



Communication

# Bioinformatics Analysis Identifies a Small ORF in the Genome of Fish Nidoviruses of Genus *Oncotshavirus* Predicted to Encode a Novel Integral Protein

Frederick Kibenge <sup>1,\*</sup> , Ashley McKibbon <sup>1</sup>, Molly Kibenge <sup>1</sup>  and Yingwei Wang <sup>2</sup>

<sup>1</sup> Department of Pathology and Microbiology, University of Prince Edward Island, Charlottetown, PE C1A 4P3, Canada; amckibbon@upei.ca (A.M.); mkibenge@upei.ca (M.K.)

<sup>2</sup> School of Mathematics and Computer Science, University of Prince Edward Island, Charlottetown, PE C1A 4P3, Canada; ywang@upei.ca

\* Correspondence: kibenge@upei.ca; Tel.: +1-(902)-566-0967

**Abstract:** Genome sequence analysis of Atlantic salmon bafinivirus (ASBV) revealed a small open reading frame (ORF) predicted to encode a Type I membrane protein with an N-terminal cleaved signal sequence (110 aa), likely an envelope (E) protein. Bioinformatic analyses showed that the predicted protein is strikingly similar to the coronavirus E protein in structure. This is the first report to identify a putative E protein ORF in the genome of members of the *Oncotshavirus* genus (subfamily Piscavirinae, family Tobaniviridae, order Nidovirales) and, if expressed would be the third family (after Coronaviridae and Arteriviridae) within the order to have the E protein as a major structural protein.

**Keywords:** fish viruses; nidovirus; integral protein; envelope



**Citation:** Kibenge, F.; McKibbon, A.; Kibenge, M.; Wang, Y. Bioinformatics Analysis Identifies a Small ORF in the Genome of Fish Nidoviruses of Genus *Oncotshavirus* Predicted to Encode a Novel Integral Protein. *Microbiol. Res.* **2021**, *12*, 753–764. <https://doi.org/10.3390/microbiolres12040055>

Academic Editor: Fernando Leal

Received: 6 September 2021

Accepted: 29 September 2021

Published: 6 October 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Nidoviruses that infect fish are classified into two families, Coronaviridae and Tobaniviridae, order Nidovirales [1]. The virus order designation is derived from Latin “nidus”, referring to the 3′-coterminal nested set of subgenomic mRNAs that characterize its genome transcription [2]. Nidovirus particles are enveloped with long single-stranded positive-sense polycistronic RNA genomes of ~12–41 kb in length—the largest among RNA viruses [3,4]. These viruses cause important diseases in many hosts, including humans, other mammals, birds, pythons, shrimp, and fish. Their relevance has skyrocketed with the emergence of coronavirus disease 2019 (COVID-19), an extremely infectious pandemic disease caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). As of 1 October 2021, this pandemic had infected ~234.8 million people worldwide, with more than 4.8 million deaths [5]. While the most currently known nidovirus species are associated with terrestrial hosts, the greatest phylogenetic diversity of nidoviruses is associated with hosts living in aquatic environments [6].

We isolated a novel salmonid nidovirus from farmed Atlantic salmon; its ultrastructural and genomic characteristics placed it in the genus *Oncotshavirus* (subfamily Piscavirinae, family Tobaniviridae, order Nidovirales), and we named it Atlantic salmon bafinivirus (ASBV) [7,8]. *Oncotshavirus* are among the least studied nidoviruses, with very little known about their host range, pathogenicity, geographical distribution, and replication kinetics. To date, genomic sequences of seven *oncotshavirus* isolated from various freshwater fish species have been deposited in the GenBank database (Table 1). All the genomes show an organization characteristic of the family Tobaniviridae [9] with a 5′ leader sequence followed by four major open reading frames (ORFs) that encode the putative replicase polyprotein (pp1ab) and the putative structural proteins spike (S), membrane (M), and nucleocapsid (N) proteins [10,11]. Consistent with the subfamily Torovirinae, the subfamily

Piscavirinae is also known to lack a homolog of the coronavirus (CoV) E protein [12]. The difference in the presence of E protein has been used to explain the structural differences between the coronaviruses and toroviruses, and E gene mutants of mouse hepatitis virus (a coronavirus) were shown to display bacilliform morphology [13] resembling that of members of the subfamilies Torovirinae and Piscavirinae [14].

**Table 1.** Oncotshavirus isolates <sup>1</sup> with complete genome sequences reported.

Virus Isolate	GenBank Accession No.	Genome Length (nt)	% Sequence Identity	Host Fish Species	Reference
CSBV NIDO	KJ681496	27,004	-	Chinook salmon, <i>Oncorhynchus tshawytscha</i>	[15]
ASBV VT01292015-09	KY130432	26,496	99.43%	Atlantic salmon, <i>Salmo salar</i>	[7]
CSBV WHQSR4345	MG600027	26,466	97.40%	Lesser spiny eel, <i>Macrogathus aculeatus</i>	[16]
CSBV HB93	MH171482	25,971	97.21%	Crucian carp, <i>Carassius auratus</i>	[17]
YCBV Shaoxing	MH822145	26,985	96.24%	Yellow catfish, <i>Pelteobagrus fulvidraco</i>	[11]
CSBV Cefas-W054	MT123520	25,969	97.99%	Goldfish, <i>Carassius auratus</i>	[10]
PFO-1 ZJLH18531	MT424676	26,996	96.14%	Yellow catfish, <i>Tachysurus fulvidraco</i>	[18]

<sup>1</sup> Oncotshavirus isolate Atlantic salmon bafinivirus (ASBV), Chinook salmon bafinivirus (CSBV), Yellow catfish bafinivirus (YCBV), and *Pelteobagrus fulvidraco* oncotshavirus-1 (PFO-1).

Here, we describe the structure and topology of a novel integral membrane protein encoded by a small ORF in the ASBV genome hitherto unknown in Tobaniviridae. We show that this ORF is present in members of the genus *Oncotshavirus* but not in *Bafinivirus*. Furthermore, the predicted protein is strikingly similar to the CoV E protein in structure. Biological validation of these predictions and elucidation of the role of E protein in the life cycle of oncotshaviruses are planned in future experiments. Thus, in future experiments, we will establish whether the ASBV E protein is produced during virus replication by investigating the kinetics of both transcript and protein production by virus-infected fish cell lines and cells transfected with recombinant plasmids encoding the ASBV E protein. The ASBV E transcripts would be quantified by RT-qPCR whereas the viral proteins would be detected by Western blotting or immunoprecipitation with antibody reagents to ASBV or ASBV E protein.

