Preprints.org

Article

Not peer-reviewed version

Potential Transcriptional Enhancers in Coronaviruses: From Infectious Bronchitis Virus to SARS-CoV-2

Roberto Patarca * and William A. Haseltine

Posted Date: 26 June 2024

doi: 10.20944/preprints202406.1845.v1

Keywords: Coronavirus; enhancer; long-range RNA-RNA interactions; IBV; SARS-CoV-2; viral attenuation; host immune evasion

Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 June 2024 doi:10.20944/preprints202406.1845.v1

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Potential Transcriptional Enhancers in Coronaviruses: From Infectious Bronchitis Virus to SARS-CoV-2

Roberto Patarca [1,2,*] and William A. Haseltine [1,2]

[1] ACCESS Health International, 384 West Lane, Ridgefield, CT 06877, USA; roberto.patarca@accessh.org; william.haseltine@accesh.org;
[2] Feinstein Institutes for Medical Research, 350 Community Dr, Manhasset, NY 11030 USA;
* Correspondence: roberto.patarca@accessh.org

Abstract: Coronaviruses constitute a global threat to human and animal health. It is essential to investigate the long-distance RNA-RNA interactions that approximate remote regulatory elements in strategies, including genome circularization, discontinuous transcription, and transcriptional enhancers, aimed at rapid replication of their large genomes, pathogenicity, and immune evasion. Based on the primary sequences and modeled RNA-RNA interactions of two experimentally defined coronaviral enhancers, we detected via in silico primary and secondary structural analysis potential enhancers in various coronaviruses, from the phylogenetically ancient avian IBV to the recently emerged SARS-CoV-2. These potential enhancers possess a core duplex-forming region that could transition between closed and open states, as molecular switches directed by viral or host factors. The open state would permit the enhancer to pair with remote sequences in the viral genome and modulate the expression of crucial genes involved in viral replication and host immune evasion. Consistently, variations in the predicted IBV enhancer region or its distant targets coincide with cases of viral attenuation possibly driven by decreased expression of the ORF3a immune evasion protein. The annotated enhancer sequences could inform structural prediction tools and antiviral interventions if validated experimentally.

Keywords: coronavirus; enhancer; long-range RNA interactions; IBV; SARS-CoV-2; viral attenuation; host immune evasion

1. Introduction

Coronaviruses employ complex replicative strategies involving long-range RNA-RNA interactions. These strategies, which include genome circularization, discontinuous transcription, and viral enhancers, approximate regulatory sequences in their large genomes influencing replication, pathogenicity, and immune evasion [1–6].

Genome circularization approximates regulatory sequences in the 5′- and 3′-untranslated regions (UTRs) via the interaction of complementary sequences in the UTRs facilitated by viral and host protein bridges during the synthesis of subgenomic negative-sense strands [5–21]. Discontinuous transcription generates a nested set of subgenomic RNAs via transcription regulatory sequence (TRS)-dependent template switching. The 5′-UTR TRS-leader interacts with homologous TRS-body elements upstream of structural and accessory genes in the last third of the genome driven by the extent of TRSs' base-pairing, viral and host protein-RNA binding, and high-order RNA-RNA interactions [6,22–25].

Regarding viral enhancer elements, a long-range interaction spanning approximately 26 kilobases was identified in the *Alphacoronavirus* transmissible gastroenteritis virus (TGEV) that was required for efficient transcription of its subgenomic mRNA encoding the *N* gene, its most abundant subgenomic mRNA despite its less robust TRS-L-TRS-B interaction (Figure 1) [24,25]. This nonanucleotide TGEV enhancer is conserved among other *Alphacoronaviruses*, such as feline infectious peritonitis virus and canine enteric coronavirus. The TGEV enhancer via long-distance RNA-RNA interactions between complementary proximal and distal elements, brings TRS-L and

© 2024 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 June 2024 doi:10.20944/preprints202406.1845.v1

TRS-B upstream of the *N* gene closer for discontinuous transcription, resulting in higher *N* gene expression levels [24,25]. The coronaviral N protein, in turn, has been postulated to act as a transactivator of gene expression similar to the lentiviral Tat protein via interactions with the 5′-UTR leader sequence and with viral and host proteins in the replication-transcription complex [26].

A nonanucleotide enhancer element (UUUAUAAAC) also was characterized in MHV, a *Betacoronavirus* (subgenus *Embecovirus*) just distal to its TRS-L [27] (Figure 1B). However, the full mechanism underlying the enhancer activity has not been characterized.

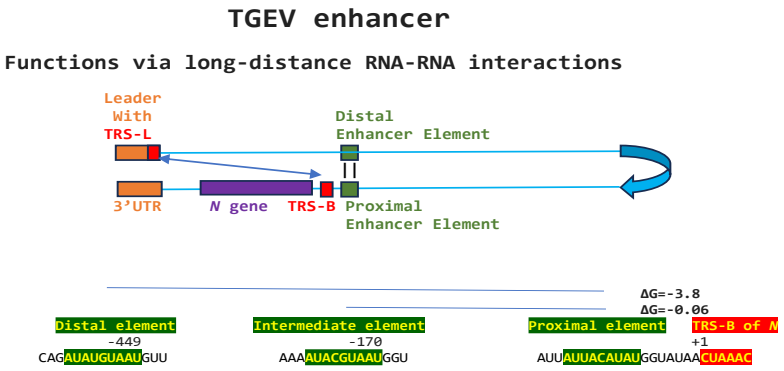


Figure 1. Nonanucleotide enhancer in transmissible gastroenteritis virus (TGEV, *Alphacoronavirus*, subgenus *Tegacovirus*, GenBank Accession: NC_038861). The TGEV enhancer upregulates the expression of the subgenomic RNA encoding the nucleocapsid (N) protein, possibly by approximating the TRS-L and TRS-B of N via duplex formation between the distal (close to the middle of the *M* gene) and proximal (7 nucleotides upstream of the N TRS-B) enhancer elements. Despite its sequence similarity to the distal element, the intermediate element shown does not contribute to enhancer activity consistently with its high minimum free energy.

The high prevalence of long-range RNA-RNA interactions in both coding and non-coding regions of coronaviruses and other RNA viruses, including insect nodaviruses [28] and plant tombusviruses [29,30], the phylogenetic conservation of the involved RNA motifs, and the considerable level of global organization [31] in replicative strategies in coronaviruses reinforces the relevance of these high-order structures for viral transcription of numerous distinct subgenomic mRNAs.

Long-distance RNA-RNA interactions contribute to RNA's three-dimensional structure complexity and gene expression regulation, translation, and resistance to degradation, among other functions [32]. Although genetic and biochemical analyses have confirmed the functional importance of many of these structures, their precise roles remain to be fully defined [33–35]. Understanding the noncoding functions of the viral RNA during viral replication provides new insights into virus-host interactions and therapeutic targets, which is the focus of this study predicting coronaviral enhancers and analyzing, in the case of IBV, their impact using published cases of viral attenuation or its reversal.

2. Results

2.1. TGEV Enhancer-Based Model for MHV Enhancer and Potentially Bovine Coronavirus

We used the TGEV-enhancer model to ascertain if the MHV enhancer could use a similar mechanism to that of TGEV in Figure 1. We searched for sequences similar to the MHV enhancer immediately distal to the 5-leader in the MHV genome and found one in the region encoding NSP14 in *ORF1ab* (Figure 2A). Therefore, the characterized 5′-UTR MHV enhancer element [27] could pair a complementary sequence in the region encoding NSP14 at the end of *ORF1ab* with a minimum free energy (ΔG) of -6.9 kcal/mol. This pairing would bring closer together the TRS-L and the TRS-B preceding the *ORF2* gene, encoding a phosphodiesterase that antagonizes host interferon (IFN)

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 June 2024 doi:10.20944/preprints202406.1845.v1

signaling via antagonism of the 2′,5′-oligoadenylate synthetase (OAS)–RNase L pathway [36–38]. NS2 is a critical determinant of MHV strain A59 liver tropism in C57BL/6 (B6) mice and is required for the virus to cause hepatitis [36,38].

Another *Embecovirus*, the bovine coronavirus, has a matching potential enhancer element to that of MHV, also immediately distal to the leader sequence (Figure 2B). As in MHV, said enhancer element could pair with more distal genomic sequences. Two distal sequences are in the *NSP3* gene, and a third is in *NSP12*, encoding the RNA-dependent RNA polymerase. The pairings involving the first *NSP3* sequence to the second more distal *NSP3* one and that in *NSP12* are more stable, i.e., have lower minimum free energy, than those involving the 5′-UTR sequence and would bring the TRS-L closer to the TRS-B preceding the segments encoding from the NSP13 helicase, NSP14 exoribonuclease, NSP15 endoribonuclease, to the NSP16 2′-O-methyltransferase, all involved in RNA genome replication and immune evasion and beyond to accessory and structural genes.

The proposed MHV and bovine coronavirus enhancer model is reminiscent of that described in the tomato bushy stunt virus, a positive-strand RNA tombusvirus [39]. In this tombusvirus, a 12-nucleotide sequence located ~1,000 nucleotides upstream from the initiation site of subgenomic mRNA2 synthesis is required for the accumulation of said RNA through a long-distance RNA-RNA base-pairing interaction with a sequence located just 5′ to the transcription initiation site [39].

Eight of the 9 nucleotides in the documented MHV and potential bovine coronavirus enhancer elements comprise a sequence that reads identically in the sense and antisense directions with complementary halves (Figure 2C).

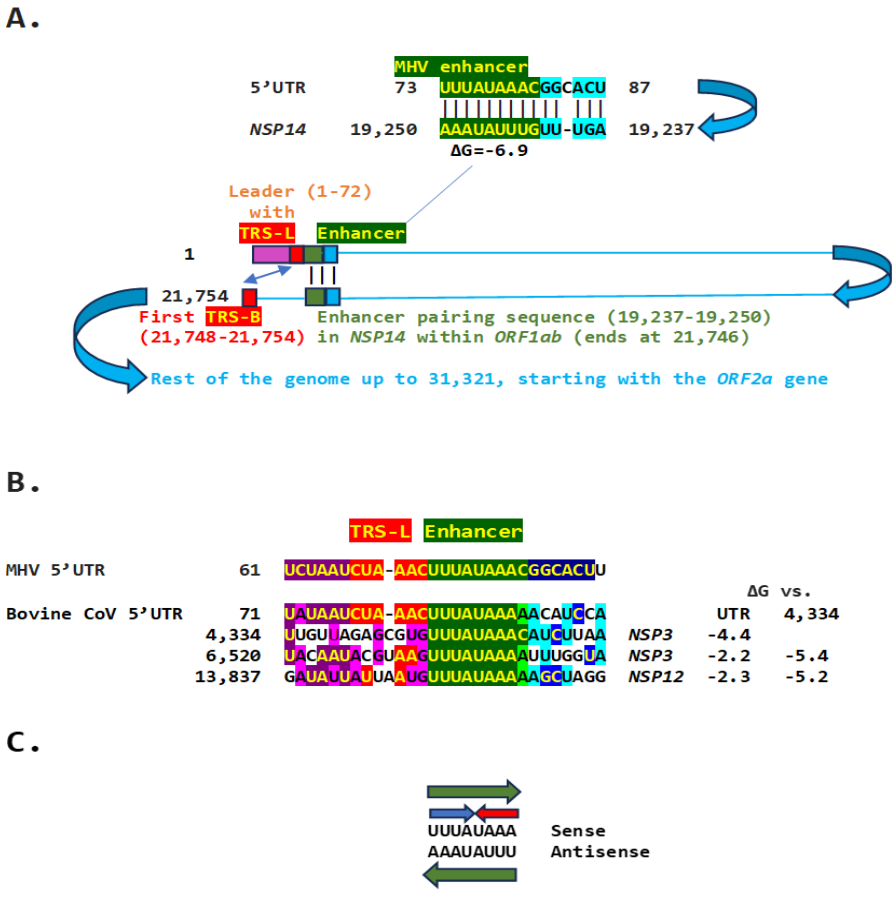


Figure 2. A. TGEV enhancer-based model for mechanism underlying MHV (*Betacoronavirus*, subgenus *Embecovirus*, GenBank Accession: NC_048217.1 GenBank Accession: NC_048217.1) enhancer function. Located immediately after the leader sequence, the MHV enhancer (green) could pair with a sequence towards the end of the genomic region encoding the ORF1b polyprotein. Because *ORF1ab* covers approximately two-thirds of the genome, the pairing would bring the TRS-L (red) closer to the first genomic TRS-B upstream of ORF2, potentially enhancing the transcription of its subgenomic RNA in a sequence-specific manner. B. Bovine CoV (*Betacoronavirus*, subgenus

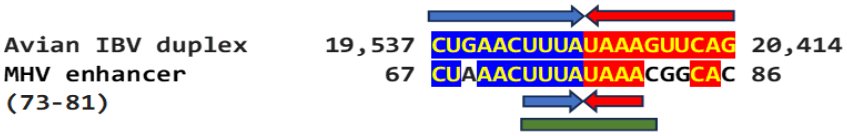
Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 June 2024 doi:10.20944/preprints202406.1845.v1

Embecovirus) has an MHV-like enhancer also immediately after the 5′-UTR leader sequence and similar sequences at three distal positions in the genome, which could pair with the 5′-UTR sequence. Minimum free energies (kcal/mol) are shown relative to the 5′-UTR sequence and the one located at position 4,434; the latter pairings are predicted to be more stable. Similarities with the leader, TRS-L, enhancer, and beyond are highlighted in purple, red, green, and blue, respectively C. Enhancer sequence shared between MHV and bovine CoV includes an octanucleotide reading the same in the sense and antisense strands (green arrows) and with complementary halves (blue and red arrows).

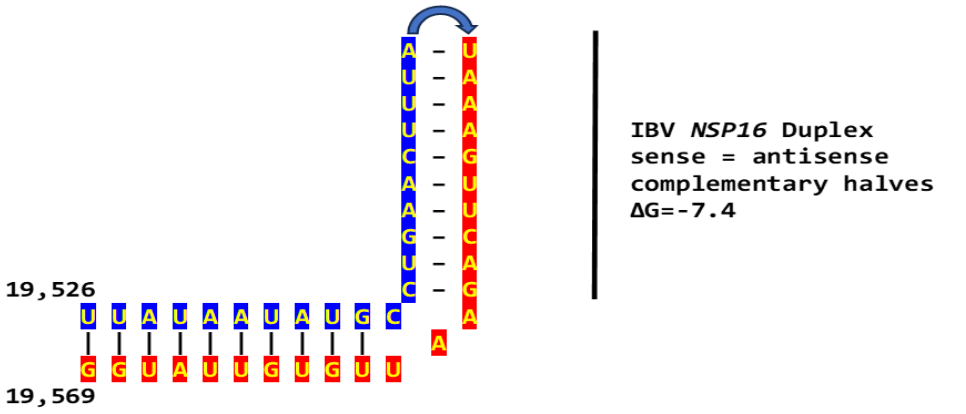
2.2. Potential Enhancer in the Infectious Bronchitis Virus (IBV)

We determined if other coronaviruses have a sequence similar in primary and secondary structure to the MHV enhancer (Figure 2C), reading the same in the sense and antisense strands with complementary halves. We found this sequence in the region encoding NSP16 of IBV (Figure 3A), a *Gammacoronavirus*, the most ancient genus among coronaviruses [40]. NSP16 and NSP7 block antigen presentation inhibiting adaptive immunity, and NSP16, together with NSP13, and NSP14, is involved as a 2′-O-methyltransferase in capping viral mRNA, which limits its recognition by the innate immune system [41]. The duplex-forming sequence reading the same in the sense and antisense direction with complementary halves that can form a duplex (ΔG=-7.4 kcal/mol) is part of an extended duplex-forming region (ΔG=-13.3 kcal/mol) (Figure 3B). Two similar subsequences (expect = 0.2) that could pair with the *NSP16* duplex are found proximal to it in the region encoding the first protein of ORF1a and distal to it in that encoding the spike (S) protein of IBV and preceding the *ORF3a* gene, which encodes an accessory protein involved in host immune evasion, specifically interferon resistance [42,43] (Figure 3C).

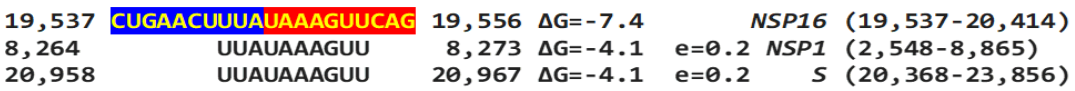
A.



B.



C.



Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 June 2024 doi:10.20944/preprints202406.1845.v1

Figure 3. A. Similarity between the *NSP16* duplex in avian infectious bronchitis virus (IBV; Gammacoronavirus, subgenus *Igacovirus*) and the MHV enhancer (green). Red and blue arrows indicate complementary halves that can form a duplex B. NSP16 duplex (with same sense and antisense sequences) and extended duplex. ΔG is the minimum free energy in kcal/mol. C. Similar distal sequences potentially pairing with the *NSP16* duplex. One sequence is proximal (region encoding first protein in ORF1a) and the other distal (region encoding spike [S]) to the duplex.

Avian IBV, including chicken, turkey, pigeon, goose, and swan coronaviruses, among others [44], is one of the major causes of highly contagious respiratory diseases in domestic fowl and a serious economic threat to the poultry industry worldwide as the second most dangerous disease after highly pathogenic influenza [42,44–46]. Live-attenuated vaccines against IBV have been generated by serial passage, from 51 or 100 passages, of a virulent isolate in chicken embryonated eggs until attenuation is achieved [46,47]. The exact mechanisms of attenuation are unknown, and vaccine strains have a risk of reversion to virulence.

Several studies have aimed at characterizing determinants of attenuation of IBV strains. For instance, the attenuated recombinant IBV M41-R was generated from the pathogenic strain MR41-CK by two amino acid changes, namely NSP10-Pro85Leu and NSP-14-Val393Leu, which were associated with a temperature-sensitive replication phenotype at 41°C in vitro [48]. Accessory genes have also been associated with attenuation in natural hosts [49–53].

Although in some studies changes in the S glycoprotein have been linked to IBV host adaptation leading to either increased or decreased pathogenicity [54,55], this has not been the case in others. For instance, even when the S glycoprotein ectodomain from the attenuated Beaudette laboratory strain was replaced with that from the pathogenic Mass41 strain, the strain remained nonpathogenic in chickens, indicating that the S glycoprotein is not the only determinant of IBV pathogenicity [45,46].

In another study, a chimeric IBV was created with the replicase genes 1a and 1ab from the attenuated Beaudette strain and all the structural genes from the pathogenic Mass 41 strain including the *S* gene. This chimeric virus was not pathogenic in chickens, indicating that the replicase proteins also appear to be determinants of the IBV pathotype [45,57]. Genetic differences reported in ORF1a and S between virulent and avirulent strains of IBV also led others to suggest that the replicase proteins, in addition to S, are involved in the pathotype of the virus [46].

Because enhancers in viruses and eukaryotes show variation with functional consequences [58], we analyzed for the presence of mutations in the extended *NSP16* duplex among avian IBV in the GenBank database and their effect on duplex minimum free energy and encoded amino acid sequence as potential attenuation determinants (Figure 4; GenBank accession numbers, geographical locations, and pathogenic characteristics are provided in Supplement 1). Some mutations did not affect the duplex minimum free energy while others decreased or increased it. Attenuation is unlikely to be driven by amino acid sequence changes because 91.8% of IBV strains in GenBank have a conserved amino acid sequence in the segment encoded by the *NSP16* extended duplex at the 5′ end of the reading frame (Figure 4).

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 June 2024 doi:10.20944/preprints202406.1845.v1

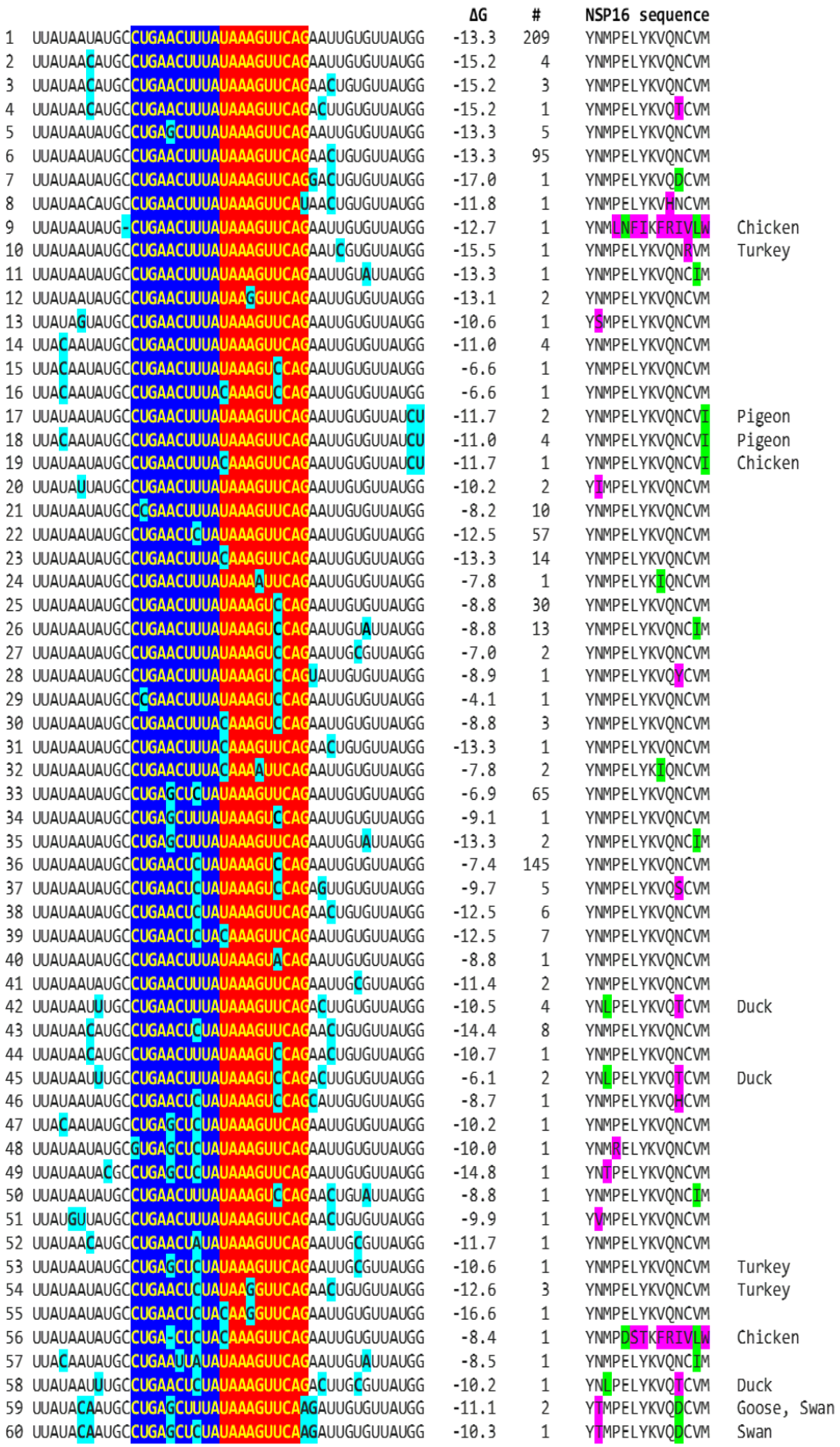


Figure 4. Mutations (highlighted in light blue) in avian IBV *NSP16* extended duplex (complementary halves in blue and red), duplex minimum free energy (ΔG), number of GenBank strains with each mutation combination (mutations in light blue), translated amino sequence (conservative substitutions in green, nonconservative ones in magenta), and avian IBV origin other than chicken.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 June 2024 doi:10.20944/preprints202406.1845.v1

2.2.1. Variation in NSP16 Duplex, Its Distal Binding Sequence in S, or Both Consistent with Reported Cases of Viral Attenuation or Its Reversal

Using published cases of viral attenuation or its reversal, we analyzed whether changes in the *NSP16* duplex, the distal binding sequence in *S*, or both would be consistent with the phenotypes described. Figure 5 provides examples of the three scenarios of such changes consistent with attenuation or its reversion documented in the literature. These findings may pinpoint the *NSP16* duplex as another potential pathogenicity determinant underlying attenuation or conversion to a pathogenic phenotype. However, these changes occurred in the context of others suggested to be associated with attenuation or its reversal.

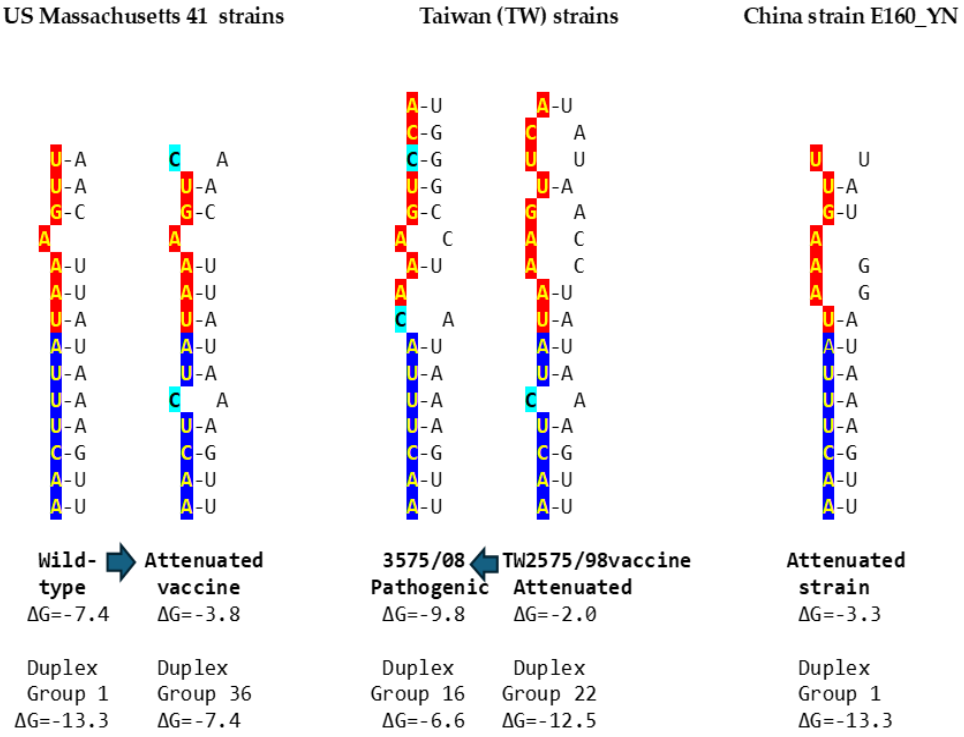


Figure 5. Examples of changes in the *NSP16* duplex, the distal binding sequence, or both, which may underlie published cases of viral attenuation (Massachusetts and China strains) or its reversal (Taiwan strains). The *NSP16* duplex is shown vertically in an open configuration with complementary halves in blue and red. Mutations are highlighted in light blue.

In the first example of attenuation of a pathogenic strain consistent with mutations in *NSP16* duplex (Figure 5), after serial passages in chicken embryos, the Mass41 attenuated vaccine strain (Mildvac-H) was derived from the Massachusetts strain Mass/Mass41/41 wild-type strain [59]. The Mass41 attenuated vaccine strain has two mutations in the extended *NSP16* duplex (duplex group 36, ΔG of -7.4 Kcal/mol, GenBank accession number GQ504725.1, Figure 4) relative to the wild-type strain (duplex group 1, ΔG of -13.3 Kcal/mol, GenBank accession number GQ504724.1). The sequence to which the open *NSP16* duplex could bind in S is conserved in both strains resulting in a less favorable minimum free energy of the pairing for the attenuated vaccine strain (ΔG = -3.8 Kcal/mol) than the wild-type strain (ΔG = -7.4 Kcal/mol). Expression of *ORF3a* would be predicted to be less or not enhanced in the attenuated strain, diminishing its interferon-resistance-mediated host immune evasion ability [43]. The amino acid sequence encoded by the extended duplex segment does not differ between strains (Figure 4), underscoring the nucleotide changes as potential pathogenic determinants.

Another publication provides the second example in Figure 5, namely of re-emergence of pathogenicity consistent with mutations in *NSP16* duplex and potential binding site in *S* in an attenuated vaccine strain. IBV 3575/08 (GenBank accession number KX266757), with strong respiratory and renal pathogenicity, was isolated from chicken broilers vaccinated with the

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 June 2024 doi:10.20944/preprints202406.1845.v1

attenuated viral vaccine derived from a Taiwan strain 2575/98 (GenBank accession number MN128087.1), with which it shares high similarity in structural proteins such as spike; however, the amino acid differences confer distinct antigenicity and low cross-protection [68]. Nonstructural proteins involved in host immune evasion, such as ORFs *3a*, *3b*, and *5* genes, are not highly similar between strains. In terms of the extended *NSP16* duplex, IBV 3575/08 has three mutations (duplex ΔG of -6.6 Kcal/mol; group 16 in Figure 4) while the attenuated parent strain 2575/98 has only one mutation (duplex ΔG = -12.5 Kcal/mol; group 22 in Figure 4). However, the free minimum energy (ΔG) of the interaction between the open *NSP16* duplex and *S* gene sequence is more favorable for the 3575/08 (ΔG = -9.8 Kcal/mol) than for the parent strain 2575/98 (ΔG = -2.0 Kcal/mol) (Figure 5). The predicted lower enhancer effect for expression of the *ORF 3a* and *3b* genes in the 2575/98 strain would be consistent with the observation that chickens infected with 3575/08 have delayed expression of a set of host's innate immune genes, i.e., better host immune evasion by the virus. In contrast, expression of host innate immune genes is quicker and more efficient in chickens infected with 2575/98, reflecting worse host immune evasion by the virus possibly secondary to lower or absent enhancer effect by the open *NSP16* duplex [68]. An alternative or additional explanation is that the decreased immune evasion capability of the 2575/98 strain relative to 3575/08 could be due to differences in the ORF 3a, 3b, and 5 protein sequences. Amino acid sequences in said nonstructural or other genes may also underlie the attenuation of the vaccine strain 2575/98 relative to the wild-type strain 2575/98 (GenBank accession number DQ646405.2) because both have the same *NSP16* extended duplex sequence (both in group 22, Figure 4) and possible binding sequence in S. All three strains have the same NSP16 amino acid sequence in the segment encoded by the extended duplex (Figure 4, two are in group 16 and one in group 22). Therefore, differences in pathogenicity are not related to variations in the NSP16 amino acid sequence in the region encoded by the extended duplex.

The third example in Figure 5 involves attenuation of pathogenic strain consistent with mutations in the potential binding site in *S*. IBV strain E160_YN (GenBank accession number MK644086.1) is a strain with an *NSP16* duplex corresponding to group 1 (ΔG = -13.3 Kcal/mol) but with mutations in the potential binding site in *S* that could reduce or obliterate enhancement of *ORF 3a* expression (low ΔG for pairing between *NSP16* core duplex and binding sequence in *S*). However, the *S* gene has a premature termination codon, and the virus does not encode ORF 5a because of an 81-nucleotide deletion [49,53], both suggested to contribute to the attenuation.

Other published cases of viral attenuation do not involve changes in the IBV *NSP16* extended duplex, such as the attenuated commercial vaccine strain TW2575/98vac derived from the wild strain TW2575/98w (both duplex group 1, Figure 4, Supplement 1)[47]; the attenuated Ark DPI 101 derived from the virulent Ark DPI 11 (both duplex group 36, Figure 4, Supplement 1) [46]; the less pathogenic Mass variant 15SK-02 relative to the more pathogenic 15AB-01 (both duplex group 1, Figure 4 Supplement 1 [60]; and the attenuated CK/CH/LDL/97I P115 strain derived from the CK/CH/LDL/97I P5 strain (both duplex group 36 with same binding sequences) [61]. These observations underscore the existence of multiple mechanisms for viral attenuation.

2.2.2. Further Analysis of NSP16 Duplex Variation and Its Possible Association with Viral Attenuation or Its Reversal

To further analyze the possible association between extended *NSP16* duplex variation and viral attenuation, we assessed the distribution of numbers of IBV strains in the GenBank database according to the ΔG of their extended *NSP16* duplex (Figure 6).

