Check for updates

Coronavirus, the King Who Wanted More Than a Crown: From Common to the Highly Pathogenic SARS-CoV-2, Is the Key in the Accessory Genes?

Nathalie Chazal*

Institut de Recherche en Infectiologie de Montpellier (IRIM), Université de Montpellier, CNRS, Montpellier, France

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), that emerged in late 2019, is the etiologic agent of the current "coronavirus disease 2019" (COVID-19) pandemic, which has serious health implications and a significant global economic impact. Of the seven human coronaviruses, all of which have a zoonotic origin, the pandemic SARS-CoV-2, is the third emerging coronavirus, in the 21st century, highly pathogenic to the human population. Previous human coronavirus outbreaks (SARS-CoV-1 and MERS-CoV) have already provided several valuable information on some of the common molecular and cellular mechanisms of coronavirus infections as well as their origin. However, to meet the new challenge caused by the SARS-CoV-2, a detailed understanding of the biological specificities, as well as knowledge of the origin are crucial to provide information on viral pathogenicity, transmission and epidemiology, and to enable strategies for therapeutic interventions and drug discovery. Therefore, in this review, we summarize the current advances in SARS-CoV-2 knowledges, in light of pre-existing information of other recently emerging coronaviruses. We depict the specificity of the immune response of wild bats and discuss current knowledge of the genetic diversity of bat-hosted coronaviruses that promotes viral genome expansion (accessory gene acquisition). In addition, we describe the basic virology of coronaviruses with a special focus SARS-CoV-2. Finally, we highlight, in detail, the current knowledge of genes and accessory proteins which we postulate to be the major keys to promote virus adaptation to specific hosts (bat and human), to contribute to the suppression of immune responses, as well as to pathogenicity.

Keywords: SARS-CoV-2, origin, biology, accessory proteins, evolution

CORONAVIRUSES: TAXONOMY AND HISTORICALLY PERSPECTIVES

Inside the *Nidovirales* order and in the sub-order *Cornidovirineae*, the highly diverse *coronaviridae* family of enveloped positive-sense single-stranded RNA viruses is divided in two sub-families *Letovirinae* and *Orthocoronavirinae*. The sub-family *Orthocoronavirinae* is composed of four genera: *alphacoronaviruses* (fourteen sub-genera), *betacoronavirus* (five sub-genera),

gammacoronavirus (three sub-genera) and *deltacoronavirus* (three sub-genera) (**Figure 1**). The human coronaviruses HCoV-229E and HCoV-NL63 are members of *alphacoronavirus* (alpha-CoVs); the Murine coronavirus (previously named Mouse Hepatitis Virus, MHV), the two human coronaviruses HCoV-OC43 and HKU1, the Severe Acute Respiratory Syndrome (SARS)-Cororonavirus (CoV)-1, the Middle East Respiratory Syndrome Coronavirus (MERS)-CoV, the bat SARS-CoV-like coronaviruses and the SARS-CoV-2 are included in the *betacoronavirus genus* (beta-CoVs); the Goose coronavirus CB17, the Avian coronavirus, the Avian coronavirus 9203, the Duck coronavirus 2714 and the Beluga whale coronavirus 1 are members of *gammacoronavirus* (gamma-CoVs); and the porcine deltacoronavirus (PVCoV) is a part of the *deltacoronaviruses* (delta-CoVs). Whereas alpha-CoVs and beta-CoVs exclusively infect mammalian species, gamma-CoVs and delta-CoVs have a wider host range.

The discovery of the first coronavirus occurred in the first third of the 20th century. In 1931, in the North Dakota, a new type of upper respiratory tract disease was described among chickens (Schalk and Hawn, 1931). Five years later, the virus was isolated and named Infectious Bronchitis Virus (IBV, later renamed as Avian coronavirus) (Beach and Schalm, 1936). Successively, in 1946 and independently in 1949 and 1951, two other animal coronaviruses were described: the transmissible gastroenteritis virus in pigs (later named transmissible gastroenteritis coronavirus, TGEV) (Doyle and Hutchings, 1946) and the murine hepatitis virus (MHV) (Cheever et al., 1949; Gledhill and Andrewes, 1951). In 1965 and 1966, for the first time, two human coronaviruses were isolated from the nasal washings of a male child collected 5 years earlier and from the respiratory tract of medical students with cold collected in 1962, respectively, and named: B814 (B814 was no longer study) and 229E (later renamed human coronavirus-229E, HCoV-229E) (Tyrrell and Bynoe, 1965; Hamre and Procknow, 1966). Based on transmission electron microscopy studies and making comparison between (IBV, MHV and HCoV-229E), by mid-1967, in London, June Dalziel Almeida and David Tyrrell recognized that these three viruses were structurally similar and decided collectively to give the name of coronavirus as all those viruses were characterized by corona-like projections (spikes) on their surface (Almeida and Tyrrell, 1967). In 1967, Kenneth McIntosh and co-workers characterized new human coronaviruses called the human "IBV-like virus" and later renamed human coronavirus-Organ Culture (OC)OC38 and -OC43 (HCoV-OC43) (McIntosh et al., 1967a,b). More recently, the diversity of *Coronaviridae* has greatly improved and a numerous of novel coronaviruses have been discovered in human: the SARS-CoV-1 in 2003, the HCoV-NL63 in 2004 (van der Hoek et al., 2004), the HCoV HKU1 in 2005 (Woo et al., 2005), the MERS-CoV in 2012 (Zaki et al., 2012), but also in a vast variety of animals especially several bat species throughout Africa, America, Asia and Europe (Poudel et al., 2020) until the discovery of SARS-CoV-2 in late 2019, which adds to the list of six known HCoVs associated with different disease phenotypes, from common to highly pathogenic.
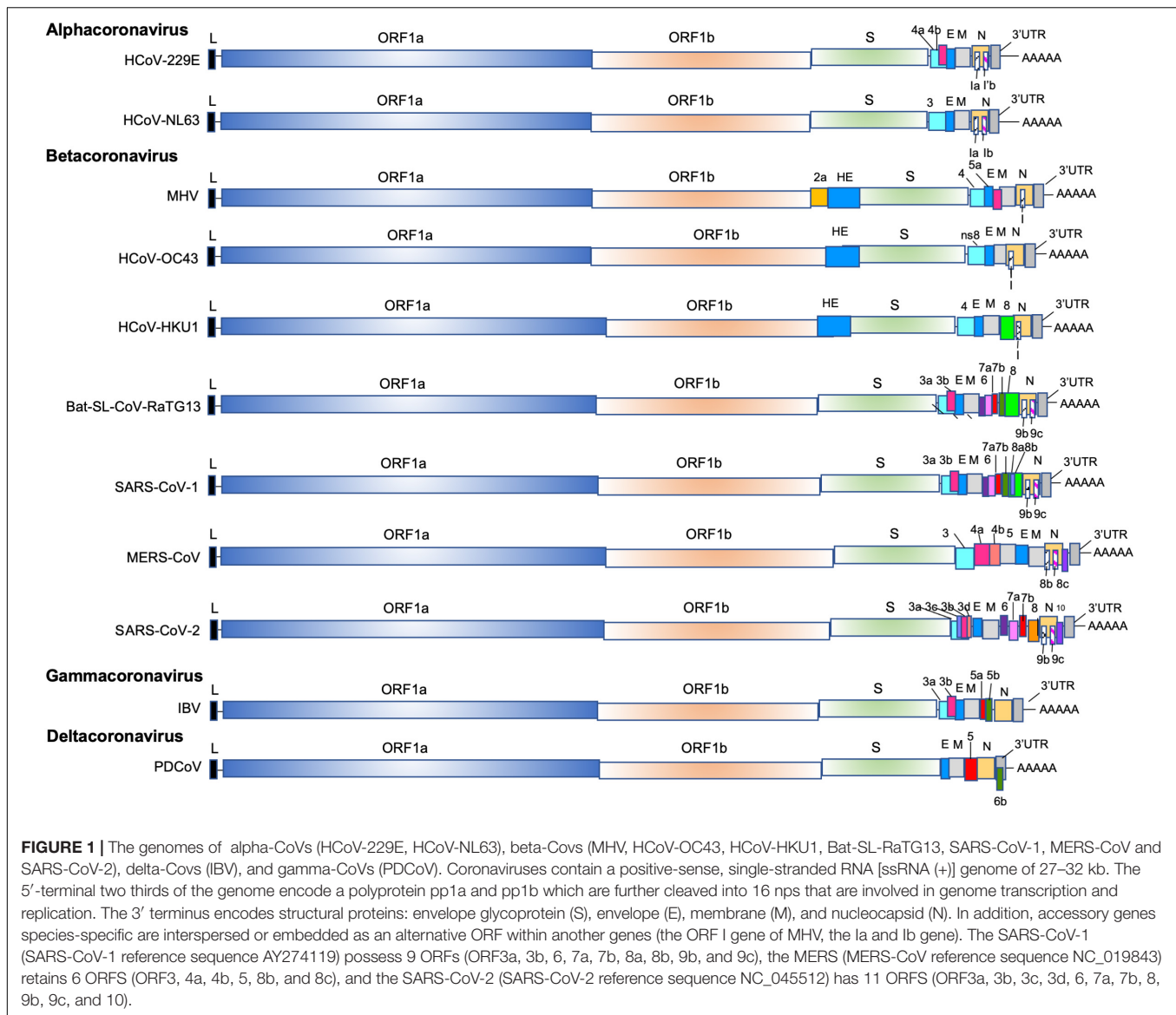
FROM COMMON TO HIGHLY PATHOGENIC HUMAN CORONAVIRUSES: OUTBREAKS AND PHYSIOPATHOLOGY OF CORONAVIRUSES INFECTION

As we saw above, coronaviruses are broadly distributed among humans, and animals (bats, birds, cats, dogs, pigs, mice, horses, whales etc.) causing respiratory, enteric, hepatic, or neurological disorders with variable severity. Animal coronaviruses were mainly studied during several decades because they cause economically significant respiratory and gastrointestinal diseases in domestic animals (20,000 pigs killed by an HKU2-related coronavirus) (Guy, 1998; Saif, 2004; Zhou et al., 2018) and because they provide unique models to study coronavirus pathogenesis. On the other hand, during a long period, the human coronavirus research remained more confidential. However, soon after the first isolation of HCoV-229E and HCoV-OC43, their cellular receptors were characterized. The HCoV-229E was shown to bind to the cellular receptor aminopeptidase N (APN), while the HCoV-OC43 was reported to utilize HLA class I molecule or sialic acids for entry (McIntosh et al., 1967a; Collins, 1993; Krempl et al., 1995). Later, the HCoV-NL63 was found to use angiotensin-converting enzyme–related carboxypeptidase (ACE2) (van der Hoek et al., 2004) and the HKU1, to exploit the HLA class I molecule or sialic acids as receptors (Woo et al., 2005; Esper et al., 2006; Lau et al., 2006; Sloots et al., 2006; Vabret et al., 2006; Chan et al., 2009; Hulswit et al., 2019). As these four HCoVs infect the respiratory tract and are known to be mostly responsible for common colds, the knowledges of these human viruses have only progressively evolved.

However, the field of human coronaviruses research took an abrupt turn when, in late 2002, a new epidemic human disease emerged in southeast China (Guangdong Province), Hong-Kong, Vietnam and subsequently spread to 32 countries, which was caused by an unknown coronavirus. This new disease was named SARS and the causative agent was quickly identified and named SARS-CoV-1. SARS-CoV-1 was transmitted by direct contacts with patients, but also *via* aerosol droplets and contaminated stool (Weinstein, 2004). Air travel by infected individuals quickly spread the disease resulting in the first severe epidemic of the twenty-first century. Over a 6-month period this epidemic led to 8098 illnesses and 774 deaths worldwide (mortality = 9.7%) (Petersen et al., 2020).

Lately in 2012, a new coronavirus from a man suffering of a severe respiratory syndrome was isolated in Saudi Arabia. This new coronavirus, named MERS-CoV that uses the dipeptidyl peptidase 4 (DPP4 also named CD26) as receptor, was responsible of an epidemic disease mostly in Arabic peninsula (Saudi Arabia) leading to 2,494 illnesses and 858 deaths (mortality = 34%) (Alfaraj et al., 2019).

In late 2019, a new strain of coronavirus named SARS-CoV-2, emerged in Wuhan, the capital of China's Hubei province, and the associated illness was named Coronavirus Disease 2019 (COVID-19) (Wu et al., 2020; Zhou et al., 2020c). Following the

FIGURE 1 | The genomes of alpha-CoVs (HCoV-229E, HCoV-NL63), beta-Covs (MHV, HCoV-OC43, HCoV-HKU1, Bat-SL-RaTG13, SARS-CoV-1, MERS-CoV and SARS-CoV-2), delta-Covs (IBV), and gamma-CoVs (PDCoV). Coronaviruses contain a positive-sense, single-stranded RNA [ssRNA (+)] genome of 27–32 kb. The 5′-terminal two thirds of the genome encode a polyprotein pp1a and pp1b which are further cleaved into 16 nps that are involved in genome transcription and replication. The 3′ terminus encodes structural proteins: envelope glycoprotein (S), envelope (E), membrane (M), and nucleocapsid (N). In addition, accessory genes species-specific are interspersed or embedded as an alternative ORF within another genes (the ORF I gene of MHV, the Ia and Ib gene). The SARS-CoV-1 (SARS-CoV-1 reference sequence AY274119) possess 9 ORFs (ORF3a, 3b, 6, 7a, 7b, 8a, 8b, 9b, and 9c), the MERS (MERS-CoV reference sequence NC_019843) retains 6 ORFS (ORF3, 4a, 4b, 5, 8b, and 8c), and the SARS-CoV-2 (SARS-CoV-2 reference sequence NC_045512) has 11 ORFS (ORF3a, 3b, 3c, 3d, 6, 7a, 7b, 8, 9b, 9c, and 10).

first case in Wuhan, the infection spread rapidly until March 11, 2020 when the World Health Organization (WHO) declared COVID-19 as a pandemic. SARS-CoV-2 is transmitted mainly through droplets, close contacts and aerosols. The COVID-19 has considerable global economic and severe health impacts around the world (Ahmad et al., 2020; Munir et al., 2020). As observed in patients infected with SARS-CoV-1 and MERS-CoV, the COVID-19 symptoms are fever, dry cough, fatigue, lymphopenia and can also lead to gastrointestinal symptoms or/and acute respiratory distress syndrome. In all three cases a strong and uncontrolled cytokine release syndrome ("cytokine storm"), responsible of the acute respiratory distress, can be observed in patients (Moore and June, 2020; Tay et al., 2020). As SARS-CoV-1 (Hamming et al., 2004; Jia et al., 2005; Xu et al., 2020), the SARS-CoV-2 uses ACE2 as entry receptor and employs the cellular serine protease TMPRSS2 for envelope protein Spike (S) priming (Li et al., 2003; Kuba et al., 2005;

Matsuyama et al., 2010; Glowacka et al., 2011; Shulla et al., 2011; Hoffmann et al., 2020). As the ACE2 receptor is widely expressed in the arterial and endothelial cells, the arterial smooth muscle cells and especially highly expressed in the lung, renal, cardiovascular, gastrointestinal tissues, but also in hematopoietic cells as macrophages and monocytes, SARS-CoV-1 was shown to infect primary human monocytes, macrophages and but also DCs (Cheung et al., 2005; Gu et al., 2005; Law et al., 2005a; Tseng et al., 2005; Yilla et al., 2005). However, SARS-CoV-1 replication was also described to be abortive in macrophages and the virus was shown be phagocytosed by these cells (Yilla et al., 2005). MERS-CoV was found to infect monocyte-derived macrophages and DCs, although low replication was observed in these cells (Zhou et al., 2015; Tynell et al., 2016). For SAR-CoV-2, as the DCs, residing notably in the lungs, express at the surface the ACE2, these cells can be directly infected by SARS-CoV-2. Moreover, as for the SARS-CoV-1, the

SARS-CoV-2 can also invade host cells via the atypic route of CD147-spike protein (SP) (Wang et al., 2020). Surprisingly, the monocytes obtained from the COVID-19 positive patients are found to be actively expressing ACE2 receptor, suggesting the possibility of direct infection of the monocytes thereby leading to abrogated viral replication and a delayed type I IFN signaling (Acharya et al., 2020). On the other hand, morphological and functional differences in monocytes derived from COVID-19 patients in comparison with monocytes from healthy donors were also observed (Zhang et al., 2021a). Recently, the SARS-CoV-2 infection of ACE2-positive macrophages, present in the spleen and lymph nodes together with alveolar type II pneumocytes and macrophages, were reported (Feng et al., 2020; Xu et al., 2020; Zhang et al., 2020b). SARS-CoV-1, MERS CoV and SARS-CoV-2 are cytopathic viruses, which induce notably lung cell death and trigger immune response (Vabret et al., 2020). Macrophages and monocytes, recruited in response to the infection, release cytokines [interferons I/III (IFN-I/III)], but also proinflammatory Tumor Necrosis Factor alpha (TNF-α), Interleukin-1 (IL-1), IL-6 and IL-18 to prime adaptive T and B cell immune responses (Tay et al., 2020). In most cases, this process is capable of resolving the infection, the immune response declines and patients recover. However, some patients develop a severe disease (3–10% of subjects requires hospitalization, with up to 20% of them experiencing severe disease and a high mortality rate), as it has been previously described for the SARS-CoV-1 and the MERS-CoV infections (Channappanavar et al., 2016, 2019), which is often associated with an impaired type I IFN signature that is blocked by multiple processes at various stages of the IFN signaling pathway (Sa Ribero et al., 2020; Vabret et al., 2020; Xia et al., 2020). This strong antagonism is supposed to assist viral replication, but also to induce an increased release of pyroptosis products that can further bring an aberrant inflammatory response. Together, in addition to the direct damage resulting from the virus replication, the dysregulated pro-inflammatory response can mediate damages in the lung, but also can cause septic shock and induce a multi-organ failure. Patients with severe COVID-19, exhibited, as we mentioned above, an immune dysfunction (dysregulation of T-cells, B-Cells, and innate immune-cells). They present a significantly impaired and delayed secretion of IFN-I and IFN-III compared with flu patient and higher blood plasma levels of IL-2, IL-7, IL-10, granulocyte colony-stimulating factor (G-CSF), IFN gamma-induced protein 10 (IP-10), monocyte chemoattractant protein 1 (MCP1), macrophage inflammatory protein 1α (MIP1α), TNF-α, IL-6, and also show, a higher percentage of CD14 + CD16+ inflammatory monocytes in peripheral blood that secrete inflammatory cytokines contributing to the cytokine storm (Blanco-Melo et al., 2020; Huang et al., 2020; Vabret et al., 2020; Zhou et al., 2020a,d). Together, as with SARS-CoV-1, the clinical manifestation of SARS-CoV-2 results from a dysregulated immune response in patients. SARS-CoV-2 can antagonize type I interferon (IFN-I) production, attenuate activation of the IFN-I signaling pathway, disrupt antigen presentation, increase the production of proinflammatory mediators, in some cases resulting in worsening of the disease. Therefore, understanding the molecular i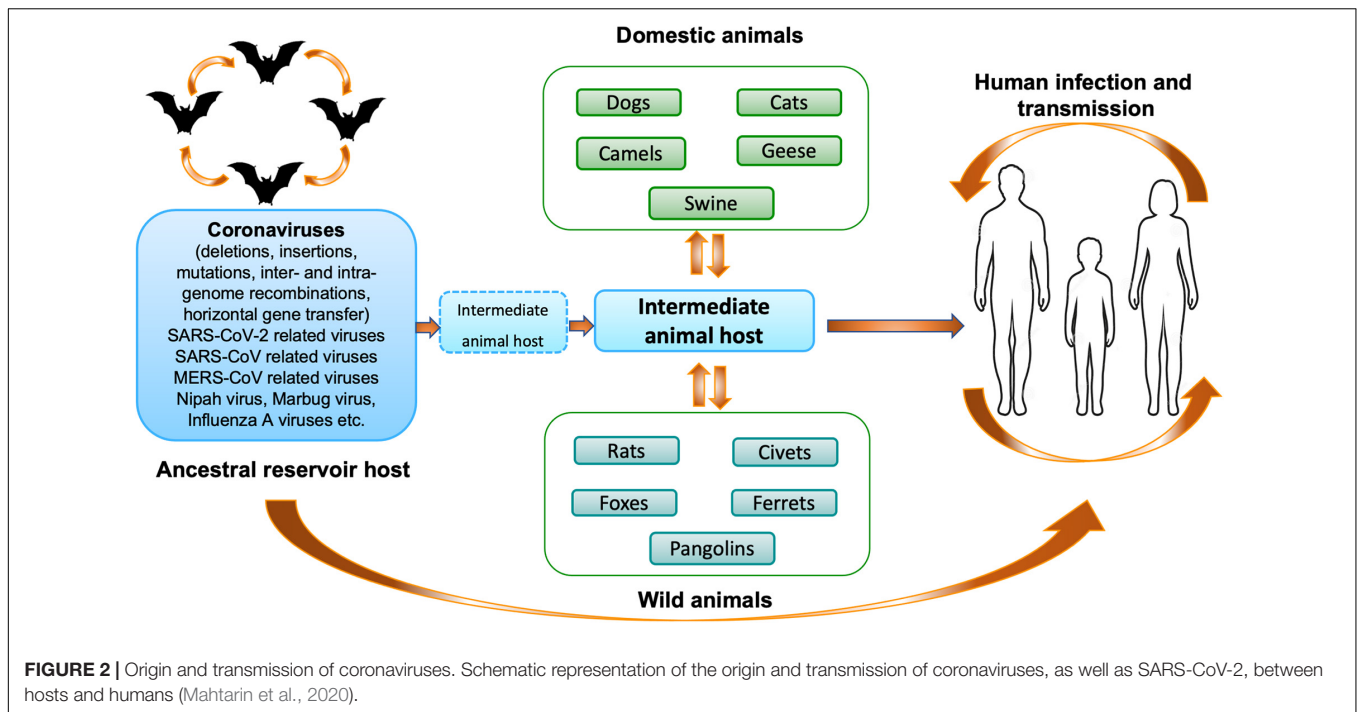nteractions between SARS-CoV-2 and the host defense machinery is crucial for improving treatment interventions. Among several actors, accessory proteins, have been found to modulate the host immune response and to be involve in the inflammatory response. To explore the role played by SARS-CoV-2 accessory proteins in the physiopathology of SARS-CoV-2, additional study remains necessary and especially on the origin highly pathogenic coronaviruses.

ORIGIN OF HUMAN CORONAVIRUSES: SARS-CoV-1, MERS-CoV, and SARS-CoV-2

Different phylogenetic studies have shown that rodents and bats may act as the main reservoir of most alpha-CoVs and beta-CoVs, while birds may be the source of gamma- and delta-CoVs (Su et al., 2016). To date, seven human-CoVs that have zoonotic origins or circulate in animals are known: two alpha-CoVs (HCoV-229E and HCoV-NL63) and five beta-CoVs (HCoV-OC43, HCoV-HKU1, SARS-CoV-1, MERS-CoV, and SARS-CoV-2). They share the common characteristic of having a bat coronavirus as a direct or indirect ancestor (Banerjee et al., 2019b; Cui et al., 2019; Letko et al., 2020). Since the emergence of SARS-CoV-1, MERS-CoV and SARS-COV-2, it becomes clear that bats are probably one of the main reservoirs of CoVs (**Figure 2**). In 2017, a study, conducted by the Goldstein's team, on 12,333 bats representing 282 species from 20 countries in Asia, Africa and Latin America described that 9% of bats carried at least one of 91 distinct coronaviruses. They estimated that there are at minimum 3,204 coronaviruses in bats, supporting the hypothesis that bats can be infected by multiple coronaviruses and that recombination events may occurred between several coronaviruses present in these animals (Anthony et al., 2017b). Interestingly, the authors observed that host switching events (inter-genus or family switching) were proportionally lower in Latin America compared with Africa and Asia without understanding why, suggesting that Africa and Asia are highly susceptible geographic zones for zoonotic emergence.

The emergence of HCoV-OC43 in humans was proposed to occur at the end of the 19th century, when a pandemic of respiratory disease was recorded in humans (Vijgen et al., 2005, 2006). This emergence was suggested to be linked to a host-switching event from an animal-CoV to a human. Indeed, HCoV-OC43 apparently emerged from a bovine coronavirus (BCoV) spillover from a single cross-species transmission event that gave rise to a human-only virus (Vijgen et al., 2005, 2006; Lau et al., 2011). However, the recombination theory from different animal coronaviruses was also retained to promote the emergence of this variant. Indeed, recently a novel canine respiratory coronavirus strain was characterized, which likely resulted from recombination between a Chinese canine respiratory coronavirus and BCoV (Lu et al., 2017).

In 2010, a bat coronavirus termed Appalachian Ridge CoV (ARCoV-2) was detected in North American tricolored bat (*Perimyotis subflavus*) in the United States and showed close relationship with HCoV-NL63 with an estimated divergence of ~550 years ago (Donaldson et al., 2010; Huynh et al., 2012;

FIGURE 2 | Origin and transmission of coronaviruses. Schematic representation of the origin and transmission of coronaviruses, as well as SARS-CoV-2, between hosts and humans (Mahtarin et al., 2020).

