Veterinary Microbiology 197 (2016) 27–38

Contents lists available at ScienceDirect

Veterinary Microbiology

journal homepage: www.elsevier.com/locate/vetmic

Full genome analysis of Australian infectious bronchitis viruses suggests frequent recombination events between vaccine strains and multiple phylogenetically distant avian coronaviruses of unknown origin

José A. Quinteros[a,*], Sang-Won Lee[a,c], Philip F. Markham[a], Amir H. Noormohammadi[b], Carol A. Hartley[a], Alistair R. Legione[a], Mauricio J.C. Coppo[a], Paola K. Vaz[a], Glenn F. Browning[a]

[a] Asia-Pacific Centre for Animal Health, Faculty of Veterinary and Agricultural Sciences, The University of Melbourne, Parkville, 3010, Victoria, Australia
[b] Asia-Pacific Centre for Animal Health, Faculty of Veterinary and Agricultural Sciences, The University of Melbourne, Werribee, 3030, Victoria, Australia
[c] College of Veterinary Medicine, Konkuk University, 120 Neungdong-ro, Gwangjin-gu, Seoul 143-701, Republic of Korea, Republic of Korea

ARTICLE INFO

Article history:
Received 1 September 2016
Received in revised form 3 November 2016
Accepted 4 November 2016

Keywords:
Infectious bronchitis virus
Recombinants
Unknown avian coronaviruses
Full genome sequences
Vaccine recombinants
Recombination

ABSTRACT

Australian strains of infectious bronchitis virus (IBV) have been evolving independently for many years, with control achieved by vaccination with local attenuated strains. Previous studies have documented the emergence of recombinants over the last 20 years, with the most recent one, Ck/Aus/N1/08, detected in 2008. These recombinants did not appear to be controlled by the vaccines currently in use.

In this study we sequenced the complete genomes of three emergent Australian strains of IBV (IBV/Ck/Aus/N1/88, IBV/Ck/Aus/N1/03 and IBV/Ck/Aus/N1/08) and a previously incompletely characterised vaccine strain, IBV/Ck/Aus/Armidale, and compared them to the genome of the vaccine strain VicS. We detected multiple recombination events throughout the genome between wild type viruses and the vaccine strains in all three emergent isolates. Moreover, we found that strain N1/88 was not entirely exogenous, as was previously hypothesised. Rather, it originated from a recombination event involving the VicS vaccine strain. The S glycoprotein genes of N1/88 and N1/03 were known to be genetically distinct from previously characterised circulating strains and from each other, and the original donors of these genes remains unknown. The S1 glycoprotein gene of N1/88, a subgroup 2 strain, shares a high nucleotide identity with the sequence of the S1 gene of the recent isolate N1/08. As the subgroup 2 strains have not been isolated for at least 20 years, it appears likely that an unknown avian coronavirus that was the donor of the S1 glycoprotein sequence of N1/88 in the 1980s is still recombining with IBV strains in the field.

1. Introduction

Infectious bronchitis (IB) is a disease mainly affecting chickens that is caused by the infectious bronchitis virus (IBV), a member of the genus Gammacoronavirus in the family Coronaviridae (International Committee on Taxonomy of Viruses, 2015), with a single stranded positive sense RNA genome approximately 27.6 kb in length (Jackwood and de Witt, 2013). It causes significant losses in meat and egg producing chickens due to its effects on growth rates, egg production and quality, and mortality in infected flocks. There is no effective treatment, and vaccination with attenuated vaccine strains is the major tool used to prevent the losses in production caused by IBV in Australia (Chousalkar et al., 2009). Inactivated vaccines are used outside Australia, normally as part of vaccination programs in combination with live attenuated vaccines, particularly in layer and breeder flocks.

There is considerable variation in the virulence and tropism of IBVs, and new strains frequently emerge, in part due to point mutations, deletions and insertions. However, one of the most

* Corresponding author at: Veterinary Preclinical Centre, Corner of Flemington Road and Park Drive, Parkville, 3010, Victoria, Australia.
E-mail address: jose.quinteros@unimelb.edu.au (J.A. Quinteros).

http://dx.doi.org/10.1016/j.vetmic.2016.11.003
0378-1135/

important mechanisms underlying the emergence of new strains is genomic recombination (Jia et al., 1995). The emergent virulent strain Ark DPI appears to have originated from recombination between four different IBV strains (Ammayappan et al., 2008). Additionally, the emergence of new strains of IBV in the USA has been reported to have arisen from recombination events between the Massachusetts (Mass), Conn and Holte strains and field strains (Thor et al., 2011).

In Australia, IBV strains have been classified into three subgroups based on the nucleotide sequences of the genes encoding their S1 protein (Ignjatovic et al., 2006): subgroup 1, which includes vaccine or vaccine-related strains; subgroup 2, a relatively new group that emerged at the end of the 1980s and the beginning of the 1990s; and subgroup 3, which includes more recently isolated strains that have been shown to be derived from recombination between subgroup 1 and 2 strains (Mardani et al., 2010).

There have been several episodes of re-emergence of IB as a clinical problem in Australian chicken flocks, and it has been suspected that the emergent viruses may have originated from recombination events. The origin of subgroup 2 strains has remained unknown, but as the S1 gene of this subgroup shares less than 64% nucleotide (nt) sequence identity with those of subgroup 1 (vaccine related) strains (Sapats et al., 1996), it has been assumed that they may have arisen from a coronavirus exogenous to chickens (Mardani et al., 2008). The subgroup 3 strains, which were first isolated in 2002–2003, have an S1 gene that differs considerably from those of subgroup 1 and 2 strains, with nucleotide identities of 61–63% and 56–59%, respectively (Ignjatovic et al., 2006). Phylogenetic and similarity plot analyses of the genomic region encoding the structural genes have suggested that this subgroup emerged as a result of recombination between subgroup 1 and 2 strains (Ignjatovic et al., 2006; Mardani et al., 2010). The strain Ck/Aus/N1/08 (N1/08) was isolated from a flock of chickens with clinical signs of respiratory disease that had previously been vaccinated against IB. Analysis of the section of the genome encoding the structural proteins found evidence of recombination in the S1 gene, with the subgroup 2 strain N1/88 and a subgroup 3 strain as possible parental viruses (Hewson et al., 2014).

Although there is evidence of involvement of recombination within the structural and accessory genes in the emergence of new strains of IBV in Australia, little is known about the genomic region that contains the polymerase genes. Here, we sequenced the complete genomes of subgroup 1, 2 and 3 strains in order to detect potential recombination hot spots in the polymerase genes, and to investigate the potential involvement of recombination in these regions in the emergence of novel IBV strains.

2. Materials and methods

2.1. Virus strains and culture

Four different Australian strains of IBV were selected for this study: one from subgroup 1 (Armidale A3), a vaccine strain from subtype (subtype C) distinct from the most commonly used vaccine strain VicS (subtype B), as assessed by virus neutralisation, (Wadey and Faragher, 1981), one from subgroup 2 (IBV/Ck/Aus/N1/88, or N1/88) (Arvidson et al., 1991), one from subgroup 3 (IBV/Ck/Aus/N1/03, or N1/03) (Ignjatovic et al., 2006), and a relatively recent recombinant strain IBV/Ck/Aus/N1/08, or N1/08 (Hewson et al., 2014).

Viral culture was performed by inoculating the allantoic cavity of 8- to 9-day-old embryonated specific pathogen-free chicken eggs. After 48 h incubation, the allantoic fluid (AF) was aseptically collected. The AF extracted from the embryonated eggs was clarified and the virus purified as described previously (Lougovskaia et al., 2002), with minor modifications. Briefly, the AF was clarified by centrifugation at 2,500 × *g* for 20 min. The supernatant was then centrifuged at 100,000 × *g* for 2 h at 4 °C. Viral pellets were then resuspended in 200 μL of Tris-buffered saline (TBS) (pH 7.4), and this viral suspension was layered over a 30% to 55% continuous sucrose gradient in TBS. The gradient was centrifuged at 100,000 × *g* for 4 h at 4 °C. The virus band was resuspended in TBS and the viral particles pelleted by centrifugation at 90,000 × *g* for 1 h and the pellet resuspended in 250 μL of TBS. The genomic RNA was extracted using RNeasy kits (Qiagen), following the manufacturer's instructions for RNA extraction from cell cultures and tissues, with minor modifications (Asia Pacific Centre of Animal Health protocol for RNA extraction from allantoic fluid). In summary, the initial denaturation was performed by mixing 100 μL of the viral pellet with 400 μL of lysis buffer plus 5 μL of β-mercaptoethanol, and the mixture was incubated at 4 °C overnight. A 300 μL volume of 70% ethanol was added to the lysate. For the steps with washing buffers, the time and speed were changed from 15 to 30 s, and from 8000 to 10,000 × *g*. The final centrifugation at full speed was performed for 2 min instead of 1 min, and the speed during the elution step was modified from 8000 to 10,000 × *g*.

