

Review

# Betacoronavirus genomes: How genomic information has been used to deal with past outbreaks and the COVID-19 pandemic

Alejandro Llanes <sup>1</sup>, Carlos M. Restrepo <sup>1</sup>, Zuleima Caballero <sup>1</sup>, Sreekumari Rajeev <sup>2</sup>, Melissa A. Kennedy <sup>3</sup> and Ricardo Lleonart <sup>1,\*</sup>

<sup>1</sup> Centro de Biología Celular y Molecular de Enfermedades, Instituto de Investigaciones Científicas y Servicios de Alta Tecnología (INDICASAT AIP), Panama City, Panama; allanes@indicasat.org.pa (A.L.); crestrepo@indicasat.org.pa (C.M.R.); zcaballero@indicasat.org.pa (Z.C.); rlleonart@indicasat.org.pa (R.L.)

<sup>2</sup> College of Veterinary Medicine, University of Florida, Gainesville, Florida, United States; sree63rajeev@gmail.com (S.R.)

<sup>3</sup> College of Veterinary Medicine, University of Tennessee, Knoxville, Tennessee, United States; mkenned2@utk.edu (M.A.K.)

\* Correspondence: rlleonart@indicasat.org.pa; Tel.: +507-517-0740.

**Abstract:** In the 21<sup>st</sup> century, three highly pathogenic betacoronaviruses have emerged, with an alarming rate of human morbidity and case fatality. Genomic information has been widely used to understand the pathogenesis, animal origin and mode of transmission of betacoronaviruses in the aftermath of the 2002-03 severe acute respiratory syndrome (SARS) and 2012 Middle East respiratory syndrome (MERS) outbreaks. Furthermore, genome sequencing and bioinformatic analysis have had an unprecedented relevance in the battle against the 2019-20 coronavirus disease 2019 (COVID-19) pandemic, the newest and most devastating outbreak caused by a coronavirus in the history of mankind, allowing the follow up of disease spread and transmission dynamics in near real time. Here, we review how genomic information has been used to tackle outbreaks caused by emerging, highly pathogenic, betacoronavirus strains, emphasizing on SARS-CoV, MERS-CoV and SARS-CoV-2.

**Keywords:** betacoronaviruses; genomics; SARS-CoV; MERS-CoV; SARS-CoV-2; COVID-19

## 1. Introduction

Coronaviruses (CoV) are important pathogens of vertebrates with the ability to cause respiratory, enteric and systemic diseases in humans and animals. They are enveloped, single-stranded, positive-sense RNA viruses belonging to subfamily *Orthocoronavirinae* of family *Coronaviridae*, order *Nidovirales*. The subfamily is further divided into four genera, namely, *Alphacoronavirus*, *Betacoronavirus*, *Gammacoronavirus* and *Deltacoronavirus*. The majority of clinically relevant coronaviruses belong to the *Alphacoronavirus* and *Betacoronavirus* genera [1]. Genus *Alphacoronavirus* comprises species infecting a diverse group of mammals, including two (229E and NL63) of the seven known species of human coronaviruses. In the case of genus *Betacoronavirus*, the International Committee on Taxonomy of Viruses (ICTV) currently divides it into five subgenera, *Embecovirus*, *Sarbecovirus*, *Merbecovirus*, *Nobecovirus* and *Hibecovirus*, established based on phylogenetic analysis of conserved protein domains (see [Basic phylogenetic relationships](#)). The first four of these subgenera were formerly known as lineages or subgroups A, B, C and D, respectively.

Subgenus *Embecovirus* includes two human coronaviruses (HKU1 and OC43), as well as several animal coronaviruses of veterinary relevance such as bovine, canine, equine, porcine and murine coronaviruses. *Sarbecovirus* comprises the severe acute respiratory syndrome (SARS)-related coronaviruses such as SARS-CoV and SARS-CoV-2 (2019-nCoV), respectively responsible for the

2002-03 SARS outbreak and the 2019-20 coronavirus disease 2019 (COVID-19) pandemic. Several SARS-related bat coronaviruses, mainly isolated from Chinese horseshoe bats (*Rhinolophus* spp.), also belong to this subgenus. Subgenus *Merbecovirus* comprises the Middle East respiratory syndrome (MERS)-related coronaviruses, including the MERS-CoV responsible for the 2012 MERS outbreak, as well as two additional species of bat coronaviruses isolated from *Tylonycteris* and *Pipistrellus* bats. Subgenera *Nobecovirus* and *Hibecovirus* comprise only bat coronaviruses, mainly isolated from *Rousettus* and *Hipposideros* bats, respectively.

Since the 2002-03 SARS outbreak, genomic information has become ever-increasingly significant to address outbreaks caused by pathogenic coronaviruses. Before the 2019-20 COVID-19 pandemic, there were ~1,200 complete genomes of betacoronaviruses deposited in the GenBank database. The number of available genomes has increased dramatically during the pandemic, with ~4,000 genomes available as of May 2019. A variety of information including phylogenetic relationships, mode of transmission, evolutionary rates and the role of mutations in infection and disease severity can be deduced from comparing multiple genomes. In this review, we focus on the genomic features of family *Coronaviridae* with special emphasis on the *Betacoronavirus* genus. We also review how genomic information can be useful to tackle epidemics caused by these viruses, including the ongoing COVID-19 pandemic and future ones, potentially caused by emerging strains.

## 2. Genome structure and protein-coding genes

Betacoronaviruses, like all other members of the *Coronaviridae* family, have relatively large RNA genomes of around 30 kb in size (Table 1). The genomes have short untranslated regions (UTR) at both ends, with a 5' methylated cap and a 3' polyadenylated tail. Typically, genomes contain 9-12 open reading frames (ORF) (Figure 1), six of which are conserved and follow the same order, namely, those encoding the replicase/transcriptase polyproteins and the spike (S), envelope (E), membrane (M) and nucleocapsid (N) structural proteins. Replicase/transcriptase is organized in two overlapping ORFs, called ORF1a (11-13 kb) and ORF1b (7-8 kb), that occupy nearly two thirds of the genome. These ORFs are translated into two polyproteins that later cleave themselves to form several nonstructural proteins (Nsps), most of them involved in genome replication and translation [2]. The remaining 3' portion of the genome encodes the structural proteins and the so-called accessory proteins, whose number and functions vary among different coronaviruses.

As in most viruses, coronavirus genomes are compact and roughly encode for the few proteins that the virus needs for its replication cycle (Figure 2). Transcription of the protein-coding genes involves the production of subgenomic mRNAs that include a common leader sequence in their 5' end. This common leader is in turn encoded near the 5' end of the genome and its fusion with subgenomic mRNAs is mediated by a conserved transcription regulatory sequence (TRS) preceding most genes [2]. The role of viral proteins in the replication cycle and their conserved domains (Figure 3) are briefly reviewed in the sections below. We have also provided the InterPro accession numbers for these domains, if available.

### 2.1. Spike (S) protein

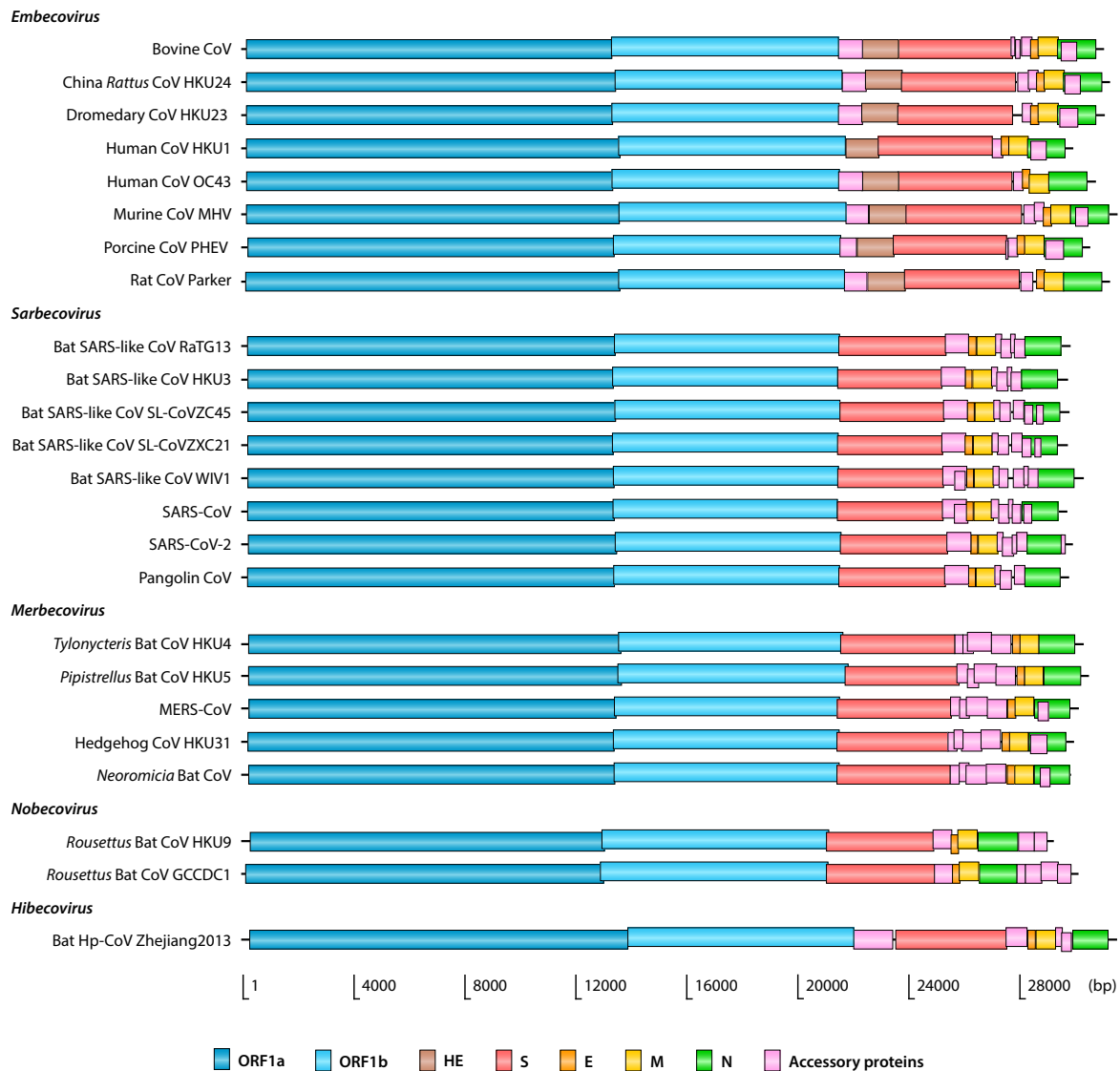
Spike (S) is a glycoprotein that recognizes the host cell receptor and allows the virus to attach to the surface of host cells. Its name refers to the spike-like structures located in the outer surface of the viral envelope, which are trimers of the S protein. After receptor recognition and attachment, the virus enters the host cell through endocytosis or by direct fusion of its envelope with the host cell plasma membrane. SARS-CoV and SARS-CoV-2 use angiotensin-converting enzyme 2 (ACE2) as their receptor [3,4], whereas MERS-CoV uses dipeptidyl peptidase 4 (DPP4) [5] and murine coronaviruses use the murine carcinoembryonic antigen-related adhesion molecule 1 (mCEACAM1a) [6,7]. Viruses from the *Embecovirus* subgenus can use certain types of sialic acids as receptors [8], due to an additional hemagglutinin esterase gene uniquely present in this subgenus (discussed below).

Table 1. Genomic features of representative betacoronaviruses.

Virus <sup>1</sup>	GenBank Accession	Size (bp)	GC%	ORFs / Accessory proteins <sup>2</sup>
<i>Embecovirus</i>				
Bovine CoV	NC_003045	31,028	37.12	12 / 5
China <i>Rattus</i> CoV HKU24	NC_026011	31,249	40.07	11 / 4
Dromedary CoV HKU23	KF906249	31,052	36.95	9 / 2
Human CoV HKU1	NC_006577	29,926	32.06	9 / 2
Human CoV OC43	NC_006213	30,741	36.79	9 / 2
Mouse hepatitis virus (MHV)	NC_001846	31,526	42.03	11 / 4
Porcine hemagglutinating encephalomyelitis virus (PHEV)	DQ011855	30,480	37.25	12 / 5
Rat CoV Parker	NC_012936	31,250	41.26	10 / 3
<i>Sarbecovirus</i>				
Bat SARS-like CoV RaTG13	MN996532	29,855	38.04	11 / 5
Bat SARS-like CoV HKU3	DQ022305	29,728	41.12	12 / 6
Bat SARS-like CoV SL-CoVZC45	MG772933	29,802	38.90	12 / 6
Bat SARS-like CoV SL-CoVZXC21	MG772934	29,732	38.82	12 / 6
Bat SARS-like CoV WIV1	KF367457	30,309	40.77	13 / 7
SARS-CoV (Human)	NC_004718	29,751	40.76	14 / 8
SARS-CoV (Civet)	AY686863	29,499	40.85	13 / 7
SARS-CoV-2 (Human)	NC_045512	29,903	37.97	12 / 6
SARS-CoV-2 (Tiger)	MT365033	29,897	37.97	11 / 5
Pangolin CoV	MT040333	29,805	38.52	10 / 4
<i>Merbecovirus</i>				
Hedgehog CoV HKU31	MK907286	29,951	37.69	10 / 4
MERS-CoV (Human)	NC_019843	30,119	41.24	11 / 5
MERS-CoV (Dromedary camel)	KF917527	29,851	41.19	10 / 4
<i>Neoromicia</i> bat CoV	MF593268	30,009	40.21	10 / 4
<i>Pipistrellus</i> bat CoV HKU5	NC_009020	30,482	43.19	10 / 4
<i>Tylonycteris</i> bat CoV HKU4	NC_009019	30,286	37.82	10 / 4
<i>Nobecovirus</i>				
<i>Rousettus</i> bat CoV GCCDC1	NC_030886	30,161	45.30	11 / 5
<i>Rousettus</i> bat CoV HKU9	NC_009021	29,114	41.05	9 / 3
<i>Hibecovirus</i>				
Bat Hp-BetaCoV Zhejiang2013	NC_025217	31,491	41.28	10 / 4

<sup>1</sup> Detailed information about the corresponding isolates is provided in supplementary table S1. <sup>2</sup> Number of open reading frames (ORFs) annotated in the corresponding GenBank entry. This number can vary among different versions of genome annotation for the same isolate.

Due to its binding specificity, the S protein determines tissue tropism and host species range of different coronaviruses. Binding specificity of the S protein is determined by its receptor-binding domain (RBD) (IPR018548), sometimes called C-domain, responsible for recognizing and binding to the host cell receptor. The sequence of the S protein is commonly divided into two sections, termed S1 and S2, corresponding to the two subunits in which the protein is cleaved by host proteases after receptor recognition, although this cleavage does not occur in all coronaviruses [9]. The RBD is located in the S1 subunit and contains a shorter receptor-binding motif (RBM) that directly interacts with the receptor. Consistent with the fact that SARS-CoV and MERS-CoV have different host cell receptors, their RBDs are structurally similar, but the corresponding RBMs differ in sequence [10]. The RBMs of SARS-CoV and SARS-CoV-2 are composed of ~70 amino acids [11,12], whereas the MERS-CoV RBM is composed of ~83 amino acids [13,14]. The S1 subunit also contains an additional N-terminal domain (NTD) (IPR032500) that has been shown to mediate binding to mCEACAM1a in murine coronaviruses [15].