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 June 2024 doi:10.20944/preprints202406.1845.v1

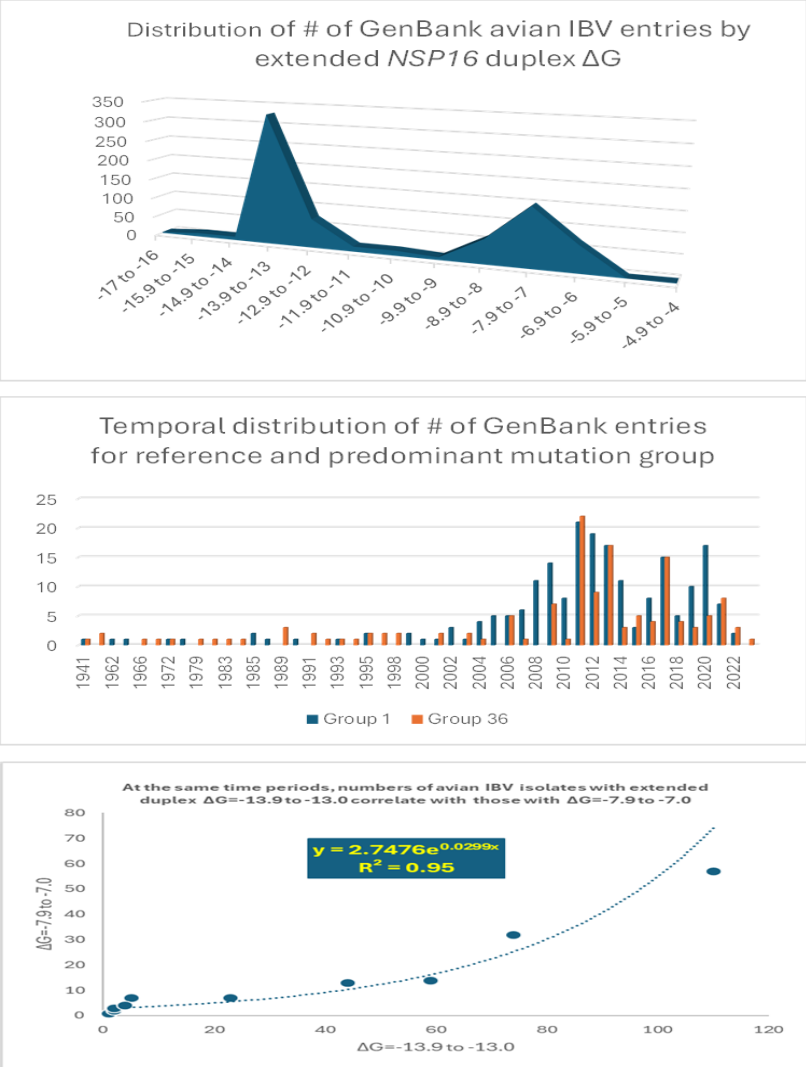


Figure 6. Analyses of possible association between extended *NSP16* duplex variation and viral attenuation or its reversal.

Overall, there is a dichotomization of the distribution of the number of GenBank entries into two clusters around the most prevalent values of minimum free energy corresponding to groups 1 (reference) and 36 in Figure 4 (Figure 6 top panel). The temporal distributions of numbers of groups 1 and 36 entries were similar (Figure 6 middle panel), with an exponential correlation favoring greater numbers of IBV strains with higher minimum free energy over time (Figure 6 bottom panel). This correlation is consistent with vaccination campaigns with attenuated strains that can spread among poultry and the emergence of novel pathogenic strains via mutations and recombination within and among wild-type and vaccine strains, including attenuated and revertants [49,61–92]. Some of these mutation and recombination events may allow IBV to evolve to infect other species, as demonstrated with primate cells in vitro [94].

The possible relationship between extended *NSP16* duplex ΔG and frequency of attenuated strains in the groups with the highest numbers of strains, i.e., those in the cluster peaks in Figure 6, was further analyzed (Figure 7). A Pearson's chi-square test of independence was performed to examine the relation between ΔG grouping and frequency of viral attenuation or vaccine derivation/relatedness. The relation between these variables was significant, with group 36 with the higher ΔG including ~54% of attenuated/vaccine-derived/vaccine revertant strains in contrast to ~1.5% in group 1 with the lower ΔG. The observation that a higher ΔG is present in half of the attenuated/vaccine-derived/vaccine revertant strains underscores the many variations unrelated to the *NSP16* duplex that may underlie attenuation. It is hard to accurately estimate IBV genetic

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 June 2024 doi:10.20944/preprints202406.1845.v1

diversity and mutation rates because attenuated vaccine strains can evolve in various ways and counteract an attenuation event, without altering its underlying mechanism, with another that causes reversal [63].

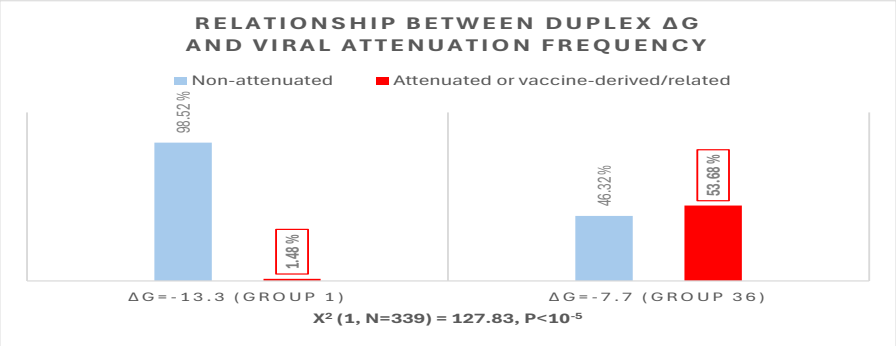


Figure 7. Relationship between extended *NSP16* duplex minimum free energy and frequency of viral attenuated/vaccine-derived/vaccine revertant strains.

2.2.3. Sequences Similar to the IBV NSP16 Extended Duplex are Present in the Rousettus Bat Betacoronaviruses (Nobecoviruses), Related to SARS-CoV-1/-2.

Using the BLAST program, we checked for sequences similar to the IBV *NSP16* extended duplex in *Viridae* and detected them only in *Rousettus* bat-coronaviruses (genus *Nobecoviruses*) (Figure 8), which are *Betacoronaviruses* related to SARS-CoV-1 and -2 and able to utilize the human ACE2 receptor for cell entry in vitro [95,96].

		ΔG	e
1	UUAUAAUAUGCCUGAACUUUAUAAAGUUCAGAAU--UGUGUUAUGG	-13.3	
2	UUACAAUAUGCCUGCACUUUAUAAAGUACAGAAUGCUGUGUUAGAA	-7.2	$2e^{-04}$
3	UUAUAAUAUGCCUGCACUUUAUAAAGUGCAAAAU	-8.2	0.008
4	UUAUAGUAUGCCUGCGCUUUAUAAAGUGCAGAAU	-10.8	0.1

Figure 8. Similarities of the IBV *NSP16* extended duplex with *Rousettus* bat beta-coronaviruses (*Nobecoviruses*), closely related to SARS-CoV-1 and -2, and can utilize the human ACE2 receptor *in vitro*. 1. Extended duplex in IBV (NC_001451). 2. OQ175246.1 (Bat CoV RlYN17 [*Rousettus leschenaultia*], China/Yunnan, 2016, isolate BtR1-BetaCoV/YN2016-Q319, toward end of ORF1ab); OQ175248.1 (Bat CoV RlYN17 [*Rousettus leschenaultia*], China/Yunnan, 2016, isolate BtR1-BetaCoV/YN2016-Q320, toward end of ORF1ab); OQ175341.1 (Bat CoV RlYN17 [*Rousettus leschenaultia*], China/Yunnan, 2017, isolate BtR1-BetaCoV/YN2017-Q321, toward end of ORF1ab starting at position 20,136). 3. MK492263.1 (*Rousettus* Bat CoV strain BtCoV92, *Cynopterus brachyotis*, Singapore, 2015). 4. OM219649.1 (Bat CoV GCCDC1, *Eonycteris spelaea*, Cambodia, 12/18,19/2010, isolate RK091); KU762332.1 (*Rousettus leschneaulti* Bat CoV isolate GCCDC1 356, China, 05/28/2014); NC_030886.1 (*Rousettus leschneaulti* Bat CoV isolate GCCDC1 356, China, 05/28/2014); KU762337.1 (*Rousettus leschneaulti* Bat CoV isolate GCCDC1 346, China, 05/28/2014); MT350598.1 (*Rousettus* bat CoV GCCDC1, *Eonycteris spelaea*, Singapore, 10/2016, beta-CoV, *Nobecovirus*); OQ175331.1 (Bat CoV EsYN16, *Eonycteris spelaea*, China/Yunnan, 2016, BtEs-13BetaCoV/YN2016-Q311); OQ175332.1 (Bat CoV EsYN17, *Eonycteris spelaea*, China/Yunnan, 2017, BtEs-13BetaCoV/YN2017-Q312); OQ175333.1 (Bat CoV EsYN17, *Eonycteris spelaea*, China/Yunnan, 2017, BtEs-13BetaCoV/YN2017-Q313); OQ175242.1 (Bat CoV EsYN17, *Eonycteris spelaea*, China/Yunnan, 2017, BtEs-13BetaCoV/YN2017-Q309). Duplex complementary halves are highlighted in blue and red. Differences in bat sequences relative to IBV extended duplex are highlighted in light blue. Nucleotides shared among sequences are highlighted in grey. Minimum free energy (ΔG) is shown for each duplex, as is the degree of similarity of duplexes relative to IBV reference duplex expressed as expect (e).

2.3. SARS-CoV-2 and Other Betacoronaviruses Infecting Humans

We then searched for sequences reading the same in the sense and antisense directions with complementary halves that could form a duplex in SARS-CoV-2, a *Betacoronavirus* (subgenus

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 June 2024 doi:10.20944/preprints202406.1845.v1

Sarbecovirus) and the most recent coronavirus infecting humans characterized. SARS-CoV-2 has a sequence with said characteristics in the region of the replicase gene encoding NSP3 (Figure 9A). The duplex-forming sequence is also part of an extended discontinuous duplex-forming region. SARS-CoV-1 has a similar duplex-forming region but as part of an extended continuous duplex (Figure 9B). These SARS-CoV regions share secondary structural similarity and a 10-nucleotide region of primary sequence similarity repeated twice with the IBV extended duplex (Figure 9C). These primary and secondary structural similarities underscore the nucleotide sequence over the encoded amino sequence of these regions as functionally relevant to potential enhancer function and pathogenicity.

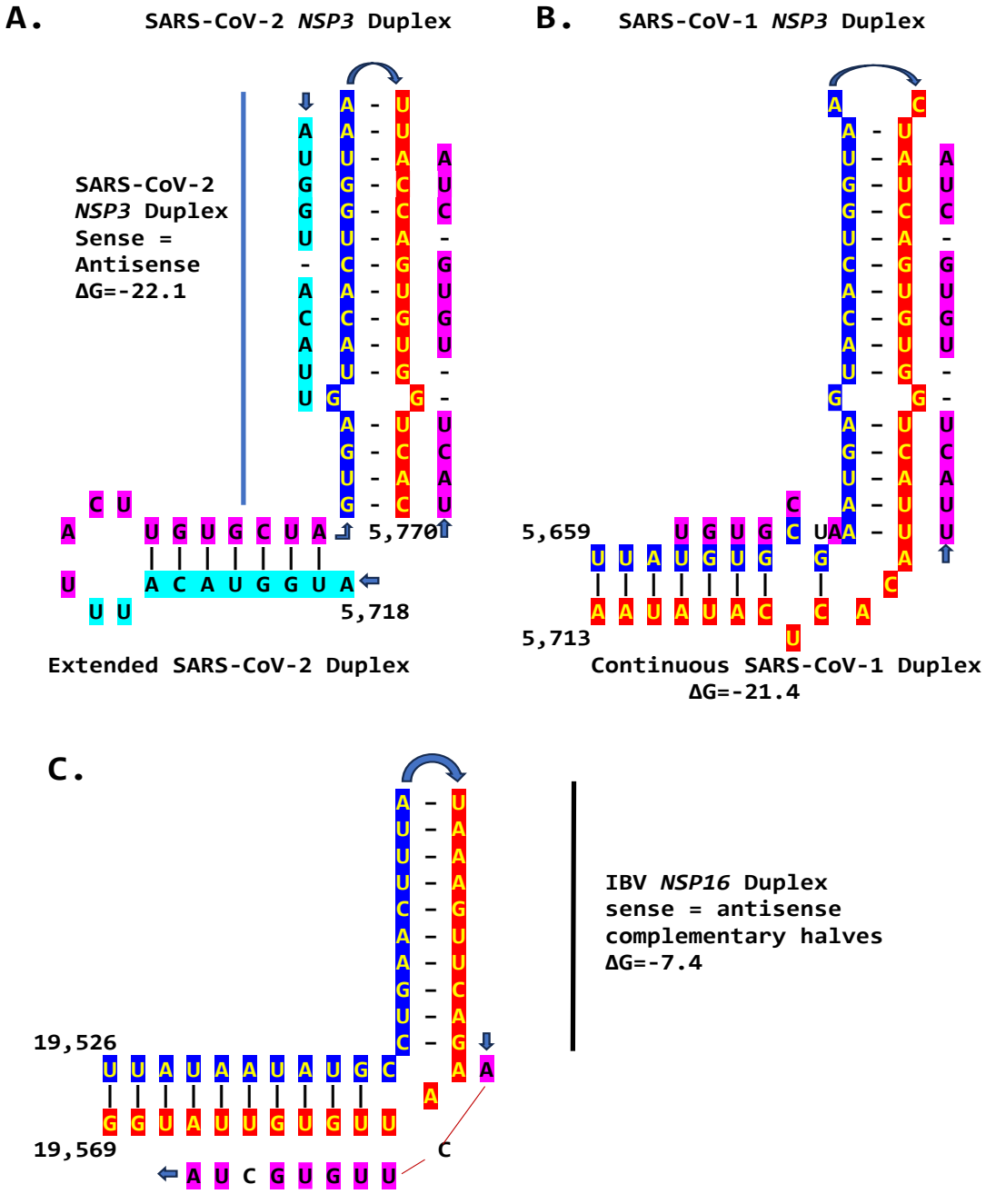
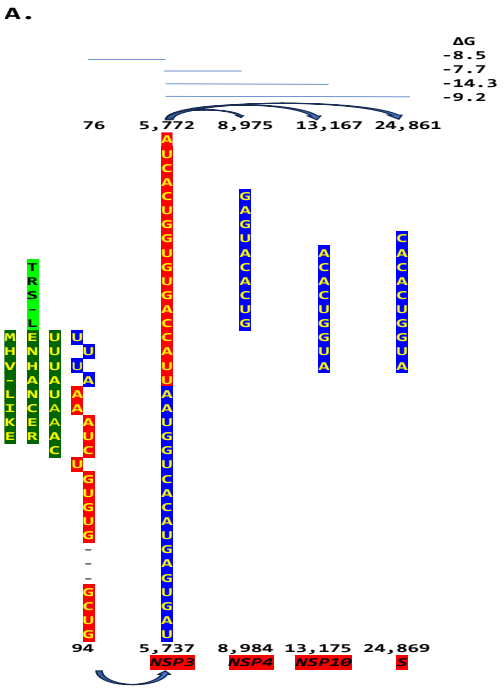
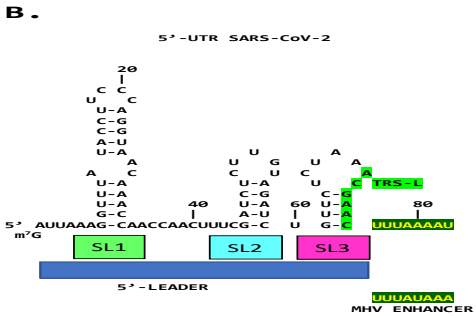


Figure 9. Duplex-forming sequences reading the same in sense and sense directions in the *NSP3* gene of SARS-CoV-2 (A) and SARS-CoV-1 (B), and in NSP16 of IBV. The complementary halves are highlighted in blue and red. Extended duplex regions are also shown. Regions of similarity within the SARS-CoV-2 extended duplex are highlighted in light blue with arrows indicating that they are in inverted orientations. Similar sequences within and among SARS-CoV-2 and SARS-CoV-1 and in IBV are highlighted in pink. Minimum free energies are shown for all duplexes.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 June 2024 doi:10.20944/preprints202406.1845.v1

We then searched in the SARS-CoV-2 genome for sequences proximal or distal to the core duplex-forming region and found several, including one immediately after the TRS-L reminiscent of the MHV enhancer, and others in *NSP4*, *NSP10*, and *S* (Figure 10A-C). As shown in Figure 10C, the first complementary half of the SARS-CoV-2 *NSP3* duplex-forming sequence shows similarities to the enhancer element proximal to the *N* gene in TGEV (Figure 1A).





