Avian Pathology, 2015
Vol. 44, No. 3, 182–191, http://dx.doi.org/10.1080/03079457.2015.1022857

Taylor & Francis
Taylor & Francis Group

ORIGINAL ARTICLE

Analysis of the complete genomic sequences of two virus subpopulations of the Australian infectious bronchitis virus vaccine VicS

José A. Quinteros[1]*, Philip F. Markham[1], Sang-Won Lee[1,2], Kylie A. Hewson[3], Carol A. Hartley[1], Alistair R. Legione[1], Mauricio J. C. Coppo[1], Paola K. Vaz[1], and Glenn F. Browning[1]

[1]*Asia-Pacific Centre for Animal Health, Faculty of Veterinary and Agricultural Sciences, The University of Melbourne, Parkville, Victoria, Australia,* [2]*College of Veterinary Medicine, Konkuk University, Seoul, Republic of Korea, and* [3]*Asia-Pacific Centre for Animal Health, Faculty of Veterinary and Agricultural Sciences, The University of Melbourne, Werribee, Victoria, Australia*

Although sequencing of the 3′ end of the genome of Australian infectious bronchitis viruses (IBVs) has shown that their structural genes are distinct from those of IBVs found in other countries, their replicase genes have not been analysed. To examine this, the complete genomic sequences of the two subpopulations of the VicS vaccine, VicS-v and VicS-del, were determined. Compared with VicS-v, the more attenuated VicS-del strain had two non-synonymous changes in the non-structural protein 6 (nsp6), a transmembrane (TM) domain that may participate in autocatalytic release of the 3-chymotrypsin-like protease, a polymorphic difference at the end of the S2 gene, which coincided with the body transcription-regulating sequence (B-TRS) of mRNA 3 and a truncated open reading frame for a peptide encoded by gene 4 (4b). These genetic differences could be responsible for the differences between these variants in pathogenicity *in vivo*, and replication *in vitro.* Phylogenetic analysis of the whole genome showed that VicS-v and VicS-del did not cluster with strains from other countries, supporting the hypothesis that Australian IBV strains have been evolving independently for some time, and analyses of individual polymerase peptide and S glycoprotein genes suggested a distant common ancestor with no recent recombination. This study suggests the potential role of the TM domain in nsp6, the integrity of the S2 protein and the B-TRS 3, and the putative accessory protein 4b, as well as the 3′ untranslated region, in the virulence and replication of IBV and has provided a better understanding of relationships between the Australian vaccine strain of IBV and those used elsewhere.

Introduction

Infectious bronchitis is a highly contagious and widespread disease of chickens and is a major ongoing problem in all countries with an intensive poultry industry (Sapats *et al.*, 1996; Casais *et al.*, 2003). The aetiological agent is infectious bronchitis virus (IBV), a member of the family *Coronaviridae* (Beaudette & Hudson, 1937; Tyrell, 1968), subfamily *Coronavirinae* and genus *Gammacoronavirus* (group 3; Cavanagh *et al.*, 1994). While IBV has been detected in other avian species, including guinea fowl (*Numida meleagridis*), partridge (*Alectoris* sp.), peafowl (*Pavo cristato*) and the blue-winged teal (*Anas* sp.), clinically detectable disease has only been observed in chickens (Liu *et al.*, 2005; Cavanagh, 2007; Sun *et al.*, 2007).

IBV has a positive sense, single-stranded RNA genome approximately 27.6 kb in length (Tannock, 1973; Watkins *et al.*, 1975; Schochetman *et al.*, 1977; Jackwood & de Witt, 2013) with the general organization 5′ untranslated region (UTR) - 1 a, ab (or polymerase genes) - S - 3 a, b, c (E) - M - 4 b, c - 5 a, b - N - 6b - 3′ UTR (Cavanagh, 2007; Ammayappan *et al.*, 2008). The genes encoding the structural proteins are S (spike glycoprotein), E (envelope glycoprotein), M (membrane glycoprotein) and N (nucleocapsid protein) (Ignjatovic *et al.*, 1997; Cavanagh, 2007), while the 3a, 3b, 4b, 4c, 5a, 5b and 6b genes encode accessory proteins (Liu *et al.*, 1991; Liu & Inglis, 1992a, b; Cao *et al.*, 2008; Hewson *et al.*, 2011; Bentley *et al.*, 2013), the functions of which are still unknown (Liu *et al.*, 2008). The S glycoprotein is a virion surface, rod-shaped protein that is post-translationally cleaved into two subunits, S1 and S2 (Cavanagh, 1981; Stern & Sefton, 1982; Cavanagh, 2007).

During IBV infection, virions bind to the target cell receptors and release the viral genome into the cytoplasm of the host cell (Ziebuhr, 2005; Knoops *et al.*, 2008). The polyproteins 1a and 1ab (pp1a and pp1ab) are then translated, and the papain-like (PLP) and 3-chymotrypsin-like (3CL) proteases are released from these polyproteins after an autocatalytic process. These proteases initiate the cleavage *in trans* of the 15 peptides contained in pp1a/pp1ab. These peptides, probably together with some cellular proteins, form

*To whom correspondence should be addressed: E-mail: jose.quinteros@unimelb.edu.au
(*Received 30 September 2014; accepted 6 February 2015*)

the replication–transcription complex. This initiates the transcription of a series of 3′ nested subgenomic RNAs that are translated into the viral structural and accessory proteins (Gorbalenya *et al.*, 1989; Liu & Brown, 1995; Tibbles *et al.*, 1996; Liu *et al.*, 1998; Ziebuhr *et al.*, 2001; Ziebuhr, 2005).

In Australia, three genotypic subgroups of IBV have been distinguished (Hewson *et al.*, 2012). Subgroup 1 or "classical" strains were isolated between 1962 and 1984 and include the subtype B vaccines, which were classified using serological methods (Wadey & Faragher, 1981; Ignjatovic & Galli, 1995) and vaccine-related strains. Subgroup 2 or "novel" strains emerged between 1988 and 1994, while subgroup 3 strains, which emerged around 2002, are recombinants derived from subgroup 1 and 2 strains (Ignjatovic *et al.*, 2006; Hewson *et al.*, 2012).

Four different IBV strains are included in the commercially attenuated vaccines available for the control of infectious bronchitis in Australia. Three of these, VicS, I (Inghams) and S (Steggles), belong to the same serotype (subtype B), while A3 (Armidale) is serotypically distinct (subtype C; Ignjatovic *et al.*, 2006; Hewson *et al.*, 2012). Previous studies (Hewson *et al.*, 2012) have demonstrated that the VicS vaccine contains two subpopulations, referred to as VicS-v (the predominant subpopulation in the vaccine) and VicS-del, which has a 40-nucleotide deletion in the 3′ UTR and multiple single-nucleotide differences in the structural protein genes compared with VicS-v (Hewson *et al.*, 2012). These viral subpopulations differ in their ability to replicate *in vitro* and in their pathogenicity *in vivo*. In VicS-v inoculated groups, all the birds exhibit clinical signs, including head shaking and eye scratching/irritation, from day 2 to day 8 post-inoculation, whereas in the VicS-del inoculated group the clinical signs are less severe and only seen between days 6 and 9 post-inoculation. Birds inoculated with VicS-v also have significantly more severe tracheal lesions than those inoculated with VicS-del.

Over the last few years, a number of complete genome sequences of many IBVs have been determined and the sequences of more than a hundred are now available in GenBank (National Center for Biotechnology Information, NCBI). However, the complete genome of an IBV isolated in Australia has not been determined. As the pathogenicity and replication of the two IBV subpopulations, VicS-v and VicS-del, in the VicS vaccine, differ *in vivo* and *in vitro* under laboratory conditions, the complete genome sequences of these two variants were compared in order to identify potential novel molecular determinants of pathogenicity and replication in IBV.

Materials and Methods

Virus. The VicS-v and VicS-del strains of IBV, purified previously in our laboratory, were propagated in specific pathogen-free embryonated hen eggs (Australian SPF Services, Woodend, Victoria, Australia), as described previously (Hewson *et al.*, 2009). After 8–9 days of incubation, the allantoic cavities of the specific pathogen-free embryonated eggs were inoculated with one of the two strains. After 48 h of incubation, the eggs were chilled for 12–18 h at 4°C and the allantoic fluid collected and stored at –80°C.