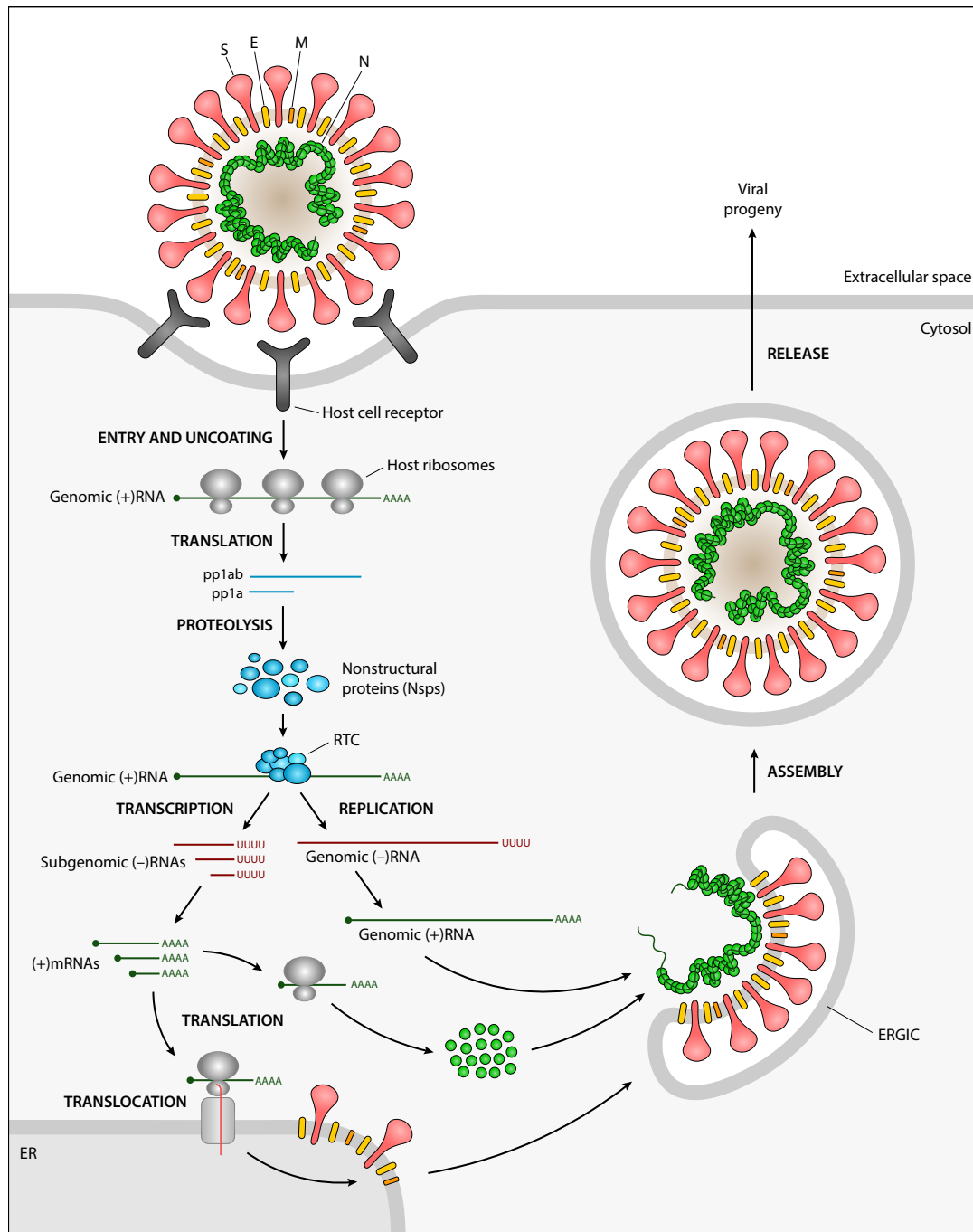


**Fig. 1. Organization of betacoronavirus genomes.** Name abbreviations are provided in Table 1.

The S2 subunit is considered to act as a class I viral fusion protein, promoting virus entry to the host cell through membrane fusion [16]. This subunit contains a fusion peptide (FP) that is believed to penetrate the host cell membrane, initiating the membrane fusion process [17]. S2 contains two additional  $\alpha$ -helical heptad repeat domains, called HR1 and HR2 (IPR027400), which interact with each other to form a coiled coil conformation, facilitating membrane fusion by bringing together the viral envelope and the host cell membrane [18]. The S2 subunit also contains a transmembrane (TM) domain that anchors the S protein to the viral envelope, as well as a short cysteine-rich endodomain, also known as CP (cytoplasmic) domain, oriented towards the interior of the viral particle.

## 2.2. Replicase/transcriptase and nonstructural proteins

Upon host cell entry, the virus is uncoated and the host ribosome then translates the first two overlapping ORFs, ORF1a and ORF1b, to generate the replicase/transcriptase polyproteins pp1a and pp1ab. The pp1a polyprotein is synthesized by translation of ORF1a, whereas the longer pp1ab polyprotein is synthesized from both ORFs, due to a ribosomal frameshifting event allowing their continuous translation [2]. These polyproteins self-cleave to produce up to 16 nonstructural proteins (Nsps) (see Snijder, Decroly and Ziebuhr [19] for a comprehensive review). Nsp1 to Nsp11 are encoded by ORF1a and are therefore present in both pp1a and pp1ab, whereas Nsp12 to Nsp16 are encoded by ORF1b and are only present in pp1ab.



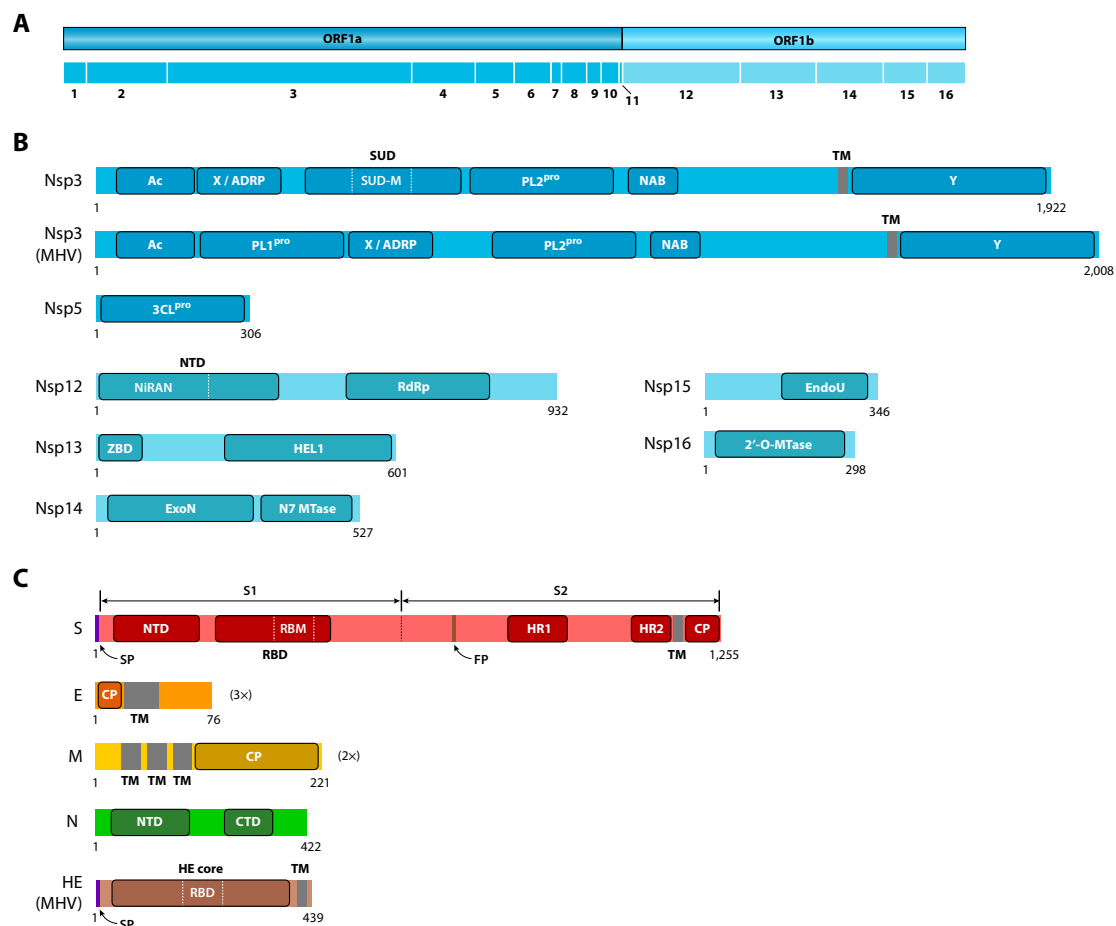
**Fig. 2. Replication cycle of a typical coronavirus.** Upon recognition of the host cell receptor, the viral particle enters the host cell and is uncoated, releasing its positive-sense genomic RNA. Host ribosomes translate polyproteins pp1a and pp1ab, which self-cleave to produce the nonstructural proteins (Nsps). Several Nsps assemble into the replicase-transcriptase complex (RTC) that generates the mRNAs for structural and accessory proteins through transcription, as well as positive-sense genomic RNAs through replication. Viral core particles are assembled within smooth vesicles derived from the endoplasmic reticulum-Golgi intermediate compartment (ERGIC). The viral progeny is ultimately released via exocytosis.

At least two Nsps are responsible for the proteolytic activity, namely, Nsp3 (papain-like protease or PL<sup>pro</sup>) and Nsp5 (3C-like protease or 3CL<sup>pro</sup>). The Nsp3 proteins from genus *Alphacoronavirus* and subgenus *Embecovirus* have two PL<sup>pro</sup> functional domains (IPR022733), respectively termed PL1<sup>pro</sup> and PL2<sup>pro</sup>. All other coronaviruses have only one domain, collinear with PL2<sup>pro</sup>. Nsp3 from SARS-related coronaviruses has several functional domains in addition to PL2<sup>pro</sup>, including an acidic (Ac) C-terminal



domain, an ADP-ribose-1''-phosphatase (ADRP) domain, a SARS-specific unique domain (SUD) (IPR024375), a nucleic acid-binding (NAB) domain (IPR032592) and a TM segment [20]. At least two of these domains seem to have affinity for single-stranded RNA [21,22].

Several Nsps (Nsp7 to Nsp16) form the active multimeric replicase/transcriptase complex (RTC). The main component of this complex is Nsp12, the RNA-dependent RNA polymerase (RdRp) that directly mediates the *de novo* primer-independent RNA synthesis during replication of the virus, as well as transcription of ORFs to produce the mRNAs for structural and accessory proteins. During replication, RdRp synthesizes a negative-sense genomic RNA by using the positive-sense genome as a template. During transcription, RdRp synthesizes negative-sense subgenomic RNAs that are subsequently transcribed into the corresponding positive-sense mRNAs. These mRNAs are then translated by the host ribosome into the structural and accessory proteins. To accomplish these functions in transcription and replication, Nsp12 has at least two well-conserved functional domains, namely, the RdRp catalytic domain (IPR007094) and a relatively large N-terminal domain (NTD) (IPR009469). This NTD is unique to the *Nidovirales* order and contains a nucleotidyltransferase subdomain called NiRAN (nidovirus RdRp-associated nucleotidyltransferase) [23].



**Fig. 3. Main functional domains in protein-coding genes. (A)** Location of Nsps along the sequence of ORF1a and ORF1b. **(B)** Functional domains of Nsp3, Nsp5, Nsp12, Nsp13, Nsp14, Nsp15 and Nsp16. **(C)** Functional domains of structural proteins. All proteins are from SARS-CoV, except for the Nsp3 and HE proteins of murine hepatitis virus (MHV), which are included for comparative purposes. Proteins are drawn to scale, except for E and M, which are drawn two (2×) and three (3×) times larger, respectively. Specific domain name abbreviations are explained in the main text. **TM**: transmembrane domain, **SP**: signal peptide, **FP**: fusion peptide, **RBD**: receptor-binding domain, **CP**: cytoplasmic domain, **NTD**: N-terminal domain, **CTD**: C-terminal domain.

The other Nsps forming the RTC assist RdRp during replication and transcription [24]. Nsp7 and Nsp8 are thought to help with the processivity of RdRp and together form the main polymerase holoenzyme [25]. Nsp13 is a highly conserved helicase subunit that is required for efficient replication of the viral genome [19]. In addition to the HEL1 helicase core domain (IPR027351), Nsp13 also has an N-terminal cysteine-rich zinc-binding domain (ZBD) (IPR027352) that appears to modulate the helicase activity [26]. Three additional nonstructural proteins, Nsp14, Nsp15 and Nsp16, have functional domains likely to be involved in RNA processing pathways. Nsp14 is a bifunctional protein that has a N7-methyltransferase domain and an ExoN domain with 3'-5' exonuclease activity. This exonuclease activity provides a proofreading function that is lacking in RdRp and enhances the fidelity of replication [27].

### 2.3. Envelope (E) and Membrane (M) proteins

Envelope (E) and membrane (M) are conserved, envelope-associated, integral membrane proteins. Proteins S, E and M are translocated into the endoplasmic reticulum (ER) of the host cell during translation. Unlike the S protein, however, E and M do not appear to have a recognizable N-terminal signal peptide [9]. Upon entry into the ER, the three proteins are integrated into the ER membrane and follow the secretory pathway towards the ER-Golgi intermediate compartment (ERGIC). There, E and M engage in several molecular interactions to facilitate assembly and release of new viral particles [28–30].

The E protein has a single TM domain and a relatively short N-terminal CP endodomain, whereas M has three TM domains and a much larger C-terminal CP endodomain. It has been suggested that both CP endodomains play a significant role in the critical functions of these proteins in assembly and release of new viral particles [28,31]. The E protein also acts as an ion channel, an activity that has been associated with its TM domain [32,33]. In SARS-CoV, this activity is not essential for replication, but it appears to be required for virulence [34].

### 2.4. Nucleocapsid (N) protein

The nucleocapsid (N) protein binds to genomic RNAs in a beads-on-a-string conformation. Unlike S, E and M, the N protein stays in the cytosol of the host cell after translation, where it binds genomic RNAs to form new nucleocapsids. These nucleocapsids travel to the ERGIC and are used for the assembly of new viral core particles. The N protein also appears to bind to Nsp3 and M, thus suggesting an important role in guiding viral RNA through replication, transcription and assembly [9]. The N protein contains two functional domains, termed N-terminal domain (NTD) (IPR037195) and C-terminal domain (CTD) (IPR037179), both of which are capable to interact with RNA [35,36].

### 2.5. Accessory proteins

Accessory proteins are genus- or species-specific and are usually dispensable for viral replication *in vitro*, but required *in vivo* [37]. The functions of accessory proteins and their pathophysiological roles are not completely understood. SARS-CoV contains at least eight ORFs encoding accessory proteins, namely, ORFs 3a, 3b, 6, 7a, 8a, 8b and 9b. Some of these proteins, particularly 6 and 7b, appear to contribute to virulence [38,39]. Most of these proteins are involved in cellular processes such as interfering with DNA synthesis, induction of caspase-dependent apoptosis, induction of proinflammatory cytokines and activation of the mitogen-activated protein kinase (MAPK) pathway [37]. Many of these accessory proteins are also incorporated into mature SARS-CoV virions, filling the role of minor structural proteins. The SARS-CoV-2 genome seems to encode a set of accessory proteins similar to that of SARS-CoV, with noticeable differences in ORFs 3a, 3b and 8b, which have been associated with interferon modulation and activation of the inflammasome [40].

Certain coronaviruses have one or two ORFs overlapping the N protein gene, although these are not always annotated in the corresponding genomes. Betacoronaviruses from the *Embecovirus* and *Merbecovirus* subgenera typically have a single overlapping ORF that has been found to encode a 23-kDa protein in mouse hepatitis virus (MHV) [41] and bovine coronavirus [42]. Experiments performed

in MHV-infected cells have demonstrated that this protein is a structural component of the MHV virion and may be involved in the processing or transport of the S protein [41]. In the *Sarbecovirus* subgenus, there are usually one or two shorter ORFs overlapping the N gene, often termed ORF9b and ORF9c. ORF9b has been shown to encode an accessory protein that is also a virion component [43] and seems to participate in the suppression of host innate immunity [44].

### 2.6. Haemagglutinin esterase (HE)

All known betacoronaviruses from the *Embecovirus* subgenus have an additional haemagglutinin esterase (HE) gene located upstream of that encoding the S protein [45]. This gene encodes a glycoprotein with neuraminidase O-acetyl esterase activity that mediates reversible attachment to O-acetylated sialic acids by acting as a receptor-binding molecule and a receptor-destroying enzyme [46]. HE is an integral membrane protein with a single TM domain and a relatively large ectodomain, which contains the esterase core domain (IPR003860) and a lectin subdomain acting as the RBD [47]. This gene is suspected to be acquired from the influenza C virus through heterologous recombination [48]. Due to the absence of this gene in all other betacoronaviruses, this event is likely to have occurred after major subgenera diverged from common ancestors.

