ELSEVIER

Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.

Review

Understanding the origin of COVID-19 requires to change the paradigm on zoonotic emergence from the spillover to the circulation model

Roger Frutos [a,*], Laurent Gavotte [b], Christian A. Devaux [c]

[a] *Cirad, UMR 17, Intertryp, Montpellier, France*
[b] *Espace-Dev, Université de Montpellier, Montpellier, France*
[c] *IHU-Méditerranée Infection and CNRS, Marseille, France*

ARTICLE INFO

ABSTRACT

While the COVID-19 pandemic continues to spread with currently more than 117 million cumulated cases and 2.6 million deaths worldwide as per March 2021, its origin is still debated. Although several hypotheses have been proposed, there is still no clear explanation about how its causative agent, SARS-CoV-2, emerged in human populations. Today, scientifically-valid facts that deserve to be debated still coexist with unverified statements blurring thus the knowledge on the origin of COVID-19. Our retrospective analysis of scientific data supports the hypothesis that SARS-CoV-2 is indeed a naturally occurring virus. However, the spillover model considered today as the main explanation to zoonotic emergence does not match the virus dynamics and somehow misguided the way researches were conducted. We conclude this review by proposing a change of paradigm and model and introduce the circulation model for explaining the various aspects of the dynamic of SARS-CoV-2 emergence in humans.

1. Introduction

Since the beginning of the COVID-19 pandemic in December 2019, a major issue has been the origin of the virus and how it was transmitted to humans. If the human-to-human transmission of SARS-CoV-2 has been rapidly demonstrated (Huang et al., 2020; Kucharski et al., 2020), explaining the magnitude of the pandemic, the origin of the virus remains elusive. COVID-19 is considered to have started at the Huanan Seafood Wholesale Market (HSWM) in Wuhan on December 8, 2019 (Huang et al., 2020; WHO, 2020). Following the description of the first clinical cases (Huang et al., 2020; Lu et al., 2020), SARS-CoV-2 was rapidly sequenced (Zhu et al., 2020) and linked to bats CoV belonging to the lineage b of betacoronavirus (Zhou et al., 2020a). RaTG13, the closest to SARS-CoV-2 genome, was isolated in 2016 from an anal sample of a horseshoe bat (*Rhinolophus affinis*) from Yunnan (Zhou et al., 2020a; Ge et al., 2016). Another batCoV sequence, RmYN02, was identified also in Yunnan in the horseshoe bat *Rhinolophus malayanus* (Zhou et al., 2020b). Since then, viruses closely related SARS-CoV-2 have been described in *Rhinolophus shameli* bats caught in 2010 in Cambodia (Hul et al., 2021) and in *Rhinolophus acuminatus* bats from Thailand (Wacharapluesadee et al., 2021). Beside laboratory accidents, there is no evidence of direct transmission of bat-CoV to humans

(Heymann et al., 2004; Watts, 2004a; Webster, 2004; WHO, 2004). Since the spillover model postulates that an animal reservoir must be at the origin of the zoonosis (Plowright et al., 2017), a hunt for this intermediary host began using *in-silico* studies. Snakes were first proposed (Ji et al., 2020) and after rejection of the hypothesis (Zhang et al., 2020a; Callaway and Cyranoski, 2020), the Malayan or Sunda pangolin, *Manis javanica*, was in turn designated as intermediate host (Xiao et al., 2020a; Zhang et al., 2020b). The *in-silico* modeling of SARS-CoV-2 spike and ACE2 receptor highlighted a potential affinity between SARS-CoV-2 spike and pangolin ACE2, but it was not limited to this species (Liu et al., 2020a; Luan et al., 2020; Shi et al., 2020). Although contradicted (Frutos et al., 2020a; Li et al., 2020; Liu et al., 2020b) the pangolin hypothesis was repeated so many times in scientific papers and on social networks that was taken for a truth even within part of the scientific community. This leads to a key feature of COVID-19: it is the very first pandemic to occur in our hyper connected society. A rush for scientific publications occurred with 111,726 Pubmed references for "COVID-19" as per March 11, 2021. A parallel 'digital pandemic' (overcommunication on more or less probable 'scientific hypothesis') developed on social networks, bringing opinions and conspiracy theories, generating anxiety and irrational behavior. However, two major issues remain unresolved: i) the origin of the virus and ii) the initial route of infection leading to the

---

pandemic.

2. The origin of SARS-CoV-2

2.1. SARS-CoV-2: the man-made virus theory

The origin of SARS-Cov-2 is still passionately debated since it makes ground for geopolitical confrontations and conspiracy theories besides scientific ones. The first hypothesis is that of a man-made virus raised first by Pradhan and colleagues who claimed to have observed the presence of HIV sequences in SARS-CoV-2, before retraction of the manuscript and then by Perez and Montagnier (2020). This hypothesis was rebutted by bioinformatic analyses showing that the similarity of those short putative HIV insertions was insufficient to support a common ancestral origin of the sequences (Liu et al., 2020c; Xiao et al., 2020b). This article has since then been retracted. Furthermore, the four insertions identified in SARS-CoV-2 occurred independently at different times of coronaviruses diversification (Sallard et al., 2020). Other hypotheses were based on the construction of infectious recombinant SARS-CoV capable to replicate in mammalian cell lines or animal models (Becker et al., 2008; Menachery et al., 2020). Before retraction of their manuscript, Pradhan et al. also claimed that the RBD of SARS-CoV-2 might have been engineered by using a RBD domain with a higher affinity for the human ACE2 receptor and by inserting the RRAR furin cleavage site downstream of the RBD, making the virus more infectious in human cells. This hypothesis was mostly motivated by the fact that this furin cleavage site is unique to SARS-CoV-2 among all Sarbecoviruses (Andersen et al., 2020; Coutard et al., 2020). However, the supposedly engineered sequences were simply natural features (Liu et al., 2020c; Andersen et al., 2020; Hao, 2020; Othman et al., 2020). Furthermore, naturally occurring polybasic furin cleavage sites have been described in other lineages of coronaviruses such as MERS-CoV, HKU1, HCoV-OC43 or IBV (Andersen et al., 2020; Huang et al., 2006; Yamada and Liu, 2009) and is a common feature in viral envelope glycoproteins (Hao, 2020; Dimitrov, 2004). The natural occurrence of furin-cleavage sites in various viruses has been documented for long. We provide a list of 50 selected references as Supplementary Data. Some linked the presence of the least preferred CGG codons in the SRAS-CoV-2 furin cleavage sites as a "proof" of engineering. A codon being least preferred does not mean it should never exist and this CGG codon present in SARS-CoV-2 is for instance present at a higher rate in MERS-CoV (Chen et al., 2017; Hou, 2020). The lower presence of CpG (intrachain Cytosine-Guanosine dinucleotide linked by a phosphate bond) in human pathogens has been shown to be a selective process. CpGs trigger direct B-Cell activation and therefore these dinucleotides provide a selective disadvantage (Krieg et al., 1995). However, they nevertheless exist in human pathogens. When considering the huge selective advantage in transmissibility brought by the furin-cleavage site and the disadvantage brought by the B-cell activation, the net result in largely in favor of the furin-cleavage site leading to the fixation of these mutations in the human populations even if it involves the rare CGG codons. This is a simple and straightforward selective and evolutionary process. RmYN02 also carries an indel at the same place as the furin site (PRRA) indel of SARS-CoV-2 with the insertion of a PAA sequence and the deletion of the immediate QTQT upstream sequence (Zhou et al., 2020b). The naturally occurring PAAR sequence displayed by RmYN02 is not active as a furin-cleavage site but is only one mutation away from the active RNNR furin cleavage site. One additional mutation turning the proline (P) into an arginine (R) will generate a RAAR active furin-cleavage site. This suggests that viruses other than SARS-CoV-2 were under similar selective pressure. The selection of SARS-CoV-2 through successive passages in cell culture was refuted (Andersen et al., 2020). Scientists at the Wuhan Institute of Virology denied having carried out engineering and gain of function experiments on SARS-CoV-2but only on SARS-CoV in published and openly displayed international collaborations (Cohen, 2020). Altogether, these elements indicate that there is no evidence to support the hypothesis of a man-made origin of SARS-CoV-2.