2.2. Complete genome sequencing

RNA was fragmented using RNase III (Life Technologies) and purified using magnetic beads. To construct the libraries, the RNA was reverse transcribed and the cDNA purified and amplified following the manufacturer's protocols, and then sequenced (Quinteros et al., 2015). Sequencing was performed using the PGM Ion Torrent™ platform (Life Technologies), with a 314 chip and the 200-base sequencing Ion OneTouch Kit v2. All the sequencing protocols were performed at the Monash Health Translation Precinct (MHTP) medical genomics facility, within the Monash Institute of Medical Research, Victoria, Australia. All reads were compared to chicken ribosomal RNA and mitochondrial genome sequences, and the matching reads were discarded. The remaining reads were mapped using the genome sequences of the Beaudette strain of IBV, two IBV vaccine strains from the United States (strains Conn46 1996 and Massachusetts; Genbank accession numbers FJ904716 and GQ504724, respectively), and two strains from China (SAIBK and SC021202; Genbank accession numbers DQ288927 and EU714029, respectively) as reference sequences, using Geneious version 6.1.4 (Biomatters). Because of the high level of sequence diversity of the S gene, the reads were also mapped using the previously determined sequences of the entire structural protein gene region (from the S glycoprotein to the 3′ end) of the Australian strains Armidale, N1/08, N1/88, VicS-del and VicS (accession numbers DQ490205, JN176213, DQ490207, JN983807 and JN176213, respectively), and the sequences of the S and N genes of N1/03 (accession numbers FJ235186 and FJ235194, respectively). All gaps and ambiguous sequences were resolved by Sanger sequencing of PCR products using BigDye® Terminator v3.1 (Applied Biosystems) kits with specific primers designed using Primer 3 (Untergasser et al., 2012). The list of primers are available upon request.

In order to confirm the sequences generated using this approach, a second assembly method was used. The reads were assembled *de novo*, using Geneious version 8.1.8 (Biomatters), allowing a maximum gaps per read of 10% and a maximum gap size of 2 nucleotides. The minimum overlap allowed was 75 nucleotides, with at least 95% identity in the overlapping region. The maximum proportion of mismatches per read allowed was 10%, with an ambiguity of 2 bases or less. All contigs were then mapped to the reference sequence, using at least 4 different levels

of stringency. These results were then aligned in order to obtain a final consensus, which was subsequently aligned to the homologous sequence that had been generated previously.

2.3. S1 glycoprotein (gp) gene phylogenetic tree

The S1 gp nucleotide sequences from every Australian IBV strain available (except for those generated in this study) were downloaded from Genbank. The amino acid sequences of the S1 proteins were then deduced using Geneious version 8.1.8, and aligned in Clustal W (Thompson et al., 1994) using the GONNET cost matrix, with the default settings (a gap opening cost of 10 and a gap extension cost of 0.1). The phylogenetic tree was constructed using the same parameters described previously (Stroehlein et al., 2015). The aligned residues were subjected to Bayesian inference analysis, using MrBayes version 3.2.2. (Ronquist et al., 2012). The posterior probabilities were calculated using a mixture of models with fixed rate matrices, generating 1,000,000 trees and sampling every 100th tree. The initial 25% of trees were discarded as burn-in, and the others were used to construct a majority rule tree. The Potential Scale Reduction Factor or PSRF (Gelman and Rubin, 1992) was 1.0, which means that the runs converged. As a comparison, the alignment was also subjected to a maximum likelihood analysis, using the Whelan and Goldman (WAG) model in MrBayes version 3.2.2 (Ronquist et al., 2012) and the parameters described above.

2.4. Detection of network relationships and recombination break points

2.4.1. Sequence alignment and network tree construction

A multiple sequence alignment of the complete genome sequences of Australian and international isolates of IBV was constructed with Clustal-Omega (Sievers et al., 2011), using the default settings. This included previously determined genomic sequences of two variants of the VicS vaccine strain (Quinteros et al., 2015). This was used to prepare a network tree using SplitsTree4 (Huson and Bryant, 2006), version 4.14.3, using the maximum likelihood (F84) distance model, the NeighborNet network algorithm, 1000 bootstrap replicates and an estimated proportion of invariant sites of 0.37. The tree included eleven sequences of strains from the United States (accession numbers GQ504725, vaccine Mass41; AJ311317, Beaudette; FJ904723, Mass41 1985; GU393332, Delaware 072; GQ504723, Georgia 1998 vaccine; GQ504721, Arkansas vaccine; AY514485, California 99; FJ904714, Cal 1995; FJ904716, Conn46 1996; GU393336, Holte; GU393338, JMK), five of which were vaccine strains. Thirteen sequences of strains isolated in China (JX195178, Ck/Ch/LDL/97I;

Table 1
Recombination breakpoints, and predicted major and minor parental sequences among other Australian infectious bronchitis virus (IBV) strains.

Potential recombinant	Breakpoints (99% CI)[c]		Genes	Major parent	Minor parent	Detection Method
	Start	End				
N1/08	1010 (937–1133)	2712 (2432–2744)	1ab (nsp2 and 3)	N1/03	Armidale	RDP, GENECONV, BootScan, MaxChi, Chimaera, SiScan, 3Seq
	19712 (19242–20175)	23273 (23259–23301)	1ab (nsp16), S1, S2	VicS-v	N1/88	RDP, GENECONV, BootScan, MaxChi, Chimaera, SiScan, 3Seq
N1/03	17055 (16740–17168)	18215 (18024–23576)	1ab (nsp14)	N1/08[a]	Armidale[a]	RDP, GENECONV, BootScan, MaxChi, Chimaera, SiScan, 3Seq
	20068 (19749–20236)	20378 (20350–20424)	1ab (nsp16), S1	Unknown (VicS-v)[a]	Armidale[a]	GENECONV, MaxChi, Chimaera, 3Seq
	21931 (21803–22342)	Indet[b]., 23255 (21803–22342)	S1 and S2	Armidale[a]	N1/08[a]	RDP, MaxChi, Chimaera, SiScan, 3Seq
	23602 (23453–23754)	24128 (23756–24161)	S2, 3a and 3b	VicS-del[a]	N1/08[a]	GENECONV, BootScan, Chimaera
N1/88	27139 (Indet[b].)	Indet[b]., 27383 (Indet[2].)	N, 3′ UTR	Unknown (N1/08)[a]	Armidale[a]	RDP, MaxChi, Chimaera, SiScan
VicS-v	Indet[b]., 3757 (Indet[b].)	Indet[b]., 4981 (Indet[b].)	1ab (nsp3)	N1/08[a]	N1/88[a]	RDP, GENECONV, BootScan, SiScan, 3Seq
	Indet[b]., 4981 (4797–4986)	12387 (12287–12476)	1ab (nsp3 to 12)	Armidale[a]	N1/88[a]	RDP, GENECONV, BootScan, MaxChi, Chimaera, SiScan, 3Seq
	Indet[b]., 24232 (23427–24383)	24589 (24515–24917)	E and M	N1/08[a]	Unknown (Armidale)[a]	GENECONV, BoostScan, MaxChi, Chimera, 3Seq
Armidale	18699 (Indet[b].)	Indet[b]., 19549 (19231–23679)	1ab (nsp15 and 16)	N1/03[a]	N1/88[a]	RDP, GENECONV, BootScan, MaxChi, Chimaera, SiScan
	23450 (18446–23583)	Indet[b]., 24166 (Indet[b].)	S2, 3a and 3b	Unknown (VicS-del)[a]	N1/08[a]	RDP, GENECONV, BootScan, MaxChi, Chimaera, 3Seq

[a] One or both parental strains may be the actual recombinant.
[b] Indet.: Indeterminate.
[c] Nucleotide numbers in the breakpoints refers to the recombinant genome sequence.

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; KP118885, Ck/Ch/LSD/110857; KP036503, Ck/Ch/LHB/121010; and HQ850618, GX-YL9), two sequences of strains from India (KT203557, B17 and KR902510, Ind/TN92/03), Italy (KP780179, Italy/I2022/13 and FN430414, Ita/90254/2005), the Netherlands (EU817497, H52 and GU393335, H120) and South Korea (JQ977698, KM91and JQ977697, SNU8067) and individual isolates from Nigeria, Sweden, Taiwan, Poland, Belgium and Ukraine (FN430415, Ibadan; JQ088078, Ck/Swe/0658946/10; DQ646405, TW2575/98; KT886454, Poland/74/2009; KR231009, B1648 and KJ135013, Ukr27-11, respectively), as well as the complete genome sequence of the vaccine 4/91 (KF377577). Three turkey coronavirus (TCoV) sequences were also included as an outgroup (EU022525, TCoV-540; EU022526, TCoV-ATCC; GQ427174, TCoV/TX-GL/01). An additional two gammacoronaviruses isolated from non-gallinaceous species (JF705860, Duck CoV; and AY641576, Peafowl) were included to compare with the Australian strains. A second network tree was also constructed using only the nucleotide sequences of the Australian strains, using the same parameters, but with an estimated proportion of invariant sites of 0.2.

2.4.2. Similarity plots

Similarity plots were created using the on-line tool MUltiple sequence Local AligNment (Mulan) (Ovcharenko et al., 2005). The evolutionarily conserved regions (ECR) in the graph were estimated using a window length of 10 bases, a similarity threshold of 90% and a bottom cut-off of 50%. The genome sequence of N1/08 was used as the query to construct the five plots.

A pairwise alignment using the structural gene nucleotide sequences of the strains Armidale, VicS, N1/88 and N1/03 was performed to create similarity plots using either the genome sequences of the Armidale or N1/03 strain as the query. The similarity plots were constructed using the maximum likelihood distance model, an empirical transition/transversion (Ts/Tv) ratio calculated for each window, a window size of 200 bp and a step size of 20 bp. The analyses were performed using SimPlot, version 3.5.1 (Lole et al., 1999).