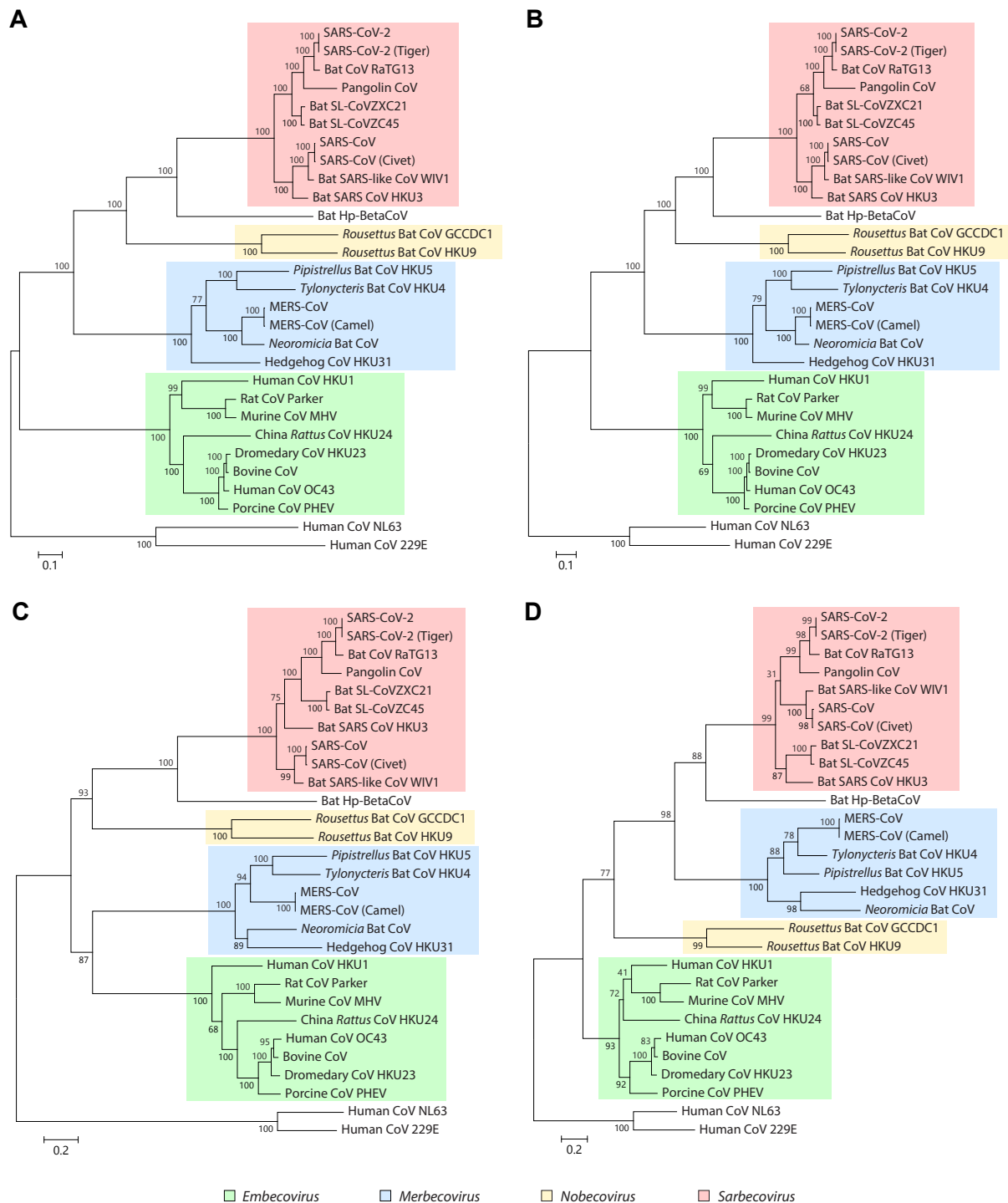
## 3. Basic phylogenetic relationships

Phylogenetic analyses help us understand the evolutionary history of viruses and provide a solid basis for their classification. A paramount application of these analyses is to make inferences about the origin of novel viral strains or species. This is particularly relevant in the context of outbreaks, to identify possible animal reservoirs involved in transmission to other susceptible hosts, including humans. To perform these analyses, researchers usually focus on genes or genomic segments conserved through all the species of interest, but with enough sequence divergence to allow their unambiguous separation in a phylogenetic tree. In the case of the coronaviruses, phylogenetic analyses are usually based on whole or partial sequences of the ORF1ab, S and N genes, whereas genes E and M are generally deemed as too short for these analyses [45].

Several genes are usually considered when exploring the phylogenetic relationships between coronaviruses, since trees built from different genomic regions often have inconsistent topologies [45] (Figure 4). One possible cause for these inconsistencies is genetic recombination, which is thought to occur frequently during evolution of coronaviruses (see [Molecular epidemiology](#)). Recombination usually involves segments totally or partially spanning the S gene, but may be associated with other regions of the genome. For instance, soon after the discovery of human coronavirus HKU1 in 2005, phylogenetic analysis suggested the existence of two putative genotypes, but conflicting results were obtained when using different regions of the genome to infer the phylogenetic relationship between these genotypes [49,50]. It was later demonstrated that these discrepancies were due to recombination between the two genotypes, with recombination breakpoints located within Nsp16 and HE [51].

For the taxonomic classification of coronaviruses, the ICTV currently recommends the use of domains 3CL<sup>pro</sup>, NiRAN, RdRp, ZBD and HEL1 of Nsp3, Nsp12 and Nsp13 [52]. These domains are conserved in all viruses of the order *Nidovirales* and can therefore be used for deeper phylogenetic analyses [53]. Recent studies exploring the phylogenetic position of SARS-CoV-2 have shown that trees built using some of these conserved domains are consistent with those based on whole genome sequences, at least at the genus and subgenus levels [4,52,54,55] (Figure 4A and 4B). However, in studies exploring shallower phylogenetic relationships, such as those focused on closely related viral strains isolated from different hosts, conserved domains may not have enough variability to ensure robust separation of some taxa. This is evidenced by the relatively low branch support estimates occasionally obtained for some subclades from the same subgenus in trees based on conserved domains, when compared to whole genome trees [4,52] (Figure 4B and 4D). More variable segments or complete genome sequences may be a better choice in these scenarios, to build more robust phylogenetic trees.





**Fig. 4. Phylogenetic analysis of representative betacoronaviruses.** Figure shows four alternative phylogenies for coronaviruses in Table 1, inferred from complete genome sequences (A), concatenated sequences of ORF1ab domains (B), whole S protein (C) and the receptor-binding domain (RBD) (D). Phylogenetic analysis was performed as previously described [4,52], briefly, sequences were aligned with MAFFT [56] and trees were built with IQ-TREE [57], with the maximum likelihood (ML) method and the GTR+G+I model. For protein sequences, amino acid alignments were converted to nucleotides with PAL2NAL [58]. Numbers above or below branches indicate branch support measures expressed as percentage and estimated using the Shimodaira-Hasegawa (SH)-like approximate likelihood ratio test (aLRT) with 1,000 replicates. Trees were rooted with human alphacoronaviruses 229E and NL63 (GenBank accession numbers NC\_002645 and NC\_005831, respectively).

Several studies addressing the origin of coronavirus species have identified bats as the natural reservoirs of alphacoronaviruses and betacoronaviruses [59,60]. This is not surprising, since bat coronaviruses are highly ubiquitous in most currently accepted taxonomic subgroups [1]. In fact, five of the seven known human coronaviruses are likely to have originated from bats, namely, NL63, 229E, MERS-CoV, SARS-CoV and SARS-CoV-2. The remaining two human coronaviruses, HKU1 and OC43, are thought to have originated from rodents [61]. In the case of SARS-CoV, its possible origin from bats was first suggested in 2005, when two studies independently reported the discovery of SARS-related coronaviruses isolated from Chinese horseshoe bats (*Rhinolophus* spp.) [62,63], with several more strains discovered in subsequent years (reviewed by Luk et al. [64]). Similar findings have been reported for MERS-CoV, which was found to be closely related to coronaviruses isolated from bamboo bats (*Tylonycteris* spp.) and pipistrelle bats (*Pipistrellus* spp.), respectively termed *Tylonycteris* bat coronavirus HKU4 (Ty-BatCoV-HKU4) and *Pipistrellus* bat coronavirus HKU5 (Pi-BatCoV-HKU5) [65]. In phylogenetic trees, these two coronaviruses separate well from MERS-related coronaviruses found in other bat species, including those recently isolated from serotine bats (*Neoromicia* spp.) in South Africa [66].

Another important conclusion drawn from exhaustive phylogenetic analyses is that, although bats appear to act as natural reservoirs of coronaviruses, intermediate animal hosts may also play a critical role in transmission to other susceptible hosts. The presumed intermediate host for SARS-CoV, the masked palm civet from the *Viverridae* family (*Paguma larvata*), was identified even before the natural bat carriers, when highly similar SARS-CoV strains were found in the civets from a wet market and in workers supposed to handle them [67]. The intermediate role of civets was suspected when comparing samples from market civets to those in the wild, which suggested that SARS-CoV was likely transmitted to the market civets by other animals [68]. A similar scenario has been reported for MERS-CoV and its intermediate host, the dromedary camel (*Camelus dromedarius*). After the 2012 MERS outbreak, MERS-CoV strains highly similar in sequence to those from human patients were isolated from camels [69,70]. In 2013, a novel dromedary camel coronavirus HKU13 was also identified [71], however, phylogenetic analysis positioned it within the *Embecovirus* subgenus and it is therefore not directly related to MERS-CoV.

Soon after the onset of the 2019-20 COVID-19 pandemic and the availability of the whole genome sequence of the novel coronavirus, phylogenetic analysis revealed that it was closely related to SARS-CoV and it was officially designated as SARS-CoV-2 [52]. Not surprisingly, it was soon reported that SARS-CoV-2 was phylogenetically related to two bat SARS-like coronaviruses, SL-CoVZC45 and SL-CoVZXC21, previously isolated from *Rhinolophus sinicus* in 2018 [54,72,73]. Bat coronavirus RaTG13, isolated from *Rhinolophus affinis* in 2013, was found to be even more closely related to SARS-CoV-2 than the first two [4]. Further studies identified coronaviruses similar to SARS-CoV-2 in Malayan pangolins (*Manis javanica*), a highly smuggled animal illegally sold in China [55,74]. Although bat coronavirus RaTG13 is phylogenetically closer to SARS-CoV-2 than the pangolin coronaviruses, the RBD of the latter are more similar to those of SARS-CoV-2, thus suggesting a possible role of pangolins as intermediate hosts in transmission to humans [75].

### 3. Molecular Epidemiology

Molecular epidemiology focuses on the contribution of genomic, genetic and other molecular factors to etiology, distribution and prevention of diseases. Central to molecular epidemiology of betacoronaviruses is their circulation among different animal hosts, as well as the evolutionary forces that facilitate these cross-species jumps. Here, we discuss how genomic information has been used to better understand the rate of evolution of betacoronaviruses and their transmission in human populations, as well as the evolutionary changes associated with host and tissue tropism.

#### 3.1 Evolutionary rates and divergence

Estimation of evolutionary rates is an important step to characterize the genetic diversity among viral lineages and to place a timescale in phylogenetic hypotheses explaining their origin and divergence. The rate of evolution of viruses is often assessed through the number of errors occurring

during replication of the viral genome (the mutation rate) and the frequency at which such mutations become fixed in the population (the substitution rate) [76]. The substitution rate depends on several factors, including the underlying mutation rate and the presence of selective forces that influence fixation of mutations in association with their fitness. Mutation rates of RNA viruses are generally higher than those of DNA viruses, due to the lack of a proofreading activity and consequent low fidelity of their RdRp [77]. However, due to the proofreading activity of Nsp14, members of the order *Nidovirales* have relatively lower mutation rates [78].

The substitution rate is often expressed in substitutions per nucleotide site per year (s/n/y) and can be estimated from phylogenetic reconstructions, when divergence time is known for particular lineages. Although several methods have been traditionally used to estimate substitution rates, including linear regression and maximum likelihood (ML), the most popular method nowadays is the Bayesian Markov chain Monte Carlo (MCMC) approach, such as that implemented in the BEAST package [79]. Globally, substitution rates of coronaviruses have been estimated to be in the order of  $10^{-3}$ – $10^{-4}$  s/n/y [80,81]. Studies conducted in SARS-CoV and MERS-CoV have estimated substitution rates for the whole genome to be between  $0.80$ – $2.38 \times 10^{-3}$  and  $0.88$ – $1.37 \times 10^{-3}$  s/n/y, respectively [82,83]. However, variation in the estimates for particular genes have been observed for both SARS-CoV [84–88] and SARS-CoV-2 [89,90], suggesting that some genes may be subjected to negative or positive selective pressure.

An important application of this type of analysis is the estimation of the time to most recent common ancestor (TMRCA) between two lineages, as an approximate measure of their time since divergence. Several studies have estimated that the SARS-CoV lineage within the SARS-related coronaviruses most probably emerged between 1961–1985, while the civet SARS-CoV strains may have originated at a time period around 1986–1995 [86–88]. TMRCA estimates for the SARS-CoV and MERS-CoV strains respectively involved in the 2002–03 and 2012 outbreaks have been roughly consistent with period of time in which the first cases were reported [83,88]. A preliminary study has estimated that the group containing SARS-CoV-2 and its closest bat coronavirus, RaTG13, may have diverged between 40–70 years ago [91].

### 3.2. Recombination, RBD mutations and host/tissue tropism

Recombination events are often inferred by comparing phylogenetic trees built from different genes or genomic regions, since occurrence of recombination often leads to inconsistent topologies in such trees (see [Basic phylogenetic relationships](#)). One of the most popular methods for detecting recombination in viral genomes is bootscan analysis [92]. To use this method, sequences of target genomes are aligned against reference sequences associated with the suspected recombination events. The alignments are divided into short sequential segments and phylogenetic trees are then built from these segments. Recombination is suspected in segments for which the trees exhibit an alternative topology, involving different reference sequences. Bootscan and other complementary methods, such as sequence similarity plots for the putative recombinant regions, are implemented in packages like SimPlot [93] or the Recombination Detection Program (RDP) [94]. Studies relying on these methods have documented the occurrence of recombination in several coronavirus genera and have also provided ample evidence supporting the important role of this process in coronavirus cross-species transmission [1,95,96].

The first reported example of natural recombination in human coronaviruses was that occurring between two different HKU1 genotypes [51]. Putative recombination events between genotypes have also been documented for human coronaviruses NL63 [97] and OC43 [98]. In the case of SARS-related and MERS-related coronaviruses, recombination appears to occur among strains infecting several animal hosts, including bats, intermediary hosts and humans (reviewed by Hu et al. [60] and Su et al. [96]). Studies considering several bat SARS-related coronaviruses have suggested the occurrence of recombination in lineages leading to human and/or civet strains of SARS-CoV, with breakpoints often located close or within the S and ORF8 genes [87,99,100]. These findings suggest that recombination between existing strains can result in new strains or species, with possible differences in host and tissue tropism. For instance, bat coronavirus strains Rs3367 (WIV1) and WIV16 have been reported to have high sequence similarity to human/civet SARS-CoV at the S gene, allowing them to use ACE2 as a

receptor for cell entry [101,102]. Recombination analysis suggested that at least one civet SARS-CoV strain (SZ3) may have originated by recombination between WIV16 and another bat SARS-CoV strain (Rf4092) [103].

In the case of SARS-CoV-2, comparison of its genome to those of other SARS-related coronaviruses did not provide enough evidence supporting recent recombination as a possible explanation for its origin [4,54,73]. However, two putative breakpoints, possibly derived from a past recombination event, were identified within the S gene, flanking its RBD [73]. Globally, the SARS-CoV-2 genome is more similar to those of bat coronaviruses SL-CoVZXC21 and SL-CoVZC45, however, the region between the two breakpoints was found to be more similar to human/civet SARS-CoV and WIV1. Putative recombination signals have also been reported between SARS-CoV-2, RaTG13 and the pangolin-related coronaviruses [55,74]. Although SARS-CoV-2 is more similar to RaTG13 than to the pangolin coronaviruses, some of the latter have higher sequence similarity to SARS-CoV-2 in the RBD. It has also been suggested that these similarities at the amino acid level may be due to convergent evolution, arising from positive selection instead of recombination [74,104].

The fact that different evolutionary events often involve the RBD is likely to be associated with the role of this domain and its RBM in receptor recognition and adaptation to different animal hosts. In SARS-CoV-like viruses, six RBD amino acids have been found to be essential for binding to ACE2, five of which differ between SARS-CoV and SARS-CoV-2 [12,105]. Particular sets of RBD mutations appear to be associated with a specific host range for each coronavirus species, as is the case of humans and civets in SARS-CoV or humans and camels in MERS-CoV. Although it has been suggested that SARS-CoV-2 is optimized for binding to human ACE2, it may also infect other animals with highly similar ACE2 homologs such as pigs, ferrets, cats and primates [75,105]. In fact, the S gene of a SARS-CoV-2 strain recently isolated from a tiger (GenBank accession number MT365033) is identical to those of human isolates, both clustering into the same branch in phylogenetic trees (Figure 4).

Comparative sequence analysis has suggested that positive selection may have a role in shaping the evolution of the S protein and the RBD of SARS-CoV and MERS-CoV [84,106–108]. The role of natural selection in the evolution of particular genes is typically inferred by computing the ratio of the rates of nonsynonymous to synonymous changes ( $K_a/K_s$ ) between groups or lineages, with a value greater than 1 indicating an overall positive selective pressure. Although evolution of the SARS-CoV genome during the 2002-03 SARS outbreak was found to be largely neutral or nearly neutral, at least six mutations occurred in the S protein during the early, middle and late phases of the outbreak, all of which were present in the epidemic strain (Urbani) [106,109]. However, the average  $K_a/K_s$  values for the early phase were found to be significantly higher than those for the middle and late phases, suggestive of initial positive selection in the S gene, followed by purifying selection and stabilization. Likewise, a recent study has suggested limited episodes of positive selection during divergence of SARS-CoV-2 from RaTG13, although there is still insufficient evidence to associate these changes with its adaptation to humans [110].

### 3.3. Genetic variation and transmission in human populations

As pathogenic viruses replicate and spread during outbreaks, their genomes accumulate random mutations that can be used to track the spread of the disease, reconstruct their transmission routes and detect lineages with different levels of virulence and transmissibility. There are several methods for tracking mutations and inferring the mode of transmission from genomic data, most of which require the alignment of sequences from new isolates to reference genomes and the subsequent identification of genetic variants such as single nucleotide variants (SNV) and insertion/deletions (indels). The simplest and fastest methods are based on pairwise distances among samples computed from these variants. However, these methods do not consider evolutionary models and can be highly inaccurate when there is substantial divergence between donor and recipients in transmission chains [111]. More advanced methods are based on ML or MCMC approaches, often within a Bayesian framework, applying an explicit model of evolution to phylogenetic estimation. When these methods are combined with sampling dates, estimations of the presence of significant molecular evolution over a sampling



period is possible [111]. Such analyses are implemented in packages like TransPhylo [112,113], Phyloscanner [114], Outbreaker2 [115] or Phybreak [113].

The main output of these methods is a transmission tree indicating which individuals infected others. Although a transmission tree cannot be directly inferred from a phylogenetic tree, it must be consistent with the underlying phylogeny. Together, phylogenetic and transmission trees help us trace back the origin of an outbreak, detect multiple introductions of a pathogen into a given territory, identify mutations that define specific lineages and predict the potential existence of unsampled individuals that may have acted as missing transmission links. Due to their relatively recent development, transmission trees based on genomic data were not widely used to study the transmission routes of previous SARS and MERS outbreaks, however, a recent study analyzing data from the 2003 SARS outbreak has provided new insights into its early stages [116].