Banerjee et al., 2019b). In 2007, sequences related to HCoV-229E were obtained from captive alpacas (*Vicugna pacos*) in California (Crossley et al., 2010). The single virus strain, termed Alpaca coronavirus (A-CoV), was sequenced and found to be highly related to known HCoV-229E strains. In 2009, close relatives of HCoV-229E were detected in bats from Ghana (Pfefferle et al., 2009b). Later on, a phylogenetic analysis suggested that the HCoV-229E possess an ancestral origin in *hipposideridae* bats from Ghana and that camelids could be potential intermediate hosts (Corman et al., 2015). Phylogenetic analyses of complete genomes placed the dromedary-camel-associated viruses in close relationship to HCoV-229E, while the A-CoV clustered with dromedary-camel viruses. This close proximity was explained by the fact that dromedaries are, on occasion, kept along with alpacas in husbandries and zoological gardens. Interestingly, seropositive samples to HCoV-229E of camels were found in Arabian Peninsula and Africa supporting the notion that HCoV-229E has an ancestral origin in bats while camelids serve as a zoonotic natural reservoir for human transmission (Corman et al., 2016). More recently, different bat-CoVs, closely related and ancestor of HCoV-NL63, were identified in *Triaenops* bats from Kenya. Interestingly, these NL63-like viruses have acquired, by a recombination event, a 229E-like S gene from 229E-like viruses circulating in *Hipposideros* bats (Tao et al., 2017). With HKU1, to date, no related virus has been isolated from any animal species. However, as HCoV-HKU1 was depicted to be most closely related to MHV and a rat coronavirus (rat sialodacryoadenitis virus), in consequence it cannot be excluded that the hosts of the ancestral viruses are carried by rodents (Woo et al., 2005; Lau et al., 2015b).

The origin of SARS-CoV-1 was first shown by the isolation and the identification of closely related viruses in Himalayan palm civets and also raccoon dog at a live-animal market in Guangdong, China (Guan, 2003). A very high genome sequence identity (99%) was described between the SARS-CoV-1 isolated from civets and SARS-CoV-1 from humans, supporting the notion that SARS-CoV-1 is of animal origin. However, subsequent studies showed that palm civets on farms and in the field were largely free from SARS-CoV-1 infection. These latest results have thus made possible to propose that palm civets, rather than a natural reservoir, play the role of an intermediate host. In subsequent studies, among different bat populations, two teams reported the discovery of CoVs related to SARS-CoV-1 in horseshoe bats (genus *Rhinolophus*) suggesting that SARS-CoV-1 could have spilled over into humans directly from bats while the civet and the raccoon dog, were incidental hosts (Lau et al., 2005; Li et al., 2005). This hypothesis was reinforced by the isolation, during surveillance campaign, of bats coronaviruses closely matching the human SARS-CoV-1 (Ge et al., 2013; Kuehn, 2013). These viruses were called "SARS-CoV-like viruses" (SL-CoV) or "SARS-CoV-related viruses". Chinese horseshoe bats (*Rhinopholus*), belonging to the family *Rhinolophidae* of the order *Chiroptera* under Microchiroptera (microbats), is an insectivorous species widely distributed notably in Chinese forests, however, SL-CoVs were also discovered in *Rhinolophus* from other geographical origins (European, African, and Southeast Asian countries). Indeed, SL-CoVs were also discovered in *rhinolophids* from Slovenia, Bulgaria, and Italy in Europe (Drexler et al., 2010; Rihtaric et al., 2010; Balboni et al., 2011). In Africa, novel beta-CoVs related to SARS-CoV-1 were also detected in bats (*Hipposideros* and *Chaerephon* genus) from Ghana, Kenya, and Nigeria (Pfefferle et al., 2009b; Tong et al., 2009; Quan et al., 2010; Hu et al., 2015). However, compared with Asian and European

bat SL-CoVs, these viruses were phylogenetically distant to SARS-CoV-1. Overall, it was postulated that the direct ancestor of SARS-CoV-1 may have arisen from sequential recombination events between the precursors of the bat SL-CoVs prior to spillover directly to humans or by an intermediate host and that humans were exposed to bat SL-CoVs prior to the SARS-CoV-1 epidemic 2002–2003. Indeed, it has been suggested that the ancestor of civet SARS-CoV-1 probably acquired the Open Reading Frame 8 (ORF8) from *Rhinopholus ferrumequinum* SL-CoVs by recombination (Lau et al., 2015a) and seroprevalence studies described a stronger antibodies response against SL-CoV isolated from Himalayan palm civet rather than human SARS-CoV-1 in sera of healthy individuals in Hong-Kong in 2001 (Zheng et al., 2004).

Concerning MERS-CoV, neutralizing antibodies against this virus were found in the sera of dromedary-camels in the Middle East (Reusken et al., 2013) and in several African countries (Chan et al., 2015). A MERS-CoV strain, identical to the human MERS-CoV, was also isolated from the nasal swabs of dromedary-camels (Raj et al., 2014), suggesting that dromedary-camels are the natural reservoir of MERS-CoV. However, various coronaviruses that are phylogenetically related to MERS-CoV, but phylogenetically distant from humans and dromedary-camels, have also been found in different species of bats, including Egyptian tomb bats (Al-Tawfiq and Memish, 2020), African bats (Ithete et al., 2013; Corman et al., 2014a; Anthony et al., 2017a), Italian bats (Lelli et al., 2013; Moreno et al., 2017), and Chinese bats (Yang et al., 2014), suggesting that these animals are the ancestral reservoir. Thus, it is possible that other, yet unidentified strains of bat viruses, are circulating in nature and may have directly contributed to the emergence of MERS-CoV in camels less than 30 years ago and later in humans (Reusken et al., 2013; Corman et al., 2014b; Müller et al., 2014; Alshukairi et al., 2018; Al-Tawfiq and Memish, 2019; Kiyong'a et al., 2020). Again, many recombination events, involving the exchange of genetic elements between different viral ancestors, including those isolated from dromedary-camels and bats, are suspected.

Two main scenarios that can plausibly explain the origin of SARS-CoV-2 are actually proposed (**Figure 2**): (i) natural selection in an animal host before zoonotic transfer; and (ii) natural selection in humans following zoonotic transfer (Andersen et al., 2020; Lam et al., 2020a; Xiao et al., 2020; Zhou et al., 2020c). SARS-CoV-2, closely related to a bat-SL-CoV-RaTG13 (96.2% nucleotide identity), which had previously been found in 2013 in bats *Rhinolophus affinis* in Pu'er (Yunnan Province, China), a bat-SL-CoV-RmYN02 collected in *Rhinolophus malayanus* in the Yunnan Province in 2019 (93.3% nucleotide identity), a bat-SL-CoV-ZXC21/ZC45 related (87.6–87.8% nucleotide identity) found in bats *Rhinolophus pusillus* from Zhoushan, China, has emerged (Hu et al., 2018; Zhang and Holmes, 2020; Zhou et al., 2020b,c). Soon after the first detected case of SARS-CoV-2 infection in humans, it was hypothesized that the first case of animal-to-human transmission of SARS-CoV-2 was linked to an animal intermediate host, which would have been among the wildlife species sold on the Huanan market in Wuhan (Huang et al., 2020). Several sequencing studies have suggested that the pangolins could

harbor ancestral beta-CoVs related to SARS-CoV-2 (85–92% nucleotide sequence homology with SARS-CoV-2) (Liu et al., 2019; Lam et al., 2020b). Indeed, Pangolin-SARS-CoV-2-related-Cov was found to share a very similar Receptor Binding Domain (RBD) with SARS-CoV-2, (97.4% amino acid sequence identity), while the RBDs of bat-SL-CoV-RaTG13 were more divergent albeit higher degree of sequence homology (Lam et al., 2020b). However, due to the sequence divergence between SARS-CoV-2 and pangolin SARS-CoV-2-related beta-CoVs, the origin of the virus has not been determined conclusively and the evolutionary pathway of SARS-CoV-2 in bats, pangolins or other mammals remains to be established. Another proposal is in favor of multiple recombination events between a pangolin SARS-CoV-2-related beta-CoV (Xiao et al., 2020) and the bat-SL-CoVs in a third wild animal species. Since the beginning of the SARS-CoV-2 pandemic, extensive discussions regarding its origin took place. The molecular specificity of the SARS-CoV-2, however, indicates that the virus is not the product of purposeful manipulation. Indeed, the RBD sequence of SARS-CoV-2 is not optimal such as the SARS-CoV-1, suggesting that a natural selection is more probable. Another hypothesis is that SARS-CoV-2 could have originated from a bat virus that has been adapted to experimental animal models or laboratory cells and that has escaped from a laboratory. In any case, further research is required before a final decision can be made on the exact origin of the SARS-CoV-2. Bats are presumed reservoirs of diverse coronaviruses including progenitors of respiratory syndrome: SARS-CoV-1, MERS-CoV and SARS-CoV-2, but also other coronaviruses as well as viruses from different families. The complete characterization of the bat immune system is a major challenge in understanding the adaptation of coronaviruses

BAT AND CORONAVIRUSES: IMMUNE RESPONSE

Bats, which harbor a large number of zoonotic viruses per species, are animals that host the larger number of zoonotic viruses after rodents, and they are the second most diverse mammalian order on earth that are distributed across every continent except Antarctica (Luis et al., 2013). The order *Chiroptera* is complex and consists of 22 families over genera and 1,400 species of bats that are distributed across most continent (Fenton and Simmons, 2015). This order is divided in two suborders: *Yinpterochiroptera* and *Yangochiroptera*. Within these suborders, bat species display considerable diversity in diets, sizes, morphologies and ecological niches. Until now, 28 different virus families (DNA and RNA viruses) have been identified in bats. Among these viruses, several RNA viruses are zoonotic viruses (coronaviruses, henipaviruses, lyssaviruses, and filoviruses). Although they do not induce serious clinical symptoms and pathogenesis in bats, as far as we know, they represent a danger of great importance for humans and veterinary health (Ge et al., 2013; Ithete et al., 2013; Anthony et al., 2017a; Hu et al., 2017). The basis of bat resistance to viral infections is still very poorly understood. Some argue that this resistance is related to their high-energy metabolism

required during their flight that would induce a high body temperature and mimic the fever phenomenon observed during the activation of the immune response, allowing a reduction of virus replication (O'Shea et al., 2014; Brook and Dobson, 2015). However, the replication of filoviruses (Marburg virus and Ebola virus), in cell lines derived from several bat species, was found to be also effective at 37°C or 41°C, contradicting this hypothesis (Miller et al., 2016). The bat resistance is likely more related, for a large part, to the presence of a specific, but not universal, immune system. In recent years, significant efforts in the study of the biology of some bat species has been made and several characterizations of bat innate immunity have given some insights into the differences between bat and human immune responses. To date, the genomes and transcriptomes of at least 18 bat species have been completed (Hawkins et al., 2019; Jebb et al., 2020). These analysis revealed that only a small proportion of the innate immune system homolog genes (in human immune-related genes represent 7% of the genome while in bat they represent only 2.75–3.5% of the transcribed genes) is detected in the bat genome (Egyptian fruit bat: *Rousettus aegyptiacus* and black flying fox: *Pteropus alecto*) (Papenfuss et al., 2012; Lee et al., 2015). Moreover, the analysis of the *Pteropus alecto* transcriptome has also allowed the identification of a significant proportion of bat-specific transcripts without homology with known annotated transcripts, which may be involved in the immune response (Papenfuss et al., 2012). Even if innate immune response has been poorly depicted in bats, the bat cells from multiple species were shown to upregulate efficiently type I IFNs and IFN-Stimulated Genes (ISGs) in response to poly(I:C) treatment, suggesting that a sensing machinery is conserved in bat cells (Omatsu et al., 2008; Crameri et al., 2009; Cowled et al., 2012; Banerjee et al., 2016, 2017, 2020). Thus, full-length transcripts for Pattern Recognition Receptors (PRRs) as Toll-Like Receptors-1 (TLR-1) to TLR-10 have been detected and a TLR-13 pseudogene have been identified in *P. alecto* (Cowled et al., 2011). However, differences have be found in TLRs sequences between bats and humans, but also within the bat species (Schad and Voigt, 2016). Contrary to TLRs, cytosolic RNA sensors (Retinoic-acid Inducible Gene I: RIG-I and Melanoma Differentiation-Associated gene 5: MDA-5) have been shown to be conserved in most bat genomes or transcriptomes (Cowled et al., 2012; Papenfuss et al., 2012; Banerjee et al., 2017). In human cells, virus recognition, and related-signaling pathways, leads to phosphorylation of transcription factors like Interferon Regulatory Factor 3 and 7 (IRF3 and IRF7). Bat IRF3 sequences has been found evolutionarily distinct from their mammalian counterparts (Banerjee et al., 2019a), but functional studies indicate that IRF3 from *Eptesicus fuscus* cells mediates antiviral signaling in response to MERS-CoV infection or poly(I:C) stimulation, while IRF7 is induced following poly(I:C) stimulation in both *E. fuscus* and *P. alecto* cells (Zhou et al., 2014; Banerjee et al., 2017). Interestingly, IRF7 mRNA has been found to be constitutively expressed in *P. alecto* cells and to have a ubiquitarian tissue distribution compared to its human counterpart (Zhou et al., 2014; Banerjee et al., 2019a). RIG-I-like receptors (RLRs) are intracellular sensors for viral RNAs that activate Mitochondrial AntiViral Signaling protein

(MAVS) leading to nuclear translocation of NF-kB and IRF3 for the induction of type I IFNs (Seth et al., 2005). Functional MAVS signaling has been established to be present in the straw-colored fruit bat (*Eidolon helvum*) and in the horseshoe bat (*Rhinolophus sinicus*). However, downstream signaling pathways and molecules involved in MAVS-mediated signaling have not yet been characterized in bats. IFN response is critical to restrain virus propagation. Constitutive expression of IFN-α transcripts and associated IFN-stimulated genes (ISGs) have been detected in unstimulated *P. alecto* cells (a contraction of the IFN locus with only 3 functional IFN-α genes exists in this bat genome compared with 7–18 IFN-α loci in other mammals and in spite of everything a constitutive and ubiquitous expression is observed in these specific bats), but not in primary cells from *R. aegyptiacus,* which have a higher expression of IFN-α/β receptors (IFNAR1 and IFNAR2) and several ISGs (Zhou et al., 2016; Pavlovich et al., 2018). In *Pteropus vampyrus* genome, coding sequences for three type III IFNs have been described, but only two of them were found to be transcribed in the closely related bat *P. alecto* (Zhou et al., 2011). Altogether, these data suggest that species-specific differences in IFN responses exist in bats. In addition to these characteristics related to the species of bat, another parameter must be considered. Indeed, viral infections of bat cells with viruses from several families have been shown to either induce or inhibit IFN production. Indeed, MERS-CoV that inhibits antiviral IFN expression in human cells, induces an increase in IFN-β and 2–5 -OligoAdenylate Synthase 1 (OAS1) transcripts in MERS-CoV-infected *E. fuscus* cells (Banerjee et al., 2019a) and a transient upregulation of Myxovirus resistance protein 1 (Mx1), ISG56 and Regulated on Activation Normal T cell Expressed and Secreted (RANTES) in infected Jamaican fruit bats (*Artibeus jamaicensis*) cells (Munster et al., 2016), whereas Nipah virus infection has been shown to inhibit IFN production in infected-*P. alecto* cells (Virtue et al., 2011). IFN signaling in *P. alecto* cells is dependent on IFNAR2 (Zhang et al., 2017). It has not yet been established how many ISGs are induced in different bat cells, but considering the vast bat species diversity, it is likely that bat species express different ISGs. Poly(I:C) stimulation was shown to induce the expression of transcripts for MDA5, RIG-I, radical S adenosylmethionine domain-containing 2 (RSAD2), IRF7, OAS1, IFN-inducible protein 6 (IFI6) and Mx1 in *E. fuscus* cells (Banerjee et al., 2017). Expression of Mx1, OAS1 and Protein Kinase R (PKR) transcripts was also found to be induced in *P. alecto* cells (Zhou et al., 2013). Even if the expression of Mx1 from six different bat species in human embryonic kidney (HEK 293T) cells was found to efficiently reduce Ebola virus and influenza A virus replication (Fuchs et al., 2017), the expression of atypical ISGs was also described in *P. vampyrus* (RND1, SERTA-domain containing 1 (SERTAD1), ChaC glutathione specific gamma-glutamyl cyclotranferase 1 (CHAC1), MORC3) (Glennon et al., 2015; Banerjee et al., 2020). While keeping IFN response or not, bats seem to support unique adaptations to control inflammation, as far as we know, thus limiting tissue damages and pathologies. This apparent lack of a strong inflammatory response in virally-infected bat immune cells has been linked, so far, to two features: (i) a dampened activation

of the NLRP3 inflammasome in bat immune cells compared to human immune cells without impacting the overall viral loads (Ahn et al., 2019) and (ii) an interaction of a transcription repressor of the NF-kB family, c-Rel, with the putative c-Rel motif located in the TNF- promoter in the big brown bat (*E. fuscus*) (Banerjee et al., 2017).
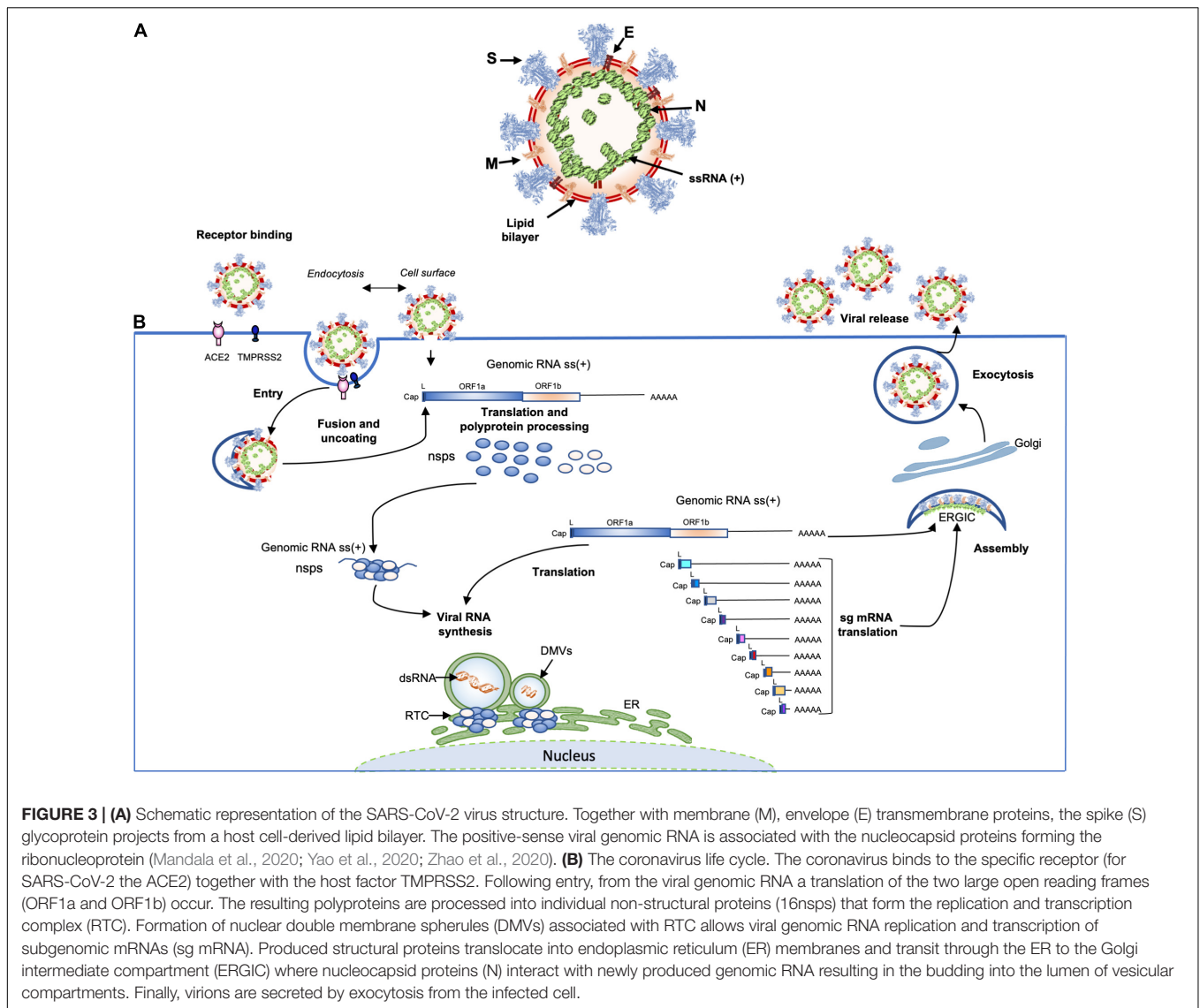
Even if studies on adaptive immunity of bats have been so far limited, some information are available. Indeed, the antibody response in bats is poorly known, however, the detection of transcripts of the four major subclasses of antibodies such as IgA, IgE, IgG, and IgM has been performed in bats cells (Baker et al., 2010; Wynne et al., 2013). Experiments on *R. aegyptiacus* with Marburg virus showed variability in antibody response. One study reported a detectable antibody response for up to 3 months (Schuh et al., 2017), while another study reported a detectable antibody response for up to 11 months (Storm et al., 2018). Moreover, studies of sera of convalescent *R. aegyptiacus* bats challenged with Ebola virus, Marburg virus and Sosuga virus, demonstrated that the clearance of these viruses does not appear to be mediated by virus-specific neutralizing antibodies and may rely on other antibody-mediated functions (antibody-dependent cellular cytotoxicity or phagocytosis) (Schuh et al., 2019). By contrast, upon experimental challenge, bats have been shown to generate virus-neutralizing antibodies to Nipah virus and rabies virus (Middleton et al., 2007; Turmelle et al., 2010). Concerning the immune cells present in bats, surprisingly, a high proportion of CD3+ T cells constitutively expressing mRNAs for IL-17A, IL-22 or Transforming Growth Factor Beta 1 (TGFB1) was detected. CD8+ T cells were also identified in the spleen as well as CD4+ T cells in the blood, lymph nodes and bone marrow and, to a lesser extent, B cells were detected in wild *P. alecto* (Martínez Gómez et al., 2016). The Major Histocompatibility Complex (MHC) molecules of bat, using *in vitro* experimental approaches, was shown to possess a unique consensus-binding motif and bind to larger peptides compared to other mammals. Finally, macrophages, monocytes and bone marrow derived DCs of bat were found to harbor phagocytic functions. Without a doubt, the long and complex history of co-evolution between bats and their viruses has played a crucial role in the construction of a bat-specific immune systems (Banerjee et al., 2019b, 2020; Letko et al., 2020). Undoubtedly, research on bats and the viruses they harbor, still requires a lot of effort. Many questions remain at the molecular and cellular levels, mainly because of to their great diversity, but also the lack of tools (*in vitro* and *in vivo* tools), the experimental limitations (e.g., limited repertoire of cell lines of different bat species, limitation in the isolation of bat viruses, propagation and virus stocks in bat cells instead of human cell lines) and the absence of experiments on live animals. It is particularly important to point out that all the current knowledge on the immune response in bats is highly fragmented and often obtained for a specific to a single pathogen in a single bat species or inferred from experiments conducted in limited cell models, so that no general conclusions should be made from any observation. The most revealing example is the primary cells of *P. alecto,* which have been described as constitutively expressing IFN-α mRNA, while it has never been found in the cells of *R. aegyptiacus*. Further studies remain necessary and essential to explore the complex relationship between bats and coronaviruses.

