Prediction of Bovine coronavirus immune epitopes using selection and phylogenetic analysis

Abstract

Bovine betacoronavirus (BoCoV) is a pneumoenteric pathogen of cattle, which is closely related to the common cold Human coronavirus OC43. Vaccines are administered to protect against diseases caused by BoCoV, but knowledge gaps exist with regards correlates of protection and the effect of immune evasion on driving evolution. Immune epitopes predicted for each structural protein were corroborated with selective pressure analysis and sequencing of genes encoding two surface proteins, spike and haemagglutinin esterase (HE). Eleven of the sixteen predicted spike immune epitopes and three of the six predicted HE immune epitopes overlapped residues under diversifying selection or residues where amino acid changes were found. Four of the nineteen residues predicted to be important for recognition of spike by neutralising antibodies (nAb) were under diversifying selection or amino acid changes were observed. Increased prevalence of amino acid changes in specific immune epitopes suggests an effective immune response recognises these regions as non-synonymous mutations would provide a selective advantage for the virus due to immune escape. Selection analysis and sequencing provided more support for MHC-II, B cell receptor and nAb epitopes and less support for MHC-I epitopes. Downregulation of MHC-I antigen presentation has been demonstrated for several coronaviruses and the reduced support for predicted MHC-I epitopes, compared to other immune epitopes, suggests this could also be true for BoCoV.

Introduction

Bovine betacoronavirus (BoCoV) is a pneumoenteric pathogen of *Bos taurus* though closely related viruses have been isolated from several other ruminant species and some non-ruminant mammals. In cattle it causes winter dysentery, calf diarrhoea [1] and is associated with the bovine respiratory disease complex, as well as being isolated from asymptomatic cattle [2]. There are full-genome sequences of 24 BoCoV isolated from cattle in 2019/20 from across Ireland [3] and ten S1 domain and one full-length spike sequences from isolates collected in the province of Munster in 2010/11 [4]. As a member of the *Embecovirus* subgenus, BoCoV encodes five structural proteins including four conserved across the *Coronaviridae* family – spike, nucleocapsid, membrane and small envelope – as well as a surface protein unique to this *Embecovirus* subgenus – haemagglutinin esterase (HE). HE is a receptor destroying enzyme, which facilitates virus entry and release by targeting the 9-O-

acetylated sialic acid residues required for BoCoV entry [5]. Spike and HE proteins stimulate the humoral immune response, and it has been shown strong antibody responses targeting these proteins can protect cattle from disease following BoCoV infection [6–8], although other studies have shown BoCoV antibody titres do not correlate with protection [9,10]. There remains several gaps or areas of uncertainty regarding knowledge of BoCoV immune epitopes; correlates of protection; and the effect of immune evasion on BoCoV evolution [7]. To identify viral proteins recognised by the immune system and assess the contribution of immune evasion to variation, predicted immune epitopes of BoCoV structural proteins were corroborated with selective pressure analysis. Phylogenetic analysis of spike and HE sequences obtained from clinical isolates collected in Ireland and the BoCoV Mebus strain used to design a live attenuated vaccine identified residues where amino acid changes have occurred compared to the Mebus strain and where amino acid changes have arisen among isolates sequenced in this study. Overlap between predicted immune epitopes and residues containing amino acid changes or residues under diversifying selection in spike and HE suggests some of the predicted immune epitopes could be recognised during clinical infections and drive variation in these proteins.

Materials and Methods

Samples and Detection of Bovine coronavirus

Nasopharyngeal samples collected from cattle displaying symptoms of respiratory illness during 2022 in Ireland were submitted to the Department of Agriculture, Food and Marine's Backweston Laboratories. Samples were tested for the presence of respiratory pathogens by qPCR. Those positive for BoCoV were identified using a previously developed TaqMan qPCR protocol [11].

Recombination and Selection analysis

BoCoV spike, HE, nucleocapsid, membrane and small envelope coding sequences were downloaded from the NCBI database. Identical and incomplete sequences were removed from datasets. Codon alignments for each gene were obtained using the Clustal Omega alignment algorithm [12,13] in Jalview [14–16]. Stop codons were deleted manually. Codon alignments, plus sequences obtained in this study, were aligned and recombination breakpoints predicted using Genetic Algorithm for Recombination Detection [17] through the data monkey webserver (datamonkey.org) [18]. Sequences arising through recombination were removed from alignments, then mixed effects model of evolution (MEME) [19] and fast unconstrained bayesian approximation (FUBAR) [20] methods were used through the data monkey web server [18] to infer site-specific selective pressures. MEME predicted sites under

diversifying episodic selection and FUBAR estimated synonymous and non-synonymous mutation rates to infer pervasive selective pressure.

Prediction of immune epitopes

Immune epitopes for BoCoV structural proteins were predicted based on three previous studies. Firstly, BoCoV spike and HE MHC-I and B cell receptor (BCR) immune epitopes were predicted bioinformatically [21]. Secondly, MHC-II immune epitopes identified for *Human betacoronavirus OC43* (HCoV OC43) by co-immunoprecipitation (co-IP) of MHC-II molecules [22] were mapped onto BoCoV spike, HE, nucleocapsid and envelope using alignments of the orthologous protein amino acid sequences. Thirdly, HCoV OC43 spike residues shown to interact with neutralising antibodies (nAb) through structural analysis or shown to be escape mutants following passage in the presence of antibodies [23] were mapped onto BoCoV spike using alignments of orthologous spike proteins.

Structures

BoCoV spike structure was predicted using SWISS-MODEL (https://swissmodel.expasy.org/) with the BoCoV Mebus spike amino acid sequence as input [24]. BoCoV HE structure had been determined by x-ray crystallography [25] and deposited on the protein data bank (https://www.rcsb.org/). These spike and HE structures were visualised and annotated in PyMOL [26].