2.4.3. Detection of recombination breakpoints

The strains isolated in Australia were examined using Recombination Detection Program 4 (RDP4, version 4.56) (Martin et al., 2010) for evidence of recombination breakpoints. The seven methods used for the analysis were RDP (Martin et al., 2005), GENECONV (Padidam et al., 1999), BootScan/Rescan (Martin et al., 2005), Maximum Chi Square (MaxChi) (Smith, 1992), Chimaera (Posada and Crandall, 2001), SiScan (Gibbs et al., 2000) and 3Seq (Boni et al., 2007). Automasking was used for the optimal detection of recombination breakpoints (Thor et al., 2011). Default settings were used for most algorithms. Changes from default settings were the window size for RDP (60) and Bootscan (500), the window and step size for SiScan (500 and 20, respectively), and the number of variable sites per window for MaxChi and Chimaera (120). The size of the window was increased from the default to reduce masking of recombination signals within IBV, which has a high mutation rate (Thor et al., 2011). The recombination breakpoint plot was

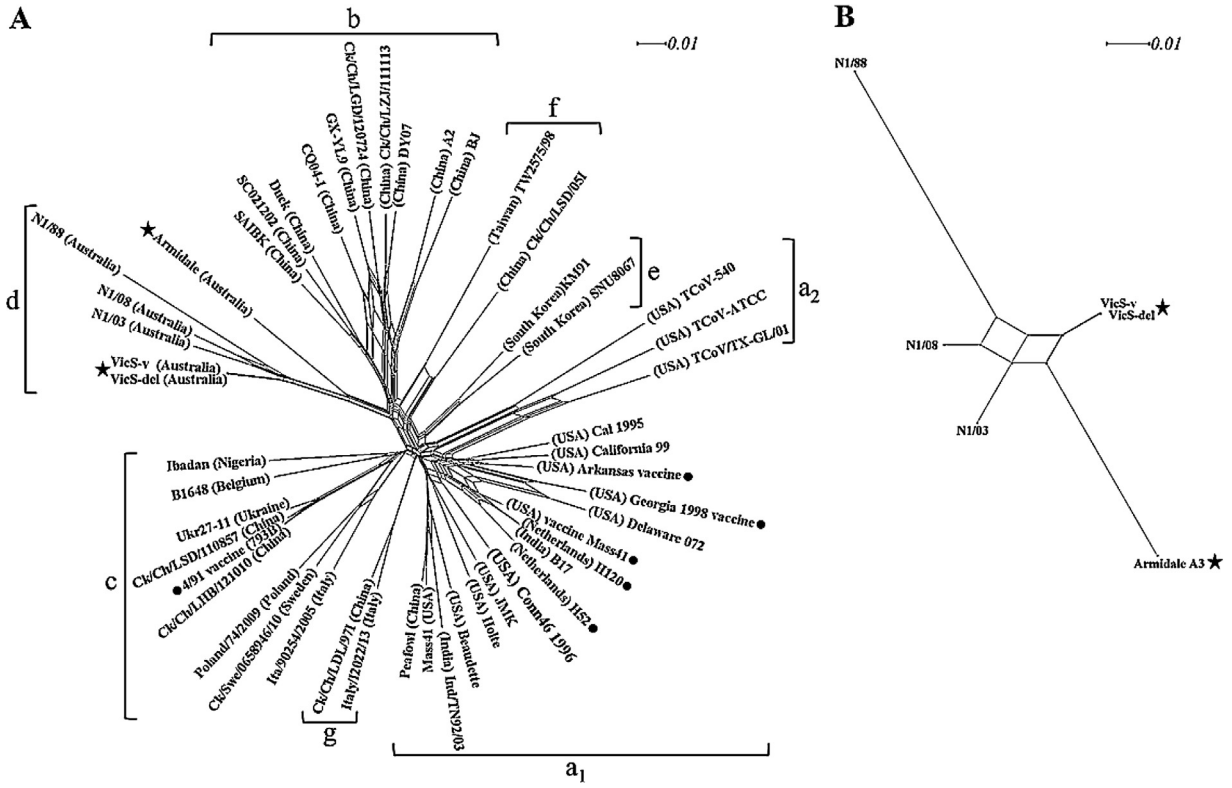


Fig. 1. Network trees (SplitsTree4, Huson and Bryant (2006), version 4.14.3) A: Constructed with the complete genome sequences of 50 IBV strains, including the 4 Australian strains sequenced for this study. B: Constructed with the complete genome sequences of Australian strains of IBV. The networked relationships indicate the presence of reticulation events. Boxes indicate the likelihood of recombination. Group a1, USA cluster, with the exceptions of Ind/TN92/03 (India), B17 (India), Peafowl (China), H52 and H120 (the Netherlands); a2, TCoV subcluster; b, Chinese cluster; c, 793B (4/91)-like strains cluster; d, Australian cluster; e, South Korean cluster; f, Taiwanese cluster; g, strains from China and from Italy, closely related to the 793B-like cluster. The stars mark the positions of Australian vaccine strains and the dots mark the positions of other vaccine strains. A phi test for each of the network trees indicated that there was a significant likelihood of recombination ($P < 0.001$), using a window size of 100.

calculated with a window of 200 nt. The breakpoint plot was constructed using a 200 bp window moved 2 nucleotides at a time along the length of the genome (Fig. 4). The breakpoints selected were those detected with 3 or more recombination detection methods at a $P < 0.05$ (Table 1).

3. Results

3.1. Genome sequences

The genome sequences of the four Australian IBV strains were determined and deposited in Genbank with the accession numbers KU556805 (Armidale), KU556804 (N1/88), KU556806 (N1/03) and KU556807 (N1/08). The genomes of these strains were 27.673, 27.165, 27.626 and 27.688 kb in lengths for Armidale, N1/88, N1/03 and N1/08, respectively. The nucleotide similarities of the consensus sequences obtained with the two assembly protocols (mapped to sequence and de novo) were: Armidale: 99.2%; N1/03: 99.8%; N1/88: 98.9%; N1/08: 98.5%. Data previously obtained for VicS-del and VicS-v were similarly analysed, and the similarities between the assemblies were 99.3% for VicS-del and 98.5% for VicS-v. The regions that were not similar corresponded to ambiguous sequences. The sequences deposited in Genbank are those deduced using the mapped to sequence protocol.

3.2. Phylogenetic tree

In the network tree (Fig. 1A), most of the IBV isolates were distributed into 7 main clusters. The six Australian strains included in this analysis formed a separate clade, with the Armidale strain branch diverging from the main cluster, with only limited recombination with any of the other five Australian strains included. This was also seen in the case of the N1/88 strain. Notably, there was little evidence of any interaction between the strains in the Australian clade and strains in clades from other countries. Furthermore, the Australian clade is equally or more divergent, compared to the other IBV genomes included in this study, than the TCoV clade (particularly the N1/88 strain). The network tree constructed using only Australian isolates (Fig. 1B) confirmed the recombination events that have been detected between these isolates and also highlighted how the Armidale and N1/88 strains have diverged. Both network trees provided strong evidence of recombination and yielded a highly significant phi test ($P < 0.001$). The strains included as outgroups (duck, turkey and peafowl coronavirus) were not phylogenetically related with the Australian strains, but with strains from China and the USA (Fig. 1A).

In the S1 phylogenetic tree, both Bayesian and maximum likelihood (WAG) models gave almost identical results (Supplementary Fig. 1), so the Bayesian phylogenetic tree was chosen for the figure. The S1 phylogenetic tree contained three clusters. The first (Fig. 2A) corresponded to vaccine and vaccine-related strains (subgroup 1), while the other two clusters (Fig. 2B and C) to subgroups 2 and 3, respectively. Three strains were removed from the alignment before the inference of the phylogenetic tree, as they were identical (100%) to other strains (026 to Strain B, and 044 and 013 strains to VicS-v).

3.3. Areas of recombination

Fig. 3 shows a schematic representation of the origin of the different sections of the genome of the three emergent recombinant Australian IBVs isolated from major outbreaks in 1988 (N1/88), 2003 (N1/03) and 2008 (N1/08).