2.2. SARS-CoV-2: the bat-pangolin recombinant virus theory

The next hypothesis to consider is the natural origin of SARS-CoV-2 as a recombinant between a Sarbecovirus from Malayan pangolins and RaTG13 from *R. affinis*. This hypothesis found its rationale in the *in-silico* analysis of sequence similarities which indicated that SARS-CoV-2 was closely related to RaTG13 but displayed specific RBD sites similar to the pangolin virus (Xiao et al., 2020a; Zhang et al., 2020b; Li et al., 2002; Lam et al., 2020). However, this recombination theory was dismissed, too many misalignments being present (Boni et al., 2020; Paraskevis et al., 2020; Chen et al., 2020). Furthermore, the detection of recombination was deduced from metagenomic data, an approach which can by itself generate artifactual recombinants. The recombinant hypothesis implies that both viruses must be at the same time in the same host cell. Not only *R. affinis* and pangolins do not share the same habitat, but no Sarbecovirus was isolated or detected in Chinese pangolins, *Manis pentadactyla* (Xiao et al., 2020a). *R. affinis* bats from China and from the Indomalayan region belong to two different subspecies, *R. affinis affinis* and *R. affinis superans*, respectively (Ith et al., 2015). The northern most limit of *R. affinis superans* is Southern Thailand (Ith et al., 2015). The RaTG13 virus was detected in an anal swab sample of a *R. affinis* bat in Yunnan (Ge et al., 2016; Cohen, 2020) whereas Sarbecoviruses were exclusively found in Malayan pangolins, *M. javanica,* smuggled from the Indomalayan region (Xiao et al., 2020a; Lam et al., 2020; Liu et al., 2019). No Sarbecovirus was detected in 334 Malayan pangolins confiscated by the Malaysian customs (Lee et al., 2020). Furthermore, SARS-CoV-2 is phylogenetically branching at an ancestral level to Sarbecoviruses from pangolins, making it impossible to be a descendant from recombination (Wenzel, 2020). The question is therefore whether the infected Malayan pangolins were natively infected or were they instead infected during the smuggling process through contact with humans or other animals? A similar analysis was performed for SARS and MERS (Frutos et al., 2021). There are currently only *in silico* predictions from metagenomic data and no factual element to support the recombinant hypothesis.

2.3. SARS-CoV-2: a naturally occurring virus

The only remaining rational option for the origin of SARS-CoV-2, is that of a naturally occurring virus circulating in the wild which came into contact with humans. However, there is still no information on where and when this contact with humans occurred at the first place. An extensive mutational bias is introduced by the host APOBEC (Apolipoproteins B mRNA editing enzyme, catalytic polypeptide-like) RNA editing system. The extensive APOBEC-driven mutational bias and adaptation suggest that SARS-CoV-2 might have circulated unnoticed in humans for a long time (Matyasek and Kovarik, 2020; Zhan, 2020), but molecular clocks indicate instead a recent introgression into humans (Lai, 2020; Zehender et al., 2020). However, the Sarbecoviruses genome is saturated with transition/transversion (Ts/Tv) ratios of 1 or below 1 for the RdRp and spike genes (Frutos et al., 2021). This indicates that the genome is highly mutated and that the linear mutation/time relationship on which molecular clock analyses are based is no longer linear and might not be significant (Frutos et al., 2021).

3. The transmission of SARS-CoV-2 to humans

3.1. The conspiracy theory of SARS-CoV-2: the voluntarily release from a laboratory

The other issue to be addressed beside the origin of SARS-CoV-2 is how this virus infected human beings at the first place. The marginal conspiracy theory of a voluntary released of an engineered virus forwarded by the press, blogs and politicians (Sutton, 2020; Everington,

2020) is not supported by any data (Calisher et al., 2020; Fowdy, 2020). This hypothesis of voluntary release has an impact on part of the population experiencing fear and distress, especially because there is still no clear explanation for the route of SARS-CoV-2 infection. There is consensus within the scientific community to consider that SARS-CoV-2 has not been engineered and is a naturally occurring virus. Therefore, it is simply impossible to voluntarily release an engineered virus which does not exist. There is thus no voluntary release (Calisher et al., 2020).

3.2. A laboratory accident?

Another hypothesis is the accidental infection of laboratory staff working on naturally occurring Sarbecoviruses. Accidents happen and have already been reported during the SARS epidemic in Taiwan, Singapore and China (Webster, 2004; WHO, 2004). This is not limited to SARS-CoV (Heymann et al., 2004). When it happened in Beijing in 2004, the information was immediately released and an investigation involving both WHO and Chinese governmental agencies was conducted, patients were identified and treated (WHO, 2004). There is today no evidence that such an accident had happened with SARS-CoV-2.. Because of the incubation period of COVID-19, the weak symptoms, the significant rate of asymptomatic patients and the low virulence (with an estimated fatality rate of 3.26%, but more likely around 1% to 2% which is significantly lower than SARS-CoV with 9.6%), an accident could have easily remained unnoticed. But staff members of the Wuhan Institute of Virology have all been tested negative indicating that no accident occurred there (Cohen, 2020). One must remember that SARS-CoV-2 was never found in the wild and that RaTG13 does not exist as real virus but instead only as a sequence in a computer (Zhou et al., 2020a; Ge et al., 2016). It is a virtual virus which thus cannot leak from a laboratory. This hypothesis has been considered as "extremely unlikely" by the official WHO investigation team (Dyer, 2021). Therefore, although a laboratory accident can never be definitively excluded, there is currently no evidence to support it.

3.3. A contamination from rural and wild environments

A more likely hypothesis is a contamination in rural and wild environments. Owing to the designation of HSWM as the official epicenter and origin of the COVID-19 pandemic, the focus was put on cities and wet markets. However, the main risk of contact and viral contamination lies in anthropized rural environments and to a lower extent in the recreational human presence in wild environments. Coronaviruses are circulating in various animal species and thus contact with animals, respiratory droplets or feces, occurring preferentially in rural areas, represent the main route of human contamination. Land conversion in these areas generates mosaic landscapes attracting wild animals and bringing them to close contact with humans (Afelt et al., 2018; Reuter, 2016). Deforestation is aggravating this phenomenon (Afelt et al., 2018). The concentration and diversity of bat-borne viruses is higher in human rural settlements than in the wild (Afelt et al., 2018; Reuter, 2016). With a growing human population, a need for more converted land for agriculture and housing, and a fast-growing deforestation, the probability of seeing further emerging coronavirus-related infectious diseases is rising.