## 2. Materials and Methods

Nucleotide and amino acid sequences of 15 selected nidoviruses were obtained from the GenBank database [19]. Similarity analysis of the DNA sequences was performed using BLAST programs available via the National Center for Biotechnology Information [20]. The phylogenetic analysis was performed using CLUSTAL X package [21–23]. Different sequence sets were explored to find stable conserved areas. Both amino acid sequences and nucleotide sequences were explored. The sequence of Equine arteritis virus (EAV) (genus *Alphaarterivirus*, family Arteriviridae) GenBank accession number: NC\_002532, was chosen as the outgroup to determine the root of the phylogenetic trees. The bootstrapping procedure was performed to estimate the confidence level on the branches of these phylogenetic trees. The number of bootstrapping trials was 1000; random numbers were

used as seeds to simulate the random processes. Transmembrane topology prediction was obtained by using DeepTMHMM [24] and TOPCONS [25], two predictors available on the internet that separate signal peptides from N-terminal transmembrane domains [26,27]. Any predicted signal peptides and the location of their cleavage sites were confirmed using SignalP 5.0 [28]. NetNGlyc-1.0 [29] was used to check for N-linked glycosylation sites. GPS-Palm [30] was used for the prediction of S-palmitoylation sites. Phyre2 and AlphaFold2 [31] were used to build a 3D model of the predicted protein structure.

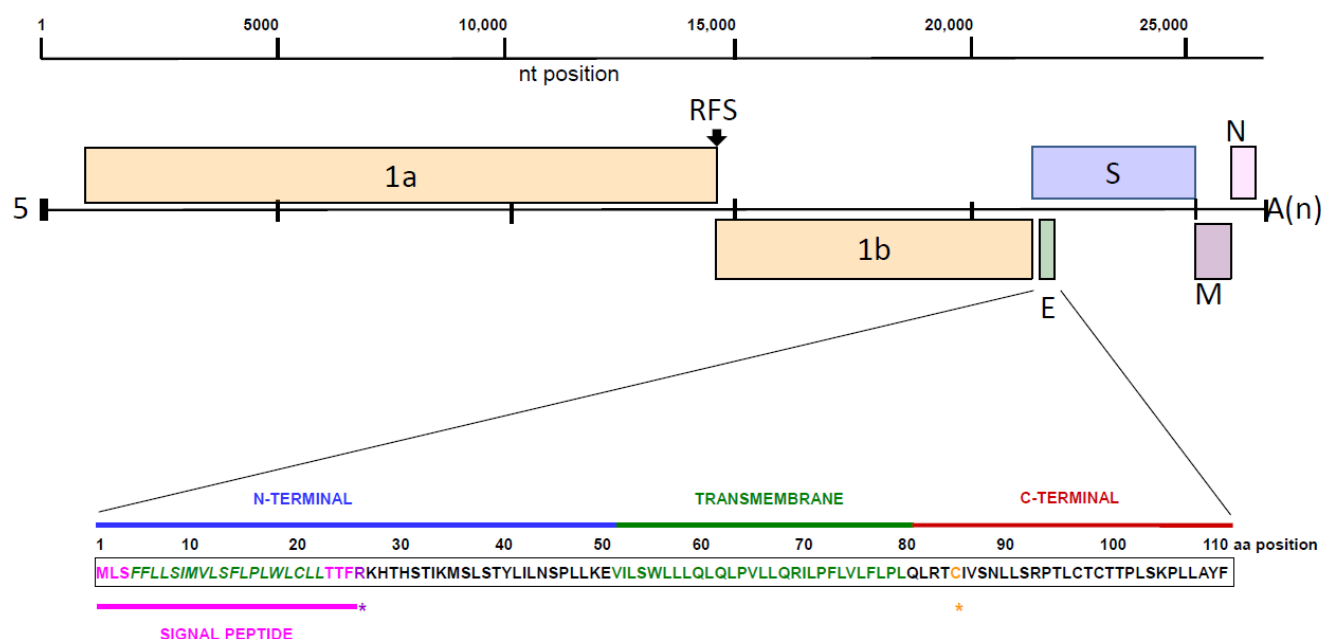
### 3. Results and Discussion

In the current study, the genomic sequence of ASBV VT01292015-09 obtained by next-generation sequencing (NGS) on the Illumina® HiSeq 2000 platform (LC Sciences, Houston, TX, USA) and completed by filling in the gaps between the assembled contigs using conventional RT-PCR and the 5'-RACE and 3'-RACE to obtain the 5' and 3' terminal sequences, respectively [7], was further analyzed. The full-length viral genome is 26,492 nt starting at the 5' GCA terminus with the untranslated region of 846 nt and the 3' end of the untranslated region of 220 nt, including a poly(A) tail of 23 nt. The four major ORFs identified using the ORF Finder program [32], which are shown in Figure 1, are in the following order: ORF 1ab encodes the putative replicase polyprotein, nt 847 to 14,361 for pp1a and nt 14,496 to 21,338 for pp1b (pp1ab, 6778 aa); S, nt 21,341 to 24,958 (1205 aa); M, nt 24,975 to 25,712 (245 aa); and N, nt 25,736 to 26,272 (178 aa). In addition, a small ORF (marked as E), nt 21,510 to 21,846 (110 aa), is present between the pp1ab and M ORFs. Similar to coronaviruses [33,34] and toroviruses [35], ASBV uses a “slippery” heptanucleotide: <sup>14,341</sup>TTTAAAC<sup>14,347</sup> at a ribosomal frameshift (RFS) site, resulting in a −1 frameshift from ORF1a to ORF1b (Figure 1).

The small ORF in the ASBV genome is predicted to encode an  $\alpha$ -helical transmembrane protein (110 aa); it is an integral viral membrane protein [36], likely a putative E protein previously overlooked in genomes of members of the family Tobamoviridae probably because of its small size. This ORF is present in members of the genus *Oncotshavirus* but not in *Bafinivirus*. Moreover, the gene order is conserved in oncotshaviruses as expected of a structural protein, although the different structural protein ORFs of oncotshaviruses have not yet been experimentally verified for expression. Relative to ASBV E protein (110 aa), the other six oncotshavirus isolates (Table S1) have a truncated E protein (86 aa, i.e., a shorter C-terminus). However, all the essential features of the putative integral membrane protein are conserved in the seven oncotshaviruses. As shown in Figure 2A, in isolate Yellow catfish bafinivirus (YCBV) Shaoxing, the proposed signal peptidase cleavage site is at residue K27 instead of R26, and in Chinook salmon bafinivirus (CSBV) HB93, the mutation at position 82 (R → I) results in a shorter transmembrane domain (TMD) (22 aa) whereas if accompanied by a mutation at position 83 (T → K) in isolates CSBV WHQSR4345 and *Pelteobagrus fulvidraco* oncotshavirus-1 (PFO-1) ZJLH18531, they result in a longer TMD (32 aa). Some of these mutations may be due to the complex transcription characteristics of nidoviruses [37–39], or they could be sequencing errors.