Sequencing spike and haemagglutinin esterase

RNA from samples positive for BoCoV was converted to cDNA using the SuperScript III Reverse Transcriptase kit (Invitrogen) following the manufacturer's protocol. cDNA was used as a template for PCR (Phusion HF kit (New England Biolabs)) with primers designed to amplify the spike and HE genes (Table 1). PCR products from reactions producing a specific fragment underwent PCR clean-up (Qiagen) and, if multiple fragments were detected, gel purification (Qiagen) from a 0.8% agarose gel was carried out. Purified fragments were sent for Sanger sequencing using appropriate forward and reverse primers (Eurofins Genomics).

Gene	Start	Finish	Primer sequences (5' – 3')		Annealing (°C)
			Forward	Reverse	
HE	-365	547	CCCTCATCACCGGCTAGACT	CCCCAAAATTAGCTTCACGAGC	65
HE	377	1353	AGGCTTGTTTTACACTCAGGT	GTACACTTTAAATCTCCTATAACAGC	61
Spike	-107	1039	CTTGGCATTCTTTTGGGTGTTGC	TAATGGAGAGGGCACCGACTT	66
Spike	782	1550	GGGTTACACCTCTCACTTCT	GCAGGACAAGTGCCTATACC	60
Spike	1460	2286	GTCCGTGTAAATTGGATGGG	TGTAGAGTAATCCACACAGT	60

Spike	1855	2731	TTACAAAAATCAAACACAGACAT	AAACTTTATTACAATCGCTTCC	60
Spike	2680	3617	TCAATTTTTCCCCTGTATTAGG	GTAGTAATAACCACTACCAGTG	59
Spike	3475	4218	TTTAGCTATGTCCCTACTAAGTA	TGTGGTAGCTATTATAATATGCTCG	60

Table 1: Primer pairs used to amplify regions of HE and spike genes. Start and finish positions of the amplicons are provided relative to the start codons of the HE or spike genes. Primers were designed for this study or taken from [27,28].

Phylogenetic analysis

Reads obtained from Sanger sequencing were trimmed to remove low quality base calls. Reads from the same sample were aligned to BoCoV spike and HE references using the Clustal Omega alignment algorithm [12,13] in Jalview [14–16]. References were selected based on the closest NCBI Blast hits for individual sequences from the same sample. Sequences were deposited on Genbank (Table 2). Nucleic acid sequences were converted to amino acid sequences then phylogenetic trees were generated using the BLOSUM62 scoring matrix [29] and neighbour-joining method [30] in Jalview [14–16]. Trees were annotated and images generated in the Interactive Tree of Life web server [31].

Accession	Location (county)	Date	Pathology	Host
OR271239	Galway	30/03/2022	ND	NA
OR271240	Waterford	11/05/2022	ND	NA
OR271241	Sligo	27/04/2022	Dead, severe bronchopneumonia with multi-focal necrosis and fibrinous pleurisy, co-infected with *Histophilus somni* and *Mannheimia haemolytica*	3 week calf
OR271242	Laois	08/04/2022	ND	2 month calf
OR271243	Laois	08/04/2022	ND	2 month calf
OR271244	Wexford	22/03/2022	Severe diarrhoea, co-infected with *Escherichia coli*	5 week calf
OR271245	Wexford	21/03/2022	Severe diarrhoea, scour and respiratory distress, co-infected with *E. coli*	10 day calf
OR271246	Kilkenny	16/03/2022	ND	Calves (pooled)
OR271247	Kilkenny	14/03/2022	Pneumonia and enteritis, Bovine herpesvirus 1 detected at high CT	4 week calf
OR271248	Cavan	23/02/2022	Co-infected with *Pasteurella multocida* and *M. haemolytica*	4 animals (pooled)
OR271249	Donegal	15/02/2022	Pneumonia	10 month weanling
OR271250	Wexford	11/03/2022	Severe pneumonia, other bacterial respiratory pathogens isolated	2 week calf
OR271251	Wexford	11/03/2022	Severe pneumonia, other bacterial respiratory pathogens isolated	2 week calf
OR271252	Sligo	06/01/2022	ND	Weanling
OR271253	Laois	08/04/2022	ND	< 1 month calf
OR271254	Wicklow	23/05/2022	Severe bronchopneumonia, co-infected with Bovine herpesvirus 4, *H. somni* and *M. haemolytica*	3 month calf
OR271255	Westmeath	2022	Dead	1 year male

OR271256	Offaly	2022	Pneumonia	8 month weanling
OR271257	Meath	2022	Pneumonia	6 week calf

Table 2: Genbank accession numbers and sample details. Abbreviations: ND, no disease information; NA, no animal information.

Results

Selection Analysis of BoCoV structural proteins

Sequences resulting from recombination were observed for the spike protein, but not other BoCoV proteins. These sequences were removed prior to selection analysis using FUBAR and MEME algorithms (Figure 1). Spike (Figure 1A) and HE (Figure 1B) were under increased diversifying selection compared to nucleocapsid (Figure 1C), membrane (Figure 1D) and small envelope (Figure 1E) proteins as demonstrated by the increased proportion of residues predicted to be under diversifying selection for spike and HE compared to other structural proteins (Table 3).

Diversifying selection within predicted BCR, MHC-I and MHC-II immune epitopes (Table 4) or residues predicted to be important in nAb recognition (Table 5) was assessed. Spike residues under diversifying selection only overlapped with one spike MHC-I immune epitope and one spike BCR immune epitope predicted computationally (Figure 2A). Five of seven spike MHC-II immune epitopes predicted by co-IP of MHC-II overlapped residues under diversifying selection (Figure 2A). A higher than expected proportion of spike residues predicted to be involved in neutralising antibody recognition were under diversifying selection (Figure 2A). 12.5% of predicted neutralising antibody immune epitopes were under diversifying selection compared to 5.3% of other residues. There was no overlap between predicted HE MHC-I and MHC-II immune epitopes (Figure 2D). Two predicted MHC-II nucleocapsid immune epitopes (Figure 1C) and one predicted MHC-II envelope immune epitope (Figure 1E) did not overlap with any residues under diversifying selection.