Progress in genome sequencing technologies has resulted in an exceptionally high number of SARS-CoV-2 genomes sequenced during the 2019-20 COVID-19 pandemic. In fact, the COVID-19 pandemic is the second one in history, after the 2009 H1N1 influenza pandemic [117,118], for which genomic data has been generated almost in a real-time fashion, allowing a very detailed reconstruction of transmission trees. Genome sequences have been made publicly accessible through several repositories, including a data sharing service hosted by the Global Initiative on Sharing All Influenza Data (GISAID) (<https://www.gisaid.org/>) [119,120]. In addition to public genome repositories, open-source platforms for real-time data visualization and analysis of genomic data are also available, including NextStrain (<https://nextstrain.org>) and CoV-GLUE (<http://cov-glue.cvr.gla.ac.uk>). NextStrain is fed with sequences from the GISAID repository and uses the Augur bioinformatics toolkit (<https://github.com/nextstrain/augur>) for tracking molecular evolution and the Auspice software (<https://nextstrain.github.io/auspice/>) for interactive visualization of phylogenomic data. Conversely, CoV-GLUE is a web application based on an integrated software environment called GLUE (Genes Linked by Underlying Evolution), designed to create bioinformatic resources based on viral genome sequences [121]. CoV-GLUE is also based on GISAID data and contains a database of replacements and indels that have been found in previously sampled SARS-CoV-2 sequences. New SARS-CoV-2 genome sequences can be loaded into the platform to identify novel or known mutations, assign them to potential lineages and visualize them in a phylogenetic context.

Several attempts have been made to classify circulating strains of SARS-CoV-2 into lineages or genotypes with potential differences in transmissibility and disease severity. Among these, a study comparing 103 SARS-CoV-2 genomes suggested the existence of two lineages, termed L and S, with potential differences in prevalence [104]. Another recent study compared 160 genomes from different countries and suggested the existence of three subtypes, A B and C, with differences in geographic distribution and prevalence [122]. However, such studies have been criticized for possible sampling biases and misinterpretation of results [123–126], stressing that caution should be taken when drawing conclusions from genomic analyses. Limited or inappropriate sampling can bias the inference of transmission networks, potentially hiding introduction events and intermediate states and resulting in inaccurate mutation rate estimates [111]. When describing new lineages based on SNVs and other genetic variants, fixation of the corresponding mutations should be first demonstrated in local populations.

#### 4. Diagnostics, drug design and vaccine candidates

Early diagnosis and rapid development of drugs and vaccines targeting emerging viruses are essential to limit their spread, but traditional development approaches are time-consuming and often inefficient. Conversely, sequence-based approaches allow rapid understanding of viral protein function and pathogenesis, as well as the identification of virus-specific factors and targets suitable for drug and vaccine design. Here, we briefly review how genomic information has fueled the development of diagnostic systems and the design of new drug and vaccine candidates. We begin with a brief introduction to the application of genomics in the development of reverse genetic systems, which are key to the previously mentioned fields.



#### 4.1. Reverse genetic systems

Genomic information has been widely used in the development of reverse genetic systems, allowing the construction of synthetic viral infectious particles and the manipulation of the genetic composition of viruses for research purposes [127]. These systems have proven indispensable for the characterization of human and animal betacoronaviruses, especially when there is limited access to clinical isolates or for research institutions that do not have the appropriate containment facilities [128]. Reverse genetics was initially difficult to implement for betacoronaviruses due to the relatively large size of their viral RNA genome, which affects transfection efficiency and stability in standard bacterial vectors [127]. Although several systems can be used to construct synthetic coronavirus genomes (reviewed by Almazán et al. [129]), most studies focusing on betacoronaviruses have used a sub-cloning strategy based on *in vitro* ligation, originally developed for the transmissible gastroenteritis virus (TGEV) [130]. This system is based on the systematic and precise assembly of complete cDNA genomes from a panel of cDNA cassettes that span the entire viral genome and that are flanked by native or engineered specific restrictions sites, allowing the construction of full-length infectious clones. This assembly strategy was rapidly deployed for the study of human pathogenic betacoronaviruses, including SARS-CoV, MERS-CoV and SARS-CoV-2 [131–133].

The use of reverse genetic systems has allowed targeted genetic manipulation of viral genes and creation of homogeneous viral stocks for running *in vitro* and *in vivo* assays. Based on their genomic organization, reporter strains of both SARS-CoV and MERS-CoV have been created by replacing the ORFs of accessory proteins with luciferase and fluorescent proteins as reporter genes [134–137]. These reporter strains, as well as mouse-adapted SARS-CoV and MERS-CoV clones, have been used to assess the role of individual mutations in host adaptation [136–139]. Similar studies based on reverse genetics have been critical in characterizing the function of several Nsps in replication and transcription, as well as modulation of host processes such as inflammatory responses during infection [19].

Reverse genetic systems are also useful to understand how the viruses evolve during outbreaks and epidemics. For instance, S protein mutations from zoonotic, early, middle and late epidemic strains of the SARS-CoV outbreak have been introduced into the S protein of the epidemic strain of SARS-CoV (Urbani) to evaluate the effect of those mutations on viral entry into human cells and viral pathogenesis in rodent and primate models [140–142]. In addition, reverse genetic systems have proven useful in analyzing the emergence and pathogenic potential of bat SARS-related and MERS-related coronaviruses [143–146]. Recombinant versions of bat betacoronaviruses can be used to evaluate the efficiency of the S protein-mediated viral entry and replication and to characterize genetic changes required for efficient infection of human cells.

#### 4.2. Diagnostics

Methods based on polymerase chain reaction (PCR) are the most frequently used for detecting highly pathogenic human betacoronaviruses. These methods have several advantages including their high sensitivity and specificity, their feasibility in settings where virus isolation is not possible due to safety concerns and their ability to detect virus presence early after infection, even before the onset of symptoms [147,148]. After the 2002-03 SARS-CoV outbreak, random-amplification deep-sequencing approaches have played a crucial role in discovery and characterization of genomic differences among SARS-related coronaviruses and identification of the emerging MERS-CoV and SARS-CoV-2 [73,149,150]. These studies allowed the rapid development of genus- and species-specific real-time PCR assays based on the genomes of these viruses. Available PCR tests for human pathogenic betacoronaviruses employ either a single or multiple primer sets targeting specific regions of the ORF1ab, E and N genes [147,148,151–155]. Although most of the assays developed using the aforementioned genes show no cross-reactivity with related species and hence, high specificity, assays targeting the N gene displayed higher sensitivity, probably because this transcript is very abundant during replication of betacoronaviruses [148,151,153,156,157]. Actually, even if relative abundance of subgenomic mRNAs is believed to be kept well controlled during the replicative cycle, an increasing

gradient of expression has been reported from 5' to 3', with the N gene exhibiting the highest expression levels in cells infected with MHV [158].

Regardless of their efficiency, PCR-based assays have several drawbacks for their massive use during outbreaks and epidemics, including requiring specialized and costly equipment and reagents, as well as having turnaround times ranging from 2 to 4 days due to the time required for sample transportation to centralized testing facilities, preparation and performance of the actual PCR test. In recent years, a rapid molecular test based on a combination of reverse transcription loop-mediated isothermal amplification and a vertical flow visualization strip (RT-LAMP-VF) have been developed to detect the N gene of MERS-CoV, exhibiting no cross-reactivity with SARS-related coronaviruses and a turnaround time of approximately 35 minutes [156]. A new assay called DNA Endonuclease-Targeted CRISPR Trans Reporter (DETECTR) has been developed for SARS-CoV-2 detection, which can be performed in less than 40 minutes [157]. This assay performs simultaneous RT-LAMP for RNA samples followed by Cas12 detection using guide RNA sequences targeting species-specific regions of the E and N genes of SARS-CoV-2.

Immunoassays based in antigen-antibody recognition are an alternative for the establishment of point-of-care tests that deliver fast results at a low-cost, and are fundamental for providing diagnostic evidence and for better understanding of the epidemiology of emerging betacoronaviruses, including the burden of asymptomatic infections and exposure. Knowledge of the genomic sequences of infecting coronaviruses has been critical for the development and validation of immunoassays that either use monoclonal antibodies (mAbs) to detect viral antigens in clinical samples or cloned viral antigens to detect patient antibodies directed against the virus [147,159–162]. Development of antigen tests requires the expression of recombinant viral proteins or fragments of them that contain potential epitopes predicted by sequence homology to previously described immunogenic motifs [163,164]. These recombinant antigens are subsequently used for the production of specific mAbs, followed by experimental validation of their affinity for viral antigens and characterization of their specific epitopes [165].

Prototypes of direct antigen tests have been developed previously for SARS-CoV and MERS-CoV, but have not received regulatory approval, whereas SARS-CoV-2 antigen tests are currently under development [147,159,161,166]. Most of these antigen-based assays have targeted the N protein, since it is probably the most convenient target for virus detection in patients due to its high abundance. Serological assays that rely on recombinant proteins as antigens have been developed for detection of SARS-CoV, MERS-CoV and SARS-CoV-2, mainly using the N and S proteins as the two major immunogenic proteins of these viruses [160,162,167,168]. Although serological assays have limited utility for diagnostic purposes due to the variable time span for antibodies to be detectable after initial infection, these assays may be especially useful for unveiling the real epidemiological impact of pandemics such as the 2019-20 COVID-19 pandemic, given the increasing evidence of highly abundant asymptomatic carriers [169–171].

#### 4.3. Drug design

When SARS-CoV suddenly emerged in late 2002, the initial approach to drug discovery was to test existing broad-spectrum antiviral drugs as potential anti-betacoronavirus candidates [172]. In addition to drug repurposing, a more rational approach widely used in the aftermath of SARS-CoV and MERS-CoV was structure-based drug design (reviewed by Hilgenfeld and Peiris [173]). Not surprisingly, similar strategies have also been explored since the beginning of the 2019-20 COVID-19 pandemic [174]. Accelerated discovery of new SARS-related betacoronaviruses and characterization of their genomes have allowed the incorporation of genomic information into drug discovery pipelines. Reverse genetic systems have made possible biological assays for characterizing the function of viral proteins, the first important step for identification of potential virus-specific drug targets [131–133]. Both genomic and functional knowledge have allowed the development of small interfering RNA (siRNA) molecules targeting specific viral proteins. This strategy has been used to design siRNA inhibitors targeting the ORF1b and S genes of SARS-CoV [175] and has also been suggested as a valid strategy against SARS-CoV-2 [176].

Genomic knowledge also serves as the basis for other ‘Omics’ such as transcriptomics, proteomics and interactomics, which have also been crucial for accelerating drug discovery against SARS-CoV-2. Based on the high sequence similarity between SARS-CoV-2 and other human betacoronaviruses, especially SARS-CoV, network proximity analysis of drug targets and virus-host interactions in the human interactome has been already used as a tool for accelerating drug repurposing [177]. In this type of analysis, proteins functionally associated with viral infection are localized in the corresponding subnetwork within the human protein-protein interaction network, and those proteins that serve as drug targets for specific diseases are selected as potential targets for antiviral drugs. These analyses are followed by bioinformatic validation of drug-induced gene signatures and human betacoronavirus-induced transcriptomics in human cell lines to inspect the postulated mechanisms of action in a specific human betacoronavirus.

A high-resolution map of the SARS-CoV-2 transcriptome and epitranscriptome has been recently elucidated [178]. Data from this study revealed a highly complex transcriptome, characterized by a large number of transcripts encoding unknown ORFs produced by fusion, deletion and/or frameshift events. Furthermore, direct RNA sequencing suggested 41 potential RNA modification sites on the viral transcripts, the majority of them containing the AAGAA motif. Functional characterization of these newly discovered ORFs and RNA modification sites may unveil key roles in viral pathogenesis, defining new potential targets for antiviral therapy. In a collaborative effort, several research groups were able to clone, tag and express 26 of the 29 viral proteins found in human cells [179]. More than 300 high-confidence SARS-CoV-2-human protein-protein interactions were further identified using affinity-purification mass spectrometry (AP-MS). Applying a combination of systematic chemoinformatic drug search and pathway centric analysis to the whole set of interactions, 66 druggable human proteins were identified that are targeted by 69 existing approved drugs and compounds in clinical and/or pre-clinical trials.

#### 4.4. Vaccine candidates

In the aftermath of previous SARS and MERS outbreaks, several laboratories around the globe pursued the development of vaccines using the traditional strategy of inactivating whole viral particles (reviewed by Roper and Rehm [180] and Zumla et al. [172]). Increasing availability of genomic information regarding SARS-CoV, MERS-CoV and other related betacoronavirus have allowed the development of other types of vaccine formulations such as live-attenuated vaccines, recombinant vector vaccines and DNA vaccines. Reverse genetic systems have been used to develop and characterize live-attenuated vaccine platforms in both SARS-CoV and MERS-CoV, based on substitutions of key residues of the Nsp16 active site [181,182], deletion of the E gene [183–186] or inactivation of the exonuclease activity of Nsp14 [187]. Although showing promising results, these live-attenuated vaccines also raised safety concerns, due to the possibility of recombination and reversal of mutations that could restore the functionality of the inactivated proteins.

As an alternative to live-attenuated vaccines, recombinant vector vaccine candidates have been developed for SARS and MERS using either adenovirus [188–194], parainfluenza [195], vesicular stomatitis virus [196], attenuated measles virus [197], baculovirus [198], vaccinia modified virus Ankara [199–201] and attenuated *Salmonella* [202] as vectors for expression of S, E, M and N proteins. These recombinant vectors express the foreign target protein in the cytoplasm of the host cell, thus inducing both cellular and humoral immune responses. Following a similar principle, DNA and RNA vaccines can induce both B- and T-cell mediated immunity without the use of any viral particle, by simply introducing into the host cell plasmids encoding proteins of the pathogen that are then endogenously produced. Several DNA vaccine formulations against SARS-CoV and MERS-CoV, mainly based on the S, M and N proteins, showed promising results in the pre-clinical phase [203–208].

Many SARS-CoV and MERS-CoV vaccine formulations evaluated in animal models protected animals from challenge with the virus, but failed to induce protective immunity in aged groups and exacerbated SARS symptoms in younger groups subsequently challenged with the virus [209–213]. These findings stress the importance of developing subunit vaccines, as these will offer targeted immunogenicity with improved safety. For this purpose, bioinformatic tools can be used to predict

potential epitopes in proteins encoded in the genomes, based on their sequence similarity to previously described immunogenic motifs or through structural methods such as molecular docking simulations. A recent study used predictive bioinformatic tools to identify potential B- and T-cell epitopes for SARS-CoV-2 in regions of its genome with high sequence similarity to SARS-CoV [164]. Epitopes derived from the S protein of human betacoronaviruses seem to be the most promising for the development of strong subunit vaccines as it has been shown that the SARS-CoV S protein can induce serum-neutralizing antibodies [195,203] and generate CD4<sup>+</sup> and CD8<sup>+</sup> T-cell responses [214].

Thanks to the availability of SARS-CoV-2 genome sequences and the previous experience with SARS-CoV and MERS-CoV, numerous vaccine projects using diverse technologies are currently in progress, with some already entering into clinical trials (reviewed by Amanat and Krammer [215]). Evidence that neutralizing antibodies against SARS-CoV cross-react with SARS-CoV-2 suggests that SARS-CoV vaccines might cross-protect against SARS-CoV-2. Unfortunately, the few SARS-CoV vaccines that made it to phase I clinical trials were not further funded due to the control of the disease [215]. Noteworthy is that, several of these abandoned projects for SARS-CoV vaccines have been reactivated and rapidly adapted to SARS-CoV-2. The relatively high sequence divergence between SARS-CoV-2 and MERS-CoV makes it unlikely that vaccines targeting MERS-CoV can induce strong cross-neutralizing antibodies against SARS-CoV-2. However, existing platforms for the development of MERS-CoV vaccines were also rapidly adapted for the production of SARS-CoV-2 vaccines, as is the case of the RNA-based vaccine developed by Moderna Therapeutics and the Vaccine Research Center at the National Institutes of Health (NIH), which encodes a segment of the S gene encapsulated in lipid-based nanoparticles (ClinicalTrials.gov: NCT04283461). The high genomic diversity observed in bat betacoronaviruses suggests that development of a pan-betacoronavirus vaccine will be unlikely, however, available technologies for vaccine production and rapid acquisition of genomic information can pave the way for the development of modular vaccine platforms that are rapidly adjustable to new antigens in potentially emerging epidemics [128].

