for Food Environmental and Occupational Health Safety (Anses) Ploufragan-PLouzane Laboratory, Avian and Rabbit Virology Immmunology and Parasitology Reserach Unit (VIPAC) ALLEE, Chantal; French Agency for Food Environmental and Occupational Health Safety (Anses) Ploufragan-PLouzane Laboratory, Avian and Rabbit Virology Immmunology and Parasitology Reserach Unit (VIPAC) BERTIN, Joel; Cooperative Le Gouessant RETAUX, Charlotte; Cooperative Le Gouessant TURBLIN, Vincent; MC Vet Conseil; present adress: Ceva Animal Health MORVAN, Hervé; Laboratoire de Développement et d'analyse des Côtes d'Armor

	ETERRADOSSI, Nicolas; French Agency for Food Environmental and Occupational Health Safety (Anses) Ploufragan-PLouzane Laboratory, Avian and Rabbit Virology Immmunology and Parasitology Reserach Unit (VIPAC)
Keywords:	turkey, coronavirus, S gene, Enteric viruses

SCHOLARONE™
Manuscripts

For Peer Review Only

Cavp-2010-0118.R2

First full length sequences of the S gene of European isolates reveal further diversity among turkey coronaviruses.

S. Maurel[1], D. Toquin[1], F.X. Briand[1], M. Quéguiner[1], C. Allée[1], J. Bertin[2], L. Ravillion[3], C. Retaux[2], V. Turblin[4], H. Morvan[5] and N. Eterradossi[1*]

[1] Anses - French agency for food, environmental and occupational health safety - Avian and Rabbit Virology, Immunology, and Parasitology Unit, B.P. 53, 22440 Ploufragan, France. [2] Coopérative Le Gouessant, B.P. 228, 22400 Lamballe, France. [3] Laboratoire BOCAVET, 79140 Cerizay, France. [4] MC Vet Conseil, 72300 Sablé-sur-Sarthe, France ; present adress : Ceva Animal Health 46300 Selangor Malaysia. [5] Laboratoire de Développement et d'analyse des Côtes d'Armor, B.P. 54, 22440 Ploufragan, France.

Short title: Molecular characterization of French turkey coronavirus.

Corresponding author contact:

Telephone number: (33) 02 96 01 62 88

Fax number: (33) 02 96 01 62 63

[*]**E-mail of corresponding author**: nicolas.eterradossi@anses.fr

Received: 24 November 2010

For Peer Review Only

Abstract

An increasing incidence of enteric disorders clinically evocative of the poult enteritis complex has been observed in turkeys in France since 2003. Using a newly designed real-time RT-PCR assay specific for the nucleocapsid (N) gene of infectious bronchitis virus (IBV) and turkey coronaviruses (TCoV), coronaviruses were identified in 37 % of the intestinal samples collected from diseased turkey flocks. The full length Spike (S) gene of these viruses was amplified, cloned and sequenced from three samples. The French S sequences shared 98 % at both the nt and aa levels, whereas they were at most 65 and 60 % identical with North American (NA) TCoV and at most 50 and 37 % identical with IBV at the nt and aa levels, respectively. Higher divergence with NA TCoV was observed in the S1-encoding domain. Phylogenetic analysis based on the S gene revealed that the newly detected viruses form a sublineage genetically related with, but significantly different from, NA TCoV. Additionally, the RNA dependent RNA polymerase (RdRp) and the N genes, located on the 5' and 3' sides of the S gene in the coronavirus genome, were partially sequenced. Phylogenetic analysis revealed that both the NA TCoV and French TCoV (Fr TCoV) lineages included some IBV relatives, which were however different in the two lineages. This suggested that different recombination events could have played a role in the evolution of the NA and Fr TCoV. The present results are the first S sequence for a European TCoV. They reveal extensive genetic variation in TCoV and suggest different evolution pathways in North America and Europe.

For Peer Review Only

Introduction

Coronaviruses (CoVs) consist of large, enveloped and positive-stranded RNA viruses within the order *Nidovirales*. They possess an approximately 30 kb-long genome from which they transcribe a set of multiple 3'-coterminal nested subgenomic mRNAs (Masters, 2006). The virions are pleomorphic enveloped particles, roughly spherical, with diameters ranging from 50 to 200 nm. They possess long petal-shaped spikes on their membrane that are responsible for the CoV typical crown-shaped morphology in electron microscopy. CoVs infect a wide variety of avian and mammalian species and cause primarily respiratory or enteric diseases, but also in some cases neurologic illness or hepatitis. In humans, the outbreak of severe acute respiratory syndrome (SARS) caused in 2003 by a CoV has led to an increased interest in family *Coronaviridae* and its potential animal reservoirs. Based on the latest proposals to the International Committee on the Taxonomy of Viruses (ICTV), family *Coronaviridae* in the *Nidovirales* order includes two sub-families, *Coronavirinae* and *Torovirinae*, the former comprising three genera, *Alpha-*, *Beta-* and *Gamma-coronaviruses* (DeGroot *et al.,* 2008). The proposed *Gammacoronavirus* genus groups mostly coronaviruses isolated from birds, with the exception of the SW1 and ALC/GX/230/06 viruses isolated from a beluga whale (Mihindukulasuriya *et al.*, 2008), and from the Asian leopard cat or Chinese ferret badger (Dong *et al.*, 2007), respectively. A first proposed species, *avian coronavirus* (AvCoV), corresponds to the former "subgroup 3a within the coronavirus genus". It groups isolates obtained from *Galliformes* (chicken, turkey, pheasant, peafowl, quail), *Columbiformes* (pigeon), *Anatidae* (teal, goose, duck, swan), *Charadriiformes* (red knot and oystercatcher) and possibly *Psittaciformes* (Cavanagh *et al.*, 2002; Jonassen *et al.*, 2005; Liu *et al.*, 2005; Qian *et al.*, 2006; Circella *et al.*, 2007; Hughes *et al.*, 2009; Gough *et al.*, 2006). The second proposed species within the *Gammacoronavirus* genus is Beluga Whale coronavirus SW1, and corresponds to the former 3b subgroup (Mihindukulasuriya *et al.*, 2008). A third group of isolates (former 3c subgroup) contains viruses obtained from passerines (bulbul, munia, and

For Peer Review Only

thrush) (Woo *et al.*, 2009) and from the Asian leopard cat and Chinese ferret badger (Dong *et al.*, 2007; Who *et al.*, 2009), however these isolates have not yet been assignated to any species within the *Gammacoronavirus* genus.

The most economically significant viruses among AvCoVs are infectious bronchitis viruses (IBV) and turkey coronaviruses (TCoV), affecting the poultry industry. IBV causes avian infectious bronchitis, a common, highly contagious, and acute viral respiratory and genital disease of domestic fowl with worldwide distribution and major economic consequences (Cavanagh & Gelb, 2008). TCoV was shown in the 1970's to be one of the causative agents of an enteric disease known as bluecomb, transmissible enteritis or coronaviral enteritis of turkeys (Guy, 2008). More recently, TCoV has been associated with poult enteritis complex (PEC) which groups several infectious intestinal disorders of young turkeys up to 7 weeks of age. Clinical signs include diarrhoea, stunting, dehydration, anorexia, weight loss and immune dysfunction. When associated with mortality, this disease is designated as poult enteritis and mortality syndrome (PEMS) (Barnes *et al.*, 2000).

CoVs share a common genomic organization which consists of two ORFs in the 5'-end that encode a single viral replicase and a varying number of genes in the 3' end that encode, among other products, the four major structural proteins (spike S, envelope E, membrane M and nucleocapsid N). The viral replicase is cleaved by proteases leading to the release of 15 or 16 proteins, among which the ORF1b-encoded RNA dependent RNA polymerase (RdRp). Among the four major structural proteins, N is a phosphoprotein of approximately 50 kDa that binds the genomic RNA. It is an immunogenic determinant for humoral and cellular immunity (Collisson *et al.*, 2000; Tang *et al.*, 2008). The S glycoprotein forms the large, club-shaped projections on the external surface of the virion envelope. S contains two subdomains S1 and S2 that are, in most Beta- and Gamma-CoVs, cleaved by a trypsin-like host protease (Masters, 2006). The S2 subdomain anchors the spike into the virus membrane whereas the S1 subdomain forms the extracellular globular portion of the spike and supports the host receptor-binding activity (Kubo *et al.*, 1994; Krempl *et al.*, 2000; Wong

For Peer Review Only

et al., 2004). In addition, the S1 glycoprotein contains major epitopes that induce neutralizing antibodies (Cavanagh *et al.*, 1988).

So far, only North American strains of TCoV (NA TCoV) have been extensively sequenced and characterized at the molecular level (Lin *et al.*, 2004; Cao *et al.*, 2008; Gomaa *et al.*, 2008; Jackwood *et al.*, 2010). These data clearly showed sequence homogeneity between the studied strains with at least 92.4 % nt identity for the full-length genomes and 91 % aa identity for the S proteins. The percent identity between TCoV and IBV isolates was also high for the full length genome (higher than 86 %) indicating a close genetic relationship between these two viruses, however the S proteins of the compared TCoV and IBV shared at most 36 % aa identity. This finding led to the hypothesis that TCoV might have emerged from IBV through recombination (Jackwood *et al*., 2010). In Europe, only short nt sequences corresponding to the 3' end of the genome of an IBV-related CoV have been detected from turkeys in the UK (Cavanagh *et al.*, 2001), in Italy (Moreno Martin *et al.*, 2002), and in Poland (Domanska-Blicharz *et al.*, 2010). However, since 2003, an increasing number of turkey flocks exhibiting clinical signs compatible with PEC have been observed in France (Germain & Rousseau, 2005). In this paper we report on the development of a Taqman quantitative RT-PCR (RT-qPCR) suitable for the detection of IBV and TCoV, on the detection in France of CoV-positive turkey flocks with signs evocative of PEC, and on the sequencing of the entire hypervariable S gene, together with fragments of the N and RdRp genes of these strains.