Viral purification and nucleic acid extraction. The allantoic fluid was clarified as previously described (Lougovskaia *et al.*, 2002), with slight modifications. The gradients were fractioned into 1 mL volumes, and the fractions containing the highest concentration of virus were determined by PCR using primers and a protocol described previously (Hewson *et al.*, 2009). After centrifugation at 90,000 × *g* for 1 h at 4°C the viral pellets were resuspended in 250–350 µL of Tris-buffered saline (pH 7.4; 0.02-M Tris–HCl, 0.15-M NaCl) and stored at –80°C. Viral RNA was extracted using RNeasy™ extraction kits (Qiagen Pty. Ltd., Chadstone, Victoria, Australia), following the manufacturer's protocol, and then stored at –80°C.

Complete genome sequencing. Complete genome sequencing of the VicS-v and VicS-del strains of IBV was performed using the PGM Ion Torrent™ platform (Life Technologies Australia Pty. Ltd., Mulgrave, Victoria, Australia). RNA was fragmented using RNase III (Life Technologies) and purified using magnetic beads. To construct the libraries, the RNA was reverse-transcribed using SuperScript™ III (Life Technologies). The cDNA was purified and amplified following the manufacturer's protocols and then sequenced using the 314 chip, 200 base sequencing Ion OneTouch™ Kit v2.

All the reads that matched to chicken ribosomal RNA and mitochondrial genome sequences were discarded. For VicS-v, the remaining reads were mapped to the genomes of the Beaudette strain [Genbank accession number (GAN) NC_001451], two vaccine strains from USA (Conn46 1996 and Massachusetts, GAN FJ904716 and GQ504724, respectively) and two strains from China (SAIBK and SC021202, GAN DQ288927 and EU714029, respectively) using Geneious™ version 6.1.4 (Biomatters Ltd., Auckland, New Zealand). Because of the high level of sequence diversity in the S gene, the readings were also mapped to the previously determined sequence of structural protein genes from the VicS vaccine (GAN JN176213). For VicS-del, the readings were mapped to the genome of VicS-v and the previously determined sequence of the structural protein gene region of VicS-del (GAN JN983807). All gaps and ambiguous sequences were corrected using Sanger sequencing and the Big Dye™ Terminator v3.1 sequencing protocol (Life Technologies), following the manufacturer's instructions.

Sequence and phylogenetic tree analysis. Twelve sequences of IBV from the USA (GAN NC_001451, Beaudette RS; GQ504725, Mass 41 vaccine; AJ311317, Beaudette CK; FJ904723, Mass 41 1985; GU393332, Delaware 072; GQ504723, Georgia 1998; GQ504721, Arkansas vaccine; AY514485, California 99; FJ904714, Cal 1995; FJ904716, Conn 46 1996; GU393336, Holte; and GU393338, JMK), five of which were vaccine strains, 11 strains from China (GAN JX195178, Ck/Ch/LDL/97I; EU637854, Ck/Ch/LSD/05I; DQ288927, SAIBK; EU714029, SC021202; AY319651, BJ; EU526388, A2; HM245923, DY07; HM245924, CQ04-1; JX195176, Ck/Ch/LZJ/111113; KC119407, Ck/Ch/LGD/120724; and HQ850618, Gx-YL9), two from The Netherlands (EU817497, H52 and GU393335, H120), and one each from South Korea (GAN JQ977697, SNU8067), Sweden (GAN JQ088078, Ck/Swe/065846/10), Nigeria (GAN FN430415, Ibadan), Italy (FN430414, Ita/90254/2005) and Taiwan (GAN DQ646405, Tw2575/98), as well as three turkey coronavirus (TCoV) sequences and sequences of coronaviruses isolated from a duck and a peafowl in China (GAN EU022525, TCoV-540; EU022526, TCoV-ATCC; GQ427174, TCoV/TX-GL/01; JF705860, Duck CoV; and AY641576, Peafowl, respectively), were used for phylogenetic analyses. The sequences were aligned using ClustalOmega (Sievers *et al.*, 2011; www.ebi.ac.uk) and the phylogenetic tree constructed using Geneious™ 6.1.4, using the nearest neighbour interchange maximum likelihood heuristic method with 100 bootstrap replications and the general time-reversible (GTR) substitution model, gamma distributed with invariant sites (G+I). The GTR+G+I substitution model was selected as the ideal model for these data using the "Find Best DNA/Protein Model" tool, available in Mega5.2 (Tamura *et al.*, 2011).

Phylogenetic trees were also constructed for individual peptides encoded by the polymerase genes, the main proteinase, also referred to as 3CL, the helicase (Hel), the PLP and the RNA-dependent RNA polymerase (RdRp), and also for the S1 glycoprotein. The sequences were aligned with ClustalW (Geneious™ 6.1.4) and the phylogenetic trees constructed as described above. In these phylogenetic trees, only the VicS-v sequence was included (except for the S1 glycoprotein tree, where both VicS-v and VicS-del were included), as the VicS-v and VicS-del sequences were highly similar (99.8%), and VicS-v is the main component of the VicS vaccine (Hewson *et al.*, 2012).

To assess the extent of potential recombination between viruses, a network tree was constructed using the same complete genome alignments as described above and SplitsTree 4.13.1 (Huson & Bryant, 2006). The distances were calculated with the uncorrected P-method and the network constructed with Neighbor-Net using 1000 bootstrap replicates.

The sequences of the complete genomes of VicS-v and VicS-del have been submitted to GenBank under accession numbers KF460437 and KF931628, respectively.

Results

Sequencing and coverage. For VicS-v and VicS-del, the Ion-Torrent 314 chip generated 52.9 and 130.3 Mb of data, comprising 503,622 and 480,037 reads, respectively, with mean read lengths of 102.9 and 124.2 nucleotides, respectively. After removal of the host-matched reads, a total of 466,235 and 456,359 reads were retained (92.57% and 95.06%), respectively. For VicS-v, a total of 212,507 readings (45.6%) were matched with the complete genome sequence of the Beaudette strain. The mean depth of coverage was 716 fold. A total of 98,124 reads could be mapped to the structural protein region, yielding a mean depth of coverage of 1200 fold and a minimum depth of 92 fold. For VicS-del, 285,759 reads (62.6%) were matched with the complete genome sequence of the VicS-v strain. The mean depth of coverage was 760 fold. A total of 95,802 reads could be mapped to the structural protein region, yielding a mean depth of coverage of 925 fold and a minimum depth of 126 fold.

The VicS-v and VicS-del genomes were 27,610 and 27,567 nucleotides in length, with a G + C content of 38.2% and 37.8%, respectively. Vics-v had the typical IBV genome organization (5′ UTR - 1a/1ab - S – 3 a, b, c (E)- 4 a, b - M – 5 a, b - N - 3′ UTR; Stadler *et al*., 2003; Cavanagh, 2007; Thor *et al.*, 2011), while VicS-del had a truncated 4b open reading frame (ORF).

Comparisons of the complete genomes of VicS-v and the more attenuated VicS-del (Figure 1) revealed several differences, which are summarized in Table 1. The most significant of these differences found in VicS-del were the absence of a glycine (Gly) in two Gly-Gly motifs in VicS-del and two insertions–deletions (indels) in the nucleotide sequence of the non-structural protein 6 (nsp6) gene, which would lead to a change in 11 amino acids. This frameshift introduced two positively charged amino acids, changing the region from predominantly hydrophobic, and resulted in the loss of a third Gly-Gly motif. There was also a single-nucleotide polymorphism (SNP) in the S gene at nucleotide 3475, close to the 3′ end of the gene, with approximately equal numbers of VicS-v genomes having either a G or a U, while in VicS-del most genomes had a U in this position. Finally, a frameshift mutation resulting in the truncation of ORF 4b and a 40-nucleotide deletion in the 3′ UTR region were recognized in the genome of VicS-del.

The SNP in the S gene generated two non-synonymous codons, GAA (glutamate) or UAA (a premature stop codon), nine codons prior to the usual S gene stop codon. This SNP also lies within the body transcription (B-TRS) for mRNA 3, which encodes the accessory proteins 3a and 3b, and the structural protein E. In VicS-v, 48.4% of the reads (595/1222) contained the codon GAA and 51% contained the UAA stop codon. In VicS-del, 5.5% of the reads contained the GAA codon and 93.6% the UAA codon (Table 1).

A second Gly-Gly motif, which has been described as a PLP cleavage site (Ziebuhr, 2005), was identified in the VicS-v sequence between nucleotides 7306 and 7311.