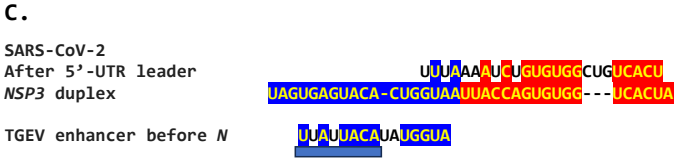


Figure 10. A. Region in SARS-CoV-2 5′-UTR with MHV-like enhancer and similarity to *NSP3* duplex, with which it can pair, leaving the duplex free to pair with three other distal genomic regions. The pairings involving the *S* gene (the first gene after *ORF1ab*; e = 2.4 for sequence similarity to the *NSP3* duplex) would approximate the TRS-L to the TRS-B of the gene encoding the viroporin *ORF3a*. The pairings involving *NSP4* and *NSP10* (e = 0.6 for both sequence comparisons to the *NSP3* duplex) would also decrease the distance between the TRS sequences. B. Sequence of the SARS-CoV-2 leader (blue box) that precedes the MHV-like enhancer (yellow letters with a dark green background) and that is added via discontinuous transcription to all accessory and structural genes. C. Similarities

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 June 2024 doi:10.20944/preprints202406.1845.v1

among SARS-CoV-2 5′-UTR sequence after leader, NSP3 duplex, and TGEV proximal enhancer element.

The SARS-CoV-2 NSP3 duplex, when not pairing with the 5′-UTR complementary sequence, can pair with a sequence in the accessory *ORF3a* gene with a minimum free energy similar to that of the within-duplex pairing (Figure 11A). The latter interaction may affect the expression of the envelope gene, encoding a viroporin [97], distal to *ORF3a* and *ORF3b*. Figure 11B depicts a model by which the *NSP3* duplex could open and interact with the complementary sequence in *ORF3a*, as it would with those in the other genomic locations shown in Figures 10 and 11C. Interactions with other potential viral or host nucleic acids and proteins could influence the closing and opening of the duplex.

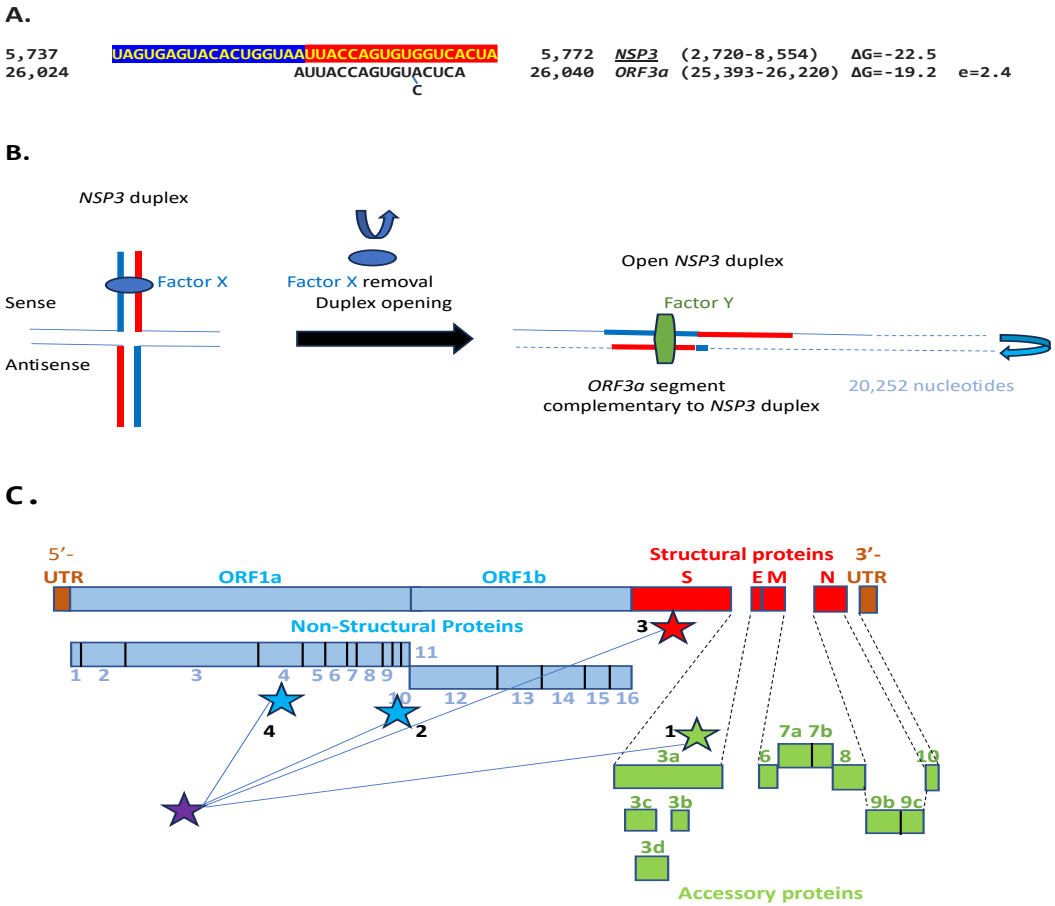


Figure 11. A. *NSP3* duplex and complementary sequence in the SARS-CoV-2 genome. The pair can form a duplex with a minimum free energy (ΔG) similar to the *NSP3* duplex. B. Switch model for the opening of the *NSP3* duplex and interaction with a 17-nucleotide complementary sequence in *ORF3a*. Initially, the duplex could be stabilized by RNA-RNA, RNA-protein, and protein-protein interactions involving undetermined factors, here termed X. Upon removal of said factors by epigenetic modification or interaction with other proteins or regulatory RNAs, the duplex could open and interact with other genomic sequences. In the illustration, the duplex and complementary sequences are 20,252 nucleotides apart, and their interaction is more stable in terms of ΔG, and length of sequences involved than those between TRS-L (6 nucleotides long) and TRS-Bs during discontinuous transcription of accessory and structural genes. A protein (named here Y or a combination of factors) could stabilize the new duplex between distant complementary sequences. C. Positions (depicted with stars) in the SARS-CoV-2 genome of the *NSP3* duplex (purple) and complementary sequences in Panel A and Figure 10. Numbers next to positions correspond from lowest to highest ΔG; e ranged from 0.6 to 2.4.

No mutations in the SARS-CoV-2 *NSP3* extended duplex were detected in a GISAID database (GISAID - Initiative) [98–100] search of all SARS-CoV-2 lineages, including Alpha (B.1.1.7 [Pango

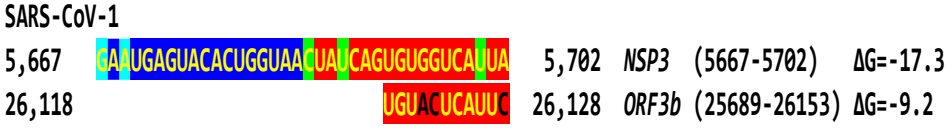
Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 June 2024 doi:10.20944/preprints202406.1845.v1

v.4.3.1 consensus call], Alpha [B.1.1.7-like] Scorpio, former VOC Alpha GRY [B.1.1.7+ Q.*]), Beta (B.1.351 + B.1.351.2 + B.1.351.3, former VOC Beta GH/501Y.V2), Gamma (P.1 + P.1.*, former VOC Gamma GR/501Y.V3), Delta (B.1.617.2, former VOC Delta GK [B.1.617.2 + AY.*]), Lambda (C.37 + C.37.1, former VOC Lambda GR/452Q.V1), Mu (B.1.621 + B.1.621.1, former VOI MuGH), Omicron (B.1.529 + BA.*, VOI GRA [JN.1 + JN.1.*]), BA.286 + BA.286.*, XBB.1.5 + XBB.1.5.*, XBB.1.16 + XBB.1.16.*, EG.5 + EG.5.*, BA.2.75 + BA.275.*, CH.1.1 + CH.1.1.*, XBB + XBB.*, XBB.1.9.1 + XBB.1.9.1.*, XBB.1.9.2 + XBB.1.9.2.*, XBB.2.3 + XBB.2.3.*, GH/490R (B.1.640 + B.1.640.*). The same applies to VOIs GRA XBB.1.5 + XBB.1.5.*, XBB.1.16 + XBB.1.16.*, EG.5 + EG.5.*, BA.2.86 + BA.2.86.* [excluding JN.1, JN.1.*], JN.1 + JN.1.*; and for VUMs GRA BA.2.75+ BA.2.75.*, CH.1.1 + CH.1.1.*, XBB + XBB.* [excluding XBB1.5, XBB1.16, XBB1.9.1], XBB 1.9.2 and XBB.2.3, XBB.1.9.1 + XBB.1.9.1.*, XBB.1.9.2 + XBB1.9.2.*, and XBB.2.3 + XBB.2.3.*.

The *NSP3* duplex-forming sequence, almost identical in the sense and antisense directions and with complementary halves, in SARS-CoV-2 is present in SARS-CoV-1 (another *Sarbecovirus*) with minor differences (Figure 5). In comparison against all *Viridae*, the 36-nucleotide *NSP3* sequence in SARS-CoV-2 is present only in the *NSP3*s of related bat Sarbecoviruses (Figure 12). The order of levels of similarity (expect, $1e^{-09}$ to $3e^{-06}$) matches that of minimum free energy (ΔG) of the duplex structure, with more favorable structures having lower ΔG. Sequences in SARS-CoV-2-related bat coronaviruses could be divided into four groups shown in Figure 12. The SARS-CoV-1 *NSP3* duplex is conserved among all SARS-CoV-1 strains in Gen Bank and related non-human coronaviruses, with a few exceptions that mostly do not affect duplex structure or render it more robust.

Figure 12. The SARS-CoV-2 *NSP3* duplex-forming 36-nucleotide sequence is present only in closely related bat *Sarbecoviruses* and SARS-CoV-1 among all *Viridae*. Nucleotide changes among all isolates are highlighted in light blue and those unique to SARS-CoV-1 are in light green. Sequences in SARS-CoV-2-related bat coronaviruses could be divided into four groups: *With an identical 36-nucleotide segment (ΔG = -22.5)*: BetaCoV_Yunnan_Rp_JCC9_2020 (OK287355.1); BANAL-20-236/Laos/2020 (MZ937003.2); BANAL-20-247/Laos/2020 (MZ937004.1); BANAL-20-116/Laos/2020 (MZ937002.1); BANAL-20-103/Laos/2020 (MZ937001.1); BANAL-20-52/Laos/2020 (MZ937000.1); RpYN06 strain bat/Yunnan/RpYN06/2020 (MZ081381.1); isolate PrC31 (MW703458.1). *With identical nucleotides 1-33 except nucleotide 3 (ΔG = -17.3)*: isolates RsHB20 BtRs-BetaCoV/HB2020-Q329 (OQ175349.1), Jingmen *Rhinolophus sinicus betacoronavirus* 1 (MZ328294.1), SC2018B (OK017846.1), and BM48-31/BGR/2008 (NC_014470.1). With identical nucleotides 4-33 (ΔG = -17.1): Horseshoe bat *Sarbecovirus* isolates Rt22QT53 (OR233321.1), Rt22QT48 (OR233320.1), Rt22QT46 (OR233319.1), Rt22QT36 (OR233318.1), Rt22QT178 (OR233317.1), Rt22QT161 (OR233316.1), Rt22QT124 (OR233300.1), Rt22QB8 (OR233299.1), and Rt22QB78 (OR233298.1); and isolates BtSY1 (OP963575.1), HN2021F (OK017835.1), and HN2021E (OK017834.1). With identical nucleotides 1-36 except nucleotides 3 and 7 (ΔG = -15.1): isolates GD2019E (OK017828.1), GD2019D (OK017827.1), GD2019B (OK017826.1), GD2019A (OK017825.1), GD2017W (OK017824.1), GD2017P (OK017822.1). SARS-CoV-1 sequence in Tor2 (NC_004718) and Urbani (MT308984) strains.

The extended *NSP3* duplex in SARS-CoV-1 could pair with a similar sequence (e = 0.15) in *ORF3b* instead of *ORF3a* as for SARS-CoV-2 (Figure 13).



Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 June 2024 doi:10.20944/preprints202406.1845.v1

Figure 13. A complementary segment (11 nucleotides, e=0.3) to the SARS-CoV-1 Tor2 (NC_004718.3) and Urbani (MT308984) strains *NSP3* duplex are similar to that in SARS-CoV-2. However, the minimum free energy is higher for the pairing between the switch duplex and the complementary sequence, rendering the pairing less stable.

SARS-CoV-1, including Tor2 (GenBank accession: NC_004718.3) and Urbani (MT308984) strains, has another duplex-forming sequence reading the same in the sense and antisense directions with complementary halves in the *ORF3a* gene that could pair with a complementary sequence in *NSP4* approximating the TRS-L and TRS-B of the *ORF3b* gene. The pairing would be less stable than that between the SARS-CoV-1 *NSP3* duplex and the *ORF3b* gene, an interferon antagonist, upstream of the envelope gene [101]. It appears that the activities of enhancer elements in SARS-CoV-2 would be stronger than those in SARS-CoV-1.

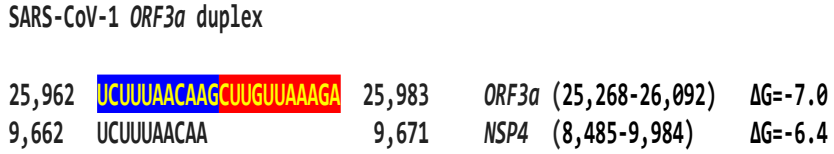


Figure 14. SARS-CoV-1 ORF3a duplex and complementary sequence in SARS-CoV-1 genome with a similar minimum free energy. In this case, the duplex switch structure is distal to the complementary sequence with which it can pair with a similar ΔG. However, the effect of reducing the distance between TRS-L and the TRS-B of the gene distal to the viroporin *ORF3a*, namely the viroporin *E*, is achieved.

Although SARS-CoV-2 does not have the duplex-forming potential core enhancer element in *ORF3a* present in SARS-CoV-1, the amino acid sequence encoded by the corresponding region is similar, with two conservative substitutions out of the three amino acid changes (Figure 15).

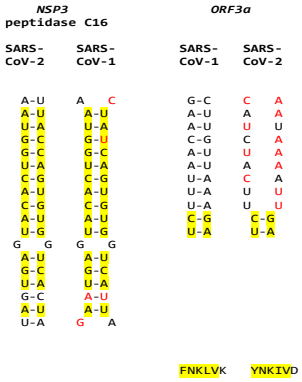


Figure 15. Comparison between the SARS-CoV-2 *NSP3* duplex and equivalent region in SARS-CoV-1, and between the SARS-CoV-1 *ORF3a* duplex and equivalent region in SARS-CoV-2. The SARS-CoV-2 *NSP3* duplex is relatively well conserved in SARS-CoV-2, while the SARS-CoV-1 *ORF3a* duplex is not conserved in SARS-CoV-2 (nucleotides differing between SARS-CoV-2 and -1 are shown in red). However, the encoded amino acid sequences are relatively well conserved (highlighted in yellow, as are similar nucleotides).

In terms of other *Betacoronaviruses*, the MERS-CoV (subgenus *Merbecovirus*), which like SARS-CoV-1 and -2 has caused epidemics in humans, has sequences in the NSP3-, NSP5-, and NSP6-encoding regions of the replicase gene (shown in Figure 16) matching those of the proximal (end of replicase gene), intermediate, and distal (near the TRS-B preceding the *N* gene) elements, respectively, of the *Alphacoronavirus* TGEV (shown in Figure 1). However, the minimum free energy for the duplex pairing of the distal and proximal sequences in MERS-CoV (ΔG of -1.3 vs. -3.8 kcal/mol in TGEV) is higher and may not be functional. The observation that all the MERS-CoV sequences similar to the TGEV enhancer elements are within the replicase gene would further argue against

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 June 2024 doi:10.20944/preprints202406.1845.v1

their functionality in altering the expression of structural and accessory genes in the last third of the viral genome.

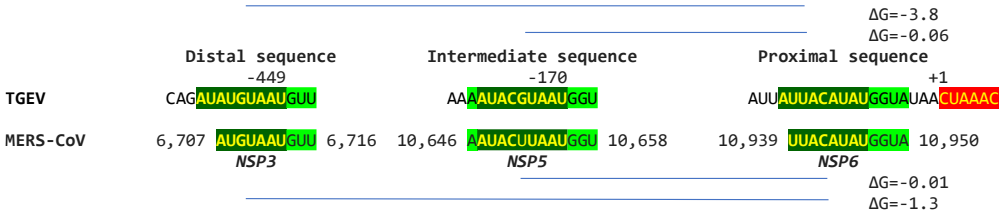


Figure 16. Similarity between the enhancer elements of TGEV and MERS-CoV sequences. The intermediate sequence is the one with the highest similarity, yet in TEGV, it does not contribute to the enhancer activity. The TEGV enhancer distal and proximal sequences are partly present. The minimum free energy of the distal-proximal element pairing is higher than that in TEGV, rendering this potential enhancer unlikely to be active in MERS-CoV.