4. The dynamic of zoonoses: definitions and concepts

Considering that SARS-CoV-2 is a naturally occurring virus, the main question is then to understand how such a virus can come into contact with humans and cause a major pandemic. However, it is important before addressing this question to review the definitions and concepts involved in this process. This an important issue because the distorted use of concepts often leads to misunderstandings. Definitions of key concepts are given below (Box 1). A first key concept to address is that of the species barrier. A species is an artificial concept, a simplistic representation of the biological entity for the purpose of classification. A species is perceived as an isolated entity, hence the concept of species barrier which considers that to move from a species to another, a virus must cross an undetermined and uncharacterized virtual barrier. These species classification criteria are visible morphological and physiological traits used for discriminating and classifying populations, but they are unrelated to traits involved in interactions with viruses. A virus is not spreading based on species and species barriers but simply based on its ability to recognize a receptor and circumvent the host immune defenses. This occurs regardless of the "species" status given by humans using classification criteria. There is no distinction between "animal hosts", "human hosts", hence there is no such thing as the crossing of the species barrier. Humans simply make one species among others manipulated as well by viruses for replication and dissemination. The concept of zoonosis is an intellectual anthropocentric construct simply differentiating humans from animals. COVID-19 itself which displayed cases of transmission from humans to animals such as cats, tigers or minks demonstrates that the circulation goes from humans to animals without any problem (Halfmann et al., 2020; Enserink, 2020). Another major source of confusion is the concept of disease. A disease is a medical concept not a biological one. It is defined by the existence of a specific pattern of symptoms physicians can name and recognize. An emerging disease is by definition unknown until a specific set of symptoms is recognized and named. An epidemic is not defined by a geographic expansion of a disease. It is the appearance of disease, i.e. a specific set of symptoms, on a large number of people at the same time or within a short period of time. Therefore, an emerging disease is recognized as such only after having reached an epidemic stage. Before that, the virus circulates in the population, most likely leading to sporadic cases which are not recognized as a novel disease but confused with a known disease, many early symptoms being indeed similar. This is a latency phase or stuttering phase during which the disease in uncharacterized and the virus is undetected (Frutos et al., 2020a; Getz and Dougherty, 2018; Hartfield and Alizon, 2013; Lo Iacono et al., 2016). This also indicates that an epidemic, and the recognition of a novel disease, does not start from few cases but when an epidemic threshold or Critical Community Size (CCS) is reached (Hartfield and Alizon, 2013). This means that a minimum percentage of the host population must be infected and actively transmit the virus to trigger the epidemic. This is for instance calculated annually for the seasonal influenza to determine the beginning of the epidemic phase (Vega et al., 2013). Hartfield and Alizon (2013) elegantly developed the equivalent concept of outbreak threshold for the development of an emerging disease epidemic. However, unlike known diseases like influenza, a novel emerging disease cannot by definition be identified until the epidemic/outbreak threshold was reached and the epidemic has actually started. The spillover concept was initially invented to represent the risk of epizootics in wildlife from livestock, the reverse being named spillback (Daszak et al., 2000). It was initially not considered a zoonotic model but an epizootic one. The spillover model was later formalized by Power and Mitchell (2004) who defined it as follows: "In some host populations, epidemic or endemic diseases may be primarily driven not by intraspecific pathogen transmission within that population but by transmission from a reservoir species that maintains a relatively high pathogen population. In such a case, the pathogen typically reaches high prevalence in the reservoir and then spills over into the other host, a process called "the spillover effect" or "pathogen spillover"". Power and Mitchell (2004) provided in their definition a rational hypothesis for crossing the outbreak threshold. The concept of spillover was later on distorted from its original meaning to become a synonym of altogether contamination, infection and transmission even though these three concepts are totally different (Borremans et al., 2019; Plowright et al., 2017). Contamination refers to the exposure to a pathogen whereas infection refers to the successful establishment and replication in a host. An infection is always preceded by a contamination whereas a contamination does not necessarily lead to an infection. Transmission is not a status in the host-pathogen

Table 1
Comparison of the spillover and circulation models key characteristics.

Model	Event	Observed/not observed
Spillover	Presence of SARS-CoV-2 in a reservoir	Not observed
	Presence of SARS-CoV-2 in an intermediate species	Not observed
	Identification of an intermediate species	Not observed
	High virus incidence in an intermediate species	Not observed
	Epizootic	Not observed
	Epidemic	Observed
	Pandemic	Observed
Circulation	Circulation of a metapopulation of SARS-related viruses	Observed
	Presence of SARS-related viruses in various hosts	Observed
	Presence of the virus in the human population before the epidemic (too early to say forSARS-CoV-2 but true for previous Epidemics or pandemic)	Observed
	Quasispecies model of virus evolution	Observed
	Intra-host evolution (mutations/variants)	Observed
	Anthropogenic amplification loops	Observed
	Epidemic	Observed
	Pandemic	Observed

relationship but a process of dissemination which is very diverse. This usage of spillover as a synonym of so many different concepts and processes is a source of confusion and makes it challenging to develop relevant models (Cross et al., 2019). Furthermore, this usage of the spillover concept does not bring any explanation on how the virus can reach the outbreak threshold. Another important concept when addressing the emergence of diseases through the spillover process, whatever the definition, is that pathogens are most of the time considered as passive elements, sometimes referred to as "propagules". A pathogen is all but a passive element. Microorganisms are not necessarily pathogens. They might be commensals or symbionts with other hosts than humans and display complex multi-hosts replication cycles. However, viruses are pathogens by nature but their virulence may vary considerably from one host to another or one population to another. Nevertheless, parasitic microorganisms are not passive but instead active in the process of transmission. It is part of their replication cycle and has evolved to use various hosts for their replication, genetic exchange and dissemination. Hosts are totally passive in terms of mechanisms of transmission. Their only active implication is on replication machinery, mobility and contact, something often manipulated by the pathogen/parasite (Poulin, 2010). A well-known example is the neuronal pathology and behavioral modification induced by the rabies virus leading to the aggressiveness and biting behavior favorable for virus transmission (Hueffer et al., 2017). The diversity of cycles, transmission processes and mechanisms of infection makes it very difficult to model and fully apprehend the process of inter-species and inter-population translation of microorganisms.

4.1. Why is the spillover model not compatible with the observed dynamic of COVID-19?

We consider here the spillover theory as formulated by Power and Mitchell (2004) which states that a zoonotic emergence is preceded by an epizootic in an animal species at such a high level that the pathogens are spilling over from this species to inundate other species. There is thus a zoonotic pressure that triggers a high-incidence infection in humans essential to reach the outbreak threshold and starts the epidemic in the human population. Another consequence of the spillover model is that there must be an "animal intermediate species" also called "reservoir" bearing the same virus as the one causing the epidemic. This theory of spillover is the reference driving strategies for preventing and controlling emerging infectious diseases at the early stage. It is at the origin of the search for intermediate species and screening projects such as the

Global Virome or PREDICT (Carroll et al., 2018; Jonas and Seifman, 2019) which objectives are to identify potential zoonotic viruses circulating in the wild. In the COVID-19 context this intermediate species is supposed to make the link between bats, the putative original virus reservoir, and humans, the final recipient host. As soon as COVID-19 broke out, the hunt for the animal reservoir started and all researches (conducted *in silico*) have thus been shaped by the spillover model. No predictions from the spillover model were confirmed (Table 1, Fig. 1). The pangolin was the main hypothesis. However, the virus sequences considered came from metagenomic analyses on smuggled Malayan pangolins confiscated by Chinese customs before the COVID-19 crisis (Zhang et al., 2020b; Liu et al., 2020b; Liu et al., 2019). The status of the pangolin as formal intermediary was only built through successive deformation of the initial reports even though articles showed that this hypothesis was not valid (Zhang et al., 2020b; Frutos et al., 2020a; Li et al., 2020; Liu et al., 2020b; Lee et al., 2020; Wenzel, 2020; Tang et al., 2020). No related epizootic was described in pangolins or other animals in China or elsewhere. Furthermore, no SARS-CoV-2 related viruses were reported in wild animals other than smuggled Malayan pangolin and *Rhinolophus* bats. Thus, to date, no experimental data support a spillover of SARS-CoV-2 from any animal species. Since all researches and strategies are based on this model, there is a risk of misled investigations. The main problem associated with the spillover model is that it is a theoretical construction and did not come from evidence.