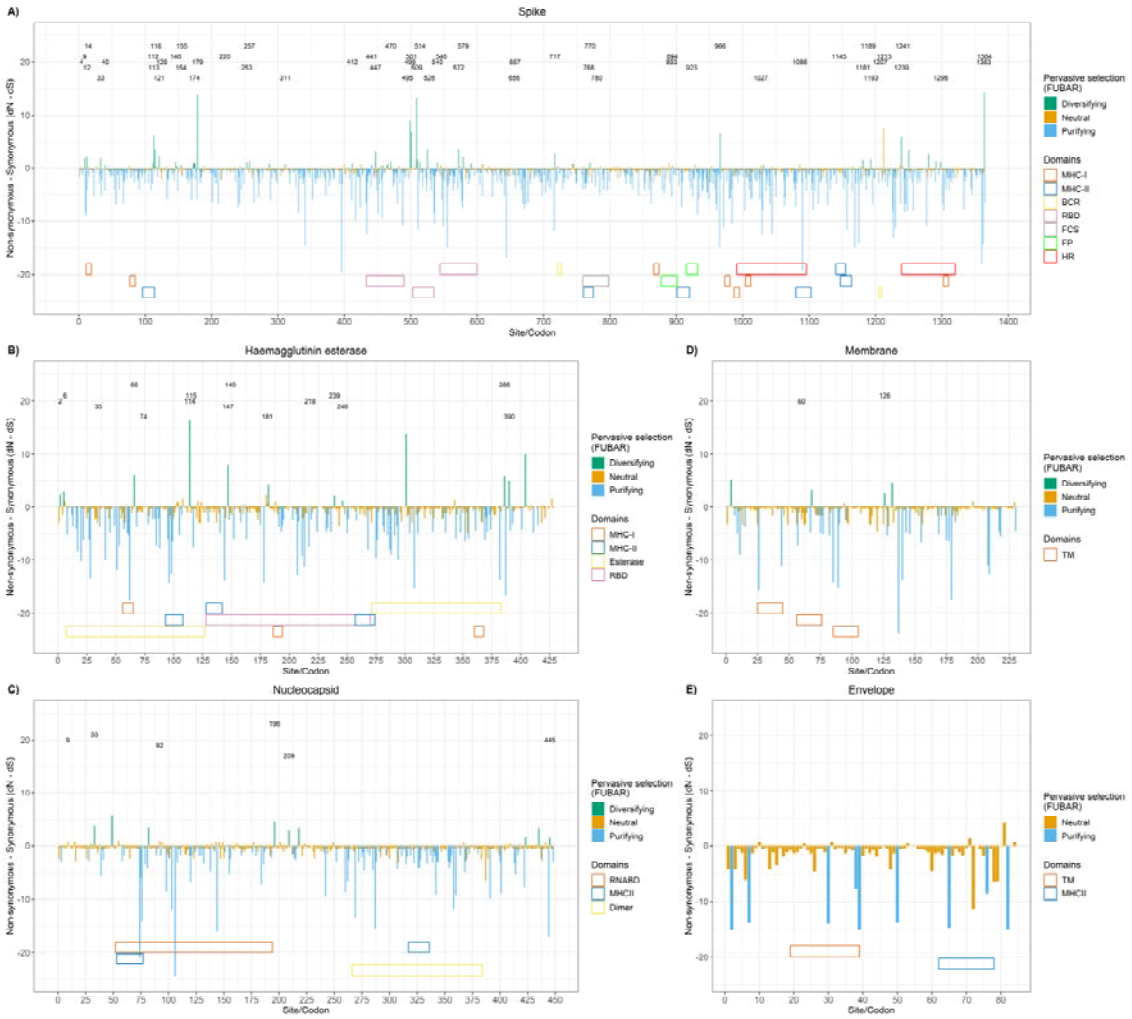

Figure 1: Selection analysis of structural BoCoV proteins. Bar plots show the results of selection analysis with FUBAR and annotated residues were predicted to be under diversifying selection using MEME for spike (A), HE (B), nucleocapsid (C), membrane (D) and small envelope (E) proteins. Locations of structural domains, functional domains and immune epitopes were annotated at the bottom of plots. Abbreviations: BCR, B-cell receptor; Dimer, interaction interface for dimerisation; FCS, Furin Cleavage Site; FP, fusion peptide; HR, heptad repeat; MHC-I, predicted MHC-I epitope; MHC-II, predicted MHC-II epitope; RBD, receptor binding domain; RNABD, RNA binding domain; and TM, transmembrane domain.

Protein	Number of residues under selection						
	FUBAR (%)		MEME (%)	FUBAR + MEME	FUBAR only	MEME only	Diversifying total (%)
	Diversifying	Purifying					
Spike	49 (3.59)	570 (41.76)	57 (4.18)	33	16	24	73 (5.35)
HE	12 (2.80)	97 (22.66)	15 (3.50)	8	4	7	19 (4.44)
NCP	9 (2.00)	100 (22.27)	6 (1.34)	4	5	2	11 (2.45)
M	4 (1.74)	36 (15.65)	2 (0.87)	1	3	1	5 (2.17)
E	0 (0)	9 (10.71)	0 (0)	0	0	0	0 (0)

Table 3: Selection analysis of structural proteins. Number and proportion of residues found to be under selective pressure for each BoCoV structural protein as determined by FUBAR and MEME.