The MERS-CoV NSP16-encoding region of the replicase gene harbors a duplex-forming sequence reading similarly (one nucleotide difference out of 26 positions) in the sense and antisense directions with complementary halves and as part of an extended duplex region that has regions of similarity with the *NSP3* duplexes of SARS-CoV-2 and -1 (Figure 17).

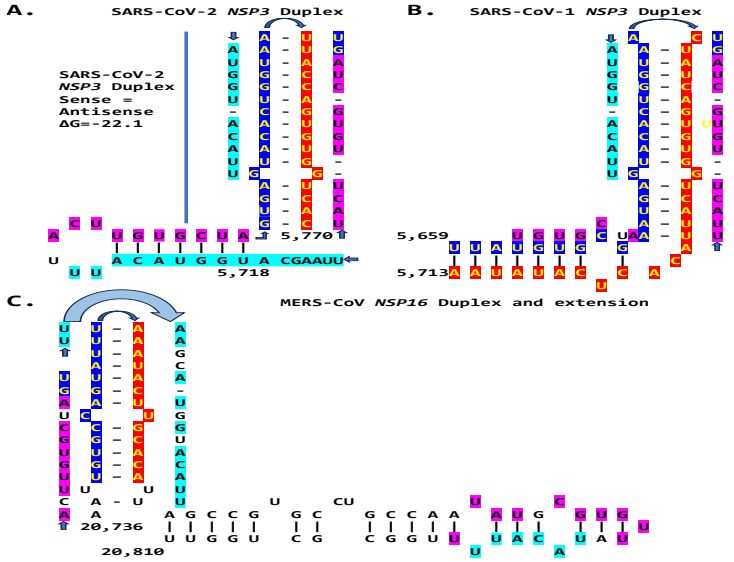


Figure 17. Comparison between NSP3 duplexes and extended duplexes of the Betacoronaviruses (subgenus *Sarbecovirus*) SARS-CoV-2 (A) and SARS-CoV-1 (B) with the NSP16 duplex and extended duplex in the *Betacorinavirus* (subgenus *Merbecovirus*) MERS-CoV. Regions of similarity are highlighted in light blue and pink, and arms of duplex reading similarly in the sense and antisense directions are highlighted in dark blue and red.

Another *Betacoronavirus* infecting humans, the hCoV-OC43 of the same *Embecovirus* subgenus as murine hepatitis virus and bovine coronavirus, has a duplex reading the same in the sense and antisense directions as part of an extended duplex in *NSP3* with sequence similarities to those in *NSP3* of SARS-CoV-2 and -1 (Figure 18).

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 June 2024 doi:10.20944/preprints202406.1845.v1

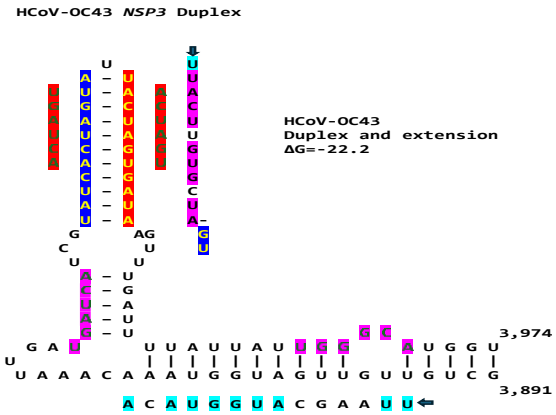
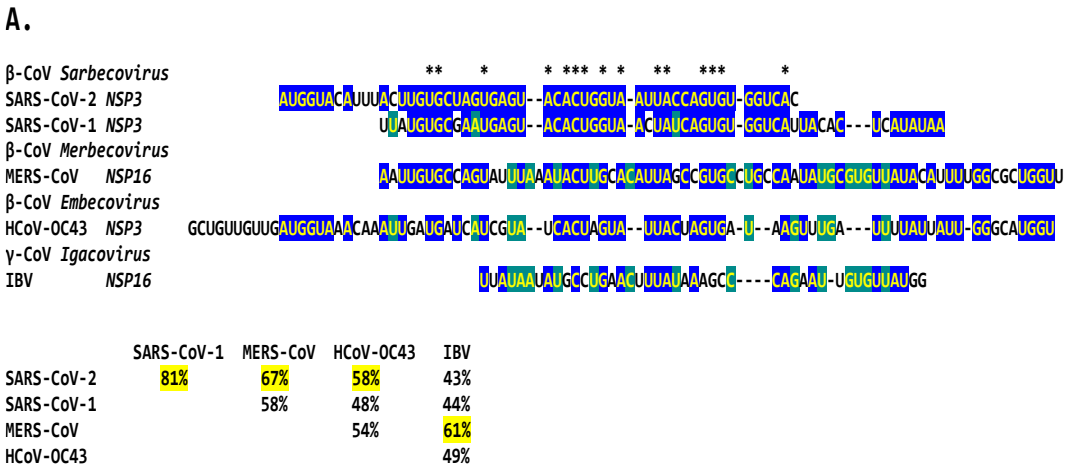
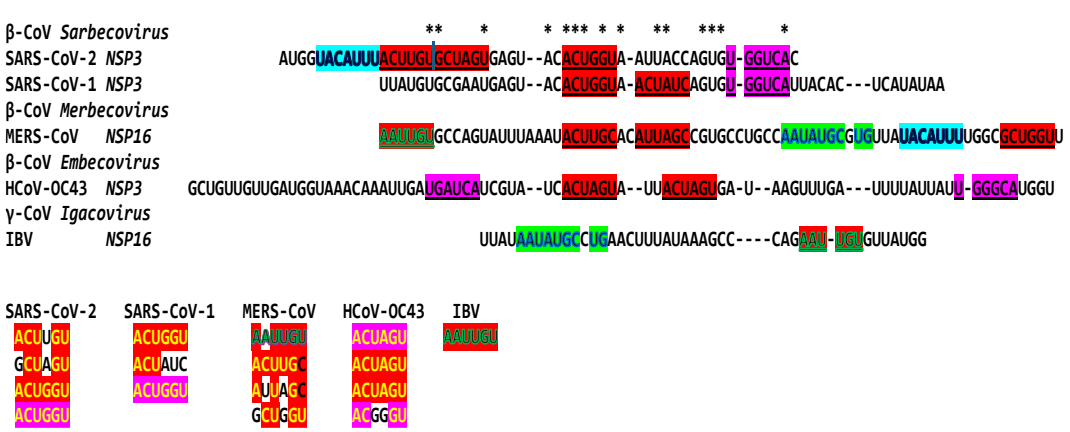


Figure 18. *NSP3* extended duplex in hCoV-OC43 (*Betacoronavirus*, subgenus *Embecovirus*). Regions of similarity with the NSP3 duplexes in SARS-CoV-2 and -1 and MERS-CoV are highlighted in light blue and pink, and arms of duplex reading the same in the sense and antisense directions are highlighted in dark blue and red. Sequences in green letters highlighted in red correspond to repeated sequences in the duplexes described in this paper and shown in Figure 19.

Finally, we analyzed similarities among all the duplexes described above for *Betacoronaviruses* infecting humans. The potential coronaviral duplex-forming enhancers contain repeats of hexanucleotide sequences, some with complementary halves (Figure 19), which might function as binding sites for host or viral proteins or RNAs, possibly regulating the closing and opening of the duplex sequences or bridging interaction with remote complementary sequences. Transcription factor binding drives enhancer activity and gene expression in various systems, and enhancers are usually intrinsically redundant in factors' binding sequences in either orientation [102].





Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 June 2024 doi:10.20944/preprints202406.1845.v1

Figure 19. Extended duplexes of the Betacoronaviruses infecting humans. A. Multiple alignments of the duplexes with similarity percentages. The highest duplex similarity was between SARS-CoV-2 and -1, while the MERS-CoV extended duplex sequence was closest to that of IBV (61% similarity), both in *NSP16* but also to the *NSP3* duplex of SARS-CoV-2 and -1 (67% and 58%). Asterisks denote nucleotides conserved among all duplexes, highlighted in either blue or green. B. Repeated sequences within duplexes and similarity among them. Repeated similar sequences within and among duplexes are highlighted in red, fuchsia (reverse of those in red), light blue (shared between SARS-CoV-2 and MERS-CoV), and green (shared between MERS-CoV and IBV), and green letters (shared between MERS-CoV and IBV). .

3. Discussion

In the present study, using the nonanucleotide-based enhanceosomes of *Alphacoronaviruses* and a *Betacoronavirus* and experience with other viral and host genome enhanceosomes, we detected potential enhancers, which vary in primary sequence and location, in the phylogenetically ancient *Gammacoronavirus* IBV and more recent *Betacoronaviruses* including SARS-CoV-2. The proposed enhancer model comprises an epistatic network with a core duplex-forming region molecular switch with identical sense and antisense sequences and complementary halves that could transition to an open configuration as dictated by a viral/host protein(s)/RNA(s), allowing robust pairing with proximal or distal remote viral genome complementary sequences to enhance specific viral gene expression.

Increased host immune evasion provides a replicative advantage to the virus. The coronaviral enhancer models proposed and reviewed here provide an alternative potential mechanism for viral attenuation, a process with multiple underlying mechanisms yet to be fully characterized. Analyses of sequenced IBV strains suggest disruptive effects of variation in the *NSP16* gene duplex-forming region, remote complementary sequences in the *Spike* gene upstream of the immune evasion-related *ORF3a*, or both, consistent with some documented cases of viral attenuation as potential novel pathogenicity determinants. The MHV and SARS-CoV-2 enhancers also are predicted to affect viral gene expression in host immune evasion. Although the duplex-forming sequence in SARS-CoV-2 appears thus far conserved among variants, its variability may adversely affect virulence. Based on the IBV experience, it may be early to detect variation in SARS-CoV-2. The same applies to SARS-CoV-1, where scarce cases of mutations do not adversely affect the minimum free energy of the duplex.

The features of the potential coronavirus enhancers are similar to those of enhancers overall. Enhancers, as cis-acting key scaffolds for the transient dynamic recruitment and assembly of transcription factor/coactivator clusters, integrate regulatory information encoded by the surrounding genome and biophysical properties of trans-acting transcription pieces of machinery, such as RNA polymerase and transcriptional coregulators [103–106]. All the coronaviral duplex-forming structures described here have repeats of similar motifs in different orientations, possibly mediating binding to viral and host factors involved in the potential enhancer function consistent with the redundancy of transcription-factor-interacting sequences within enhancers.

Enhancers control gene expression location, level, and timing [58,107,108]. In mammalian systems, enhancers determine spatiotemporal gene expression programs by engaging distant promoters over long genomic ranges [109–113]. For instance, some enhancer-promoter RNA interaction sites involve pairwise interacting Alu and non-Alu RNA sequences that tend to be complementary and potentially form duplexes [109].

Beyond the potential coronaviral enhancers analyzed here, we annotated duplex-forming regions reading similarly in the sense and antisense directions with complementary halves in MHV and bovine coronavirus, up to 83 nucleotides long (Figure 20). The latter two viruses illustrate the presence of possible networks of different potential enhancer elements with more than one long-range RNA-RNA interaction. They are reminiscent of the nested epistasis enhancer networks for robust genome regulation reported in mammalian genomes [114].

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 June 2024 doi:10.20944/preprints202406.1845.v1

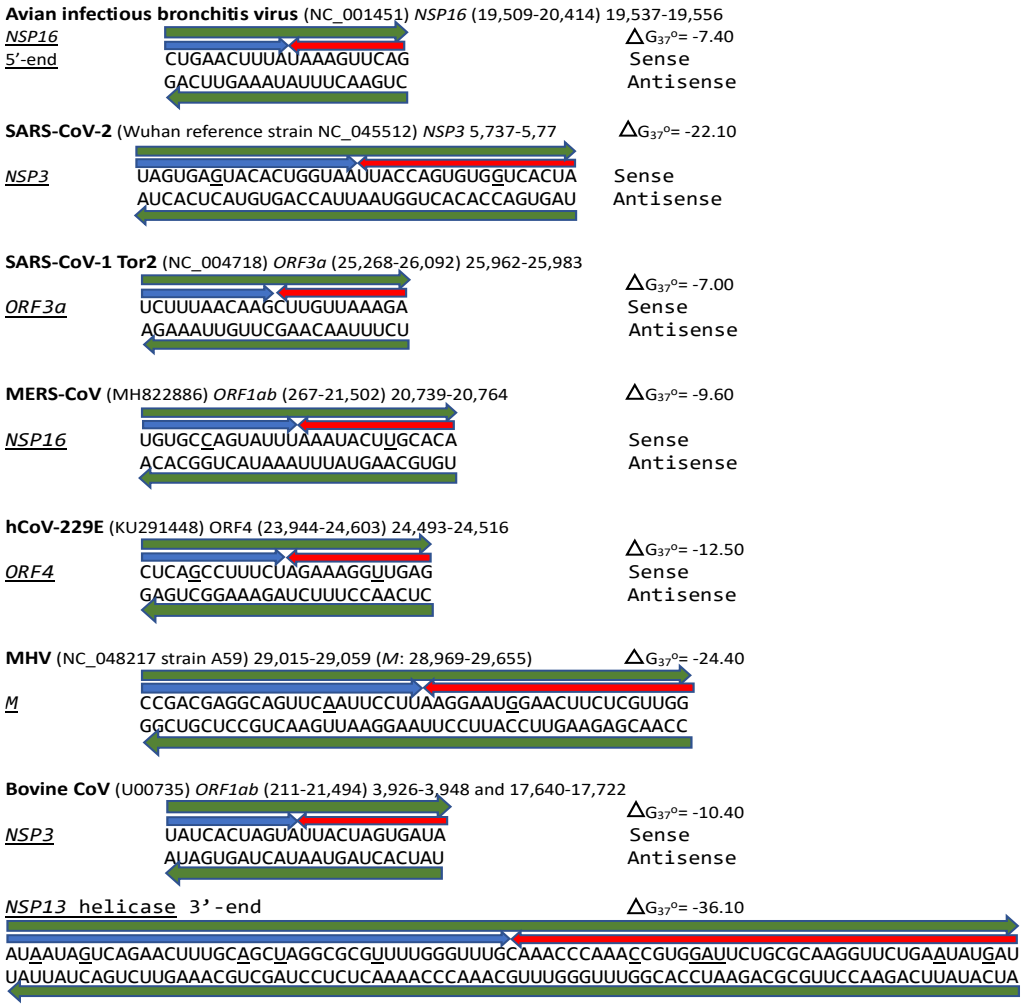


Figure 20. Annotated coronaviral duplex-forming sequences that read similarly in the sense and antisense directions with complementary halves. Minimum free energies (ΔG) are shown next to each hairpin. GenBank accession numbers are shown in parentheses. Positions that do not read the same in the sense and antisense directions are underlined. The MERS-CoV *NSP16* duplex is present in bat *Merbecoviruses* and pangolin CoV HKU4 (GenBank OM009282.1). Most of the MHV *M* duplex is in the bovine coronavirus M gene (OP866729.1), and the one in the bovine CoV *NSP13* helicase is in the canine respiratory coronavirus (ON133844.1), also an *Embecovirus*. The sequences annotated here do not comprise an exhaustive list, and other coronaviruses that share some of these duplexes were mentioned in the text.

In terms of the origin of the possible origin of the duplex-forming regions reading similarly in the sense and antisense directions with complementary halves (inverted repeats) and potential enhancer function, one could propose a template-switching mutation mechanism during RNA replication via RNA-dependent RNA polymerase. This mechanism would be similar to that postulated by Mönttinen et al. [115], in which a template-switching mutation mechanism during DNA replication by DNA-dependent DNA polymerase could generate hairpin structures via inverted repeats or inverted and direct repeats that could evolve into novel microRNAs [115]. The observation by Mönttinen et al. [115] of template-switching mutation-driven evolution at the nucleic acid level occurring without affecting the protein sequence encoded is consistent, for instance, with the observation here that although the duplex in SARS-CoV-1 *ORF3a* lacks a homolog in SARS-CoV-2, the encoded peptide sequences are similar. The same applies to variation in the potential IBV enhancer element at the nucleotide level, with the amino acid sequences encoded by the duplex region being conserved among strains.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 June 2024 doi:10.20944/preprints202406.1845.v1

The observation that the sense and antisense strands have the same or almost the same sequence in the duplexes described also raises the possibility that the enhancer activity may extend to gene expression by the antisense strand. As is the case for the transcriptomes of cytomegaloviruses [116], retroviruses [117–119], and prokaryotes [120,121], that of SARS-CoV-2 might include RNAs that are transcribed from the negative-sense genomic RNA and encode functional proteins or nucleic acids involved in RNA regulation [122]. For instance, HTLV-1 antisense strand-encoded mRNA interacts with the promoter and enhances transcription of the C-C chemokine receptor type 4 (*CCR4*) gene to support the proliferation of HTLV-1-infected cells, and HIV-1 antisense mRNA is recruited to the viral long terminal repeat and inhibits sense mRNA expression to maintain the latency of HIV-1 infection [123].