Nucleotide alignment and phylogenetic analyses. Phylogenetic analysis assigned most of the genomes into two main clusters (Figure 2). The first major cluster (a) included predominantly USA strains, except for the Ck/Ch/LDL/97I (China) and peafowl (China) strains. The peafowl (China) isolate lay within the Beaudette subcluster and has been described previously as very similar to a Mass 41 strain isolated in 1985 (Liu *et al.*, 2005). The USA cluster can be subdivided into the Mass 41 vaccine, Beaudette and Holte subclusters. The second cluster (b) contained only strains of Chinese origin, including a strain isolated from a duck, and these were subdivided in two subclusters (b1 and b2 in Figure 2). TCoV strains lay in a separate cluster, closely related to USA IBV strains. There was a high level of similarity between the Ita/90254/2005 (Italy) and Ck/Swe/0658946/10 (Sweden) strains, which have been described previously as European QX-like strains, a Chinese origin genotype widely

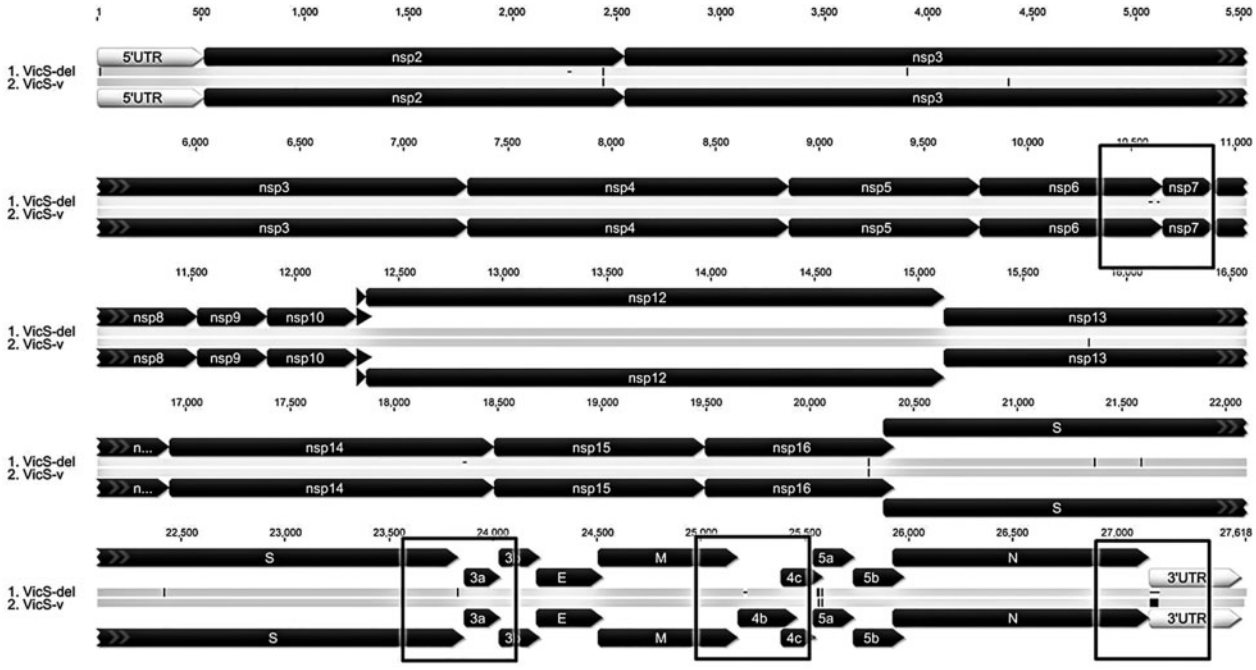

Figure 1. *Nucleotide sequence alignment of the VicS-v and VicS-del genomes (GAN KF460437 and KF931628). ORF 4b is absent in VicS-del and there is a 40-nucleotide deletion in the 3′ UTR of VicS-del. There are also non-synonymous changes at the end of nsp6 and a SNP at the end of the S glycoprotein sequences (boxed). Non-structural protein 1 (nsp1), present in other members of family* Coronaviridae*, is absent in IBV. Vertical lines indicate differences in the nucleotide sequences. Dashes indicate deletions in the sequence.*

Table 1. *Nucleotide differences detected between the genome sequences of strains VicS-del and VicS-v*[a].

Site (nucleotide)	Peptide[b]	Change[c]	Sequence	Effect
2,268–2,270	nsp2	D	GGGGGU → GG - - -U	Gly-Gly → Gly
2,434	nsp2	S	CAU → GAU	His → Asp
3,894	nsp3	S	AAU → AAC	Synonymous
4,380	nsp3	S	AAU → AAG	Lys → Asn
10,585	nsp6	D	GCU → - CU	W E **V F A T N I L I Q G I G** G D →
10,621–10,622	nsp6	D	GGGGGU → GG - - - U	W E **S L R Q I Y L Y K E L** G D[d]
15,816	nsp13	S	GUU → UUU	Val → Phe
18,341–18,343	nsp14	D	GGGGGC → GG - - - C	Gly-Gly → Gly
20,279	nsp16	S	CCA → CCG	Synonymous
21,365	Sgp	S	UGU → CGU	Cys → Arg
22,421	Sgp	S	CAU → CCU	His → Pro
23,830	Sgp	P	GAA (48.4%) → GAA (5.5%) UAA (51.0%) → UAA (93.6%)	Glu → Stop codon
25,208	4b	D	UUG → U - G	Premature stop codon (25,232-25,234) Alternative start codon (25,206-25,208)
25,555	5a	S	AGU → AUU	Ile → Ser
25,563 & 25,565	5a	S	GGA → AGC	Gly → Ser
25,582	5a	S	UAU → UGU	Tyr → Cys
27,157–27,196	3′ UTR	D	NA	40 nucleotide deletion

[a]All changes indicated are from VicS-v to VicS-del.
[b]Peptide in which the change occurred; nsp2, 3, 6, 13, 14, 16, non-structural proteins 2, 3, 13, 14, 16; Sgp, S glycoprotein; 3′ UTR, 3′ untranslated region.
[c]D, deletion; S, substitution; P, polymorphism.
[d]Amino acid residues in bold are those altered by reading frame shifts, with the VicS-v sequence on the top and the VicS-del sequence below.

distributed in Europe (Ducatez *et al.*, 2009; Abro *et al.*, 2012). The strains Ck/Ch/LSD/05I (China) and Tw2575/98 (Taiwan) also shared a cluster. The VicS-v and VicS-del strains (Australia) form their own cluster, while Ibadan (Nigeria) and SNU8067 (South Korea) strains did not fall into any cluster.

In the phylogenetic trees of the 3CL, Hel, PLP and RdRp genes (Figure 3), the Chinese and USA clusters were

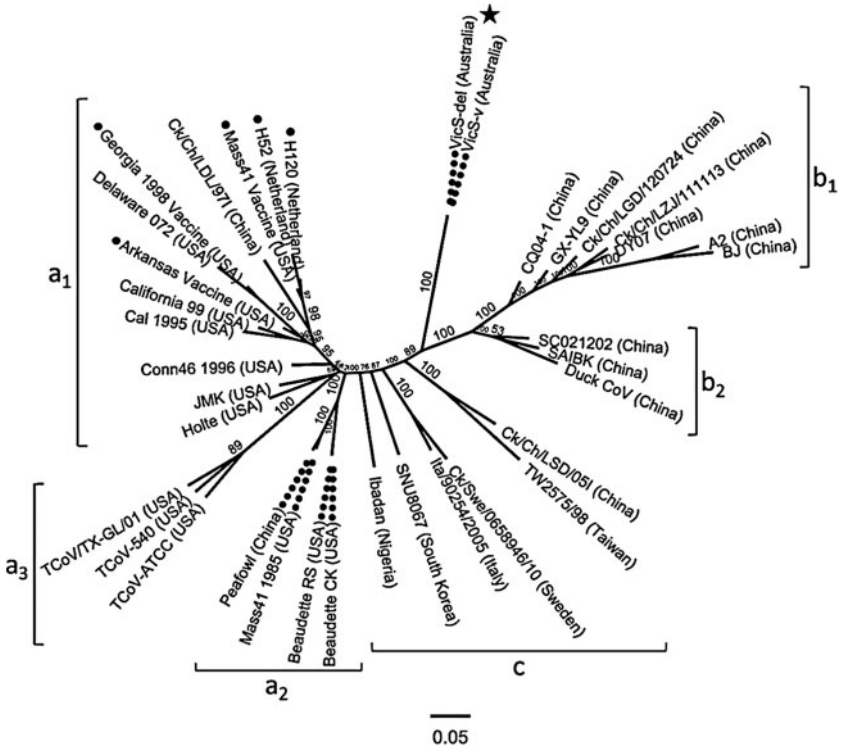

Figure 2. *Unrooted phylogenetic tree of 35 complete genomes of IBV strains isolated in different countries. The nucleotide alignment was performed using Clustal-Omega (Sievers et al., 2011) (http://www.ebi.ac.uk) and the phylogenetic tree inferred in Geneious 6.1.4, using the nearest neighbour interchange (NNI) maximum likelihood heuristic method with 100 bootstrap replications and a general time-reversible (GTR) substitution model, gamma distributed with invariant sites (G+I);* $\mathbf{a}_1$ *and* $\mathbf{a}_2$*, subclusters within the USA cluster;* $\mathbf{a}_3$*, turkey coronavirus (TCoV) cluster;* $\mathbf{b}_1$ *and* $\mathbf{b}_2$*, subclusters of the Chinese cluster;* **c***, strains outside the main clusters, including a QX-like strain (Ck/Swe/0658946/10 and Ita/90254/2005). The star highlights VicS-v and VicS-del, and the dots indicate other vaccine strains (H120, H52, Mass 41, Georgia 1998 and Arkansas).*