For Peer Review Only

Materials and Methods

Viruses and clinical samples. The IBV and TCoV strains used in the validation of the RT-qPCR are listed in Table 1. Twenty-one viruses not belonging to the *Gammacoronavirus* genus but representing avian metapneumoviruses, para- and orthomyxoviruses, adenoviruses, reoviruses, avibirnaviruses, rotavirus and astroviruses were also tested (Guionie *et al.*, 2007).

Clinical samples (duodenum, jejunum, cæcum, kidney, spleen, liver, bursa of Fabricius or/and thymus) were received for virological investigation from 81 turkey flocks collected in Western France and presenting with signs evocative of PEC between April 2007 and October 2009. Twenty three digestive contents collected before 1988 from meat turkey flocks with enteric disorders (Andral *et al.*, 1985) where also investigated. The pre-1988 samples had been kept frozen at – 20 °C since harvest. All samples were ground and suspended w/v in PBS. The suspensions were centrifuged at 3000 x g for 15 min and the supernatants were stored at – 70 °C prior to RNA extraction.

RNA extraction. RNA was extracted from 140 µl of viral suspension using the QIAamp® Viral RNA Mini Kit (Qiagen, Hilden, Germany), according to the manufacturer's instructions. The purified RNA was eluted in 60 µl of AVE buffer and stored at – 70 °C prior to RT-qPCR.

Development and validation of the RT-qPCR. To easily screen the field samples and avoid a lack of amplification of some European viruses due to their possible genetic variation, it was decided not to rely on the methods previously developed for TCoV but to develop a broad spectrum RT-qPCR specific for both IBV and TCoV. The oligonucleotide primers and Taqman probe were designed with the Primer Express® software (Applied Biosystems, Warrington, UK) from a conserved region in the middle of the N gene. The primers and probe were screened with the NCBI BLAST programme to check for the lack of cross reaction with previously released cellular, viral or bacterial sequences. The forward primer Ncor800f, reverse primer Ncor860r and the Taqman® dual-labeled probe Ncor830p (Table 2) were synthesized by Applied Biosystems and amplified a 66-bp cDNA. The one-step Taqman® AvCoV RT-qPCR assay was run using the Quantitect™ probe RT-PCR Kit (Qiagen, Hiden, Germany). Briefly, each mix contained 1X (12.5 µl) of Quantitect™ RT-PCR master mix, 0.25 µl of Quantitect RT mix, 300 nM of each primer, 100 nM of probe, 5 µl of template RNA and RNase-free water to a final volume of 25 µl. RT-qPCR was performed using a

For Peer Review Only

Taqman 7000 thermocycler (Applied Biosystems) under the following cycling conditions: 48 °C for 30 min (reverse transcription), 95 °C for 10 min (RT inactivation and activation of the HotStartTaq DNA polymerase); 40 cycles combining 95 °C for 15 sec (denaturation) and 60 °C for 1 min (annealing, extension step, and fluorescence data collection). Data were analysed with the SDS software, version 1.3.1 (Applied Biosystems).

To assess the analytical sensitivity of detection, a fragment of the N gene encompassing the region amplified by the newly developed RT-qPCR assay was amplified with the Ncor1f and Ncor860r primers (Table 2), then was cloned into the pcDNA 3.1©Directional TOPO expression kit (Invitrogen, Carlsbad, USA) according to the manufacturer's procedure. The resulting plasmid was linearized at the *Eco RV* (New England Biolabs, Hitchin, UK) restriction site located downstream of the insert. Runoff transcripts were synthesized using the T7 RNA polymerase (Promega, Madison, USA) according to the standard protocol. Following the transcription reaction, the DNA templates were removed by digestion with the RQ1 RNase-free DNase (Promega, Madison, USA). The lack of residual contaminating DNA was assessed by checking negativity in a qPCR assay (RT-qPCR without the reverse transcriptase enzyme). The *in vitro* transcripts were extracted with phenol/chloroform, resuspended in nuclease-free water, aliquoted and stored at –70 °C. They were quantified by measuring the A_{260} with a spectrophotometer to calculate the number of copies. To evaluate the amplification efficiency, the RT-qPCR assay was tested using 10-fold dilution series from $5x10^2$ to $5x10^9$ copies per reaction of RNA runoff transcripts.

Full length amplification and sequencing of the S gene. Some field samples with a positive RT-qPCR result were used for the full length amplification and sequencing of the S gene. To amplify the full length S gene, RT-PCR was performed using primers defined in conserved regions flanking the S gene, namely primer Scor+260f in the 3' end of gene 1b and primer Scor-150r in the 5' end of gene 3a (Table 2). RNA was reverse transcribed at 42 °C for 75 min using Superscript II reverse transcriptase (Invitrogen, Carlsbad, USA) and primers Scor-150r. Four microliters of cDNA were amplified with primer pairs Scor+260f / Scor-150r

For Peer Review Only

(Table 2). PCR was performed with the Expand High Fidelity PCR kit (Roche, Mannheim, Germany) according to the manufacturer's instructions. The PCR products were purified using the Gene Elute kit (Qiagen, Hiden, Germany) and cloned into the pGEM®-T easy vector (Promega, Madison, USA) according to the manufacturer's instructions. At least 3 positive clones for each sample were sequenced in both directions. The primers defined in the newly determined sequence were used according to a gene-walking strategy (primers available upon request). PCR and sequencing were performed with the Big Dye® terminator cycle kit and the Genetic Analyzer 3130 system according to the manufacturer's recommendations (Applied Biosystems).

Partial amplification and sequencing of the N and RdRp genes. Some field samples with a positive RT-qPCR result were also used for the partial amplification and sequencing of both the N and RdRp genes. RNA was reverse transcribed at 42 °C for 75 min using Superscript II reverse transcriptase (Invitrogen, Carlsbad, USA) and primers Ncor860r or POLcor2590r for the N or RdRp genes, respectively. Four microliters of cDNA were amplified with primer pairs Ncor1f / Ncor860r or POLcor1900f / POLcor2590r (Table 2), for the N or RdRp genes, respectively. PCR was performed with the Expand High Fidelity PCR kit (Roche, Mannheim, Germany) according to the manufacturer's instructions. Amplicons were purified after agarose gel electrophoresis using the Gene Elute kit (Qiagen, Hiden, Germany). The purified amplicons were sequenced as described for the S gene, using the PCR primers.

Similarity searches, phylogenetic analysis and protein motif predictions. The nt-nt or protein-protein BLAST search analyses to define the closest IBV or NA-TCoV relatives of the newly determined sequences were performed on-line at the National Center of Biotechnology Information (http://blast.ncbi.nlm.nih.gov/Blast.cgi). Percent nt identities in pairwise alignments were determined with the MEGA software (Kumar *et al.*, 2008). For phylogenetic analyses, coronavirus sequences were retrieved from databanks that were closely related to TCoV, as detected by Blast, or were representative of IBV isolates from the USA,

For Peer Review Only

Europe or China and had their S, N and RdRp genes sequenced, or were representative of other virus species within the *Gammacoronavirus genus* (Table 3). The nt sequences were first translated and the deduced protein sequences were aligned using the BLOSUM matrix. The alignment of nt sequences was then deduced from the alignment of protein sequences. Phylogenetic analysis were finally inferred using the Neighbor-Joining (with the Kimura 2-parameter model), using the MEGA software, and with the maximum likelihood method using the PhyML software (Guindon & Gascuel, 2003). All methods were implemented with bootstrap on 1000 replicates and using the SARS-CoV (Beta-CoV) sequence as an outgroup for maximum likelihood method.

Concerning the motif predictions from the amino acid sequence of the S protein, we used the ProP server (http://www.cbs.dtu.dk/services/ProP/) to detect putative peptide cleavage sites, the TopPred website (http://mobyle.pasteur.fr/cgi-bin/portal.py?form=toppred) to predict hydrophobicity profiles, and the COILS program (http://www.ch.embnet.org/software/COILS_form.html) to deduce heptad repeat regions relative to coil-coiled structure.

Accession numbers for the reported sequences. The gene sequences for FR070341j, FR080147c, FR080183j and IBV-4/91 were submitted to the EMBL database and have been assigned accession numbers: Full spike gene of FR070341j, FR080147c and FR080183j are under accession number GQ411201, FN434203, FN545819, respectively. The accession numbers for the partial N gene sequences of FR070341j, FR080147c and FR080183j are FN665664 to FN665666, respectively. The accession numbers for the partial RdRp gene sequences of FR070341j, FR080147c, FR080183j and IBV-4/91 are FN811144 to FN811147, respectively.

Results

For Peer Review Only

Specificity and sensitivity of the RT-qPCR. Primers Ncor800f and Ncor860r and probe Ncor830p detected efficiently the genome of both TCoV-INP4 and the 14 strains of IBV tested in this study (Table 1). All the non-CoV controls (Guionie *et al.*, 2007) produced cycle threshold (Ct) higher than 35.

A linear standard curve of template copy number against Ct value was established using serial 10-fold dilutions of in vitro transcripts. The detection range was between $10^{2.0}$ and $10^{9.0}$ copies per µl, with a slope of - 3.48 $\pm$ 0.09 indicating an amplification efficiency of 94.0 % $\pm$ 3.3 % and a coefficient of correlation (R^2) of 0.999 $\pm$ 0.001 for eighteen independent repeated runs (data not shown). The linear (quantification) range was between $10^{3.0}$ and $10^{9.0}$ copies per µl.