CORONAVIRUS BIOLOGY

Coronavirus Structure

The coronavirus virions are spherical and ellipsoidal enveloped particles with an average diameter of 60–160 nm and a characteristic crown of spike proteins (**Figure 3**). The main structural components of coronaviruses are the spike (S), the envelop protein (E), the transmembrane protein (M) and the nucleoprotein (N) which forms a viral ribonucleoprotein (vRNP) complex with the viral RNA (vRNA). The first electron microscopy analysis on SARS-CoV-2 particles revealed that SARS-CoV-2 is consistent with these characteristics with diameters ranging from 53 to 110 nm including the contribution of the spikes ($\cong$23 nm) (Klein et al., 2020; Yao et al., 2020; Zhou et al., 2020c). The S glycoprotein of SARS-COV-2, composed of S1 and S2 domains, which forms trimer ($\sim$600-kDa), is heavily glycosylated with 66 N-linked glycans and found at average of 26 ± 15 distributed on each virion, half as much as SARS-CoV-1, which was estimated to possess an average of $\sim$50–100 S per virion (Neuman et al., 2011). The S1 glycoprotein trimer is the primary determinant of host interaction with three RBDs responsible for the binding of the virus to its receptor ACE2, while the S2 is involved in the membrane fusion process. An envelope protein (E) is also found on the surface of the virion of coronaviruses, it is the smallest of all the structural proteins (8–12 kDa). The E protein is an intrinsic membrane protein that has been described to be involved in a wide spectrum of functions. Indeed, the SARS-CoV-1 E protein has been found to be engaged in several steps of the viral life cycle, such as assembly, budding but also in pathogenesis by activating the host inflammasome (Schoeman and Fielding, 2019). The structure of the SARS-CoV-2 E protein (75 residues) has been solved by RMN and its influence on the capacity of the S protein to promote assembly of SARS-CoV-2 virus-like-particles was described (Boson et al., 2020; Mandala et al., 2020). The most abundant structural protein, the M multi-spanning membrane protein, defines the shape of the viral envelope, it is also the central organizer of coronaviruses assembly, interacting with all other major coronaviral structural proteins (Masters, 2006). During assembly of the virion, the M interacts with itself, with the nucleocapsid protein N, with E and with the S protein (Kuo and Masters, 2002; Boscarino et al., 2008; Arndt et al., 2010). SARS-CoV-2 M protein has recently been found to act as an antagonist of both types I and III IFNs by affecting the formation of the RIG-I/MDA-5–MAVS–TNF Receptor-Associated Factor 3 (TRAF3)–TANK-binding kinase 1 (TBK1) signalosome (Zheng et al., 2020). Moreover, the M protein was found to interact with the central adaptor protein MAVS impairing MAVS aggregation and recruitment of downstream TRAF3, TBK1, and IRF3, leading innate antiviral response attenuation. SARS-CoV-2 membrane glycoprotein M antagonizes the MAVS-mediated innate antiviral response (Fu et al., 2020). In the viral lumen, oligomers of N protein interact with the $\sim$30-kb-long single-stranded vRNA to

FIGURE 3 | (A) Schematic representation of the SARS-CoV-2 virus structure. Together with membrane (M), envelope (E) transmembrane proteins, the spike (S) glycoprotein projects from a host cell-derived lipid bilayer. The positive-sense viral genomic RNA is associated with the nucleocapsid proteins forming the ribonucleoprotein (Mandala et al., 2020; Yao et al., 2020; Zhao et al., 2020). **(B)** The coronavirus life cycle. The coronavirus binds to the specific receptor (for SARS-CoV-2 the ACE2) together with the host factor TMPRSS2. Following entry, from the viral genomic RNA a translation of the two large open reading frames (ORF1a and ORF1b) occur. The resulting polyproteins are processed into individual non-structural proteins (16nsps) that form the replication and transcription complex (RTC). Formation of nuclear double membrane spherules (DMVs) associated with RTC allows viral genomic RNA replication and transcription of subgenomic mRNAs (sg mRNA). Produced structural proteins translocate into endoplasmic reticulum (ER) membranes and transit through the ER to the Golgi intermediate compartment (ERGIC) where nucleocapsid proteins (N) interact with newly produced genomic RNA resulting in the budding into the lumen of vesicular compartments. Finally, virions are secreted by exocytosis from the infected cell.

form vRNPs (Kang et al., 2020). Until now, very little information on the structure of RNPs has been available. However, by studying the SARS-CoV-2 virus, the native conformation of vRNPs and their high order assembling were revealed for the first-time (Klein et al., 2020; Yao et al., 2020). An average of 38 vRNPs per virion was described. It has been proposed that vRNPs would be linked to other neighboring vRNPs as "beads on a string" in order to maintain a relative flexibility and allowing an efficient packaging of the very large 30 kb vRNA in the viral lumen of 80 nm diameter (Klein et al., 2020; Yao et al., 2020).

Coronaviruses, Genome Organization and Viral Proteins

Coronaviruses have the largest RNA genomes (27–32 kb; SARS-CoV-2 reference sequence NC_045512: 29,903 kb) of any known virus family (**Figure 1**). The single-stranded positive sense vRNA genome contains a 5′ methyl-guanosine cap and a 3′ poly(A) tail.

The order of genes is highly conserved among all coronaviruses. The genome comprises a basic set of genes, 5′-replicase-S-E-M-N-3′. The replicase gene includes two sub-units: ORF1a and ORF1b, that occupy two thirds of the viral genome and are translated into two polyproteins pp1a (nsp-1-11) and pp1ab (nsp-1-10, nsp-12-16) that are cleaved by viral proteases into 16 proteins involved in genome replication and in the subgenomic mRNAs (sg mRNAs) synthesis (**Figure 3**) (Yao et al., 2020). The two polyproteins are the only proteins translated from the genome whereas the products of all downstream ORFs are derived from sg mRNAs. Pp1a is cleaved into 11 nsps: nsp1 (19.6 kDa, 180 amino acids, host mRNA degradation, translation inhibition), nsp2 (70.5 kDa, 639 amino acids, unknown function), nsp3 (217 kDa, 1946 amino acids, polyprotein processing, de-ADP-ribosylation, deubiquitination, IFN antagonist), nsp4 [56 kDa, 501 amino acids, double membrane vesicles (DMV formation)], nsp5 (33.7 kDa, 307 amino acids, polyprotein processing, inhibition of IFN signaling), nsp6 (DMV formation),

nsp7 (9.2 kDa, 84 amino acids, co-factor for RNA dependent RNA polymerase), nsp8 (21.8 kDa, 199 amino acids, primase or 3′ terminal adenylyltransferase, cofactor for RNA dependent RNA polymerase), nsp9 (12.3 kDa, 114 amino acids, RNA-binding protein), nsp10 (14.7 kDa, 140 amino acids, cofactor for nsp14 and 16), nsp11 (1.3 kDa, 13 amino acids, unknown function), whereas pp1ab is processed into nsp1-10, plus nsp12 (106.6 kDa, 932 amino acids RNA-dependent polymerase RNA polymerase, nucleotidyltransferase), nsp13 (66.8 kDa, 601 amino acids, helicase, RNA 5′ triphosphatase), nsp14 (59.8 kDa, 527 amino acids, 3′–5′ exonuclease, proofreading, RNA cap formation, guanosine N7-methyltransferase), nsp15 (38.8 kDa, 346 amino acids, endoribonuclease, evasion of immune response) and nsp16 (RNA-cap formation, ribose 2′-O-methyltransferase). The other one-third of the viral genome contains genes coding for virion structural proteins, S-E-M and N, described above. In addition to the basic genes downstream of the replicase gene, supplementary Open Reading Frames (ORFs), named "accessory" genes are interspersed between the main genes or embedded as alternative ORF within another gene [e.g., the internal (I) gene of MHV]. These "accessory" genes, appear to be specific within each coronavirus genus. They are labeled according to the smallest transcript in which they fall, so there is usually no relatedness. Some of these extra ORFs are thought to have been acquired through ancestral recombination with RNA from heterologous viral sources. Many of these genes are maintained in the genome of coronaviruses suggesting that they might play an important role in viral cycle or/and pathogenesis. The SARS-CoV-2 genome encodes at least for 9 non-structural ORFs: ORF3a, 3b 6, 7a, 7b, 8, 9b, 9c, and 10. Some of these ORFs are translated into accessory proteins: ORF3a, 6, 7a, 7b, 8 while others ORF3b and ORF10 might not be translated. These non-structural ORFs and the accessory proteins will be further described in more details.

CORONAVIRUS REPLICATION CYCLE

Coronavirus replication cycle is initiated by the binding of virions to cellular attachment factors and then to the specific interaction between S and the cellular receptor, such as ACE2 for SARS-CoV-2, together with host factor, such as TMPRSS2 which promotes viral uptake and fusion at the cellular membrane or endocytosed into endosomes where spike is processed by cathepsin L (**Figure 3**) (Murgolo et al., 2021). Following delivery and uncoating of the viral nucleocapsid to the cytoplasm, the translation of the two ORF1a and ORF1b from the genomic RNA start and, as we discussed above, two polyproteins are product (pp1a and pp1ab). The latter results from a programmed −1 ribosomal frameshift at the overlap of ORF1a et ORF1b (the coronavirus frameshifting stimulation element-stem-loop 1 of SARS-CoV-2, reference strain: NC_045512, is located between nucleotides 13,476–13,503) (Finkel et al., 2020). Thus, Pp1a and pp1ab are then processed by two cysteine proteases located within the nsp3 (papain-like protease) and nsp5 (chymotrypsin-like protease) respectively, into 16 mature products (nsp1 to nsp16). Rapidly after proteolytic cleavage, the nsp1 targets the host-cell translation (Schubert et al., 2020; Thoms et al., 2020), while

the other processes nsps (nsp2-16) together with characteristic perinuclear double membrane vesicles (DMVs) that assemble to form a Replication and Transcription Complex (RTC) (Romano et al., 2020). The establishment of the RTCs is supposed to offer an optimal environment for RNA synthesis and prevent the exposure of viral replication intermediates to cytosolic innate immune sensors. However, a recent direct visualization of SARS-CoV-2 RTCs, using cryo-electron microscopy, revealed that the main component of RTCs interior is branched to double stranded RNA filaments and not proteins (Klein et al., 2020). This suggests that viral replicase complex is not located inside the RTCs, but instead, is associated with DMVs, and that viral RNAs intermediates accumulate into the RTC via a molecular pore complex newly described, interconnecting RTC interior and cytoplasm (Wolff et al., 2020). This viral replicase complex is based, on one hand, on nsp2-11 (factor assisting viral replication, host immune evasion or recruitment of intracellular membranes) and, on the other hand, on nsp12-16 that provide enzymatic functions, but also of several cellular proteins (V'kovski et al., 2019). Following the RTCs formation, the replication of genomic RNA and the transcription of multiple subgenomic mRNAs (1 h post-infection, the genome and sg mRNAs are detected) begin. The viral genomic replication starts by the synthesis of a full-length-negative-sense genomic copies that serve as template for the generation of new positive-sense genomic RNA. These new genomes are then packaged into new virions or used for translation to produce more nsps. During negative-strand RNA synthesis, a discontinuous transcription mechanism takes also place. The evidence for canonical discontinuous synthesis was first proposed following the discovery of a specific RNA sequence on the viral plus strand in addition to the poly(A) tract located at the 3′ end. Indeed, it has been shown that the genome and all sg mRNAs possessed, notably, at the end of the leader and before the body of each ORF, an "intergenic sequence" (IGS) or "transcription regulating sequence" [TRS, TRS-Body (TRS-B), TRS-Leader (TRS-L), which served to redirect the viral polymerase to an internal site and, then, served as a primer for elongation (van der Most et al., 1994; van Marle et al., 1995)]. During this event of (-) sg RNAs synthesis, when the TRS-body is reached, the transcription is interrupted and re-initiated at the TRS adjacent or switch to amplify the leader sequence (∼70 nt) at the 5′ end of the genome guided by complementarity of the TRS-B to the leader TRS (TRS-L). Following this process, a set of (-) sg RNAs is produced and used to synthetize a set of positive-sense sg mRNAs. For both human alpha-CoVs (HCoV-229E and HCoV-NL63), as well as the beta-CoVs (MHV, HCoV-OC43, and HCoV-HKU1), the TRS sequence is 5′ UCUAAAC3′, except for HCoV-OC43, where a C is replaced by a U (UCUAAAU). For SARS-CoV-1 and SARS-CoV-2, as well as bat SL-CoVs, the TRS sequence is 5′AAACGAAC3′. In addition to this canonical sg RNAs synthesis, a non-canonical sg RNAs production was recently described for SARS-CoV-2 and showed that, in addition to the genomic and 9 sg mRNAs, SARS-CoV-2 produces transcripts encoding unknown ORFs with fusion, deletion, and/or frameshift (Kim et al., 2020). Moreover, RNA modification sites on viral transcripts were also described, with the most frequent motif, AAGAA. The

modification sites on the 'AAGAA-like' motif (including AAGAA and other A/G-rich sequences) are found throughout the viral genome but particularly enriched in genomic positions 28,500–29,500. Long viral transcripts (gRNA, S, 3a, E, and M) are more frequently modified than shorter RNAs (6, 7a, 7b, 8, and N), suggesting a modification mechanism that is specific for particular RNA species (Kim et al., 2020). Following transcription and translation, the structural proteins M, S, E, and N translocate into endoplasmic reticulum (ER) and transit through the ER-to-Golgi Intermediate Compartment (ERGIC) where N proteins interact with genomic RNA and condense with the M proteins, the envelope components to form virions, which then bud into lumen of secretory vesicular compartments. Finally, virions are secreted from the infected cells by exocytosis (V'kovski et al., 2020). As we have seen above, the biology of coronaviruses is complex and further studies on these viruses, their diversity and evolution are crucial today to better understand and anticipate other zoonotic events.

DIVERSITY OF CORONAVIRUSES

Coronaviruses can infect a variety of animals and humans. The intra- and interspecies transmissions of coronaviruses form a complex ecosystem promoting evolution and diversity. Previous studies on CoVs genomes have shown a higher composition in AU than in GC [selective advantage of the abundance of the A nucleotide (Kustin and Stern, 2021)]. For the SARS-CoV-2, in all the structural genes, either A or U nucleotides are the most predominant nucleotides and, in addition, A or U nucleotides are also the most predominant nucleotides at the 3rd position of codons implying a higher gene expression efficiency of these SARS−CoV−2 proteins (Kandeel et al., 2020; Sheikh et al., 2020). In term of evolution, two genes of SARS-CoV-2, M and E, tend to evolve slowly by accumulating nucleotide mutations while genes encoding N, viral replicase and S, tend to evolve faster (Dilucca et al., 2020). The diversity of coronaviruses can be explained at least by three main mechanisms (Anthony et al., 2017b). First, the error rate of the viral RNA polymerase. Indeed, in contrast to other RNA viruses that have high mutation rates (around 10^{-4} to 10^{-6} errors per nucleotide, which is equivalent to approximately one mutation per genome per replication cycle), coronaviruses replication fidelity is determined by the $3'$-to-$5'$ exoribonuclease activity encoded in nsp14 that proofreads RNA during replication through excision of mismatched incorporated nucleotides. But mutations are still introduced (Eckerle et al., 2007, 2010; Ferron et al., 2018). For SARS-CoV-2, the genome mutation rate was evaluated at one or two mutations per month (Kupferschmidt, 2020). Second, adenosine-to-inosine and cytosine-to-uracil transitions can occur spontaneously through oxidative damage by free radicals or through the action of the Adenosine Deaminases that Act on RNA (ADAR) [that target double-stranded RNA (dsRNA) for deamination of adenines into inosines] and the Apolipoprotein B mRNA Editing Catalytic subunit (APOBEC) (that target single-stranded nucleic acids). These transitions, mostly occurring during the RNA plus strand replication and translation, cannot be corrected by the proof-reading machinery of coronaviruses. Thus, this leads to amino-acid (a.a) substitutions that can be quite drastic for the fitness of a virus or may help the virus to escape the immune response (Di Giorgio et al., 2020; Khrustalev et al., 2020; Ratcliff and Simmonds, 2021). Third, coronaviruses have the capacity to undergo both inter-molecular recombination between 2 distinct molecules and intra-molecular recombination within the same molecule. Thus, co-infection with related strains of MHV give birth to chimeric viral genomes by inter-molecular recombination (Makino et al., 1986; Keck et al., 1988). The civet SARS-CoV-1 strain-SZ3 arose through recombination of two existing bat strains, WIV16 and Rf4092 (Hu et al., 2017). As we discussed above, strong evidence also suggest that a recombination event occurred between HCoV-229E-like viruses found in *Hipposideros* bats and HCoV-NL63-like viruses found in *Triaenops afer* bats, where the S gene is more closely related to the HCoV-229E virus (Tao et al., 2017). Recombination is also implicated in the emergence of SARS-CoV-1 and MERS-CoV (Li et al., 2005; Hon et al., 2008; Lau et al., 2015a; Anthony et al., 2017a). In addition to inter-molecular recombination, intra-molecular recombination, at virus-specific transcription regulatory sequences (TRSs), to generate a set of sg mRNAs (sg mRNAs) with common $5'$ and $3'$ ends have been also observed (Dufour et al., 2011; Sola et al., 2015). The majority of recombination events, identified in coronaviruses isolated from bats, suggest recombination hotspots around the S gene (Hu et al., 2017). For the SARS-CoV-2, a study of genetic variation in patient samples has suggested that recombination may be occurring during infections in humans (Yi, 2020) and the most frequent recombination breakpoint have been located within the S gene, as previously found for other coronaviruses (Graham and Baric, 2010; Bobay et al., 2020; Tagliamonte et al., 2020; Zhu et al., 2020). In relation to this issue, at the whole genome level, the SARS-CoV-2, as discussed above, has been suspected to possibly originate from a recombination between bat and pangolin coronaviruses. Indeed, SARS-CoV-2 is 96.2% identical to bat-SL-CoV-RaGT13, however, the S region resembles to the divergent SARS strain isolated from Pangolin-CoV-2019 (Pangolin-CoV/hCoV-19/Pangolin/Guandong 2019). Indeed, the S gene of Pangolin-CoV-2019 shared 97.5% identity with S gene of SARS-CoV-2 while, at the whole genome level, the identity between SARS-CoV-2 and Pangolin-CoV-2019 showed lower sequence identity (91.02%) (Lam et al., 2020a; Wu et al., 2020; Xiao et al., 2020; Zhang et al., 2020a; Zhou et al., 2020c). Recently, Zhou et al. (2020b) discovered a new viral strain, RmYN02, from the bat *R. malayanus*, with 97.2% identity in the ORF1ab gene to SARS-CoV-2 (and only 61.3% similarity to in the RBD motif), which harbors multiple a.a insertions at the S1/S2 cleavage site in the S protein. Altogether, CoVs mutations and recombination can offer a rapid means of acquiring gene variants that facilitate host switching, but also induce vaccine failure, as it has been already previously observed (Chen et al., 2017; Feng et al., 2018). The adaptive capacity of CoVs largely results from their large genomes, which reduce the risk of deleterious mutational errors and facilitate recombination events. Large CoV genomes are made possible by the unique proofreading capacity encoded for their

RNA-dependent polymerase. Overall, the study of recombination events and mutation processes are a key consideration in the context of the current SARS-CoV-2 pandemic as well as for future animal and zoonotic coronaviruses, especially considering the acquisition of accessory genes.

CORONAVIRUS ACCESSORY PROTEINS

Coronavirus accessory proteins have been less studied than other proteins of the virus for two main reasons (**Figure 1**). First, they are often dispensable for *in vitro* viral replication, but play a crucial role in pathogenesis and virus fitness under the natural environment of the host (Narayanan et al., 2008; Menachery et al., 2017). Second, bioinformatics approaches to detect accessory proteins are generally strong challenges and particularly for overlapping genes, which encode for accessory proteins that are often overlooked during genome annotation (Pavesi et al., 2018). Accessory genes are present in greater or lesser number in various CoVs and their functions are still incompletely understood. One of the main characteristics of the SARS-CoV-1, MERS-CoV, SARS-CoV-2, and bat SL-CoVs is the large number of accessory genes in comparison to other coronaviruses. For example, the SARS-CoV-1 retains 9 ORFs while the HCoV-HUK1 contains only three accessory ORFs (ORF4, ORF8 and one putative ORFI). The origin of these accessory genes, remains an open question as most of them have no identifiable orthologs in protein databases. It is possible that some of them have evolved through intragenomic recombination, as suggested for several of the accessory genes of SARS-CoV-1, but it is also probable that some of these extra ORFs have been acquired through ancestral recombination with RNA from cellular or heterologous viral sources. One of the best examples of this possible horizontal gene transfer is the hemagglutinin esterase (HE) gene, which is homologous to an influenza C virus gene (Luytjes et al., 1988; Zeng et al., 2008). Other candidates have also been depicted: the 2a gene found in MHV, HCoV-OC43, but also MERS-CoV (named ORF4b in MERS-CoV), which encodes a $2'5'$ phosphodiesterase that antagonizes IFN (Goldstein and Weiss, 2017), but also the ORF6 and the ORF10 of beluga whale coronavirus that encoded proteins possessing a.a similarity to human astrovirus capsid proteins and uridine kinase that, until now, has never been described in viruses (Mihindukulasuriya et al., 2008). Mutational knockdown or deletion of accessory genes has revealed, as described above, that none of them are essential for viral replication *per se*, however, deletion of some of these genes can have a profound effect on viral pathogenesis, including impacts on host innate immunity (de Haan et al., 2002). Indeed, the deletion of the gene 7 of TGEV was shown to accelerate growth kinetics and pathology (Cruz et al., 2011). This observation was also reinforced when the ORF6 of SARS-CoV-1 was transfer to a related mouse virus. The replication of the recombinant MHV was enhanced (Tangudu et al., 2007).

The SARS-CoV-1 genome possesses 9 ORFs coding for accessory proteins in the 3'-end of its genome. Two are located between the S and the E genes (ORFs 3a, 3b), five between the M

and the N genes (6, 7a, 7b, 8a, and 8b) and two within the N gene (ORF9b and a putative ORF9c). Although some of these genes are presents among human and some animal coronaviruses, no high similarities with accessory genes of other coronaviruses belonging to other genera was found (Lai and Cavanagh, 1997). Since 2003, the field of coronavirus accessory proteins has gained significant attention and reverse genetics approaches gave some clues concerning the role of some of them. For the SARS-CoV-2, a controversy still persists. Indeed, some reports described 11 predicted accessory protein ORFs (ORF3a, 3b, 3c, 3d, 6, 7a, 7b, 8, 9b, 9c, and 10) while others concluded that the SARS-CoV-2 express only ORF3a, 6, 7a, 7b, 8 or more or less ORF3b, 3c, 9b, 9c, and 10 (Gordon et al., 2020; Jungreis et al., 2020; Kim et al., 2020; Michel et al., 2020).

Accessory Proteins of SARS-CoV-1 and SARS-CoV-2

Previous studies have been shown that deletion of six of the eight ORFs of SARS-CoV-1 (ORF3a, ORF3b, ORF6, ORF7a, ORF7b, and ORF9b), alone or in combination, does not dramatically influence the level of RNA or the replication efficiency of SARS-CoV-1 *in vitro* or *in vivo* in a mouse model (Yount et al., 2005; von Brunn et al., 2007). Although most of the accessory proteins do not appear to play a central role in viral replication, their functions in the virus-host interaction and pathogenicity are determinant.

ORF3a and 3b Proteins

The SARS-CoV-1 encodes ORF3a and ORF3b proteins that can be both detected in SARS-CoV-1-infected tissues (Chan et al., 2005). The ORF3a of SARS-CoV-1 is an *O*-glycosylated 274 a.a long protein that contains three transmembrane domains and, in the C-terminus part, two intracellular protein sorting and trafficking signals (YXX and di-acidic motifs), which are important for the transport of ORF3a protein to the cell surface (**Figure 4A** and **Supplementary Table 1**). Multiple functions have been assigned to the ORF3a protein. As the ORF3a protein was shown to form a homo-tetramer complex, hypothesis have been made about a possible formation of a potassium-permeable channel-like structure. In addition of this function, the ORF3a protein was found to interact with the structural proteins, S, E, M, and the ORF7a in the Golgi apparatus proximal to the site of virus assembly and budding (Tan et al., 2004b; Zeng et al., 2004; Yuan et al., 2005a). Two studies demonstrated that the ORF3a protein is a *bona fide* viral structural protein (Ito et al., 2005; Shen et al., 2005), and that, although it is incorporated into virus-like particles (VLPs), it is not necessary for the particles formation. *In vitro* expression studies revealed that the ORF3a protein induces, G1 phase cell cycle arrest by reducing cyclin D3 expression and inhibiting retinoblastoma protein (Rb) phosphorylation (Yuan et al., 2007), and also apoptosis in Vero E6 cells (Law et al., 2005b). Moreover, electron microscopy analyses showed that the ORF3a protein induced Golgi fragmentation and an accumulation of intracellular vesicles (Freundt et al., 2010). In addition to an intracellular effect, the ORF3a protein was also described to be released from protein

FIGURE 4 | Amino acid alignment of ORFs encoding accessory proteins. **(A)** Amino acid alignment of ORF3a and 3b sequences of SARS-CoV-1 (AY274119), Bat-SL-CoV-Rp3 (DQ071615), SARS-CoV-2 (NC_045512), Bat-SL-CoV-RaTG13 (MN996532), Bat-SL-CoV-RmYN02 (JX993988), Pangolin-CoV-2019 (MT121216). No ORF3b is found in the Bat-SL-CoV-RmYN02 and Pangolin-CoV-2019 sequences, the stars indicate the stop codons. Amino acid alignment of ORF3c and 3d sequences. ORF3d is only found in the SARS-CoV-2. The asterisks indicate the stop codons. **(B)** Amino acid alignment of ORF6, 7a and 7b sequences. **(C)** Amino acid alignment of ORF8, 9b, 9c, and 10 sequences. The SARS-CoV-1 ORF8 went through a gradual deletion over the course of the epidemic and at the end of the outbreak ORF8 was divided in 2 ORFS (ORF8a and 8b). ORF10 is only found in SARS-CoV-2, Bat-SL-CoV-RaTG13 and Pangolin-CoV-2019. Sequences were analyzed using Unipro UGENE: a unified bioinformatics tollkit Okonechnikov; Golosova; Fursov. Bioinformatics 2012 28: 1,166–1,167. Amino-acids are color-coded according Clustal X. For each ORF, the SARS-CoV-2 sequence (NC_045512) was used as a reference sequence to perform the alignment.

expressing cells (Huang et al., 2006b). Although the biological role of this release remains unclear, some observations about the presence of a strong, and maybe protective, humoral response against this protein in SARS-CoV-1 patients, suggest that this ORF protein is playing an important role in viral pathogenesis (Zhong et al., 2006). Indeed, the immunization of rabbits with a synthetic peptide corresponding to the N amino terminal domain of the ORF3a protein results in the induction of neutralizing antibodies that inhibit SARS-CoV-1 infection in VeroE6 cells. Finally, the ORF3a protein of SARS-CoV-1 was also shown to activate NF-kB, a critical transcription factor involved in the activation of pro-inflammatory genes (Kanzawa et al., 2006), and the NLRP3 inflammasome by promoting TRAF3-dependent ubiquitination of adaptor protein apoptosis-associated speck-like protein containing a caspase recruitment domain (Siu et al., 2019). The SARS-CoV-2 ORF3a protein, which presents 72.4% a.a identity with the SARS-CoV-1 ORF3a protein and 97.8% with the bat-SL-CoV-RaTG13 ORF3a, retains six functional domains (I–VI) (**Figure 4A** and **Supplementary Table 1**). These functional domains are linked to virulence, infectivity, ion channel formation and virus release (Issa et al., 2020; Kern et al., 2020) and as the ORF3a protein of SARS-CoV-1, the ORF3a protein SARS-CoV-2 retains the capacity to induce apoptosis (Ren et al., 2020) and also to inhibit fusion autophagosomes with lysosomes (Zhang et al., 2021b).