The genome of strain N1/88 was highly similar (over 95% identity) to that of strain VicS from the 5′ end of the genome to the end of ORF 1a (nt 12,000). There were two small segments with high similarity to VicS in ORF 1ab, between nt 12,500 and 13,200, and between nt 16,800 and 17,800 (98.1% and 99.5% identity, respectively). There was a small segment at the 3′ end of ORF 1ab,

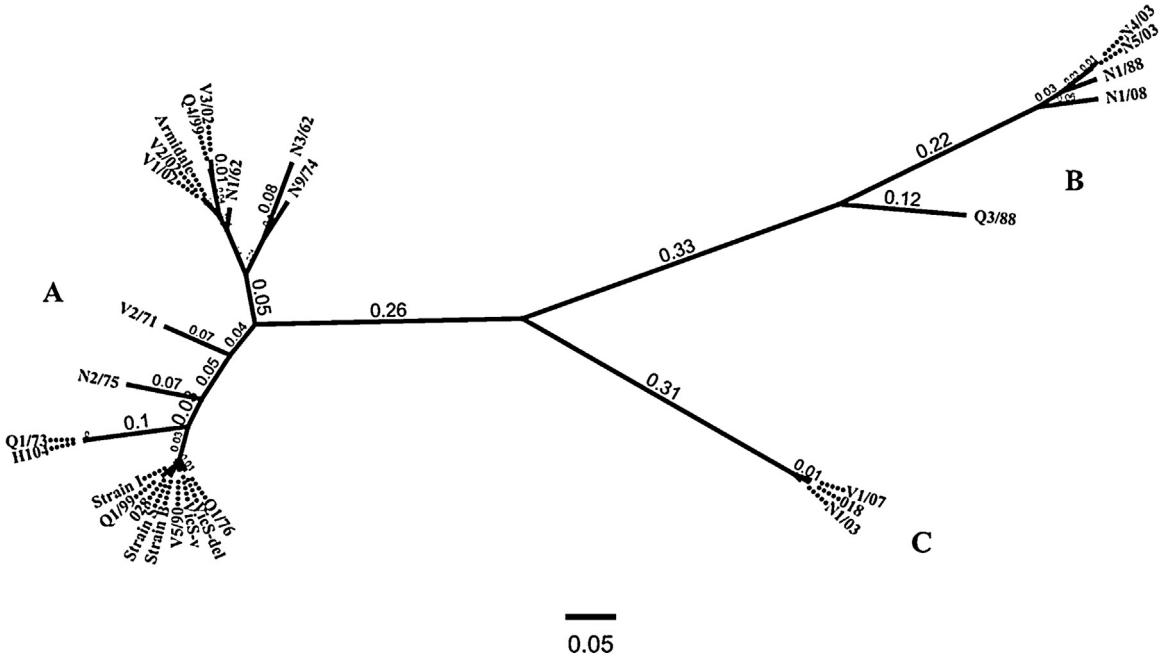
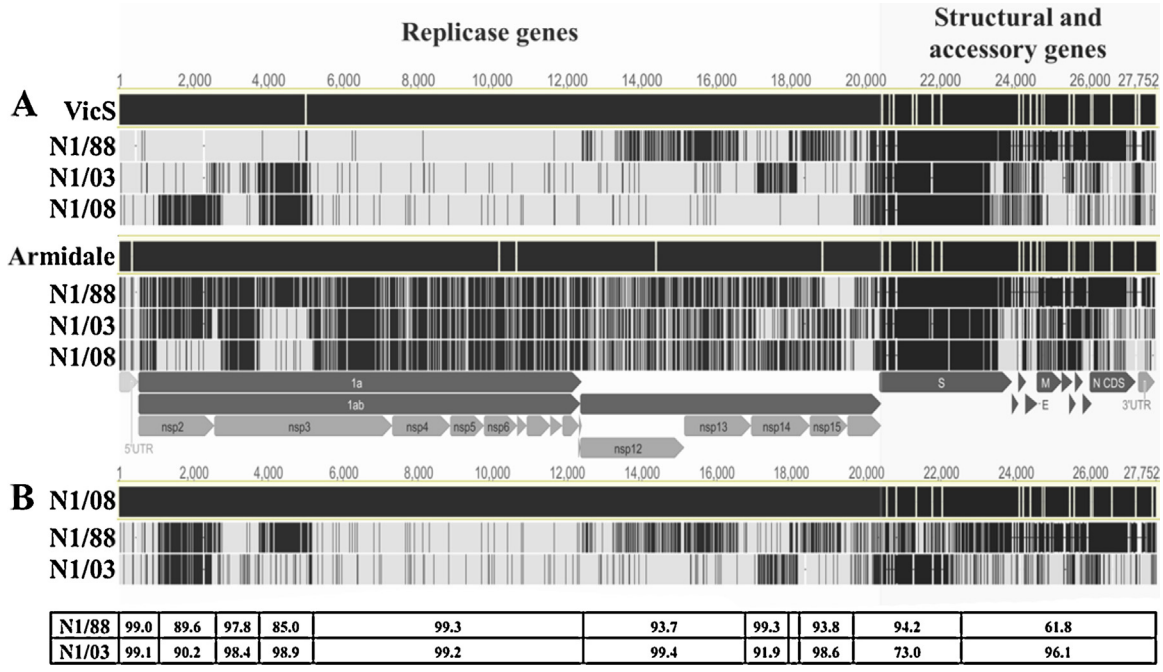
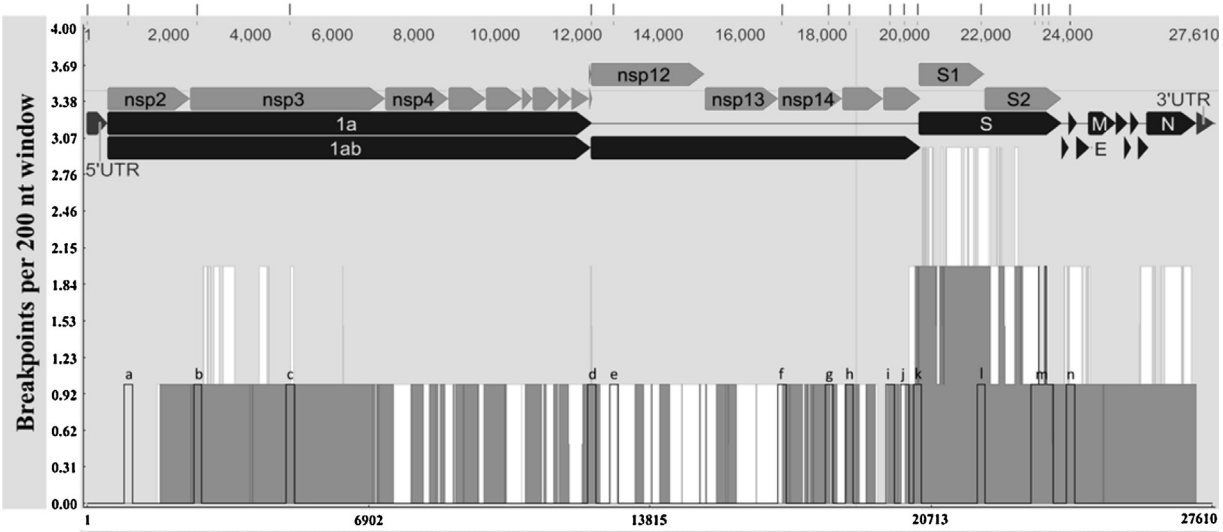


Fig. 2. Phylogenetic tree constructed with the S1 gp deducted amino acid sequences of the Australian strains of IBV. The alignment was constructed using Clustal W (Thompson et al., 1994), with the GONNET cost matrix. The aligned sequences were subjected to Bayesian inference analysis, using MrBayes v.3.2.2 (Ronquist et al., 2012), generating 1,000,000 trees. A: subgroup 1 or classical strains (vaccine and vaccine-related strains); B: subgroup 2 or novel strains; C: subgroup 3 strains. The dots located at the end of the tree branches were added to increase clarity.



N1/88	99.0	89.6	97.8	85.0	99.3	93.7	99.3	93.8	94.2	61.8
N1/03	99.1	90.2	98.4	98.9	99.2	99.4	91.9	98.6	73.0	96.1

Fig. 3. A: Nucleotide alignments, with VicS (on top) and Armidale (below) as the reference sequences, demonstrating the origins of sections of the genomes of three emergent recombinant infectious bronchitis viruses isolated from major outbreaks in 1988 (N1/88), 2003 (N1/03) and 2008 (N1/08). Nucleotide disagreements with the reference sequences are highlighted in black. Major sections of these viruses were derived from the two vaccine strains used in Australia, VicS and Armidale. B: Nucleotide alignment, with N1/08 set as the reference sequence, demonstrating the nt similarities of different sections of the genome of the emergent recombinant infectious bronchitis virus N1/08 with those of N1/88 and N1/03. Nucleotide disagreements with the reference sequences are highlighted in black. The table indicates the level of similarity (%) in each section. The arrows below each reference strain indicate the positions of the open reading frames and small peptides of the IBV genome. The nucleotide numbers refer to the VicS alignment.



Interval values for each breakpoint

a	911-1109	b	2613-2811	c	4881-5081	d	12289-12487	e	12833-13031	f	16953-17151	g	18113-18311
h	18605-18803	i	19607-19805	j	19969-20167	k	20279-20477	l	21845-22043	m	23163-23707	n	24035-24233

Fig. 4. Recombination breakpoint distribution plot for the Australian IBV strains VicS-v, VicS-del, Armidale, N1/88, N1/03 and N1/08, predicted using RDP4 (Martin et al., 2015) showing the detectable recombination breakpoints. The predicted recombination breakpoints are marked at the top of the figure. The dark grey and white areas are 95% and 99% confidence intervals, respectively, for local breakpoint clusters. The black lines indicate the breakpoint count within the 200 bp window, and are marked with the letters a to n. The interval values for each breakpoint are indicated in the table below the figure. The panel on top of the figure indicates the positions of the open reading frames of the IBV genome. The light grey arrows indicate the regions encoding the mature peptides, while the dark grey arrows indicate the CDSs. The names of some genes have been removed for clarity (nsp5 to 11, 15 and 16, and accessory proteins 3a, 3b, 4b, 4c, 5a and 5b).

between nt 19,000 and 19,500, that had over 99% sequence identity with the Armidale strain. The remainder of the genome, encoding all the structural and accessory genes, had low levels of similarity with both of the vaccine strains.