RT-qPCR applied to clinical samples. Out of 52 flocks with clinical signs evocative of PEC/PEMS, 19 flocks (36.5 %) were positive, whereas only 5 out of 29 (17.2 %) were positive among flocks without enteritis or with an enteritis not evocative of PEC/PEMS ($p \leq 0.07$, Chi² test). Positive RT-qPCR results were obtained predominantly from the content of jejunum (41.3 %) and cæcum (39.1 %) and in the bursa of Fabricius (38.6 %). RT-qPCR did not reveal any positive out of twenty three clinical samples that had been collected during previous surveys of turkey enteric disorders performed before 1988 (Andral *et al.*, 1985).

Genetic analysis of the CoV strains based on the complete spike sequences. The S ORF, as amplified from the genomes of the French TCoV (Fr TCoV) FR070341j, FR080147c and FR080183j viruses, was 3597 nt-long (including the stop codon). This size was intermediate between that of TCoV (3612-3681 nts) and IBV (3489-3510 nts) S genes. A typical transcription-regulating sequence (TRS) motif (Zuniga *et al.*, 2004), CUGAACAA, was identified 52 nt upstream of the S ORF initiation codon, as in all other IBV and TCoV genes sequenced to date. The French S sequences shared 98 % nt identity, whereas they were at most 65 % identical with TCoV-ATCC, one of their closest NA TCoV relatives, and merely 50 % identical with IBV-California99 (also originating from the USA), one of their closest

For Peer Review Only

IBV relatives, both as detected with the Blast programme. The sequence differences were most obvious in the region encoding the S1 subdomain, where the French sequences shared 97 % nt identity, but at most only 56 and 38 % nt identity with NA TCoV and IBV strains, respectively.

The consensus tree resulting from the phylogenetic analysis of the nt sequences encoding the S1 subunit of gamma- CoVs is shown in Figure 1A. The tree corresponding to the full S gene had a similar topology with significant bootstrap support (data not shown). A first significant cluster (99 % boootstrap value) grouped all the selected sequences from IBV and TCoV and was consistent with isolates belonging to the AvCoV species. The SW1 virus from a beluga whale made up a second lineage, which was consistent with this virus representing a separate species. The three viruses, ThCoV, BuCoV, and MuCoV detected in thrush, bulbul, and munia, respectively, clustered with the asian leopard cat virus in a third significant cluster (100 % bootstrap value). Within AvCoVs, two major genetic lineages were supported by a 100 % bootstrap value: the first one grouped all selected IBV whereas the second grouped all NA TCoV sequenced to date, the sequence from a quail CoV (QCoV) recently identified in Italy (Circella *et al.*, 2007) and the FR070341j, FR080147c and FR080183j sequences. However, in agreement with the percent nt identity discussed above, the three French sequences clustered significantly apart in a significant sublineage (100 % bootstrap value) (Figure 1A).

The deduced putative S protein of FR070341j, FR080147c and FR080183j was 1198 aa-long. The pairwise comparisons of these aa sequences revealed 98 % aa identity between the three French strains, which shared at best 61 % overall identity with the NA TCoV and 37 % aa acid identity with the IBV strains used in this study. The S1 subdomain (aa 1-529) of the three French strains shared 96 % aa identity, whereas maximum aa identity was only 42 and 18 % with NA TCoV and with the closest IBV relatives, respectively. The alignment of the S1 subdomains of the French sequences with that of NA TCoV revealed two zones with different levels of aa identity (Figure 2). Indeed, positions 1 to 196 (S1a) exhibited only 35 (18 %) fully conserved residues shared by all strains from turkeys or quail, whereas aa

For Peer Review Only

positions 197 to 529 (S1b) contained 144 (43 %) such residues. It is tempting to speculate that high aa divergence in the S1a region supports an antigenically significant hypervariable region (HVR) similar to the HVR1 and HVR2 regions described in IBV S1 (Cavanagh *et al.*, 1988). However, further molecular and antigenic studies on TCoV will be required to substantiate this hypothesis. At the carboxy terminal end of the S1 subdomain, the three French strains shared the same putative glycoprotein furin cleavage recognition site, TRSRR/S (aa positions 525-529), which appeared to be unique among IBV and TCoV strains. The S2 subdomain, carboxyterminal to the cleavage site, proved more conserved (99 %, 76 % and 52 % maximum aa identity among French sequences, between Fr and NA TCoV and between Fr TCoV and IBV sequences, respectively). Its predicted hydrophobicity profile was consistent with three extracellular, transmembrane and intracytoplasmic domains spanning aa positions 530 to 1120, 1121 to 1159, and 1160 to 1198, respectively. As recently reported with S2 of other CoVs (Yamada & Liu, 2009), the extracellular domain of the newly determined TCoV S2 sequences exhibited the consensus SAIEDLLF aa stretch (aa positions 694-701), located immediately carboxy terminal to the SGKPQGR sequence. This sequence did not correspond to any furin cleavage site, unlike that observed in several IBV isolates, however it was identified as the second most probable protein cleavage site (after the S1/S2 cleavage site) in the TCoV S protein, and it further fits the XXXR/S model recently proposed for a second cleavage of the S protein, which appears critical for CoV-induced cell-to-cell fusion in cultured cells (Yamada & Liu, 2009). The search of the S2 subdomain for heptad repeats (HR), possibly indicative of a coiled-coil tertiary structure, revealed two compatible areas spanning aa positions 792 to 948 and 1056 to 1141, respectively. These positions were consistent in the alignment of the S proteins with the regions previously proposed as encompassing HR1 and HR2 in IBV (Bosch *et al.*, 2004). Interestingly, the alignment demonstrated that all TCoV sequences exhibited an exact insertion of two HR (14 aa) in both HR1 and HR2. These insertions occurred in close vicinity (only 4 aa apart) from the sites where similar insertions had been previously recognized (although with different inserted aa) as characteristic of HR1 and HR2 in all alpha-CoVs (Bosch *et al.*, 2004). The second

For Peer Review Only

predicted HR2 in the TCoV extracellular domain involved the most amino terminal aa of the transmembrane domain, which also belonged to the YIKWPWYVWL aa stretch that is highly conserved among the three CoV genera (Rota *et al.*, 2003). Finally, the predicted S intracellular domain was also highly conserved between TCoV and IBV and included the YYTTF signal involved in the retention of IBV S at a late Golgi compartment (Winter *et al.*, 2008).

Genetic analysis of the CoV strains based on the nucleocapsid and polymerase sequences. For each analyzed gene, all methods used for phylogenetic analysis resulted in trees with a consistent topology (Figures 1B and 1C). Irrespective of the analyzed gene, the SW1 CoV and the group made of ThCoV, BuCoV, MuCoV and the Asian leopard cat CoV always represented significant separated genetic lineages, whereas IBV and TCoV strains always grouped significantly together within a cluster consistent with the AvCoV species. In the N gene, sequences FR070341j, FR080147c and FR080183j shared 94 to 97 % nt identity. They shared 95 to 99 % and 91 to 92 % nt identity with their closest IBV (strain 4/91) or TCoV (strain IN-517/945) relatives (as identified with the Blast programme), respectively. Within AvCoV, and as already observed with the S1 gene, distinct sublineages were apparent, containing either the NA or the Fr TCoV sequences, as supported by 84 % and 100 % bootstrap values, respectively (Figure 1B). However, unlike previously observed with the S1 gene, the IBV sequences did not cluster in a specific sublineage. Indeed, different IBV strains with close phylogenetic relationships with TCoVs were apparent, either in the significant sublineage containing all NA TCoVs (IBV strains CK/CH/LSD/05I, ArkDPI101 and California99, 84 % booststrap value), or in the sublineage containing the three newly determined sequences (IBV-4/91, 100 % bootstrap value with TCoV-FR070341j). Additional IBV strains were also present in the AvCoV lineage, but without any significant connection to the two above mentioned sublineages (IBV strains M41, ITA/90254/2005 and BJ).

In the RdRp gene, sequences FR070341j, FR080147c and FR080183j shared 95 to 98 % nt identity, and at most 95 and 94 % nt identity with their closest IBV (strain ITA/90254/2005) and TCoV (strain IN-517/94) relatives, respectively. Within the AvCoV

For Peer Review Only

lineage, the three French strains clustered significantly together again, in a sublineage that also contained IBV strain ITA/90254/2005 (90 % bootstrap value in Figure 1C). As already observed with the S1 or N genes, a significant cluster (92 % bootstrap value) grouped some NA TCoV (TX-1038/98, VA-74/03, MG10) with IBV strains (California99 and ArkDPI101). However, unlike observed previously, the bootstrap values did not support the clustering of all NA TCoV sequences into it, as IN-517/94 and TCoV-ATCC were only loosely related (bootstrap values < 75%). In addition, the phylogenetic analysis suggested that IBV CK/CH/LSD/05l branched significantly apart from all other studied AvCoV RdRp sequences (95 % bootstrap value), whereas this IBV strain had previously exhibited a N sequence genetically related to NA TCoVs (see above and Figure 1B).

For Peer Review Only

Discussion

The aim of this study was to further investigate the possible role of TCoV in digestive disorders of commercial turkeys in France and to refine the molecular characterization of the detected viruses.

As a first step, we therefore decided to develop a RT-qPCR assay specific for the most economically significant AvCoV (i.e. IBV and TCoV). We selected the N gene to design primers and probe, as it is present in all genomic and subgenomic CoV RNA transcribed in infected cells, thus ensuring sensitivity of detection, and also contains highly conserved regions allowing to avoid false negatives due to genetic variation. Indeed, it has been reported for different avian RNA viruses that the genetic lineages that circulate in the USA and Europe may exhibit a very significant degree of genetic variation (Webster *et al.*, 1992; Toquin *et al.*, 2006). The authors are aware of only one other published broad spectrum RT-PCR assay suitable for the detection of both IBV and TCoV (Callison *et al.*, 2006). This assay however targeted the 3'UTR region of AvCoV, and not the N gene as targeted here.