The ORF3b of SARS-CoV-1 encodes for a protein (154 a.a) that has been shown to be localized to the nucleolus and/or mitochondria (Yuan et al., 2005b, 2006a; Kopecky-Bromberg et al., 2007). *In vitro* expression studies showed that the ORF3b protein of SARS-CoV-1 induces a cell growth arrest in G0/G1 phase (Yuan et al., 2005b) and can stimulate both apoptosis and necrosis in Vero E6 cells (Khan et al., 2006). The ORF3b protein of SARS-CoV-1 was described to possess, in its C-terminal part, a nuclear localization signal (a.a 134–154). Interestingly, the ORF3b homologs identified from three bat-SL-CoV strains (Bat-SL-CoV-Rf1, -Rm1 and -Rp3) were C-terminally truncated, lacked the C-terminal NLS of SARS-CoV-1 and display a different cellular localization (ORF3b of bat-SL-CoV-Rf1, -Rm1 and -Rp3 were only found in cytoplasm). The ORF3b of SARS-CoV-1 and bat-SL-CoV (Bat-SL-CoV-Rm1) were proposed to antagonize IFN by modulating the activity of IRF3 (Zhou et al., 2012), while the ORF3b protein of bat-SL-CoV-Rp3 showed no IFN antagonism. This observation is particularly important. Indeed, it is tempting to speculate that the ORF3b derived from bat SL-CoVs could have a different IFN antagonism profiles in bat cells just like the ORF3b of SARS-CoV-1. Recently, it was described the presence of a premature stop codon in the ORF3b gene of SARS-CoV-2 (**Figure 4A** and **Supplementary Table 1**). Furthermore, SARS-CoV-2 natural variants, isolated from two patients, in which a longer ORF3b reading frame was shown to be associated with severe disease and increased ability to suppress IFN response (Konno et al., 2020). As describe (**Figure 4A** and **Supplementary Table 1**), the ORF3b of SARS-CoV-2 has a high identity percentage (93.3%) with the ORF3b of the bat-SL-CoV-RaTG13 and does not contain the C-terminal part (which encompass a NLS). However, stop codons are retained in both sequences (**Figure 4A**).

Finally, some additional ORFs have been proposed for the SARS-CoV-2: ORF3c and ORF3d (**Figure 4A** and **Supplementary Table 1**). Indeed, a putative ORF, named ORF3c was proposed to overlap the ORF3a in an alternative reading frame and, recently, a ribosome profiling study confirmed that the ORF3c is indeed translated during infection. Moreover, the ORF3c (40–41 a.a predicted transmembrane protein) was found to be conserved across the subgenus *Sarbecovirus* (Firth, 2020). The ORF3d was also detected in Guangxi pangolin-CoVs and SARS-CoV-2, but not in other closely related pangolin-CoVs or bat-SL-CoVs (Chan et al., 2020; Nelson et al., 2020; Pavesi, 2020). To note, the ORF3d has been mistaken with the previously documented ORF3b in several studies (Fung et al., 2020; Ge et al., 2020; Hachim et al., 2020; Helmy et al., 2020; Yi et al., 2020; Jungreis et al., 2021).

ORF6 Protein

The SARS-CoV-1 ORF6 protein is a 63 a.a, membrane associated protein (**Figure 4B** and **Supplementary Table 1**). Its expression was detected in virus-infected Vero E6 cells as well as in the lung and intestine tissues of SARS-CoV-1 patients (Geng et al., 2005; Pewe et al., 2005). The ORF6 protein was described to mainly localized in the ER and Golgi compartments in virus-infected cells, but was also found to be incorporated into viral particles and released (Huang et al., 2007). In studies using recombinant MHV, the ORF6 protein of SARS-CoV-1 was shown to be associated with viral RNA, co-localized with replicating viral RNA on cytoplasmic vesicles and enhanced MHV viral RNA and protein synthesis (Frieman et al., 2007). Furthermore, studies showing the interaction of the ORF6 protein with the nsp8 support the notion that the ORF6 protein could play a role in virus replication (Tangudu et al., 2007). Several studies revealed that the ORF6 protein blocked IFN-induced STAT1 nuclear translocation by retention of the nuclear import-adaptor-molecule karyopherin alpha 2 in the cytoplasm (Kumar et al., 2007). Finally, the ORF6 protein of SARS-CoV-1 was shown to be required for optimal replication. As for the ORF3a and ORF3b proteins, the ORF6 protein of SARS-CoV-2 presented a strong similarity with the ORF6 protein of SARS-CoV-1 and bat-CoV-SL-RmYN02 (68.9 and 70.5%, respectively), but even more with the ORF6 protein of the bat-SL-CoV-RaTG13 and Pangolin-CoV-2019 (100 and 96.7%) (**Figure 4B** and **Supplementary Table 1**). Recently, SARS-CoV-2 ORF6 deletion variants were isolated and characterized (26-nt deletion and 34-nt deletion, respectively) (Quéromès et al., 2021) or carrying a nucleotide mutation leading to a stop codon in the ORF6 protein (Delbue et al., 2021). No significant difference in term of viral replication, as well as ISG expression, was observed between these two SARS-CoV-2 deletion variants and the reference strain. However, an upregulation of several genes, all involved in the NF-kB pathway, was noted after infection with these variants in comparison to the reference strain, suggesting that the truncated ORF6 proteins could play a role in the inflammatory host-response (Xia et al., 2020). Lately, the SARS-CoV-2 ORF6 protein function as a potent IFN antagonist (Yuen et al., 2020; Kimura et al., 2021) and also was found to block nuclear import of a wide range

of host factors through interactions with RAE1 and NUP98 (Addetia et al., 2021).

ORF 7a and ORF 7b Proteins

Both SARS-CoV-1 ORF7a and ORF7b proteins have been shown to be dispensable for virus replication (**Figure 4B**). The 122 a.a SARS-CoV-1 ORF7a protein is a type I transmembrane protein containing a signal peptide sequence (15 amino-acid), an 81 a.a lumenal domain, a 21 a.a transmembrane domain and a short C-terminal tail (**Figure 4B**). The ORF7a protein was found to be present in the perinuclear region in SARS-CoV-1-infected cells, in the ER or the ERGIC (Sims et al., 2005; Yount et al., 2005), the *trans*-Golgi network (Fielding et al., 2004; Nelson et al., 2005) and also to be incorporated into purified SARS-CoV-1 particles (Huang et al., 2006a). Several biological functions have been assigned to the SARS-CoV-1 ORF7a protein. These include the induction of apoptosis through a caspase-dependent pathway (Tan et al., 2004a; Schaecher et al., 2007), the inhibition of cellular protein synthesis, the activation of p38 mitogen-activated protein kinase (Kopecky-Bromberg et al., 2006) and the cell cycle arrest at the G0/G1 phase (Yuan et al., 2006b). The ORF7a protein was also shown to enhance pro-inflammatory cytokine production through the activation of NF-kB and JNK in A549 cells (Kanzawa et al., 2006). Finally, the ORF7a protein was found to inhibit BST-2 glycosylation, leading to a loss of its antiviral function (Kanzawa et al., 2006; Taylor et al., 2015). The detection of anti-ORF7a antibody in SARS-CoV-1 patient serum suggests that it is expressed in infected patients (Guo et al., 2004).

For the SARS-CoV-2, the ORF7a protein has been found to interacts with the ribosomal transport proteins HEATDR3 and the AAA-ATPase Midasin-1 (MDN1) and to antagonize type I IFN signaling (Gordon et al., 2020; Xia et al., 2020). Moreover, the SARS-CoV-2 ORF7a protein was depicted to be an immunomodulating factor for immune cell binding (HLA-DR/DP/DQ) and to trigger dramatic inflammatory responses (upregulation of proinflammatory cytokines, including IL-6, IL-1β, IL-8, and 38 TNF-α) (Zhou et al., 2021). A deletion of 27 a.a, which maps to a putative signal peptide within the ORF7a protein, has recently been reported in two SARS-CoV-2 isolates from two separate patients (Holland et al., 2020). This large deletion of the ORF7a is proposed to prevent redundant functions on inhibition of cellular protein synthesis also found associated with the ORF6 protein that, as we have seen above, interacts with the mRNA export proteins NUP98 and RAE1 and may inhibit cellular translation (Addetia et al., 2020, 2021). The a.a sequence of the SARS-COV-2 ORF7a protein is very close to the ORF7a of bat-SL-CoV-RaTG13 ORF7a like that of the Pangolin-CoV-2019, but much more distant than that of the bat-SL-CoV-RmYN02 (97.5,97.5, and 40% a.a identity, respectively).

The ORF7b of SARS-CoV-1 was predicted to encode an extremely hydrophobic 44 a.a protein (**Figure 4B**) (Pekosz et al., 2006). During the outbreak of SARS-CoV-1 in 2002–2003, a virus was isolated from a patient in Frankfurt (strain Frankfurt-1). Frankfurt-1 isolate was described to possess a 45-nucleotide deletion in the transmembrane domain of the ORF7b protein

(Pfefferle et al., 2009a). Interestingly, when expressed in HEK 293T cells, the full-length protein, but not the ORF7b with the deletion caused IFN-β induction and cleavage of procaspase 3. In CaCo-2 and HUH7 cells, but not in Vero cells, the Frankfurt-1 carrying the ORF7b deletion was shown to present a replicative advantage in regard to the virus containing an intact ORF7b. This effect was neither associated with changes in the induction or secretion of type I IFN, nor with altered induction of apoptosis in cell culture. However, pretreatment of cells with IFN-β caused the deleted virus to replicate to higher titers than the parental strain (Pfefferle et al., 2009a).

The ORF7b protein of SARS-CoV-2 presents a strong similarity with the ORF7b of bat-SL-CoV-RaTG13 and Pangolin-CoV-2019 (97.7 and 95.3% a.a identity, respectively) while some a.a variations are observed in the C-terminal part of the SARS-CoV-1 ORF7b protein of, bat-CoV-SL-Rp3 and bat-CoV-SL-RmYN02 (85.4, 85.7, and 85.4% a.a identity, respectively) (**Figure 4B** and **Supplementary Table 1**). The ORF7b protein of SARS-CoV-2 has been shown to inhibit type I IFN signaling by inhibiting the STAT1 phosphorylation (Xia et al., 2020).

ORF8 Proteins of SARS-CoV-1 and SARS-CoV-2

During the outbreak of SARS-CoV-1 in 2002–2003, the coding sequence of SARS-CoV-1 ORF8 (122 a.a) went through a gradual deletion over the course of the epidemic and most of the human isolates of SARS-CoV-1 presented a naturally deletion in ORF8 (deletion of 29 nucleotides) at the end of the outbreak, resulting in two ORFs: ORF8a (39 a.a) and ORF8b (84 a.a) (**Figure 4C**). The expression of the ORF8a and ORF8b proteins, produced from SARS-CoV-1 sg RNA 8, was later confirmed in infected cells and insertion of the 29 missing nucleotides into human isolates, which merges ORF8a and 8b into a continuous ORF8, showed little impact on virus growth and on RNA replication, suggesting that the deletion might not be responsible for the increased pathogenicity (Keng et al., 2006). However, lately, replication in primate cell lines, as cell line generated from the lung of a *rhinolophid* bat as well as human airway epithelial cultures, revealed that the 29nucleotides deletion conferred an attenuation of replication level (Yount et al., 2005). Less frequent deletion, which led to large deletions (82 nt or 415 nt) or complete loss of ORF8, were also observed (Muth et al., 2018). In the expression studies, the SARS-CoV-1 ORF8a protein was shown to interact with the S protein, while the ORF8b protein interacts with M, E, ORF3a and ORF7a proteins. The SARS-CoV-1 ORF8a protein was also proposed to form an ion channel and the ORF8b protein was showed to down-regulated the E protein level (Chen et al., 2011). Finally, recently, the aggregation of the ORF8b proteins of SARS-CoV-1 was shown to induce ER stress, lysosomal damage, and subsequent activation of autophagy and lysosome machinery (Keng et al., 2006; Shi et al., 2019).

For the SARS-CoV-2, the protein encodes by this ORF8 is similar to the ORF8 protein of bat-SL-CoVs (Bat-CoV-SL-RaTG13 and Pangolin-CoV-2019: 95% a.a identity for both sequences), but presents a low homology with the SARS-CoV-1,

bat-CoV-SL-RmYN02 and bat-SL-CoV-Rp3 ORF8 protein (26.2, 57.9, and 56.2%) (**Figure 4C** and **Supplementary Table 1**). The ORF8 protein of SARS-CoV-2 has been shown to strongly inhibit the IFN-stimulated response element (ISRE) after IFN-β treatment. The crystallization of SARS-CoV-2 ORF8 revealed a structure similar to the one of SARS-CoV-1 ORF7a, with the addition of two dimerization interfaces unique to SARS-CoV-2 ORF8 (Li et al., 2020, p. 8). In contrast to the SARS-CoV-1, the SARS-CoV-2 isolates presenting a large deletion that truncates the ORF7b and ORF8 (382-nt deletion) was showed to induce a significantly higher replication level than the wild type (Flower et al., 2020), but less severe infection and lower concentrations of proinflammatory cytokines, chemokines, and growth factors (Young et al., 2020). Recently, the ORF8 protein was shown to contributes to cytokine storm during SARS-CoV-2 infection by activating IL-17 pathway (Lin et al., 2021).

ORF9b and 9c Proteins

Overlapping reading frames have been discovered in many organisms, but they are a common feature in viruses (Pavesi et al., 2018). In this context, the same nucleotide sequence codes for more than one protein in a different reading frame. These overlaps are typically assumed to be a form of genome compression, allowing the virus to increase its repertoire of proteins without increasing its genome length (Pavesi et al., 2018; Pavesi, 2019). Proteins created by gene overlaps (sometimes called "overprinting" or "overprinted," or "novel") are typically accessory proteins that play a role in viral pathogenicity or spreading (Rancurel et al., 2009; Pavesi, 2019). Several alpha-CoVs (as HCoV-229E and HCoV-NL63), but also beta-CoVs (as MHV, BCoV, HCoV-OC43, and HCoV-HKU1) have one or several alternative ORFs (from one to three) embedded entirely within their nucleocapsid gene (N). In the beta-CoVs genus, there is a relative sequence conservation of one ORF named "I" in the two coronaviruses MHV and BCoV (44.1% a.a identity), but a very low similarity is detected with the SARS-CoV-1 or SARS-CoV-2 ORF-9b (**Supplementary Table 2**). Functionally, the ORF-9b homolog, named "I," in the +1 reading frame with respect to the N gene, of MHV was shown to act as an accessory structural protein that is not essential for viral infection, but confers a growth advantage (Fischer et al., 1997). Endemic human alpha-CoVs (HCoV-229E and HCoV-NL63) and human beta-CoVs (HCoV-OC43 and HCoV-HKU1) also present alternative ORFs embedded entirely within their nucleocapsid gene (N), but again a very low similarity with the ORFs Ia or I of HCoV-229E, HCoV-NL63, HCoV-OC43, HCoV-HKU1 was observed with the ORF-9b SARS-CoV-1 or SARS-CoV-2 (**Supplementary Table 2**). The SARS-CoV-1 ORF9b protein (98 a.a residues, **Figure 4C** and **Supplementary Table 1**) has been shown to be expressed in infected cells, from an internal ORF in the N gene by a leaky ribosomal scanning (Xu et al., 2009), but was also identified in the intestinal surface enterocytes and pneumocytes (Chan et al., 2005). Viruses are known to operate with several non-canonical translational mechanisms (re-initiation, ribosomal shunt, internal initiation, and leaky scanning), which may be

programmed or incidental. The latter can be considered as translation noise, not subject to strong purifying selection, and therefore generally not phylogenetically preserved. However, in some cases, programmed exceptions tend to be subjected to purifying selection and phylogenetic preservation. Leaky scanning is the most common process used by RNA viruses to facilitate the access to ORFs. In this mechanism, a significant proportion of scanning ribosomes fail to initiate at the first AUG codon (suboptimal context surrounding the first AUG codon). In mammals, the optimal context for recognition of the AUG Start codon is GCCRCCAUGG, including a purine (R) at position −3 and a G at position +4 highly conserved and is especially important in the absence of an A at position −3 (Kozak, 1986, 2002) but instead, continue scanning until they reach an alternative initiation codon further downstream (**Figure 5A**). Leaky scanning is employed by numerous viruses (plants viruses and mammalian viruses). Besides suboptimal context surrounding the first AUG codon, leaky scanning may also be promoted by a number of other mechanisms. If an AUG codon is very close to the 5′ end of the transcript, then it is often not recognized efficiently (Sedman et al., 1990; Kozak, 1991). The close proximity of a downstream AUG codon to a preceding AUG codon (e.g., within approximatively 10 nt) can also increase the efficiency of leaky scanning (Williams and Lamb, 1989; Matsuda and Dreher, 2006). Leaky scanning can also be promoted by short upstream ORFs. In SARS-CoV-1, as well as SARS-COV-2, bat-CoV-SL-Rp3, -RaTG13, -RmNY02 and Pangolin-CoV-2019, the start codon for ORF9b is very close to the start codon for ORFN, with only 10 nucleotides in between (**Figure 5B**). Antibodies against the ORF9b protein were detected in the sera of SARS-CoV-1 patients demonstrating that this protein is produced during infection (Chow et al., 2006). The expression level of the SARS-CoV-1 ORF9b was shown to be relatively weaker when compared to that of the N protein (Xu et al., 2009) and to be dispensable for virus replication *in vitro*. The crystal structure of the SARS-CoV-1 ORF9b protein was solved in 2006 and revealed a novel dimeric tent-like β structure with an amphipathic surface and a central hydrophobic cavity, which binds lipid molecules that probably allows its association with intracellular vesicles (Meier et al., 2006). This hypothesis was reinforced by the observation, in mammalian cells, of the SARS-CoV-1 ORF-9b protein distributed diffusely within cytoplasm and nucleus (HeLa cells) (von Brunn et al., 2007) or associated with intracellular vesicles (HEK 293T cells) (Meier et al., 2006). Intraviral protein interactions study, using a yeast-two-hybrid approach, revealed that ORF9b protein is the most dominant interactor of the accessory proteins (The ORF9b protein showed interactions with at least 10 nsps, the E protein and four accessory proteins: the ORF8a, ORF8b, ORF9b, and ORF9c) (von Brunn et al., 2007). Currently, there is no explanation for this large number of interactions and the precise function of the SARS-CoV-1 ORF9b protein is still not understood, although it has been proposed that the ORF9b protein might contribute to virus assembly as a membrane-attachment point for other viral proteins, like N protein (Meier et al., 2006). As some other ORF proteins, the SARS-CoV-1 ORF9b protein was found to be a virion-associated protein
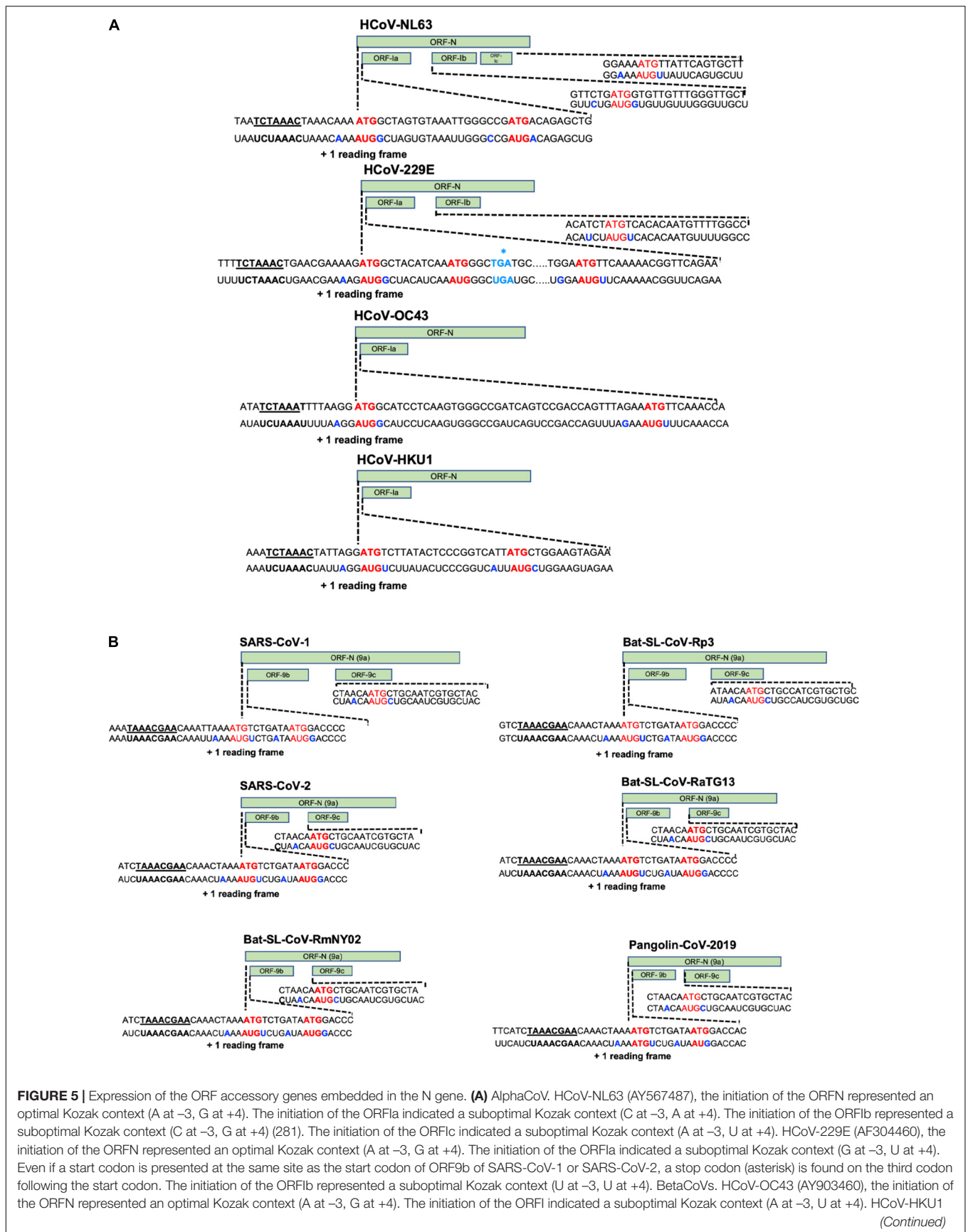
FIGURE 5 | Expression of the ORF accessory genes embedded in the N gene. **(A)** AlphaCoV. HCoV-NL63 (AY567487), the initiation of the ORFN represented an optimal Kozak context (A at −3, G at +4). The initiation of the ORFIa indicated a suboptimal Kozak context (C at −3, A at +4). The initiation of the ORFIb represented a suboptimal Kozak context (C at −3, G at +4) (281). The initiation of the ORFIc indicated a suboptimal Kozak context (A at −3, U at +4). HCoV-229E (AF304460), the initiation of the ORFN represented an optimal Kozak context (A at −3, G at +4). The initiation of the ORFIa indicated a suboptimal Kozak context (G at −3, U at +4). Even if a start codon is presented at the same site as the start codon of ORF9b of SARS-CoV-1 or SARS-CoV-2, a stop codon (asterisk) is found on the third codon following the start codon. The initiation of the ORFIb represented a suboptimal Kozak context (U at −3, U at +4). BetaCoVs. HCoV-OC43 (AY903460), the initiation of the ORFN represented an optimal Kozak context (A at −3, G at +4). The initiation of the ORFI indicated a suboptimal Kozak context (A at −3, U at +4). HCoV-HKU1

