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.

Biochemical and Biophysical Research Communications 538 (2021) 14–23

Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

journal homepage: www.elsevier.com/locate/ybbrc

History of the COVID-19 pandemic: Origin, explosion, worldwide spreading

Sara Platto [a, b], Yanqing Wang [b], Jinfeng Zhou [b], Ernesto Carafoli [c, *]

[a] Department of Biotechnology, College of Life Sciences, Jianghan University, Wuhan, China
[b] China Biodiversity Conservation and Green Development Foundation, Beijing, China
[c] Venetian Institute of Molecular Medicine, University of Padova, Italy

ARTICLE INFO

Article history:
Received 14 October 2020
Accepted 20 October 2020
Available online 6 November 2020

Keywords:
Adenosine analogs
Antiviral therapy
Remdesivir
RNA dependent RNA polymerase

ABSTRACT

The SARS-CoV-2 virus of the COVID-19 pandemic, that is presently devastating the entire world, had been active well before January of this year, when its pathogenic potential exploded full force in Wuhan. It had caused the onset of small disease outbreaks in China, and probably elsewhere as well, which failed to reach epidemic potential. The distant general origin of its zoonosis can be traced back to the ecosystem changes that have decreased biodiversity, greatly facilitating the contacts between humans and the animal reservoirs that carry pathogens, including SARS-CoV-2. These reservoirs are the bats. The transition between the limited outbreaks that had occurred through 2019 and the epidemic explosion of December–January was made possible by the great amplification of the general negative conditions that had caused the preceding small outbreaks. In the light of what we have now learned, the explosion was predictable, and could have happened wherever the conditions that had allowed it, could be duplicated. What could not have been predicted was the second transition, from epidemic to pandemic. Research has now revealed that the globalization of the infection appears to have been caused by a mutation in the spike protein of the SARS-CoV-2, that has dramatically increased its transmissibility.

1. Introduction

Today, nine or ten months after the recognition that a new coronavirus had caused the sudden appearance of a large number of atypical pneumonias in Wuhan, the city is still largely considered the place of origin of the new viral disease. However, as documented in a recent contribution [1], the virus and its pathologies had been present in China well ahead of the Wuhan epidemic explosion. They had actually been active outside China as well, for instance in Europe [2]. Somehow, however, for many months before Wuhan the new virus, now known as SARS-CoV-2, had remained dormant, as it had not encountered conditions that would have led to its epidemic awakening. It had somehow still managed to cause some small outbreaks, but they had remained limited in space and time. Gradually, however, in the last decades of past century conditions have matured that have favored the chances of all sorts of pathogens, including viruses like SARS-CoV-2, to become dangerous and even lethal to animals and humans. These conditions have been described in detail in an accompanying contribution in this Special Issue [3]. They are essentially related to the loss of ecosystem biodiversity caused by the conversion of natural habitats to agricultural and/or urban ecosystems, that have increased the contacts between humans and wildlife, and among it, reservoirs of potential zoonoses. Biodiversity reduces pathogen transmission by a "*dilution effect*", that acts when in an ecosystem - the large variety of species present "buffers" the persistence of pathogens. Human activities like agricultural practices, hunting, uncontrolled urbanization, have disturbed the ecosystems, interfering with the *dilution effect,* decreasing it, and increasing risks of pathogen of all types, including viruses to spill lover to humans.

These considerations on biodiversity have general character, and apply to all sort of zoonoses. In the case of COVID-19, the ecosystem alteration that has favored its grand scale awakening as a dangerous human pathogen has used bats as essential mediators. Bats are the second most numerous mammals after the rodents. They have long been known as reservoirs of coronaviruses, and host more zoonotic viruses than any other mammalian order [4]. They adapt easily to anthropized environments, and the anthropogenic

* Corresponding author.
E-mail addresses: saraplatto@cbcgdf.org (S. Platto), v52@cbcgdf.org (Y. Wang), jz@cbcgdf.org (J. Zhou), ernestocarafoli@gmail.com (E. Carafoli).

https://doi.org/10.1016/j.bbrc.2020.10.087
0006-291X/

ecosystem changes mentioned above have greatly increased their promiscuity with domestic animals and humans. Evolution has gradually shaped their metabolism and their immunological system to make them insensitive to the dangerous weapons of pathogenic coronaviruses, making them important reservoirs of them [3]. Importantly, bats are also a locus for coronavirus evolution, as the recombination frequency of coronavirus is high. Virus variants can thus appear and be maintained in bats with properties that increase greatly their aggressiveness: a prominent example is the development of the ability of the virus to use the ACE2 receptor as an entry path into host cells.