4.2. The circulation model: an evidence-based model

The need was thus to build an evidence-based model starting from actual observations in order to describe the COVID-19 dynamic of emergence (Table 1, Fig. 1). This process led to the development of the circulation model. It is an integrative model which considers all components of emerging diseases: the different hosts, the virus and the societal factors allowing to pass through the outbreak threshold into a pandemic. There is thus no wonder that the circulation matches all observed traits since it was built after them to match them all. The circulation model starts from the concept that viruses (and other pathogenic microorganisms) are not passive transportable elements but instead organisms which have evolved to use hosts for multiplication and dissemination as part of their replication cycle (Poulin, 2010). Viruses are naturally circulating from host to host only on the basis of compatibility, i.e. receptor recognition, and manipulate their hosts for mobility and dispersion, i.e. contact and transmission if the host in contact is susceptible. The circulation model thus considers that viruses circulate broadly within the animal kingdom regardless of the animal or human status and the classification within a given species. Another observation is that the contamination most of the time does not necessarily result in an epidemic. The compatibility is a matter of physiology, evolution and adaptation whereas the contact is mostly a matter of human activities. Considering COVID-19 under the light of the circulation model allows to understand/explain many observations (Table 1, Fig. 1). There is no reservoir but simply viruses circulating from one compatible host to another compatible host upon contact. This circulation has no specific animal to human orientation. A virus can be present in a host long before any epidemic outbreak. SARS-CoV has been detected on human samples from 2001 indicating that humans might have been exposed long before the SARS outbreak (Zheng et al., 2004). This was also observed with earlier pandemics like HIV/AIDS (Worobey et al., 2016). It is also in agreement with reports of SARS-CoV-2 adaptation and circulation in humans before December 2019 (Matyasek and Kovarik, 2020; Frutos et al., 2020b). This is a consequence of the mode of evolution of COVID-19 and other RNA viruses. These viruses evolve differently in each host due to the quasispecies process and the specific selective pressure generated by each host. It is thus not possible to find in another host species the same virus as that causing the epidemic, but only related viruses. Indeed, it has until now not been possible to find in the wild the same virus as that causing an epidemic. Furthermore, the
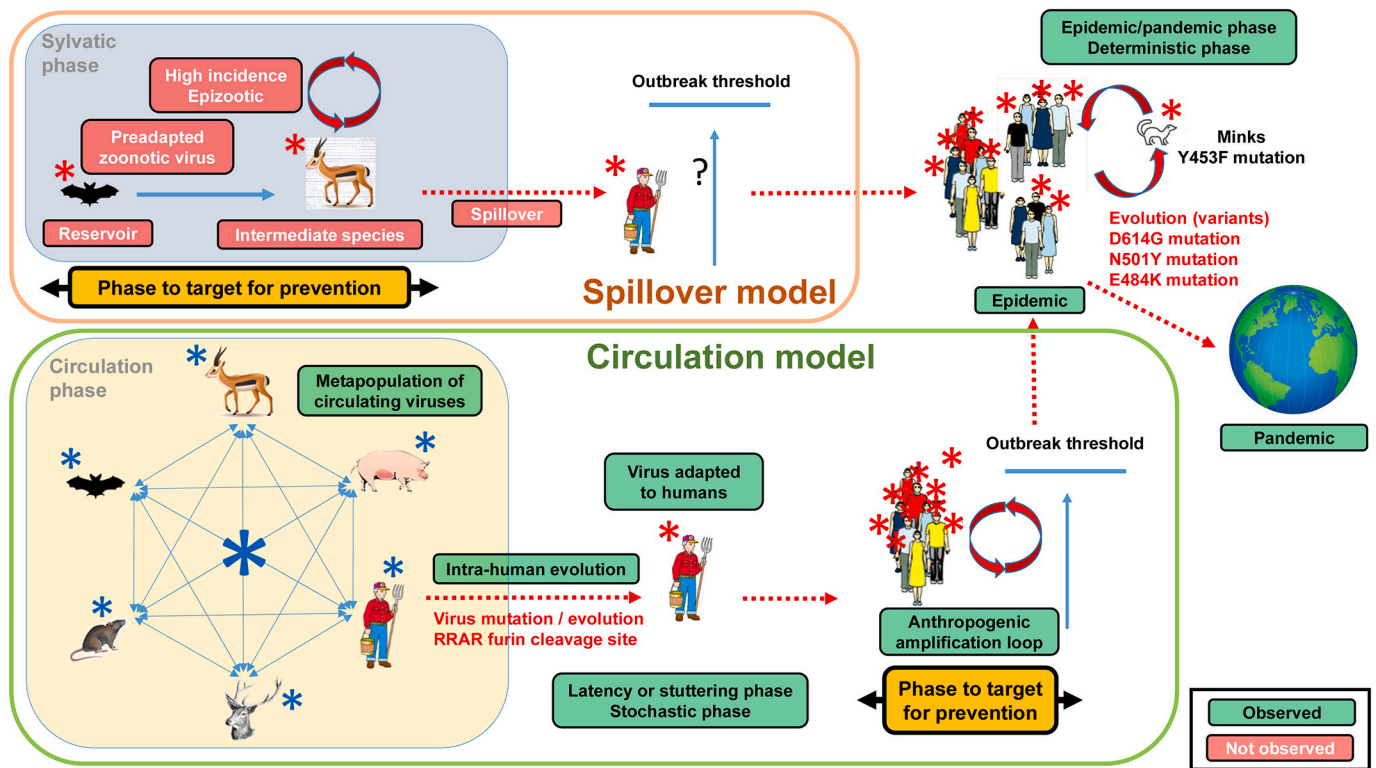
Fig. 1. Comparison of the spillover and circulation models.

a. Spillover model.

According to the spillover model, a preadapted SARS-CoV-2 virus is present in a reservoir which is not in contact with human populations. An intermediate species in contact with the human population is transmitting this virus to humans. According to the definition by Power and Mitchell (2004), this intermediate species is experiencing a high virus incidence leading to the spilling over of the virus into the human population. This obligatorily translates into an epizootic. A high incidence of a virus is obligatorily associated with a disease, i.e. a set of specific symptoms, since a virus is an intracellular pathogens destroying host cells. This would explain how the virus population can undergo the growth necessary to cross the outbreak threshold. If no epizootic occurs, the question is thus to explain how the virus can reach the outbreak threshold needed to trigger an epidemic (represented in the figure by a question mark). The epidemic of COVID-19 which started locally in a human population is expanding into a pandemic owing to international mobility. Under the spillover model, the phase to target to prevent an epidemic is the sylvatic phase with the search for the reservoir and intermediate species being key issues. Red boxes indicate elements which have not been observed under actual conditions. Green boxes indicate prerequisites and elements which have been observed under actual conditions. Red viruses represent viruses adapted to humans.

b. Circulation model.