Gene	Immune epitope	Location		Diversifying selection	Amino acid changes		Ref
		Start	Finish		Compared to Mebus	Among Irish sequences	
Spike	MHC-I	11	19	2	1	0	19
		77	85	0	0	1	
		866	874	0	0	0	
		973	981	0	0	0	
		987	995	0	1	0	
		1004	1012	0	0	0	
		1302	1310	0	0	0	
	MHC-II	96	114	2	2	1	20
		267	292	1	0	1	
		760	775	2	0	1	
		901	920	0	2	1	
		1080	1103	1	1	0	
		1140	1155	1	1	1	
		1147	1164	0	0	0	
	BCR	721	727	0	0	1	19
		1205	1209	1	0	0	
HE	MHC-I	56	65	0	0	0	19
		186	194	0	0	0	
		360	368	0	2	2	
	MHC-II	93	108	0	0	1	20
		128	142	0	0	1	
		257	274	0	0	0	

Table 4: Immune epitopes and overlap with residues under diversifying selection and where amino acid changes were observed.

Location	Selection	Amino acid changes	
		Compared to Mebus	Among Irish Sequences
400	Neutral	No change	No change
401	Neutral	No change	No change
416	Neutral	No change	No change
461	Neutral	No change	No change
548	Purifying	No change	No change
553	Neutral	No change	No change
581	Purifying	No change	No change
582	Neutral	No change	No change
33	Diversifying	A33V	No change
34	Purifying	No change	No change
35	Purifying	No change	S35L
243	Neutral	No change	No change
256	Purifying	M256L	No change
267	Neutral	No change	No change
459	Purifying	No change	No change
483	Neutral	No change	No change
497	Purifying	No change	No change
501	Diversifying	P501S	S501F
503	Neutral	No change	No change

Table 5: Residues predicted to be important to nAb recognition and overlap with selection or amino acid changes.

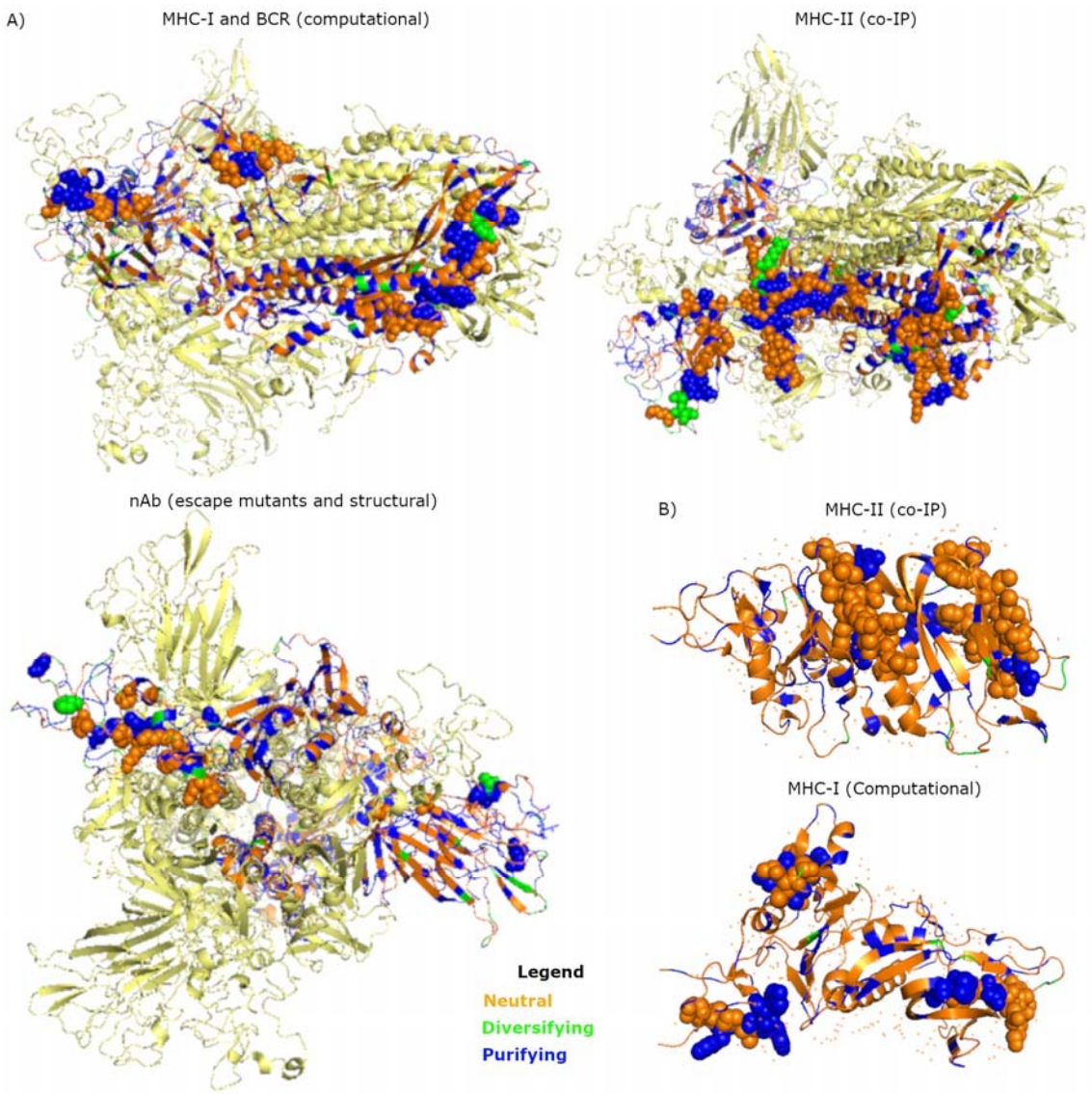

Figure 2: Spike trimer and HE structures with residues coloured based on results of FUBAR and MEME. Only one spike monomer was coloured. A) Spike immune epitopes predicted computationally, by co-IP of HCoV OC43 and by structural analysis of HCoV OC43 nAb-protein interactions were shown as spheres. B) Predicted HE MHC-II and MHC-I immune epitopes shown as spheres. Abbreviations: BCR, B cell receptor; co-IP, co-immunoprecipitation; MHC-I, major histocompatibility complex I; MHC-II, major histocompatibility complex II; and nAb, neutralising antibody.