(Continued)

(Xu et al., 2009). More recently, it has been shown that the SARS-CoV-1 ORF9b protein limits host cell IFN responses by targeting several proteins of the MAVS signalosome to degradation. Finally, transient ORF-9b expression was showed to result in a strong induction of autophagy (Shi et al., 2014). Another overlapping gene was proposed in the SARS-CoV-1 genome, the ORF9c. It has been reported within the nucleocapsid gene and coding for a predicted protein of 70 a.a (Marra et al., 2003; Rota et al., 2003). Until now, the SARS-CoV-1 ORF9c expression has not been demonstrated (**Figure 4C** and **Supplementary Table 1**). Since the SARS-CoV-2 outbreak, Ribo-seq and sequencing to systematically delineate the landscape of translated ORFs of SARS-CoV-2 and their expression during infection were done (Gordon et al., 2020; Kim et al., 2020). RNA-seq and ribosome profiling analyses of SARS-CoV-2 infected Vero E6 cells revealed out of frame internal initiations within ORFN (ORF9b and ORF9c) (Finkel et al., 2020) supporting the conclusion that these two SARS-CoV-2 proteins are expressed in SARS-CoV-2-infected cells. In addition, Gordon et al. (2020) demonstrated that stable expression of the SARS-CoV-2 ORF9b and ORF9c can occur in transfected HEK 293T and interactome studies showed that the ORFs proteins interact with several proteins as the mitochondrial import receptor subunit TOM70 (TOMM70) and the mitochondrial chaperone BCS1 (BCS1L), respectively. Finally, transcriptomic and proteomic studies, revealed that the ORF9b transcript and ORF9b protein (97 a.a) are detected in of SARS-CoV-2-infected cells (Bojkova et al., 2020; Nomburg et al., 2020) supporting, one more time, the notion that ORF9b is a *bona fide* SARS-CoV-2 protein. As for the SARS-CoV-1, antibodies against SARS-CoV-2 ORF9b were detected in the sera of convalescent patients (Personal communication) (Guo et al., 2004; Qiu et al., 2005; Jiang et al., 2020a). The SARS-CoV-2 ORF9b protein was depicted to inhibit the type I IFN response through its interaction with the outer membrane mitochondrial adaptor TOM70 (Jiang et al., 2020b). Recently, the crystal structure of ORF9b in complex with the human TOM70 was resolved (Gao et al., 2021). Moreover, SARS-CoV-2 ORF9b antagonizes types I and III interferons by targeting multiple components of the RIG-I/MDA-5-MAVS, TLR3-TRIF, and cGAS-STING signaling pathways (Han et al., 2021). As others ORFs, the SARS-CoV-2 ORF9b protein was found to be very similar to the ORF9b of bat-CoV-RaTG13 and the Pangolin-CoV-2019 (92.8% a.a identity), but more distant from the SARS-CoV-1, bat-SL-CoV-Rp3 and bat-SL-CoV-RmYN02 (77.3, 72.2, and 73.5% a.a identity, respectively, **Supplementary Table 1**). In addition to the ORF9b, a predicted ORF9c retained in the SARS-CoV-1, bat SL-CoVs and Pangolin-CoV-2019 has

also been described in the SARS-CoV-2 genome (**Figure 4C**). Dominguez Andres et al. (2020) report that the viral SARS-CoV-2 ORF9c protein (also named ORF14) is unstable and has acquired a transmembrane domain. They found that, when expressed in a human lung epithelial cell line (A549), the SARS-CoV-2 ORF9c interfered with IFN signaling, and several other pathways. A strong similarity is observed between the ORF9c proteins of SARS-CoV-2 (73 a.a), bat-SL-CoV-RatGT13 and Pangolin-CoV-2019, where the SARS-CoV-1 ORF9c protein appears to be closer to the ORF9c of bat-SL-CoV-Rp3.

ORF10 Protein

Unlike the SARS-CoV-1, the bat-CoV-SL-RaTG13, the Pangolin-CoV-2019 and the SARS-CoV-2 possess a supplementary putative extra ORF named ORF10, which may encode for a protein of 38 a.a (**Figure 4C** and **Supplementary Table 1**). Recent studies demonstrated that very few transcripts containing ORF10 are detected in SARS-CoV-2-infected cells and several teams have reported that ORF10 is not a coding region (Caly et al., 2020; Davidson et al., 2020; Kim et al., 2020). The disease induced by the SARS-CoV-2 variant strains, in which the ORF10 gene was prematurely terminated, was not attenuated and *in vitro* replication analysis showed no difference compared to the wild-type virus (Jungreis et al., 2020; Pancer et al., 2020).

CONCLUSION

Highly pathogenic coronaviruses have emerged regularly since the beginning of the 21st century, resulting in epidemics: SARS-CoV-1 in 2002–2003 and MERS-CoV in 2012 and a SARS-CoV-2 pandemic since 2019. The SARS-CoV-1, MERS-CoV and SARS-CoV-2 are all human coronaviruses closely related to bat-SL-CoVs. It can be tempting to speculate that the direct ancestor of these viruses may have arisen from sequential recombination events between the precursors of the bat-SL-CoVs prior to spillover to intermediate hosts or directly to humans. However, it is also important to note that after the SARS-CoV-1 outbreak in 2002–2003, bats were sampled extensively for coronaviruses and other viruses, although the latter probably represent a major initial reservoir. It is also possible that a similar, as yet unexplored, diversity of viruses exists in other animals. So, continuing surveillance in bats, but also in others animals (reservoir hosts or intermediate hosts), is vital to uncover the origin of these viruses but also to prevent the emergence of a new

coronavirus disease. A better understanding of the bat coronavirus ecology, physiology and immunity response is also critical to elucidate the origin, the diversity and the emergence of coronaviruses. As we discussed above, the SARS-CoV-1, MERS-CoV, and SARS-CoV-2 encode several accessory proteins. Reverse genetic studies have demonstrated that none of them are essential for virus replication *per se*, but most of them appear to play a crucial role in virus pathogenesis. As the pathophysiological significances of these accessory proteins are poorly understood, further structural, evolution and functional studies are required to shed more-light on their precise functions and also to propose new antiviral targets. Increasing evidences show that preserving natural habitats reduces the risk of emergence or re-emergence of pathogens from wildlife. Ecosystems that are heavily used by humans (modern agriculture, deforestation, and urbanization) contain a tremendous variety of wildlife that carry pathogens (for many, still totally unknown) capable of infecting people. As the human population expands in an increasingly globalized world together with socio-economical changes, human contact with wildlife will continue inevitably to increase. This increases the risk of zoonotic viruses of emerge, including CoVs. To date, coronavirus surveillance has been almost entirely restricted to China. More vigorous field research efforts tracking the circulation of beta-CoVs more generally is needed across a broader global range if we are to avoid future repeats of the COVID-19 pandemic. Changes in our behaviors, but also integrative research across disciplines, by applying a "One Health" concept from field to lab, are urgently needed to prevent future zoonotic events and the repetition of the current episode.

AUTHOR CONTRIBUTIONS

Conceptualization of the article and writing of the original draft was done by NC.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.682603/full#supplementary-material

Supplementary Table 1 | Sequences were analyzed using Unipro UGENE: a unified bioinformatics tollkit Okonechnikov; Golosova; Fursov. Bioinformatics 2012 28: 1,166–1,167. For each ORF, the SARS-CoV-2 sequence (NC_045512) was used as a reference sequence to perform the alignment. Identity was calculated using SIM – Alignment Tool for protein sequences (www.expasy.org/).

Supplementary Table 2 | Sequences were analyzed using Unipro UGENE: a unified bioinformatics tollkit Okonechnikov; Golosova; Fursov. Bioinformatics 2012 28: 1,166–1,167. For each ORF, the SARS-CoV-2 sequence (NC_045512) was used as a reference sequence to perform the alignment. Identity was calculated using SIM - Alignment Tool for protein sequences (www.expasy.org/).

REFERENCES

Acharya, D., Liu, G., and Gack, M. U. (2020). Dysregulation of type I interferon responses in COVID-19. *Nat. Rev. Immunol.* 20, 397–398. doi: 10.1038/s41577-020-0346-x

Addetia, A., Lieberman, N. A. P., Phung, Q., Hsiang, T.-Y., Xie, H., Roychoudhury, P., et al. (2021). SARS-CoV-2 ORF6 Disrupts bidirectional nucleocytoplasmic transport through interactions with Rae1 and Nup98. *mBio* 12, e65–e21. doi: 10.1128/mBio.00065-21

Addetia, A., Xie, H., Roychoudhury, P., Shrestha, L., Loprieno, M., Huang, M.-L., et al. (2020). Identification of multiple large deletions in ORF7a resulting in in-frame gene fusions in clinical SARS-CoV-2 isolates. *J. Clin. Virol.* 129:104523. doi: 10.1016/j.jcv.2020.104523

Ahmad, T., Haroon, H., Baig, M., and Hui, J. (2020). Coronavirus Disease 2019 (COVID-19) pandemic and economic impact. *Pak. J. Med. Sci.* 36, S73–S78. doi: 10.12669/pjms.36.COVID19-S4.2638

Ahn, M., Anderson, D. E., Zhang, Q., Tan, C. W., Lim, B. L., Luko, K., et al. (2019). Dampened NLRP3-mediated inflammation in bats and implications for a special viral reservoir host. *Nat. Microbiol.* 4, 789–799. doi: 10.1038/s41564-019-0371-3

Alfaraj, S. H., Al-Tawfiq, J. A., Assiri, A. Y., Alzahrani, N. A., Alanazi, A. A., and Memish, Z. A. (2019). Clinical predictors of mortality of Middle East Respiratory Syndrome Coronavirus (MERS-CoV) infection: a cohort study. *Travel Med. Infect. Dis.* 29, 48–50. doi: 10.1016/j.tmaid.2019.03.004

Almeida, J. D., and Tyrrell, D. A. J. (1967). The morphology of three previously uncharacterized human respiratory viruses that grow in organ culture. *J. Gen. Virol.* 1, 175–178. doi: 10.1099/0022-1317-1-2-175

Alshukairi, A. N., Zheng, J., Zhao, J., Nehdi, A., Baharoon, S. A., Layqah, L., et al. (2018). High Prevalence of MERS-CoV infection in camel workers in Saudi Arabia. *mBio* 9:e01985–18. doi: 10.1128/mBio.01985-18

Al-Tawfiq, J. A., and Memish, Z. A. (2019). Middle East respiratory syndrome coronavirus in the last two years: health care workers still at risk. *Am. J. Infect. Control.* 47, 1167–1170. doi: 10.1016/j.ajic.2019.04.007

Al-Tawfiq, J. A., and Memish, Z. A. (2020). Middle east respiratory syndrome coronavirus and severe acute respiratory syndrome coronavirus. *Semin. Respir. Crit. Care Med.* 41, 568–578. doi: 10.1055/s-0040-1709160

Andersen, K. G., Rambaut, A., Lipkin, W. I., Holmes, E. C., and Garry, R. F. (2020). The proximal origin of SARS-CoV-2. *Nat. Med.* 26, 450–452. doi: 10.1038/s41591-020-0820-9

Anthony, S. J., Gilardi, K., Menachery, V. D., Goldstein, T., Ssebide, B., Mbabazi, R., et al. (2017a). Further evidence for bats as the evolutionary source of middle east respiratory syndrome coronavirus. *mBio* 8, e00373–17. doi: 10.1128/mBio.00373-17

Anthony, S. J., Johnson, C. K., Greig, D. J., Kramer, S., Che, X., Wells, H., et al. (2017b). Global patterns in coronavirus diversity. *Virus Evol.* 3:vex012. doi: 10.1093/ve/vex012

Arndt, A. L., Larson, B. J., and Hogue, B. G. (2010). A conserved domain in the coronavirus membrane protein tail is important for virus assembly. *J. Virol.* 84, 11418–11428. doi: 10.1128/JVI.01131-10

Baker, M. L., Tachedjian, M., and Wang, L.-F. (2010). Immunoglobulin heavy chain diversity in Pteropid bats: evidence for a diverse and highly specific antigen binding repertoire. *Immunogenetics* 62, 173–184. doi: 10.1007/s00251-010-0425-4

Balboni, A., Palladini, A., Bogliani, G., and Battilani, M. (2011). Detection of a virus related to betacoronaviruses in Italian greater horseshoe bats. *Epidemiol. Infect.* 139, 216–219. doi: 10.1017/S0950268810001147

Banerjee, A., Baker, M. L., Kulcsar, K., Misra, V., Plowright, R., and Mossman, K. (2020). Novel insights into immune systems of bats. *Front. Immunol.* 11:26. doi: 10.3389/fimmu.2020.00026

Banerjee, A., Falzarano, D., Rapin, N., Lew, J., and Misra, V. (2019a). Interferon Regulatory Factor 3-mediated signaling limits middle-east respiratory syndrome (MERS) coronavirus propagation in cells from an insectivorous bat. *Viruses* 11:152. doi: 10.3390/v11020152

Banerjee, A., Kulcsar, K., Misra, V., Frieman, M., and Mossman, K. (2019b). Bats and coronaviruses. *Viruses* 11:41. doi: 10.3390/v11010041

Banerjee, A., Rapin, N., Bollinger, T., and Misra, V. (2017). Lack of inflammatory gene expression in bats: a unique role for a transcription repressor. *Sci. Rep.* 7:2232. doi: 10.1038/s41598-017-01513-w

Banerjee, A., Rapin, N., Miller, M., Griebel, P., Zhou, Y., Munster, V., et al. (2016). Generation and Characterization of Eptesicus fuscus (Big brown bat) kidney cell lines immortalized using the Myotis polyomavirus large T-antigen. *J. Virol. Methods* 237, 166–173. doi: 10.1016/j.jviromet.2016.09.008

Beach, J. R., and Schalm, O. W. (1936). A filterable virus, distinct from that of laryngotracheitis, the cause of a respiratory disease of chicks. *Poult. Sci.* 15, 199–206. doi: 10.3382/ps.0150199

Blanco-Melo, D., Nilsson-Payant, B. E., Liu, W.-C., Uhl, S., Hoagland, D., Møller, R., et al. (2020). Imbalanced host response to SARS-CoV-2 drives development of COVID-19. *Cell* 181, 1036–1045.e9. doi: 10.1016/j.cell.2020.04.026

Bobay, L.-M., O'Donnell, A. C., and Ochman, H. (2020). Recombination events are concentrated in the spike protein region of Betacoronaviruses. *PLoS Genet.* 16:e1009272. doi: 10.1371/journal.pgen.1009272

Bojkova, D., Klann, K., Koch, B., Widera, M., Krause, D., Ciesek, S., et al. (2020). Proteomics of SARS-CoV-2-infected host cells reveals therapy targets. *Nature* 583, 469–472. doi: 10.1038/s41586-020-2332-7

Boscarino, J. A., Logan, H. L., Lacny, J. J., and Gallagher, T. M. (2008). Envelope protein palmitoylations are crucial for murine coronavirus assembly. *J. Virol.* 82, 2989–2999. doi: 10.1128/JVI.01906-07

Boson, B., Legros, V., Zhou, B., Siret, E., Mathieu, C., Cosset, F.-L., et al. (2020). The SARS-CoV-2 Envelope and Membrane proteins modulate maturation and retention of the Spike protein, allowing assembly of virus-like particles. *J. Biol. Chem.* 296:100111. doi: 10.1074/jbc.RA120.016175

Brook, C. E., and Dobson, A. P. (2015). Bats as 'special' reservoirs for emerging zoonotic pathogens. *Trends Microbiol.* 23, 172–180. doi: 10.1016/j.tim.2014.12.004

Caly, L., Druce, J., Roberts, J., Bond, K., Tran, T., Kostecki, R., et al. (2020). Isolation and rapid sharing of the 2019 novel coronavirus (SARS-CoV-2) from the first patient diagnosed with COVID-19 in Australia. *Med. J. Aust.* 212, 459–462. doi: 10.5694/mja2.50569

Chan, C. M., Lau, S. K. P., Woo, P. C. Y., Tse, H., Zheng, B.-J., Chen, L., et al. (2009). Identification of major histocompatibility complex class i c molecule as an attachment factor that facilitates coronavirus HKU1 spike-mediated infection. *J. Virol.* 83, 1026–1035. doi: 10.1128/JVI.01387-08

Chan, J. F.-W., Kok, K.-H., Zhu, Z., Chu, H., To, K. K.-W., Yuan, S., et al. (2020). Genomic characterization of the 2019 novel human-pathogenic coronavirus isolated from a patient with atypical pneumonia after visiting Wuhan. *Emerg. Microbes Infect.* 9, 221–236. doi: 10.1080/22221751.2020.1719902

Chan, J. F. W., Lau, S. K. P., To, K. K. W., Cheng, V. C. C., Woo, P. C. Y., and Yuen, K.-Y. (2015). Middle east respiratory syndrome coronavirus: another zoonotic betacoronavirus causing SARS-Like Disease. *Clin. Microbiol. Rev.* 28, 465–522. doi: 10.1128/CMR.00102-14

Chan, W. S., Wu, C., Chow, S. C. S., Cheung, T., To, K.-F., Leung, W.-K., et al. (2005). Coronaviral hypothetical and structural proteins were found in the intestinal surface enterocytes and pneumocytes of severe acute respiratory syndrome (SARS). *Mod. Pathol.* 18, 1432–1439. doi: 10.1038/modpathol.3800439

Channappanavar, R., Fehr, A. R., Vijay, R., Mack, M., Zhao, J., Meyerholz, D. K., et al. (2016). Dysregulated type i interferon and inflammatory monocyte-macrophage responses cause lethal pneumonia in SARS-CoV-Infected Mice. *Cell Host Microbe* 19, 181–193. doi: 10.1016/j.chom.2016.01.007

Channappanavar, R., Fehr, A. R., Zheng, J., Wohlford-Lenane, C., Abrahante, J. E., Mack, M., et al. (2019). IFN-I response timing relative to virus replication determines MERS coronavirus infection outcomes. *J. Clin. Invest.* 129, 3625–3639. doi: 10.1172/JCI126363

Cheever, F. S., Daniels, J. B., Pappenheimer, A. M., and Bailey, O. T. (1949). A murine virus (JHM) causing disseminated encephalomyelitis with extensive destruction of myelin. *J. Exp. Med.* 90, 181–194. doi: 10.1084/jem.90.3.181

Chen, C.-C., Krüger, J., Sramala, I., Hsu, H.-J., Henklein, P., Chen, Y.-M. A., et al. (2011). ORF8a of SARS-CoV forms an ion channel: experiments and molecular dynamics simulations. *Biochim. Biophys. Acta* 1808, 572–579. doi: 10.1016/j.bbamem.2010.08.004

Chen, N., Li, S., Zhou, R., Zhu, M., He, S., Ye, M., et al. (2017). Two novel porcine epidemic diarrhea virus (PEDV) recombinants from a natural recombinant and distinct subtypes of PEDV variants. *Virus Res.* 242, 90–95. doi: 10.1016/j.virusres.2017.09.013

Cheung, C. Y., Poon, L. L. M., Ng, I. H. Y., Luk, W., Sia, S.-F., Wu, M. H. S., et al. (2005). Cytokine Responses in Severe Acute Respiratory Syndrome Coronavirus-Infected Macrophages In Vitro: Possible Relevance to Pathogenesis. *J. Virol.* 79, 7819–7826. doi: 10.1128/JVI.79.12.7819-7826.2005

Chow, S. C. S., Ho, C. Y. S., Tam, T. T. Y., Wu, C., Cheung, T., Chan, P. K. S., et al. (2006). Specific epitopes of the structural and hypothetical proteins elicit variable humoral responses in SARS patients. *J. Clin. Pathol.* 59, 468–476. doi: 10.1136/jcp.2005.029868

Collins, A. R. (1993). HLA class i antigen serves as a receptor for human coronavirus OC43. *Immunol. Invest.* 22, 95–103. doi: 10.3109/08820139309063393

Corman, V. M., Baldwin, H. J., Tateno, A. F., Zerbinati, R. M., Annan, A., Owusu, M., et al. (2015). Evidence for an ancestral association of human coronavirus 229E with Bats. *J. Virol.* 89, 11858–11870. doi: 10.1128/JVI.01755-15

Corman, V. M., Eckerle, I., Memish, Z. A., Liljander, A. M., Dijkman, R., Jonsdottir, H., et al. (2016). Link of a ubiquitous human coronavirus to dromedary camels. *Proc. Natl. Acad. Sci. U.S.A.* 113, 9864–9869. doi: 10.1073/pnas.1604472113

Corman, V. M., Ithete, N. L., Richards, L. R., Schoeman, M. C., Preiser, W., Drosten, C., et al. (2014a). Rooting the phylogenetic tree of middle east respiratory syndrome coronavirus by characterization of a conspecific virus from an african bat. *J. Virol.* 88, 11297–11303. doi: 10.1128/JVI.01498-14

Corman, V. M., Jores, J., Meyer, B., Younan, M., Liljander, A., Said, M. Y., et al. (2014b). Antibodies against MERS coronavirus in dromedary camels, Kenya, 1992–2013. *Emerg. Infect. Dis.* 20, 1319–1320. doi: 10.3201/eid2008.140596

Cowled, C., Baker, M., Tachedjian, M., Zhou, P., Bulach, D., and Wang, L.-F. (2011). Molecular characterisation of Toll-like receptors in the black flying fox Pteropus alecto. *Dev. Comp. Immunol.* 35, 7–18. doi: 10.1016/j.dci.2010.07.006

Cowled, C., Baker, M. L., Zhou, P., Tachedjian, M., and Wang, L.-F. (2012). Molecular characterisation of RIG-I-like helicases in the black flying fox, Pteropus alecto. *Dev. Comp. Immunol.* 36, 657–664. doi: 10.1016/j.dci.2011.11.008

Crameri, G., Todd, S., Grimley, S., McEachern, J. A., Marsh, G. A., Smith, C., et al. (2009). Establishment, immortalisation and characterisation of pteropid bat cell lines. *PLoS One* 4:e8266. doi: 10.1371/journal.pone.0008266

Crossley, B. M., Barr, B. C., Magdesian, K. G., Ing, M., Mora, D., Jensen, D., et al. (2010). Identification of a novel coronavirus possibly associated with acute respiratory syndrome in Alpacas (*Vicugna Pacos*) in California, 2007. *J. VET Diagn. Invest.* 22, 94–97. doi: 10.1177/104063871002200118

Cruz, J. L. G., Sola, I., Becares, M., Alberca, B., Plana, J., Enjuanes, L., et al. (2011). Coronavirus gene 7 counteracts host defenses and modulates virus virulence. *PLoS Pathog.* 7:e1002090. doi: 10.1371/journal.ppat.1002090

Cui, J., Li, F., and Shi, Z.-L. (2019). Origin and evolution of pathogenic coronaviruses. *Nat. Rev. Microbiol.* 17, 181–192. doi: 10.1038/s41579-018-0118-9

Davidson, A. D., Williamson, M. K., Lewis, S., Shoemark, D., Carroll, M. W., Heesom, K. J., et al. (2020). Characterisation of the transcriptome and proteome of SARS-CoV-2 reveals a cell passage induced in-frame deletion of the furin-like cleavage site from the spike glycoprotein. *Genome Med.* 12:68. doi: 10.1186/s13073-020-00763-0

de Haan, C. A. M., Masters, P. S., Shen, X., Weiss, S., and Rottier, P. J. M. (2002). The group-specific murine coronavirus genes are not essential, but their deletion, by reverse genetics, is attenuating in the natural host. *Virology* 296, 177–189. doi: 10.1006/viro.2002.1412

Delbue, S., D'Alessandro, S., Signorini, L., Dolci, M., Pariani, E., Bianchi, M., et al. (2021). Isolation of SARS-CoV-2 strains carrying a nucleotide mutation, leading to a stop codon in the ORF 6 protein. *Emerg. Microbes Infect.* 10, 252–255. doi: 10.1080/22221751.2021.1884003

Di Giorgio, S., Martignano, F., Torcia, M. G., Mattiuz, G., and Conticello, S. G. (2020). Evidence for host-dependent RNA editing in the transcriptome of SARS-CoV-2. *Sci. Adv.* 6:eabb5813. doi: 10.1126/sciadv.abb5813

Dilucca, M., Forcelloni, S., Georgakilas, A. G., Giansanti, A., and Pavlopoulou, A. (2020). Codon usage and phenotypic divergences of SARS-CoV-2 Genes. *Viruses* 12:498. doi: 10.3390/v12050498

Dominguez Andres, A., Feng, Y., Campos, A. R., Yin, J., Yang, C.-C., James, B., et al. (2020). SARS-CoV-2 ORF9c Is a Membrane-associated protein that suppresses antiviral responses in cells. *bioRxiv* [Preprint]. doi: 10.1101/2020.08.18.256776

Donaldson, E. F., Haskew, A. N., Gates, J. E., Huynh, J., Moore, C. J., and Frieman, M. B. (2010). Metagenomic analysis of the viromes of three North American bat species: viral diversity among different bat species that share a common habitat. *J. Virol.* 84, 13004–13018. doi: 10.1128/JVI.01255-10

Doyle, L. P., and Hutchings, L. M. (1946). A transmissible gastroenteritis in pigs. *J. Am. Vet. Med. Assoc.* 108, 257–259.