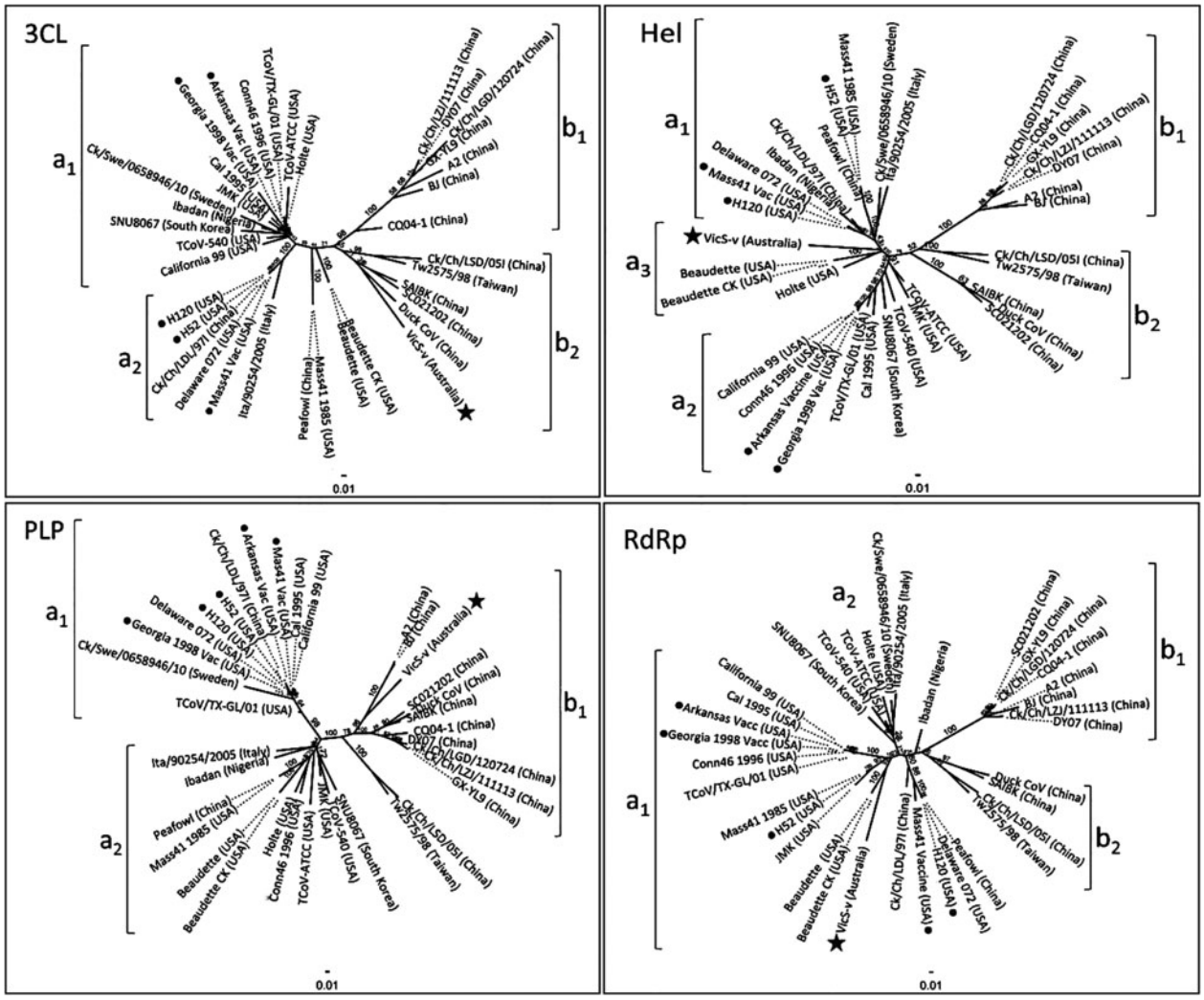

Figure 3. *Phylogenetic trees constructed with the nucleotide sequence alignments of the 3CL, helicase (Hel), papain-like proteinase (PLP) and RNA-dependent RNA polymerase (RdRp) genes inferred using the nearest neighbour interchange (NNI) maximum likelihood heuristic method and 100 bootstrap replications, and a general time-reversible (GTR) substitution model, gamma distributed with invariant sites (G+I);* $\mathbf{a}_1$ *and* $\mathbf{a}_2$ *indicate the two USA subclusters,* $\mathbf{b}_1$ *and* $\mathbf{b}_2$ *the two Chinese subclusters. The stars mark the position of VicS-v and the dots the position of other vaccine strains (H120, H52, Mass 41, Georgia 1998 and Arkansas).*

conserved. However, VicS-v grouped within the Chinese cluster in the 3CL and PLP gene phylogenetic trees and within the USA cluster in the Hel and RdRp gene phylogenetic trees.

In the S glycoprotein gene trees (Figure 4) the isolates fell into distinct Chinese and USA clusters. In these trees, VicS-v, VicS-del and Ibadan lay closer to the USA cluster, while the South Korean, Taiwanese and European strains were positioned closer to the Chinese cluster. The Georgia 1998 and Delaware 072 strains were distinct from all other strains, and were located between the IBV and TCoV groups. A phylogenetic tree inferred using the amino acid sequences had the same cluster organization (Figure 5).

Network tree analysis. In the analysis shown in Figure 6, the boxes indicate the likelihood of recombination events between strains of IBV. The tree has the same cluster organization seen in the phylogenetic tree using whole genome sequences (Figure 2). The VicS-v and VicS-del branch splits from all clusters close to the centre of the tree, suggesting limited and historically distant recombination with other strains.

Discussion

Although there has been considerable investigation of IBV and of coronaviruses in general, the determinants of their virulence and their capacity to replicate *in vivo* remain uncertain. As recent studies have suggested that the virulence determinants may not be located in the structural protein region of some IBVs (Armesto *et al.*, 2009), it can be inferred that these determinants may be encoded within the 1a and 1ab polyproteins. Therefore, studies of complete genome sequences could provide useful information about the position and characteristics of potential virulence determinants.

In this study, the genome sequences of the strains VicS-v and VicS-del were completed. The 1a and 1ab polyproteins of IBV are post-translationally cleaved by virus-encoded proteinases into 15 peptides (nsp2 to nsp16). Nsp1, which is found in all other coronaviruses, is absent in IBV, but nsp2 is considerably larger (Liu *et al.*, 1998; Snijder *et al.*, 2003; Stadler *et al.*, 2003; Gorbalenya *et al.*, 2004; Ziebuhr, 2005; Fang *et al.*, 2007; Sawicki *et al.*, 2007; Carbajo-Lozoya *et al.*, 2012). In order to determine the location of the 15 peptides encoded by the ORFs in the 1a