There was increased prevalence of purifying selection observed for residues within functional domains such as the fusion peptides of spike (Figure 1A), the esterase domains of HE (Figure 1B), the

RNA binding and dimerisation domains of nucleocapsid (Figure 1C) and the transmembrane domains of membrane (Figure 1D) and envelope (Figure 1E). Selection at the spike and HE receptor binding domains (RBD) was of interest as nAb often recognise epitopes in the RBD, but they are also involved in protein-protein interactions. There was an increased prevalence of diversifying (9.5% residues within RBD under diversifying selection compared to 4.8% of residues outside RBD) and purifying (48.0% within versus 41.0% outside) selection for the RBD of spike (Figure 3A). For HE RBD there was increased prevalence of diversifying selection (4.9% within versus 4.2% outside) but decreased purifying selection (19.6% within versus 24.2% outside) (Figure 3B).

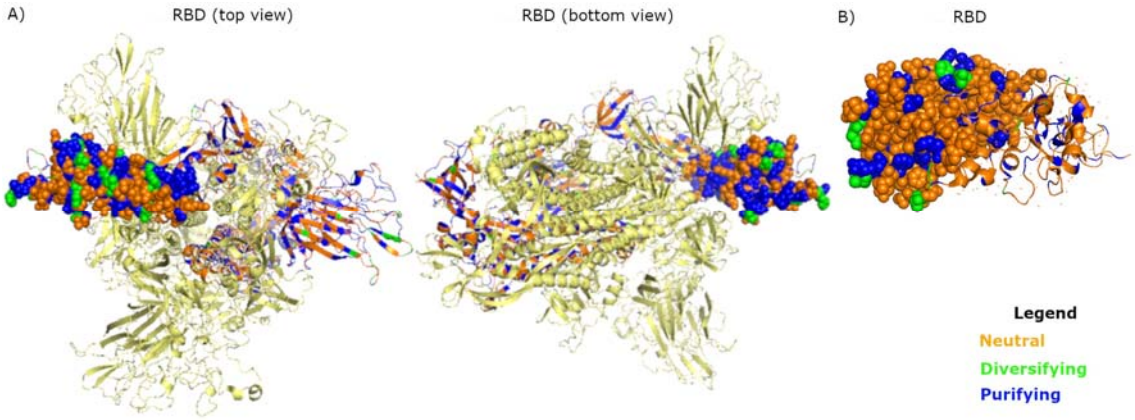

Figure 3: RBD of spike (A) and HE (B) shown as spheres with residues coloured based on selection analysis.

Sequencing and phylogenetic analysis of BoCoV spike and HE

The spike and HE genes of BoCoV clinical isolates collected from cattle presenting with respiratory and sometimes enteric symptoms in Ireland were sequenced. Spike nucleotide sequences from this study fell into one of two clades (Figure 4A), while most HE sequences were split between two subclades with Sample 117 the only exception (Figure 4B). As has previously been shown, spike amino acid sequences diverged based on location with one clade of Asian and American isolates and a second clade of European isolates containing the Irish sequences from this study (Figure 4C). Some of the spike sequences from this study were in a subclade containing other Irish isolates collected in 2019/20, whilst others were in a subclade of French isolates from 2013/14 (Figure 4C). An enteric Irish isolate from 2011 diverged from other sequenced BoCoV isolates in its own subclade (Figure 4C). Most HE sequences from this study were closest to Irish isolates collected in 2019/20, but still formed their own subclade, while Sample 117 was closest to enteric isolates from China collected in 2018 and 2020 (Figure 4D). Differences in the spike and HE phylogenetic trees suggests recombination between BoCoV circulating in Ireland, which was supported by the detection of four

recombination breakpoints at positions 1594, 2344, 4552 and 4979 of the HE and spike coding region alignment, which all overlapped with spike. Three breakpoints were also predicted within the HE and spike alignment of sequences downloaded from NCBI at positions 1563, 3424 and 4756.