The genome of strain N1/03 had over 94% similarity with VicS from the 5′ end of the genome to nt 3800. From this point to nt 5000, it was 99.5% identical to the Armidale strain, after which it was again very similar to VicS, between nt 5000 and 17,000. At this point there was a third crossover and the similarity to Armidale increased to over 95%. The similarities changed again at nt 18,200, with similarity to VicS increasing to over 99% (all inside ORFs 1a and 1ab). Between nt 20,000 and 23,500, the similarity of the sequence of N1/03 to either vaccine strain was less than 77%. From this point to the 3′ end of the genome, corresponding to the region in which all the structural and accessory genes, except for the peplomer, are encoded, it was equally similar to both strains (93.5% and 93.4% identity with VicS-v and Armidale, respectively).

The genome of strain N1/08 had a high level of sequence identity with that of N1/03 (95.2%), except between nt 1000 and 2700, where N1/08 had a high level of sequence identity with the Armidale strain rather than VicS, and between nt 17,000 and 18,200, where it had 99.9% identity with VicS, and another small region between nt 19,500 and 20,100, where it had 99.3% identity with Armidale.

Fig. 4 depicts the recombination breakpoints as calculated using RDP4. There were 14 recombination points detected along the genomes of these six strains. These breakpoints were located in genes 1ab (nsps 2, 3, 10, 11, 12, 14, 15 and 16), S1, S2, 3b and E, and were consistent with the data previously described and depicted in Fig. 3.

3.4. N1/08 full-genome sequence analysis

The similarity plot (Fig. 5) of isolate N1/08 with each of the other 5 Australian strains revealed the considerable complexity of the recombination events that have generated strain N1/08. Its genome shared a high level of identity with strain N1/03 in the regions at the 5′ end of the genome (from nt 1 to 940; over 98%), in the area encoding non-structural protein (nsp) 3 to nsp13 (from nt 2741 to 17,001; over 94%), part of nsp14, nsp15 and the beginning of nsp16 (from nt 18,221 to 19,601; over 98%), and from the region encoding the S2 glycoprotein (S2gp) until the 3′ end of the genome (from nt 22,481; over 88%). The sections of the genome between nt 2741 and 19,601, and from nt 22,481 to the 3′ end were also highly

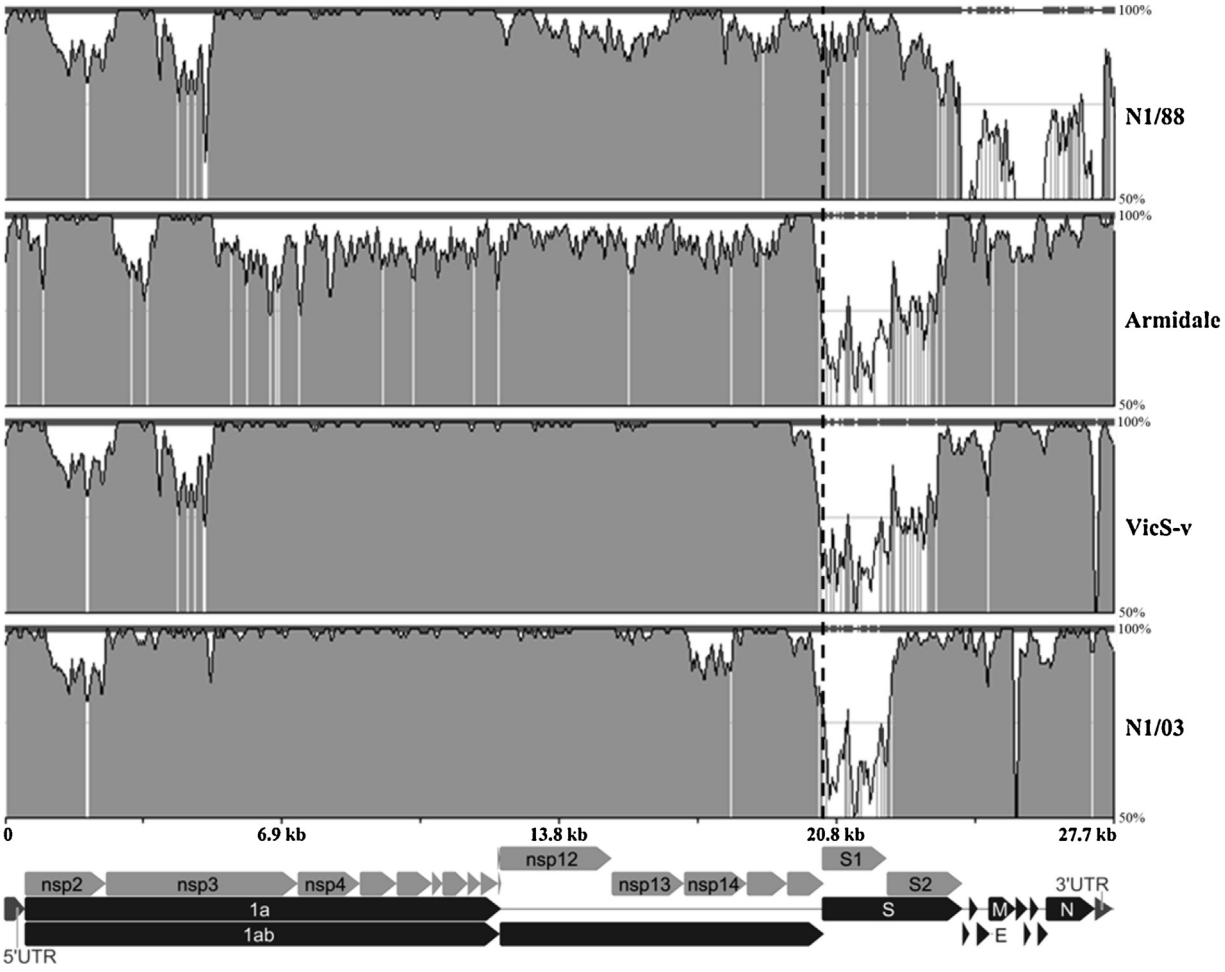


Fig. 5. Similarity plots of the full genome sequences of six Australian IBV strains. N1/08 was used as the query sequence. The area under the line indicates the level of sequence identity (%) between each strain and the query along the length of the genome. The dark and light arrows at the bottom indicate the positions of the CDSs and the mature peptides, respectively. The names of some genes have been removed for clarity (nsp5 to 11, 15 and 16, and accessory proteins 3a, 3b, 4b, 4c, 5a and 5b). The dashed line indicates the boundary between the polymerase genes and the structural genes. The nucleotide alignment and the similarity plots were constructed with multiple sequence local alignment (Mulan) (Ovcharenko et al., 2005), with a sliding window (ECR length) of 10 bases, an ECR similarity of 90% as the minimum limit, and a lower cut-off of 50%.

similar to VicS-v (98.5% and 92.1% identity, respectively). It also shared a high level of identity with the Armidale strain from nsp2 to the beginning of nsp3, between nts 1061 and 2621 (over 98%) and at the end of the nsp genes, between nts 19,581 and 20,121 (over 99%), immediately upstream of the S1gp gene. The level of identity with VicS-v and VicS-del was over 98% in the nsp14 gene region, between nt 17,021 and 18,241. N1/08 had a high level of nt sequence identity with N1/88 in the region encoding nsp14 (97.4%) between nt 16,926 and 18,488, especially in the region encoding the N-terminus between nt 16,926 and 17,934 (99.4%). N1/88 also had a high level of identity with N1/08 in the region encoding the S1gp (in Fig. 5, the first half of the S CDS), between nt 20,141 and 22,941 (88%), while the other strains included in this study had sequence identities with N1/08 below 68% in this region (Fig. 5). These observations were supported by the RDP4 analyses, which detected two recombination breakpoints in this region (Table 1). The first was located between nt 1010 (99% CI; 937–1133) and 2712 (99% CI; 2432–2744), with N1/03 and Armidale proposed as the major and minor parent strains. The second breakpoint was located between nt 19,712 (99% CI; 19,242–20,175) and 23,273 (99% CI; 23,259–23,301), with VicS-v and N1/88 as the proposed parent strains.