## Concluding remarks

Since the 2002-03 SARS outbreak, genomic information has been crucial to tackle epidemics caused by betacoronaviruses. During the 2019-20 COVID-19 pandemic, quick availability of genomic data has allowed a very rapid, detailed and accurate follow-up of disease progression worldwide and has tremendously supported the development of diagnostic systems, drug candidates and vaccines. Full viral genome analysis has swiftly changed the way scientists deal with epidemic viruses in two main ways. First, the speed that allows the description and classification of the responsible pathogen in a record timeframe, and second, the ability to generate massive amounts of viral genome data, contributing to establishing sound hypotheses on evolution and transmission. It is remarkable that genome analyses at such scale are now increasingly feasible, without having to culture the viruses, many of which are classified as Biosafety Level 3 agents. However, it is important to stress that genomic information must be used carefully when drawing conclusions related to human and animal health. Sampling bias, selection of inadequate bioinformatic tools and misinterpretation of results can all lead to unreliable conclusions. Furthermore, the quality of genomic sequences used in comparative analyses is also crucial to establish sound conclusions. Currently, there are several platforms used for genome sequencing, each of them with their own patterns of systematic sequencing errors. Data curation and normalization is extremely important before conducting further analyses, particularly when comparing data from different sequencing platforms. If analyzed properly, genomic data will indisputably serve as strong basis for addressing future outbreaks caused by highly pathogenic emerging viruses such as SARS-CoV-2.

**Supplementary Materials:** Table S1: Additional information on selected betacoronavirus genomes.

**Author Contributions:** Conceptualization, A.L. and C.M.R.; formal analysis, A.L., C.M.R and Z.C.; investigation, A.L., C.M.R and Z.C.; writing—original draft preparation, A.L., C.M.R, Z.C., S.R., M.A.K and R.L.; writing—review and editing, A.L., C.M.R, Z.C., S.R., M.A.K and R.L. All authors have read and agreed to the published version of the manuscript.



**Funding:** This research received no external funding.

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

## References

1. Cui, J.; Li, F.; Shi, Z.L. Origin and evolution of pathogenic coronaviruses. *Nat. Rev. Microbiol.* **2019**, *17*, 181–192.
2. Sola, I.; Almazán, F.; Zúñiga, S.; Enjuanes, L. Continuous and Discontinuous RNA Synthesis in Coronaviruses. *Annu. Rev. Virol.* **2015**, *2*, 265–288.
3. Li, W.; Moore, M.J.; Vasllieva, N.; Sui, J.; Wong, S.K.; Berne, M.A.; Somasundaran, M.; Sullivan, J.L.; Luzuriaga, K.; Greeneugh, T.C.; et al. Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. *Nature* **2003**, *426*, 450–454.
4. Zhou, P.; Yang, X. Lou; Wang, X.G.; Hu, B.; Zhang, L.; Zhang, W.; Si, H.R.; Zhu, Y.; Li, B.; Huang, C.L.; et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **2020**, *579*, 270–273.
5. Raj, V.S.; Mou, H.; Smits, S.L.; Dekkers, D.H.W.; Müller, M.A.; Dijkman, R.; Muth, D.; Demmers, J.A.A.; Zaki, A.; Fouchier, R.A.M.; et al. Dipeptidyl peptidase 4 is a functional receptor for the emerging human coronavirus-EMC. *Nature* **2013**, *495*, 251–254.
6. Williams, R.K.; Jiang, G.S.; Holmes, K. V. Receptor for mouse hepatitis virus is a member of the carcinoembryonic antigen family of glycoproteins. *Proc. Natl. Acad. Sci. U. S. A.* **1991**, *88*, 5533–5536.
7. Nédellec, P.; Dveksler, G.S.; Daniels, E.; Turbide, C.; Chow, B.; Basile, A.A.; Holmes, K. V.; Beauchemin, N. Bgp2, a new member of the carcinoembryonic antigen-related gene family, encodes an alternative receptor for mouse hepatitis viruses. *J. Virol.* **1994**, *68*, 4525–4537.
8. Schultze, B.; Cavanagh, D.; Herrler, G. Neuraminidase treatment of avian infectious bronchitis coronavirus reveals a hemagglutinating activity that is dependent on sialic acid-containing receptors on erythrocytes. *Virology* **1992**, *189*, 792–794.
9. Fehr, A.R.; Perlman, S. Coronaviruses: An overview of their replication and pathogenesis. *Methods Mol. Biol.* **2015**, *1282*, 1–23.
10. Song, Z.; Xu, Y.; Bao, L.; Zhang, L.; Yu, P.; Qu, Y.; Zhu, H.; Zhao, W.; Han, Y.; Qin, C. From SARS to MERS, Thrusting Coronaviruses into the Spotlight. *Viruses* **2019**, *11*.
11. Li, F.; Li, W.; Farzan, M.; Harrison, S.C. Structural biology: Structure of SARS coronavirus spike receptor-binding domain complexed with receptor. *Science (80-. )*. **2005**, *309*, 1864–1868.
12. Tai, W.; He, L.; Zhang, X.; Pu, J.; Voronin, D.; Jiang, S.; Zhou, Y.; Du, L. Characterization of the receptor-binding domain (RBD) of 2019 novel coronavirus: implication for development of RBD protein as a viral attachment inhibitor and vaccine. *Cell. Mol. Immunol.* **2020**, 1–8.
13. Chen, Y.; Rajashankar, K.R.; Yang, Y.; Agnihothram, S.S.; Liu, C.; Lin, Y.-L.; Baric, R.S.; Li, F. Crystal Structure of the Receptor-Binding Domain from Newly Emerged Middle East Respiratory Syndrome Coronavirus. *J. Virol.* **2013**, *87*, 10777–10783.
14. Wang, N.; Shi, X.; Jiang, L.; Zhang, S.; Wang, D.; Tong, P.; Guo, D.; Fu, L.; Cui, Y.; Liu, X.; et al. Structure of MERS-CoV spike receptor-binding domain complexed with human receptor DPP4. *Cell Res.* **2013**, *23*, 986–993.
15. Peng, G.; Sun, D.; Rajashankar, K.R.; Qian, Z.; Holmes, K. V.; Li, F. Crystal structure of mouse coronavirus receptor-binding domain complexed with its murine receptor. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 10696–10701.



16. Bosch, B.J.; van der Zee, R.; de Haan, C.A.M.; Rottier, P.J.M. The Coronavirus Spike Protein Is a Class I Virus Fusion Protein: Structural and Functional Characterization of the Fusion Core Complex. *J. Virol.* **2003**, *77*, 8801–8811.
17. Sainz, B.; Rausch, J.M.; Gallaher, W.R.; Garry, R.F.; Wimley, W.C. Identification and Characterization of the Putative Fusion Peptide of the Severe Acute Respiratory Syndrome-Associated Coronavirus Spike Protein. *J. Virol.* **2005**, *79*, 7195–7206.
18. Liu, S.; Xiao, G.; Chen, Y.; He, Y.; Niu, J.; Escalante, C.R.; Xiong, H.; Farmar, J.; Debnath, A.K.; Tien, P.; et al. Interaction between heptad repeat 1 and 2 regions in spike protein of SARS-associated coronavirus: Implications for virus fusogenic mechanism and identification of fusion inhibitors. *Lancet* **2004**, *363*, 938–947.
19. Snijder, E.J.; Decroly, E.; Ziebuhr, J. The Nonstructural Proteins Directing Coronavirus RNA Synthesis and Processing. In *Advances in Virus Research*; 2016; Vol. 96, pp. 59–126.
20. Snijder, E.J.; Bredenbeek, P.J.; Dobbe, J.C.; Thiel, V.; Ziebuhr, J.; Poon, L.L.M.; Guan, Y.; Rozanov, M.; Spaan, W.J.M.; Gorbalenya, A.E. Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage. *J. Mol. Biol.* **2003**, *331*, 991–1004.
21. Chatterjee, A.; Johnson, M.A.; Serrano, P.; Pedrini, B.; Joseph, J.S.; Neuman, B.W.; Saikatendu, K.; Buchmeier, M.J.; Kuhn, P.; Wüthrich, K. Nuclear Magnetic Resonance Structure Shows that the Severe Acute Respiratory Syndrome Coronavirus-Unique Domain Contains a Macrodomain Fold. *J. Virol.* **2009**, *83*, 1823–1836.
22. Serrano, P.; Johnson, M.A.; Chatterjee, A.; Neuman, B.W.; Joseph, J.S.; Buchmeier, M.J.; Kuhn, P.; Wüthrich, K. Nuclear Magnetic Resonance Structure of the Nucleic Acid-Binding Domain of Severe Acute Respiratory Syndrome Coronavirus Nonstructural Protein 3. *J. Virol.* **2009**, *83*, 12998–13008.
23. Lehmann, K.C.; Gulyaeva, A.; Zevenhoven-Dobbe, J.C.; Janssen, G.M.C.; Ruben, M.; Overkleeft, H.S.; Van Veelen, P.A.; Samborskiy, D. V.; Kravchenko, A.A.; Leontovich, A.M.; et al. Discovery of an essential nucleotidylating activity associated with a newly delineated conserved domain in the RNA polymerase-containing protein of all nidoviruses. *Nucleic Acids Res.* **2015**, *43*, 8416–8434.
24. Ogando, N.S.; Ferron, F.; Decroly, E.; Canard, B.; Posthuma, C.C.; Snijder, E.J. The Curious Case of the Nidovirus Exoribonuclease: Its Role in RNA Synthesis and Replication Fidelity. *Front. Microbiol.* **2019**, *10*, 1–17.
25. Kirchdoerfer, R.N.; Ward, A.B. Structure of the SARS-CoV nsp12 polymerase bound to nsp7 and nsp8 co-factors. *Nat. Commun.* **2019**, *10*.
26. Seybert, A.; Posthuma, C.C.; van Dinten, L.C.; Snijder, E.J.; Gorbalenya, A.E.; Ziebuhr, J. A Complex Zinc Finger Controls the Enzymatic Activities of Nidovirus Helicases. *J. Virol.* **2005**, *79*, 696–704.
27. Denison, M.R.; Graham, R.L.; Donaldson, E.F.; Eckerle, L.D.; Baric, R.S. Coronaviruses: an RNA proofreading machine regulates replication fidelity and diversity. *RNA Biol.* **2011**, *8*, 270–9.
28. Hurst, K.R.; Kuo, L.; Koetzner, C.A.; Ye, R.; Hsue, B.; Masters, P.S. A Major Determinant for Membrane Protein Interaction Localizes to the Carboxy-Terminal Domain of the Mouse Coronavirus Nucleocapsid Protein. *J. Virol.* **2005**, *79*, 13285–13297.
29. Ye, Y.; Hogue, B.G. Role of the Coronavirus E Viroprotein Protein Transmembrane Domain in Virus Assembly. *J. Virol.* **2007**, *81*, 3597–3607.
30. Neuman, B.W.; Kiss, G.; Kunding, A.H.; Bhella, D.; Baksh, M.F.; Connelly, S.; Droese, B.; Klaus, J.P.; Makino, S.; Sawicki, S.G.; et al. A structural analysis of M protein in coronavirus assembly and morphology. *J. Struct. Biol.* **2011**, *174*, 11–22.

31. Kuo, L.; Hurst-Hess, K.R.; Koetzner, C.A.; Masters, P.S. Analyses of Coronavirus Assembly Interactions with Interspecies Membrane and Nucleocapsid Protein Chimeras. *J. Virol.* **2016**, *90*, 4357–4368.
32. Wilson, L.; Mckinlay, C.; Gage, P.; Ewart, G. SARS coronavirus E protein forms cation-selective ion channels. *Virology* **2004**, *330*, 322–331.
33. Madan, V.; García, M.D.J.; Sanz, M.A.; Carrasco, L. Viroporin activity of murine hepatitis virus E protein. *FEBS Lett.* **2005**, *579*, 3607–3612.
34. Nieto-Torres, J.L.; DeDiego, M.L.; Verdiá-Báguena, C.; Jimenez-Guardeño, J.M.; Regla-Nava, J.A.; Fernandez-Delgado, R.; Castaño-Rodriguez, C.; Alcaraz, A.; Torres, J.; Aguilera, V.M.; et al. Severe Acute Respiratory Syndrome Coronavirus Envelope Protein Ion Channel Activity Promotes Virus Fitness and Pathogenesis. *PLoS Pathog.* **2014**, *10*.
35. Kuo, L.; Koetzner, C.A.; Masters, P.S. A key role for the carboxy-terminal tail of the murine coronavirus nucleocapsid protein in coordination of genome packaging. *Virology* **2016**, *494*, 100–107.
36. Jayaram, H.; Fan, H.; Bowman, B.R.; Ooi, A.; Jayaram, J.; Collisson, E.W.; Lescar, J.; Prasad, B.V.V. X-Ray Structures of the N- and C-Terminal Domains of a Coronavirus Nucleocapsid Protein: Implications for Nucleocapsid Formation. *J. Virol.* **2006**, *80*, 6612–6620.
37. Liu, D.X.; Fung, T.S.; Chong, K.K.L.; Shukla, A.; Hilgenfeld, R. Accessory proteins of SARS-CoV and other coronaviruses. *Antiviral Res.* **2014**, *109*, 97–109.
38. Pewe, L.; Zhou, H.; Netland, J.; Tangudu, C.; Olivares, H.; Shi, L.; Look, D.; Gallagher, T.; Perlman, S. A Severe Acute Respiratory Protein Enhances Virulence of an Attenuated Murine Coronavirus. *J. Virol.* **2005**, *79*, 11335–42.
39. Pfefferle, S.; Krähling, V.; Ditt, V.; Grywna, K.; Mühlberger, E.; Drosten, C. Reverse genetic characterization of the natural genomic deletion in SARS-Coronavirus strain Frankfurt-1 open reading frame 7b reveals an attenuating function of the 7b protein in-vitro and in-vivo. *Virol. J.* **2009**, *6*, 1–17.
40. Fung, S.Y.; Yuen, K.S.; Ye, Z.W.; Chan, C.P.; Jin, D.Y. A tug-of-war between severe acute respiratory syndrome coronavirus 2 and host antiviral defence: lessons from other pathogenic viruses. *Emerg. Microbes Infect.* **2020**, *9*, 558–570.
41. Fischer, F.; Peng, D.; Hingley, S.T.; Weiss, S.R.; Masters, P.S. The internal open reading frame within the nucleocapsid gene of mouse hepatitis virus encodes a structural protein that is not essential for viral replication. *J. Virol.* **1997**, *71*, 996–1003.
42. Senanayake, S.D.; Brian, D.A. Bovine coronavirus I protein synthesis follows ribosomal scanning on the bicistronic N mRNA. *Virus Res.* **1997**, *48*, 101–105.
43. Xu, K.; Zheng, B.J.; Zeng, R.; Lu, W.; Lin, Y.P.; Xue, L.; Li, L.; Yang, L.L.; Xu, C.; Dai, J.; et al. Severe acute respiratory syndrome coronavirus accessory protein 9b is a virion-associated protein. *Virology* **2009**, *388*, 279–285.
44. Shi, C.-S.; Qi, H.-Y.; Boularan, C.; Huang, N.-N.; Abu-Asab, M.; Shelhamer, J.H.; Kehrl, J.H. SARS-Coronavirus Open Reading Frame-9b Suppresses Innate Immunity by Targeting Mitochondria and the MAVS/TRAF3/TRAF6 Signalingosome. *J. Immunol.* **2014**, *193*, 3080–3089.
45. Woo, P.C.Y.; Huang, Y.; Lau, S.K.P.; Yuen, K.Y. Coronavirus genomics and bioinformatics analysis. *Viruses* **2010**, *2*, 1805–1820.
46. Vlasak, R.; Luytjes, W.; Leider, J.; Spaan, W.; Palese, P. The E3 protein of bovine coronavirus is a receptor-destroying enzyme with acetylsterase activity. *J. Virol.* **1988**, *62*, 4686–4690.