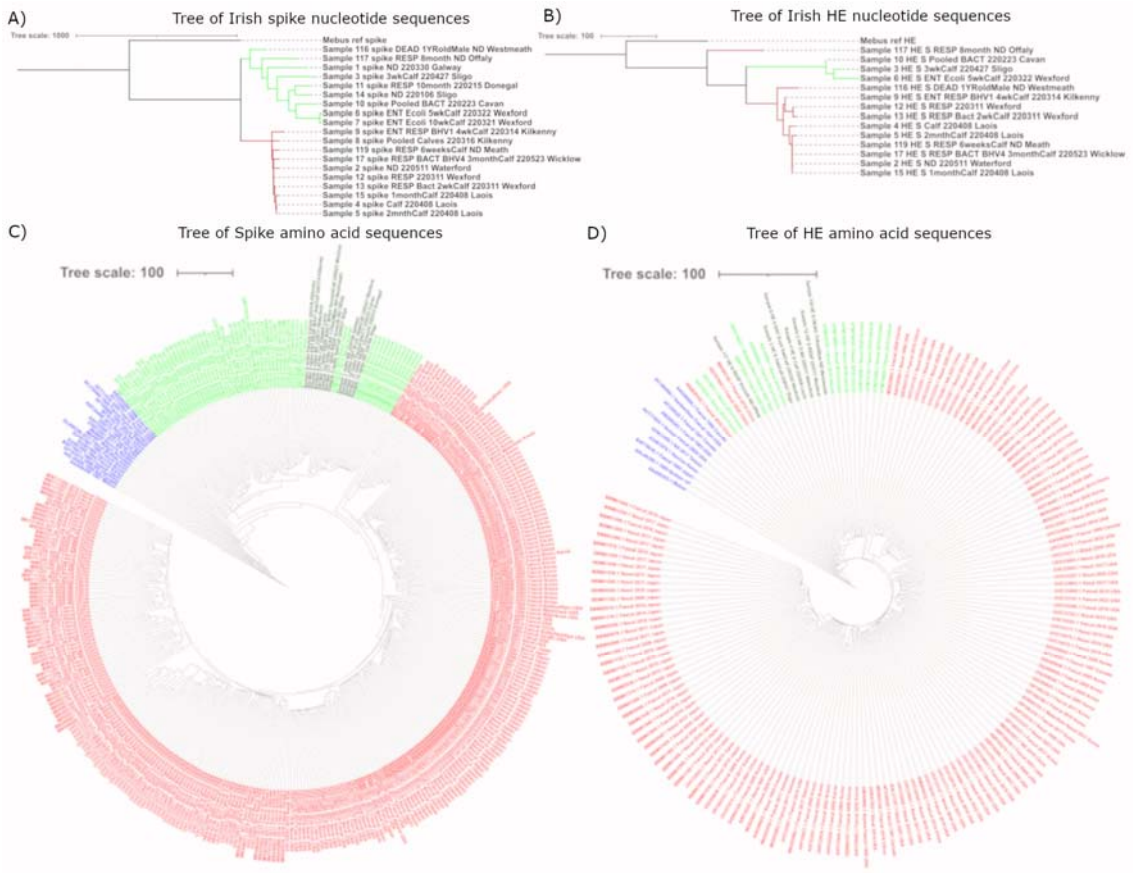

Figure 4: Phylogenetic trees of BoCoV spike and HE. A) Nucleotide sequences of spike for isolates from this study and the Mebus strain with different coloured branches indicating two clades. B) Same as A), but for HE. C) All unique spike amino acid sequences downloaded from genbank, with labels colours indicating location of isolation with Asian/American in red, European in green and old isolates/vaccine strains in blue. Sequences from this study are shown in darker green. D) The same as C), but for HE.

There were amino acid changes in 54 (4.0%) spike residues when comparing Mebus spike with the sequences from this study. Among the spike sequences obtained in this study there were amino acid changes in 42 (3.1%) residues. For HE, there were amino acid changes in 10 (2.4%) residues compared to Mebus HE and among the sequences from this study amino acid changes in 17 (4.0%) residues. Residues containing amino acid changes were grouped based on those different from Mebus (Mebus mutations) and those where variation was observed when comparing sequences from this study (Ireland mutations). Some amino acid changes overlapped with predicted immune

epitopes (Table 4) or occurred at residues potentially involved in nAb recognition (Table 5). Mebus mutations overlapped with two MHC-I epitopes, four MHC-II epitopes and three residues predicted to be important to nAb recognition in spike (Figure 5A). Ireland mutations overlapped with one BCR, one MHC-I and five MHC-II epitopes plus two residues predicted to be important to nAb recognition in spike (Figure 5A). One predicted HE MHC-I epitopes overlapped with Mebus and Ireland mutations while two predicted MHC-II epitopes overlapped with Ireland mutations only (Figure 5B).

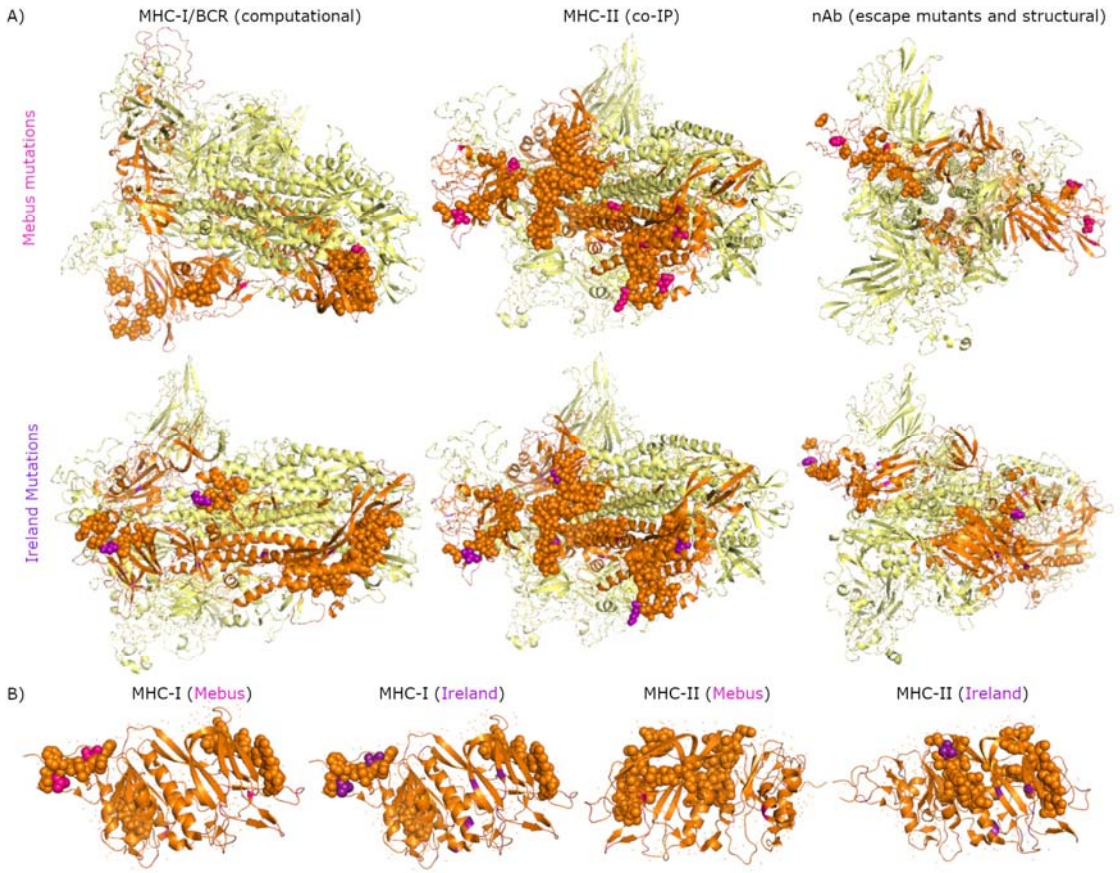

Figure 5: Overlap between residues with amino acid changes and predicted immune epitopes. A) The spike trimer is presented as a cartoon, with predicted immune epitopes presented as spheres on one spike monomer in orange. Mebus mutations are shown in pink and Ireland mutations in purple. B) Same as A) but for HE.