According to the circulation model, a metapopulation of SARS-related viruses circulate in various susceptible hosts depending upon contact. Different virus populations are found in each host due to the quasispecies evolutionary process. Humans represent on host among other and participate to the global circulation. These infections are under a stochastic process and do not trigger epidemics or epizootics. Within the human population, under host pressure, the virus population is acquiring the mutations characteristic of SARS-CoV-2, in particular an increased transmissibility. At this stage no epidemic exists, no disease is described and SARS-CoV-2 is still in the stochastic phase. The human society is providing though gatherings, meetings, markets, etc. the amplification loop needed to reach the outbreak threshold. Once the outbreak threshold is reached, an epidemic outbreak occurs leading to the description of the COVID-19 disease. The virus is now in the deterministic phase. The epidemic of COVID-19 which started locally in a human population is expanding into a pandemic owing to international mobility. Under the circulation model, the phase to target to prevent an epidemic is the amplification loop with is linked to human activities. Red boxes indicate elements which have not been observed under actual conditions. Green boxes indicate prerequisites and elements which have been observed under actual conditions. Red viruses represent viruses adapted to humans. Blue viruses represent viruses not adapted to humans.

quasispecies process is an excellent means of quickly adapting to another host. These viruses are by nature multi-hosts viruses and being multi-hosts obligatorily require regular passages from one susceptible species to another susceptible species upon contact. The quasispecies evolution of these viruses also explains the pre-epidemic process and early in-host presence. Owing to the intra-host evolution mechanism of the quasispecies, host-specific mutations appear after the initial infection, not before. There is no preadaptation. This explains why the same virus cannot be found in the wild and looking for a virus preadapted to humans as in the spillover model cannot succeed. The RNA virus present in humans corresponds to the human avatar of the metapopulation of viruses. In other animals, the virus will evolve differently because of the selective pressure is different. What leads to an epidemic is not the presence of an already human pathogenic virus or a zoonotic pressure in an intermediate species delivering enough viruses to start the outbreak in a new population, but the occurrence of a double accident. The first

accident is the occurrence during this intra-host evolution of a mutation in a virus already circulating in humans making it more virulent or more transmissible. This was for instance exemplified with influenza viruses, chikungunya or Zika (Webster et al., 1982; Tsetsarkin and Weaver, 2011; Yuan et al., 2017). In the case of SARS-CoV-2, this primary accident could well be the generation of the RRAR furin cleavage site, making the virus capable of efficiently replicate in humans (Johnson et al., 2020), directly as an indel or by mutation of an existing indel as that of RmYN02. It also corresponds to the D614G mutation in the spike protein which also increases the infectivity of SARS-CoV-2 (Korber et al., 2020) or to the more recent N501Y mutation corresponding to the variant called B.1.1.7-SARS-CoV-2 also known as the "British variant" which increases strongly the transmissibility of SARS-CoV-2 (Leung et al., 2021). The RRAR sequence is not necessary for infection of human cells (Wong et al., 2020) indicating that an ancestral virus with or without the indel but lacking the furin cleavage site could have infected

---

Box 1. Definitions

Spillover

The term "spillover" was initially developed by Daszak et al. (2000) to characterize the transmission of infectious diseases from domestic animals to wildlife. The reverse was named "spillback". Initially, "spillover" was not associated to zoonoses. The concept of "spillover" was formally defined by Power and Mitchell (2004) as follows: "In some host populations, epidemic or endemic diseases may be primarily driven not by intraspecific pathogen transmission within that population but by transmission from a reservoir species that maintains a relatively high pathogen population. In such a case, the pathogen typically reaches high prevalence in the reservoir and then spills over into the other host, a process called "the spillover effect" or "pathogen spillover". The term spillover was later used as a synonym of any kind of contamination, infection or transmission (Plowright et al., 2017) losing thus any specific meaning. In this work, we use the definition given by Power and Mitchell (2004).

Pathogen

As previously reported (Devaux, 2012): "Pathogen developed strategies aimed at: (1) maximizing invasion rate; (2) selecting host traits that can reduce their impact on host life span and fertility; (3) ensuring timely replication and survival both within host and between hosts; and (4) facilitating reliable transmission to progeny".

Emerging infectious disease

As previously reported (Devaux, 2012): "Emerging infectious diseases (EIDs) belong to a nosological entity whose nature is proved infectious, regardless of the pathogen, or only suspected in case of novelty and until the agent is identified. It is understood that the identification of a 'new' pathogen is compatible with a previously undisclosed preexisting one. By extension, it is assumed that the EIDs include infectious diseases known endemic showing an obvious resurgence. An EID can affect all types of eukaryotic organisms. The EIDs generally have a high social impact and economic consequences. An EID is obviously unusual; it is surrounded by uncertainty and anxiety, real or perceived, as to its evolutionary potential, its impact on health and the ability of leaders and stakeholders to control the phenomenon. These emerging diseases are therefore: (1) the development of a new disease, a consequence of a new pathogen (in its nature, its mode of transmission, its expression and/or its adaptation to host species); and (2) a disease previously identified but whose manifestations are new (associated with a sudden increase in the incidence, severity or geographical area within a time span of a few weeks/months to one or several decades)".

Species

A species is an intellectual construction aiming at defining the organization of wildlife for the purpose of classification based on morphological and/or genomic similarity, and on reproductive isolation. Individuals from two different species cannot reproduce or cannot give a fertile offspring. The species is the human representation of the organization of life, not the actual organization. The species is a static and deterministic concept.

Species barrier

The species barrier is a concept which postulates that an undefined and uncharacterized "barrier" exists between two species preventing pathogens to move from one species to another. In order for a pathogen to cross the barrier and move from a "donor species" to a "recipient species", it must be preadapted to the latter.

Metapopulation

The metapopulation was defined by Richard Levins in 1969. A metapopulation is a population of populations. A metapopulation is made of spatially distributed populations sharing most of their genomic background, allowing thus crossfertility, but displaying also significant genomic differences. The metapopulation is biological concept closest to the intellectual concept of species but unlike species it is a dynamic concept.

Quasispecies

The quasispecies model is a model of evolution specifically developed for RNA viruses which utilize a polymerase, the RNA dependent RNA polymerase (RdRp), which is characterized by a low fidelity and a high rate of mutations (Xu et al., 2004; Briese et al., 2014; Andino and Domingo, 2015; Karamitros et al., 2020). The quasispecies model of evolution is based on the concept of the "flattest" in which numerous variants generated at each cycle, and covering all together the whole spectrum of possible mutations (sequence space), cooperate to allow the survival of the virus in a host. There is no fixation of mutations. This is opposed to the concept of "fittest" in which one or a limited number of fixated mutations preadapt a virus to a host. In the quasispecies model of evolution, mutations occur after infection as a consequence of the host pressure, mostly to avoid host immune defenses. These mutations are host dependent.

Disease

A disease is a medical concept defining a pathology characterized by a specific set of symptoms to which a name is given.

Epidemic

An epidemic is the occurrence of a significant number of cases in the human population displaying the same symptoms (disease) in a short period of time. The equivalent in an animal population is called an epizootic.

Epidemic threshold

The epidemic threshold is a deterministic concept. It is the minimal number of infected individuals in a population needed for an epidemic to start. It is for instance calculated annually to determine the beginning of the epidemic phase of seasonal influenza.

Outbreak threshold

The outbreak threshold was defined by Hartfield and Alizon (2013) as the equivalent of the epidemic threshold for an emerging disease for which, by definition, there is no pre-established set of symptoms to identify it.

Stochastic phase

The stochastic or probabilistic phase is the phase of a pathogen cycle during which it is exposed to many factors and drifts which result in irregular infection frequency or incidence and which could yield to the extinction of the pathogen population. During this phase, corresponding to the "latency phase" or "stuttering phase" during which no disease is characterized and the infection remains often unnoticed or confuses with another disease.

Deterministic phase

The deterministic phase of a pathogen replication cycle is the phase following the crossing of the epidemic or outbreak threshold. At this stage the pathogen cannot go extinct by chance and an epidemic is triggered.