Several lines of evidence lend support to the view that the ASBV small ORF is authentic and encodes an integral membrane protein. Bioinformatic analyses of the amino acid sequence using the latest and most accurate softwares, DeepTMHMM [24] and TOPCONS [25], show that the predicted protein is strikingly similar to the CoV E protein in structure [40–42], albeit unique to members of genus *Oncotshavirus*. The CoV E protein is a Type I membrane protein with a single TMD [40,43,44] and does not have a canonical cleaved signal sequence [40,45] whereas, as defined by Goder and Spiess [46], the ASBV E protein is predicted to be a Type I membrane protein with an N-terminal cleaved signal sequence as shown in Figures 1 and 2. Thus the N-terminus consists of a signal peptide (SP) (aa positions 4–24) followed by 25 amino acids and then a TMD of 29 residues (aa positions 51–79) and C-terminus approximately 30 amino acids (aa positions 80–110) (Figure 1). The signal peptide was confirmed using SignalP 5.0 software [28]. The proposed topology of the putative ASBV E protein is illustrated in Figure 2B. The hydrophobic regions in the

SP and the TMD impose a hairpin topology for the ASBV E protein before the N-terminal SP being cleaved off at residue R26. The cleavage site at R26 was confirmed using PeptideCutter software [47] and is +3 amino acids from the hydrophobic segment of the SP, which conforms to the structure of a typical cleavable amino-terminal signal sequence [28]. Moreover, the location and size of the predicted SP region are characteristic of signal sequences in eukaryotes (usually 16 to 30 amino acid residues in length and comprising a hydrophilic, usually positively charged N-terminal region, a central hydrophobic domain, and a C-terminal region with the cleavage site for signal peptidase [48]). The mature protein would have the N-terminus on the ER lumen/virion exterior and the C-terminus exposed to the cytoplasmic side/virion interior, as illustrated by Goder and Spiess [46] and Figure 2B. The predicted 3D structure of ASBV E protein shown in Figure 2C removes any ambiguity because it is similar to that of SARS-CoV-2 E protein [49].

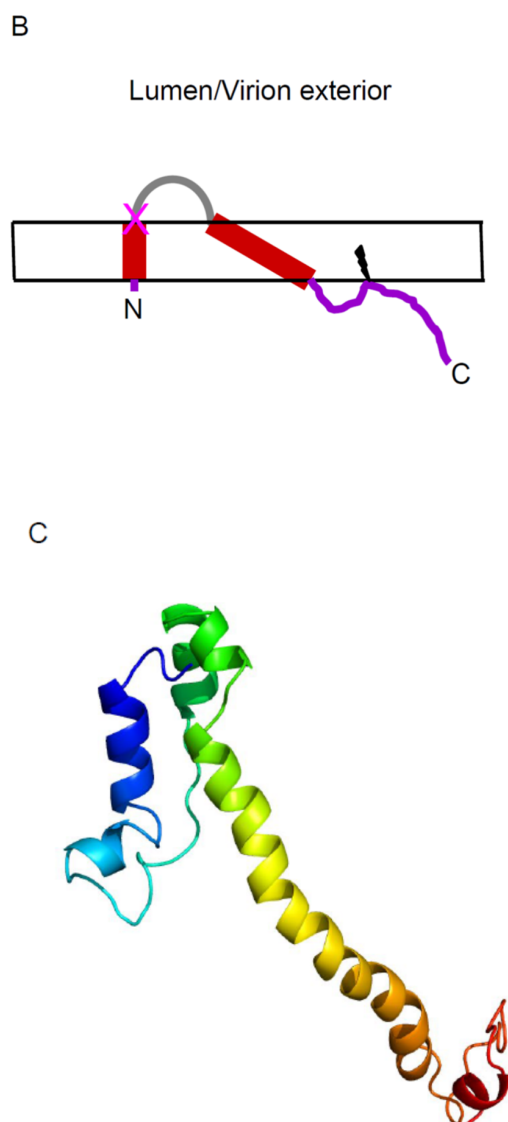


**Figure 1.** The genome architecture of ASBV VT01292015-09 with canonical features of nidovirus genomes annotated. The full-length RNA genome is 26,496 nucleotides. Boxes represent major open reading frames (ORFs). The proteins encoded by the ORFs are also indicated within, above, or below the boxes. The arrow indicates the position of the putative ribosomal frameshift (RFS) “slippery” sequence (putative replicase polyprotein 1a and 1b; putative envelope protein E (p12.7); S, spike protein; M, membrane protein; and N, nucleocapsid). The 5′ capped mRNA with a leader sequence is depicted by a small black box. The poly(A) tail is indicated by A(n). The predicted structure of the ASBV E protein consisting of three domains: the amino-terminal (N-terminal at aa positions 1–50), the putative transmembrane  $\alpha$ -helical hydrophobic domain (at aa positions 51–79), and the carboxy-terminal (C-terminal at positions 80–110) is shown. The ASBV E protein is predicted to be a Type I membrane protein with an N-terminal cleaved signal sequence. The residues in the two hydrophobic domains are in green, one in italics (aa positions 4–22) is in the signal peptide region (pink at aa positions 1–25), and another in the transmembrane domain (green at aa positions 51–79), which would impose a hairpin topology for the ASBV E protein before the N-terminal signal peptide is cleaved off. A signal peptidase cleavage site is marked with a star (\*) at residue R26 in purple (proposed to be in the ER lumen/virion exterior resulting in a Type I membrane protein (i.e., a mature protein with N-terminus on the ER lumen/virion exterior and the C-terminus exposed to the cytoplasmic side/virion interior). The predicted palmitoylation site at residue C84 is marked with a star (\*) in orange.