Discussion

Pressure to evade the immune response is one driver of diversifying selection in *Coronaviridae* [32,33], although there are other drivers of diversifying selection, such as adaptation to a new host [34]. There was increased diversifying selection for spike and HE compared to other structural proteins suggesting these could contain immune epitopes. Seven of sixteen predicted immune epitopes and

two of the nineteen residues potentially important to nAb recognition overlapped residues under diversifying selection.

As expected, there was overlap between residues under purifying selection and several structural and functional domains along BoCoV proteins. The high level of selection observed for spike RBS is expected because its interaction with the entry receptor would drive purifying selection [32,35], while RBD often contain nAb epitopes, which would drive diversifying selection [36,37]. Several residues within the HE RBD were also under selection with an increased prevalence of diversifying selection compared to the rest of the protein. The high level of purifying selection in the esterase domains, which do not overlap the RBD, means there was not an increased prevalence of purifying selection for HE RBD.

Embecoviruses, including BoCoV and HCoV OC43, might bind MHC-I and heparan sulfate as well as 9-O-acetylated sialic acid residues to facilitate cell entry [38,39] suggesting there could be a second, and potentially third, RBD on spike. The increased diversifying selection observed between residues 100-200 and predicted exposure of this region on the surface of the virion suggest it could represent another RBD and potentially another region containing immune epitopes though further investigation is required to determine this.

Phylogenetic analysis of spike and HE genes of BoCoV isolates collected from Ireland in 2022 and previously sequenced BoCoV isolates corroborated previous studies demonstrating divergence of isolates from America or Asia and Europe, although HE sequences of some Asian isolates were in the European clade, which could have been caused by recombination. Although this would require a co-infection of European or Asian cattle with BoCoV originating from the two regions. Furthermore, differences between the spike and HE phylogenetic trees for isolates sequenced in this study could indicate recombination occurring between BoCoV circulating in Ireland. Both potential recombination events predicted from phylogenetic tree structures were supported by GARD, which predicted several recombination breakpoints including at the 5' end of spike.

Spike sequences from this study were in one of two clades suggesting isolates descending from two distinct introductions of the virus into Ireland have been sequenced or diverged along two separate branches following a single introduction. There were multiple residues with amino acid changes when comparing the clinical isolates sequenced in this study with the Mebus strain as well as among the sequences from this study. These overlapped with ten of the sixteen predicted immune epitopes and four of the nineteen residues predicted to be important to nAb recognition. Amino acid changes in immune epitopes could facilitate immune escape and provide the mutated virus with a selective advantage.

There was more overlap with diversifying selection and amino acid changes for predicted MHC-II and BCR immune epitopes as well as residues potentially involved in nAb recognition when compared with predicted MHC-I immune epitopes. The reduced support for MHC-I epitopes by selection analysis and analysis of sequence data could be due to downregulation of the MHC-I presentation by BoCoV. HCoV OC43 [22] and other coronaviruses have been shown to suppress MHC-I presentation [40,41] suggesting BoCoV could also downregulate MHC-I. This study has predicted several BoCoV immune epitopes and corroborated several of these in spike and HE with selection analysis and sequence data.

References

1. Gomez, D. E. & Weese, J. S. Viral enteritis in calves. *Can. Vet. J.* **58**, 1267 (2017).

2. Fulton, R. W. Viruses in Bovine Respiratory Disease in North America: Knowledge Advances Using Genomic Testing. *Vet. Clin. North Am. Food Anim. Pract.* **36**, 321 (2020).

3. Dhufaigh, K. N. *et al.* Genome Sequence of Bovine Coronavirus Variants from the Nasal Virome of Irish Beef Suckler and Pre-weaned Dairy Calves Clinically Diagnosed with Bovine Respiratory Disease. *Microbiol. Resour. Announc.* **11**, (2022).

4. Gunn, L., Collins, P. J., O'Connell, M. J. & O'Shea, H. Phylogenetic investigation of enteric bovine coronavirus in Ireland reveals partitioning between European and global strains. *Ir. Vet. J.* **68**, (2015).

5. Vlasak,', R., Luytjes, W., Leider,', J., Spaan, W. & Palese1, P. The E3 protein of bovine coronavirus is a receptor-destroying enzyme with acetylesterase activity. *J. Virol.* **62**, 4686–4690 (1988).

6. Lin, X., O'Reilly, K. L., Burrell, M. L. & Storz, J. Infectivity-neutralizing and hemagglutinin-inhibiting antibody responses to respiratory coronavirus infections of cattle in pathogenesis of shipping fever pneumonia. *Clin. Diagn. Lab. Immunol.* **8**, 357–362 (2001).

7. Vlasova, A. N. & Saif, L. J. Bovine Coronavirus and the Associated Diseases. *Front. Vet. Sci.* **8**, 293 (2021).

8. Milane, G., Kourtesis, A. B. & Dea, S. Characterization of monoclonal antibodies to the hemagglutinin-esterase glycoprotein of a bovine coronavirus associated with winter dysentery and cross-reactivity to field isolates. *J. Clin. Microbiol.* **35**, 33–40 (1997).

9. Workman, A. M. *et al.* Evaluation of the effect of serum antibody abundance against bovine coronavirus on bovine coronavirus shedding and risk of respiratory tract disease in beef calves from birth through the first five weeks in a feedlot. *Am. J. Vet. Res.* **78**, 1065–1076 (2017).

10. Workman, A. M., Kuehn, L. A., McDaneld, T. G., Clawson, M. L. & Loy, J. D. Longitudinal study of humoral immunity to bovine coronavirus, virus shedding, and treatment for bovine respiratory disease in pre-weaned beef calves. *BMC Vet. Res.* **15**, 1–15 (2019).