Amplification loop

An amplification loop defines the set of events, essentially anthropogenic, which lead to pathogen population growth allowing to cross the epidemic of outbreak threshold and to move a pathogen from the stochastic phase to the deterministic phase.

---

humans. The rest of the dynamic can be explained by in-host evolution. SARS-CoV-2 and coronaviruses are evolving according to the quasispecies model (Song et al., 2005; Karamitros et al., 2020; Chaudhry et al., 2020). The host-driven quasispecies evolutionary process facilitates the acquisition of the furin cleavage site (Chaudhry et al., 2020). The initial infection of humans by a parental virus which acquired the RRAR sequence later as part of the in-host evolution under the combined effect of host editing systems and quasispecies dynamic can explain why SARS-CoV-2 looked immediately adapted to humans while displaying a high replication efficiency (Zhan, 2020; Zhang and Holmes, 2020). The RRAR-variant displays a tremendous advantage over any other variants and is strongly selected. This corresponds to a selective sweep and explains the low variability of SARS-CoV-2 and the apparent lack of evolution displayed by SARS-CoV-2 (Zhan, 2020). The furin-cleavage site might have also been acquired partially or completely before the primary human infection. The emergence of SARS-CoV-2 in humans can be explained under the circulation model through well-known evolutionary mechanisms and in host evolution without resorting to passages through cell culture, genetic engineering, laboratory accidents or deliberate creation of a human-pathogenic virus. The very same natural process occurred for instance in the chikungunya virus with the A226V mutation in the envelop gene which created a variant with modified vector specificity and increased epidemic potential (Tsetsarkin and Weaver, 2011; Kumar et al., 2008). The secondary accident is of a totally different nature and corresponds to the stochastic occurrence of a conjunction of events leading to the epidemic (Frutos et al., 2020b). Since circulating viruses do not generate an incidence high enough to reach the outbreak threshold they must go through an amplification stage to go above that threshold. Until now all events are stochastic whereas the outbreak threshold is a determinist event. In this stochastic phase, the drift is important and the infection can go extinct simply by chance (Hartfield and Alizon, 2013). This stochastic phase corresponds to the latency phase or stuttering phase during which only sporadic and uncharacterized cases can be observed. Furthermore, the pathogen population growth even during an epidemic does not follow an exponential growth but a polynomial growth (Chowell et al., 2016), making it unlikely to reach the outbreak threshold from this stochastic latency phase. An amplification loop must take place to provide the necessary population growth. This amplification loop is provided by societal events (Frutos et al., 2020b). It is the accidental conjunction of such events which provides the high concentration of hosts needed to obtain a polynomial growth of the already mutated viral population high enough to reach the outbreak threshold and triggers the epidemic. In the case of COVID-19, this amplifying loop was the conjunction of major celebrations during the new year period in Wuhan combined with a high concentration and mobility of people which made this amplification loop (Frutos et al., 2020a). This accidental nature of such conjunctions of natural and societal events explains why pandemic are so rare.

4.3. The circulation model: a change of paradigm

The search for the origin of SARS-CoV-2, following the assumptions from the spillover model is focusing on wildlife as the obvious source of the zoonosis. Teams worldwide are therefore desperately searching for the animal reservoir and for the virus similar to the one having emerged. However, no epizootic, no animal reservoir and SARS-CoV-2 virus have ever been identified. Incidentally, this failure in identifying the virus and the reservoir species in the natural environment facilitated the development of conspiracy theories linking SARS-CoV-2 to genetic engineering. More importantly, the spillover model leads to the paradigm in prevention that viruses have to be identified in the wild along with the intermediate species in order to prevent pandemics. The consequence is often the useless culling and slaughter of animals accused of being responsible. It happened with the culling of civets in China during the SARS epidemic (Watts, 2004b) or recently during the COVID-19 crisis, with the slaughter of minks in mass rearing in Denmark (Enserink, 2020; Oreshkova et al., 2020; Frutos et al., 2020a, 2020b). None of the prerequisites of the spillover model have been verified and this defense strategy is misleading by focusing on the wrong segment and wrong dynamic of pathogen transmission, leaving thus humanity vulnerable to further epidemics and pandemics. With a growing human population and an ever growing impact on the environment we can expect other infectious diseases to emerge and other pandemics to occur, simply because the probability of occurrence of such events is increasing. There is a need for a change of paradigm in defense strategy. What the circulation model, developed from actual observations, tells us is that nothing can be done at the level of the sylvatic cycle to prevent epidemics or pandemics. Furthermore, the current COVID-19 crisis shows that reactions after the appearance of the pandemic cannot stop it. It is too late. The circulation model tells us that the focus should be on the human activities making the amplification loops leading the growth allowing the virus to go over the outbreak threshold (Frutos et al., 2020a, 2020b). Targeting these human activities is essential because owing to their nature, they can be modelled, analyzed, managed, controlled and regulated in order to prevent pathogens to reach the outbreak threshold and to give rise to epidemic. Furthermore, whatever the virus which may emerge from the wild, it will have to through these routes and amplification loops shaped by human activities to reach the outbreak threshold. Regulating and controlling the human activities making these amplification loops will thus allow to prevent any emergence of viral diseases with no need to identify the viruses. Owing to the accidental nature of the emergence of a disease, i.e. a unique combination of biological processes (quasispecies evolution and mutations) and anthropogenic factors (amplification loops), an emerging disease cannot be predicted. However, there is no need for prediction. The need is to redirect efforts on the societal dimension of the dynamic of disease emergence to control key factors involved in the society-driven amplification loops (Frutos et al., 2020a; Dykstra et al., 2020).

Authors' contributions

All authors contributed to the manuscript. All authors participated in the writing and correction of the manuscript. All authors read and agreed with the manuscript.

Ethics declarations

No human samples or clinical data were used.

Funding

The only funding sources involved are the institutions of affiliation of each author: CIRAD (Centre de Coopération Internationale en Recherche Agronomique pour le Développement) for RF, University of Montpellier for LG, and CNRS (Centre National de la Recherche Scientifique) and Mediterranee-infection for CAD. The involvement of the funding sources was limited to the salaries of the authors, with no other role or involvement.

Declaration of Competing Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.meegid.2021.104812.

References

Afelt, A., et al., 2018. Bats, coronaviruses, and deforestation: toward the emergence of novel infectious diseases? Front. Microbiol. 9, 702.

Andersen, K.G., Rambaut, A., Lipkin, W.I., Holmes, E.C., Garry, R.F., 2020. The proximal origin of SARS-CoV-2. Nat. Med. 26, 450–452.

Andino, R., Domingo, E., 2015. Viral quasispecies. Virology 479, 46–51.

Becker, M.M., et al., 2008. Synthetic recombinant bat SARS-like coronavirus is infectious in cultured cells and in mice. Proc. Natl. Acad. Sci. U. S. A. 105, 19944–19949.

Boni, M.F., et al., 2020. Evolutionary origins of the SARS-CoV-2 sarbecovirus lineage responsible for the COVID-19 pandemic. Nat. Microbiol. 5, 1408–1417.

Borremans, B., Faust, C., Manlove, K.R., Sokolow, S.H., Lloyd-Smith, J.O., 2019. Cross-species pathogen spillover across ecosystem boundaries: mechanisms and theory. Philos. Trans. R. Soc. B 374, 20180344.

Briese, T., et al., 2014. Middle East respiratory syndrome coronavirus quasispecies that include homologues of human isolates revealed through whole-genome analysis and virus cultured from dromedary camels in Saudi Arabia. MBio 5, e01146-14.