2. Variants of the SARS-CoV-2 and their geographical distribution

SARS-CoV-2 belongs to the *Coronaviridae* subfamily (CoV) of the *Coronavirinae* family of RNA viruses. The subfamily includes alpha, beta, gamma, and delta coronaviruses. Seven beta coronaviruses most commonly infect humans: apart from SARS-CoV-2, two more have caused serious diseases: SARS-CoV which was responsible for the SARS outbreak of 2002–2003, and MERS, which caused the Middle East MERS CoV outbreak of 2012. HCUI, NL63, OC43, and 229E have instead only caused mild disease symptoms.

The number of infected cases and the mortality rates related to COVID-19 vary in different geographical areas. Phylogenetic network analysis on 160 complete SARS-CoV-2 genomes [5] has revealed 3 central variants (A, B, and C) distinguished by amino acid differences. Variant A is the closest to that discovered in bats, that had 96.2% sequence similarity to the human virus [6], and was thus considered the original human virus genome. Although present in Wuhan, where the explosion of the pandemic had taken place, in the city it was not the predominant genome. Large number of A genomes were instead found in Americans who had lived in Wuhan and in the USA and Australia. The predominant genome in Wuhan and East Asia was variant B, which is separated from variant A by two mutations: the finding that it did not travel much beyond China and South East Asia without further mutations, implies a founder event in Wuhan and/or East Asia. Variant C differs from its parent variant B by a G-V change: it is the major European type, initially found in patients from France, Italy, England, and Sweden. It is absent from mainland China, but has been found in Singapore, Hong Kong, Taiwan and South Korea: the earliest introductions of the variant in Italy actually occurred from a first documented German case in January 27, 2020, and from the Singapore cluster. The German infection had occurred from an employee of the Webasto Company in Munich, who had contracted the disease from a Chinese colleague in Shanghai, who had in turn received a visit by her Wuhan parents (from Italy the variant had then spread to Brazil and Mexico by people who had visited Italy): during its travel from its origin in Wuhan to Italy, and to Brazil and Mexico the virus had undergone 10 mutations [5].

3. SARS-CoV-2: attack of the host cell

Coronaviruses contain the largest known RNA genomes (26–32 kilobases). In SARS-CoV-2, the RNA genome codes for four structural and 16 non-structural proteins (nps). The structural proteins are the S, M, N, and E proteins. The E and M proteins form the envelope and the membrane that coat the virus, the N protein binds the RNA genome, and the S (Spike) protein, which forms the characteristic protrusions from the surface of the virus, interacts with the plasma membrane receptor of the target cell to mediate the penetration of the virus into it.

During the SARS-CoV 2002–2003 outbreak the receptor for the S-protein was identified as the transmembrane metallocarboxypeptidase angiotensin converting enzyme ACE2 [7]. The identification was based on a number of findings: ACE2 was isolated from SARS-CoV −permissive VeroE6 cells, and found to bind efficiently the S1 segment of the virus S-protein. The virus replicated efficiently in 293T cells transfected with ACE2, but not in mock-transfected cells, and, finally, anti ACE2 antibodies inhibited the replication of the virus in VeroE6 cells. ACE2 was later identified as the receptor for COVID-19 as well by showing that it used ACE2 to gain entry into HeLa cells from humans and a number of animals transfected with ACE2 [6,8,9] (interestingly, the betacoronavirus responsible for the MERS epidemic of 2012 used instead dipeptidyl peptidase 4 (DPP4) as receptor that mediated its entry into the target cells [10].

In a simplified form, the sequence of events that follow the invasion of the cell by the virus by endocytosis can be summarized as follows: the endocytosed virus moves from early endosomes to late endosomes. It eventually ends up into the lysosomes, where their hydrolytic enzymes are activated by the low lysosomal pH to uncoat the viral genome. Once liberated in the cytoplasm, the RNA genome joins the rough endoplasmic reticulum of the host cell where its open reading frames *ORF 1a* and *ORF 1b* are translated into *pp1a* and *pp1ab* proteins, which are then cleaved by cell proteases to yield the 16 nsps. Some of these form the replicase/transcriptase complex: *nsp 12* is the RNA-dependent RNA polymerase that will produce the new viral genome. To function properly, the polymerase needs the assistance of *nsp7* and *nsp8* and, possibly, of other non-structural proteins: the triad *nsp12, nsp7, and nsp 8* is the minimal complex necessary to polymerize the RNA. The newly synthesized RNA is then translated into the four structural proteins, which, after suitable modification of the nascent S-protein in the host cell endoplasmic reticulum (ER) and the Golgi complex with N-linked glycans (presumably to protect it against neutralizing antibodies), form the mature new virus particle. The exocytosis vesicle that contains it travels to the surface of the host cells to be eventually excreted. In the end, one virus unit that penetrates the host cell yields a myriad of new viruses that propagate the infection to neighbouring cells, the preferred targets of the excreted viruses being macrophages and T-lymphocytes. The N structural protein of the virus promotes the synthesis of Interleukin 6 (IL6) by binding to the promoter region of the IL6 gene of target cells. IL6 is the main actor in the *"cytokine storm"* triggered by the coronavirus [11].