47. Zeng, Q.; Langereis, M.A.; Van Vliet, A.L.W.; Huizinga, E.G.; De Groot, R.J. Structure of coronavirus hemagglutinin-esterase offers insight into corona and influenza virus evolution. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 9065–9069.
48. Luytjes, W.; Bredenbeek, P.J.; Noten, A.F.H.; Horzinek, M.C.; Spaan, W.J.M. Sequence of mouse hepatitis virus A59 mRNA 2: Indications for RNA recombination between coronaviruses and influenza C virus. *Virology* **1988**, *166*, 415–422.
49. Woo, P.C.Y.; Lau, S.K.P.; Tsoi, H.; Huang, Y.; Poon, R.W.S.; Chu, C.; Lee, R.A.; Luk, W.; Wong, G.K.M.; Wong, B.H.L.; et al. Clinical and Molecular Epidemiological Features of Coronavirus HKU1–Associated Community-Acquired Pneumonia. *J. Infect. Dis.* **2005**, *192*, 1898–1907.
50. Lau, S.K.P.; Woo, P.C.Y.; Yip, C.C.Y.; Tse, H.; Tsoi, H.W.; Cheng, V.C.C.; Lee, P.; Tang, B.S.F.; Cheung, C.H.Y.; Lee, R.A.; et al. Coronavirus HKU1 and other coronavirus infections in Hong Kong. *J. Clin. Microbiol.* **2006**, *44*, 2063–2071.
51. Woo, P.C.Y.; Lau, S.K.P.; Yip, C.C.Y.; Huang, Y.; Tsoi, H.-W.; Chan, K.-H.; Yuen, K.-Y. Comparative Analysis of 22 Coronavirus HKU1 Genomes Reveals a Novel Genotype and Evidence of Natural Recombination in Coronavirus HKU1. *J. Virol.* **2006**, *80*, 7136–7145.
52. Gorbalenya, A.E.; Baker, S.C.; Baric, R.S.; de Groot, R.J.; Drosten, C.; Gulyaeva, A.A.; Haagmans, B.L.; Lauber, C.; Leontovich, A.M.; Neuman, B.W.; et al. The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. *Nat. Microbiol.* **2020**, *5*, 536–544.
53. Saberi, A.; Gulyaeva, A.A.; Brubacher, J.L.; Newmark, P.A.; Gorbalenya, A.E. A planarian nidovirus expands the limits of RNA genome size. *PLoS Pathog.* **2018**, *14*, 368.
54. Lu, R.; Zhao, X.; Li, J.; Niu, P.; Yang, B.; Wu, H.; Wang, W.; Song, H.; Huang, B.; Zhu, N.; et al. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet* **2020**, *395*, 565–574.
55. Zhang, T.; Wu, Q.; Zhang, Z. Probable Pangolin Origin of SARS-CoV-2 Associated with the COVID-19 Outbreak. *Curr. Biol.* **2020**, *30*, 1346–1351.e2.
56. Katoh, K.; Standley, D.M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **2013**, *30*, 772–80.
57. Nguyen, L.-T.; Schmidt, H.A.; von Haeseler, A.; Minh, B.Q. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **2015**, *32*, 268–74.
58. Suyama, M.; Torrents, D.; Bork, P. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Res.* **2006**, *34*, W609–12.
59. Woo, P.C.Y.; Lau, S.K.P.; Lam, C.S.F.; Lau, C.C.Y.; Tsang, A.K.L.; Lau, J.H.N.; Bai, R.; Teng, J.L.L.; Tsang, C.C.C.; Wang, M.; et al. Discovery of Seven Novel Mammalian and Avian Coronaviruses in the Genus Deltacoronavirus Supports Bat Coronaviruses as the Gene Source of Alphacoronavirus and Betacoronavirus and Avian Coronaviruses as the Gene Source of Gammacoronavirus and Deltacoronavi. *J. Virol.* **2012**, *86*, 3995–4008.
60. Hu, B.; Ge, X.; Wang, L.F.; Shi, Z. Bat origin of human coronaviruses Coronaviruses: Emerging and re-emerging pathogens in humans and animals Susanna Lau Positive-strand RNA viruses. *Virol. J.* **2015**, *12*, 1–10.
61. Forni, D.; Cagliani, R.; Clerici, M.; Sironi, M. Molecular Evolution of Human Coronavirus Genomes. *Trends Microbiol.* **2017**, *25*, 35–48.
62. Li, W.; Shi, Z.; Yu, M.; Ren, W.; Smith, C.; Epstein, J.H.; Wang, H.; Crameri, G.; Hu, Z.; Zhang, H.; et al. Bats are natural reservoirs of SARS-like coronaviruses. *Science (80-. ).* **2005**, *310*, 676–679.

63. Lau, S.K.P.; Woo, P.C.Y.; Li, K.S.M.; Huang, Y.; Tsoi, H.W.; Wong, B.H.L.; Wong, S.S.Y.; Leung, S.Y.; Chan, K.H.; Yuen, K.Y. Severe acute respiratory syndrome coronavirus-like virus in Chinese horseshoe bats. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 14040–14045.
64. Luk, H.K.H.; Li, X.; Fung, J.; Lau, S.K.P.; Woo, P.C.Y. Molecular epidemiology, evolution and phylogeny of SARS coronavirus. *Infect. Genet. Evol.* **2019**, *71*, 21–30.
65. Van Boheemen, S.; Graaf, M. De; Lauber, C.; Bestebroer, T.M.; Raj, V.S.; Zaki, A.M.; Osterhaus, A.D.M.E.; Haagmans, B.L.; Gorbalenya, A.E.; Snijder, E.; et al. Genomic Characterization of a Newly Discovered Coronavirus. *MBio* **2012**, *3*, 1–9.
66. Geldenhuys, M.; Mortlock, M.; Weyer, J.; Bezuidt, O.; Seamark, E.C.J.; Kearney, T.; Gleasner, C.; Erkkila, T.H.; Cui, H.; Markotter, W. A metagenomic viral discovery approach identifies potential zoonotic and novel mammalian viruses in Neoromicia bats within South Africa. *PLoS One* **2018**, *13*, 1–27.
67. Guan, Y.; Zheng, B.J.; He, Y.Q.; Liu, X.L.; Zhuang, Z.X.; Cheung, C.L.; Luo, S.W.; Li, P.H.; Zhang, L.J.; Guan, Y.J.; et al. Isolation and characterization of viruses related to the SARS coronavirus from animals in Southern China. *Science (80-. )*. **2003**, *302*, 276–278.
68. Kan, B.; Wang, M.; Jing, H.; Xu, H.; Jiang, X.; Yan, M.; Liang, W.; Zheng, H.; Wan, K.; Liu, Q.; et al. Molecular Evolution Analysis and Geographic Investigation of Severe Acute Respiratory Syndrome Coronavirus-Like Virus in Palm Civets at an Animal Market and on Farms. *J. Virol.* **2005**, *79*, 11892–11900.
69. Haagmans, B.L.; Al Dhahiry, S.H.S.; Reusken, C.B.E.M.; Raj, V.S.; Galiano, M.; Myers, R.; Godeke, G.J.; Jonges, M.; Farag, E.; Diab, A.; et al. Middle East respiratory syndrome coronavirus in dromedary camels: An outbreak investigation. *Lancet Infect. Dis.* **2014**, *14*, 140–145.
70. Stalin Raj, V.; Farag, E.A.B.A.; Reusken, C.B.E.M.; Lamers, M.M.; Pas, S.D.; Voermans, J.; Smits, S.L.; Osterhaus, A.D.M.E.; Al-Mawlawi, N.; Al-Romaihi, H.E.; et al. Isolation of MERS coronavirus from dromedary camel, Qatar, 2014. *Emerg. Infect. Dis.* **2014**, *20*, 1339–1342.
71. Woo, P.C.Y.; Lau, S.K.P.; Wernery, U.; Wong, E.Y.M.; Tsang, A.K.L.; Johnson, B.; Yip, C.C.Y.; Lau, C.C.Y.; Sivakumar, S.; Cai, J.P.; et al. Novel betacoronavirus in dromedaries of the Middle East, 2013. *Emerg. Infect. Dis.* **2014**, *20*, 560–572.
72. Benvenuto, D.; Giovanetti, M.; Ciccozzi, A.; Spoto, S.; Angeletti, S.; Ciccozzi, M. The 2019-new coronavirus epidemic: Evidence for virus evolution. *J. Med. Virol.* **2020**, *92*, 455–459.
73. Wu, F.; Zhao, S.; Yu, B.; Chen, Y.M.; Wang, W.; Song, Z.G.; Hu, Y.; Tao, Z.W.; Tian, J.H.; Pei, Y.Y.; et al. A new coronavirus associated with human respiratory disease in China. *Nature* **2020**, *579*, 265–269.
74. Lam, T.T.Y.; Shum, M.H.H.; Zhu, H.C.; Tong, Y.G.; Ni, X.B.; Liao, Y.S.; Wei, W.; Cheung, W.Y.M.; Li, W.J.; Li, L.F.; et al. Identifying SARS-CoV-2 related coronaviruses in Malayan pangolins. *Nature* **2020**.
75. Andersen, K.G.; Rambaut, A.; Lipkin, W.I.; Holmes, E.C.; Garry, R.F. The proximal origin of SARS-CoV-2. *Nat. Med.* **2020**, *26*, 450–452.
76. Duffy, S.; Shackelton, L.A.; Holmes, E.C. Rates of evolutionary change in viruses: Patterns and determinants. *Nat. Rev. Genet.* **2008**, *9*, 267–276.
77. Sanjuán, R.; Nebot, M.R.; Chirico, N.; Mansky, L.M.; Belshaw, R. Viral Mutation Rates. *J. Virol.* **2010**, *84*, 9733–9748.
78. Gorbalenya, A.E.; Enjuanes, L.; Ziebuhr, J.; Snijder, E.J. Nidovirales: Evolving the largest RNA virus genome. *Virus Res.* **2006**, *117*, 17–37.
79. Suchard, M.A.; Lemey, P.; Baele, G.; Ayres, D.L.; Drummond, A.J.; Rambaut, A. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evol.* **2018**, *4*, 1–5.

80. Sánchez, C.M.; Gebauer, F.; Suñé, C.; Mendez, A.; Dopazo, J.; Enjuanes, L. Genetic evolution and tropism of transmissible gastroenteritis coronaviruses. *Virology* **1992**, *190*, 92–105.
81. Vijgen, L.; Keyaerts, E.; Moës, E.; Thoelen, I.; Wollants, E.; Lemey, P.; Vandamme, A.-M.; Van Ranst, M. Complete Genomic Sequence of Human Coronavirus OC43: Molecular Clock Analysis Suggests a Relatively Recent Zoonotic Coronavirus Transmission Event. *J. Virol.* **2005**, *79*, 1595–1604.
82. Zhao, Z.; Li, H.; Wu, X.; Zhong, Y.; Zhang, K.; Zhang, Y.P.; Boerwinkle, E.; Fu, Y.X. Moderate mutation rate in the SARS coronavirus genome and its implications. *BMC Evol. Biol.* **2004**, *4*.
83. Cotten, M.; Watson, S.J.; Zumla, A.I.; Makhdoom, H.Q.; Palser, A.L.; Ong, S.H.; Al Rabeeah, A.A.; Alhakeem, R.F.; Assiri, A.; Al-Tawfiq, J.A.; et al. Spread, circulation, and evolution of the Middle East respiratory syndrome coronavirus. *MBio* **2014**, *5*.
84. Song, H.D.; Zhang, G.W.; Cheng, F.; Pan, C.M.; Chen, S.J.; Tu, C.C.; Lei, L.C.; Gao, Y.W.; Xiang, H.; Xuan, H.; et al. Cross-host evolution of severe acute respiratory syndrome coronavirus in palm civet and human. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 2430–2435.
85. Salemi, M.; Fitch, W.M.; Ciccozzi, M.; Ruiz-Alvarez, M.J.; Rezza, G.; Lewis, M.J. Severe Acute Respiratory Syndrome Coronavirus Sequence Characteristics and Evolutionary Rate Estimate from Maximum Likelihood Analysis. *J. Virol.* **2004**, *78*, 1602–1603.
86. Vijaykrishna, D.; Smith, G.J.D.; Zhang, J.X.; Peiris, J.S.M.; Chen, H.; Guan, Y. Evolutionary Insights into the Ecology of Coronaviruses. *J. Virol.* **2007**, *81*, 4012–4020.
87. Hon, C.-C.; Lam, T.-Y.; Shi, Z.-L.; Drummond, A.J.; Yip, C.-W.; Zeng, F.; Lam, P.-Y.; Leung, F.C.-C. Evidence of the Recombinant Origin of a Bat Severe Acute Respiratory Syndrome (SARS)-Like Coronavirus and Its Implications on the Direct Ancestor of SARS Coronavirus. *J. Virol.* **2008**, *82*, 1819–1826.
88. Lau, S.K.P.; Li, K.S.M.; Huang, Y.; Shek, C.-T.; Tse, H.; Wang, M.; Choi, G.K.Y.; Xu, H.; Lam, C.S.F.; Guo, R.; et al. Ecoepidemiology and Complete Genome Comparison of Different Strains of Severe Acute Respiratory Syndrome-Related Rhinolophus Bat Coronavirus in China Reveal Bats as a Reservoir for Acute, Self-Limiting Infection That Allows Recombination Events. *J. Virol.* **2010**, *84*, 2808–2819.
89. Wang, C.; Liu, Z.; Chen, Z.; Huang, X.; Xu, M.; He, T.; Zhang, Z. The establishment of reference sequence for SARS-CoV-2 and variation analysis. *J. Med. Virol.* **2020**, *92*, 667–674.
90. Pachetti, M.; Marini, B.; Benedetti, F.; Giudici, F.; Mauro, E.; Storici, P.; Masciovecchio, C.; Angeletti, S.; Ciccozzi, M.; Gallo, R.C.; et al. Emerging SARS-CoV-2 mutation hot spots include a novel RNA-dependent-RNA polymerase variant. *J. Transl. Med.* **2020**, *18*.
91. Boni, M.F.; Lemey, P.; Jiang, X.; Lam, T.T.-Y.; Perry, B.; Castoe, T.; Rambaut, A.; Robertson, D.L. Evolutionary origins of the SARS-CoV-2 sarbecovirus lineage responsible for the COVID-19 pandemic. *bioRxiv* **2020**, 2020.03.30.015008.
92. Salminen, M.O.; Carr, J.K.; Burke, D.S.; McCutchan, F.E. Identification of breakpoints in intergenotypic recombinants of HIV type 1 by bootscanning. *AIDS Res. Hum. Retroviruses* **1995**, *11*, 1423–5.
93. Lole, K.S.; Bollinger, R.C.; Paranjape, R.S.; Gadkari, D.; Kulkarni, S.S.; Novak, N.G.; Ingersoll, R.; Sheppard, H.W.; Ray, S.C. Full-length human immunodeficiency virus type 1 genomes from subtype C-infected seroconverters in India, with evidence of intersubtype recombination. *J. Virol.* **1999**, *73*, 152–60.
94. Martin, D.P.; Lemey, P.; Lott, M.; Moulton, V.; Posada, D.; Lefeuve, P. RDP3: a flexible and fast computer program for analyzing recombination. *Bioinformatics* **2010**, *26*, 2462–3.
95. Graham, R.L.; Baric, R.S. Recombination, Reservoirs, and the Modular Spike: Mechanisms of Coronavirus Cross-Species Transmission. *J. Virol.* **2010**, *84*, 3134–3146.



96. Su, S.; Wong, G.; Shi, W.; Liu, J.; Lai, A.C.K.; Zhou, J.; Liu, W.; Bi, Y.; Gao, G.F. Epidemiology, Genetic Recombination, and Pathogenesis of Coronaviruses. *Trends Microbiol.* **2016**, *24*, 490–502.
97. Pyrc, K.; Dijkman, R.; Deng, L.; Jebbink, M.F.; Ross, H.A.; Berkhout, B.; van der Hoek, L. Mosaic structure of human coronavirus NL63, one thousand years of evolution. *J. Mol. Biol.* **2006**, *364*, 964–73.
98. Lau, S.K.P.; Lee, P.; Tsang, A.K.L.; Yip, C.C.Y.; Tse, H.; Lee, R.A.; So, L.-Y.; Lau, Y.-L.; Chan, K.-H.; Woo, P.C.Y.; et al. Molecular epidemiology of human coronavirus OC43 reveals evolution of different genotypes over time and recent emergence of a novel genotype due to natural recombination. *J. Virol.* **2011**, *85*, 11325–37.
99. Lau, S.K.P.; Feng, Y.; Chen, H.; Luk, H.K.H.; Yang, W.-H.; Li, K.S.M.; Zhang, Y.-Z.; Huang, Y.; Song, Z.-Z.; Chow, W.-N.; et al. Severe Acute Respiratory Syndrome (SARS) Coronavirus ORF8 Protein Is Acquired from SARS-Related Coronavirus from Greater Horseshoe Bats through Recombination. *J. Virol.* **2015**, *89*, 10532–10547.
100. Wu, Z.; Yang, L.; Ren, X.; Zhang, J.; Yang, F.; Zhang, S.; Jin, Q. ORF8-related genetic evidence for Chinese horseshoe bats as the source of human severe acute respiratory syndrome coronavirus. *J. Infect. Dis.* **2016**, *213*, 579–583.
101. Ge, X.Y.; Li, J.L.; Yang, X. Lou; Chmura, A.A.; Zhu, G.; Epstein, J.H.; Mazet, J.K.; Hu, B.; Zhang, W.; Peng, C.; et al. Isolation and characterization of a bat SARS-like coronavirus that uses the ACE2 receptor. *Nature* **2013**, *503*, 535–538.
102. Yang, X.-L.; Hu, B.; Wang, B.; Wang, M.-N.; Zhang, Q.; Zhang, W.; Wu, L.-J.; Ge, X.-Y.; Zhang, Y.-Z.; Daszak, P.; et al. Isolation and Characterization of a Novel Bat Coronavirus Closely Related to the Direct Progenitor of Severe Acute Respiratory Syndrome Coronavirus. *J. Virol.* **2015**, *90*, 3253–6.
103. Hu, B.; Zeng, L.P.; Yang, X. Lou; Ge, X.Y.; Zhang, W.; Li, B.; Xie, J.Z.; Shen, X.R.; Zhang, Y.Z.; Wang, N.; et al. Discovery of a rich gene pool of bat SARS-related coronaviruses provides new insights into the origin of SARS coronavirus. *PLoS Pathog.* **2017**, *13*, 1–27.
104. Tang, X.; Wu, C.; Li, X.; Song, Y.; Yao, X.; Wu, X.; Duan, Y.; Zhang, H.; Wang, Y.; Qian, Z.; et al. On the origin and continuing evolution of SARS-CoV-2. *Natl. Sci. Rev.* **2020**, nwaa036.
105. Wan, Y.; Shang, J.; Graham, R.; Baric, R.S.; Li, F. Receptor Recognition by the Novel Coronavirus from Wuhan: an Analysis Based on Decade-Long Structural Studies of SARS Coronavirus. *J. Virol.* **2020**, *94*, 1–9.
106. He, J.F.; Peng, G.W.; Min, J.; Yu, D.W.; Liang, W.J.; Zhang, S.Y.; Xu, R.H.; Zheng, H.Y.; Wu, X.W.; Xu, J.; et al. Molecular Evolution of the SARS Coronavirus, during the Course of the SARS Epidemic in China. *Science (80-. )*. **2004**, *303*, 1666–1669.
107. Zhang, C.Y.; Wei, J.F.; He, S.N. Adaptive evolution of the spike gene of SARS coronavirus: Changes in positively selected sites in different epidemic groups. *BMC Microbiol.* **2006**, *6*, 88.
108. Zhang, Z.; Shen, L.; Gu, X. Evolutionary Dynamics of MERS-CoV: Potential Recombination, Positive Selection and Transmission. *Sci. Rep.* **2016**, *6*, 25049.
109. Bolles, M.; Donaldson, E.; Baric, R. SARS-CoV and emergent coronaviruses: Viral determinants of interspecies transmission. *Curr. Opin. Virol.* **2011**, *1*, 624–634.
110. Cagliani, R.; Forni, D.; Clerici, M.; Sironi, M. Computational inference of selection underlying the evolution of the novel coronavirus, SARS-CoV-2. *J. Virol.* **2020**, *In Press*.
111. Lam, T.T.Y.; Hon, C.C.; Tang, J.W. *Use of phylogenetics in the molecular epidemiology and evolutionary studies of viral infections*; 2010; Vol. 47; ISBN 1040836100.