Drexler, J. F., Gloza-Rausch, F., Glende, J., Corman, V. M., Muth, D., Goettsche, M., et al. (2010). Genomic characterization of severe acute respiratory syndrome-related coronavirus in european bats and classification of coronaviruses based on partial RNA-Dependent RNA polymerase gene sequences. *J. Virol.* 84, 11336–11349. doi: 10.1128/JVI.00650-10

Dufour, D., Mateos-Gomez, P. A., Enjuanes, L., Gallego, J., and Sola, I. (2011). Structure and functional relevance of a transcription-regulating sequence involved in coronavirus discontinuous RNA synthesis. *J. Virol.* 85, 4963–4973. doi: 10.1128/JVI.02317-10

Eckerle, L. D., Becker, M. M., Halpin, R. A., Li, K., Venter, E., Lu, X., et al. (2010). Infidelity of SARS-CoV Nsp14-exonuclease mutant virus replication is revealed by complete genome sequencing. *PLoS Pathog.* 6:e1000896. doi: 10.1371/journal.ppat.1000896

Eckerle, L. D., Lu, X., Sperry, S. M., Choi, L., and Denison, M. R. (2007). High fidelity of murine hepatitis virus replication is decreased in nsp14 exoribonuclease mutants. *J. Virol.* 81, 12135–12144. doi: 10.1128/JVI.01296-07

Esper, F., Weibel, C., Ferguson, D., Landry, M. L., and Kahn, J. S. (2006). Coronavirus HKU1 Infection in the United States. *Emerg. Infect. Dis.* 12, 775–779. doi: 10.3201/eid1205.051316

Feng, K. Y., Chen, T., Zhang, X., Shao, G. M., Cao, Y., Chen, D. K., et al. (2018). Molecular characteristic and pathogenicity analysis of a virulent recombinant avain infectious bronchitis virus isolated in China. *Poult. Sci.* 97, 3519–3531. doi: 10.3382/ps/pey237

Feng, Z., Diao, B., Wang, R., Wang, G., Wang, C., Tan, Y., et al. (2020). The novel severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) directly decimates human spleens and lymph nodes. *medRxiv* [Preprint]. doi: 10.1101/2020.03.27.20045427

Fenton, M. B., and Simmons, N. B. (2015). *Bats: A World of Science and Mystery*. Chicago, IL: University of Chicago Press, doi: 10.7208/chicago/9780226065267.001.0001

Ferron, F., Subissi, L., Silveira De Morais, A. T., Le, N. T. T., Sevajol, M., Gluais, L., et al. (2018). Structural and molecular basis of mismatch correction and ribavirin excision from coronavirus RNA. *Proc. Natl. Acad. Sci. U.S.A.* 115, E162–E171. doi: 10.1073/pnas.1718806115

Fielding, B. C., Tan, Y.-J., Shuo, S., Tan, T. H. P., Ooi, E.-E., Lim, S. G., et al. (2004). Characterization of a unique group-specific protein (U122) of the severe acute respiratory syndrome coronavirus. *J. Virol.* 78, 7311–7318. doi: 10.1128/JVI.78.14.7311-7318.2004

Finkel, Y., Mizrahi, O., Nachshon, A., Weingarten-Gabbay, S., Morgenstern, D., Yahalom-Ronen, Y., et al. (2020). The coding capacity of SARS-CoV-2. *Nature* 589, 125–130. doi: 10.1038/s41586-020-2739-1

Firth, A. E. (2020). A putative new SARS-CoV protein, 3c, encoded in an ORF overlapping ORF3a. *J. Gen. Virol.* 101, 1085–1089. doi: 10.1099/jgv.0.001469

Fischer, F., Peng, D., Hingley, S. T., Weiss, S. R., and Masters, P. S. (1997). 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.* 71, 996–1003. doi: 10.1128/JVI.71.2.996-1003.1997

Flower, T. G., Buffalo, C. Z., Hooy, R. M., Allaire, M., Ren, X., and Hurley, J. H. (2020). Structure of SARS-CoV-2 ORF8, a rapidly evolving coronavirus protein implicated in immune evasion. *bioRxiv* [Preprint]. doi: 10.1101/2020.08.27.270637

Freundt, E. C., Yu, L., Goldsmith, C. S., Welsh, S., Cheng, A., Yount, B., et al. (2010). The open reading frame 3a protein of severe acute respiratory syndrome-associated coronavirus promotes membrane rearrangement and cell death. *J. Virol.* 84, 1097–1109. doi: 10.1128/JVI.01662-09

Frieman, M., Yount, B., Heise, M., Kopecky-Bromberg, S. A., Palese, P., and Baric, R. S. (2007). Severe acute respiratory syndrome coronavirus ORF6 antagonizes STAT1 function by sequestering nuclear import factors on the rough endoplasmic reticulum/Golgi membrane. *J. Virol.* 81, 9812–9824. doi: 10.1128/JVI.01012-07

Fu, Y.-Z., Wang, S.-Y., Zheng, Z.-Q., Huang, Y., Li, W.-W., Xu, Z.-S., et al. (2020). SARS-CoV-2 membrane glycoprotein M antagonizes the MAVS-mediated innate antiviral response. *Cell. Mol. Immunol.* 18, 613–620. doi: 10.1038/s41423-020-00571-x

Fuchs, J., Hölzer, M., Schilling, M., Patzina, C., Schoen, A., Hoenen, T., et al. (2017). Evolution and antiviral specificities of interferon-induced Mx Proteins of Bats against Ebola, Influenza, and Other RNA Viruses. *J. Virol.* 91:e00361–17. doi: 10.1128/JVI.00361-17

Fung, S.-Y., Yuen, K.-S., Ye, Z.-W., Chan, C.-P., and Jin, D.-Y. (2020). A tug-of-war between severe acute respiratory syndrome coronavirus 2 and host antiviral defence: lessons from other pathogenic viruses. *Emerg. Microbes. Infect.* 9, 558–570. doi: 10.1080/22221751.2020.1736644

Gao, X., Zhu, K., Qin, B., Olieric, V., Wang, M., and Cui, S. (2021). Crystal structure of SARS-CoV-2 Orf9b in complex with human TOM70 suggests unusual virus-host interactions. *Nat. Commun.* 12:2843. doi: 10.1038/s41467-021-23118-8

Ge, H., Wang, X., Yuan, X., Xiao, G., Wang, C., Deng, T., et al. (2020). The epidemiology and clinical information about COVID-19. *Eur. J. Clin. Microbiol. Infect. Dis.* 39, 1011–1019. doi: 10.1007/s10096-020-03874-z

Ge, X.-Y., Li, J.-L., Yang, X.-L., Chmura, A. A., Zhu, G., Epstein, J. H., et al. (2013). Isolation and characterization of a bat SARS-like coronavirus that uses the ACE2 receptor. *Nature* 503, 535–538. doi: 10.1038/nature12711

Geng, H., Liu, Y.-M., Chan, W.-S., Lo, A. W.-I., Au, D. M.-Y., Waye, M. M.-Y., et al. (2005). The putative protein 6 of the severe acute respiratory syndrome-associated coronavirus: expression and functional characterization. *FEBS Lett.* 579, 6763–6768. doi: 10.1016/j.febslet.2005.11.007

Gledhill, A. W., and Andrewes, C. H. (1951). A hepatitis virus of mice. *Br. J. Exp. Pathol.* 32, 559–568.

Glennon, N. B., Jabado, O., Lo, M. K., and Shaw, M. L. (2015). Transcriptome Profiling of the Virus-Induced Innate Immune Response in Pteropus vampyrus and Its Attenuation by Nipah Virus Interferon Antagonist Functions. *J. Virol.* 89, 7550–7566. doi: 10.1128/JVI.00302-15

Glowacka, I., Bertram, S., Muller, M. A., Allen, P., Soilleux, E., Pfefferle, S., et al. (2011). Evidence that TMPRSS2 activates the severe acute respiratory syndrome coronavirus spike protein for membrane fusion and reduces viral control by the humoral immune response. *J. Virol.* 85, 4122–4134. doi: 10.1128/JVI.02232-10

Goldstein, S. A., and Weiss, S. R. (2017). Origins and pathogenesis of Middle East respiratory syndrome-associated coronavirus: recent advances. *F1000Res* 6:1628. doi: 10.12688/f1000research.11827.1

Gordon, D. E., Jang, G. M., Bouhaddou, M., Xu, J., Obernier, K., White, K. M., et al. (2020). A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature* 583, 459–468. doi: 10.1038/s41586-020-2286-9

Graham, R. L., and Baric, R. S. (2010). Recombination, reservoirs, and the modular spike: mechanisms of coronavirus cross-species transmission. *J. Virol.* 84, 3134–3146. doi: 10.1128/JVI.01394-09

Gu, J., Gong, E., Zhang, B., Zheng, J., Gao, Z., Zhong, Y., et al. (2005). Multiple organ infection and the pathogenesis of SARS. *J. Exp. Med.* 202, 415–424. doi: 10.1084/jem.20050828

Guan, Y. (2003). Isolation and characterization of viruses related to the SARS Coronavirus from Animals in Southern China. *Science* 302, 276–278. doi: 10.1126/science.1087139

Guo, J.-P., Petric, M., Campbell, W., and McGeer, P. L. (2004). SARS corona virus peptides recognized by antibodies in the sera of convalescent cases. *Virology* 324, 251–256. doi: 10.1016/j.virol.2004.04.017

Guy, J. S. (1998). Virus infections of the gastrointestinal tract of poultry. *Poult. Sci.* 77, 1166–1175. doi: 10.1093/ps/77.8.1166

Hachim, A., Kavian, N., Cohen, C. A., Chin, A. W. H., Chu, D. K. W., Mok, C. K. P., et al. (2020). ORF8 and ORF3b antibodies are accurate serological

markers of early and late SARS-CoV-2 infection. *Nat. Immunol.* 21, 1293–1301. doi: 10.1038/s41590-020-0773-7

Hamming, I., Timens, W., Bulthuis, M. L. C., Lely, A. T., Navis, G. J., and van Goor, H. (2004). Tissue distribution of ACE2 protein, the functional receptor for SARS coronavirus. A first step in understanding SARS pathogenesis. *J. Pathol.* 203, 631–637. doi: 10.1002/path.1570

Hamre, D., and Procknow, J. J. (1966). A new virus isolated from the human respiratory tract. *Exp. Biol. Med.* 121, 190–193. doi: 10.3181/00379727-121-30734

Han, L., Zhuang, M., Deng, J., Zheng, Y., Zhang, J., Nan, M., et al. (2021). SARS−CoV−2 ORF9b antagonizes type I and III interferons by targeting multiple components of the RIG−I/MDA−5−MAVS, TLR3−TRIF, and cGAS−STING signaling pathways. *J. Med. Virol.* doi: 10.1002/jmv.27050 [Epub ahead of print].

Hawkins, J. A., Kaczmarek, M. E., Müller, M. A., Drosten, C., Press, W. H., and Sawyer, S. L. (2019). A metaanalysis of bat phylogenetics and positive selection based on genomes and transcriptomes from 18 species. *Proc. Natl. Acad. Sci. U.S.A.* 116, 11351–11360. doi: 10.1073/pnas.1814995116

Helmy, Y. A., Fawzy, M., Elaswad, A., Sobieh, A., Kenney, S. P., and Shehata, A. A. (2020). The COVID-19 Pandemic: a comprehensive review of taxonomy, genetics, epidemiology, diagnosis, treatment, and control. *J. Clin. Med.* 9:1225. doi: 10.3390/jcm9041225

Hoffmann, M., Kleine-Weber, H., Schroeder, S., Krüger, N., Herrler, T., Erichsen, S., et al. (2020). SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is blocked by a clinically proven protease inhibitor. *Cell* 181, 271–280.e8. doi: 10.1016/j.cell.2020.02.052

Holland, L. A., Kaelin, E. A., Maqsood, R., Estifanos, B., Wu, L. I., Varsani, A., et al. (2020). An 81-Nucleotide Deletion in SARS-CoV-2 ORF7a identified from sentinel surveillance in arizona (January to March 2020). *J. Virol.* 94, e711–e720. doi: 10.1128/JVI.00711-20

Hon, C.-C., Lam, T.-Y., Shi, Z.-L., Drummond, A. J., Yip, C.-W., Zeng, F., et al. (2008). Evidence of the recombinant origin of a bat severe acute respiratory syndrome (SARS)-like coronavirus and its implications on the direct ancestor of SARS coronavirus. *J. Virol.* 82, 1819–1826. doi: 10.1128/JVI.01926-07

Hu, B., Ge, X., Wang, L.-F., and Shi, Z. (2015). Bat origin of human coronaviruses. *Virol. J.* 12:221. doi: 10.1186/s12985-015-0422-1

Hu, B., Zeng, L.-P., Yang, X.-L., Ge, X.-Y., Zhang, W., Li, B., et al. (2017). Discovery of a rich gene pool of bat SARS-related coronaviruses provides new insights into the origin of SARS coronavirus. *PLoS Pathog.* 13:e1006698. doi: 10.1371/journal.ppat.1006698

Hu, D., Zhu, C., Ai, L., He, T., Wang, Y., Ye, F., et al. (2018). Genomic characterization and infectivity of a novel SARS-like coronavirus in Chinese bats. *Emerg. Microb. Infect.* 7, 1–10. doi: 10.1038/s41426-018-0155-5

Huang, C., Ito, N., Tseng, C.-T. K., and Makino, S. (2006a). Severe respiratory syndrome coronavirus 7a accessory protein is a viral structural protein. *J. Virol.* 80, 7287–7294. doi: 10.1128/JVI.00414-06

Huang, C., Narayanan, K., Ito, N., Peters, C. J., and Makino, S. (2006b). Severe acute respiratory syndrome coronavirus 3a protein is released in membranous structures from 3a protein-expressing cells and infected cells. *J. Virol.* 80, 210–217. doi: 10.1128/JVI.80.1.210-217.2006

Huang, C., Peters, C. J., and Makino, S. (2007). Severe acute respiratory syndrome coronavirus accessory protein 6 is a virion-associated protein and is released from 6 protein-expressing cells. *J. Virol.* 81, 5423–5426. doi: 10.1128/JVI.02307-06

Huang, C., Wang, Y., Li, X., Ren, L., Zhao, J., Hu, Y., et al. (2020). Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* 395, 497–506. doi: 10.1016/S0140-6736(20)30183-5

Hulswit, R. J. G., Lang, Y., Bakkers, M. J. G., Li, W., Li, Z., Schouten, A., et al. (2019). Human coronaviruses OC43 and HKU1 bind to 9- O -acetylated sialic acids via a conserved receptor-binding site in spike protein domain A. *Proc. Natl. Acad. Sci. U.S.A.* 116, 2681–2690. doi: 10.1073/pnas.1809667116

Huynh, J., Li, S., Yount, B., Smith, A., Sturges, L., Olsen, J. C., et al. (2012). Evidence supporting a zoonotic origin of human coronavirus strain NL63. *J. Virol.* 86, 12816–12825. doi: 10.1128/JVI.00906-12

Issa, E., Merhi, G., Panossian, B., Salloum, T., and Tokajian, S. (2020). SARS-CoV-2 and ORF3a: nonsynonymous mutations, functional domains, and viral pathogenesis. *mSystems* 5, e00266–20. doi: 10.1128/mSystems.00266-20

Ithete, N. L., Stoffberg, S., Corman, V. M., Cottontail, V. M., Richards, L. R., Schoeman, M. C., et al. (2013). Close Relative of Human middle east respiratory syndrome coronavirus in bat, South Africa. *Emerg. Infect. Dis.* 19, 1697–1699. doi: 10.3201/eid1910.130946

Ito, N., Mossel, E. C., Narayanan, K., Popov, V. L., Huang, C., Inoue, T., et al. (2005). Severe acute respiratory syndrome coronavirus 3a protein is a viral structural protein. *J. Virol.* 79, 3182–3186. doi: 10.1128/JVI.79.5.3182-3186.2005

Jebb, D., Huang, Z., Pippel, M., Hughes, G. M., Lavrichenko, K., Devanna, P., et al. (2020). Six reference-quality genomes reveal evolution of bat adaptations. *Nature* 583, 578–584. doi: 10.1038/s41586-020-2486-3

Jia, H. P., Look, D. C., Shi, L., Hickey, M., Pewe, L., Netland, J., et al. (2005). ACE2 receptor expression and severe acute respiratory syndrome coronavirus infection depend on differentiation of human airway epithelia. *J. Virol.* 79, 14614–14621. doi: 10.1128/JVI.79.23.14614-14621.2005

Jiang, H., Li, Y., Zhang, H., Wang, W., Yang, X., Qi, H., et al. (2020a). SARS-CoV-2 proteome microarray for global profiling of COVID-19 specific IgG and IgM responses. *Nat. Commun.* 11:3581. doi: 10.1038/s41467-020-17488-8

Jiang, H.-W., Zhang, H.-N., Meng, Q.-F., Xie, J., Li, Y., Chen, H., et al. (2020b). SARS-CoV-2 Orf9b suppresses type I interferon responses by targeting TOM70. *Cell. Mol. Immunol.* 17, 998–1000. doi: 10.1038/s41423-020-0514-8

Jungreis, I., Nelson, C. W., Ardern, Z., Finkel, Y., Krogan, N. J., Sato, K., et al. (2021). Conflicting and ambiguous names of overlapping ORFs in the SARS-CoV-2 genome: a homology-based resolution. *Virology* 558, 145–151. doi: 10.1016/j.virol.2021.02.013

Jungreis, I., Sealfon, R., and Kellis, M. (2020). SARS-CoV-2 gene content and COVID-19 mutation impact by comparing 44 Sarbecovirus genomes. *Res. Sq.* [Preprint]. doi: 10.21203/rs.3.rs-80345/v1

Kandeel, M., Ibrahim, A., Fayez, M., and Al−Nazawi, M. (2020). From SARS and MERS CoVs to SARS−CoV−2: Moving toward more biased codon usage in viral structural and nonstructural genes. *J. Med. Virol.* 92, 660–666. doi: 10.1002/jmv.25754

Kang, S., Yang, M., Hong, Z., Zhang, L., Huang, Z., Chen, X., et al. (2020). Crystal structure of SARS-CoV-2 nucleocapsid protein RNA binding domain reveals potential unique drug targeting sites. *Acta Pharm. Sin B* 10, 1228–1238. doi: 10.1016/j.apsb.2020.04.009

Kanzawa, N., Nishigaki, K., Hayashi, T., Ishii, Y., Furukawa, S., Niiro, A., et al. (2006). Augmentation of chemokine production by severe acute respiratory syndrome coronavirus 3a/X1 and 7a/X4 proteins through NF-kappaB activation. *FEBS Lett.* 580, 6807–6812. doi: 10.1016/j.febslet.2006.11.046

Keck, J. G., Matsushima, G. K., Makino, S., Fleming, J. O., Vannier, D. M., Stohlman, S. A., et al. (1988). In vivo RNA-RNA recombination of coronavirus in mouse brain. *J. Virol.* 62, 1810–1813. doi: 10.1128/JVI.62.5.1810-1813.1988

Keng, C.-T., Choi, Y.-W., Welkers, M. R. A., Chan, D. Z. L., Shen, S., Gee Lim, S., et al. (2006). The human severe acute respiratory syndrome coronavirus (SARS-CoV) 8b protein is distinct from its counterpart in animal SARS-CoV and down-regulates the expression of the envelope protein in infected cells. *Virology* 354, 132–142. doi: 10.1016/j.virol.2006.06.026

Kern, D. M., Sorum, B., Hoel, C. M., Sridharan, S., Remis, J. P., Toso, D. B., et al. (2020). Cryo-EM structure of the SARS-CoV-2 3a ion channel in lipid nanodiscs. *bioRxiv* [Preprint]. doi: 10.1101/2020.06.17.156554

Khan, S., Fielding, B. C., Tan, T. H. P., Chou, C.-F., Shen, S., Lim, S. G., et al. (2006). Over-expression of severe acute respiratory syndrome coronavirus 3b protein induces both apoptosis and necrosis in Vero E6 cells. *Virus Res.* 122, 20–27. doi: 10.1016/j.virusres.2006.06.005

Khrustalev, V. V., Giri, R., Khrustaleva, T. A., Kapuganti, S. K., Stojarov, A. N., and Poboinev, V. V. (2020). Translation-Associated mutational U-pressure in the first ORF of SARS-CoV-2 and other coronaviruses. *Front. Microbiol.* 11:559165. doi: 10.3389/fmicb.2020.559165

Kim, D., Lee, J.-Y., Yang, J.-S., Kim, J. W., Kim, V. N., and Chang, H. (2020). The Architecture of SARS-CoV-2 Transcriptome. *Cell* 181, 914–921.e10. doi: 10.1016/j.cell.2020.04.011

Kimura, I., Konno, Y., Uriu, K., Hopfensperger, K., Sauter, D., Nakagawa, S., et al. (2021). Sarbecovirus ORF6 proteins hamper induction of interferon signaling. *Cell Rep.* 34:108916. doi: 10.1016/j.celrep.2021.108916

Kiyong'a, A. N., Cook, E. A. J., Okba, N. M. A., Kivali, V., Reusken, C., Haagmans, B. L., et al. (2020). Middle east respiratory syndrome coronavirus (MERS-CoV) Seropositive Camel Handlers in Kenya. *Viruses* 12:396. doi: 10.3390/v12040396

Klein, S., Cortese, M., Winter, S. L., Wachsmuth-Melm, M., Neufeldt, C. J., Cerikan, B., et al. (2020). SARS-CoV-2 structure and replication characterized by in situ cryo-electron tomography. *Nat. Commun.* 11:5885. doi: 10.1038/s41467-020-19619-7

Konno, Y., Kimura, I., Uriu, K., Fukushi, M., Irie, T., Koyanagi, Y., et al. (2020). SARS-CoV-2 ORF3b Is a potent interferon antagonist whose activity is increased by a naturally occurring elongation variant. *Cell Rep.* 32:108185. doi: 10.1016/j.celrep.2020.108185

Kopecky-Bromberg, S. A., Martínez-Sobrido, L., Frieman, M., Baric, R. A., and Palese, P. (2007). Severe acute respiratory syndrome coronavirus open reading frame (ORF) 3b, ORF 6, and nucleocapsid proteins function as interferon antagonists. *J. Virol.* 81, 548–557. doi: 10.1128/JVI.01782-06

Kopecky-Bromberg, S. A., Martinez-Sobrido, L., and Palese, P. (2006). 7a protein of severe acute respiratory syndrome coronavirus inhibits cellular protein synthesis and activates p38 mitogen-activated protein kinase. *J. Virol.* 80, 785–793. doi: 10.1128/JVI.80.2.785-793.2006

Kozak, M. (1986). Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell* 44, 283–292. doi: 10.1016/0092-8674(86)90762-2

Kozak, M. (1991). A short leader sequence impairs the fidelity of initiation by eukaryotic ribosomes. *Gene Expr.* 1, 111–115.