The potential enhancer elements and models discussed here require direct experimental validation. However, delineating possible long-range RNA-RNA interactions can enrich prediction tools and analysis of RNA folding and potential for novel structures, which are challenges for in silico prediction tools, none of which can offer full accuracy [32,124].

The experience with IBV strains provides fertile ground for these explorations illustrating the well-documented intricacies and limitations of ascribing viral attenuation to a particular mechanism given the diversity and frequency of mutations and recombination events among an ever-growing number of variants. However, the characterization of all possible attenuation and enhancement mechanisms of viral pathogenesis bodes well for the rational development of preventive and therapeutic strategies for all coronaviruses. Similar to using combination therapies against retroviruses [122], the combination of attenuation mechanisms may allow the development of more effective coronaviral vaccines [125–127].

4. Materials and Methods

4.1. Detection of Duplex-Forming Genomic Sequences Reading Similarly in the Sense and Antisense Directions with Complementary Halves as Potential Core Enhancer Elements

We compared coronaviral genomic sequences against themselves using the BLASTN program (nucleotide collection [nr/nt]; expect threshold: 0.5; mismatch scores: 2, -3; gap costs: linear) [128], which evaluates sense and antisense strands, to determine the presence of genomic sequences reading the same, or almost the same, in the sense and antisense directions. These sequences have complementary halves because they are inverted repeats. The representative reference sequences (RefSeq) for each coronaviral genus were obtained from the International Committee on Taxonomy of Viruses (ICTV) taxonomy Coronaviridae Study Group [129], and the GenBank and NCBI Virus databases.

4.2. Detection of Coronaviral Genomic Sequences that Could Pair to a Core Enhancer Element

We used the BLASTN program (nucleotide collection [nr/nt]; expect threshold: 10; mismatch scores: 2, -3; gap costs: linear) [128] to determine if the MHV enhancer nonanucleotide could pair with genomic sequences, following TGEV's enhancer model to affect subgenomic mRNA expression. The locations of the detected genomic sequences, and the boundaries of nonstructural, structural, and accessory open reading frames were determined based on GenBank annotation and manual inspection of multiple alignments and sequence similarities. The same approach was used with all potential core enhancer elements that, similar to the MHV enhancer, consisted of a duplex-forming sequence reading similarly in the sense and antisense directions with complementary halves.

4.3. Characterization of Extended Duplexes and Visualization of RNA Secondary Structures and Estimation of Minimum Free Energies of Duplexes

Sequences of up to 50 nucleotides before and after the core enhancer duplexes were analyzed for the possibility of their forming duplexes beyond the core duplex reading the same or similarly in the sense and antisense directions. RNA secondary structures were visualized using forna, a force directed graph layout (ViennaRNA Web services) [130]. Optimal secondary structures were also

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 June 2024 doi:10.20944/preprints202406.1845.v1

visualized using the RNAfold webserver, which was used to estimate the minimum free energy of core and extended duplexes and their pairing to other genomic sequences [131,132].

4.4. Determination of Similarities between Coronaviral Extended Duplexes with Viridae Sequences in GenBank and Analysis of Possible Nucleotide and Amino Acid Sequence Variations Being Consistent with Attenuation

The nucleotide sequences of the reference extended duplexes were compared against those entered under *Viridae* in the GenBank database. This allowed us to detect similarities between IBV or SARS-CoV-2 duplex-forming potential core enhancer elements and sequences in bat coronaviruses related to SARS-coronaviruses. It also allowed the detection of primary sequence variations in the extended duplex among IBV strains in the Gen Bank database. The Supplementary section includes the accession numbers, collection site, and date of the IBV strains.

The nucleotide sequences of the extended duplex of IBV variants were translated using the insilico (DNA to protein translation (ehu.es) [133] and Expasy (ExPASy - Translate tool [134]) tools to determine if amino acid sequence variation was consistent with viral attenuation.

The GISAID database (GISAID - Initiative) [98–100] was used to search for nucleotide variation in the *NSP3* SARS-CoV-2 extended duplex region of all SARS-CoV-2 lineages.

4.5. Assessment of Sequence Similarities in Multiple Alignments of Potential Coronaviral Extended Duplex Core Enhancer Elements

Multiple potential coronaviral extended duplex core enhancer elements alignments were evaluated using CLUSTAL Omega (https://www.ebi.ac.uk/services) [135] to search for similarities, including short repeats in different orientations, which were confirmed by visual inspection.

4.5. Assessment of Attenuation Status of IBV Strains

To analyze the relationship between IBV duplex variation and viral attenuation, we checked publications listed in GenBank for the nucleotide sequences included and publications listing attenuated vaccine strains or strains derived from them (Supplement 1). When no information could be obtained on attenuation status, strains were categorized as non-attenuated. The analysis here focused on the strains in groups 1 and 36 in Figure 4 which were the most representative in terms of the number of entries of the two clusters apparent by the ΔG distribution of the *NSP16* duplexes (as shown in Figure 5).

Author Contributions: Both authors (R.P. and W.A.H.) contributed to conceptualization; methodology; validation; formal analysis; investigation; data curation; writing—original draft preparation; writing—review and editing; and visualization. Both authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: All new data created, and datasets analyzed are specified in the manuscript text.

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

References

1. Gorbalenya, A.E., Enjuanes, L., Ziebuhr, J., Snijder, E.J. Nidovirales: evolving the largest RNA virus genome. *Virus Res* **2006**, *117*, 17-37. https://doi.org/10.1016/j.virusres.2006.01.017
2. Lauber, C., Goeman, J.J., Parquet Mdel, C., Nga, P.T., Snijder, E.J., Morita, K., Gorbalenya, A.E. The footprint of genome architecture in the largest genome expansion in RNA viruses. *PLoS Pathog* **2013**, *9*, e1003500. https://doi.org/10.1371/journal.ppat.1003500
3. Masters, P.S. The molecular biology of coronaviruses. *Adv Virus Res* **2006**, *66*, 193-292. https://doi.org/10.1016/S0065-3527(06)66005-3
4. Forni, D., Cagliani, R., Clerici, M., Sironi, M. Molecular Evolution of Human Coronavirus Genomes. *Trends Microbiol* **2017**, *25*, 35-48. https://doi.org/10.1016/j.tim.2016.09.001
5. Ziv, O., Gabryelska, M.M., Lun, A.T.L., Gebert, L.F.R., Sheu-Gruttadauria, J., Meredith, L.W., Liu, Z.Y., Kwok, C.K., Qin, C.F., MacRae, I.J., Goodfellow, I., Marioni, J.C., Kudla, G., Miska, E.A. COMRADES

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 June 2024 doi:10.20944/preprints202406.1845.v1

determines in vivo RNA structures and interactions. *Nat Methods* **2018**, *15*, 785-788. https://doi.org/10.1038/s41592-018-0121-0

6. Chen, S.C., Olsthoorn, RCL. Group-specific structural features of the 5′-proximal sequences of coronavirus genomic RNAs. *Virology* **2010**, 401, 29-41. https://doi.org/10.1016/j.virol.2010.02.007
7. Wang, D., Jiang, A., Feng, J., Li, G., Guo, D., Sajid, M., Wu, K., Zhang, Q., Ponty, Y., Will, S., Liu, F., Yu, X., Li, S., Liu, Q., Yang, X.L., Guo, M., Li, X., Chen, M., Shi, Z.L., Lan, K., Chen, Y., Zhou, Y. The SARS-CoV-2 subgenome landscape and its novel regulatory features. *Molecular cell* **2021**, 81, 2135–2147.e5. https://doi.org/10.1016/j.molcel.2021.02.036
8. Zúñiga, S., Cruz, J.L., Sola, I., Mateos-Gómez P.A., Palacio, L., Enjuanes, L. Coronavirus nucleocapsid protein facilitates template switching and is required for efficient transcription. *Journal of virology* **2010**, *84*, 2169-2175. https://doi.org/10.1128/JVI.02011-09
9. Li, L., Kang, H., Liu, P., Makkinje, N., Williamson, S.T., Leibowitz, J.L., Giedroc, D.P. Structural lability in stem-loop 1 drives a 5′ UTR-3′ UTR interaction in coronavirus replication. *J Mol Biol* **2008**, *377*, 790-803. https://doi.org/10.1016/j.jmb.2008.01.068
10. Miao, Z., Tidu, A., Eriani, G., Martin, F. Secondary structure of the SARS-CoV-2 5′-UTR. *RNA biology* **2021**, *18*, 447-456. https://doi.org/10.1080/15476286.2020.1814556
11. Hahn, C.S., Hahn, Y.S., Rice, C.M., Lee, E., Dalgarno, L., Strauss, E.G., Strauss, J.H. Conserved elements in the 3′ untranslated region of flavivirus RNAs and potential cyclization sequences. *J Mol Biol* **1987**, *198*, 33-41. https://doi.org/10.1016/0022-2836(87)90455-4
12. Lo, C.Y., Tsai, T.L., Lin, C.N., Lin, C.H., Wu, H.Y. Interaction of coronavirus nucleocapsid protein with the 5′- and 3′-ends of the coronavirus genome is involved in genome circularization and negative-strand RNA synthesis. *The FEBS journal* **2019**, *286*, 3222-3239. https://doi.org/10.1111/febs.14863
13. Goebel, S.J., Hsue, B., Dombrowski, T.F., Masters, P.S. Characterization of the RNA components of a putative molecular switch in the 3′ untranslated region of the murine coronavirus genome. *J Virol* **2004**, *78*, 669-682. https://doi.org/10.1128/jvi.78.2.669-682.2004
14. Xue, X., Yang, H., Shen, W., Zhao, Q., Li, J., Yang, K., Chen, C., Jin, Y., Bartlam, M., Rao, Z. Production of authentic SARS-CoV M(pro) with enhanced activity: application as a novel tag-cleavage endopeptidase for protein overproduction. *Journal of molecular biology* **2007**, *366*, 965-975. https://doi.org/10.1016/j.jmb.2006.11.073
15. Jonassen, C.M., Jonassen, T.O., Grinde, B. A common RNA motif in the 3′ end of the genomes of astroviruses, avian infectious bronchitis virus and an equine rhinovirus. *The Journal of general virology* **1998**, *79*, 715-718. https://doi.org/10.1099/0022-1317-79-4-715
16. Robertson, M.P., Igel, H., Baertsch, R., Haussler, D., Ares, M. Jr., Scott, W.G. The structure of a rigorously conserved RNA element within the SARS virus genome. *PLoS biology* **2005**, *3*, e5. https://doi.org/10.1371/journal.pbio.0030005
17. Spagnolo, J.F., Hogue, B.G. Host protein interactions with the 3′ end of bovine coronavirus RNA and the requirement of the poly(A) tail for coronavirus defective genome replication. *Journal of virology* **2000**, *74*, 5053-5065. https://doi.org/10.1128/jvi.74.11.5053-5065.2000
18. Tarun, S.Z. Jr., Wells, S.E., Deardorff, J.A., Sachs, A.B. Translation initiation factor eIF4G mediates in vitro poly(A) tail-dependent translation. *Proceedings of the National Academy of Sciences of the United States of America* **1997**, *94*, 9046-9051. https://doi.org/10.1073/pnas.94.17.9046
19. Viswanathan, T., Arya, S., Chan, S.H., Qi, S., Dai, N., Misra, A., Park, J.G., Oladunni, F., Kovalskyy, D., Hromas, R.A., Martinez-Sobrido, L., Gupta, Y.K. Structural basis of RNA cap modification by SARS-CoV-2. *Nat Commun* **2020**, *11*, 3718. https://doi.org/10.1038/s41467-020-17496-8
20. Thiel, V., Ivanov, K.A., Putics, Á., Hertzig, T., Schelle, B., Bayer, S., Weißbrich, B., Snijder, E.J., Rabenau, H., Doerr, H.W., Gorbalenya, A.E., Ziebuhr, J.Á. Mechanisms and enzymes involved in SARS coronavirus genome expression. *The Journal of general virology* **2003**, *84*, 2305-2315. https://doi.org/10.1099/vir.0.19424-0
21. Lai, M.M. Cellular factors in the transcription and replication of viral RNA genomes: a parallel to DNA-dependent RNA transcription. *Virology* **1998**, *244*, 1-12. https://doi.org/10.1006/viro.1998.9098
22. Bentley, K., Keep, S.M., Armesto, M., Britton, P. Identification of a noncanonically transcribed subgenomic mRNA of infectious bronchitis virus and other gammacoronaviruses. *Journal of virology* **2013**, *87*, 2128-2136. https://doi.org/10.1128/JVI.02967-12
23. Wang, D., Jiang, A., Feng, J., Li, G., Guo, D., Sajid, M., Wu, K., Zhang, Q., Ponty, Y., Will, S., Liu, F., Yu, X., Li, S., Liu, Q., Yang, X.L., Guo, M., Li, X., Chen, M., Shi, Z.L., Lan, K., Chen, Y., Zhou, Y. The SARS-CoV-2 subgenome landscape and its novel regulatory features. *Mol Cell* **2021**, *81*, 2135-2147.e5. https://doi.org/10.1016/j.molcel.2021.02.036
24. Moreno, J.L., Zúñiga, S., Enjuanes, L., Sola, I. Identification of a coronavirus transcription enhancer. *J Virol* **2008**, *82*, 3882-3893. http://doi.org/10.1128/JVI.02622-07
25. Mateos-Gómez, P.A., Morales, L., Zuñiga, S., Enjuanes, L., Sola, I. Long-distance RNA-RNA interactions in the coronavirus genome form high-order structures promoting discontinuous RNA synthesis during transcription. *Journal of virology* **2013**, *87*, 177-186. https://doi.org/10.1128/JVI.01782-12

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 June 2024 doi:10.20944/preprints202406.1845.v1