11. Decaro, N. *et al.* Detection of bovine coronavirus using a TaqMan-based real-time RT-PCR assay. *J. Virol. Methods* **151**, 167 (2008).

12. Sievers, F. & Higgins, D. G. Clustal Omega for making accurate alignments of many protein

sequences. *Protein Sci.* **27**, 135–145 (2018).

13. Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539 (2011).

14. Procter, J. B. *et al.* Alignment of Biological Sequences with Jalview. *Methods Mol. Biol.* **2231**, 203–224 (2021).

15. Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M. & Barton, G. J. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189–1191 (2009).

16. Troshin, P. V., Procter, J. B. & Barton, G. J. Java bioinformatics analysis web services for multiple sequence alignment—JABAWS:MSA. *Bioinformatics* **27**, 2001–2002 (2011).

17. Kosakovsky Pond, S. L., Posada, D., Gravenor, M. B., Woelk, C. H. & Frost, S. D. W. GARD: a genetic algorithm for recombination detection. *Bioinformatics* **22**, 3096–3098 (2006).

18. Weaver, S. *et al.* Datamonkey 2.0: A Modern Web Application for Characterizing Selective and Other Evolutionary Processes. *Mol. Biol. Evol.* **35**, 773–777 (2018).

19. Murrell, B. *et al.* Detecting Individual Sites Subject to Episodic Diversifying Selection. *PLOS Genet.* **8**, e1002764 (2012).

20. Murrell, B. *et al.* FUBAR: A Fast, Unconstrained Bayesian AppRoximation for Inferring Selection. *Mol. Biol. Evol.* **30**, 1196–1205 (2013).

21. Awadelkareem, E. A. & Hamdoun, S. Insilco Vaccine Design of spike and hemagglutinin esterase proteins of Bovine Coronavirus. (2022) doi:10.21203/RS.3.RS-1791707/V1.

22. Becerra-Artiles, A. *et al.* Immunopeptidome profiling of human coronavirus OC43-infected cells identifies CD4 T cell epitopes specific to seasonal coronaviruses or cross-reactive with SARS-CoV-2. *bioRxiv* (2022) doi:10.1101/2022.12.01.518643.

23. Wang, C. *et al.* Antigenic structure of the human coronavirus OC43 spike reveals exposed and occluded neutralizing epitopes. *Nat. Commun.* **13**, (2022).

24. Waterhouse, A. *et al.* SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* **46**, W296 (2018).

25. Zeng, Q., Langereis, M. A., Van Vliet, A. L. W., Huizinga, E. G. & De Groot, R. J. Structure of coronavirus hemagglutinin-esterase offers insight into corona and influenza virus evolution. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 9065 (2008).

26. Schrödinger, L., & DeLano, W. PyMOL. (2020).

27. Hasoksuz, M., Sreevatsan, S., Cho, K. O., Hoet, A. E. & Saif, L. J. Molecular analysis of the S1 subunit of the spike glycoprotein of respiratory and enteric bovine coronavirus isolates. *Virus Res.* **84**, 101 (2002).

28. Martínez, N. *et al.* Molecular and phylogenetic analysis of bovine coronavirus based on the spike glycoprotein gene. *Infect. Genet. Evol.* **12**, 1870 (2012).

29. Henikoff, S. & Henikoff, J. G. Amino acid substitution matrices from protein blocks. *Proc. Natl. Acad. Sci.* **89**, 10915–10919 (1992).

30. Saitou, N. & Nei, M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**, 406–425 (1987).

31. Letunic, I. & Bork, P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* **49**, W293–W296 (2021).

32. Franzo, G. *et al.* Bovine Coronavirus: Variability, Evolution, and Dispersal Patterns of a No Longer Neglected Betacoronavirus. *Viruses 2020, Vol. 12, Page 1285* **12**, 1285 (2020).

33. Kistler, K. E. & Bedford, T. Evidence for adaptive evolution in the receptor-binding domain of seasonal coronaviruses OC43 and 229E. *Elife* **10**, 1–35 (2021).

34. Zhang, C. Y., Wei, J. F. & He, S. N. Adaptive evolution of the spike gene of SARS coronavirus: Changes in positively selected sites in different epidemic groups. *BMC Microbiol.* **6**, 1–10 (2006).

35. Li, X. *et al.* Emergence of SARS-CoV-2 through recombination and strong purifying selection. *Sci. Adv.* **6**, (2020).

36. Dumonteil, E. & Herrera, C. Polymorphism and Selection Pressure of SARS-CoV-2 Vaccine and Diagnostic Antigens: Implications for Immune Evasion and Serologic Diagnostic Performance. *Pathog. 2020, Vol. 9, Page 584* **9**, 584 (2020).

37. Armijos-Jaramillo, V., Yeager, J., Muslin, C. & Perez-Castillo, Y. SARS-CoV-2, an evolutionary perspective of interaction with human ACE2 reveals undiscovered amino acids necessary for complex stability. *Evol. Appl.* **13**, 2168–2178 (2020).

38. Szczepanski, A. *et al.* Canine Respiratory Coronavirus, Bovine Coronavirus, and Human Coronavirus OC43: Receptors and Attachment Factors. *Viruses 2019, Vol. 11, Page 328* **11**, 328 (2019).

39. Collins, A. R. HLA Class I Antigen Serves as a Receptor for Human Coronavirus OC43. *http://dx.doi.org/10.3109/08820139309063393* **22**, 95–103 (2009).

40. Zhang, F. *et al.* Inhibition of major histocompatibility complex-I antigen presentation by sarbecovirus ORF7a proteins. *Proc. Natl. Acad. Sci. U. S. A.* **119**, (2022).

41. Arshad, N. *et al.* SARS-CoV-2 accessory proteins ORF7a and ORF3a use distinct mechanisms to downregulate MHC-I surface expression. *bioRxiv* 2022.05.17.492198 (2022) doi:10.1101/2022.05.17.492198.