In France, the most studied AvCoV has been so far IBV, with outbreaks of infectious bronchitis in the late 1980's leading to the first detection of the CR88 genotype (also named 793/B, or 4/91) (Cavanagh *et al.*, 1998) that replaced the previously prevalent D274 and Massachusetts serotypes (Davelaar *et al.*, 1984; Picault *et al.*, 1986). More recently, two new genotypes, QX and Italy02 have emerged in Europe and in France (Worthington *et al.*, 2008). There has been so far no report in France of any turkey infection by AvCoV. In Great Britain however, Cavanagh *et al.* (2001) reported the presence of a coronavirus in turkeys with enteritis.

Clinical data from the investigated flocks suggest a trend for this virus to be more frequently detected (36.5 % flock prevalence) in flocks with digestive disorders ($p \leq 0.07$, Chi^2 test). In North America, PEMS has often been associated with the presence of TCoV (Brown *et al.*, 1997; Guy *et al.*, 1997; Loa *et al.*, 2000). The virus was detected in the epithelia of the intestinal tract and bursa of Fabricius (Nagaraja & Pomeroy, 1997). These anatomical locations correlate well with the samples identified in the present study as most contaminated by Fr TCoV (jejunum, cæcum and bursa of Fabricius, which thus appear as the samples to be preferred for the molecular diagnosis of Fr TCoV). Many viruses other than TCoV have also been reported to be associated with enteritis in turkeys. These include, both in the Americas and in Europe, type 2 turkey astroviruses, picorna-like viruses, reoviruses, rotaviruses, and adenoviruses (Andral *et al.*, 1985; Reynolds & Saif, 1986; Gough & Drury, 1998; Guy, 1998; Koci *et al.*, 2000; Cattoli *et al.*, 2007; Da Silva *et al.*, 2008; Pantin-Jackwood *et al.*, 2008a; Woolcock & Shivaprasad, 2008; Jindal *et al.*, 2009b). However, experimental studies with these agents inoculated alone (Schultz-Cherry *et al.*, 2000; Heggen-Peay *et al.*, 2002; Pantin-Jackwood *et al.*, 2008b; Gomaa *et al.*, 2009) or in combinations (Yu *et al.*, 2000; Jindal *et al.*, 2009a; Spackman *et al.*, 2010) seldom reproduce the whole range of signs observed in the diseased flocks. Determining whether the newly detected Fr TCoV plays a significant role in the increased incidence of digestive disorders reported in French turkey flocks since 2003 (Germain & Rousseau, 2005) will require isolation of the virus and further *in vivo* experimental studies. Conversely, whether the IBV isolates that appear

For Peer Review Only

phylogenetically related with Fr TCoV could possibly infect turkeys might deserve further experimental testing.

In order to evaluate precisely the genetic relationships between the newly detected sequences and other avian CoVs, the entire S protein gene was sequenced. To the authors' knowledge these are the first full length sequences of the S gene in a European TCoV. The three newly determined S sequences were highly homogeneous (98 % nt and aa identity) and the BLAST search revealed highest similarity with the published sequences of the S gene of NA TCoVs (approx. 65 % and 60 % nt and aa identity, respectively, E value = 0). Consistently, these French sequences clustered significantly with previously published NA TCoV sequences in all phylogenetic analysis (100 % bootstrap value in Figure 1A and data not shown). This confirmed that the newly detected viruses can indeed be considered as *bona fide* Fr TCoV strains.

Although significantly related to the previously published NA TCoV sequences, the Fr TCoV sequences exhibited striking original features. Indeed, the genetic region encoding the hypervariable S1 subunit proved very different from all sequences previously published, as NA TCoVs shared at least 86 % aa identity, but at best 42 % aa identity with Fr TCoVs. Fr TCoV sequences also appeared strikingly different from the recently released partial sequence of QCoV from Italy (Circella *et al.*, 2007). As shown in Figure 2, genetic divergence between Fr TCoV and previously released sequences was especially apparent between aa positions 1 to 196 with only 18 % aa identity with NA TCoVs or QCoV and several gaps necessary to maintain alignment. Such a region with low aa conservation is highly evocative of strong antigenic variation between North American and French viruses, as observed in IBV HVRs (Cavanagh *et al.*, 1988). Interestingly, a lack of cross neutralization in spite of high aa identity (96 - 98 %) in the S gene has been reported between NA TCoV strains VA-74/03, TX-1038/98, and IN-517/94 (Jackwood *et al.*, 2010). Such a result further supports the hypothesis of large antigenic variation among North American and French turkey viruses, as these two groups of viruses share a low aa identity, but it also indicates that cross neutralization studies between French viruses might be interesting in spite of the fact that the three Fr TCoV

For Peer Review Only

sequences share 96 % aa identity in their S1 region. The originality of the Fr TCoV sequences was confirmed by all phylogenetic analysis performed with the S, S1 and even S2 sequences, as Fr TCoVs always grouped as a very significant subcluster (100 % bootstrap value in Figure 1A and data not shown) apart from NA TCoVs, in contrast with the previously identified monophyletic TCoV group (Jackwood *et al.*, 2010). Although based on a limited number of S sequences from Fr TCoV, these results suggest that different lineages have evolved in North America and Europe. Such a genetic drift had already been suspected by comparing partial sequences encompassing several genes at the 3'end of the genome of NA TCoVs with the homologous sequences detected in turkeys in the UK (Loa *et al.*, 2006). Only 247-bp, spanning from the 3'-end of S gene to the 3' end of 3a gene, overlap between the previously released UK sequences (Cavanagh *et al.*, 2001) and those determined for the present study. The high identity scores between the UK sequence and Fr TCoVs (96 - 99 %) suggest both viruses could belong to the same group.

Possible recombination with IBV has been suggested recently as an evolutive mechanism important in the molecular epidemiology of NA TCoVs (Jackwood *et al.*, 2010). To test this hypothesis in Fr TCoVs, we sequenced fragments of genes upstream (RdRp) and downstream (N) of the S gene. Their phylogenetic analysis further supported that Fr TCoV represent a genetic lineage significantly different from NA TCoVs (Figures 1B and 1C). However, in contrast with the phylogenetic trees derived from the S sequences, which all grouped NA and Fr TCoVs in significant clusters formed of turkey viruses only, the trees derived from the N or RdRp sequences always included some IBV strains clustering within the NA or Fr TCoV lineages. The IBV strains most phylogenetically related to either NA TCoVs or Fr TCoVs were different, thus indicating that whereas recombination events could have occurred, they might have involved different IBV partners in the American or European TCoV lineages. Interestingly, the IBVs most related with the N and RdRp genes from Fr TCoVs appeared to be 4/91 and QX (ITA/90254/2005), respectively. These strains represented the last two emerged IBV serotypes in Europe (Worthington *et al.*, 2008). It is not known whether this observation is consistent with the lack of detection of Fr TCoV in

For Peer Review Only

samples collected before 1988 and indicates a possible recent recombination event involving the recent IBV serotypes.

Acknowledgements

The authors thank Prof Y.M. Saif (OARDC, Ohio, USA) and Prof C.C. Wu (Purdue University, Indiana, USA) for providing reference TCoV material and helpful advices on the project, as well as Dr JP. Picault and Mrs J. Lamandé (AFSSA Ploufragan) for constructive discussions on AvCoVs and providing IBV strains, respectively. The authors acknowledge the financial support of Conseil Régional de Bretagne, Conseil Régional des Pays de Loire, FranceAgrimer, and Comité Interprofessionnel de la Dinde Française (CIDEF), and the help of Pôle Agronomique de l'Ouest for project coordination.

For Peer Review Only

References

Andral, B., Toquin, D., L'Haridon, R., Jestin, A., Metz, M.H. & Rose, R. (1985). Les diarrhees du dindonneau: un bilan des recherches virales affectuees (rotavirus, reovirus, adenovirus, pseudopicornavirus). *Avian Pathology*, *14*, 147-162.

Ammayappan, A., Upadhyay, C., Gelb, J. Jr. & Vakharia,V.N. (2009). Identification of sequence changes responsible for the attenuation of avian infectious bronchitis virus strain Arkansas DPI. *Archives of Virology*, *154*, 495-499.

Barnes, H.J., Guy, J.S. & Vaillancourt, J.P. (2000). Poult enteritis complex. *Revue Scientifique et Technique Office Internationale Epizooties, 19*, 565-588.

Bosch, B.J., Martina, B.E., Van Der Zee, R., Lepault, J., Haijema, B.J., Versluis, C., Heck, A.J., De Groot, R., Osterhaus, A.D. & Rottier, P.J. (2004). Severe acute respiratory syndrome coronavirus (SARS-CoV) infection inhibition using spike protein heptad repeat-derived peptides. *Proceedings of the National Academy of Sciences of the United States of America*, *101*, 8455-8460.

Breslin, J.J., Smith, L.G., Fuller, F.J. & Guy,J.S. (1999). Sequence analysis of the turkey coronavirus nucleocapsid protein gene and 3' untranslated region identifies the virus as a close relative of infectious bronchitis virus. *Virus Research*, *65*, 187-193.

Brown, T.P., Garcia, A. & Kelley, L. (1997). Spiking mortality of turkey poults: 1. Experimental reproduction in isolation facilities. *Avian Diseases*, *41,* 604-609.

Callison, S.A., Hilt, D.A., Boynton, T.O., Sample, B.F., Robison, R., Swayne, D.E. & Jackwood, M.W. (2006). Development and evaluation of a real-time Taqman RT-PCR assay for the detection of infectious bronchitis virus from infected chickens. *Journal of Virological Methods*, *138*, 60-65.