The similarity plots (Fig. 5) comparing N1/88 (upper panel) and N1/03 (lower panel) with N1/08 show that N1/88 was equally or less similar to N1/08 than N1/03 was to N1/08 across the genes encoding the polymerase genes, except for the nsp14 gene, where N1/88 was more similar to N1/08 than N1/03. Analysis of the predicted amino acid sequences of nsp14 (Fig. 6) revealed that

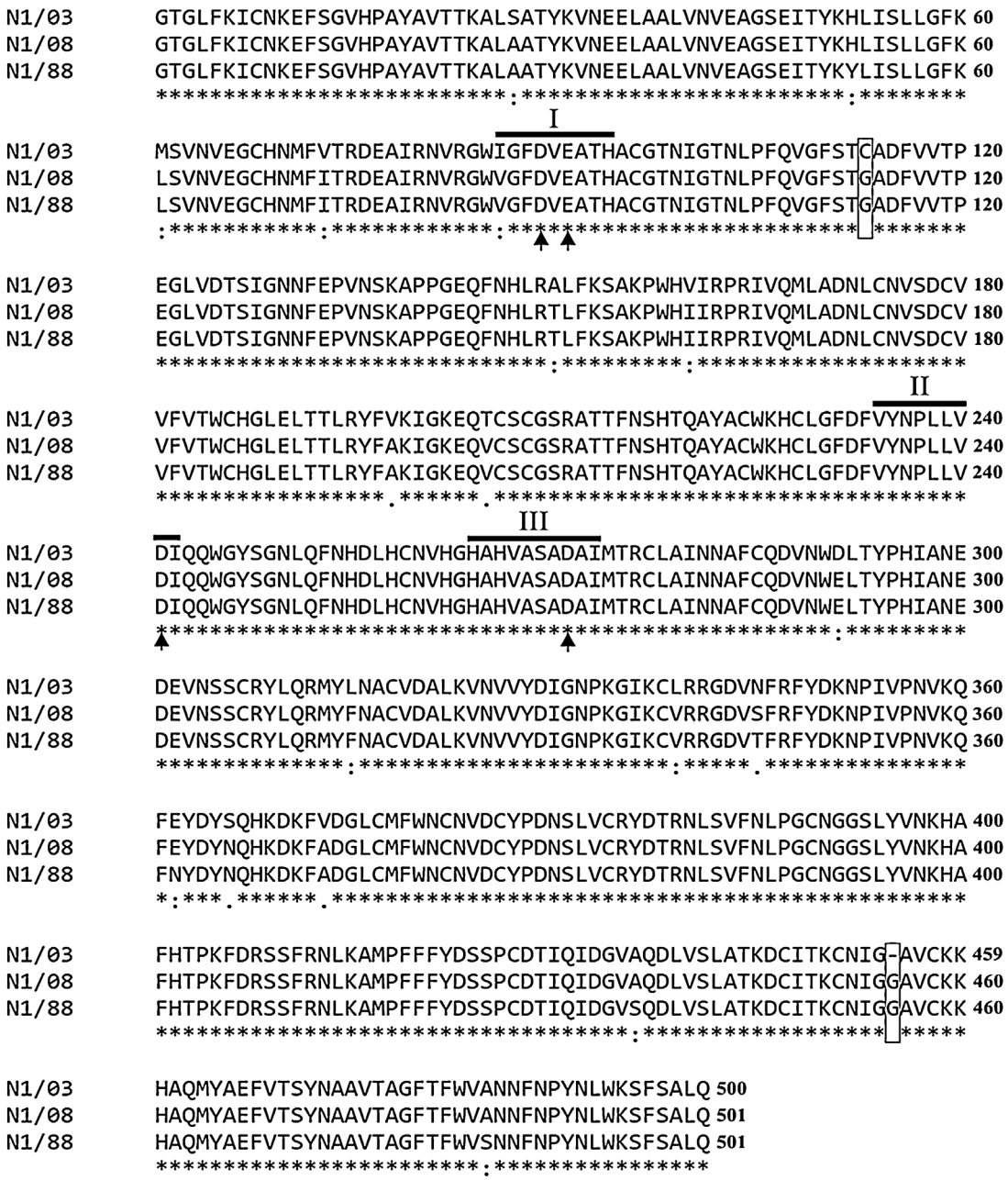


Fig. 6. Multiple alignment of the amino acid residues encoded by the non-structural protein 14 from the strains N1/88, N1/03 and N1/08, constructed with Clustal Omega (http://www.ebi.ac.uk/Tools/msa/clustalo), using the default settings. The black lines highlight the location of motifs I, II and III. The arrows mark the amino acid residues D-E-D-D. The grey boxes highlight the location of the main amino acid differences.

there were differences within motif I, where the first amino acid residue in N1/88 and N1/08 was valine, while in N1/03 it was isoleucine, and at residue 113, where N1/88 and N1/08 both have cysteines, whilst N1/03 has a glycine, and residue 455, where there is a glycine deletion in N1/03.

4. Discussion

Infectious bronchitis virus is a highly prevalent virus, causing economic losses in the chicken industry globally (Jackwood and de Witt, 2013). Infectious bronchitis virus strains have been continuously monitored in Australian chickens since the first characterisation and isolation of the aetiological agent of 'nephrosis' syndrome in the 1960s (Cumming, 1962). One of the principal milestones of this monitoring was the discovery of the subgroup 2 strains at the end of the 1980s and the beginning of the 1990s (Sapats et al., 1996). These viruses had S1gp and N protein gene sequences that were very distinct from those of the strains that had been encountered previously. This clade of strains also has a translocation and deletion of some accessory protein genes (Mardani et al., 2008; Hewson et al., 2011). However, since only the structural and accessory genes in the 7.5 kb region at the 3′ end of the genome had been sequenced previously, the available information suggested that the origin of the subgroup 2 strains was from a coronavirus completely distinct from the strains already circulating in Australian chicken flocks. The determination of the complete sequences of the polymerase genes presented in this study now suggests that N1/88 originated in part from IBV strains already circulating in the field, as well as from an avian coronavirus of unknown origin. Phylogenetic analysis of different strains of IBVs isolated from different countries (Fig. 1) supported previous suggestions that Australian strains of IBV have been diverging as an independent cluster (Quinteros et al., 2015). However, the representation of genetic divergence by a phylogenetic tree alone may not be accurate, because once a branch is created no further potential interactions between the branches are represented (Thor et al., 2011). In network trees, boxes represent the likelihood that there has been recombination between two different branches. The IBV Network Splitstree (Fig. 1) suggested considerable intra-clade interactions, with a significantly lower number of inter-clade interactions, supporting suggestions that Australian IBVs have been genetically isolated for a considerable time (Ignjatovic et al., 2006; Quinteros et al., 2015). This could be explained by the geographic location of Australia as an island and the enclosed nature of the Australian ecosystem, and also by the biosecurity barriers imposed by Australian authorities on biological imports. It is likely that IBV vaccine strains may have played a significant role in intra-clade recombination events in Australia. Because of the high reliance on live attenuated vaccines in Australia, there is a high probability that co-infection with a field strain and a vaccine strain could occur. The chimaeric pattern of relationships between the three Australian isolates of IBV characterised here and the two most distinct vaccine strains in use suggests that this has happened multiple times, as summarized in Fig. 3.

As has been reported previously, there is a high level of divergence between the sequence of N1/88 and the vaccine strains in the region of the genome encoding the structural and accessory proteins (downstream of nt 20,000) (Mardani et al., 2008). In the region encoding the polymerase genes, the genome sequence was highly similar to VicS. This suggests that, for the emergent N1/88 strain, the donor of the polymerase genes was most likely to have been VicS or a VicS-like strain, while the donor of the structural protein genes remains unknown, as reported previously (Mardani et al., 2008). This may suggest that the polymerase genes from the parental virus that served as the source of the structural genes were poorly adapted to chicken cells, and that the emergence of N1/88 was driven not only by the antigenic differences between this virus and other circulating IBVs, but also by an enhanced capacity of the recombinant virus to replicate in chickens. A similar scenario could be behind the emergence of N1/08. A previous study comparing the sequences of the N1/08 structural genes suggested that the emergence of this virus could be due to a recombination event between a subgroup 2 strain (N1/88) and a subgroup 3 strain (N1/03) (Hewson et al., 2014) (Fig. 4; last third of the genome, from nt 20,000 to the 3′ end). However, it is possible that N1/08 could have resulted from an independent recombination event with VicS-v, rather than with N1/03, with recombination breakpoints located between 181 and 1114 bp upstream of the start codon of S1gp (nt 19,242–20,175) and between 1278 and 1320 bp downstream of the protease cleavage site (nt 23,259–23,301) for the S1 and S2 gp genes (Table 1, Fig. 4, letters j and m). However an N1/03-like strain, which could be a recombinant derived from N1/03 and Armidale, could also be one of the parental strains, with N1/08 derived from recombination between this virus, VicS and N1/88 (Fig. 5 and Table 1). It is difficult to infer the exact order of these recombination events, as the recombinational history of N1/08 appears to be extremely complex.

Although the S1gp gene sequence of N1/08 is very similar to that of N1/88, the subgroup 2 strains have not been isolated from any IB outbreak in Australia for over 20 years. Therefore, it is possible that the unknown avian coronavirus donor of the structural and accessory genes of N1/88 is continuing to recombine with IBV vaccine strains, leading to the emergence of new strains. The detection of this "unknown avian coronavirus" donor will be crucial to better understanding the origin and evolution of these new recombinant strains. There is prior evidence of interspecies recombination in avian coronaviruses, in the emergence of turkey coronavirus (TCoV). The genome of TCoV is very similar to that of IBV, except for the gene for the S gp, and, as in the case with the emergence of N1/88, the parental donor of the spike glycoprotein gene remains unknown (Lin et al., 2004; Jackwood et al., 2010). These interspecies recombination processes have also been described in other members of the family *Coronaviridae.* It has been suggested that Rp3, a severe-acute respiratory syndrome (SARS)–like coronavirus from bats, may have originated by recombination between another SARS-like coronavirus (Bt-SLCoV Rm1) and a human SARS coronavirus (Hu-SCoV) (Hon et al., 2008). Interestingly, the recombination breakpoint in Rp3 is located around the ORF 1b/S junction, which is the site at which recombination between the VicS-like parental virus and the unknown parental virus of N1/88 occurred (Fig. 3).

As it can be seen in the phylogenetic tree of S1 (Fig. 2), the Australian strains are divided in three different clusters, and N1/88 (subgroup 2), N1/03 (subgroup 3) and N1/08 (subgroup undetermined) belong to 2 clusters separated from the main one, where all the vaccines and vaccine related strains can be found. N1/08 lies in the same cluster as N1/88, implying that the unknown avian coronavirus donor of the S1 gene of N1/88 could still be recombining with IBV vaccines or field strains to generate new recombinants.