Kozak, M. (2002). Pushing the limits of the scanning mechanism for initiation of translation. *Gene* 299, 1–34. doi: 10.1016/s0378-1119(02)01056-9

Krempl, C., Schultze, B., and Herrler, G. (1995). "Analysis of cellular receptors for human coronavirus OC43," in *Corona- and Related Viruses Advances in Experimental Medicine and Biology*, eds P. J. Talbot and G. A. Levy (Boston, MA: Springer US), 371–374. doi: 10.1007/978-1-4615-1899-0_60

Kuba, K., Imai, Y., Rao, S., Gao, H., Guo, F., Guan, B., et al. (2005). A crucial role of angiotensin converting enzyme 2 (ACE2) in SARS coronavirus–induced lung injury. *Nat. Med.* 11, 875–879. doi: 10.1038/nm1267

Kuehn, B. M. (2013). More evidence emerges that bats may have spread SARS. *JAMA* 310:2138. doi: 10.1001/jama.2013.283495

Kumar, P., Gunalan, V., Liu, B., Chow, V. T. K., Druce, J., Birch, C., et al. (2007). The nonstructural protein 8 (nsp8) of the SARS coronavirus interacts with its ORF6 accessory protein. *Virology* 366, 293–303. doi: 10.1016/j.virol.2007.04.029

Kuo, L., and Masters, P. S. (2002). Genetic evidence for a structural interaction between the carboxy termini of the membrane and nucleocapsid proteins of mouse hepatitis virus. *J. Virol.* 76, 4987–4999. doi: 10.1128/jvi.76.10.4987-4999.2002

Kupferschmidt, K. (2020). Genome analyses help track coronavirus' moves. *Science* 367, 1176–1177. doi: 10.1126/science.367.6483.1176

Kustin, T., and Stern, A. (2021). Biased mutation and selection in RNA Viruses. *Mol. Biol. Evol.* 38, 575–588. doi: 10.1093/molbev/msaa247

Lai, M. M., and Cavanagh, D. (1997). The molecular biology of coronaviruses. *Adv. Virus Res.* 48, 1–100. doi: 10.1007/978-1-4615-9373-7_1

Lam, T. T.-Y., Jia, N., Zhang, Y.-W., Shum, M. H.-H., Jiang, J.-F., Zhu, H.-C., et al. (2020a). Identifying SARS-CoV-2-related coronaviruses in Malayan pangolins. *Nature* 583, 282–285. doi: 10.1038/s41586-020-2169-0

Lam, T. T.-Y., Shum, M. H.-H., Zhu, H.-C., Tong, Y.-G., Ni, X.-B., Liao, Y.-S., et al. (2020b). Identification of 2019-nCoV related coronaviruses in Malayan pangolins in southern China. *bioRxiv* [Preprint]. doi: 10.1101/2020.02.13.945485

Lau, S. K. P., Feng, Y., Chen, H., Luk, H. K. H., Yang, W.-H., Li, K. S. M., et al. (2015a). Severe acute respiratory syndrome (SARS) Coronavirus ORF8 protein is acquired from SARS-related coronavirus from greater horseshoe bats through recombination. *J. Virol.* 89, 10532–10547. doi: 10.1128/JVI.01048-15

Lau, S. K. P., Lee, P., Tsang, A. K. L., Yip, C. C. Y., Tse, H., Lee, R. A., et al. (2011). 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.* 85, 11325–11337. doi: 10.1128/JVI.05512-11

Lau, S. K. P., Woo, P. C. Y., Li, K. S. M., Huang, Y., Tsoi, H.-W., Wong, B. H. L., et al. (2005). Severe acute respiratory syndrome coronavirus-like virus in Chinese horseshoe bats. *Proc. Natl. Acad. Sci. U.S.A.* 102, 14040–14045. doi: 10.1073/pnas.0506735102

Lau, S. K. P., Woo, P. C. Y., Li, K. S. M., Tsang, A. K. L., Fan, R. Y. Y., Luk, H. K. H., et al. (2015b). Discovery of a novel coronavirus, china rattus coronavirus HKU24, from norway rats supports the murine origin of Betacoronavirus 1 and Has Implications for the Ancestor of Betacoronavirus Lineage A. *J. Virol.* 89, 3076–3092. doi: 10.1128/JVI.02420-14

Lau, S. K. P., Woo, P. C. Y., Yip, C. C. Y., Tse, H., Tsoi, H. -w, Cheng, V. C. C., et al. (2006). Coronavirus HKU1 and other coronavirus infections in Hong Kong. *J. Clin. Microbiol.* 44, 2063–2071. doi: 10.1128/JCM.02614-05

Law, H. K. W., Cheung, C. Y., Ng, H. Y., Sia, S. F., Chan, Y. O., Luk, W., et al. (2005a). Chemokine up-regulation in SARS-coronavirus–infected, monocyte-derived human dendritic cells. *Blood* 106, 2366–2374. doi: 10.1182/blood-2004-10-4166

Law, P. T. W., Wong, C.-H., Au, T. C. C., Chuck, C.-P., Kong, S.-K., Chan, P. K. S., et al. (2005b). The 3a protein of severe acute respiratory syndrome-associated coronavirus induces apoptosis in Vero E6 cells. *J. Gen. Virol.* 86, 1921–1930. doi: 10.1099/vir.0.80813-0

Lee, A. K., Kulcsar, K. A., Elliott, O., Khiabanian, H., Nagle, E. R., Jones, M. E. B., et al. (2015). De novo transcriptome reconstruction and annotation of the Egyptian rousette bat. *BMC Genomics* 16:1033. doi: 10.1186/s12864-015-2124-x

Lelli, D., Papetti, A., Sabelli, C., Rosti, E., Moreno, A., and Boniotti, M. (2013). Detection of coronaviruses in bats of various species in Italy. *Viruses* 5, 2679–2689. doi: 10.3390/v5112679

Letko, M., Seifert, S. N., Olival, K. J., Plowright, R. K., and Munster, V. J. (2020). Bat-borne virus diversity, spillover and emergence. *Nat. Rev. Microbiol.* 18, 461–471. doi: 10.1038/s41579-020-0394-z

Li, J.-Y., Liao, C.-H., Wang, Q., Tan, Y.-J., Luo, R., Qiu, Y., et al. (2020). The ORF6, ORF8 and nucleocapsid proteins of SARS-CoV-2 inhibit type I interferon signaling pathway. *Virus Res.* 286:198074. doi: 10.1016/j.virusres.2020.198074

Li, W., Moore, M. J., Vasilieva, N., Sui, J., Wong, S. K., Berne, M. A., et al. (2003). Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. *Nature* 426, 450–454. doi: 10.1038/nature02145

Li, W., Shi, Z., Yu, M., Ren, W., Smith, C., Epstein, J. H., et al. (2005). Bats are natural reservoirs of SARS-like coronaviruses. *Science* 310, 676–679. doi: 10.1126/science.1118391

Lin, X., Fu, B., Yin, S., Li, Z., Liu, H., Zhang, H., et al. (2021). ORF8 contributes to cytokine storm during SARS-CoV-2 infection by activating IL-17 pathway. *iScience* 24:102293. doi: 10.1016/j.isci.2021.102293

Liu, P., Chen, W., and Chen, J.-P. (2019). Viral metagenomics revealed sendai virus and coronavirus infection of malayan pangolins (Manis javanica). *Viruses* 11:979. doi: 10.3390/v11110979

Lu, S., Wang, Y., Chen, Y., Wu, B., Qin, K., Zhao, J., et al. (2017). Discovery of a novel canine respiratory coronavirus support genetic recombination among betacoronavirus1. *Virus Res.* 237, 7–13. doi: 10.1016/j.virusres.2017.05.006

Luis, A. D., Hayman, D. T. S., O'Shea, T. J., Cryan, P. M., Gilbert, A. T., Pulliam, J. R. C., et al. (2013). A comparison of bats and rodents as reservoirs of zoonotic viruses: are bats special? *Proc. Biol. Sci.* 280:20122753. doi: 10.1098/rspb.2012.2753

Luytjes, W., Bredenbeek, P. J., Noten, A. F., Horzinek, M. C., and Spaan, W. J. (1988). Sequence of mouse hepatitis virus A59 mRNA 2: indications for RNA recombination between coronaviruses and influenza C virus. *Virology* 166, 415–422. doi: 10.1016/0042-6822(88)90512-0

Mahtarin, R., Islam, S., Islam, M. J., Ullah, M. O., Ali, M. A., and Halim, M. A. (2020). Structure and dynamics of membrane protein in SARS-CoV-2. *J. Biomol. Struct. Dyn* doi: 10.1080/07391102.2020.1861983 [Epub ahead of print].

Makino, S., Keck, J. G., Stohlman, S. A., and Lai, M. M. (1986). High-frequency RNA recombination of murine coronaviruses. *J. Virol.* 57, 729–737. doi: 10.1128/JVI.57.3.729-737.1986

Mandala, V. S., McKay, M. J., Shcherbakov, A. A., Dregni, A. J., Kolocouris, A., and Hong, M. (2020). Structure and Drug Binding of the SARS-CoV-2 envelope protein in phospholipid bilayers. *Res. Sq.* [Preprint]. doi: 10.21203/rs.3.rs-77124/v1

Marra, M. A., Jones, S. J. M., Astell, C. R., Holt, R. A., Brooks-Wilson, A., Butterfield, Y. S. N., et al. (2003). The Genome sequence of the SARS-associated coronavirus. *Science* 300, 1399–1404. doi: 10.1126/science.1085953

Martínez Gómez, J. M., Periasamy, P., Dutertre, C.-A., Irving, A. T., Ng, J. H. J., Crameri, G., et al. (2016). Phenotypic and functional characterization of the major lymphocyte populations in the fruit-eating bat Pteropus alecto. *Sci. Rep.* 6:37796. doi: 10.1038/srep37796

Masters, P. S. (2006). The molecular biology of coronaviruses. *Adv. Virus Res.* 66, 193–292. doi: 10.1016/S0065-3527(06)66005-3

Matsuda, D., and Dreher, T. W. (2006). Close spacing of AUG initiation codons confers dicistronic character on a eukaryotic mRNA. *RNA* 12, 1338–1349. doi: 10.1261/rna.67906

Matsuyama, S., Nagata, N., Shirato, K., Kawase, M., Takeda, M., and Taguchi, F. (2010). Efficient activation of the severe acute respiratory syndrome coronavirus spike protein by the transmembrane protease TMPRSS2. *J. Virol.* 84, 12658–12664. doi: 10.1128/JVI.01542-10

McIntosh, K., Becker, W. B., and Chanock, R. M. (1967a). Growth in suckling-mouse brain of "IBV-like" viruses from patients with upper respiratory tract disease. *Proc. Natl. Acad. Sci. U.S.A.* 58, 2268–2273. doi: 10.1073/pnas.58.6.2268

McIntosh, K., Dees, J. H., Becker, W. B., Kapikian, A. Z., and Chanock, R. M. (1967b). Recovery in tracheal organ cultures of novel viruses from patients with respiratory disease. *Proc. Natl. Acad. Sci. U.S.A.* 57, 933–940. doi: 10.1073/pnas.57.4.933

Meier, C., Aricescu, A. R., Assenberg, R., Aplin, R. T., Gilbert, R. J. C., Grimes, J. M., et al. (2006). The crystal structure of ORF-9b, a lipid binding protein from the SARS Coronavirus. *Structure* 14, 1157–1165. doi: 10.1016/j.str.2006.05.012

Menachery, V. D., Mitchell, H. D., Cockrell, A. S., Gralinski, L. E., Yount, B. L., Graham, R. L., et al. (2017). MERS-CoV Accessory ORFs Play Key Role for Infection and Pathogenesis. *mBio* 8, e00665–17. doi: 10.1128/mBio.00665-17

Michel, C. J., Mayer, C., Poch, O., and Thompson, J. D. (2020). Characterization of accessory genes in coronavirus genomes. *Virol. J.* 17:131. doi: 10.1186/s12985-020-01402-1

Middleton, D. J., Morrissy, C. J., van der Heide, B. M., Russell, G. M., Braun, M. A., Westbury, H. A., et al. (2007). Experimental Nipah virus infection in pteropid bats (*Pteropus poliocephalus*). *J. Comp. Pathol.* 136, 266–272. doi: 10.1016/j.jcpa.2007.03.002

Mihindukulasuriya, K. A., Wu, G., St Leger, J., Nordhausen, R. W., and Wang, D. (2008). Identification of a novel coronavirus from a beluga whale by using a panviral microarray. *J. Virol.* 82, 5084–5088. doi: 10.1128/JVI.02722-07

Miller, M. R., McMinn, R. J., Misra, V., Schountz, T., Müller, M. A., Kurth, A., et al. (2016). Broad and temperature independent replication potential of filoviruses on cells derived from old and new world bat species. *J. Infect. Dis.* 214, S297–S302. doi: 10.1093/infdis/jiw199

Moore, J. B., and June, C. H. (2020). Cytokine release syndrome in severe COVID-19. *Science* 368, 473–474. doi: 10.1126/science.abb8925

Moreno, A., Lelli, D., de Sabato, L., Zaccaria, G., Boni, A., Sozzi, E., et al. (2017). Detection and full genome characterization of two beta CoV viruses related to Middle East respiratory syndrome from bats in Italy. *Virol. J.* 14:239. doi: 10.1186/s12985-017-0907-1

Müller, M. A., Corman, V. M., Jores, J., Meyer, B., Younan, M., Liljander, A., et al. (2014). MERS coronavirus neutralizing antibodies in camels, Eastern Africa, 1983–1997. *Emerg. Infect. Dis.* 20:1026. doi: 10.3201/eid2012.141026

Munir, K., Ashraf, S., Munir, I., Khalid, H., Muneer, M. A., Mukhtar, N., et al. (2020). Zoonotic and reverse zoonotic events of SARS-CoV-2 and their impact on global health. *Emerg. Microb. Infect.* 9, 2222–2235. doi: 10.1080/22221751.2020.1827984

Munster, V. J., Adney, D. R., van Doremalen, N., Brown, V. R., Miazgowicz, K. L., Milne-Price, S., et al. (2016). Replication and shedding of MERS-CoV in Jamaican fruit bats (*Artibeus jamaicensis*). *Sci. Rep.* 6:21878. doi: 10.1038/srep21878

Murgolo, N., Therien, A. G., Howell, B., Klein, D., Koeplinger, K., Lieberman, L. A., et al. (2021). SARS-CoV-2 tropism, entry, replication, and propagation: considerations for drug discovery and development. *PLoS Pathog.* 17:e1009225. doi: 10.1371/journal.ppat.1009225

Muth, D., Corman, V. M., Roth, H., Binger, T., Dijkman, R., Gottula, L. T., et al. (2018). Attenuation of replication by a 29 nucleotide deletion in SARS-coronavirus acquired during the early stages of human-to-human transmission. *Sci. Rep.* 8:15177. doi: 10.1038/s41598-018-33487-8

Narayanan, K., Huang, C., and Makino, S. (2008). SARS coronavirus accessory proteins. *Virus Res.* 133, 113–121. doi: 10.1016/j.virusres.2007.10.009

Nelson, C. A., Pekosz, A., Lee, C. A., Diamond, M. S., and Fremont, D. H. (2005). Structure and intracellular targeting of the SARS-coronavirus Orf7a accessory protein. *Structure* 13, 75–85. doi: 10.1016/j.str.2004.10.010

Nelson, C. W., Ardern, Z., Goldberg, T. L., Meng, C., Kuo, C.-H., Ludwig, C., et al. (2020). Dynamically evolving novel overlapping gene as a factor in the SARS-CoV-2 pandemic. *eLife* 9:e59633. doi: 10.7554/eLife.59633

Neuman, B. W., Kiss, G., Kunding, A. H., Bhella, D., Baksh, M. F., Connelly, S., et al. (2011). A structural analysis of M protein in coronavirus assembly and morphology. *J. Struct. Biol.* 174, 11–22. doi: 10.1016/j.jsb.2010.11.021

Nomburg, J., Meyerson, M., and DeCaprio, J. A. (2020). Pervasive generation of non-canonical subgenomic RNAs by SARS-CoV-2. *Genome Med.* 12:108. doi: 10.1186/s13073-020-00802-w

Omatsu, T., Bak, E.-J., Ishii, Y., Kyuwa, S., Tohya, Y., Akashi, H., et al. (2008). Induction and sequencing of Rousette bat interferon α and β genes. *Vet. Immunol. Immunopathol.* 124, 169–176. doi: 10.1016/j.vetimm.2008.03.004

O'Shea, T. J., Cryan, P. M., Cunningham, A. A., Fooks, A. R., Hayman, D. T. S., Luis, A. D., et al. (2014). Bat flight and zoonotic viruses. *Emerg. Infect. Dis.* 20, 741–745. doi: 10.3201/eid2005.130539

Pancer, K., Milewska, A., Owczarek, K., Dabrowska, A., Kowalski, M., Łabaj, P. P., et al. (2020). The SARS-CoV-2 ORF10 is not essential *in vitro* or *in vivo* in humans. *PLoS Pathog.* 16:e1008959. doi: 10.1371/journal.ppat.1008959

Papenfuss, A. T., Baker, M. L., Feng, Z.-P., Tachedjian, M., Crameri, G., Cowled, C., et al. (2012). The immune gene repertoire of an important viral reservoir, the Australian black flying fox. *BMC Genomics* 13:261. doi: 10.1186/1471-2164-13-261

Pavesi, A. (2019). Asymmetric evolution in viral overlapping genes is a source of selective protein adaptation. *Virology* 532, 39–47. doi: 10.1016/j.virol.2019.03.017

Pavesi, A. (2020). New insights into the evolutionary features of viral overlapping genes by discriminant analysis. *Virology* 546, 51–66. doi: 10.1016/j.virol.2020.03.007

Pavesi, A., Vianelli, A., Chirico, N., Bao, Y., Blinkova, O., Belshaw, R., et al. (2018). Overlapping genes and the proteins they encode differ significantly in their sequence composition from non-overlapping genes. *PLoS One* 13:e0202513. doi: 10.1371/journal.pone.0202513

Pavlovich, S. S., Lovett, S. P., Koroleva, G., Guito, J. C., Arnold, C. E., Nagle, E. R., et al. (2018). The Egyptian rousette genome reveals unexpected features of bat antiviral immunity. *Cell* 173, 1098–1110.e18. doi: 10.1016/j.cell.2018.03.070

Pekosz, A., Schaecher, S. R., Diamond, M. S., Fremont, D. H., Sims, A. C., and Baric, R. S. (2006). Structure, expression, and intracellular localization of the SARS-CoV accessory proteins 7a and 7b. *Adv. Exp. Med. Biol.* 581, 115–120. doi: 10.1007/978-0-387-33012-9_20

Petersen, E., Koopmans, M., Go, U., Hamer, D. H., Petrosillo, N., Castelli, F., et al. (2020). Comparing SARS-CoV-2 with SARS-CoV and influenza pandemics. *Lancet Infect. Dis.* 20, e238–e244. doi: 10.1016/S1473-3099(20)30484-9

Pewe, L., Zhou, H., Netland, J., Tangudu, C., Olivares, H., Shi, L., et al. (2005). A severe acute respiratory syndrome-associated coronavirus-specific protein enhances virulence of an attenuated murine coronavirus. *J. Virol.* 79, 11335–11342. doi: 10.1128/JVI.79.17.11335-11342.2005

Pfefferle, S., Krähling, V., Ditt, V., Grywna, K., Mühlberger, E., and Drosten, C. (2009a). 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.* 6:131. doi: 10.1186/1743-422X-6-131

Pfefferle, S., Oppong, S., Drexler, J. F., Gloza-Rausch, F., Ipsen, A., Seebens, A., et al. (2009b). Distant relatives of severe acute respiratory syndrome coronavirus and close relatives of human coronavirus 229E in Bats, Ghana. *Emerg. Infect. Dis.* 15, 1377–1384. doi: 10.3201/eid1509.090224

Poudel, U., Subedi, D., Pantha, S., and Dhakal, S. (2020). Animal coronaviruses and coronavirus disease 2019: lesson for one health approach. *Open Vet. J.* 10, 239–251. doi: 10.4314/ovj.v10i3.1

Qiu, M., Shi, Y., Guo, Z., Chen, Z., He, R., Chen, R., et al. (2005). Antibody responses to individual proteins of SARS coronavirus and their neutralization activities. *Microbes Infect.* 7, 882–889. doi: 10.1016/j.micinf.2005.02.006

Quan, P.-L., Firth, C., Street, C., Henriquez, J. A., Petrosov, A., Tashmukhamedova, A., et al. (2010). Identification of a Severe Acute Respiratory Syndrome Coronavirus-Like Virus in a Leaf-Nosed Bat in Nigeria. *mBio* 1, e208–e210. doi: 10.1128/mBio.00208-10

Quéromès, G., Destras, G., Bal, A., Regue, H., Burfin, G., Brun, S., et al. (2021). Characterization of SARS-CoV-2 ORF6 deletion variants detected in a nosocomial cluster during routine genomic surveillance, Lyon, France. *Emerg. Microbes Infect.* 10, 167–177. doi: 10.1080/22221751.2021.1872351

Raj, V. S., Farag, E. A. B. A., Reusken, C. B. E. M., Lamers, M. M., Pas, S. D., Voermans, J., et al. (2014). Isolation of MERS coronavirus from a dromedary camel, Qatar, 2014. *Emerg. Infect. Dis.* 20, 1339–1342. doi: 10.3201/eid2008.140663

Rancurel, C., Khosravi, M., Dunker, A. K., Romero, P. R., and Karlin, D. (2009). Overlapping genes produce proteins with unusual sequence properties and offer insight into de novo protein creation. *J. Virol.* 83, 10719–10736. doi: 10.1128/JVI.00595-09

Ratcliff, J., and Simmonds, P. (2021). Potential APOBEC-mediated RNA editing of the genomes of SARS-CoV-2 and other coronaviruses and its impact on their longer term evolution. *Virology* 556, 62–72. doi: 10.1016/j.virol.2020.12.018

Ren, Y., Shu, T., Wu, D., Mu, J., Wang, C., Huang, M., et al. (2020). The ORF3a protein of SARS-CoV-2 induces apoptosis in cells. *Cell. Mol. Immunol.* 17, 881–883. doi: 10.1038/s41423-020-0485-9

Reusken, C. B., Haagmans, B. L., Müller, M. A., Gutierrez, C., Godeke, G.-J., Meyer, B., et al. (2013). Middle East respiratory syndrome coronavirus neutralising serum antibodies in dromedary camels: a comparative serological study. *Lancet Infect. Dis.* 13, 859–866. doi: 10.1016/S1473-3099(13)70164-6

Rihtaric, D., Hostnik, P., Steyer, A., Grom, J., and Toplak, I. (2010). Identification of SARS-like coronaviruses in horseshoe bats (*Rhinolophus hipposideros*) in Slovenia. *Arch. Virol.* 155, 507–514. doi: 10.1007/s00705-010-0612-5

Romano, M., Ruggiero, A., Squeglia, F., Maga, G., and Berisio, R. (2020). A Structural View of SARS-CoV-2 RNA Replication Machinery: RNA synthesis, proofreading and final capping. *Cells* 9:1267. doi: 10.3390/cells9051267

Rota, P. A., Oberste, M. S., Monroe, S. S., Nix, W. A., Campagnoli, R., Icenogle, J. P., et al. (2003). Characterization of a novel coronavirus associated with severe acute respiratory syndrome. *Science* 300, 1394–1399. doi: 10.1126/science.1085952

Sa Ribero, M., Jouvenet, N., Dreux, M., and Nisole, S. (2020). Interplay between SARS-CoV-2 and the type I interferon response. *PLoS Pathog.* 16:e1008737. doi: 10.1371/journal.ppat.1008737

Saif, L. J. (2004). Animal coronaviruses: what can they teach us about the severe acute respiratory syndrome? *Rev. Sci. Tech.* 23, 643–660. doi: 10.20506/rst.23.2.1513

Schad, J., and Voigt, C. C. (2016). Adaptive evolution of virus-sensing toll-like receptor 8 in bats. *Immunogenetics* 68, 783–795. doi: 10.1007/s00251-016-0940-z

Schaecher, S. R., Touchette, E., Schriewer, J., Buller, R. M., and Pekosz, A. (2007). Severe acute respiratory syndrome coronavirus gene 7 products contribute to virus-induced apoptosis. *J. Virol.* 81, 11054–11068. doi: 10.1128/JVI.01266-07

Schalk, A. E., and Hawn, M. C. (1931). An apparently new respiratory disease of chicks. *J. Am. Vet. Med. Assoc.* 78, 413–422.