The S protein deserves a more detailed discussion, as it is responsible for the attack of the target cells by the virus.It contains a signal peptide and the receptor binding domain (RBD) [12,13] that will recognize the receptor of the host cell and interact with it (Fig. 1). Following binding to the receptor, the S protein is cleaved by host proteases into S1 and S2 segments, which do not separate: they remain bound non-covalently, with the C-terminal S2 segment anchoring the protein to the viral membrane. The cleavage of the S protein is performed by a number of proteases of the host cell: both endosomal cathepsins B and L, furin, and the transmembrane serine protease TMPRSS2 can perform the cleavage, but TMPRSS2 appears to predominate [14]. it has been suggested that furin could prepare the S-protein for the cleavage by TMPRSS2. The cleavage is essential to "prime" the S -protein [15], as it exposes a peptide within the S2 segment that promotes the fusion of the S protein with the plama membrane of the host cell and the eventual internalization of the virus into it.

Endocytosis is generally assumed to be the classical mechanism by which the SARS-CoV-2 virus invades target cells. However, coronaviruses, including SARS-CoV-2, may also penetrate into cells by an endocytosis-independent mechanism: which, incidentally, is a mechanism used by a number of other viruses. The RBD of the S-protein recognizes the ACE2 receptor, but another domain of the5S-protein interacts with a close-by area of the cell surface that, in a