112. Didelot, X.; Gardy, J.; Colijn, C. Bayesian inference of infectious disease transmission from whole-genome sequence data. *Mol. Biol. Evol.* **2014**, *31*, 1869–1879.
113. Klinkenberg, D.; Backer, J.A.; Didelot, X.; Colijn, C.; Wallinga, J. *Simultaneous inference of phylogenetic and transmission trees in infectious disease outbreaks*; 2017; Vol. 13; ISBN 1111111111.
114. Wymant, C.; Hall, M.; Ratmann, O.; Bonsall, D.; Golubchik, T.; De Cesare, M.; Gall, A.; Cornelissen, M.; Fraser, C. PHYLOSCANNER: Inferring transmission from within- and between-host pathogen genetic diversity. *Mol. Biol. Evol.* **2018**, *35*, 719–733.
115. Campbell, F.; Didelot, X.; Fitzjohn, R.; Ferguson, N.; Cori, A.; Jombart, T. outbreaker2: A modular platform for outbreak reconstruction. *BMC Bioinformatics* **2018**, *19*.
116. Jombart, T.; Cori, A.; Didelot, X.; Cauchemez, S.; Fraser, C.; Ferguson, N. Bayesian reconstruction of disease outbreaks by combining epidemiologic and genomic data. *PLoS Comput. Biol.* **2014**, *10*, e1003457.
117. Fraser, C.; Donnelly, C.A.; Cauchemez, S.; Hanage, W.P.; Van Kerkhove, M.D.; Hollingsworth, T.D.; Griffin, J.; Baggaley, R.F.; Jenkins, H.E.; Lyons, E.J.; et al. Pandemic potential of a strain of influenza A (H1N1): Early findings. *Science (80-. )*. **2009**, *324*, 1557–1561.
118. Smith, G.J.D.; Vijaykrishna, D.; Bahl, J.; Lycett, S.J.; Worobey, M.; Pybus, O.G.; Ma, S.K.; Cheung, C.L.; Raghvani, J.; Bhatt, S.; et al. Origins and evolutionary genomics of the 2009 swine-origin H1N1 influenza a epidemic. *Nature* **2009**, *459*, 1122–1125.
119. Elbe, S.; Buckland-Merrett, G. Data, disease and diplomacy: GISAID's innovative contribution to global health. *Glob. Challenges* **2017**, *1*, 33–46.
120. Shu, Y.; McCauley, J. GISAID: Global initiative on sharing all influenza data – from vision to reality. *Eurosurveillance* **2017**, *22*, 2–4.
121. Singer, J.B.; Thomson, E.C.; McLauchlan, J.; Hughes, J.; Gifford, R.J. GLUE: A flexible software system for virus sequence data. *BMC Bioinformatics* **2018**, *19*, 1–18.
122. Forster, P.; Forster, L.; Renfrew, C.; Forster, M. Phylogenetic network analysis of SARS-CoV-2 genomes. *Proc. Natl. Acad. Sci.* **2020**, 202004999.
123. Ruffell, D. Coronavirus SARS-CoV-2: filtering fact from fiction in the infodemic: Q&A with virologist Professor Urs Greber. *FEBS Lett.* **2020**, *594*, 1127–1131.
124. Sánchez-Pacheco, S.J.; Kong, S.; Pulido-Santacruz, P.; Murphy, R.W.; Kubatko, L. Median-joining network analysis of SARS-CoV-2 genomes is neither phylogenetic nor evolutionary. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *0*, 4–5.
125. Mavian, C.; Pond, S.K.; Marini, S.; Magalis, B.R.; Vandamme, A.-M.; Dellicour, S.; Scarpino, S. V.; Houldcroft, C.; Villabona-Arenas, J.; Paisie, T.K.; et al. Sampling bias and incorrect rooting make phylogenetic network tracing of SARS-COV-2 infections unreliable. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *0*, 2–3.
126. Chookajorn, T. Evolving COVID-19 conundrum and its impact. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *0*, 4–5.
127. Cockrell, A.S.; Beall, A.; Yount, B.; Baric, R. Efficient Reverse Genetic Systems for Rapid Genetic Manipulation of Emergent and Preemergent Infectious Coronaviruses. *Methods Mol. Biol.* **2017**, *1602*, 59–81.
128. Totura, A.L.; Bavari, S. Broad-spectrum coronavirus antiviral drug discovery. *Expert Opin. Drug Discov.* **2019**, *14*, 397–412.
129. Almazán, F.; Sola, I.; Zuñiga, S.; Marquez-Jurado, S.; Morales, L.; Becares, M.; Enjuanes, L. Coronavirus reverse genetic systems: Infectious clones and replicons. *Virus Res.* **2014**, *194*, 67–75.

130. Yount, B.; Curtis, K.M.; Baric, R.S. Strategy for Systematic Assembly of Large RNA and DNA Genomes: Transmissible Gastroenteritis Virus Model. *J. Virol.* **2000**, *74*, 10600–10611.
131. Scobey, T.; Yount, B.L.; Sims, A.C.; Donaldson, E.F.; Agnihothram, S.S.; Menachery, V.D.; Graham, R.L.; Swanstrom, J.; Bove, P.F.; Kim, J.D.; et al. Reverse genetics with a full-length infectious cDNA of the Middle East respiratory syndrome coronavirus. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 16157–16162.
132. Xie, X.; Muruato, A.; Lokugamage, K.G.; Narayanan, K.; Zhang, X.; Zou, J.; Liu, J.; Schindewolf, C.; Bopp, N.E.; Aguilar, P. V.; et al. An Infectious cDNA Clone of SARS-CoV-2. *Cell Host Microbe* **2020**, 1–8.
133. Yount, B.; Curtis, K.M.; Fritz, E.A.; Hensley, L.E.; Jahrling, P.B.; Prentice, E.; Denison, M.R.; Geisbert, T.W.; Baric, R.S. Reverse genetics with a full-length infectious cDNA of severe acute respiratory syndrome coronavirus. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 12995–13000.
134. Sims, A.C.; Baric, R.S.; Yount, B.; Burkett, S.E.; Collins, P.L.; Pickles, R.J. Severe Acute Respiratory Syndrome Coronavirus Infection of Human Ciliated Airway Epithelia: Role of Ciliated Cells in Viral Spread in the Conducting Airways of the Lungs. *J. Virol.* **2005**, *79*, 15511–15524.
135. Yount, B.; Roberts, R.S.; Lindesmith, L.; Baric, R.S. Rewiring the severe acute respiratory syndrome coronavirus (SARS-CoV) transcription circuit: Engineering a recombination-resistant genome. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 12546–12551.
136. Frieman, M.; Yount, B.; Agnihothram, S.; Page, C.; Donaldson, E.; Roberts, A.; Vogel, L.; Woodruff, B.; Scorpio, D.; Subbarao, K.; et al. Molecular Determinants of Severe Acute Respiratory Syndrome Coronavirus Pathogenesis and Virulence in Young and Aged Mouse Models of Human Disease. *J. Virol.* **2012**, *86*, 884–897.
137. Cockrell, A.S.; Yount, B.L.; Scobey, T.; Jensen, K.; Douglas, M.; Beall, A.; Tang, X.C.; Marasco, W.A.; Heise, M.T.; Baric, R.S. A mouse model for MERS coronavirus-induced acute respiratory distress syndrome. *Nat. Microbiol.* **2016**, *2*, 1–11.
138. Roberts, A.; Deming, D.; Paddock, C.D.; Cheng, A.; Yount, B.; Vogel, L.; Herman, B.D.; Sheahan, T.; Heise, M.; Genrich, G.L.; et al. A mouse-adapted SARS-coronavirus causes disease and mortality in BALB/c mice. *PLoS Pathog.* **2007**, *3*, 0023–0037.
139. Day, C.W.; Baric, R.; Cai, S.X.; Frieman, M.; Kumaki, Y.; Morrey, J.D.; Smee, D.F.; Barnard, D.L. A new mouse-adapted strain of SARS-CoV as a lethal model for evaluating antiviral agents in vitro and in vivo. *Virology* **2009**, *395*, 210–222.
140. Sheahan, T.; Rockx, B.; Donaldson, E.; Sims, A.; Pickles, R.; Corti, D.; Baric, R. Mechanisms of Zoonotic Severe Acute Respiratory Syndrome Coronavirus Host Range Expansion in Human Airway Epithelium. *J. Virol.* **2008**, *82*, 2274–2285.
141. Rockx, B.; Sheahan, T.; Donaldson, E.; Harkema, J.; Sims, A.; Heise, M.; Pickles, R.; Cameron, M.; Kelvin, D.; Baric, R. Synthetic Reconstruction of Zoonotic and Early Human Severe Acute Respiratory Syndrome Coronavirus Isolates That Produce Fatal Disease in Aged Mice. *J. Virol.* **2007**, *81*, 7410–7423.
142. Rockx, B.; Feldmann, F.; Brining, D.; Gardner, D.; LaCasse, R.; Kercher, L.; Long, D.; Rosenke, R.; Virtaneva, K.; Sturdevant, D.E.; et al. Comparative pathogenesis of three human and zoonotic sars-cov strains in cynomolgus macaques. *PLoS One* **2011**, *6*.
143. Becker, M.M.; Graham, R.L.; Donaldson, E.F.; Rockx, B.; Sims, A.C.; Sheahan, T.; Pickles, R.J.; Corti, D.; Johnston, R.E.; Baric, R.S.; et al. Synthetic recombinant bat SARS-like coronavirus is infectious in cultured cells and in mice. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 19944–19949.
144. Agnihothram, S.; Yount, B.L.; Donaldson, E.F. A Mouse Model for Betacoronavirus Subgroup 2c Using a Bat. *MBio* **2014**, *5*, 1–12.

145. Menachery, V.D.; Yount, B.L.; Debbink, K.; Agnihothram, S.; Gralinski, L.E.; Plante, J.A.; Graham, R.L.; Scobey, T.; Ge, X.Y.; Donaldson, E.F.; et al. A SARS-like cluster of circulating bat coronaviruses shows potential for human emergence. *Nat. Med.* **2015**, *21*, 1508–1513.
146. Menachery, V.D.; Yount, B.L.; Sims, A.C.; Debbink, K.; Agnihothram, S.S.; Gralinski, L.E.; Graham, R.L.; Scobey, T.; Plante, J.A.; Royal, S.R.; et al. SARS-like WIV1-CoV poised for human emergence. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, 3048–3053.
147. Loeffelholz, M.J.; Tang, Y.W. Laboratory diagnosis of emerging human coronavirus infections—the state of the art. *Emerg. Microbes Infect.* **2020**, *9*, 747–756.
148. Corman, V.M.; Landt, O.; Kaiser, M.; Molenkamp, R.; Meijer, A.; Chu, D.K.; Bleicker, T.; Brünink, S.; Schneider, J.; Schmidt, M.L.; et al. Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. *Eurosurveillance* **2020**, *25*, 1.
149. Zaki, A.M.; van Boheemen, S.; Bestebroer, T.M.; Osterhaus, A.D.M.E.; Fouchier, R.A.M. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. *N. Engl. J. Med.* **2012**, *367*, 1814–1820.
150. Briese, T.; Mishra, N.; Jain, K.; East, M.; Syndrome, R.; Quasispecies, C.; Include, T.; Revealed, H.I.; Analysis, W.; Cultured, V.; et al. Dromedary Camels in Saudi Arabia Include Homologues of Human Isolates Revealed through Whole-Genome analysis etc. *MBio* **2014**, *5*, 1–5.
151. Hui, R.K.H.; Zeng, F.; Chan, C.M.N.; Yuen, K.Y.; Peiris, J.S.M.; Leung, F.C.C. Reverse Transcriptase PCR Diagnostic Assay for the Coronavirus Associated with Severe Acute Respiratory Syndrome. *J. Clin. Microbiol.* **2004**, *42*, 1994–1999.
152. Corman, V.M.; Eckerle, I.; Bleicker, T.; Zaki, A.; Landt, O.; Eschbach-Bludau, M.; van Boheemen, S.; Gopal, R.; Ballhause, M.; Bestebroer, T.M.; et al. Detection of a novel human coronavirus by real-time reverse-transcription polymerase chain reaction. *Eurosurveillance* **2012**, *17*, 1–6.
153. Lu, X.; Whitaker, B.; Sakthivel, S.K.K.; Kamili, S.; Rose, L.E.; Lowe, L.; Mohareb, E.; Ellassal, E.M.; Al-sanouri, T.; Haddadin, A.; et al. Real-time reverse transcription-pcr assay panel for middle east respiratory syndrome coronavirus. *J. Clin. Microbiol.* **2014**, *52*, 67–75.
154. Holshue, M.L.; DeBolt, C.; Lindquist, S.; Lofy, K.H.; Wiesman, J.; Bruce, H.; Spitters, C.; Ericson, K.; Wilkerson, S.; Tural, A.; et al. First case of 2019 novel coronavirus in the United States. *N. Engl. J. Med.* **2020**, *382*, 929–936.
155. Wang, Y.; Kang, H.; Liu, X.; Tong, Z. Combination of RT-qPCR testing and clinical features for diagnosis of COVID-19 facilitates management of SARS-CoV-2 outbreak. *J. Med. Virol.* **2020**, 538–539.
156. Huang, P.; Wang, H.; Cao, Z.; Jin, H.; Chi, H.; Zhao, J.; Yu, B.; Yan, F.; Hu, X.; Wu, F.; et al. A rapid and specific assay for the detection of MERS-CoV. *Front. Microbiol.* **2018**, *9*, 1–9.
157. Broughton, J.P.; Deng, X.; Yu, G.; Fasching, C.L.; Servellita, V.; Singh, J.; Miao, X.; Streithorst, J.A.; Granados, A.; Sotomayor-Gonzalez, A.; et al. CRISPR–Cas12-based detection of SARS-CoV-2. *Nat. Biotechnol.* **2020**, 10.1038/s4.
158. Irigoyen, N.; Firth, A.E.; Jones, J.D.; Chung, B.Y.W.; Siddell, S.G.; Brierley, I. High-Resolution Analysis of Coronavirus Gene Expression by RNA Sequencing and Ribosome Profiling. *PLoS Pathog.* **2016**, *12*.
159. Lau, S.K.P.; Woo, P.C.Y.; Wong, B.H.L.; Tsoi, H.W.; Woo, G.K.S.; Poon, R.W.S.; Chan, K.H.; Wei, W.I.; Malik Peiris, J.S.; Yuen, K.Y. Detection of severe acute respiratory syndrome (SARS) coronavirus nucleocapsid protein in SARS patients by enzyme-linked immunosorbent assay. *J. Clin. Microbiol.* **2004**, *42*, 2884–2889.