Cao, J., Wu, C.C. & Lin, T.L. (2008). Complete nucleotide sequence of polyprotein gene 1 and genome organization of turkey coronavirus. *Virus Research*, *136*, 43-49.

Cattoli, G., De Battisti, C., Toffan, A., Salviato, A., Lavazza, A., Cerioli, M. & Capua, I. (2007). Co-circulation of distinct genetic lineages of astroviruses in turkeys and guinea fowl. *Archives of Virology*, *152*, 595-602.

Cavanagh, D., Davis, P.J. & Mockett, A.P. (1988). Amino acids within hypervariable region 1 of avian coronavirus IBV (Massachusetts serotype) spike glycoprotein are associated with neutralization epitopes. *Virus Research*, *11,* 141-150.

Cavanagh, D. & Gelb, J. (2008). Infectious bronchitis. In Y.M. Saif, A.M. Fadly, J.R. Glisson, L.R. McDougald, L.K. Nolan & D.E. Swayne (Eds.). *Diseases of Poultry* 12th edn (pp.117-135). Ames: Blackwell Publishing Professional.

Cavanagh, D., Mawditt, K., Gough, R., Picault, J.P. & Britton, P. (1998). Sequence analysis of strains of the 793/B genotype (CR88, 4/91) of IBV isolated between 1985 and 1997. In E.F. Kaleta & U. Heffels-Redmann (Eds.). Proceedings of the *International Symposium on Infectious Bronchitis and Pneumovirus Infections in Poultry* (p.252). Rauischholzhausen, Germany.

Cavanagh, D., Mawditt, K., Sharma, M., Drury, S.E., Ainsworth, H.L., Britton, P. & Gough, R.E. (2001). Detection of a coronavirus from turkey poults in Europe genetically related to infectious bronchitis virus of chickens. *Avian Pathology*, *30*, 355-368.

Cavanagh, D., Mawditt, K., Welchman Dde, B., Britton, P. & Gough, R.E. (2002). Coronaviruses from pheasants (Phasianus colchicus) are genetically closely related to coronaviruses of domestic fowl (infectious bronchitis virus) and turkeys. *Avian Pathology*, *31,* 81-93.

Circella, E., Camarda, A., Martella, V., Bruni, G., Lavazza, A. & Buonavoglia, C. (2007). Coronavirus associated with an enteric syndrome on a quail farm. *Avian Pathology*, *36,* 251-258.

Collisson, E.W., Pei, J., Dzielawa, J. & Seo, S.H. (2000). Cytotoxic T lymphocytes are critical in the control of infectious bronchitis virus in poultry. *Developmental and Comparative Immunology*, *24,* 187-200.

For Peer Review Only

Da Silva, S.E., Bonetti, A.M., Petrocelli, A.T., Ferrari, H.F., Luvizotto, M.C. & Cardoso, T.C. (2008). Detection of Turkey astrovirus in young poults affected with poult enteritis complex in Brazil. *The Journal of Veterinary Medical Science*, *70,* 629-631.

Davelaar, F.G., Kouwenhoven, B. & Burger, A.G. (1984). Occurrence and significance of infectious bronchitis virus variant strains in egg and broiler production in the Netherlands. *The Veterinary Quarterly*, *6*, 114-120.

de Groot, R. J., J. Ziebuhr, L. L. Poon, P. C. Woo, P. J. Talbot, P. J. M. Rottier, K. V. Holmes, R. Baric, S. Perlman, L. Enjuanes, & A. E. Gorbalenya (2008). Taxonomic proposal to the ICTV Executive Committee: Revision of the family 25 Coronaviridae. *http://talk.ictvonline.org/files/ictv_official_taxonomy_updates_since_the_8th_report/m/vertebrate-2008/1230/download.aspx.*

Domanska-Blicharz, K., Seroka, A., Lisowska, A., Tomczyk, G. & Minta, Z. (2010). Experimental infection with turkey coronavirus isolate. In H.M. Hafez (Ed.). Proceedings of the *8th International Symposium on Turkey Diseases* (p.). Berlin, Germany.

Dong, B.Q., Liu, W., Fan, X.H., Vijaykrishna, D., Tang, X.C., Gao, F., Li, F.L., Li, G.J., Zhang, J.X., Yang, L.Q., Poon, L.L.M., Zhang, S.Y., Peiris, J.S.M., Smith, G.J.D., Chen, H., & Guan, Y. (2007). Detection of a novel and highly divergent coronavirus from Asian Leopard Cats and Chinese Ferret Badgers in Southern China. *Journal of Virology*, *81*, 6920-6926.

Ducatez, M.F., Martin, A.M., Owoade, A.A., Olatoye, I.O., Alkali, B.R., Maikano, I., Snoeck, C.J., Sausy, A., Cordioli, P. & Muller, C.P. (2009). Characterization of a new genotype and serotype of infectious bronchitis virus in Western Africa. *Journal of General Virology*, *90,* 2679-2685.

Germain, M. & Rousseau, A. (2005). Syndrome enterite, frilosite, baisse de la croissance de la dinde: etude de cas terrain en France en 2004. Proceedings of the *6emes Journees de la Recherche Avicole* (p.408). St-Malo, France.

Gomaa, M.H., Barta, J.R., Ojkic, D. & Yoo, D. (2008). Complete genomic sequence of turkey coronavirus. *Virus Research*, *135,* 237-246.

Gomaa, M.H., Yoo, D., Ojkic, D. & Barta, J.R. (2009). Infection with a pathogenic turkey coronavirus isolate negatively affects growth performance and intestinal morphology of young turkey poults in Canada. *Avian Pathology*, *38*, 279-286.

Gonzalez, J.M., Gomez-Puertas, P., Cavanagh, D., Gorbalenya, A.E. & Enjuanes, L. (2003). A comparative sequence analysis to revise the current taxonomy of the family Coronaviridae. *Archives of Virology*, *148,* 2207-2235.

Gough, R.E. & Drury, S.E. (1998). Detection of viruses associated with enteric problems in turkeys. Proceedings of the *1st International Symposium on Turkey Diseases* (p.218). Berlin, Germany.

Gough, R.E., Drury, S.E., Culver, F., Britton, P. & Cavanagh, D. (2006). Isolation of a coronavirus from a green-cheeked Amazon parrot *(Amazon viridigenalis Cassin). Avian Pathology*, *35*, 122-126.

Guindon, S. & Gascuel, O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology*, *52*, 696-704.

Guionie, O., Toquin, D., Sellal, E., Bouley, S., Zwingelstein, F., Allee, C., Bougeard, S., Lemiere, S. & Eterradossi, N. (2007). Laboratory evaluation of a quantitative real-time reverse transcription PCR assay for the detection and identification of the four subgroups of avian metapneumovirus. *Journal of Virological Methods*, *139,* 150-158.

Guy, J.S. (1998). Virus infections of the gastrointestinal tract of poultry. *Poultry Science*, *77*, 1166-1175.

Guy, J.S. (2008). Turkey coronavirus enteritis. In Y.M. Saif, A.M. Fadly, J.R. Glisson, L.R. McDougald, L.K. Nolan & D.E. Swayne (Eds.). *Diseases of Poultry* 12th edn (pp.330-338). Ames: Blackwell Publishing Professional.

Guy, J.S., Barnes, H.J., Smith, L.G. & Breslin, J. (1997). Antigenic characterization of a turkey coronavirus identified in poult enteritis- and mortality syndrome-affected turkeys. *Avian Diseases*, *41*, 583-590.

For Peer Review Only

Heggen-Peay, C.L., Qureshi, M.A., Edens, F.W., Sherry, B., Wakenell, P.S., O'Connell, P.H. & Schat, K.A. (2002). Isolation of a reovirus from poult enteritis and mortality syndrome and its pathogenicity in turkey poults. *Avian Diseases*, *46*, 32-47.

Hughes, L.A., Savage, C., Naylor, C., Bennett, M., Chantrey, J. & Jones, R. (2009). Genetically diverse coronaviruses in wild bird populations of northern England. *Emerging Infectious Diseases*, *15*, 1091-1094.

Jackwood, M.W., Boynton, T.O., Hilt, D.A., McKinley, E.T., Kissinger, J.C., Paterson, A.H., Robertson, J., Lemke, C., McCall, A.W., Williams, S.M., Jackwood, J.W. & Byrd, L.A. (2010). Emergence of a group 3 coronavirus through recombination. *Virology*, *398,* 98-108.

Jindal, N., Patnayak, D.P., Ziegler, A.F., Lago, A. & Goyal, S.M. (2009a). Experimental reproduction of poult enteritis syndrome: clinical findings, growth response, and microbiology. *Poultry Science*, *88*, 949-958.

Jindal, N., Patnayak, D.P., Ziegler, A.F., Lago, A. & Goyal, S.M. (2009b). A retrospective study on poult enteritis syndrome in Minnesota. *Avian Diseases*, *53*, 268-275.

Jonassen, C.M., Kofstad, T., Larsen, I.L., Lovland, A., Handeland, K., Follestad, A. & Lillehaug, A. (2005). Molecular identification and characterization of novel coronaviruses infecting graylag geese (Anser anser), feral pigeons (Columbia livia) and mallards (Anas platyrhynchos). *Journal of General Virology*, *86*, 1597-1607.

Jungherr, E.L., Chomiak, T.W. & Luginbuhl, R.E. (1956). Immunologic differences in strains of infectious bronchitis. Proceedings of the *60th Annual Meeting of the United States Livestock Sanitary Association* (p.203). Chicago, IL.

Koci, M.D., Seal, B.S. & Schultz-Cherry, S. (2000). Molecular characterization of an avian astrovirus. *Journal of Virology*, *74,* 6173-6177.