Schoeman, D., and Fielding, B. C. (2019). Coronavirus envelope protein: current knowledge. *Virol. J.* 16:69. doi: 10.1186/s12985-019-1182-0

Schubert, K., Karousis, E. D., Jomaa, A., Scaiola, A., Echeverria, B., Gurzeler, L.-A., et al. (2020). SARS-CoV-2 Nsp1 binds the ribosomal mRNA channel to inhibit translation. *Nat. Struct. Mol. Biol.* 27, 959–966. doi: 10.1038/s41594-020-0511-8

Schuh, A. J., Amman, B. R., Sealy, T. K., Kainulainen, M. H., Chakrabarti, A. K., Guerrero, L. W., et al. (2019). Antibody-Mediated virus neutralization is not a universal mechanism of marburg, ebola, or sosuga virus clearance in egyptian rousette bats. *J. Infect. Dis.* 219, 1716–1721. doi: 10.1093/infdis/jiy733

Schuh, A. J., Amman, B. R., Sealy, T. K., Spengler, J. R., Nichol, S. T., and Towner, J. S. (2017). Egyptian rousette bats maintain long-term protective immunity against Marburg virus infection despite diminished antibody levels. *Sci. Rep.* 7:8763. doi: 10.1038/s41598-017-07824-2

Sedman, S. A., Gelembiuk, G. W., and Mertz, J. E. (1990). Translation initiation at a downstream AUG occurs with increased efficiency when the upstream AUG is located very close to the 5' cap. *J. Virol.* 64, 453–457. doi: 10.1128/JVI.64.1.453-457.1990

Seth, R. B., Sun, L., Ea, C.-K., and Chen, Z. J. (2005). Identification and characterization of MAVS, a mitochondrial antiviral signaling protein that activates NF-kappaB and IRF 3. *Cell* 122, 669–682. doi: 10.1016/j.cell.2005.08.012

Sheikh, A., Al-Taher, A., Al-Nazawi, M., Al-Mubarak, A. I., and Kandeel, M. (2020). Analysis of preferred codon usage in the coronavirus N genes and their implications for genome evolution and vaccine design. *J. Virol. Methods* 277:113806. doi: 10.1016/j.jviromet.2019.113806

Shen, S., Lin, P.-S., Chao, Y.-C., Zhang, A., Yang, X., Lim, S. G., et al. (2005). The severe acute respiratory syndrome coronavirus 3a is a novel structural protein. *Biochem. Biophys. Res. Commun.* 330, 286–292. doi: 10.1016/j.bbrc.2005.02.153

Shi, C.-S., Nabar, N. R., Huang, N.-N., and Kehrl, J. H. (2019). SARS-coronavirus open reading frame-8b triggers intracellular stress pathways and activates NLRP3 inflammasomes. *Cell Death Discov.* 5:101. doi: 10.1038/s41420-019-0181-7

Shi, C.-S., Qi, H.-Y., Boularan, C., Huang, N.-N., Abu-Asab, M., Shelhamer, J. H., et al. (2014). SARS-coronavirus open reading Frame-9b suppresses innate immunity by targeting mitochondria and the MAVS/TRAF3/TRAF6 Signalosome. *J.Immunol.* 193, 3080–3089. doi: 10.4049/jimmunol.1303196

Shulla, A., Heald-Sargent, T., Subramanya, G., Zhao, J., Perlman, S., and Gallagher, T. (2011). A transmembrane serine protease is linked to the severe acute respiratory syndrome coronavirus receptor and activates virus entry. *J. Virol.* 85, 873–882. doi: 10.1128/JVI.02062-10

Sims, A. C., Baric, R. S., Yount, B., Burkett, S. E., Collins, P. L., and Pickles, R. J. (2005). 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.* 79, 15511–15524. doi: 10.1128/JVI.79.24.15511-15524.2005

Siu, K.-L., Yuen, K.-S., Castaño-Rodriguez, C., Ye, Z.-W., Yeung, M.-L., Fung, S.-Y., et al. (2019). Severe acute respiratory syndrome coronavirus ORF3a protein activates the NLRP3 inflammasome by promoting TRAF3-dependent ubiquitination of ASC. *FASEB J.* 33, 8865–8877. doi: 10.1096/fj.201802418R

Sloots, T., Mcerlean, P., Speicher, D., Arden, K., Nissen, M., and Mackay, I. (2006). Evidence of human coronavirus HKU1 and human bocavirus in Australian children. *J. Clin. Virol.* 35, 99–102. doi: 10.1016/j.jcv.2005.09.008

Sola, I., Almazán, F., Zúñiga, S., and Enjuanes, L. (2015). Continuous and discontinuous RNA synthesis in coronaviruses. *Annu. Rev. Virol.* 2, 265–288. doi: 10.1146/annurev-virology-100114-055218

Storm, N., Jansen Van Vuren, P., Markotter, W., and Paweska, J. (2018). Antibody responses to marburg virus in Egyptian Rousette bats and their role in protection against infection. *Viruses* 10:73. doi: 10.3390/v10020073

Su, S., Wong, G., Shi, W., Liu, J., Lai, A. C. K., Zhou, J., et al. (2016). Epidemiology, genetic recombination, and pathogenesis of coronaviruses. *Trends Microbiol.* 24, 490–502. doi: 10.1016/j.tim.2016.03.003

Tagliamonte, M. S., Abid, N., Borocci, S., Sangiovanni, E., Ostrov, D. A., Kosakovsky Pond, S. L., et al. (2020). Multiple recombination events and strong purifying selection at the origin of SARS-CoV-2 spike glycoprotein increased correlated dynamic movements. *Int. J. Mol. Sci.* 22:80. doi: 10.3390/ijms22010080

Tan, Y.-J., Fielding, B. C., Goh, P.-Y., Shen, S., Tan, T. H. P., Lim, S. G., et al. (2004a). Overexpression of 7a, a protein specifically encoded by the severe acute respiratory syndrome coronavirus, induces apoptosis via a caspase-dependent pathway. *J. Virol.* 78, 14043–14047. doi: 10.1128/JVI.78.24.14043-14047.2004

Tan, Y.-J., Teng, E., Shen, S., Tan, T. H. P., Goh, P.-Y., Fielding, B. C., et al. (2004b). A novel severe acute respiratory syndrome coronavirus protein, U274, is transported to the cell surface and undergoes endocytosis. *J. Virol.* 78, 6723–6734. doi: 10.1128/JVI.78.13.6723-6734.2004

Tangudu, C., Olivares, H., Netland, J., Perlman, S., and Gallagher, T. (2007). Severe acute respiratory syndrome coronavirus protein 6 accelerates murine coronavirus infections. *J. Virol.* 81, 1220–1229. doi: 10.1128/JVI.01515-06

Tao, Y., Shi, M., Chommanard, C., Queen, K., Zhang, J., Markotter, W., et al. (2017). Surveillance of bat coronaviruses in kenya identifies relatives of human coronaviruses NL63 and 229E and Their Recombination History. *J. Virol.* 91:e01953–16. doi: 10.1128/JVI.01953-16

Tay, M. Z., Poh, C. M., Rénia, L., MacAry, P. A., and Ng, L. F. P. (2020). The trinity of COVID-19: immunity, inflammation and intervention. *Nat. Rev. Immunol.* 20, 363–374. doi: 10.1038/s41577-020-0311-8

Taylor, J. K., Coleman, C. M., Postel, S., Sisk, J. M., Bernbaum, J. G., Venkataraman, T., et al. (2015). Severe Acute Respiratory Syndrome Coronavirus ORF7a inhibits bone marrow stromal antigen 2 virion tethering through a novel

mechanism of glycosylation interference. *J. Virol.* 89, 11820–11833. doi: 10. 1128/JVI.02274-15

Thoms, M., Buschauer, R., Ameismeier, M., Koepke, L., Denk, T., Hirschenberger, M., et al. (2020). Structural basis for translational shutdown and immune evasion by the Nsp1 protein of SARS-CoV-2. *Science* 369, 1249–1255. doi: 10.1126/science.abc8665

Tong, S., Conrardy, C., Ruone, S., Kuzmin, I. V., Guo, X., Tao, Y., et al. (2009). Detection of Novel SARS-like and other coronaviruses in bats from kenya. *Emerg. Infect. Dis.* 15, 482–485. doi: 10.3201/eid1503.081013

Tseng, C.-T. K., Perrone, L. A., Zhu, H., Makino, S., and Peters, C. J. (2005). Severe acute respiratory syndrome and the innate immune responses: modulation of effector cell function without productive infection. *J. Immunol.* 174, 7977–7985. doi: 10.4049/jimmunol.174.12.7977

Turmelle, A. S., Jackson, F. R., Green, D., McCracken, G. F., and Rupprecht, C. E. (2010). Host immunity to repeated rabies virus infection in big brown bats. *J. Gen. Virol.* 91, 2360–2366. doi: 10.1099/vir.0.020073-0

Tynell, J., Westenius, V., Rönkkö, E., Munster, V. J., Melén, K., Österlund, P., et al. (2016). Middle East respiratory syndrome coronavirus shows poor replication but significant induction of antiviral responses in human monocyte-derived macrophages and dendritic cells. *J. Gen. Virol.* 97, 344–355. doi: 10.1099/jgv.0.000351

Tyrrell, D. A. J., and Bynoe, M. L. (1965). Cultivation of a novel type of common-cold virus in organ cultures. *BMJ* 1, 1467–1470. doi: 10.1136/bmj.1.5448.1467

Vabret, A., Dina, J., Gouarin, S., Petitjean, J., Corbet, S., and Freymuth, F. (2006). Detection of the new human coronavirus HKU1: a report of 6 cases. *Clin. Infect. Dis.* 42, 634–639. doi: 10.1086/500136

Vabret, N., Britton, G. J., Gruber, C., Hegde, S., Kim, J., Kuksin, M., et al. (2020). Immunology of COVID-19: current state of the science. *Immunity* 52, 910–941. doi: 10.1016/j.immuni.2020.05.002

van der Hoek, L., Pyrc, K., Jebbink, M. F., Vermeulen-Oost, W., Berkhout, R. J. M., Wolthers, K. C., et al. (2004). Identification of a new human coronavirus. *Nat. Med.* 10, 368–373. doi: 10.1038/nm1024

van der Most, R. G., de Groot, R. J., and Spaan, W. J. (1994). Subgenomic RNA synthesis directed by a synthetic defective interfering RNA of mouse hepatitis virus: a study of coronavirus transcription initiation. *J. Virol.* 68, 3656–3666. doi: 10.1128/JVI.68.6.3656-3666.1994

van Marle, G., Luytjes, W., van der Most, R. G., van der Straaten, T., and Spaan, W. J. (1995). Regulation of coronavirus mRNA transcription. *J. Virol.* 69, 7851–7856. doi: 10.1128/JVI.69.12.7851-7856.1995

Vijgen, L., Keyaerts, E., Lemey, P., Maes, P., Van Reeth, K., Nauwynck, H., et al. (2006). Evolutionary history of the closely related group 2 coronaviruses: porcine hemagglutinating encephalomyelitis virus, bovine coronavirus, and human coronavirus OC43. *J. Virol.* 80, 7270–7274. doi: 10.1128/JVI.02675-05

Vijgen, L., Keyaerts, E., Moës, E., Thoelen, I., Wollants, E., Lemey, P., et al. (2005). Complete genomic sequence of human coronavirus OC43: molecular clock analysis suggests a relatively recent zoonotic coronavirus transmission event. *J. Virol.* 79, 1595–1604. doi: 10.1128/JVI.79.3.1595-1604.2005

Virtue, E. R., Marsh, G. A., Baker, M. L., and Wang, L.-F. (2011). Interferon production and signaling pathways are antagonized during henipavirus infection of fruit bat cell lines. *PLoS One* 6:e22488. doi: 10.1371/journal.pone.0022488

V'kovski, P., Gerber, M., Kelly, J., Pfaender, S., Ebert, N., Braga Lagache, S., et al. (2019). Determination of host proteins composing the microenvironment of coronavirus replicase complexes by proximity-labeling. *Elife* 8:e42037. doi: 10.7554/eLife.42037

V'kovski, P., Kratzel, A., Steiner, S., Stalder, H., and Thiel, V. (2020). Coronavirus biology and replication: implications for SARS-CoV-2. *Nat. Rev. Microbiol.* 19, 155–170. doi: 10.1038/s41579-020-00468-6

von Brunn, A., Teepe, C., Simpson, J. C., Pepperkok, R., Friedel, C. C., Zimmer, R., et al. (2007). Analysis of intraviral protein-protein interactions of the SARS Coronavirus ORFeome. *PLoS One* 2:e459. doi: 10.1371/journal.pone.0000459

Wang, K., Chen, W., Zhou, Y.-S., Lian, J.-Q., Zhang, Z., Du, P., et al. (2020). SARS-CoV-2 invades host cells via a novel route: CD147-spike protein. *bioRxiv* [Preprint]. doi: 10.1101/2020.03.14.988345

Weinstein, R. A. (2004). Planning for epidemics — the lessons of SARS. *N. Engl. J. Med.* 350, 2332–2334. doi: 10.1056/NEJMp048082

Williams, M. A., and Lamb, R. A. (1989). Effect of mutations and deletions in a bicistronic mRNA on the synthesis of influenza B virus NB and NA glycoproteins. *J. Virol.* 63, 28–35. doi: 10.1128/JVI.63.1.28-35.1989

Wolff, G., Limpens, R. W. A. L., Zevenhoven-Dobbe, J. C., Laugks, U., Zheng, S., de Jong, A. W. M., et al. (2020). A molecular pore spans the double membrane of the coronavirus replication organelle. *Science* 369, 1395–1398. doi: 10.1126/science.abd3629

Woo, P. C. Y., Lau, S. K. P., Chu, C., Chan, K., Tsoi, H., Huang, Y., et al. (2005). Characterization and complete genome sequence of a novel coronavirus, coronavirus HKU1, from Patients with Pneumonia. *J. Virol.* 79, 884–895. doi: 10.1128/JVI.79.2.884-895.2005

Wu, F., Zhao, S., Yu, B., Chen, Y.-M., Wang, W., Song, Z.-G., et al. (2020). A new coronavirus associated with human respiratory disease in China. *Nature* 579, 265–269. doi: 10.1038/s41586-020-2008-3

Wynne, J. W., Di Rubbo, A., Shiell, B. J., Beddome, G., Cowled, C., Peck, G. R., et al. (2013). Purification and Characterisation of Immunoglobulins from the Australian Black Flying Fox (*Pteropus alecto*) using anti-fab affinity chromatography reveals the low abundance of IgA. *PLoS One* 8:e52930. doi: 10.1371/journal.pone.0052930

Xia, H., Cao, Z., Xie, X., Zhang, X., Chen, J. Y.-C., Wang, H., et al. (2020). Evasion of Type I interferon by SARS-CoV-2. *Cell Rep.* 33:108234. doi: 10.1016/j.celrep.2020.108234

Xiao, K., Zhai, J., Feng, Y., Zhou, N., Zhang, X., Zou, J.-J., et al. (2020). Isolation of SARS-CoV-2-related coronavirus from Malayan pangolins. *Nature* 583, 286–289. doi: 10.1038/s41586-020-2313-x

Xu, H., Zhong, L., Deng, J., Peng, J., Dan, H., Zeng, X., et al. (2020). High expression of ACE2 receptor of 2019-nCoV on the epithelial cells of oral mucosa. *Int. J. Oral. Sci.* 12:8. doi: 10.1038/s41368-020-0074-x

Xu, K., Zheng, B.-J., Zeng, R., Lu, W., Lin, Y.-P., Xue, L., et al. (2009). Severe acute respiratory syndrome coronavirus accessory protein 9b is a virion-associated protein. *Virology* 388, 279–285. doi: 10.1016/j.virol.2009.03.032

Yang, L., Wu, Z., Ren, X., Yang, F., Zhang, J., He, G., et al. (2014). MERS–related betacoronavirus in vespertilio superans Bats, China. *Emerg. Infect. Dis.* 20, 1260–1262. doi: 10.3201/eid2007.140318

Yao, H., Song, Y., Chen, Y., Wu, N., Xu, J., Sun, C., et al. (2020). Molecular architecture of the SARS-CoV-2 Virus. *Cell* 183, 730–738.e13. doi: 10.1016/j.cell.2020.09.018

Yi, H. (2020). 2019 novel coronavirus is undergoing active recombination. *Clin. Infect. Dis.* 71, 884–887. doi: 10.1093/cid/ciaa219

Yi, Y., Lagniton, P. N. P., Ye, S., Li, E., and Xu, R.-H. (2020). COVID-19: what has been learned and to be learned about the novel coronavirus disease. *Int. J. Biol. Sci.* 16, 1753–1766. doi: 10.7150/ijbs.45134

Yilla, M., Harcourt, B. H., Hickman, C. J., McGrew, M., Tamin, A., Goldsmith, C. S., et al. (2005). SARS-coronavirus replication in human peripheral monocytes/macrophages. *Virus Res.* 107, 93–101. doi: 10.1016/j.virusres.2004.09.004

Young, B. E., Fong, S.-W., Chan, Y.-H., Mak, T.-M., Ang, L. W., Anderson, D. E., et al. (2020). Effects of a major deletion in the SARS-CoV-2 genome on the severity of infection and the inflammatory response: an observational cohort study. *Lancet* 396, 603–611. doi: 10.1016/S0140-6736(20)31757-8

Yount, B., Roberts, R. S., Sims, A. C., Deming, D., Frieman, M. B., Sparks, J., et al. (2005). Severe acute respiratory syndrome coronavirus group-specific open reading frames encode nonessential functions for replication in cell cultures and mice. *J. Virol.* 79, 14909–14922. doi: 10.1128/JVI.79.23.14909-14922.2005

Yuan, X., Li, J., Shan, Y., Yang, Z., Zhao, Z., Chen, B., et al. (2005a). Subcellular localization and membrane association of SARS-CoV 3a protein. *Virus Res.* 109, 191–202. doi: 10.1016/j.virusres.2005.01.001

Yuan, X., Shan, Y., Yao, Z., Li, J., Zhao, Z., Chen, J., et al. (2006a). Mitochondrial location of severe acute respiratory syndrome coronavirus 3b protein. *Mol. Cells* 21, 186–191.

Yuan, X., Wu, J., Shan, Y., Yao, Z., Dong, B., Chen, B., et al. (2006b). SARS coronavirus 7a protein blocks cell cycle progression at G0/G1 phase via the cyclin D3/pRb pathway. *Virology* 346, 74–85. doi: 10.1016/j.virol.2005.10.015

Yuan, X., Yao, Z., Shan, Y., Chen, B., Yang, Z., Wu, J., et al. (2005b). Nucleolar localization of non-structural protein 3b, a protein specifically encoded by the severe acute respiratory syndrome coronavirus. *Virus Res.* 114, 70–79. doi: 10.1016/j.virusres.2005.06.001

Yuan, X., Yao, Z., Wu, J., Zhou, Y., Shan, Y., Dong, B., et al. (2007). G1 phase cell cycle arrest induced by SARS-CoV 3a protein via the cyclin D3/pRb pathway. *Am. J. Respir. Cell. Mol. Biol.* 37, 9–19. doi: 10.1165/rcmb.2005-0345RC

Yuen, C.-K., Lam, J.-Y., Wong, W.-M., Mak, L.-F., Wang, X., Chu, H., et al. (2020). SARS-CoV-2 nsp13, nsp14, nsp15 and orf6 function as potent interferon antagonists. *Emerg. Microbes Infect.* 9, 1418–1428. doi: 10.1080/22221751.2020.1780953

Zaki, A. M., van Boheemen, S., Bestebroer, T. M., Osterhaus, A. D. M. E., and Fouchier, R. A. M. (2012). Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. *N. Engl. J. Med.* 367, 1814–1820. doi: 10.1056/NEJMoa1211721

Zeng, Q., Langereis, M. A., van Vliet, A. L. W., Huizinga, E. G., and de Groot, R. J. (2008). Structure of coronavirus hemagglutinin-esterase offers insight into corona and influenza virus evolution. *Proc. Natl. Acad. Sci. U.S.A.* 105, 9065–9069. doi: 10.1073/pnas.0800502105

Zeng, R., Yang, R.-F., Shi, M.-D., Jiang, M.-R., Xie, Y.-H., Ruan, H.-Q., et al. (2004). Characterization of the 3a protein of SARS-associated coronavirus in infected vero E6 cells and SARS patients. *J. Mol. Biol.* 341, 271–279. doi: 10.1016/j.jmb.2004.06.016

Zhang, D., Guo, R., Lei, L., Liu, H., Wang, Y., Wang, Y., et al. (2021a). Frontline Science: COVID-19 infection induces readily detectable morphologic and inflammation-related phenotypic changes in peripheral blood monocytes. *J. Leukoc. Biol.* 109, 13–22. doi: 10.1002/JLB.4HI0720-470R

Zhang, Q., Zeng, L.-P., Zhou, P., Irving, A. T., Li, S., Shi, Z.-L., et al. (2017). IFNAR2-dependent gene expression profile induced by IFN-α in Pteropus alecto bat cells and impact of IFNAR2 knockout on virus infection. *PLoS One* 12:e0182866. doi: 10.1371/journal.pone.0182866

Zhang, T., Wu, Q., and Zhang, Z. (2020a). Probable pangolin origin of SARS-CoV-2 associated with the COVID-19 outbreak. *Curr. Biol.* 30, 1346–1351.e2. doi: 10.1016/j.cub.2020.03.022

Zhang, X., Song, K., Tong, F., Fei, M., Guo, H., Lu, Z., et al. (2020b). First case of COVID-19 in a patient with multiple myeloma successfully treated with tocilizumab. *Blood Adv.* 4, 1307–1310. doi: 10.1182/bloodadvances.2020001907

Zhang, Y., Sun, H., Pei, R., Mao, B., Zhao, Z., Li, H., et al. (2021b). The SARS-CoV-2 protein ORF3a inhibits fusion of autophagosomes with lysosomes. *Cell Discov.* 7:31. doi: 10.1038/s41421-021-00268-z

Zhang, Y.-Z., and Holmes, E. C. (2020). A genomic perspective on the origin and emergence of SARS-CoV-2. *Cell* 181, 223–227. doi: 10.1016/j.cell.2020.03.035

Zhao, J., Cui, W., and Tian, B. (2020). The potential intermediate hosts for SARS-CoV-2. *Front. Microbiol.* 11:580137. doi: 10.3389/fmicb.2020.580137

Zheng, B. J., Wong, K. H., Zhou, J., Wong, K. L., Young, B. W. Y., Lu, L. W., et al. (2004). SARS-related virus predating SARS outbreak, Hong Kong. *Emerg. Infect. Dis.* 10, 176–178. doi: 10.3201/eid1002.030533

Zheng, Y., Zhuang, M.-W., Han, L., Zhang, J., Nan, M.-L., Zhan, P., et al. (2020). Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) membrane (M) protein inhibits type I and III interferon production by targeting RIG-I/MDA-5 signaling. *Signal Transduct Target Ther.* 5:299. doi: 10.1038/s41392-020-00438-7

Zhong, X., Guo, Z., Yang, H., Peng, L., Xie, Y., Wong, T.-Y., et al. (2006). Amino terminus of the SARS coronavirus protein 3a elicits strong, potentially protective humoral responses in infected patients. *J. Gen. Virol.* 87, 369–373. doi: 10.1099/vir.81078-0

Zhou, F., Yu, T., Du, R., Fan, G., Liu, Y., Liu, Z., et al. (2020a). Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet* 395, 1054–1062. doi: 10.1016/S0140-6736(20)30566-3

Zhou, H., Chen, X., Hu, T., Li, J., Song, H., Liu, Y., et al. (2020b). A novel bat coronavirus closely related to SARS-CoV-2 contains natural insertions at the S1/S2 cleavage site of the spike protein. *Curr. Biol.* 30, 2196–2203.e3. doi: 10.1016/j.cub.2020.05.023

Zhou, J., Chu, H., Chan, J. F.-W., and Yuen, K.-Y. (2015). Middle East respiratory syndrome coronavirus infection: virus-host cell interactions and implications on pathogenesis. *Virol. J.* 12:218. doi: 10.1186/s12985-015-0446-6

Zhou, P., Cowled, C., Mansell, A., Monaghan, P., Green, D., Wu, L., et al. (2014). IRF7 in the Australian black flying fox, Pteropus alecto: evidence for a unique expression pattern and functional conservation. *PLoS One* 9:e103875. doi: 10.1371/journal.pone.0103875

Zhou, P., Cowled, C., Todd, S., Crameri, G., Virtue, E. R., Marsh, G. A., et al. (2011). Type III IFNs in pteropid bats: differential expression patterns provide evidence for distinct roles in antiviral immunity. *J. Immunol.* 186, 3138–3147. doi: 10.4049/jimmunol.1003115

Zhou, P., Cowled, C., Wang, L.-F., and Baker, M. L. (2013). Bat Mx1 and Oas1, but not Pkr are highly induced by bat interferon and viral infection. *Dev. Comp. Immunol.* 40, 240–247. doi: 10.1016/j.dci.2013.03.006

Zhou, P., Fan, H., Lan, T., Yang, X.-L., Shi, W.-F., Zhang, W., et al. (2018). Fatal swine acute diarrhoea syndrome caused by an HKU2-related coronavirus of bat origin. *Nature* 556, 255–258. doi: 10.1038/s41586-018-0010-9

Zhou, P., Li, H., Wang, H., Wang, L.-F., and Shi, Z. (2012). Bat severe acute respiratory syndrome-like coronavirus ORF3b homologues display different interferon antagonist activities. *J. Gen. Virol.* 93, 275–281. doi: 10.1099/vir.0.033589-0

Zhou, P., Tachedjian, M., Wynne, J. W., Boyd, V., Cui, J., Smith, I., et al. (2016). Contraction of the type I IFN locus and unusual constitutive expression of IFN-α in bats. *Proc. Natl. Acad. Sci. U.S.A.* 113, 2696–2701. doi: 10.1073/pnas.1518240113

Zhou, P., Yang, X.-L., Wang, X.-G., Hu, B., Zhang, L., Zhang, W., et al. (2020c). A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* 579, 270–273. doi: 10.1038/s41586-020-2012-7

Zhou, Y., Fu, B., Zheng, X., Wang, D., Zhao, C., Qi, Y., et al. (2020d). Pathogenic T-cells and inflammatory monocytes incite inflammatory storms in severe COVID-19 patients. *Natl. Sci. Rev.* 7, 998–1002. doi: 10.1093/nsr/nwaa041

Zhou, Z., Huang, C., Zhou, Z., Huang, Z., Su, L., Kang, S., et al. (2021). Structural insight reveals SARS-CoV-2 ORF7a as an immunomodulating factor for human CD14+ monocytes. *iScience* 24:102187. doi: 10.1016/j.isci.2021.102187

Zhu, Z., Meng, K., and Meng, G. (2020). Genomic recombination events may reveal the evolution of coronavirus and the origin of SARS-CoV-2. *Sci. Rep.* 10:21617. doi: 10.1038/s41598-020-78703-6