Family *Tobaniviridae*, subfamily *Piscavirinae*,  
genus *Oncotshavirus*:

Figure 2. *Cont.*





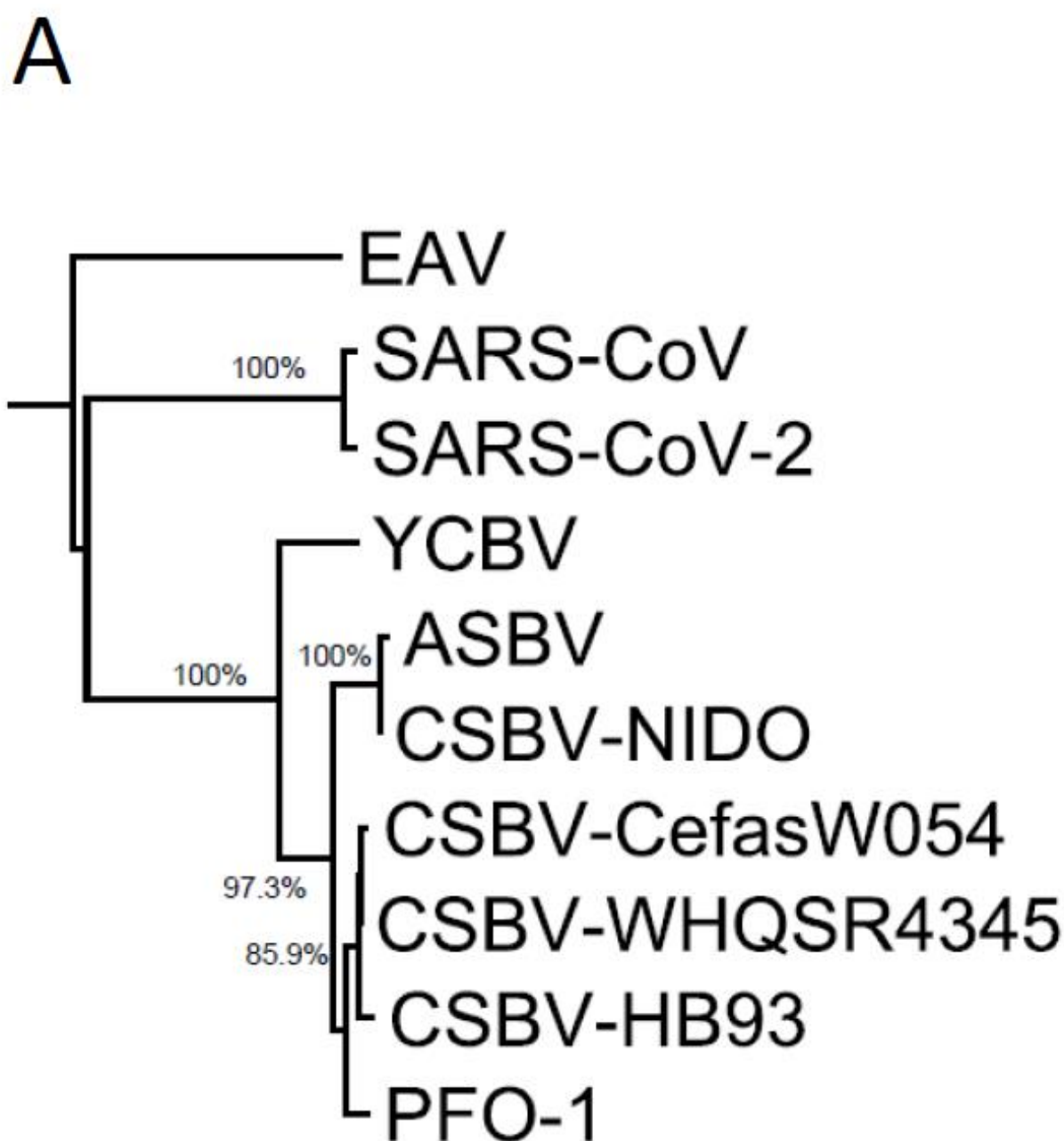
**Figure 2.** (A) Amino acid sequences and topologies of the envelope (E) protein for 15 selected nidoviruses using DeepTMHMM Topology Prediction ([24], DeepTMHMM. <https://biolib.com/DTU/DeepTMHMM>) (accessed on 9 July 2021). Abbreviations: S = Signal peptide (the putative  $\alpha$ -helical hydrophobic domain at aa 4–22 is highlighted in bold in blue and the signal peptidase cleavage site is in bold in red as residue R26 or K27); O = outside (lumen side/virion exterior); M = Transmembrane domain/region (is highlighted in bold in cyan); I = inside (cytoplasmic side/virion interior); Virus names: ASBV = Atlantic salmon bafinivirus VT01292015-09; CSBV-NIDO = Chinook salmon bafinivirus NIDO; CSBV Cefas-W054; CSBV WHQSR4345; CSBV HB93; PFO-1 = *Pelteobagrus fulvidraco* oncotshavirus-1 ZJLH18531; YCBV = Yellow catfish bafinivirus Shaoxing; TGEV = Transmissible gastroenteritis virus Purdue; MHV = Murine hepatitis virus A59; MERS-CoV = Middle Eastern respiratory syndrome coronavirus HCoV-EMC/2012; SARS-CoV = Severe acute respiratory syndrome coronavirus Tor2; SARS-CoV-2 Wuhan-Hu-1; IBV = Infectious bronchitis virus Beaudette; PorCoV = Porcine coronavirus HKU15; and EAV = Equine arteritis virus Bucyrus. (B) The proposed topology of ASBV E protein is predicted to be a Type I membrane protein with an N-terminal cleaved signal sequence. The hydrophobic regions in the signal peptide and the transmembrane domains impose a hairpin topology for the ASBV E protein before the N-terminal signal peptide cleaves off. The mature protein would have the N-terminus on the ER Lumen/virion exterior, and the C-terminus exposed to the cytoplasmic side/virion interior with a putative palmitoylation site at residue C84 in the C-terminus. The proposed signal peptidase cleavage site at residue R26 is marked with X. (C) The predicted 3D structure of ASBV E protein using Phyre2 and AlphaFold2 [31] and homology modelling to the SARS-CoV-2 E protein structure [49]. The ASBV E protein model includes the cleavable signal peptide which is absent in SARS-CoV-2 E. Image coloured by rainbow N  $\rightarrow$  C terminus. Model dimensions (Å): X:45.023 Y:73.238 Z:39.549.