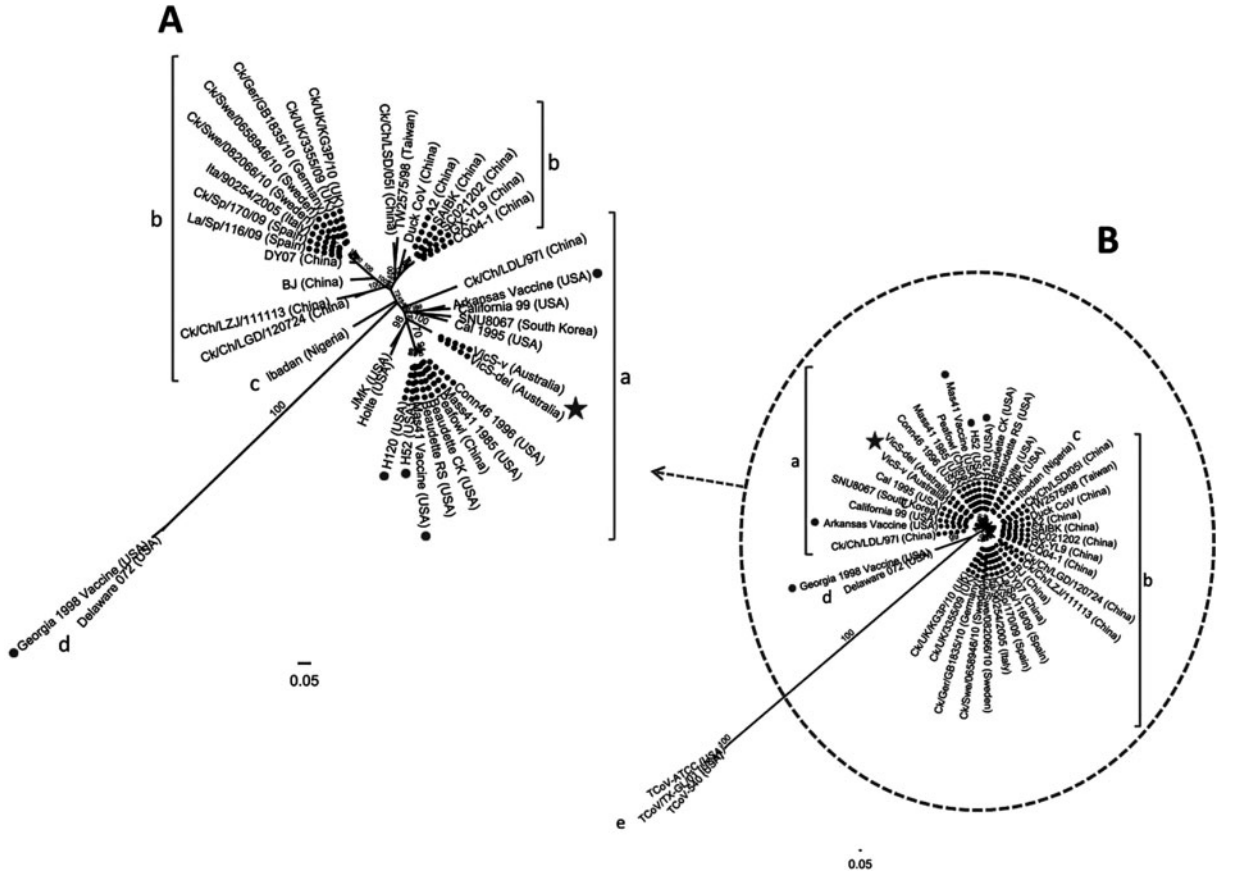

Figure 4. *Phylogenetic trees constructed with the nucleotide sequence alignments of the S1 glycoprotein genes, and with Geneious 6.1.4, using the nearest neighbour interchange (NNI) maximum likelihood heuristic method with 100 bootstrap replications and the general time-reversible (GTR) substitution model, gamma distributed with invariant sites (G+I):* **A**, *tree with TCoV strains removed (TCoV included in* **B***);* **a**, *USA cluster;* **b**, *Chinese cluster, including QX strains;* **c**, *Ibadan strain;* **d**, *Georgia and Delaware cluster;* **e**, *TCoV strains. The star marks the position of VicS-v and VicS-del, and the dots the position of other vaccine strains.*

and 1ab polyproteins, we used the annotations of the Beaudette strain (GAN NC_001451). We complemented these annotations with those of Ziebuhr (2005). The locations of these peptides within ORFs 1a and 1ab and their first and last amino acid residues are summarized in Table 2.

As discussed above, VicS-v exhibits higher rates of replication *in vitro* and causes more severe clinical signs and tracheal lesions *in vivo* than VicS-del. As the genomes of these two viruses have very high levels of nucleotide sequence identity, it can be inferred that any differences in their genomic sequences may influence these differences in virulence and replication.

As described previously, Gly-Gly is the cleavage motif for PLP (Lim *et al.*, 2000; Ziebuhr, 2005). VicS-del lacks three Gly-Gly motifs seen in VicS-v that are located in the middle of nsp2, nsp6 and nsp14. These motifs may result in additional proteolytic processing of these three nsps, and the absence of them in VicS-del could alter their function. Nsp6 and nsp14 are reported to contain transmembrane (TM) and exonuclease domains (Ziebuhr, 2005), while the function of nsp2 remains undetermined. Highly conserved Gly-Gly motifs were also found between nucleotides 2044 and 2049 (within nsp2), 3667 and 3672, 4426 and 4431 (within nsp3), 9199 and 9204, and 9391 and 9396 (within nsp5) in both VicS-v and VicS-del and in other IBVs. These motifs may also be further processed by PLP, although the previously characterized PLP cleavage sites (between nsp2 and 3, and nsp3 and 4) have an extended motif of Lys-Ala-Gly-Gly, which has been identified in all the complete genome sequences analysed in this study; this extended motif is not seen at these additional Gly-Gly sites.

Table 2. *Positions*[a] *of the non-structural proteins in ORFs 1a and 1ab*[b].

Non-structural proteins	VicS-v	VicS-del
Nsp2	1 M-674 G	1 M-673 G
Nsp3	675 G-2264 G	674 G-2263 G
Nsp4	2265 G-2778 Q	2264 G-2777 Q
Nsp5	2779 A-3085 Q	2778 A-3084 Q
Nsp6	3086 S-3379 Q	3085 S-3377 Q
Nsp7	3380 S-3462 Q	3378 S-3460 Q
Nsp8	3463 S-3672 Q	3461 S-3670 Q
Nsp9	3673 N-3783 Q	3671 N-3781 Q
Nsp10	3784 S-3928 Q	3782 S-3926 Q
Nsp11	3929 S-3951 G	3927 S-3949 G
Nsp12	3929 S-4868 Q	3927 S-4866 Q
Nsp13	4869 S-5468 Q	4867 S-5466 Q
Nsp14	5469 G-5989 Q	5467 G-5986 Q
Nsp15	5990 S-6327 Q	5987 S-6324 Q
Nsp16	6328 S-6629 M	6325 S-6626 M

[a]Positions were based on 1ab from IBV Beaudette strain (accession number NC_001451), and modified according to Ziebuhr (2005), with position 1 the methionine at the beginning of ORF 1a and 1ab.
[b]Based on Phillips *et al.* (2012).