Previous experimental inoculation studies in chickens have revealed that the N1/08 strain has tropism for tracheal tissue, inducing only mild lesions, as was seen in earlier experiments with the subgroup 2 strain N1/88 (Ignjatovic et al., 1997; Hewson et al., 2014). In contrast, the subgroup 3 strain N1/03 causes moderate to severe tracheal lesions and clinical signs, together with an increase in mortality and reduced growth rates (Ignjatovic et al., 2006). The analysis of the nt similarities between N1/88 and N1/08, and between N1/03 and N1/08, to determine genetic factors that may be associated with these differences in phenotype revealed a high similarity between the S1gp nt sequences of N1/88 and N1/08

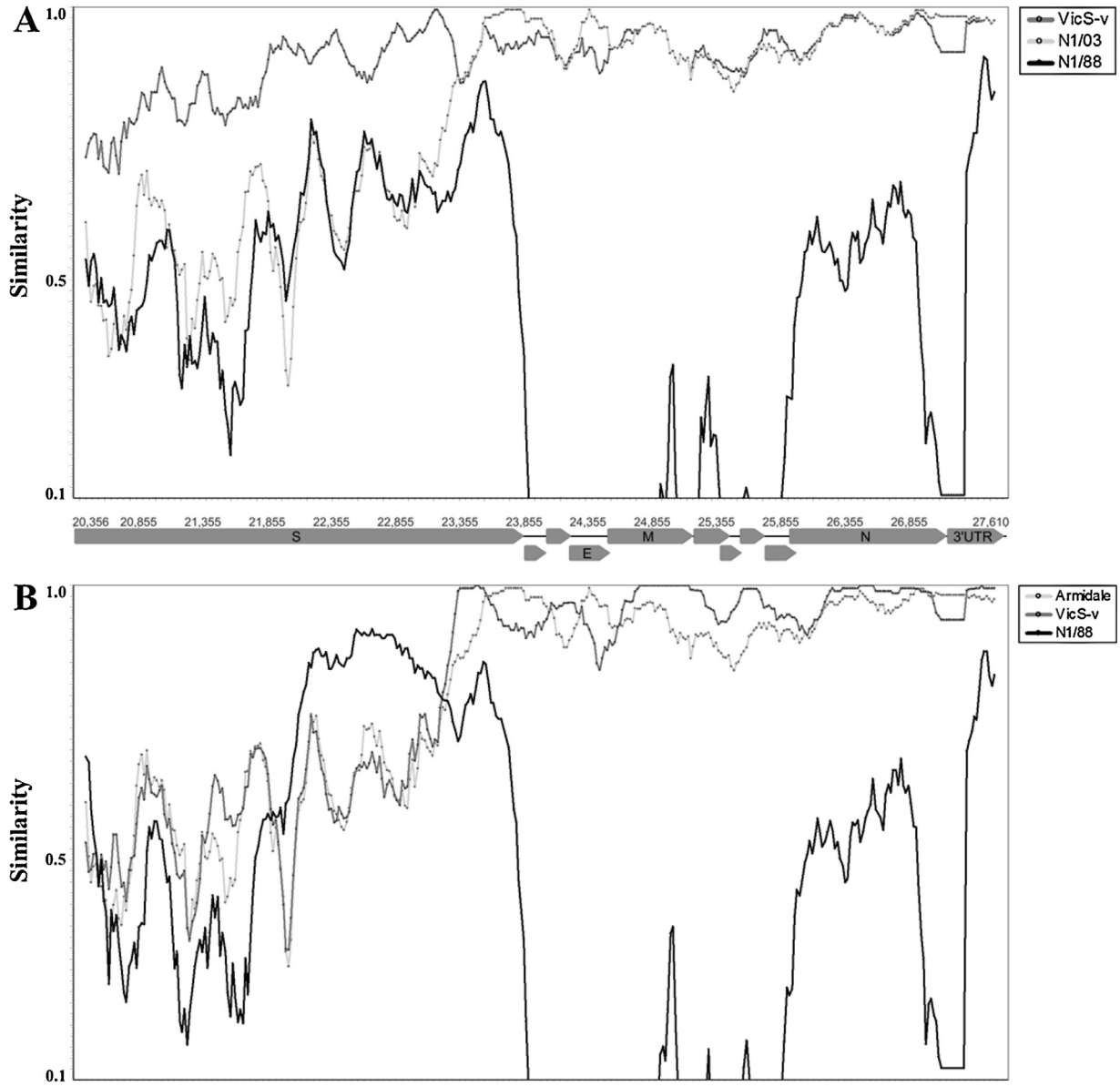


Fig. 7. Similarity plots of the structural gene region of Armidale and VicS (subgroup 1), N1/88 (subgroup 2) and N1/03 (subgroup 3) strains created with SimPlot, version 3.5.1 (Lole et al., 1999). A: Armidale strain sequence used as the query, as in Mardani et al. (2010). B: N1/03 strain sequence used as the query. The arrows indicate the positions of the CDSs and the 3′ UTR. The 3a, 3b, 4b, 4c, 5a, 5b and 6b genes are not annotated to aid clarity. The positions of genes in N1/88 differs from this depiction, as some of them have been translocated within the genome (Hewson et al., 2011).

(93.8%), and much lower similarity with the S1 gp gene sequence of N1/03 (65.3 (Hewson et al., 2014). However, it has been demonstrated that the replicase genes, rather than the structural genes, may be more closely associated with the level of virulence of an IBV strain (Armesto et al., 2009). The only peptide gene in the polymerase region in which N1/08 is more similar to N1/88 than to N1/03 is the nsp14 gene. Nsp14 has been shown to have 3′-to-5′ exoribonuclease (ExoN) and guanine-N7-methyltransferase (N7-MTase) activity in other members of the family *Coronaviridae* (Chen et al., 2009, 2013; Smith and Denison, 2012). The ExoN belongs to the DEDD superfamily (Asp-Glu-Asp-Asp), and these conserved residues (Smith and Denison, 2012) are found in the nsp14 of 75 strains of IBV at positions 89, 91, 241, 271 respectively (data not shown). There are also 3 highly conserved amino acid motifs (I, II and III) in the nsp14 of all coronaviruses with the DEDD signature (Minskaia et al., 2006). Studies in SARS-CoV have revealed that changes in the amino acid residues within motif I impair viral growth rates and lead to higher levels of viral mutations in progeny after several passages in cell culture (Eckerle et al., 2010). The predicted amino acid sequence alignment of N1/88, N1/03 and N1/08 revealed differences in motif I and in the residues at positions 113 and 455 (Fig. 6). It is possible that these changes could lead to conformational changes in this protein, especially the cysteine to glycine substitution, which could alter the function of either the ExoN or N7-MTase. However, it is more likely that these changes, in combination with other small changes in the polymerase peptides are jointly, rather than individually, responsible for the differences in virulence in these IBVs.

A previous study based on analysis of the S gp gene suggested that subgroup 3 viruses could have emerged as a result of recombination between subgroup 1 and 2 strains (Mardani et al., 2010). However, similarity plot analysis of the isolates indicates

that this is not likely to be the case. The nucleotide sequence of the S gene of N1/03 was equally dissimilar to both subgroup 1 and 2 strains. When the Armidale strain (subgroup 1) was used as the query, the S1 genes of N1/88 and N1/03 appeared to be similar. However, as shown in Fig. 7B, when N1/03 (subgroup 3) was used as the query the sequence of its S gene was found to be quite distant from that of strain N1/88, especially in the S1gp region (the first half of the S gene), with a similarity below 50%. Thus the origin of the subgroup 3 strain is most likely due to a recombination event between subgroup 1 strains and an unknown S1 gene donor that was quite distant from the donor of the same gene in the subgroup 2 strains, probably another unknown avian coronavirus, suggesting that there are multiple distant avian coronaviruses with the capacity to recombine with IBVs and contribute to future IBV diversity.

The genome sequence of the structural proteins of N1/08 suggests that, if it was used as a vaccine strain, could offer broader immunological protection than the Australian vaccine strains currently in use. The protein that induces neutralising antibodies against IBV is the S1gp, and these antibodies confer protection against disease (Cavanagh et al., 1986; Ignjatovic and Galli, 1994). All the Australian vaccines belong to subgroup 1 and the amino acid sequence of the S1gp of this subgroup differs from those of subgroup 2 and 3 viruses (Ignjatovic et al., 2006; Hewson et al., 2011). Although antibodies induced by the N protein are not protective (Ignjatovic and Galli, 1994), peptides derived from it are recognised by T cells, and it is thus believed to play a role in the induction of cell mediated immunity (Boots et al., 1991; Saif, 1993; Seo et al., 1997). N1/08 is the only strain characterised thus far that shares a high sequence identity in either its S1 and/or N genes with members of all the Australian subgroups of IBV. The potential for this virus to induce cross protection against members of all other subgroups needs to be assessed to validate this hypothesis.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.vetmic.2016.11.003.

References

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. J. Virol. 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 (10), e7384.

Arvidson, Y., Tannock, G.A., Zerbes, M., Ignjatovic, J., 1991. Efficacy of Australian vaccines against recent isolates of avian infectious bronchitis viruses. Aust. Vet. J. 68 (6), 211–212.