The E protein amino acid sequence found in different CoVs is variable [40], but the predicted structure and functional properties are highly conserved [42]. In this study, we were unable to construct phylogenetic trees to include the E gene sequences of all 15 selected nidoviruses (Figure 2A) as we could not find evidence that they were related. However, we were able to generate two almost identical phylogenetic trees, one based on amino acid sequences and another on nucleotide sequences, revealing two major clades consisting of the seven members of genus *Oncotshavirus* (Table 1) and two members of genus *Betacoronavirus*, lineage B (SARS-CoV and SARS-CoV-2) [1] (Figure 3). The oncotshavirus group of seven sequences and the betacoronavirus group of two sequences are stable groups; we are confident in them. For the rest of the sequences, their relationships with these two groups cannot be determined through the analyses of the E protein. Such relationships might be found by analyzing other protein sequences or nucleotide sequences such as the replicase polyprotein and nucleoprotein sequences, and any failures demonstrate limitations of these phylogenetic approaches in dealing with the diversity of nidoviruses [6].

The ASBV E protein is non-glycosylated. To date, among the CoV E proteins, only the SARS-CoV E protein has been glycosylated (residue N66), although this constituted an alternative minor form [42,50]; the glycosylation of SARS-CoV E protein during actual infection and its biological function remain to be further investigated [43]. A post-translational modification of more functional importance for E proteins is palmitoylation. Bioinformatic analysis of the ASBV E amino acid sequence using the GPS-Palm software to predict palmitoylation sites in proteins [29] demonstrated that, under the high threshold, residue C84 in the C-terminus is modified by palmitoylation. Of the CoV E proteins, only infectious bronchitis virus (IBV), SARS-CoV, and mouse hepatitis virus (MHV) are palmitoylated [51–53].

Similar to the CoV E protein [48], the ASBV E protein may belong to the class of small viral integral proteins called ‘vioporins’ that oligomerize within the membrane bilayer to form channels that facilitate the transport of ions or small molecules. They include the M2 protein of influenza virus, the 6K protein of alphaviruses, Vpu of HIV-1 [36], p7 ion channel of Hepatitis C virus (HCV) [54,55], as well as protein 3A of CoVs [56]. They are thought to function in various ways to facilitate the assembly and release of new viral particles from the infected cells [36].

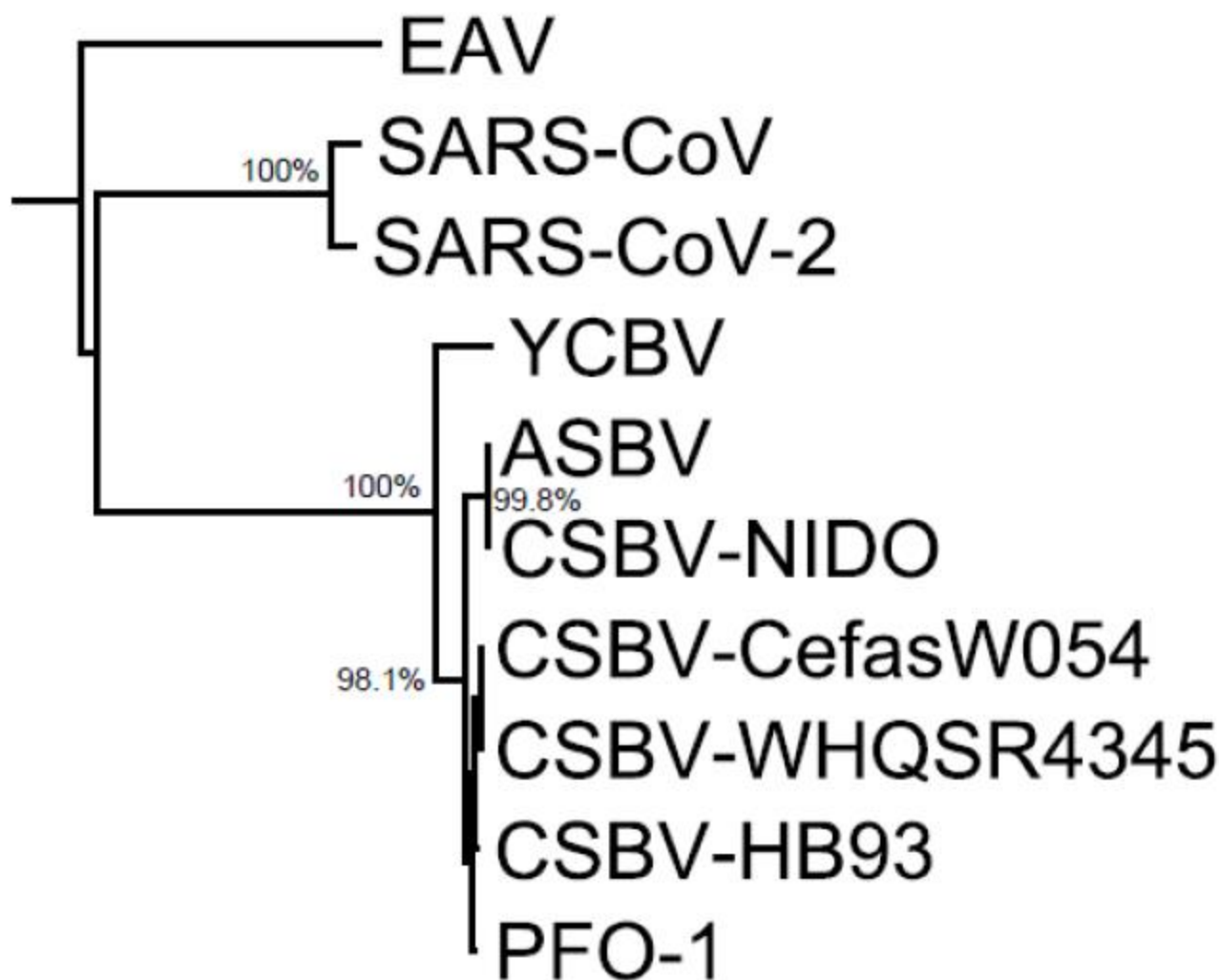
In summary, this is the first report to identify a putative E protein ORF in Tobaniviridae, and if expressed, would be the third family (after Coronaviridae and Arteriviridae) within the order Nidovirales to have the E protein as a major structural protein. Experiments to biologically validate these predictions and elucidate the role of E protein in the life cycle of oncotshaviruses are planned.