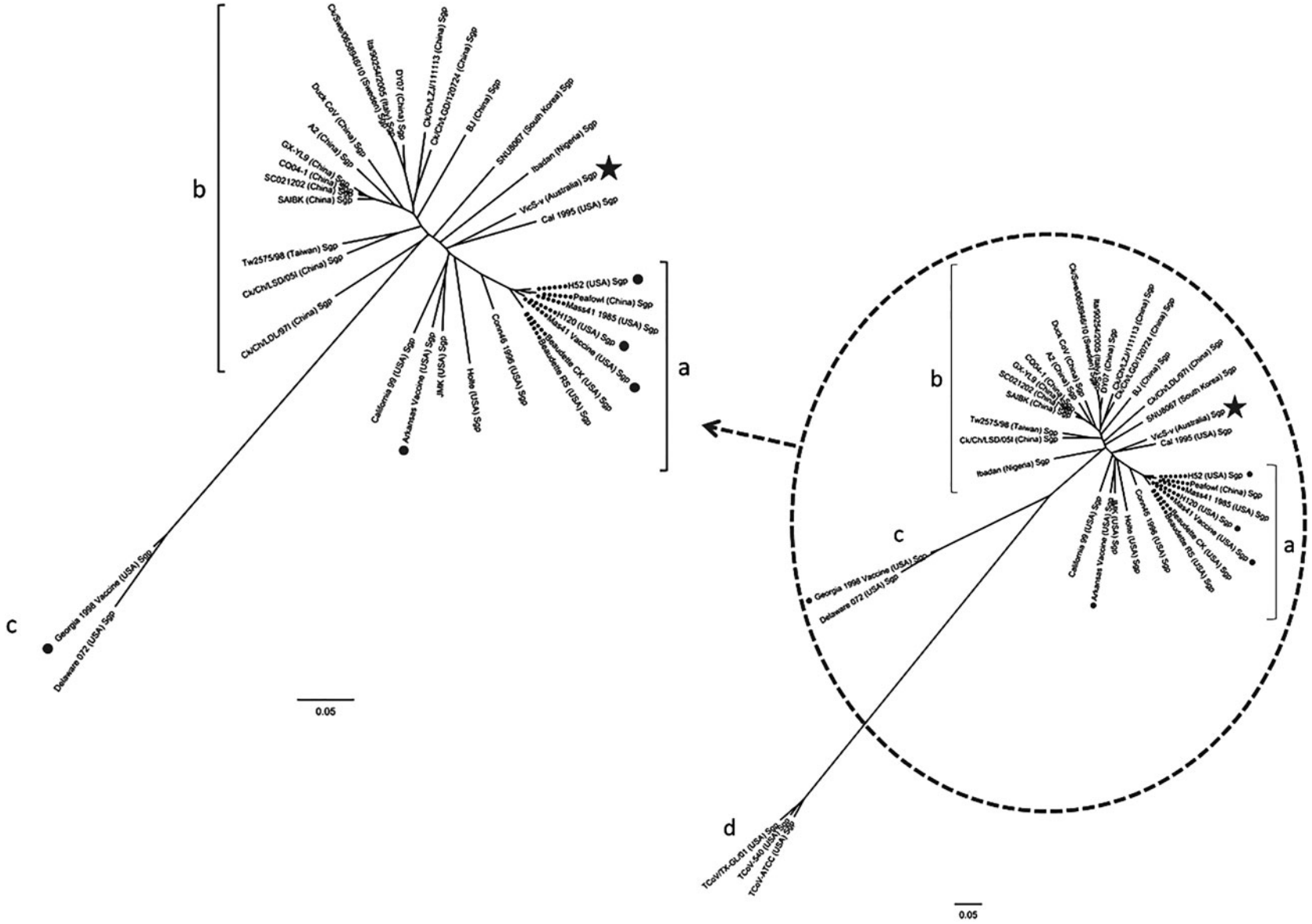

Figure 5. *Phylogenetic trees constructed with the amino acid sequences alignment of the S1 glycoprotein genes, and with Geneious 6.1.4, using the nearest neighbour interchange (NNI) maximum likelihood heuristic method with 100 bootstrap replications and the general time-reversible (GTR) substitution model, gamma distributed with invariant sites (G+I):* **A**, *tree with TCoV strains removed (TCoV included in* **B***);* **a**, *USA cluster;* **b**, *Chinese cluster;* **c**, *Georgia and Delaware cluster;* **d**, *TCoV strains. The star marks the position of VicS-v and the dots the position of other vaccine strains.*

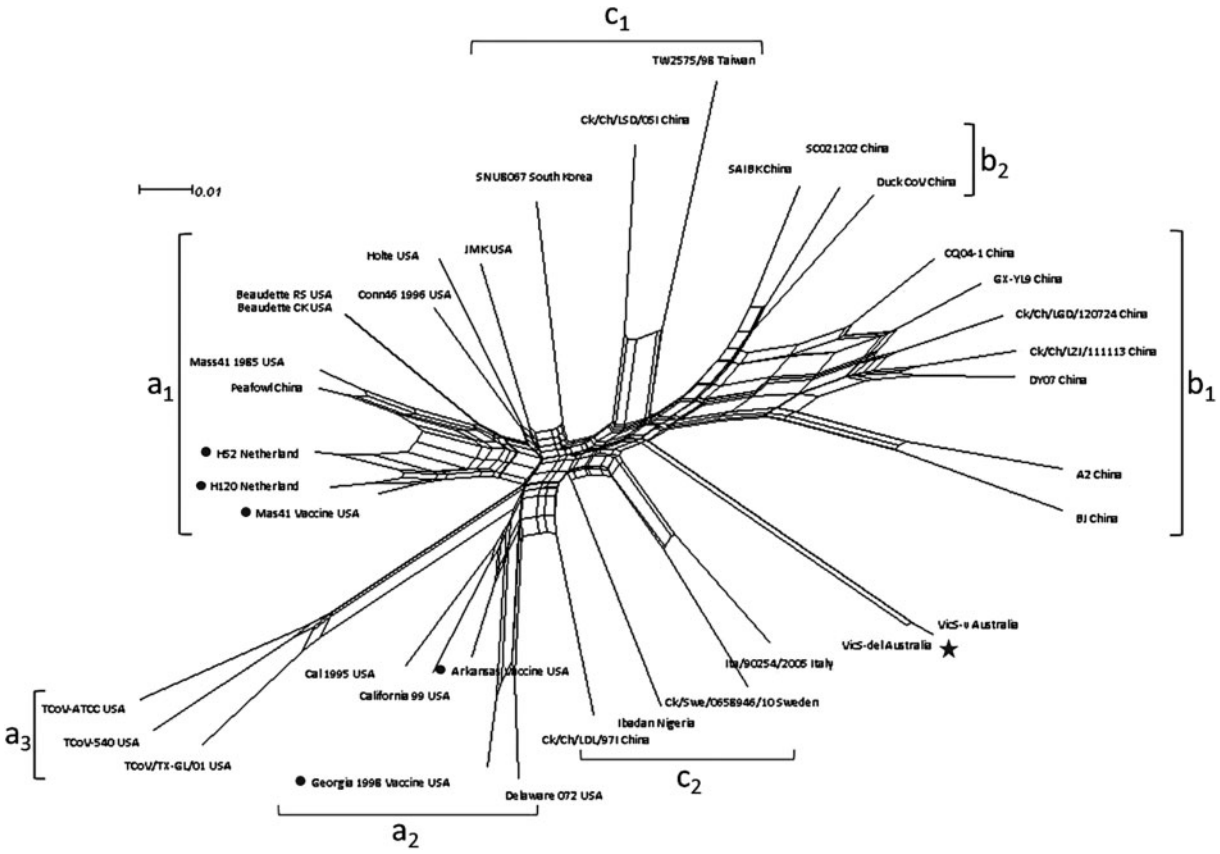

Figure 6. *Network tree analysis constructed with the nucleotide sequence alignments of the full genome of different IBVs (SplitsTree4):* $\mathbf{a}_{1+2}$*, USA IBV subclusters,* $\mathbf{a}_3$ *TCoV subcluster;* $\mathbf{b}_1$ *and* $\mathbf{b}_2$*, Chinese subclusters;* $\mathbf{c}_1$ *and* $\mathbf{c}_2$*, strains outside the main clusters, including the QX-like strains (Ck/Swe/0658946/10 and Ita/90254/2005). Boxes indicate the likelihood of recombination between strains. The star marks the position of the Australian vaccine strain and the dots the position of other vaccine strains.*

The mutations in nsp6 of VicS-del (Table 1) may have implications for the replication of this virus, which would be consistent with data from previous *in vivo* studies (Hewson *et al.*, 2012). During infection 3CL, the main protease, is auto-catalytically released (Ziebuhr *et al.*, 2001; Ziebuhr, 2005). There may be an interaction between the 3CL (nsp5), the TM2 (nsp4) and the TM3 (nsp6) hydrophobic domains (Ziebuhr, 2005), which are also referred to as MP1 and MP2 (Tibbles *et al.*, 1996), during this proteolytic process. The importance of TM3/MP2, the transmembrane domain downstream of the 3CL proteinase, in the autocatalytic processing of 3CL has been suggested by *in vitro* assays (Tibbles *et al.*, 1996). In this experiment, the further processing of a 66-kDa immature protein product into smaller, post-translationally cleaved products of 59, 42, 35 and 24 kDa (with the 35-kDa peptide the size predicted for the fully processed form of 3CL protease) did not occur if these two hydrophobic domains were not associated with microsomal membranes. Changes in TM3/MP2, such as those detected in VicS-del, could affect the membrane association of this domain and thus the conformation of the protein, and the autocatalytic release of 3CL, which is essential for the initial stages of replication (Ziebuhr, 2005).

The SNP detected at the end of the S gene is also notable. The difference in the proportions of each of the non-synonymous codons between the VicS-v and VicS-del populations, with a greater prevalence of the truncation phenotype in the VicS-del population, suggests that truncation of the S2 protein may be related to the difference in virulence between these strains. The carboxyl terminal end of S2 anchors the S protein to the virion and also assists in membrane fusion with target cells (Cavanagh, 2007). Alternatively, the differing B-TRS sequences, CUGAA-CAA and CUUAACAA, which have both been identified previously in IBV (Mondal & Cardona, 2007; Woo *et al.*, 2010), may differ in their efficiency in regulating replication and transcription and thus may influence production of mRNA 3, which encodes the E protein. In the Middle East respiratory syndrome coronavirus the deletion of the E protein results in a replication-competent, propagation-defective virus (Almazán *et al.*, 2013). A nine amino acid truncation, resulting from the nucleotide change from U to G, has been described previously in an M41 strain variant, M41H (Kusters *et al.*, 1989). While these alternative codons have also been found in other IBV strains, the methods used for sequence determination would have precluded identification of differing subpopulations within these strains, as cloned, single cDNA molecules were used for sequencing.