Krempl, C., Ballesteros, M.L., Zimmer, G., Enjuanes, L., Klenk, H.D. & Herrler, G. (2000). Characterization of the sialic acid binding activity of transmissible gastroenteritis coronavirus by analysis of haemagglutination-deficient mutants. *Journal of General Virology*, *81*, 489-496.

Kubo, H., Yamada, Y.K. & Taguchi, F. (1994). Localization of neutralizing epitopes and the receptor-binding site within the amino-terminal 330 amino acids of the murine coronavirus spike protein. *Journal of Virology*, *68*, 5403-5410.

Kumar, S., Nei, M., Dudley, J. & Tamura, K. (2008). MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences. *Briefings in Bioinformatics*, *9*, 299-306.

Lin, T.L., Loa, C.C. & Wu, C.C. (2004). Complete sequences of 3' end coding region for structural protein genes of turkey coronavirus. *Virus Research*, *106*, 61-70.

Liu, S., Chen, J., Kong, X., Shao, Y., Han, Z., Feng, L., Cai, X., Gu, S. & Liu, M. (2005). Isolation of avian infectious bronchitis coronavirus from domestic peafowl (Pavo cristatus) and teal (Anas). *Journal of General Virology*, *86,* 719-725.

Loa, C.C., Lin, T.L., Wu, C.C., Bryan, T.A., Thacker, H.L., Hooper, T. & Schrader, D. (2000). Detection of antibody to turkey coronavirus by antibody-capture enzyme-linked immunosorbent assay utilizing infectious bronchitis virus antigen. *Avian Diseases*, *44,* 498-506.

Loa, C.C., Wu, C.C. & Lin, T.L. (2006). Comparison of 3'-end encoding regions of turkey coronavirus isolates from Indiana, North Carolina, and Minnesota with chicken infectious bronchitis coronavirus strains. *Intervirology*, *49,* 230-238.

Masters, P.S. (2006). The molecular biology of coronaviruses. *Advances in Virus Research*, *66,* 193-292.

Meir, R., Maharat, O., Farnushi, Y. & Simanov, L. (2010). Development of a real-time TaqMan RT-PCR assay for the detection of infectious bronchitis virus in chickens, and comparison of RT-PCR and virus isolation. *Journal of Virological Methods*, *163,* 190-194.

Mihindukulasuriya, K.A., Wu, G., St Leger, J., Nordhausen, R.W. & Wang, D. (2008). Identification of a novel coronavirus from a beluga whale by using a panviral microarray. *Journal of Virology*, *82*, 5084-5088.

For Peer Review Only

Mondal, S.P. & Cardona, C.J. (2007). Genotypic and phenotypic characterization of the California 99 (Cal99) variant of infectious bronchitis virus. *Virus Genes*, *34,* 327-341.

Moreno Martin, A., Vinco, L., Cordioli, P. & Lavazza, A. (2002). Diagnosis of turkey viral enteritic diseases by electron microscopy and identification of coronavirus in a case of turkey enteritis. In H.M. Hafez (Ed.). Proceedings of the *4th International Symposium on Turkey Diseases* (p.114). Berlin, Germany.

Nagaraja, K.V. & Pomeroy, B.S. (1997). Coronaviral enteritis of turkeys (bluecomb disease). In B.W. Calnek, H.J. Barnes, C.W. Beard, L.R. McDougald & Y.M. Saif (Eds.). *Diseases of Poultry* 10th edn (pp.686-692). Ames: Iowa State University Press.

Pantin-Jackwood, M.J., Day, J.M., Jackwood, M.W. & Spackman, E. (2008a). Enteric viruses detected by molecular methods in commercial chicken and turkey flocks in the United States between 2005 and 2006. *Avian Diseases*, *52,* 235-244.

Pantin-Jackwood, M.J., Spackman, E. & Day, J.M. (2008b). Pathogenesis of type 2 turkey astroviruses with variant capsid genes in 2-day-old specific pathogen free poults. *Avian Pathology*, *37*, 193-201.

Picault, J.P., Drouin, P., Guittet, M., Bennejean, G., Protais, J., L'Hospitalier, R., Gillet, J.P., Lamande, J. & Bachelier, A.L. (1986). Isolation, characterisation and preliminary cross-protection studies with a new pathogenic avian infectious bronchitis virus (strain PL-84084). *Avian Pathology*, *15,* 367-383.

Picault, J.P. (1987). Contribution à l'étude des coronaviroses dans l'espèce *Gallus gallus* en France. *PhD Thesis, University Claude Bernard Lyon I* (p.162). Lyon, France.

Qian, D.H., Zhu, G.J., Wu, L.Z. & Hua, G.X. (2006). Isolation and characterization of a coronavirus from pigeons with pancreatitis. *American Journal of Veterinary Research*, *67*, 1575-1579.

Reynolds, D.L. & Saif, Y.M. (1986). Astrovirus: a cause of an enteric disease in turkey poults. *Avian Diseases*, *30,* 728-735.

Rota, P.A., Oberste, M.S., Monroe, S.S., Nix, W.A., Campagnoli, R., Icenogle, J.P., Penaranda, S., Bankamp, B., Maher, K., Chen, M.H., Tong, S., Tamin, A., Lowe, L., Frace, M., DeRisi, J.L., Chen, Q., Wang, D., Erdman, D.D., Peret, T.C., Burns, C., Ksiazek, T.G., Rollin, P.E., Sanchez, A., Liffick, S., Holloway, B., Limor, J., McCaustland, K., Olsen-Rasmussen, M., Fouchier, R., Gunther, S., Osterhaus, A.D., Drosten, C., Pallansch, M.A., Anderson, L.J. & Bellini, W.J. (2003). Characterization of a novel coronavirus associated with severe acute respiratory syndrome. *Science*, *300*, 1394-1399.

Schultz-Cherry, S., Kapczynski, D.R., Simmons, V.M., Koci, M.D., Brown, C. & Barnes, H.J. (2000). Identifying agent(s) associated with poult enteritis mortality syndrome: importance of the thymus. *Avian Diseases*, *44,* 256-265.

Spackman, E., Day, J.M. & Pantin-Jackwood, M.J. (2010). Astrovirus, reovirus, and rotavirus concomitant infection causes decreased weight gain in broad-breasted white poults. *Avian Diseases*, *54,* 16-21.

Tang, M., Wang, H., Zhou, S. & Tian, G. (2008). Enhancement of the immunogenicity of an infectious bronchitis virus DNA vaccine by a bicistronic plasmid encoding nucleocapsid protein and interleukin-2. *Journal of Virological Methods*, *149*, 42-48.

Toquin, D., Guionie, O., Jestin, V., Zwingelstein, F., Allee, C. & Eterradossi, N. (2006). European and American subgroup C isolates of avian metapneumovirus belong to different genetic lineages. *Virus Genes*, *32*, 97-103.

Webster, R.G., Bean, W.J., Gorman, O.T., Chambers, T.M. & Kawaoka, Y. (1992). Evolution and ecology of influenza A viruses. *Microbiological Reviews, 56,* 152-179.

Winter, C., Schwegmann-Wessels, C., Neumann, U. & Herrler, G. (2008). The spike protein of infectious bronchitis virus is retained intracellularly by a tyrosine motif. *Journal of Virology*, *82*, 2765-2771.

Wong, S.K., Li, W., Moore, M.J., Choe, H. & Farzan, M. (2004). A 193-amino acid fragment of the SARS coronavirus S protein efficiently binds angiotensin-converting enzyme 2. *Journal of Biological Chemistry*, *279,* 3197-3201.

Woo, P.C., Lau, S.K., Lam, C.S., Lai, K.K., Huang, Y., Lee, P., Luk, G.S., Dyrting, K.C., Chan, K.H. & Yuen, K.Y. (2009). Comparative analysis of complete genome

For Peer Review Only

sequences of three avian coronaviruses reveals a novel group 3c coronavirus. *Journal of Virology*, *83,* 908-917.

Woolcock, P.R. & Shivaprasad, H.L. (2008). Electron microscopic identification of viruses associated with poult enteritis in turkeys grown in California 1993-2003. *Avian Diseases*, *52,* 209-213.

Worthington, K.J., Currie, R.J. & Jones, R.C. (2008). A reverse transcriptase-polymerase chain reaction survey of infectious bronchitis virus genotypes in Western Europe from 2002 to 2006. *Avian Pathology*, *37*, 247-257.

Yamada, Y. & Liu, D.X. (2009). Proteolytic activation of the spike protein at a novel RRRR/S motif is implicated in furin-dependent entry, syncytium formation, and infectivity of coronavirus infectious bronchitis virus in cultured cells. *Journal of Virology*, *83,* 8744-8758.

Yu, M., Ismail, M.M., Qureshi, M.A., Dearth, R.N., Barnes, H.J. & Saif, Y.M. (2000). Viral agents associated with poult enteritis and mortality syndrome: the role of a small round virus and a turkey coronavirus. *Avian Diseases*, *44,* 297-304.

Zuniga, S., Sola, I., Alonso, S. & Enjuanes, L. (2004). Sequence motifs involved in the regulation of discontinuous coronavirus subgenomic RNA synthesis. *Journal of Virology*, *78*, 980-994.