Calisher, C., et al., 2020. Statement in support of the scientists, public health professionals, and medical professionals of China combatting COVID-19. Lancet 395, e42–e3.

Callaway, E.D., Cyranoski, D., 2020. Why snakes probably aren't spreading the new China virus. Nature 577, 1.

Carroll, D., et al., 2018. The global virome project. Science 359, 872–874.

Chaudhry, M.Z., et al., 2020. SARS-CoV-2 quasispecies mediate rapid virus evolution and adaptation. bioRxiv, 2020.08.10.241414.

Chen, Y., et al., 2017. Analysis of the codon usage pattern in Middle East Respiratory Syndrome Coronavirus. Oncotarget 8, 110337.

Chen, N., et al., 2020. How related is SARS-CoV-2 to other coronaviruses? Vet. Rec. 186, 496.

Chowell, G., et al., 2016. Mathematical models to characterize early epidemic growth: a review. Phys Life Rev 18, 66–97.

Cohen, J., 2020. Wuhan coronavirus hunter Shi Zhengli speaks out. Science 369, 487–488.

Coutard, B., et al., 2020. The spike glycoprotein of the new coronavirus 2019-nCoV contains a furin-like cleavage site absent in CoV of the same clade. Antivir. Res. 176, 104742.

Cross, P.C., et al., 2019. Confronting models with data: the challenges of estimating disease spillover. Philos. Trans. R. Soc. B 374, 20180435.

Daszak, P., Cunningham, A.A., Hyatt, A.D., 2000. Emerging infectious diseases of wildlife-threats to biodiversity and human health. Science 287, 443–449.

Devaux, C.A., 2012. Emerging and re-emerging viruses: a global challenge illustrated by Chikungunya virus outbreaks. World J. Virol. 1, 1.

Dimitrov, D.S., 2004. Virus entry: molecular mechanisms and biomedical applications. Nat. Rev. Microbiol. 2, 109–122.

Dyer, O., 2021. Covid-19: WHO says laboratory escape theory is "extremely unlikely" after mission to China. BMJ 372.

Dykstra, P., et al., 2020. Improving pandemic preparedness and management: Lessons learned and ways forward: independent expert report. Expert Rep. Eur. Commiss.

Enserink, M., 2020. Coronavirus rips through Dutch mink farms, triggering culls. Science 368, 1169.

Everington, K., 2020. Chinese Virologist Claims Coronavirus Derived From 'Zhoushan Bat Virus'.

Fowdy, T., 2020. The COVID-19 Whithleblower That Never Was.

Frutos, R., et al., 2020a. COVID-19: the conjunction of events leading to the coronavirus pandemic and lessons to learn for future threats. Front. Med. 7, 223.

Frutos, R., et al., 2020b. COVID-19: time to exonerate the pangolin from the transmission of SARS-CoV-2 to humans. Infect. Genet. Evol. 84, 104493.

Frutos, R., et al., 2021. Emergence of bat-related betacoronaviruses: hazard and risks. Front. Microbiol. 12 (437).

Ge, X.-Y., et al., 2016. Coexistence of multiple coronaviruses in several bat colonies in an abandoned mineshaft. Virol. Sin. 31, 31–40.

Getz, W.M., Dougherty, E.R., 2018. Discrete stochastic analogs of Erlang epidemic models. J. Biol. Dyn. 12, 16–38.

Halfmann, P.J., et al., 2020. Transmission of SARS-CoV-2 in domestic cats. N. Engl. J. Med. 383, 592–624.

Hao, P., 2020. Is SARS-CoV-2 originated from laboratory? A rebuttal to the claim of formation via laboratory recombination. Emerg. Microb. Infect. 9, 545–547.

Hartfield, M., Alizon, S., 2013. Introducing the outbreak threshold in epidemiology. PLoS Pathogens 9, e1003277.

Heymann, D.L., et al., 2004. Dangerous pathogens in the laboratory: from smallpox to today's SARS setbacks and tomorrow's polio-free world. Lancet 363, 1566–1568.

Hou, W., 2020. Characterization of codon usage pattern in SARS-CoV-2. Virol. J. 17, 1–10.

Huang, I.C., et al., 2006. SARS coronavirus, but not human coronavirus NL63, utilizes cathepsin L to infect ACE2-expressing cells. J. Biol. Chem. 281, 3198–3203.

Huang, C., et al., 2020. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet 395, 497–506.

Hueffer, K., Khatri, S., Rideout, S., Harris, M.B., Papke, R.L., Stokes, C., Schulte, M.K., 2017. Rabies virus modifies host behaviour through a snake-toxin like region of its glycoprotein that inhibits neurotransmitter receptors in the CNS. Sci. Rep. 7, 1–8.

Hul, V., et al., 2021. A novel SARS-CoV-2 related coronavirus in bats from Cambodia. BioRxiv.

Ith, S., et al., 2015. Taxonomic implications of geographical variation in *Rhinolophus affinis* (Chiroptera: Rhinolophidae) in mainland Southeast Asia. Zool. Stud. 54, e31.

Ji, W., et al., 2020. Cross-species transmission of the newly identified coronavirus 2019-nCoV. J. Med. Virol. 92, 433–440.

Johnson, B.A., et al., 2020. Furin cleavage site is key to SARS-CoV-2 pathogenesis. bioRxiv (2020.08.26.268854v1).

Jonas, O., Seifman, R., 2019. Do we need a Global Virome Project? Lancet Glob. Health 7, e1314–e1316.

Karamitros, T., et al., 2020. SARS-CoV-2 exhibits intra-host genomic plasticity and low-frequency polymorphic quasispecies. J. Clin. Virol. 131, 104585.

Korber, B., et al., 2020. Tracking changes in SARS-CoV-2 spike: evidence that D614G increases infectivity of the COVID-19 virus. Cell. 182 (812–27), e19.

Krieg, A.M., et al., 1995. CpG motifs in bacterial DNA trigger direct B-cell activation. Nature 374, 546–549.

Kucharski, A.J., et al., 2020. Early dynamics of transmission and control of COVID-19: a mathematical modelling study. Lancet Infect. Dis. 20, 553–558.

Kumar, N.P., et al., 2008. A226V mutation in virus during the 2007 chikungunya outbreak in Kerala, India. J. Gen. Virol. 89, 1945–1948.

Lai, A., 2020. Early phylogenetic estimate of the effective reproduction number of SARS-CoV-2. J. Med. Virol. 92, 675–679.

Lam, T.T., et al., 2020. Identifying SARS-CoV-2-related coronaviruses in Malayan pangolins. Nature. 583, 282–285.

Lee, J., Hughes, T., Lee, M.H., Field, H., Rovie-Ryan, J.J., Sitam, F.T., Sipangkui, S., Nathan, S.K.S.S., Ramirez, D., Kumar, S.V., Lasimbang, H., Epstein, J.H., Daszak, P., 2020. No evidence of coronaviruses or other potentially zoonotic viruses in Sunda pangolins (*Manis javanica*) entering the wildlife trade via Malaysia. Ecohealth 17, 406–418.

Leung, K., et al., 2021. Early transmissibility assessment of the N501Y mutant strains of SARS-CoV-2 in the United Kingdom, October to November 2020. Eurosurveillance 26, 2002106.

Li, X., et al., 2020. Evolutionary history, potential intermediate animal host, and cross-species analyses of SARS-CoV-2. J. Med. Virol. 92, 602–611.

Liu, P., et al., 2019. Viral metagenomics revealed sendai virus and coronavirus infection of Malayan Pangolins (*Manis javanica*). Viruses 11, 979.