The ORFs 4b and 4c have been described in IBV (Hewson *et al.*, 2011, 2012; Thor *et al.*, 2011; Phillips *et al.*, 2012; Bentley *et al.*, 2013), as well as in TCoV (Cao *et al.*, 2008), and have been synthesized *in vitro* (Bentley *et al.*, 2013), but the products of these ORFs are yet to be identified in infected cells, and their function remains unknown. Both these ORFs could be identified in the VicS-v genome, but in VicS-del, 4b was divided into two smaller ORFs due to a single-nucleotide deletion.

The VicS vaccine is known to induce some clinical signs after vaccination. Hewson *et al.* (2012) showed that VicS-v caused more severe vaccine reactions than VicS-del, had a higher replication rate in the trachea, kidneys and caecal

tonsils, and induced more severe clinical signs. The putative 4b and 4c proteins have two interesting features: their high content of positively charged amino acids (Lys and Arg) and the presence of phosphorylation motifs, suggesting a possible role in the packaging of the genome, which may influence the rate of viral replication and therefore pathogenicity.

Phylogenetic analyses showed that the Australian strains of IBV are distinct from strains isolated elsewhere in the world and that the most commonly used Australian vaccine strain is very distinct from other IBV vaccine strains across the entire genome. This divergence has been shown previously in the structural genes (Yu *et al*., 2001; Bochkov *et al*., 2007), but this is the first study to show that the divergence can be seen consistently throughout the genome. However, most of the fully sequenced IBV genomes are of strains with an origin in the USA or China, as only a few QX-type and QX-like strains (which now predominate in Europe) have been fully sequenced, and this bias could have influenced the topology of the tree. These European strains were represented in this tree by strains from Italy (Ita/90254/2005) and Sweden (Ck/Swe/0658946/10). The phylogenetic tree constructed with the nucleotide sequences of the S1 protein (Figure 4) includes more QX-type and QX-like strains, including Ck/UK/KG3P/10 and Ck/UK/3355/09 (UK), Ck/Ger/GB1835/10 (Germany), Ck/Swe/0658946/10 and Ck/Swe/082066/10 (Sweden), Ita/90254/2005 (Italy) and Ck/Sp/170/09 and La/Sp/116/09 (Spain). Phylogenetic trees of individual ORFs (Figures 3, 4 and 5) indicate that different regions of the VicS-v genome varied in their similarity to the two major clusters, suggesting historical recombination events. However, this seems very unlikely to have happened recently, based on the results from the network tree analysis (Figure 6). VicS-v diverged and evolved as a distinct cluster, with the length of the VicS-v branches, especially in the 3CL, Hel and RdRp phylogenetic trees, suggesting more rapid evolution than most other strains.

In conclusion, full genome analysis of VicS-v and VicS-del has shown that the Australian IBV vaccine VicS has been evolving independently of strains elsewhere in the world for some time. The differences in the replication rate and pathogenicity of these two viral subpopulations may be attributable to the mutation in the TM3 domain in nsp6 and the SNP found at the end of the S glycoprotein gene, described in this study for the first time. These newly described differences are additional to the previously described frameshift/truncation of the ORF 4b and the 40-nucleotide deletion in the 3′ UTR in VicS-del (Hewson *et al*., 2012). Further analyses of the complete genome sequences of other Australian IBV strains may reveal additional information about relationships between these distinct vaccine strains and IBVs within the USA and Chinese clusters. Additional investigation of the effect of the differences between the VicS-v and VicS-del strains described in this study may shed more light on the replication and pathogenicity of IBV and may also indicate whether the naturally attenuated strain VicS-del and other strains with similar mutations have potential as novel vaccine candidates for the control of IBV infections.

References

Abro, S., Renström, L.M., Ullman, K., Belák, S. & Baule, C. (2012). Characterization and analysis of the full-length genome of a strain of the European QX-like genotype of infectious bronchitis virus. *Archives of Virology*, 157, 1211–1215.

Almazán, F., DeDiego, M.L., Sola, I., Zuñiga, S., Nieto-Torres, J.L., Marquez-Jurado, S., Andrés, G. & Enjuanes, L. (2013). Engineering a replication-competent, propagation-defective Middle East respiratory syndrome coronavirus as a vaccine candidate. *mBio*, 4, e00650–00613.

Ammayappan, A., Upadhyay, C., Gelb, Jr. J. & Vakharia, V. N. (2008). Complete genomic sequence analysis of infectious bronchitis virus Ark DPI strain and its evolution by recombination. *Virology Journal*, 5, 157.

Armesto, M., Cavanagh, D. & Britton, P. (2009). The replicase gene of avian coronavirus infectious bronchitis virus is a determinant of pathogenicity. *PLoS ONE*, 4, e7384.

Beaudette, F. & Hudson, C. (1937). Cultivation of the virus of infectious bronchitis. *Journal of the American Veterinary Medical Association*, 90, 51–60.

Bentley, K., Keep, S.M., Armesto, M. & Britton, P. (2013). Identification of a noncanonically transcribed subgenomic mRNA of infectious bronchitis virus and other gammacoronaviruses. *Journal of Virology*, 87, 2128–2136.

Bochkov, Y.A., Tosi, G., Massi, P. & Drygin, V. (2007). Phylogenetic analysis of partial S1 and N gene sequences of infectious bronchitis virus isolates from Italy revealed genetic diversity and recombination. *Virus Genes*, 35, 65–71.

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.

Carbajo-Lozoya, J., Muller, M.A., Kallies, S., Thiel, V., Drosten, C. & von Brunn, A. (2012). Replication of human coronaviruses SARS-CoV, HCoV-NL63 and HCoV-229E is inhibited by the drug FK506. *Virus Research*, 165, 112–117.

Casais, R., Dove, B., Cavanagh, D. & Britton, P. (2003). Recombinant avian infectious bronchitis virus expressing a heterologous spike gene demonstrates that the spike protein is a determinant of cell tropism. *Journal of Virology*, 77, 9084–9089.

Cavanagh, D. (1981). Structural polypeptides of coronavirus IBV. *Journal of General Virology*, 53, 93–103.

Cavanagh, D. (2007). Coronavirus avian infectious bronchitis virus. *Veterinary Research*, 38, 281–297.

Cavanagh, D., Brien, D.A., Brinton, M., Enjuanes, L., Holmes, K.V., Horzinek, M.C., Lai, M.M.C., Laude, H., Plagemann, P.G.W., Siddell, S., Spaan, W.J.M., Taguchi, F. & Talbot, P.J. (1994). Revision of the taxonomy of the Coronavirus, Torovirus and Arterivirus genera. *Archives of Virology*, 135, 227–237.

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.

Fang, S., Chen, B., Tay, F.P.L., Ng, B.S. & Liu, D.X. (2007). An arginine-to-proline mutation in a domain with undefined functions within the helicase protein (Nsp13) is lethal to the coronavirus infectious bronchitis virus in cultured cells. *Virology*, 358, 136–147.

Gorbalenya, A.E., Koonin, E.V., Donchenko, A.P. & Blinov, V.M. (1989). Coronavirus genome: prediction of putative functional domains in the non-structural polyprotein by comparative amino acid sequence analysis. *Nucleic Acids Research*, 17, 4847–4861.

Gorbalenya, A.E., Snijder, E.J. & Spaan, W.J.M. (2004). Severe acute respiratory syndrome coronavirus phylogeny: toward consensus. *Journal of Virology*, 78, 7863–7866.

Hewson, K.A., Ignjatovic, J., Browning, G.F., Devlin, J.M. & Noormohammadi, A.H. (2011). Infectious bronchitis viruses with naturally occurring genomic rearrangement and gene deletion. *Archives of Virology*, 156, 245–252.

Hewson, K.A., Noormohammadi, A.H., Devlin, J.M., Mardani, K. & Ignjatovic, J. (2009). Rapid detection and non-subjective characterisation of infectious bronchitis virus isolates using high-resolution melt curve analysis and a mathematical model. *Archives of Virology*, 154, 649–660.