**Figure 3.** Phylogeny of selected nidoviruses based on the amino acid and nucleotide sequences of the E proteins. Phylogenetic tree based on (A) amino acid sequences and (B) nucleotide sequences, of seven members of genus *Oncotshavirus* (listed in Table 1) and two members of genus *Betacoronavirus*, lineage B (SARS-CoV and SARS-CoV-2) with Equine arteritis virus (EAV) E protein sequence GenBank accession number: NC\_002532) as the outgroup to determine its root. The graphical editor for phylogenetic trees, TreeGraph 2 [57] was used to produce the figures. Each bootstrapping value corresponds to the branch on the same vertical level. Only the bootstrapping supports higher than 70% are marked in the phylogenetic trees. Abbreviations at the end of each sequence correspond to the virus names: ASBV = Atlantic salmon bafinivirus VT01292015-09; CSBV-NIDO = Chinook salmon bafinivirus NIDO; CSBV Cefas-W054; CSBV WHQSR4345; CSBV HB93; YCBV = Yellow catfish bafinivirus Shaoxing; PFO-1 = *Pelteobagrus fulvidraco* oncotshavirus-1 ZJLH18531; SARS-CoV = Severe acute respiratory syndrome coronavirus Tor2; SARS-CoV-2 Wuhan-Hu-1; and EAV = Equine arteritis virus Bucyrus.



B

Figure 3. *Cont.*

**Supplementary Materials:** The following is available online at <https://www.mdpi.com/article/10.3390/microbiolres12040055/s1>, Table S1: Membrane topologies of E proteins of 15 selected nidoviruses predicted using DeepTMHMM <https://biolib.com/DTU/DeepTMHMM> (accessed on 9 July 2021).

**Author Contributions:** Conceptualization, F.K.; methodology, F.K., A.M., M.K. and Y.W.; software, F.K. and Y.W.; validation, F.K., M.K. and Y.W.; formal analysis, F.K. and Y.W.; investigation, F.K., A.M., M.K. and Y.W.; resources, F.K. and Y.W.; data curation, F.K. and Y.W.; writing—original draft preparation, F.K.; writing—review and editing, F.K., A.M., M.K. and Y.W.; visualization, F.K., A.M., M.K. and Y.W.; supervision, F.K. and Y.W.; project administration, F.K.; funding acquisition, F.K. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Natural Sciences and Engineering Research Council (NSERC), funding reference number RGPIN-2016-04462. Cette Recherche a Été Financée par le Conseil de Recherches en Sciences Naturelles et en Genie du Canada (CRSNG), Numéro de Reference RGPIN-2016-04462.

**Institutional Review Board Statement:** The in-vitro studies were approved by the UPEI Biosafety Committee, approval number 6008719.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The full-length viral genome sequence of ASBV VT01292015-09 has been deposited in GenBank (accession no. KY130432.2).

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. International Committee on Taxonomy of Viruses. ICTV Master Species List 2019 v2. Checklist Dataset. 2019. Available online: <https://www.gbif.org/dataset/e01b0cbb-a10a-420c-b5f3-a3b20cc266ad> (accessed on 3 July 2020).
2. Sola, I.; Almazan, F.; Zúñiga, S.; Enjuanes, L. Continuous and Discontinuous RNA Synthesis in Coronaviruses. *Annu. Rev. Virol.* **2015**, *2*, 265–288. [CrossRef] [PubMed]
3. Boon, J.A.D.; Snijder, E.J.; Chirnside, E.D.; A de Vries, A.; Horzinek, M.C.; Spaan, W.J. Equine arteritis virus is not a togavirus but belongs to the coronaviruslike superfamily. *J. Virol.* **1991**, *65*, 2910–2920. [CrossRef]
4. Saberi, A.; Gulyaeva, A.A.; Brubacher, J.L.; Newmark, P.A.; Gorbalenya, A.E. A planarian nidovirus expands the limits of RNA genome size. *PLoS Pathog.* **2018**, *14*, e1007314. [CrossRef] [PubMed]
5. Worldometer. Coronavirus Update (Live). Available online: <https://www.worldometers.info/coronavirus/> (accessed on 1 October 2021).
6. Bukhari, K.; Mulley, G.; Gulyaeva, A.A.; Zhao, L.; Shu, G.; Jiang, J.; Neuman, B.W. Description and initial characterization of metatranscriptomic nidovirus-like genomes from the proposed new family Abyssoviridae, and from a sister group to the Coronavirinae, the proposed genus *Alphaletovirus*. *Virology* **2018**, *524*, 160–171. [CrossRef] [PubMed]
7. GenBank Database. Available online: <https://www.ncbi.nlm.nih.gov/nuccore/KY130432> (accessed on 9 July 2021).
8. McKibbin, A. Analysis of Structural Proteins of Atlantic Salmon Bafinivirus—A Coronavirus of Fish. Master's Thesis, University of Prince Edward Island, Charlottetown, PE, Canada, 21 May 2020.
9. Schütze, H. Coronaviruses in aquatic organisms. In *Aquaculture Virology*; Kibenge, F.S.B., Godoy, M.M., Eds.; Elsevier: Amsterdam, The Netherlands, 2016; pp. 323–332.
10. Cano, I.; Stone, D.; Savage, J.; Wood, G.; Mulhearn, B.; Gray, J.; Stinton, N.; Ross, S.; Bonar, M.; Taylor, N.G.H.; et al. Isolation of a Chinook Salmon Bafinivirus (CSBV) in Imported Goldfish *Carassius auratus* L. in the United Kingdom and Evaluation of Its Virulence in Resident Fish Species. *Viruses* **2020**, *12*, 578. [CrossRef]
11. Zhang, X.; Shen, W.; Xu, C.; Wang, Y.; Xu, H.; Liu, X.; Wei, Y. Discovery of a novel Piscanivirus in yellow catfish (*Pelteobagrus fulvidraco*) in China. *Infect. Genet. Evol.* **2019**, *74*, 103924. [CrossRef]
12. De Groot, R.J.; Baker, S.C.; Baric, R.; Enjuanes, L.; Gorbalenya, A.E.; Holmes, K.V.; Perlman, S.; Poon, L.; Rottier, P.J.M.; Talbot, P.J.; et al. Family Coronaviridae. In *Virus Taxonomy: Classification and Nomenclature of Viruses: Ninth Report of the International Committee on Taxonomy of Viruses*; King, A.M.Q., Adams, M.J., Carstens, E.B., Lefkowitz, E.J., Eds.; Elsevier: Amsterdam, The Netherlands, 2012; pp. 806–828.
13. Fischer, F.; Stegen, C.F.; Masters, P.S.; Samsonoff, W.A. Analysis of constructed E gene mutants of mouse hepatitis virus confirms a pivotal role for E protein in coronavirus assembly. *J. Virol.* **1998**, *72*, 7885–7894. [CrossRef]
14. MacLachlan, N.J.; Dubovi, E.J. Chapter 24, Coronaviridae. In *Fenner's Veterinary Virology*, 5th ed.; MacLachlan, N.J., Dubovi, E.J., Eds.; Academic Press: Amsterdam, The Netherlands, 2017; pp. 393–413.
15. GenBank Database. Available online: <https://www.ncbi.nlm.nih.gov/nuccore/KJ681496> (accessed on 9 July 2021).
16. Shi, M.; Lin, X.-D.; Chen, X.; Tian, J.-H.; Chen, L.-J.; Li, K.; Wang, W.; Eden, J.-S.; Shen, J.-J.; Liu, L.; et al. The evolutionary history of vertebrate RNA viruses. *Nat. Cell Biol.* **2018**, *556*, 197–202. [CrossRef]
17. Chen, X.-Y.; Zhou, Y.; Chen, X.; Zhang, J.; Zeng, X.-D.; Ji, F.; Xu, L.-M. Isolation and genetic analysis of a nidovirus from crucian carp (*Carassius auratus*). *Arch. Virol.* **2019**, *164*, 1651–1654. [CrossRef]
18. GenBank Database. Available online: <https://www.ncbi.nlm.nih.gov/nuccore/MT424676> (accessed on 9 July 2021).
19. Sayers, E.W.; Cavanaugh, M.; Clark, K.; Ostell, J.; Pruitt, K.D.; Karsch-Mizrachi, I. GenBank. *Nucleic Acids Res.* **2019**, *48*, D84–D86. [CrossRef]
20. Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J. Basic local alignment search tool. *J. Mol. Biol.* **1990**, *215*, 403–410. [CrossRef]
21. Higgins, D.; Sharp, P. CLUSTAL: A package for performing multiple sequence alignment on a microcomputer. *Gene* **1988**, *73*, 237–244. [CrossRef]
22. Larkin, M.A.; Blackshields, G.; Brown, N.P.; Chenna, R.; McGettigan, P.A.; McWilliam, H.; Valentin, F.; Wallace, I.M.; Wilm, A.; Lopez, R.; et al. Clustal W and Clustal X version 2.0. *Bioinformatics* **2007**, *23*, 2947–2948. [CrossRef]
23. Thompson, J.D.; Gibson, T.J.; Plewniak, F.; Jeanmougin, F.; Higgins, D.G. The CLUSTAL\_X windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **1997**, *25*, 4876–4882. [CrossRef] [PubMed]
24. DeepTMHMM. Available online: <https://biolib.com/DTU/DeepTMHMM> (accessed on 9 July 2021).