For Peer Review Only

Table 1. *Strains used for testing the specificity of the AvCoV RT-qPCR*

Virus	Origin	Reference in authors' laboratory	Reference or source
IBV strain 84084	France	PL 84084/5.5	Picault *et al.*, (1986)
IBV strain 84221	France	CR 84221/6.1	Picault, (1987)
IBV strain 84222	France	CR 84222/6.1	Picault, (1987)
IBV strain 85131	France	CR 85131/13.1	Picault, (1987)
IBV strain 88061	France	CR 88061/8.3	Picault *et al.*, (1988)
IBV strain 88121	France	CR 88121/11.4	Picault *et al.*, (1988)
IBV vaccine (Massachusetts type)	-	CR B48/1.1	Cevac® Mass L (CEVA Santé animale, France)
IBV strain 4/91	UK	CR 4-91/3.1	Parsons *et al.* (1992)
IBV vaccine (D274b type)	Netherlands	CR D274b/6.3	Davelaar *et al.*, (1984)
IBV strain D212	Netherlands	CR D212/56	Davelaar *et al.*, (1984)
IBV strain D3128	Netherlands	CR D3128/64	Davelaar *et al.*, (1984)
IBV strain D3896	Netherlands	CR D3896/52	Davelaar *et al.*, (1984)
IBV strain Connecticut	USA	CR Conn/5.1	Jungherr *et al.*, (1956)
IBV strain Beaudette	USA	VBBPa/3.8	Picault, (1987)
TCoV	USA	TCoV INP4/1.1	Kind gift from Profs. CC Wu & YM Saïf

ew Only

Table 2. *Sequence and position of the oligonucleotides used in this study*

Name[a]	Sequence (5' → 3')[b]	Position in TCoV genome[c]	Use
Ncor800f	CGTGTTACGGCAATGCTCAA	26852-26871	RT-qPCR
Ncor860r	CGTCACTCTGCTTCCAAAAAGAC	26917-26895	RT-qPCR and RT/PCR for sequencing N gene
Ncor830p	fam-CCCTAGCAGCCATGC-mgb	26878-26892	RT-qPCR probe
Ncor1f	CACCATGGCAAGCGGTAAGGC	26050-26070	PCR for sequencing N gene
Scor+260f	GAATGCGTCTTCTTCAGAAGC	20096-20116	PCR for cloning S gene
Scor-150r	TGTGCCAAAGCAGAAGTCTAG	24150-24130	RT/PCR for cloning S gene
POLcor1900f	AAGTGTGATAGAGCAATGCC	14196-14215	PCR for sequencing RdRp gene
POLcor2590r	CTCCATAACAGCCACAGG	14906-14889	RT/PCR for sequencing RdRp gene

[a] f, forward primers; r, reverse primers; p, probe.

[b] fam, 6-carboxyfluorescein; mgb, minor groove binder.

[c] Relative to the genome of TCoV (NC_010800).

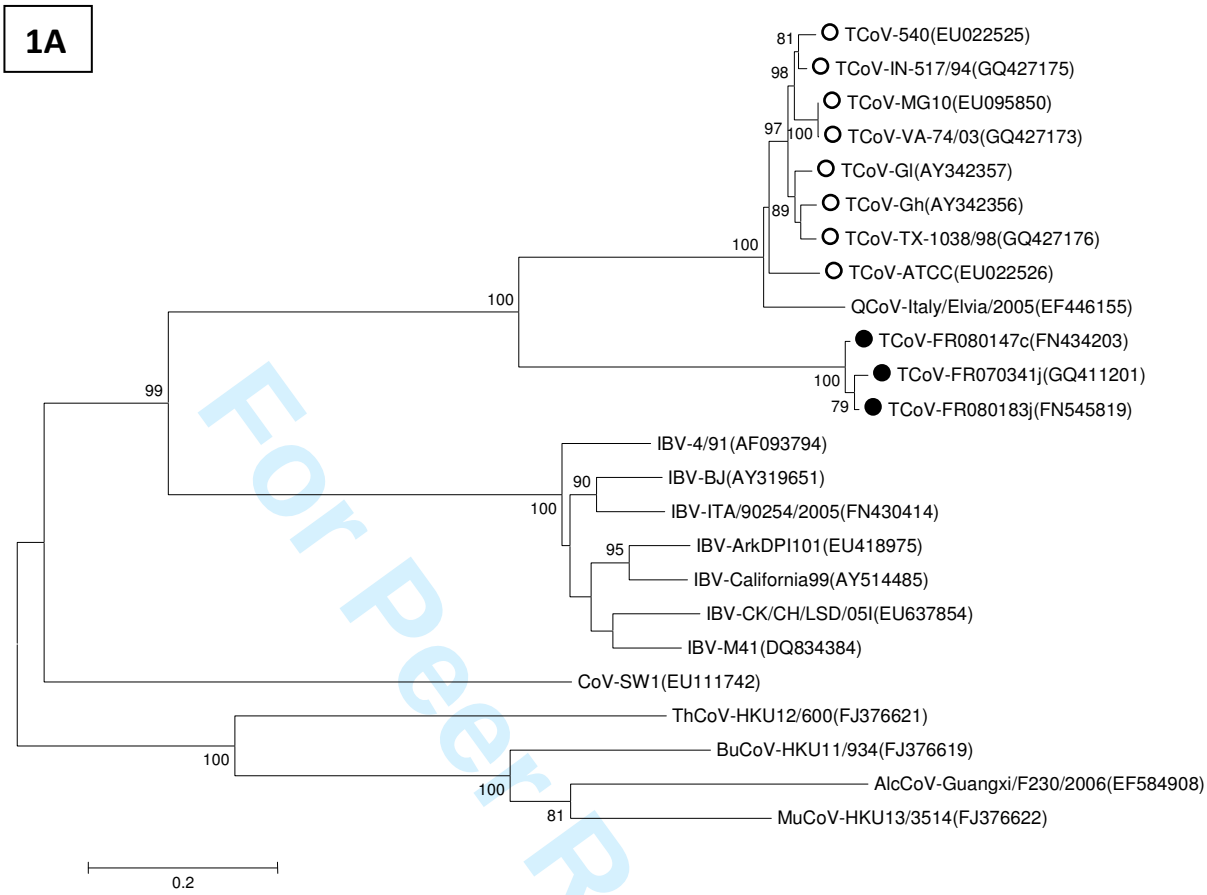
Review Only

Table 3. *Virus sequences included in the phylogenetic study.*

Virus	Origin	Accession No[a]	Reference
QCoV Italy/Elvia/2005	Italy	EF446155 /na[b]/na	Circella *et al.*, 2007
IBV strain 4/91	UK	AF093794/EU780081/FN811147	Callisson *et al.*, 2001 / Meir *et al.*, 2010 / this study
IBV strain BJ	China	AY319651	Jin *et al.*, unpubl.
IBV strain ITA/90254/2005	(Italy ~~?~~)	FN430414	Ducatez *et al.*, 2009
IBV strain Ark DPI101	USA	EU418975	Ammayapan *et al.*, 2009
IBV strain California99	USA	AY514485	Mondal *et al.*, 2007
IBV strain CK/CH/LSD/05I	China	EU637854	Wang *et al.*, unpubl.
IBV strain M41	USA	DQ834384	Mondal *et al.*, unpubl.
TCoV strain 540	USA	EU022525	Cao *et al.*, 2008
TCoV strain IN517/94	USA	GQ427175	Jackwood *et al.*, 2010
TCoV strain MG10	USA	EU095850	Gooma *et al.*, 2008
TCoV strain VA74/03	USA	GQ427173	Jackwood *et al.*, 2010
TCoV strain GI	USA	AY342357	Jackwood & Hilt, unpubl.
TCoV strain Gh	USA	AY342356	Jackwood & Hilt, unpubl.
TCoV strain TX-1038/98	USA	GQ427176	Jackwood *et al.,* 2010
TCoV strain ATCC	USA	EU022526	Cao *et al.*, 2008
TCoV strain NC95	USA	na/ AF111997/ na	Breslin *et al.*, 1999
CoV SW1	-	EU111742	Mihindukulasuriya *et al.*, 2008
ThCoV strain HKU12/600	Hong Kong	FJ376621	Woo *et al.*, 2009
BuCoV strain HKU11/934	Hong Kong	FJ376619	Woo *et al.*, 2009
MuCoV strain HKU13/3514	Hong Kong	FJ376622	Woo *et al.*, 2009
AlcCoV-Guangxi / F230/2006	China	EF584908	Dong *et al.,* 2007

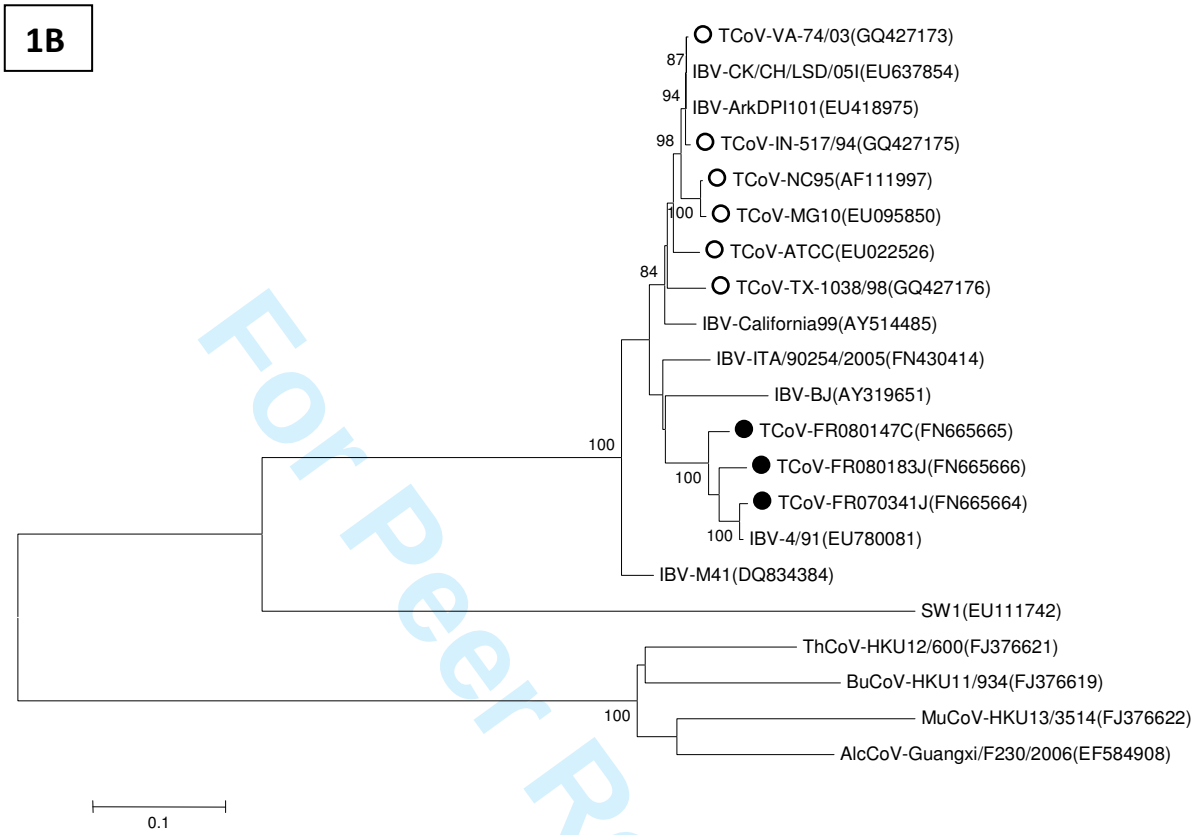
[a] A single Acc No (whole genome) or three Acc Nos (in the S/N/RdRp order) are indicated
[b] na = sequence is not available