160. Qiu, M.; Shi, Y.; Guo, Z.; Chen, Z.; He, R.; Chen, R.; Zhou, D.; Dai, E.; Wang, X.; Si, B.; et al. Antibody responses to individual proteins of SARS coronavirus and their neutralization activities. *Microbes Infect.* **2005**, *7*, 882–889.
161. Chen, Y.; Chan, K.H.; Kang, Y.; Chen, H.; Luk, H.K.H.; Poon, R.W.S.; Chan, J.F.W.; Yuen, K.Y.; Xia, N.; Lau, S.K.P.; et al. A sensitive and specific antigen detection assay for Middle East respiratory syndrome coronavirus. *Emerg. Microbes Infect.* **2015**, *4*, e26.
162. Li, Z.; Yi, Y.; Luo, X.; Xiong, N.; Liu, Y.; Li, S.; Sun, R.; Wang, Y.; Hu, B.; Chen, W.; et al. Development and clinical application of a rapid IgM-IgG combined antibody test for SARS-CoV-2 infection diagnosis. *J. Med. Virol.* **2020**.
163. Potocnakova, L.; Bhide, M.; Pulzova, L.B. An Introduction to B-Cell Epitope Mapping and in Silico Epitope Prediction. *J. Immunol. Res.* **2016**, *2016*, 6760830.
164. Grifoni, A.; Sidney, J.; Zhang, Y.; Scheuermann, R.H.; Peters, B.; Sette, A. A Sequence Homology and Bioinformatic Approach Can Predict Candidate Targets for Immune Responses to SARS-CoV-2. *Cell Host Microbe* **2020**, *27*, 671–680.e2.
165. Abbott, W.M.; Damschroder, M.M.; Lowe, D.C. Current approaches to fine mapping of antigen-antibody interactions. *Immunology* **2014**, *142*, 526–535.
166. Cheng, M.P.; Papenburg, J.; Desjardins, M.; Kanjilal, S.; Quach, C.; Libman, M.; Dittrich, S.; Yansouni, C.P. Diagnostic Testing for Severe Acute Respiratory Syndrome–Related Coronavirus-2. *Ann. Intern. Med.* **2020**, M20-1301.
167. Harvey, R.; Mattiuzzo, G.; Hassall, M.; Sieberg, A.; Müller, M.A.; Drosten, C.; Rigsby, P.; Oxenford, C.J.; Caly, L.; Li, C.; et al. Comparison of serologic assays for Middle East respiratory syndrome coronavirus. *Emerg. Infect. Dis.* **2019**, *25*, 1878–1883.
168. Zhang, W.; Zheng, X.S.; Agwanda, B.; Ommeh, S.; Zhao, K.; Lichoti, J.; Wang, N.; Chen, J.; Li, B.; Yang, X. Lou; et al. Serological evidence of MERS-CoV and HKU8-related CoV co-infection in Kenyan camels. *Emerg. Microbes Infect.* **2019**, *8*, 1528–1534.
169. Hu, Z.; Song, C.; Xu, C.; Jin, G.; Chen, Y.; Xu, X.; Ma, H.; Chen, W.; Lin, Y.; Zheng, Y.; et al. Clinical characteristics of 24 asymptomatic infections with COVID-19 screened among close contacts in Nanjing, China. *Sci. China Life Sci.* **2020**, *63*, 706–711.
170. Li, R.; Pei, S.; Chen, B.; Song, Y.; Zhang, T.; Yang, W.; Shaman, J. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV2). *Science* **2020**, *3221*, 1–9.
171. Ye, F.; Xu, S.; Rong, Z.; Xu, R.; Liu, X.; Deng, P.; Liu, H.; Xu, X. Delivery of infection from asymptomatic carriers of COVID-19 in a familial cluster. *Int. J. Infect. Dis.* **2020**, *94*, 133–138.
172. Zumla, A.; Chan, J.F.W.; Azhar, E.I.; Hui, D.S.C.; Yuen, K.Y. Coronaviruses-drug discovery and therapeutic options. *Nat. Rev. Drug Discov.* **2016**, *15*, 327–347.
173. Hilgenfeld, R.; Peiris, M. From SARS to MERS: 10 years of research on highly pathogenic human coronaviruses. *Antiviral Res.* **2013**, *100*, 286–295.
174. Liu, C.; Zhou, Q.; Li, Y.; Garner, L. V.; Watkins, S.P.; Carter, L.J.; Smoot, J.; Gregg, A.C.; Daniels, A.D.; Jervey, S.; et al. Research and Development on Therapeutic Agents and Vaccines for COVID-19 and Related Human Coronavirus Diseases. *ACS Cent. Sci.* **2020**, *6*, 315–331.
175. Li, B.J.; Tang, Q.; Cheng, D.; Qin, C.; Xie, F.Y.; Wei, Q.; Xu, J.; Liu, Y.; Zheng, B.J.; Woodle, M.C.; et al. Using siRNA in prophylactic and therapeutic regimens against SARS coronavirus in Rhesus macaque. *Nat. Med.* **2005**, *11*, 944–951.



176. Chen, W.; Feng, P.; Liu, K.; Wu, M.; Lin, H. Computational Identification of Small Interfering RNA Targets in SARS-CoV-2. *Virol. Sin.* **2020**.
177. Zhou, Y.; Hou, Y.; Shen, J.; Huang, Y.; Martin, W.; Cheng, F. Network-based drug repurposing for novel coronavirus 2019-nCoV/SARS-CoV-2. *Cell Discov.* **2020**, *6*.
178. Kim, D.; Lee, J.; Yang, J.; Kim, J.W.; Kim, V.N.; Chang, H. The architecture of SARS-CoV-2 transcriptome. *Cell* **2020**, *181*, 1–8.
179. Gordon, D.E.; Jang, G.M.; Bouhaddou, M.; Xu, J.; Obernier, K.; White, K.M.; O'Meara, M.J.; Rezelj, V. V.; Guo, J.Z.; Swaney, D.L.; et al. A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature* **2020**, 1–13.
180. Roper, R.L.; Rehm, K.E. SARS vaccines: Where are we? *Expert Rev. Vaccines* **2009**, *8*, 887–898.
181. Menachery, V.D.; Yount, B.L.; Josset, L.; Gralinski, L.E.; Scobey, T.; Agnihothram, S.; Katze, M.G.; Baric, R.S. Attenuation and Restoration of Severe Acute Respiratory Syndrome Coronavirus Mutant Lacking 2'-O-Methyltransferase Activity. *J. Virol.* **2014**, *88*, 4251–4264.
182. Menachery, V.D.; Gralinski, L.E.; Mitchell, H.D.; Dinno, K.H.; Leist, S.R.; Yount, B.L.; Graham, R.L.; McAnarney, E.T.; Stratton, K.G.; Cockrell, A.S.; et al. Middle East Respiratory Syndrome Coronavirus Nonstructural Protein 16 Is Necessary for Interferon Resistance and Viral Pathogenesis. *mSphere* **2017**, *2*, 1–12.
183. DeDiego, M.L.; Álvarez, E.; Almazán, F.; Rejas, M.T.; Lamirande, E.; Roberts, A.; Shieh, W.-J.; Zaki, S.R.; Subbarao, K.; Enjuanes, L. A Severe Acute Respiratory Syndrome Coronavirus That Lacks the E Gene Is Attenuated In Vitro and In Vivo. *J. Virol.* **2007**, *81*, 1701–1713.
184. Netland, J.; DeDiego, M.L.; Zhao, J.; Fett, C.; Álvarez, E.; Nieto-Torres, J.L.; Enjuanes, L.; Perlman, S. Immunization with an attenuated severe acute respiratory syndrome coronavirus deleted in E protein protects against lethal respiratory disease. *Virology* **2010**, *399*, 120–128.
185. Almazán, F.; DeDiego, M.L.; Sola, I.; Zuñiga, S.; Nieto-Torres, J.L.; Marquez-Jurado, S.; Andrés, G.; Enjuanes, L. Engineering a replication-competent, propagation-defective middle east respiratory syndrome coronavirus as a vaccine candidate. *MBio* **2013**, *4*.
186. Fett, C.; DeDiego, M.L.; Regla-Nava, J.A.; Enjuanes, L.; Perlman, S. Complete Protection against Severe Acute Respiratory Syndrome Coronavirus-Mediated Lethal Respiratory Disease in Aged Mice by Immunization with a Mouse-Adapted Virus Lacking E Protein. *J. Virol.* **2013**, *87*, 6551–6559.
187. Graham, R.L.; Becker, M.M.; Eckerle, L.D.; Bolles, M.; Denison, M.R.; Baric, R.S. A live, impaired-fidelity coronavirus vaccine protects in an aged, immunocompromised mouse model of lethal disease. *Nat. Med.* **2012**, *18*, 1820–1826.
188. Gao, W.; Tamin, A.; Soloff, A.; D'Aiuto, L.; Nwanegbo, E.; Robbins, P.D.; Bellini, W.J.; Barratt-Boyes, S.; Gambotto, A. Effects of a SARS-associated coronavirus vaccine in monkeys. *Lancet* **2003**, *362*, 1895–1896.
189. Hogan, R.J.; Gao, G.; Rowe, T.; Bell, P.; Flieder, D.; Paragas, J.; Kobinger, G.P.; Wivel, N.A.; Crystal, R.G.; Boyer, J.; et al. Resolution of Primary Severe Acute Respiratory Syndrome-Associated Coronavirus Infection Requires Stat1. *J. Virol.* **2004**, *78*, 11416–11421.
190. Zakhartchouk, A.N.; Viswanathan, S.; Mahony, J.B.; Glaudei, J.; Babiuk, L.A. Severe acute respiratory syndrome coronavirus nucleocapsid protein expressed by an adenovirus vector is phosphorylated and immunogenic in mice. *J. Gen. Virol.* **2005**, *86*, 211–215.

191. See, R.H.; Zakhartchouk, A.N.; Petric, M.; Lawrence, D.J.; Mok, C.P.Y.; Hogan, R.J.; Rowe, T.; Zitzow, L.A.; Karunakaran, K.P.; Hitt, M.M.; et al. Comparative evaluation of two severe acute respiratory syndrome (SARS) vaccine candidates in mice challenged with SARS coronavirus. *J. Gen. Virol.* **2006**, *87*, 641–650.
192. See, R.H.; Petric, M.; Lawrence, D.J.; Mok, C.P.Y.; Rowe, T.; Zitzow, L.A.; Karunakaran, K.P.; Voss, T.G.; Brunham, R.C.; Gauldie, J.; et al. Severe acute respiratory syndrome vaccine efficacy in ferrets: Whole killed virus and adenovirus-vectored vaccines. *J. Gen. Virol.* **2008**, *89*, 2136–2146.
193. Kim, E.; Okada, K.; Kenniston, T.; Raj, V.S.; AlHajri, M.M.; Farag, E.A.B.A.; AlHajri, F.; Osterhaus, A.D.M.E.; Haagmans, B.L.; Gambotto, A. Immunogenicity of an adenoviral-based Middle East Respiratory Syndrome coronavirus vaccine in BALB/c mice. *Vaccine* **2014**, *32*, 5975–5982.
194. Guo, X.; Deng, Y.; Chen, H.; Lan, J.; Wang, W.; Zou, X.; Hung, T.; Lu, Z.; Tan, W. Systemic and mucosal immunity in mice elicited by a single immunization with human adenovirus type 5 or 41 vector-based vaccines carrying the spike protein of Middle East respiratory syndrome coronavirus. *Immunology* **2015**, *145*, 476–484.
195. Bukreyev, A.; Lamirande, E.W.; Buchholz, U.J.; Vogel, L.N.; Elkins, W.R.; St Claire, M.; Murphy, B.R.; Subbarao, K.; Collins, P.L. Mucosal immunisation of African green monkeys (*Cercopithecus aethiops*) with an attenuated parainfluenza virus expressing the SARS coronavirus spike protein for the prevention of SARS. *Lancet* **2004**, *363*, 2122–2127.
196. Kapadia, S.U.; Simon, I.D.; Rose, J.K. SARS vaccine based on a replication-defective recombinant vesicular stomatitis virus is more potent than one based on a replication-competent vector. *Virology* **2008**, *376*, 165–172.
197. Liniger, M.; Zuniga, A.; Tamin, A.; Azzouz-Morin, T.N.; Knuchel, M.; Marty, R.R.; Wiegand, M.; Weibel, S.; Kelvin, D.; Rota, P.A.; et al. Induction of neutralising antibodies and cellular immune responses against SARS coronavirus by recombinant measles viruses. *Vaccine* **2008**, *26*, 2164–2174.
198. Bai, B.; Lu, X.; Meng, J.; Hu, Q.; Mao, P.; Lu, B.; Chen, Z.; Yuan, Z.; Wang, H. Vaccination of mice with recombinant baculovirus expressing spike or nucleocapsid protein of SARS-like coronavirus generates humoral and cellular immune responses. *Mol. Immunol.* **2008**, *45*, 868–875.
199. Czub, M.; Weingartl, H.; Czub, S.; He, R.; Cao, J. Evaluation of modified vaccinia virus Ankara based recombinant SARS vaccine in ferrets. *Vaccine* **2005**, *23*, 2273–2279.
200. Song, F.; Fux, R.; Provacia, L.B.; Volz, A.; Eickmann, M.; Becker, S.; Osterhaus, A.D.M.E.; Haagmans, B.L.; Sutter, G. Middle East Respiratory Syndrome Coronavirus Spike Protein Delivered by Modified Vaccinia Virus Ankara Efficiently Induces Virus-Neutralizing Antibodies. *J. Virol.* **2013**, *87*, 11950–11954.
201. Volz, A.; Kupke, A.; Song, F.; Jany, S.; Fux, R.; Shams-Eldin, H.; Schmidt, J.; Becker, C.; Eickmann, M.; Becker, S.; et al. Protective Efficacy of Recombinant Modified Vaccinia Virus Ankara Delivering Middle East Respiratory Syndrome Coronavirus Spike Glycoprotein. *J. Virol.* **2015**, *89*, 8651–8656.
202. Luo, F.; Feng, Y.; Liu, M.; Li, P.; Pan, Q.; Jeza, V.T.; Xia, B.; Wu, J.; Zhang, X.L. Type IVB pilus operon promoter controlling expression of the severe acute respiratory syndrome-associated coronavirus nucleocapsid gene in *Salmonella enterica* serovar Typhi elicits full immune response by intranasal vaccination. *Clin. Vaccine Immunol.* **2007**, *14*, 990–997.
203. Yang, Z.Y.; Kong, W.P.; Huang, Y.; Roberts, A.; Murphy, B.R.; Subbarao, K.; Nabel, G.J. A DNA vaccine induces SARS coronavirus neutralization and protective immunity in mice. *Nature* **2004**, *428*, 561–564.

204. Zeng, F.; Chow, K.Y.C.; Hon, C.C.; Law, K.M.; Yip, C.W.; Chan, K.H.; Peiris, J.S.M.; Leung, F.C.C. Characterization of humoral responses in mice immunized with plasmid DNAs encoding SARS-CoV spike gene fragments. *Biochem. Biophys. Res. Commun.* **2004**, *315*, 1134–1139.
205. Wang, Z.; Yuan, Z.; Matsumoto, M.; Hengge, U.R.; Chang, Y.F. Immune responses with DNA vaccines encoded different gene fragments of severe acute respiratory syndrome coronavirus in BALB/c mice. *Biochem. Biophys. Res. Commun.* **2005**, *327*, 130–135.
206. Wang, L.; Shi, W.; Joyce, M.G.; Modjarrad, K.; Zhang, Y.; Leung, K.; Lees, C.R.; Zhou, T.; Yassine, H.M.; Kanekiyo, M.; et al. Evaluation of candidate vaccine approaches for MERS-CoV. *Nat. Commun.* **2015**, *6*.
207. Wang, X.; Xu, W.; Tong, D.; Ni, J.; Gao, H.; Wang, Y.; Chu, Y.; Li, P.; Yang, X.; Xiong, S. A chimeric multi-epitope DNA vaccine elicited specific antibody response against severe acute respiratory syndrome-associated coronavirus which attenuated the virulence of SARS-CoV in vitro. *Immunol. Lett.* **2008**, *119*, 71–77.
208. Zhu, M.S.; Pan, Y.; Chen, H.Q.; Shen, Y.; Wang, X.C.; Sun, Y.J.; Tao, K.H. Induction of SARS-nucleoprotein-specific immune response by use of DNA vaccine. *Immunol. Lett.* **2004**, *92*, 237–243.
209. Bolles, M.; Deming, D.; Long, K.; Agnihothram, S.; Whitmore, A.; Ferris, M.; Funkhouser, W.; Gralinski, L.; Totura, A.; Heise, M.; et al. A Double-Inactivated Severe Acute Respiratory Syndrome Coronavirus Vaccine Provides Incomplete Protection in Mice and Induces Increased Eosinophilic Proinflammatory Pulmonary Response upon Challenge. *J. Virol.* **2011**, *85*, 12201–12215.
210. Sheahan, T.; Whitmore, A.; Long, K.; Ferris, M.; Rockx, B.; Funkhouser, W.; Donaldson, E.; Gralinski, L.; Collier, M.; Heise, M.; et al. Successful Vaccination Strategies That Protect Aged Mice from Lethal Challenge from Influenza Virus and Heterologous Severe Acute Respiratory Syndrome Coronavirus. *J. Virol.* **2011**, *85*, 217–230.
211. Agrawal, A.S.; Tao, X.; Algaissi, A.; Garron, T.; Narayanan, K.; Peng, B.H.; Couch, R.B.; Tseng, C.T.K. Immunization with inactivated Middle East Respiratory Syndrome coronavirus vaccine leads to lung immunopathology on challenge with live virus. *Hum. Vaccines Immunother.* **2016**, *12*, 2351–2356.
212. Wang, Q.; Zhang, L.; Kuwahara, K.; Li, L.; Liu, Z.; Li, T.; Zhu, H.; Liu, J.; Xu, Y.; Xie, J.; et al. Immunodominant SARS coronavirus epitopes in humans elicited both enhancing and neutralizing effects on infection in non-human primates. *ACS Infect. Dis.* **2016**, *2*, 361–376.
213. Houser, K. V.; Broadbent, A.J.; Gretebeck, L.; Vogel, L.; Lamirande, E.W.; Sutton, T.; Bock, K.W.; Minai, M.; Orandle, M.; Moore, I.N.; et al. Enhanced inflammation in New Zealand white rabbits when MERS-CoV reinfection occurs in the absence of neutralizing antibody. *PLoS Pathog.* **2017**, *13*, 1–25.
214. Huang, J.; Cao, Y.; Du, J.; Bu, X.; Ma, R.; Wu, C. Priming with SARS CoV S DNA and boosting with SARS CoV S epitopes specific for CD4+ and CD8+ T cells promote cellular immune responses. *Vaccine* **2007**, *25*, 6981–6991.
215. Amanat, F.; Krammer, F. Perspective SARS-CoV-2 Vaccines : Status Report. *Immunity* **2020**, *52*, 583–589.