25. Tsirigos, K.; Peters, C.; Shu, N.; Käll, L.; Elofsson, A. The TOPCONS web server for consensus prediction of membrane protein topology and signal peptides. *Nucleic Acids Res.* **2015**, *43*, W401–W407. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Käll, L. Prediction of Transmembrane Topology and Signal Peptide Given a Protein's Amino Acid Sequence. *Methods Mol. Biol.* **2010**, *673*, 53–62. [\[CrossRef\]](#) [\[PubMed\]](#)
27. Lao, D.M.; Arai, M.; Ikeda, M.; Shimizu, T. The presence of signal peptide significantly affects transmembrane topology prediction. *Bioinformatics* **2002**, *18*, 1562–1566. [\[CrossRef\]](#) [\[PubMed\]](#)
28. SignalP-5.0 Signal Peptide and Cleavage Sites in Gram+, Gram-and Eukaryotic Amino Acid Sequences. Available online: <https://services.healthtech.dtu.dk/service.php?SignalP-5.0>. (accessed on 9 July 2021).
29. Gupta, R.; Brunak, S. Prediction of glycosylation across the human proteome and the correlation to protein function. *Pac. Symp. Biocomput.* **2002**, *7*, 310–322.
30. Ning, W.; Jiang, P.; Guo, Y.; Wang, C.; Tan, X.; Zhang, W.; Peng, D.; Xue, Y. GPS-Palm: A deep learning-based graphic presentation system for the prediction of S-palmitoylation sites in proteins. *Brief. Bioinform.* **2021**, *22*, 1836–1847. [\[CrossRef\]](#) [\[PubMed\]](#)
31. Kelley, L.A.; Mezulis, S.; Yates, C.M.; Wass, M.N.; Sternberg, M.J.E. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **2015**, *10*, 845–858. [\[CrossRef\]](#)
32. Stothard, P. The Sequence Manipulation Suite: JavaScript programs for analyzing and for-matting protein and DNA sequences. *Biotechniques* **2000**, *28*, 1102–1104. Available online: <https://www.bioinformatics.org/sms2/> (accessed on 9 July 2021). [\[CrossRef\]](#) [\[PubMed\]](#)
33. Brierley, I.; Bournsnel, M.E.; Binns, M.M.; Bilimoria, B.; Blok, V.C.; Brown, T.D.; Inglis, S.C. An efficient ribosomal frame-shifting signal in the polymerase-encoding region of the coronavirus IBV. *EMBO J.* **1987**, *6*, 3779–3785. [\[CrossRef\]](#) [\[PubMed\]](#)
34. Brierley, I.; Digard, P.; Inglis, S.C. Characterization of an efficient coronavirus ribosomal frameshifting signal: Requirement for an RNA pseudoknot. *Cell* **1989**, *57*, 537–547. [\[CrossRef\]](#)
35. Brierley, I. Ribosomal frameshifting on viral RNAs. *J. Gen. Virol.* **1995**, *76*, 1885–1892. [\[CrossRef\]](#)
36. Lenard, J. Viral membranes. In *Encyclopedia of Virology*, 3rd ed.; Mahy, B.W.J., Van Regenmortel, M.H.V., Eds.; Academic Press: Cambridge, MA, USA, 2008; pp. 308–314.
37. Kim, D.; Lee, J.-Y.; Yang, J.-S.; Kim, J.W.; Kim, V.N.; Chang, H. The Architecture of SARS-CoV-2 Transcriptome. *Cell* **2020**, *181*, 914–921.e10. [\[CrossRef\]](#)
38. Pasternak, A.; Spaan, W.J.M.; Snijder, E. Nidovirus transcription: How to make sense ... ? *J. Gen. Virol.* **2006**, *87*, 1403–1421. [\[CrossRef\]](#)
39. Stewart, H.; Brown, K.; Dinan, A.M.; Irigoyen, N.; Snijder, E.J.; Firth, A.E. Transcriptional and Translational Landscape of Equine Torovirus. *J. Virol.* **2018**, *92*, e00589-18. [\[CrossRef\]](#)
40. Duarte, G.; García-Murriat, M.J.; Graut, B.; Acosta-Cáceres, J.M.; Martínez-Gil, L.; Mingarro, I. SARS-CoV-2 envelope protein topology in eukaryotic membranes. *Open Biol.* **2020**, *10*, 200209. [\[CrossRef\]](#)
41. Nieto-Torres, J.L.; DeDiego, M.L.; Álvarez, E.; Jiménez-Guardeño, J.M.; Regla-Nava, J.A.; Llorente, M.; Kremer, L.; Shuo, S.; Enjuanes, L. Subcellular location and topology of severe acute respiratory syndrome coronavirus envelope protein. *Virology* **2011**, *415*, 69–82. [\[CrossRef\]](#) [\[PubMed\]](#)
42. Schoeman, D.; Fielding, B.C. Coronavirus envelope protein: Current knowledge. *Virol. J.* **2019**, *16*, 1–22. [\[CrossRef\]](#) [\[PubMed\]](#)
43. Fung, T.S.; Liu, D.X. Post-translational modifications of coronavirus proteins: Roles and function. *Futur. Virol.* **2018**, *13*, 405–430. [\[CrossRef\]](#)
44. Ruch, T.R.; Machamer, C.E. The Coronavirus E Protein: Assembly and Beyond. *Viruses* **2012**, *4*, 363–382. [\[CrossRef\]](#)
45. Raamsman, M.J.B.; Locker, J.K.; de Hoo, A.; de Vries, A.A.F.; Griffiths, G.; Vennema, H.; Rottier, P.J.M. Characterization of the Coronavirus Mouse Hepatitis Virus Strain A59 Small Membrane Protein E. *J. Virol.* **2000**, *74*, 2333–2342. [\[CrossRef\]](#)
46. Goder, V.; Spiess, M. Topogenesis of membrane proteins: Determinants and dynamics. *FEBS Lett.* **2001**, *504*, 87–93. [\[CrossRef\]](#)
47. Gasteiger, E.; Hoogland, C.; Gattiker, A.; Duvaud, S.; Wilkins, M.R.; Appel, R.D.; Bairoch, A. Protein Identification and Analysis Tools on the ExPASy Server. In *The Proteomics Protocols Handbook*; Walker, J.M., Ed.; Humana Press: Totowa, NJ, USA, 2005; pp. 571–607. Available online: <http://www.expasy.org/tools/peptidecutter/> (accessed on 9 July 2021).
48. Kapp, K.; Schrempf, S.; Lemberg, M.K.; Dobberstein, B. Chapter 1, Post-Targeting Functions of Signal Peptides. In *Protein Transport into the Endoplasmic Reticulum*; Zimmermann, R., Ed.; Landes Bioscience: Austin, TX, USA, 2009. Available online: <https://www.ncbi.nlm.nih.gov/books/NBK6322/> (accessed on 9 July 2021).
49. Mandala, V.S.; McKay, M.J.; Shcherbakov, A.A.; Dregni, A.J.; Kolocouris, A.; Hong, M. Structure and drug binding of the SARS-CoV-2 envelope protein transmembrane domain in lipid bilayers. *Nat. Struct. Mol. Biol.* **2020**, *27*, 1202–1208. [\[CrossRef\]](#)
50. Yuan, Q.; Liao, Y.; Torres, J.; Tam, J.P.; Liu, D. Biochemical evidence for the presence of mixed membrane topologies of the severe acute respiratory syndrome coronavirus envelope protein expressed in mammalian cells. *FEBS Lett.* **2006**, *580*, 3192–3200. [\[CrossRef\]](#)
51. Corse, E.; Machamer, C.E. The Cytoplasmic Tail of Infectious Bronchitis Virus E Protein Directs Golgi Targeting. *J. Virol.* **2002**, *76*, 1273–1284. [\[CrossRef\]](#) [\[PubMed\]](#)
52. Liao, Y.; Yuan, Q.; Torres, J.; Tam, J.P.; Liu, D. Biochemical and functional characterization of the membrane association and membrane permeabilizing activity of the severe acute respiratory syndrome coronavirus envelope protein. *Virology* **2006**, *349*, 264–275. [\[CrossRef\]](#) [\[PubMed\]](#)