Hewson, K.A., Scott, P.C., Devlin, J.M., Ignjatovic, J. & Noormohammadi, A.H. (2012). The presence of viral subpopulations in an infectious bronchitis virus vaccine with differing pathogenicity-a preliminary study. *Vaccine*, 30, 4190–4199.

Huson, D.H. & Bryant, D. (2006). Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution*, 23, 254–267.

Ignjatovic, J. & Galli, U. (1995). Immune responses to structural proteins of avian infectious bronchitis virus. *Avian Pathology*, 24, 313–332.

Ignjatovic, J., Gould, G. & Sapats, S. (2006). Isolation of a variant infectious bronchitis virus in Australia that further illustrates diversity among emerging strains. *Archives of Virology*, 151, 1567–1585.

Ignjatovic, J., Sapats, S.I. & Ashton, F. (1997). A long-term study of Australian infectious bronchitis viruses indicates a major antigenic change in recently isolated strains. *Avian Pathology*, 26, 535–552.

Jackwood, M.W. & de Wit, J.J. (2013). Infectious Bronchitis. In D.E. Swayne, J.R. Glisson, L.R. McDougald, L.K. Nolan, D.L. Suarez & V.L. Nair. (2013). *Diseases of Poultry* 13th edn (pp. 139–159). Ames: John Wiley & Sons.

Knoops, K., Kikkert, M., Worm, S.H.E. v. d., Zevenhoven-Dobbe, J. C., van der Meer, Y., Koster, A. J., Mommaas, A. M. & Snijder, E. J. (2008). SARS-coronavirus replication is supported by a reticulovesicular network of modified endoplasmic reticulum. *PLoS Biology*, 6, e226.

Kusters, J.G., Niesters, H.G.M., Lenstra, J.A., Horzinek, M.C. & van der Zeijst, B.A.M. (1989). Phylogeny of antigenic variants of avian coronavirus IBV. *Virology*, 169, 217–221.

Lim, K.P., Ng, L.F.P. & Liu, D.X. (2000). Identification of a novel cleavage activity of the first papain-like proteinase domain encoded by open reading frame 1a of the coronavirus avian infectious bronchitis virus and characterization of the cleavage products. *Journal of Virology*, 74, 1674–1685.

Liu, D.X. & Brown, T.D.K. (1995). Characterisation and mutational analysis of an ORF 1a-encoding proteinase domain responsible for proteolytic processing of the infectious bronchitis virus 1a/1b polyprotein. *Virology*, 209, 420–427.

Liu, D.X. & Inglis, S.C. (1992a). Identification of two new polypeptides encoded by mRNA5 of the coronavirus infectious bronchitis virus. *Virology*, 186, 342–347.

Liu, D.X. & Inglis, S.C. (1992b). Internal entry of ribosomes on a tricistronic mRNA encoded by infectious bronchitis virus. *Journal of Virology*, 66, 6143–6154.

Liu, D.X., Cavanagh, D., Green, P. & Inglis, S.C. (1991). A polycistronic mRNA specified by the coronavirus infectious bronchitis virus. *Virology*, 184, 531–544.

Liu, D.X., Shen, S., Xu, X.Y. & Brown, T.D.K. (1998). Proteolytic processing of the polyprotein encoded by ORF1b of the coronavirus infectious bronchitis virus (IBV). In L. Enjuanes. (1998). *Coronaviruses and Arteriviruses* (pp. 149–159). New York: Springer.

Liu, S., Chen, J., 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.

Liu, S., Zhang, Q., Chen, J., Han, Z., Shao, Y., Kong, X. & Tong, G. (2008). Identification of the avian infectious bronchitis coronaviruses with mutations in gene 3. *Gene*, 412, 12–25.

Lougovskaia, N.N., Lougovskoi, A.A., Bochkov, Y.A., Batchenko, G.V., Mudrak, N.S., Drygin, V.V., Borisov, A.V., Borisov, V.V. & Gusev, A.A. (2002). Detection and estimation of avian infectious bronchitis virus antigen by a novel indirect liquid-phase blocking enzyme-linked immunosorbent assay using chicken and rabbit affinity purified immunoglobulins. *Avian Pathology*, 31, 549–557.

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

Phillips, J.E., Jackwood, M.W., McKinley, E.T., Thor, S.W., Hilt, D.A., Acevedol, N.D., Williams, S.M., Kissinger, J.C., Paterson, A.H., Robertson, J.S. & Lemke, C. (2012). Changes in nonstructural protein 3 are associated with attenuation in avian coronavirus infectious bronchitis virus. *Virus Genes*, 44, 63–74.

Sapats, S.I., Ashton, F., Wright, P.J. & Ignjatovic, J. (1996). Novel variation in the N protein of avian infectious bronchitis virus. *Virology*, 226, 412–417.

Sawicki, S.G., Sawicki, D.L. & Siddell, S.G. (2007). A contemporary view of coronavirus transcription. *Journal of Virology*, 81, 20–29.

Schochetman, G., Stevens, R.H. & Simpson, R.W. (1977). Presence of infectious polyadenylated RNA in the coronavirus avian bronchitis virus. *Virology*, 77, 772–782.

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Soding, J., Thompson, J.D. & Higgins, D.G. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology*, 7, 539.

Snijder, E.J., Bredenbeek, P.J., Dobbe, J.C., Thiel, V., Ziebuhr, J., Poon, L. L.M., Guan, Y., Rozanov, M., Spaan, W.J.M. & Gorbalenya, A.E. (2003). Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage. *Journal of Molecular Biology*, 331, 991–1004.

Stadler, K., Masignani, V., Eickmann, M., Becker, S., Abrignani, S., Klenk, H.-D. & Rappuoli, R. (2003). SARS-beginning to understand a new virus. *Nature Reviews Microbiology*, 1, 209–218.

Stern, D.F. & Sefton, B.M. (1982). Coronavirus proteins: structure and function of the oligosaccharides of the avian infectious bronchitis virus glycoproteins. *Journal of Virology*, 44, 804–812.

Sun, L., Zhang, G.-H., Jiang, J.-W., Fu, J.-D., Ren, T., Cao, W.-S., Xin, C.-A., Liao, M. & Liu, W.-J. (2007). A Massachusetts prototype like coronavirus isolated from wild peafowls is pathogenic to chickens. *Virus Research*, 130, 121–128.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. & Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution*, 28, 2731–2739.

Tannock, G. A. (1973). The nucleic acid of infectious bronchitis virus. *Archiv fur die Gesamte Virusforschung*, 43, 259–271.

Thor, S.W., Hilt, D.A., Kissinger, J.C., Paterson, A.H. & Jackwood, M.W. (2011). Recombination in avian gamma-coronavirus infectious bronchitis virus. *Viruses*, 3, 1777–1799.

Tibbles, K.W., Brierley, I., Cavanagh, D. & Brown, T.D. (1996). Characterization *in vitro* of an autocatalytic processing activity associated with the predicted 3C-like proteinase domain of the coronavirus avian infectious bronchitis virus. *Journal of Virology*, 70, 1923–1930.

Tyrell, D.A.J. (1968). Coronaviruses. *Nature*, 220, 650.

Wadey, C.N. & Faragher, J.T. (1981). Australian infectious bronchitis viruses: identification of nine subtypes by a neutralisation test. *Research in Veterinary Science*, 30, 70–74.

Watkins, H., Reeve, P. & Alexander, D.J. (1975). The ribonucleic acid of infectious bronchitis virus. *Archives of Virology*, 47, 279–286.

Woo, P.C.Y., Huang, Y., Lau, S.K.P. & Yuen, K.-Y. (2010). Coronavirus genomics and bioinformatics analysis. *Viruses*, 2, 1804–1820.

Yu, L., Jiang, Y., Sharon, L., Wang, Z., Nam, S.J., Liu, W. & Kwang, J. (2001). Characterization of three infectious bronchitis virus isolates from China associated with proventriculus in vaccinated chickens. *Avian Diseases*, 45, 416–424.

Ziebuhr, J. (2005). The coronavirus replicase. In L. Enjuanes. (2005). *Coronavirus Replication and Reverse Genetics* (Vol. 287, pp. 57–94). Berlin: Springer.

Ziebuhr, J., Thiel, V. & Gorbalenya, A.E. (2001). The autocatalytic release of a putative RNA virus transcription factor from its polyprotein precursor involves two paralogous papain-like proteases that cleave the same peptide bond. *Journal of Biological Chemistry*, 276, 33220–33232.