Liu, Z., et al., 2020a. Composition and divergence of coronavirus spike proteins and host ACE2 receptors predict potential intermediate hosts of SARS-CoV-2. J. Med. Virol. 92, 595–601.

Liu, P., et al., 2020b. Are pangolins the intermediate host of the 2019 novel coronavirus (SARS-CoV-2)? PLoS Pathog. 16, e1008421.

Liu, S.L., et al., 2020c. No credible evidence supporting claims of the laboratory engineering of SARS-CoV-2. Emerg. Microb. Infect. 9, 505–507.

Lo Iacono, G., et al., 2016. A unified framework for the infection dynamics of zoonotic spillover and spread. PLoS Neglect. Trop. Dis. 10, e0004957.

Lu, H., et al., 2020. Outbreak of pneumonia of unknown etiology in Wuhan, China: the mystery and the miracle. J. Med. Virol. 92, 401–402.

Luan, J., et al., 2020. Spike protein recognition of mammalian ACE2 predicts the host range and an optimized ACE2 for SARS-CoV-2 infection. Biochem. Biophys. Res. Commun. 526, 165–169.

Matyasek, R., Kovarik, A., 2020. Mutation patterns of human SARS-CoV-2 and bat RaTG13 coronavirus genomes are strongly biased towards C>U transitions, indicating rapid evolution in their hosts. Genes 11, 761.

Menachery, V.D., et al., 2020. Trypsin treatment unlocks barrier for zoonotic bat coronavirus infection. J. Virol. 94, e01774-19.

Oreshkova, N., et al., 2020. SARS-CoV-2 infection in farmed minks, the Netherlands, April and May 2020. Euro Surveill. 25 (pii=2001005).

Othman, H., et al., 2020. Interaction of the spike protein RBD from SARS-CoV-2 with ACE2: similarity with SARS-CoV, hot-spot analysis and effect of the receptor polymorphism. Biochem. Biophys. Res. Commun. 527, 702–708.

Paraskevis, D., et al., 2020. Full-genome evolutionary analysis of the novel corona virus (2019-nCoV) rejects the hypothesis of emergence as a result of a recent recombination event. Infect. Genet. Evol. 79, 104212.

Perez, J.C., Montagnier, L., 2020. COVID-19, SARS and bats coronaviruses genomes peculiar homologous RNA sequences. Int. J. Res. 8, 217–263.

Plowright, R.K., et al., 2017. Pathways to zoonotic spillover. Nat. Rev. Microbiol. 15, 502–510.

Poulin, R, 2010. Parasite manipulation of host behavior: an update and frequently asked questions. Adv. Study Behav. 41, 151–186.

Power, A.G., Mitchell, C.E., 2004. Pathogen spillover in disease epidemics. Am. Nat. 164, S79–S89.

Reuter, K.E., 2016. Using stable isotopes to infer the impacts of habitat change on the diets and vertical stratification of frugivorous bats in Madagascar. PLoS One 11, e0153192.

Sallard, E., et al., 2020. Tracing the origins of SARS-COV-2 in coronavirus phylogenies. Med. Sci. 36, 783–796.

Shi, J., et al., 2020. Susceptibility of ferrets, cats, dogs, and other domesticated animals to SARS–coronavirus 2. Science 368, 1016–1020.

Song, H.D., et al., 2005. Cross-host evolution of severe acute respiratory syndrome coronavirus in palm civet and human. Proc. Natl. Acad. Sci. U. S. A. 102, 2430–2435.

Sutton, C., 2020. Doctor Who Fled China to Reveal Virus Truth.

Tang, X., et al., 2020. On the origin and continuing evolution of SARS-CoV-2. Natl. Sci. Rev. 7, 1012–1023.

Tsetsarkin, K.A., Weaver, S.C., 2011. Sequential adaptive mutations enhance efficient vector switching by Chikungunya virus and its epidemic emergence. PLoS Pathog. 7, e1002412.

Vega, T., B., et al., 2013. Influenza surveillance in Europe: establishing epidemic thresholds by the moving epidemic method. Influenza Other Respir. Viruses 7, 546–558.

Wacharapluesadee, S., et al., 2021. Evidence for SARS-CoV-2 related coronaviruses circulating in bats and pangolins in Southeast Asia. Nat. Commun. 12, 972.

Watts, J., 2004a. China culls wild animals to prevent new SARS threat. Lancet 363, 134.

Watts, J., 2004b. SARS under control, but lab-safety questions remain. Lancet 363, 1780.

Webster, R.G., 2004. Wet markets-a continuing source of severe acute respiratory syndrome and influenza? Lancet 363, 234–236.

Webster, R.G., et al., 1982. Molecular mechanisms of variation in influenza viruses. Nature 296, 115–121.

Wenzel, J., 2020. Origins of SARS-CoV-1 and SARS-CoV-2 are often poorly explored in leading publications. Cladistics 36, 374–379.

WHO, 2004. China Confirms SARS Infection in Another Previously Reported Case (accessed October 19, 2020).

WHO, 2020. Novel Coronavirus (2019-nCoV) Situation Report-1.

Wong, Y.C., et al., 2020. Natural transmission of bat-like severe acute respiratory syndrome coronavirus 2 without proline-arginine-arginine-alanine variants in coronavirus disease 2019 patients. Clin. Infect. Dis. (ciaa953 2020).

Worobey, M., et al., 2016. 1970s and 'Patient 0' HIV-1 genomes illuminate early HIV/AIDS history in North America. Nature 539, 98–101.

Xiao, K., et al., 2020a. Isolation and characterization of 2019-nCoV-like coronavirus from Malayan Pangolins. bioRxiv (doi: 2020.02.17.951335).

Xiao, C., et al., 2020b. HIV-1 did not contribute to the 2019-nCoV genome. Emer. Microb. Infect. 9, 378–381.

Xu, D., Zhang, Z., Wang, F.S., 2004. SARS-associated coronavirus quasispecies in individual patients. N. Engl. J. Med. 350, 1366–1367.

Yamada, Y., Liu, D.X., 2009. Proteolytic activation of the spike protein at a novel RRRR/S motif is implicated in furin-dependent entry, syncytium formation, and infectivity of coronavirus infectious bronchitis virus in cultured cells. J. Virol. 83, 8744–8758.

Yuan, L., et al., 2017. A single mutation in the prM protein of Zika virus contributes to fetal microcephaly. Science 358, 933–936.

Zehender, G., et al., 2020. Genomic characterization and phylogenetic analysis of SARS-COV-2 in Italy. J. Med. Virol. 92, 1637–1640.

Zhan, S.H., 2020. SARS-CoV-2 is well adapted for humans. What does this mean for re-emergence? BioRxiv (doi: 2020.05.01.073262).

Zhang, Y.Z., Holmes, E.C., 2020. A genomic perspective on the origin and emergence of SARS-CoV-2. Cell 181, 223–227.

Zhang, C., et al., 2020a. Protein structure and sequence reanalysis of 2019-nCoV genome refutes snakes as its intermediate host and the unique similarity between its spike protein insertions and HIV-1. J. Proteome Res. 19, 1351–1360.

Zhang, T., et al., 2020b. Probable pangolin origin of SARS-CoV-2 associated with the COVID-19 outbreak. Curr. Biol. 30, 1346–1351 (e2).

Zheng, B.J., et al., 2004. SARS-related virus predating SARS outbreak, Hong Kong. Emerg. Infect. Dis. 10, 176–178.

Zhou, P., et al., 2020a. A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature 579, 270–273.

Zhou, H., 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).

Zhu, N., et al., 2020. A novel coronavirus from patients with pneumonia in China, 2019. N. Engl. J. Med. 382, 727–733.