26. Patarca, R, Haseltine, W.A. Bioinformatics Insights on Viral Gene Expression Transactivation: From HIV-1 to SARS-CoV-2. *International Journal of Molecular Sciences* **2024**, *25*, 3378. https://doi.org/10.3390/ijms25063378
27. Wang, Y., Zhang, X. The leader RNA of coronavirus mouse hepatitis virus contains an enhancer-like element for subgenomic mRNA transcription. *J Virol* **2000**, *74*, 10571-10580. https://doi.org/10.1128/jvi.74.22.10571-10580.2000
28. Lindenbach, B.D., Sgro, J.Y., Ahlquist, P. Long-distance base pairing in flock house virus RNA1 regulates subgenomic RNA3 synthesis and RNA2 replication. *J Virol* **2002**, *76*, 3905-3919. https://doi.org/10.1128/jvi.76.8.3905-3919.2002
29. Choi, I.R., White, K.A. An RNA activator of subgenomic mRNA1 transcription in tomato bushy stunt virus. *J Biol Chem* **2002**, *277*(5):3760-6. https://doi.org/10.1074/jbc.M109067200
30. Lin, H.X., White, K.A. A complex network of RNA-RNA interactions controls subgenomic mRNA transcription in a tombusvirus. *EMBO J* **2004**, *23*, 3365-3374. https://doi.org/10.1038/sj.emboj.7600336
31. Wu, B., Grigull, J., Ore, M.O., Morin, S., White, K.A. Global organization of a positive-strand RNA virus genome. *PLoS Pathog* **2013**, *9*, e1003363. https://doi.org/10.1371/journal.ppat.1003363
32. Gong, T., Ju, F., Bu, D. Accurate prediction of RNA secondary structure including pseudoknots through solving minimum-cost flow with learned potentials. *Commun Biol* **2024**, *7*, 297. https://doi.org/10.1038/s42003-024-05952-w
33. Liu, Y., Wimmer, E., Paul, A.V. Cis-acting RNA elements in human and animal plus-strand RNA viruses. *Biochim Biophys Acta* **2009**, *1789*, 495-517. https://doi.org/10.1016/j.bbagrm.2009.09.007
34. Pathak, K.B., Pogany, J., Nagy, P.D. Non-template functions of the viral RNA in plant RNA virus replication. *Curr Opin Virol* **2011**, *1*, 332-338. https://doi.org/10.1016/j.coviro.2011.09.011
35. Sztuba-Solińska, J., Stollar, V., Bujarski, J.J. Subgenomic messenger RNAs: mastering regulation of (+)-strand RNA virus life cycle. *Virology* **2011**, *412*, 245-255. https://doi.org/10.1016/j.virol.2011.02.007
36. Goldstein, S.A., Thornbrough, J.M., Zhang, R., Jha, B.K., Li, Y., Elliott, R., Quiroz-Figueroa, K., Chen, A.I., Silverman, R.H., Weiss, S.R. Lineage A Betacoronavirus NS2 Proteins and the Homologous Torovirus Berne pp1a Carboxy-Terminal Domain Are Phosphodiesterases That Antagonize Activation of RNase L. *J Virol* **2017**, *91*, e02201-16. https://doi.org/10.1128/JVI.02201-16
37. Goldstein, S.A., Elde, N.C. Recurrent viral capture of cellular phosphodiesterases that antagonize OAS-RNase L. *Proc Natl Acad Sci U S A* **2024**, *121*, e2312691121. https://doi.org/10.1073/pnas.2312691121
38. Zhao, L., Jha, B.K., Wu, A., Elliott, R., Ziebuhr, J., Gorbalenya, A.E., Silverman, R.H., Weiss, S.R. Antagonism of the interferon-induced OAS-RNase L pathway by murine coronavirus ns2 protein is required for virus replication and liver pathology. *Cell Host Microbe* **2012**, *11*, 607-616. https://doi.org/10.1016/j.chom.2012.04.011
39. Zhang, G., Slowinski, V., White, K.A. Subgenomic mRNA regulation by a distal RNA element in a (+)-strand RNA virus. *RNA* **1999**, *5*, 550-561. https://doi.org/10.1017/s1355838299982080
40. Wertheim, J.O., Chu, D.K., Peiris, J.S., Kosakovsky Pond, S.L., Poon, L.L. A case for the ancient origin of coronaviruses. *J Virol* **2013**, *87*, 7039-7045. https://doi.org/10.1128/JVI.03273-12
41. Peng, S., Wang, Y., Zhang, Y., Song, X., Zou, Y., Li, L., Zhao, X., Yin, Z. Current Knowledge on Infectious Bronchitis Virus Non-structural Proteins: The Bearer for Achieving Immune Evasion Function. *Front Vet Sci* **2022**, *9*, 820625. https://doi.org/10.3389/fvets.2022.820625
42. Cavanagh, D., Casais, R., Armesto, M., Hodgson, T., Izadkhasti, S., Davies, M., Lin, F., Tarpey, I., Britton, P. Manipulation of the infectious bronchitis coronavirus genome for vaccine development and analysis of the accessory proteins. *Vaccine* **2007**, *25*, 5558-62. https://doi.org/10.1016/j.vaccine.2007.02.046
43. Kint, J., Dickhout, A., Kutter, J., Maier, H.J., Britton, P., Koumans, J., Pijlman, G.P., Fros, J.J., Wiegertjes, G.F., Forlenza, M. Infectious Bronchitis Coronavirus Inhibits STAT1 Signaling and Requires Accessory Proteins for Resistance to Type I Interferon Activity. *J Virol* **2015**, *89*, 12047-12057. https://doi.org/10.1128/JVI.01057-15
44. de Wit, J.J.S., Cook, J.K.A. Spotlight on avian pathology: infectious bronchitis virus. *Avian Pathol* **2019**, *48*, 393-395. https://doi.org/10.1080/03079457.2019.1617400
45. Cavanagh, D. Coronavirus avian infectious bronchitis virus. *Vet Res* **2007**, *38*, 281-297. https://doi.org/10.1051/vetres:2006055
46. Ammayappan, A., Upadhyay, C., Gelb, J.Jr., Vakharia, V.N. Identification of sequence changes responsible for the attenuation of avian infectious bronchitis virus strain Arkansas DPI. *Arch Virol* **2009**, *154*, 495-499. https://doi.org/10.1007/s00705-009-0325-9
47. Tsai, C.T., Wu, H.Y., Wang, C.H. Genetic sequence changes related to the attenuation of avian infectious bronchitis virus strain TW2575/98. *Virus Genes* **2020**, *56*, 369-379. https://doi.org/10.1007/s11262-020-01753-5
48. Keep, S., StevensonLeggett, P., Dowgier, G., Foldes, K., Webb, I., Fones, A., Littolff, K., Everest, H., Britton, P., Bickerton, E. A Temperature-Sensitive Recombinant of Avian Coronavirus Infectious Bronchitis Virus

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 June 2024 doi:10.20944/preprints202406.1845.v1

Provides Complete Protection against Homologous Challenge. *J Virol* **2002**, *96*, e01100-22. https://doi.org/10.1128/jvi.01100-22

49. Zhao, Y., Liang, R., Cheng, J., Zhao, J., Xue, J., Zhang, G. Attenuated Viral Replication of Avian *Infectious Bronchitis Virus* with a Novel 82-Nucleotide Deletion in the 5a Gene Indicates a Critical Role for 5a in Virus-Host Interactions. *Microbiol Spectr* **2022**, *10*, e0140522. https://doi.org/10.1128/spectrum.01405-22
50. Haijema, B.J., Volders, H., Rottier, P.J. Live, attenuated coronavirus vaccines through the directed deletion of group-specific genes provide protection against feline infectious peritonitis. *J Virol* **2004**, *78*, 3863-3871. https://doi.org/10.1128/jvi.78.8.3863-3871.2004
51. de Haan, C.A., Masters, P.S., Shen, X., Weiss, S., Rottier, P.J. The group-specific murine coronavirus genes are not essential, but their deletion, by reverse genetics, is attenuating in the natural host. *Virology* **2002**, 296, 177-189. https://doi.org/10.1006/viro.2002.1412
52. Laconi, A., van Beurden, S.J., Berends, A.J., Krämer-Kühl, A., Jansen, C.A., Spekreijse, D., Chénard, G., Philipp, H.C., Mundt, E., Rottier, P.J.M., Hélène Verheije, M. Deletion of accessory genes 3a, 3b, 5a or 5b from avian coronavirus infectious bronchitis virus induces an attenuated phenotype both in vitro and in vivo. *J Gen Virol* **2018**, *99*, 1381–1390. https://doi.org/10.1099/jgv.0.001130
53. Zhao, Y., Cheng, J., Yan, S., Jia, W., Zhang, K., Zhang, G. S gene and 5a accessory gene are responsible for the attenuation of virulent infectious bronchitis coronavirus. *Virology* **2019**, *533*, 12-20. https://doi.org/10.1016/j.virol.2019.04.014
54. Huang, Y.P., Wang, C.H. Sequence changes of infectious bronchitis virus isolates in the 3′ 7.3 kb of the genome after attenuating passage in embryonated eggs. *Avian Pathol* **2007**, *36*, 59-67. https://doi.org/10.1080/03079450601110015
55. Liu, S., Han, Z., Chen, J., Liu, X., Shao, Y., Kong, X., Tong, G., Rong, J. S1 gene sequence heterogeneity of a pathogenic infectious bronchitis virus strain and its embryo-passaged, attenuated derivatives. *Avian Pathol* **2007**, *36*, 231-234. https://doi.org/10.1080/03079450701338730
56. Hodgson, T., Casais, R., Dove, B., Britton, P., Cavanagh, D. Recombinant infectious bronchitis coronavirus Beaudette with the spike protein gene of the pathogenic M41 strain remains attenuated but induces protective immunity. *J Virol* **2004**, *78*, 13804-13811. https://doi.org/10.1128/JVI.78.24.13804-13811.2004
57. Armesto, M., Cavanagh, D., Britton, P. The replicase gene of avian coronavirus infectious bronchitis virus is a determinant of pathogenicity. PLoS One **2009**, 4, e7384. https://doi.org/10.1371/journal.pone.0007384
58. Jindal, G.A., Bantle, A.T., Solvason, J.J., Grudzien, J.L., D'Antonio-Chronowska, A., Lim, F., Le, S.H., Song, B.P., Ragsac, M.F., Klie, A., Larsen. R.O., Frazer, K.A., Farley, E.K. Single-nucleotide variants within heart enhancers increase binding affinity and disrupt heart development. *Dev Cell* **2023**, *58*, 2206-2216.e5. https://doi.org/10.1016/j.devcel.2023.09.005
59. Phillips, J.E., Jackwood, M.W., McKinley, E.T., Thor, S.W., Hilt, D.A., Acevedol, N.D., Williams, S.M., Kissinger, J.C., Paterson, A.H., Robertson, J.S., Lemke, C. Changes in nonstructural protein 3 are associated with attenuation in avian coronavirus infectious bronchitis virus. *Virus Genes* **2012**, *44*, 63-74. https://doi.org/10.1007/s11262-011-0668-7
60. Amarasinghe, A., De Silva Senapathi, U., Abdul-Cader, M.S., Popowich, S., Marshall, F., Cork, S.C., van der Meer, F., Gomis, S., Abdul-Careem, M.F. Comparative features of infections of two Massachusetts (Mass) infectious bronchitis virus (IBV) variants isolated from Western Canadian layer flocks. *BMC Vet Res* **2018**, *14*, 391. https://doi.org/10.1186/s12917-018-1720-9
61. Zhao, F., Han, Z., Zhang, T., Shao, Y., Kong, X., Ma, H., Liu, S. Genomic characteristics and changes of avian infectious bronchitis virus strain CK/CH/LDL/971 after serial passages in chicken embryos. *Intervirology* **2014**, *57*, 319-330. https://doi.org/10.1159/000365193
62. Liu, X.L., Su, J.L., Zhao, J.X., Zhang, G.Z. Complete genome sequence analysis of a predominant infectious bronchitis virus (IBV) strain in China. *Virus Genes* **2009**, *38*, 56-65. https://doi.org/10.1007/s11262-008-0282-5
63. Han, Z., Sun, C., Yan, B., Zhang, X., Wang, Y., Li, C., Zhang, Q., Ma, Y., Shao, Y., Liu, Q., Kong, X., Liu, S. A 15-year analysis of molecular epidemiology of avian infectious bronchitis coronavirus in China. *Infect Genet Evol* **2011**, *11*, 190-200. https://doi.org/10.1016/j.meegid.2010.09.002
64. McKinley, E.T., Jackwood, M.W., Hilt, D.A., Kissinger, J.C., Robertson, J.S., Lemke, C., Paterson, A.H. Attenuated live vaccine usage affects accurate measures of virus diversity and mutation rates in avian coronavirus infectious bronchitis virus. *Virus Res* **2011**, *158*, 225-234. https://doi.org/10.1016/j.virusres.2011.04.006
65. Thor, S.W., Hilt, D.A., Kissinger, J.C., Paterson, A.H., Jackwood, M.W. Recombination in avian gamma-coronavirus infectious bronchitis virus. *Viruses* **2011**, *3*, 1777-1799. https://doi.org/10.3390/v3091777
66. Jackwood, M.W., Hall, D., Handel, A. Molecular evolution and emergence of avian gammacoronaviruses. *Infect Genet Evol* **2012**, *12*, 1305-1311. https://doi.org/10.1016/j.meegid.2012.05.003
67. Reddy, V.R., Theuns, S., Roukaerts, I.D., Zeller, M., Matthijnssens, J., Nauwynck, H.J. Genetic Characterization of the Belgian Nephropathogenic Infectious Bronchitis Virus (NIBV) Reference Strain B1648. *Viruses* **2015**, *7*, 4488-506. https://doi.org/10.3390/v7082827

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 June 2024 doi:10.20944/preprints202406.1845.v1

68. Lin, S.Y., Li, Y.T., Chen, Y.T., Chen, T.C., Hu, C.J., Chen, H.W. Identification of an infectious bronchitis coronavirus strain exhibiting a classical genotype but altered antigenicity, pathogenicity, and innate immunity profile. *Sci Rep* **2016**, *6*, 37725. https://doi.org/10.1038/srep37725
69. Lin, S.Y., Chen, H.W. Infectious Bronchitis Virus Variants: Molecular Analysis and Pathogenicity Investigation. *Int J Mol Sci* **2017**, 18, 2030. https://doi.org/10.3390/ijms18102030
70. Bande, F., Arshad, S.S., Omar, A.R., Hair-Bejo, M., Mahmuda, A., Nair, V. Global distributions and strain diversity of avian infectious bronchitis virus: a review. *Anim Health Res Rev* **2017**, *18*, 70-83. https://doi.org/10.1017/S1466252317000044
71. Liu, X., Shao, Y., Ma, H., Sun, C., Zhang, X., Li, C., Han, Z., Yan, B., Kong, X., Liu, S. Comparative analysis of four Massachusetts type infectious bronchitis coronavirus genomes reveals a novel Massachusetts type strain and evidence of natural recombination in the genome. *Infect Genet Evol* **2013**, 14, 29-38. https://doi.org/10.1016/j.meegid.2012.09.016
72. Zhou, S., Tang, M., Jiang, Y., Chen, X., Shen, X., Li, J., Dai, Y., Zou, J. Complete genome sequence of a novel infectious bronchitis virus strain circulating in China with a distinct S gene. *Virus Genes* **2014**, *49*, 152-156. https://doi.org/10.1007/s11262-014-1063-y
73. Zhang, T., Han, Z., Xu, Q., Wang, Q., Gao, M., Wu, W., Shao, Y., Li, H., Kong, X., Liu, S. Serotype shift of a 793/B genotype infectious bronchitis coronavirus by natural recombination. *Infect Genet Evol* **2015**, 32, 377-387. https://doi.org/10.1016/j.meegid.2015.03.034
74. Jakhesara, S.J., Nath, B., Pal, J.K., Joshi, C.G., Kumar, S. Emergence of a genotype I variant of avian infectious bronchitis virus from Northern part of India. *Acta Trop* **2018**, *183*, 57-60. https://doi.org/10.1016/j.actatropica.2018.04.004
75. Jiang, L., Han, Z., Chen, Y., Zhao, W., Sun, J., Zhao, Y., Liu, S. Characterization of the complete genome, antigenicity, pathogenicity, tissue tropism, and shedding of a recombinant avian infectious bronchitis virus with a ck/CH/LJL/140901-like backbone and an S2 fragment from a 4/91-like virus. *Virus Res* **2018**, 244:99-109. https://doi.org/10.1016/j.virusres.2017.11.007
76. Al-Jallad, T., Kassouha, M., Salhab, M., Alomar, A., Al-Masalma, M., Abdelaziz, F. Molecular characterization of isolated infectious bronchitis viruses from affected vaccinated broiler flocks in Syria. *BMC Vet Res* **2020**, *16*, 449. https://doi.org/10.1186/s12917-020-02672-1
77. Dinan, A.M., Keep, S., Bickerton, E., Britton, P., Firth, A.E., Brierley, I. Comparative Analysis of Gene Expression in Virulent and Attenuated Strains of Infectious Bronchitis Virus at Subcodon Resolution. *J Virol* **2019**, *93*, e00714-19. https://doi.org/10.1128/JVI.00714-19
78. Hassan, M.S.H., Ojkic, D., Coffin, C.S., Cork, S.C., van der Meer, F., Abdul-Careem, M.F. Delmarva (DMV/1639) Infectious Bronchitis Virus (IBV) Variants Isolated in Eastern Canada Show Evidence of Recombination. *Viruses* **2019**, *11*, 1054. https://doi.org/10.3390/v11111054
79. Goraichuk, I.V., Davis, J.F., Kulkarni, A.B., Afonso, C.L., Suarez, D.L. A 25-Year-Old Sample Contributes the Complete Genome Sequence of *Avian Coronavirus* Vaccine Strain ArkDPI, Reisolated from Commercial Broilers in the United States. *Microbiol Resour Announc* **2020**, *9*, e00067-20. https://doi.org/10.1128/MRA.00067-20
80. Guzmán, M., Hidalgo, H. Live Attenuated Infectious Bronchitis Virus Vaccines in Poultry: Modifying Local Viral Populations Dynamics. *Animals (Basel)* **2020**, *10*, 2058. https://doi.org/10.3390/ani10112058
81. Legnardi, M., Tucciarone, C.M., Franzo, G., Cecchinato, M. Infectious Bronchitis Virus Evolution, Diagnosis and Control. *Vet Sci* **2020**, *7*, 79. https://doi.org/10.3390/vetsci7020079
82. Bali, K., Bálint, Á., Farsang, A., Marton, S., Nagy, B., Kaszab, E., Belák, S., Palya, V., Bányai, K. Recombination Events Shape the Genomic Evolution of Infectious Bronchitis Virus in Europe. *Viruses* **2021**, *13*, 535. https://doi.org10.3390/v13040535
83. Quinteros, J.A., Ignjatovic, J., Chousalkar, K.K., Noormohammadi, A.H., Browning, G.F. Infectious bronchitis virus in Australia: a model of coronavirus evolution - a review. *Avian Pathol* **2021**, 50, 295-310. https://doi.org/10.1080/03079457.2021.1939858
84. Quinteros, J.A., Noormohammadi, A.H., Lee, S.W., Browning, G.F., Diaz-Méndez, A. Genomics and pathogenesis of the avian coronavirus infectious bronchitis virus. *Aust Vet J* **2022**, *100*, 496-512. https://doi.org/10.1111/avj.13197
85. Lisowska, A., Pikuła, A., Opolska, J., Jasik, A., Kycko, A., Domańska-Blicharz, K. Virulence Properties of GI-23 Infectious Bronchitis Virus Isolated in Poland and Efficacy of Different Vaccination strategies. *Pathogens* **2021**, *10*, 522. https://doi.org/10.3390/pathogens10050522
86. Houta, M.H., Hassan, K.E., El-Sawah, A.A., Elkady, M.F., Kilany, W.H., Ali, A., Abdel-Moneim, A.S. The emergence, evolution and spread of infectious bronchitis virus genotype GI-23. *Arch Virol* **2021**, *166*, 9-26. https://doi.org/10.1007/s00705-020-04920-z
87. Lv, D., Dong, Z.H., Fan, W.S., Tang, N., Wang, L., Wei, L.P., Ji, Z.H., Tang, J.W., Lin, L.T., Wei, T.C., Huang, T., Wei, P., Mo, M.L. Identification of a novel avian coronavirus infectious bronchitis virus variant with three-nucleotide-deletion in nucleocapsid gene in China. *J Vet Med Sci* **2021**, 83, 1608-1619. https://doi.org/10.1292/jvms.21-0351