- 
53. Lopez, L.A.; Riffle, A.J.; Pike, S.L.; Gardner, D.; Hogue, B.G. Importance of Conserved Cysteine Residues in the Coronavirus Envelope Protein. *J. Virol.* **2008**, *82*, 3000–3010. [[CrossRef](#)] [[PubMed](#)]
  54. Breiting, U.; Farag, N.; Ali, N.K.; Breiting, H.-G. Patch-clamp study of Hepatitis C p7 channels reveals genotype-specific sensitivity to inhibitors. *Biophys. J.* **2016**, *110*, 2419–2429. [[CrossRef](#)] [[PubMed](#)]
  55. Farag, N.S.; Breiting, U.; El-Azizi, M.; Breiting, H.-G. The p7 viroporin of hepatitis C virus contributes to liver inflammation by stimulating production of Interleukin-1 $\beta$ . *Biochim. Biophys. Acta Mol. Basis Dis.* **2017**, *1863*, 712–720. [[CrossRef](#)] [[PubMed](#)]
  56. Lu, W.; Zheng, B.J.; Xu, K.; Schwarz, W.; Du, L.; Wong, C.K.L.; Chen, J.; Duan, S.; Deubel, V.; Sun, B. Severe acute respiratory syndrome-associated coronavirus 3a protein forms an ion channel and modulates virus release. *Proc. Natl. Acad. Sci. USA.* **2006**, *103*, 12540–12545. [[CrossRef](#)]
  57. Stöver, B.C.; Müller, K.F. TreeGraph 2: Combining and visualizing evidence from different phylogenetic analyses. *BMC Bioinform.* **2010**, *11*, 7. Available online: <http://www.biomedcentral.com/1471-2105/11/7> (accessed on 3 October 2021). [[CrossRef](#)] [[PubMed](#)]