Boni, M.F., Posada, D., Feldman, M.W., 2007. An exact nonparametric method for inferring mosaic structure in sequence triplets. Genetics 176 (2), 1035–1047.

Boots, A.H., Kusters, J., Zeijst, B.M., Hensen, E., 1991. The nucleocapsid protein of IBV comprises immunodominant determinants recognized by T-cells. In: Cavanagh, D., Brown, T.D. (Eds.), Coronaviruses and their Diseases, vol. 276. Springer, US, pp. 189–197.

Cavanagh, D., Davis, P.J., Darbyshire, J.H., Peters, R.W., 1986. Coronavirus IBV: virus retaining spike glycopolypeptide S2 but not S1 is unable to induce virus-neutralizing or haemagglutination-inhibiting antibody, or induce chicken tracheal protection. J. Gen. Virol. 67 (7), 1435–1442.

Chen, Y., Cai, H., a, Pan J., Xiang, N., Tien, P., Ahola, T., Guo, D., 2009. Functional screen reveals SARS coronavirus nonstructural protein nsp14 as a novel cap N7 methyltransferase. Proc. Natl. Acad. Sci. U. S. A. 106 (9), 3484–3489.

Chen, Y., Tao, J., Sun, Y., Wu, A., Su, C., Gao, G., Cai, H., Qiu, S., Wu, Y., Ahola, T., Guo, D., 2013. Structure-function analysis of severe acute respiratory syndrome coronavirus RNA cap guanine-N7-methyltransferase. J. Virol. 87 (11), 6296–6305.

Chousalkar, K.K., Cheetham, B.F., Roberts, J.R., 2009. Effects of infectious bronchitis virus vaccine on the oviduct of hens. Vaccine 27 (10), 1485–1489.

Cumming, R.B., 1962. The etiology of uraemia of chickens. Aust. Vet. J. 38 (11) (554-554).

Eckerle, L.D., Becker, M.M., Halpin, R.A., Li, K., Venter, E., Lu, X., Scherbakova, S., Graham, R.L., Baric, R.S., Stockwell, T.B., Spiro, D.J., Denison, M.R., 2010. Infidelity of SARS-CoV nsp14-exonuclease mutant virus replication is revealed by complete genome sequencing. PLoS Pathog. 6 (5), e1000896.

Gelman, A., Rubin, D.B., 1992. Inference from iterative simulation using multiple sequences. Stat. Sci. 7 (4), 457–472.

Gibbs, M.J., Armstrong, J.S., Gibbs, A.J., 2000. Sister-scanning: a Monte Carlo procedure for assessing signals in recombinant sequences. Bioinformatics 16 (7), 573–582.

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. Arch. Virol. 156 (2), 245–252.

Hewson, K.A., Noormohammadi, A.H., Devlin, J.M., Browning, G.F., Schultz, B.K., Ignjatovic, J., 2014. Evaluation of a novel strain of infectious bronchitis virus emerged as a result of spike gene recombination between two highly diverged parent strains. Avian Pathol. 43 (3), 249–257.

Hon, C.-C., Lam, T.-Y., Shi, Z.-L., Drummond, A.J., Yip, C.-W., Zeng, F., Lam, P.-Y., Leung, F.C.-C., 2008. Evidence of the recombinant origin of a bat severe acute respiratory syndrome (SARS)-like coronavirus and its implications on the direct ancestor of SARS coronavirus. J. Virol. 82 (4), 1819–1826.

Huson, D.H., Bryant, D., 2006. Application of phylogenetic networks in evolutionary studies. Mol. Biol. Evol. 23 (2), 254–267.

Ignjatovic, J., Galli, L., 1994. The S1 glycoprotein but not the N or M proteins of avian infectious bronchitis virus induces protection in vaccinated chickens. Arch. Virol. 138 (1–2), 117–134.

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 Pathol. 26 (3), 535–552.

Ignjatovic, J., Gould, G., Sapats, S., 2006. Isolation of a variant infectious bronchitis virus in Australia that further illustrates diversity among emerging strains. Arch. Virol. 151 (8), 1567–1585.

International Committee on Taxonomy of Viruses, 2015. Retrieved from http://www.ictvonline.org/virusTaxonomy.asp.

Jackwood, M.W., de Witt, J.J., 2013. Infectious bronchitis, In: Swayne, D.E., Glisson, J. R., McDougald, L.R., Nolan, L.K., Suarez, D.L., Nair, V.L. (Eds.), Diseases of Poultry. 13th ed. John Wiley & Sons, Inc., pp. 39–159.

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 (1), 98–108.

Jia, W., Karaca, K., Parrish, C.R., Naqi, S.A., 1995. A novel variant of avian infectious bronchitis virus resulting from recombination among three different strains. Arch. Virol. 140 (2), 259–271.

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 Res. 106 (1), 61–70.

Lole, K.S., Bollinger, R.C., Paranjape, R.S., Gadkari, D., Kulkarni, S.S., Novak, N.G., Ingersoll, R., Sheppard, H.W., Ray, S.C., 1999. Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. J. Virol. 73 (1), 152–160.

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 Pathol. 31 (6), 549–557.

Mardani, K., Noormohammadi, A.H., Hooper, P., Ignjatovic, J., Browning, G.F., 2008. Infectious bronchitis viruses with a novel genomic organization. J. Virol. 82 (4), 2013–2024.

Mardani, K., Noormohammadi, A.H., Ignjatovic, J., Browning, G., 2010. Naturally occurring recombination between distant strains of infectious bronchitis virus. Arch. Virol. 155 (10), 1581–1586.

Martin, D., Posada, D., Crandall, K., Williamson, C., 2005. A modified bootscan algorithm for automated identification of recombinant sequences and recombination breakpoints. AIDS Res. Hum. Retrovir. 21 (1), 98–102.

Martin, D.P., Lemey, P., Lott, M., Moulton, V., Posada, D., Lefeuvre, P., 2010. RDP3: a flexible and fast computer program for analyzing recombination. Bioinformatics 26 (19), 2462–2463.

Martin, D.P., Murrell, B., Golden, M., Khoosal, A., Muhire, B., 2015. RDP4: Detection and analysis of recombination patterns in virus genomes. Virus Evol. 1 (1), 1–5.

Minskaia, E., Hertzig, T., Gorbalenya, A.E., Campanacci, V., Cambillau, C., Canard, B., Ziebuhr, J., 2006. Discovery of an RNA virus 3'→5' exoribonuclease that is critically involved in coronavirus RNA synthesis. Proc. Natl. Acad. Sci. U. S. A. 103 (13), 5108–5113.

Ovcharenko, I., Loots, G.G., Giardine, B.M., Hou, M., Ma, J., Hardison, R.C., Stubbs, L., Miller, W., 2005. Mulan: multiple-sequence local alignment and visualization for studying function and evolution. Genome Res. 15 (1), 184–194.

Padidam, M., Sawyer, S., Fauquet, C.M., 1999. Possible emergence of new geminiviruses by frequent recombination. Virology 265 (2), 218–225.

Posada, D., Crandall, K.A., 2001. Evaluation of methods for detecting recombination from DNA sequences: computer simulations. Proc. Natl. Acad. Sci. U. S. A. 98 (24), 13757–13762.

Quinteros, J.A., Markham, P.F., Lee, S.-W., Hewson, K.A., Hartley, C.A., Legione, A.R., Coppo, M.J.C., Vaz, P.K., Browning, G.F., 2015. Analysis of the complete genomic sequences of two virus subpopulations of the Australian infectious bronchitis virus vaccine VicS. Avian Pathol. 44 (3), 182–191.

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., Huelsenbeck, J.P., 2012. MrBayes 3.2: efficient bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. 61 (3), 539–542.

Saif, L.J., 1993. Coronavirus immunogens. Vet. Microbiol. 37 (3–4), 285–297.

Sapats, S.I., Ashton, F., Wright, P.J., Ignjatovic, J., 1996. Sequence analysis of the S1 glycoprotein of infectious bronchitis viruses: identification of a novel genotypic group in Australia. J. Gen. Virol. 77, 413–418.

Seo, S.H., Wang, L., Smith, R., Collisson, E.W., 1997. The carboxyl-terminal 120-residue polypeptide of infectious bronchitis virus nucleocapsid induces cytotoxic T lymphocytes and protects chickens from acute infection. J. Virol. 71 (10), 7889–7894.

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J.D., Higgins, D.G., 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol. Syst. Biol. 7 (1) (539).

Smith, E.C., Denison, M.R., 2012. Implications of altered replication fidelity on the evolution and pathogenesis of coronaviruses. Curr. Opin. Virol. 2 (5), 519–524.

Smith, J.M., 1992. Analyzing the mosaic structure of genes. J. Mol. Evol. 34 (2), 126–129.

Stroehlein, A.J., Young, N.D., Jex, A.R., Sternberg, P.W., Tan, P., Boag, P.R., Hofmann, A., Gasser, R.B., 2015. Defining the Schistosoma haematobium kinome enables the prediction of essential kinases as anti-schistosome drug targets. Sci. Rep. 5, 17759.

Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22 (22), 4673–4680.

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 (9), 1777–1799.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M., Rozen, S. G., 2012. Primer3—new capabilities and interfaces. Nucl. Acids Res. 40 (15) (e115-e115).

Wadey, C.N., Faragher, J.T., 1981. Australian infectious bronchitis viruses: identification of nine subtypes by a neutralisation test. Res. Vet. Sci. 30, 70–74.