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 June 2024 doi:10.20944/preprints202406.1845.v1

88. da Silva, A.P., Jude, R., Gallardo, R.A. Infectious Bronchitis Virus: A Comprehensive Multilocus Genomic Analysis to Compare DMV/1639 and QX Strains. *Viruses* **2022**, *14*, 1998. https://doi.org/10.3390/v14091998
89. Marandino, A., Mendoza-González, L., Panzera, Y., Tomás, G., Williman, J., Techera, C., Gayosso-Vázquez, A., Ramírez-Andoney, V., Alonso-Morales, R., Realpe-Quintero, M., Pérez, R. Genome Variability of Infectious Bronchitis Virus in Mexico: High Lineage Diversity and Recurrent Recombination. *Viruses* **2023**, *15*, 1581. https://doi.org/10.3390/v15071581
90. Rafique, S., Jabeen, Z., Pervaiz, T., Rashid, F., Luo, S., Xie, L., Xie, Z. Avian infectious bronchitis virus (AIBV) review by continent. *Front Cell Infect Microbiol* **2024**, *14*, 1325346. https://doi.org/10.3389/fcimb.2024.1325346
91. Rohaim, M.A., El Naggar, R.F., Abdelsabour, M.A., Mohamed, M.H.A., El-Sabagh, I.M., Munir, M. Evolutionary Analysis of Infectious Bronchitis Virus Reveals Marked Genetic Diversity and Recombination Events. *Genes (Basel)* **2020**, *11*, 605. https://doi.org/10.3390/genes11060605
92. Wang, C., Hou, B. A pathogenic and recombinant infectious bronchitis virus variant (CK/CH/GX/202109) with multiorgan tropism. *Vet Res* **2023**, *54*, 54. https://doi.org/10.1186/s13567-023-01182-w
93. Li, H., Liu, G., Zhou, Q., Yang, H., Zhou, C., Kong, W., Su, J., Li, G., Si, H., Ou, C. Which strain of the avian coronavirus vaccine will become the prevalent one in China next? *Front Vet Sci* **2023**, *10*, 1139089. https://doi.org/10.3389/fvets.2023.1139089
94. Fang SG, Shen S, Tay FP, Liu DX. Selection of and recombination between minor variants lead to the adaptation of an avian coronavirus to primate cells. *Biochem Biophys Res Commun* **2005**, *336*, 417-423. https://doi.org/10.1016/j.bbrc.2005.08.105
95. Lim, X.F., Lee, C.B., Pascoe, S.M., How, C.B., Chan, S., Tan, J.H., Yang, X., Zhou, P., Shi, Z., Sessions, O.M., Wang, L.F., Ng, L.C., Anderson, D.E., Yap, G. Detection and characterization of a novel bat-borne coronavirus in Singapore using multiple molecular approaches. *J Gen Virol* **2019**, *100*, 1363-1374. https://doi.org/10.1099/jgv.0.001307
96. Wang, J., Pan, Y.F., Yang, L.F., Yang, W.H., Lv, K., Luo, C.M., Wang, J., Kuang, G.P., Wu, W.C., Gou, Q.Y., Xin, G.Y., Li, B., Luo, H.L., Chen, S., Shu, Y.L., Guo, D., Gao, Z.H., Liang, G., Li, J., Chen, Y.Q., Holmes, E.C., Feng, Y., Shi, M. Individual bat virome analysis reveals co-infection and spillover among bats and virus zoonotic potential. *Nat Commun* **2023**, *14*, 4079. https://doi.org/10.1038/s41467-023-39835-1
97. Cao, Y., Yang, R., Lee, I., Zhang, W., Sun, J., Wang, W., Meng, X. Characterization of the SARS-CoV-2 E Protein: Sequence, Structure, Viroporin, and Inhibitors. *Protein Sci* **2021**, *30*, 1114-1130. https://doi.org/10.1002/pro.4075
98. Khare, S., Gurry, C.; Freitas, L.; Schultz, M.B.; Bach, G.; Diallo, A.; Akite, N; Ho, J.; Lee, R.T.; Yeo, W.; Curation Team GC; Maurer-Stroh, S. GISAID's Role in Pandemic Response. *China CDC Weekly* **2021**, *3*, 1049-1051. https://doi.org/10.46234/ccdcw2021.55
99. Elbe, S., Buckland-Merrett, G. Data, disease and diplomacy: GISAID's innovative contribution to global health. *Global Challenges* **2017**, *1*, 33-46. https://doi.org/10.1002/gch2.1018
100. Shu, Y., McCauley, J. GISAID: from vision to reality. *EuroSurveillance* **2017**, *22*, https://doi.org/10.2807/1560-7917.ES.2017.22.13.30494
101. Konno, Y., Kimura, I., Uriu, K., Fukushi, M., Irie, T., Koyanagi, Y., Sauter, D., Gifford, R.J.; USFQ-COVID19 Consortium; Nakagawa, S., Sato, K. SARS-CoV-2 ORF3b Is a Potent Interferon Antagonist Whose Activity Is Increased by a Naturally Occurring Elongation Variant. *Cell Rep* **2020**, *32*, 108185. https://doi.org/10.1016/j.celrep.2020.108185
102. Lim, F., Solvason, J.J., Ryan, G.E., Le, S.H., Jindal, G.A., Steffen, P., Jandu, S.K., Farley, E.K. Affinity-optimizing enhancer variants disrupt development. *Nature* **2024**, 626, 151-159. https://doi.org/10.1038/s41586-023-06922-8
103. Vo Ngoc, L., Wang, Y.L., Kassavetis, G.A., Kadonaga, J.T. The punctilious RNA polymerase II core promoter. *Genes Dev* **2017**, *31*, 1289-1301. https://doi.org/10.1101/gad.303149.117
104. FANTOM Consortium and the RIKEN PMI and CLST (DGT). A promoter-level mammalian expression atlas. *Nature* **2014**, *507*, 462-470. https://doi.org/10.1038/nature13182
105. Xi, H., Yu, Y., Fu, Y., Foley, J., Halees, A., Weng, Z. Analysis of overrepresented motifs in human core promoters reveals dual regulatory roles of YY1. *Genome Res* **2007**, *17*, 798-806. https://doi.org/10.1101/gr.5754707
106. Dudnyk, K., Cai, D., Shi, C., Xu, J., Zhou, J. Sequence basis of transcription initiation in the human genome. *Science* **2024**, 26, 384(6694):eadj0116. https://doi.org/10.1126/science.adj0116
107. Kawasaki, K., Fukaya, T. Regulatory landscape of enhancer-mediated transcriptional activation. Trends Cell Biol **2024**, Feb 13, S0962-8924(24)00020-5. https://doi.org/10.1016/j.tcb.2024.01.008
108. Han, Z., Li, W. Enhancer RNA: What we know and what we can achieve. *Cell Prolif* **2022**, *55*, e13202. https://doi.org/10.1111/cpr.13202
109. Liang, L., Cao, C., Ji, L., Cai, Z., Wang, D., Ye, R., Chen, J., Yu, X., Zhou, J., Bai, Z., Wang, R., Yang, X., Zhu, P., Xue, Y. Complementary Alu sequences mediate enhancer-promoter selectivity. *Nature* **2023**, *619*, 868-875. https://doi.org/10.1038/s41586-023-06323-x

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 June 2024 doi:10.20944/preprints202406.1845.v1

110. Schoenfelder, S., Fraser, P. Long-range enhancer-promoter contacts in gene expression control. *Nat Rev Genet* **2019**, 20, 437-455. https://doi.org/10.1038/s41576-019-0128-0
111. Pombo, A., Dillon, N. Three-dimensional genome architecture: players and mechanisms. *Nat Rev Mol Cell Biol* **2015**, 16, 245-257. https://doi.org/10.1038/nrm3965
112. Long, H.K., Prescott, S.L., Wysocka, J. Ever-Changing Landscapes: Transcriptional Enhancers in Development and Evolution. *Cell* **2016**, *167*, 1170-1187. https://doi:.org/10.1016/j.cell.2016.09.018
113. Furlong, E.E.M., Levine, M. Developmental enhancers and chromosome topology. *Science* **2018**, *361*, 1341-1345. https://doi.org/10.1126/science.aau0320
114. Lin, X., Liu, Y., Liu, S., Zhu, X., Wu, L., Zhu, Y., Zhao, D., Xu, X., Chemparathy, A., Wang, H., Cao, Y., Nakamura, M., Noordermeer, J.N., La Russa, M., Wong, W.H., Zhao, K., Qi, L.S. Nested epistasis enhancer networks for robust genome regulation. *Science* **2022**, *377*, 1077-1085. https://doi.org/10.1126/science.abk3512
115. Mönttinen, H.A.M., Frilander, M.J., Löytynoja, A. Generation of de novo miRNAs from template switching during DNA replication. *Proc Natl Acad Sci U S A* **2023**, *120*, e2310752120. https://doi.org/10.1073/pnas.2310752120
116. Lau, B., Kerr, K., Camiolo, S., Nightingale, K., Gu, Q., Antrobus, R., Suárez, N.M., Loney, C., Stanton, R.J., Weekes, M.P., Davison, A.J. Human Cytomegalovirus RNA2.7 Is Required for Upregulating Multiple Cellular Genes To Promote Cell Motility and Viral Spread Late in Lytic Infection. *J Virol* **2021**, *95*, e0069821. https://doi.org/10.1128/JVI.00698-21
117. Miller, R.H., Zimmer, A., Moutot, G., Mesnard, J.M., Chazal, N. Retroviral Antisense Transcripts and Genes: 33 Years after First Predicted, a Silent Retroviral Revolution? *Viruses* 2021, *13*, 2221. https://doi.org/10.3390/v13112221
118. Toyoda, K., Matsuoka, M. Functional and Pathogenic Roles of Retroviral Antisense Transcripts. *Front Immunol* **2022**, *13*, 875211. https://doi.org/10.3389/fimmu.2022.875211
119. Lin, E., Panfil, A.R., Sandel, G., Jain, P. Novel perspectives on antisense transcription in HIV-1, HTLV-1, and HTLV-2. *Front Microbiol* **2022**, *13*, 1042761. https://doi.org/10.3389/fmicb.2022.1042761
120. Georg, J., Hess, W.R. Widespread Antisense Transcription in Prokaryotes. *Microbiol Spectr* **2018**, *6*. https://doi.org/10.1128/microbiolspec.RWR-0029-2018
121. Werner, A., Kanhere, A., Wahlestedt, C., Mattick, J.S. Natural antisense transcripts as versatile regulators of gene expression. *Nat Rev Genet* **2024**, Apr 17. https://doi.org/10.1038/s41576-024-00723-z
122. Patarca R, Haseltine WA. Forty years of HIV research inspires the development of SARS-CoV-2 therapy. *J Mol Cell Biol* **2024**, *15*, mjad065. https://doi.org/10.1093/jmcb/mjad065
123. Li, R., Caico, I., Xu, Z., Iqbal, M.S., Romerio, F. Epigenetic Regulation of HIV-1 Sense and Antisense Transcription in Response to Latency-Reversing Agents. *Noncoding RNA* **2023**, *9*, 5. https://doi.org/10.3390/ncrna9010005
124. Ward, S., Childs, A., Staley, C., Waugh, C., Watts, J.A., Kotowska, A.M., Bhosale, R., Borkar, A.N. Integrating cryo-OrbiSIMS with computational modelling and metadynamics simulations enhances RNA structure prediction at atomic resolution. *Nat Commun* **2024**, *15*, 4367. https://doi.org/10.1038/s41467-024-48694-3
125. Enjuanes, L., Zuñiga, S., Castaño-Rodriguez, C., Gutierrez-Alvarez, J., Canton, J., Sola, I. Molecular Basis of Coronavirus Virulence and Vaccine Development. *Adv Virus Res* **2016**, *96*, 245-286. https://doi.org/10.1016/bs.aivir.2016.08.003
126. Menachery, V.D., Gralinski, L.E., Mitchell, H.D., Dinnon, K.H.3rd, Leist, S.R., Yount, B.L.Jr, McAnarney, E.T., Graham, R.L., Waters, K.M., Baric, R.S. Combination Attenuation Offers Strategy for Live Attenuated Coronavirus Vaccines. *J Virol* **2018**, *92*, e00710-18. https://doi.org/10.1128/JVI.00710-18
127. van Beurden, S.J., Berends, A.J., Krämer-Kühl, A., Spekreijse, D., Chénard, G., Philipp, H.C., Mundt, E., Rottier, P.J.M., Verheije, M.H. A reverse genetics system for avian coronavirus infectious bronchitis virus based on targeted RNA recombination. *Virol J* **2017**, *14*, 109. https://doi.org/10.1186/s12985-017-0775-8
128. Johnson, M.; Zaretskaya, I.; Raytselis, Y.; Merezhuk, Y.; McGinnis, S.; Madden, T.L. NCBI BLAST: a better web interface. *Nucleic Acids Res* **2008**, *36*(Web Server issue): W5-W9. https://doi.org/10.1093/nar/gkn201
129. ICTV Coronaviridae Study Group. International Committee on Taxonomy of Viruses (ICTV). 2021. Available from: https://talk.ictvonline.org/ictv-reports/ictv_9th_report/positive-sense-rna-viruses-2011/w/posrna_viruses/223/coronaviridae-figures.
130. Kerpedjiev, P.; Hammer, S.; Hofacker, I.L. Forna (force-directed RNA): Simple and effective online RNA secondary structure diagrams. *Bioinformatics* **2015**, *31*, 3377-3379. https://doi.org/10.1093/bioinformatics/btv372
131. Gruber, A.R.; Lorenz, R.; Bernhart, S.H.; Neuböck, R.; Hofacker, I.L. The Vienna RNA Websuite. *Nucleic Acids Research* **2008**, *36*, W70-W74. https://doi.org10.1093/nar/gkn188
132. Lorenz, R., Bernhart, S.H., Höner Zu Siederdissen, C., Tafer, H., Flamm, C., Stadler, P.F., Hofacker, IL. ViennaRNA Package 2.0. *Algorithms Mol Biol* **2011**, 6, 26. https://doi.org/10.1186/1748-7188-6-26

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 June 2024 doi:10.20944/preprints202406.1845.v1

133. Bikandi, J., San Millán, R., Rementeria, A., Garaizar, J. In silico analysis of complete bacterial genomes: PCR, AFLP-PCR and endonuclease restriction. *Bioinformatics* **2004**, *20*, 798-799. https://doi.org/10.1093/bioinformatics/btg491
134. Duvaud, S., Gabella, C., Lisacek, F., Stockinger, H., Ioannidis, V., Durinx, C. Expasy, the Swiss Bioinformatics Resource Portal, as designed by its users. *Nucleic Acids Res* **2021**, *49*, W216-W227. https://doi.org/10.1093/nar/gkab225
135. Madeira, F.; Pearce, M.; Tivey, A.R.N.; Basutkar, P.; Lee, J.; Edbali, O.; Madhusoodanan, N.; Kolesnikov, A.; Lopez, R. Search and sequence analysis tools services from EMBL-EBI in 2022. *Nucleic Acids Research* **2022**, *50*, W276-W279. https://doi.org/10.1093/nar/gkac240

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.