1A



For Peer Review Only

1B



For Peer Review Only

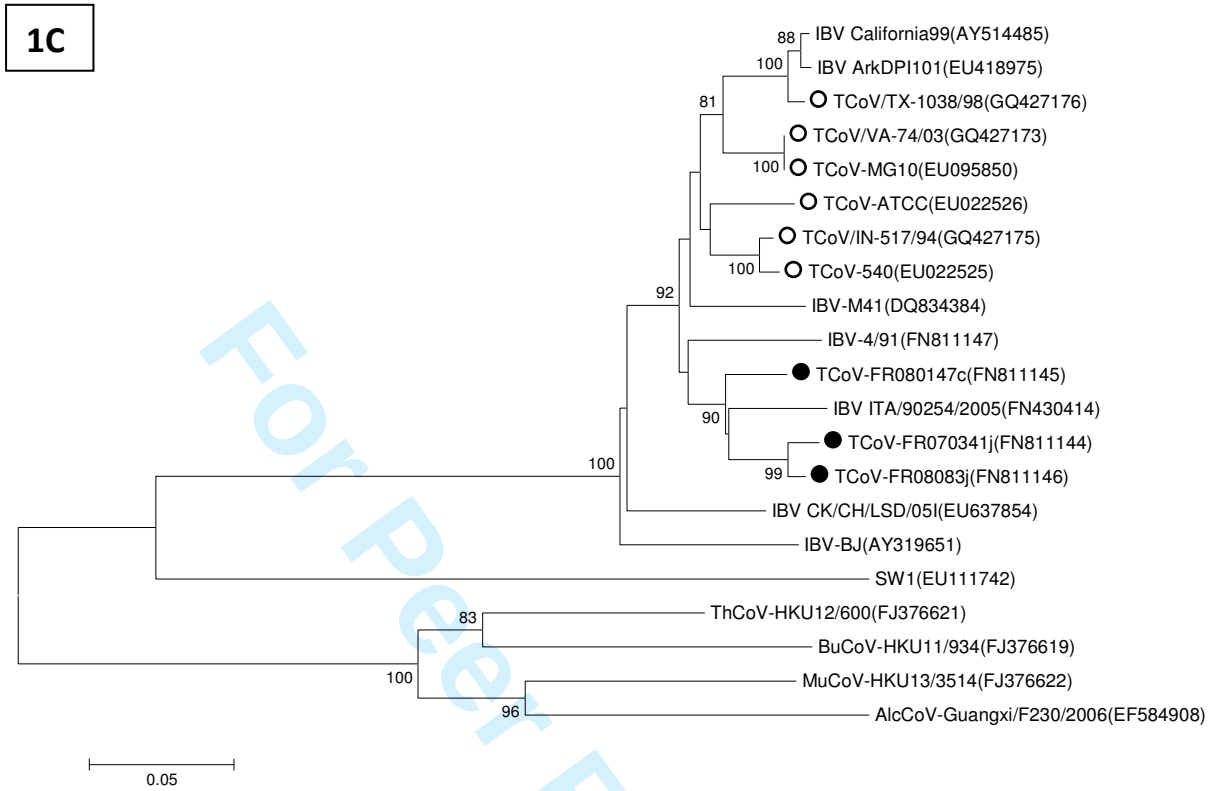


Figure 1. *Consensus phylogenetic tree resulting from the analysis of the nucleotide sequences encoding the S1 subunit of the spike glycoprotein (1A), of a 795 bp fragment of the N gene (1B), or of a 615 bp fragment of RdRp gene (1C) in Fr TCoV strains (black dots), the NA TCoV (white dots) and other representative gammacoronaviruses (accession number in parenthesis). The trees were constructed by using the neighbor-joining method and Kimura 2-parameter model. Bootstrap values and the resulting consensus tree were calculated from 1000 trees. Only bootstrap values > 75 % are indicated. The geographical origin and literature reference of the sequences are indicated in Table 3.*

For Peer Review Only

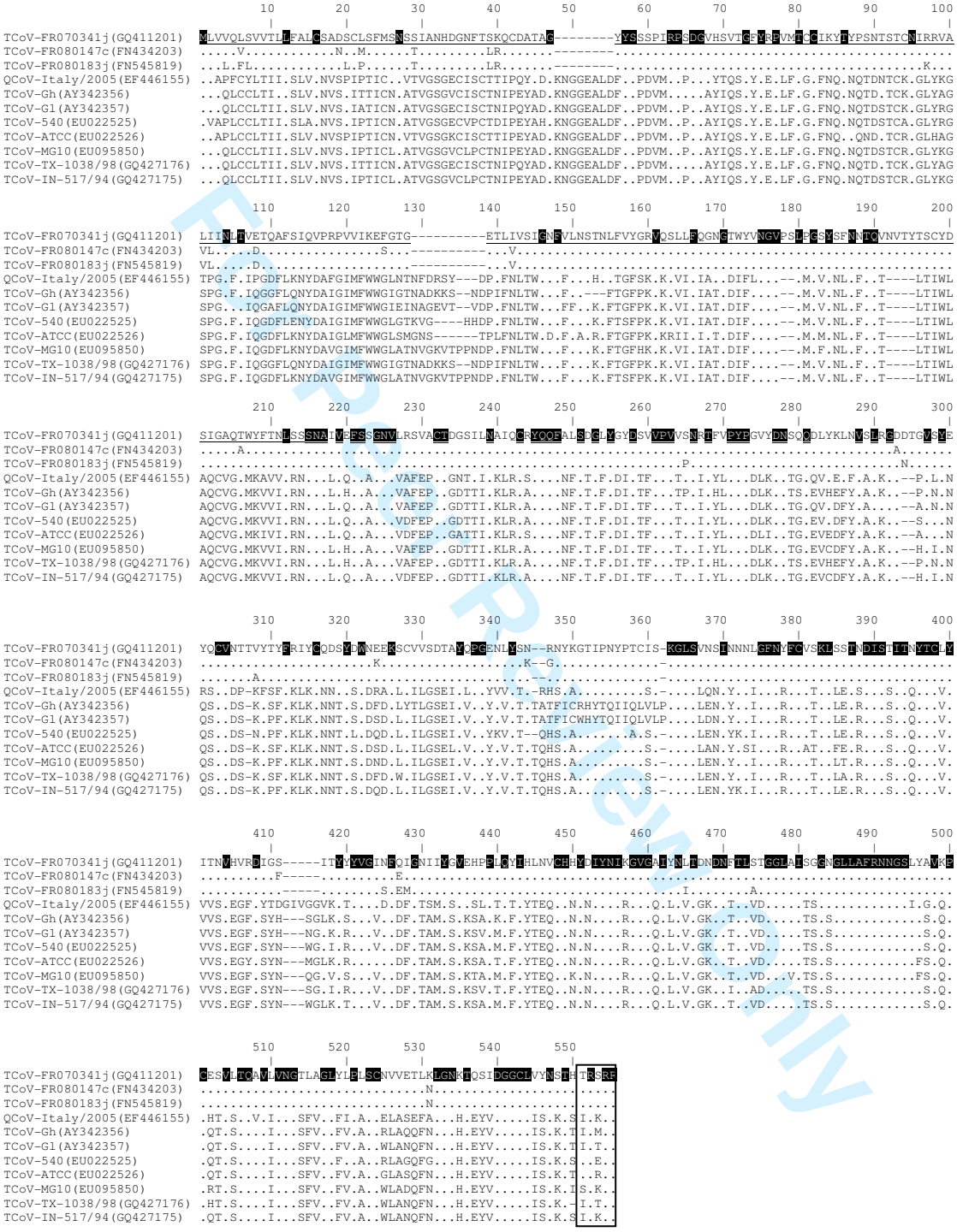


Figure 2. *Amino acid alignment of the S1 subunit of TCoV-FR070341j in comparison with that of the two other Fr TCoV, QCoV and several NA TCoV isolates. Black shading represents conserved amino acids among all TCoV strains. Dots indicate aa residues that are*

identical with the FR070341j sequence. The highly variable S1a region is underlined and the putative cleavage site XRXRR stretch is boxed.

For Peer Review Only