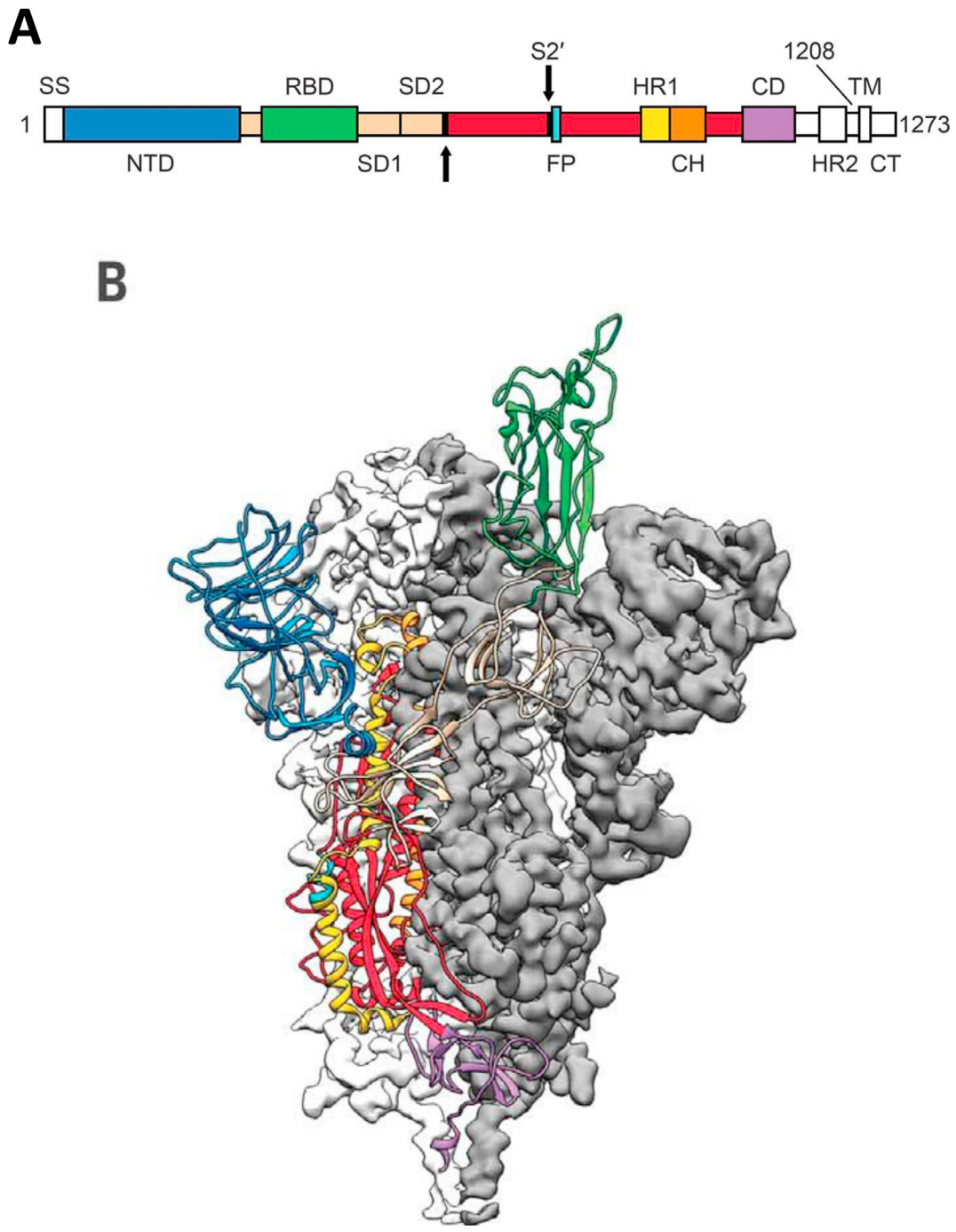

Fig. 1. Structure of the trimeric Spike protein of COVID-19. A: primary structure colored by domain. SS, signal sequence; S2′, protease cleavage site; FP, fusion peptide; HR1, heptad repeat 1; CH, central helix; CD, connector domain; HRE, heptad repeat 2; TM, transmembrane domain; CT, cytoplasmic tail; RDB, receptor binding domain; arrows indicate protease cleavage sites. B side and top views of the structure of the protein with a single RBD in the up conformation. The two RBD down protomers are shown in white or gray. The RBD protomer is shown in ribbons colored as shown in A. Modified from D Wrapp et al., 2020 [13].

sense, could act functionally as a co-receptor: recent work by Fantini and coworkers [16] has identified this domain of the plasma membrane of the host cell as a ganglioside-rich area that would interact with the N-terminal domain (NTD) of the S-protein. The virus would fuse with the plasma membrane of the target cell, leaving most of its components in it, an letting into the cytosol only the RNA genome bound to the N-nucleoprotein: which would travel with the RNA genome to the ER ribosomes independently of the lysosomal pathway.

4. The spike-protein: SARS-CoV and SARS-CoV-2

As mentioned, the S-protein recognizes the cell receptor, interacts with it, and undergoes the proteolytic processing that is essential for its fusion with the plasma membrane of the host cell and for its eventual penetration into it. A comparison of the S-proteins of the SARS-CoV and of SARS-CoV-2 may provide clues on the greater pathogenicity of the latter, which spread to pandemic proportions, whereas SARS-CoV) didn't. Of the total 380 amino

acids that differ between SARS-CoV and SARS-CoV-2, 27 belong to the S-proteins [16]. Six of the 60 that form the RBD of the S-protein of SARS-CoV-2 (Y442, L472, N479, D480, T487, and Y4911) are critical for the interaction with the ACE2 receptor, and five of them differ from the corresponding residues of SARS-CoV (the RBD is the most variable part of the protein) [17]. The amino acids of SARS-CoV-2 allow increased hydrophobic interactions and salt bridge formations, and stabilize binding hotspots at the RBD-ACE2 interface: the RBD of SARS-CoV-2 forms a larger binding interface with ACE2, and establish more contacts with it. These changes may be expected to increase the affinity of the RBD of the S-protein of SARS-CoV-2 for the ACE2receptor [18,19] which has indeed been found to by higher than that of the SARS-CoV. Surface plasmon resonance measurements have revealed that ACE2 binds to the S-protein of SARS-CoV-2 with a K_d of about 15 nM, which is 15–20 folds higher than that of the binding to SARS-CoV [20].

Another difference between the S-proteins of SARS-CoV-2 and SARS-CoV is the cleavage site at the S1–S2 junction [17,21]. A distinctive feature of the S-protein of SARS-CoV-2 is the four amino acid insert PRRA at the S1/S2 junction (Fig. 2). This is a potential cleavage site for the protease furin, and is normally referred to as a "*polybasic site*" (21): furin has been mentioned above as a protease that could "prepare" the S-protein for the cleavage by the trypsin-like protease TMPRSS2. As seen in Fig. 2, the site is not present in SARS-CoV, and furin, as expected, indeed does not cleave SARS-CoV [21] (Fig. 3). SARS-CoV-2 has evolved the ability to utilize a large number of host proteases, from cathepsins B and L, to PC, to trypsin. They cleave the S-protein of SARS-CoV-2 much more efficiently than the S-protein of SARS-CoV the only difference being cathepsin L, that instead cleaves SARS-CoV S-protein more efficiently (Fig. 3),. It could also be mentioned that the leading proline of the SARS-CoV-2 insert has been predicted to induce a turn that may facilitate the addition of O-glycans to S and T residues that flank the cleavage site. They have been suggested to positively influence virulence but it has not been demonstrated that the glycosylation site is utilized.

The structural and functional differences in the S-proteins of SARS-CoV and SARS-CoV-2 are evident and significant: clearly, the SARS_CoV attacks the target cell less efficiently than SARS-CoV-2. It seems reasonable to suggest that the differences have been responsible for the failure of SARS-CoV to reach pandemic dimensions.

4.1. Spike protein: the D614G mutation

Mutations in the S-protein are rare: the GISAID SARS-CoV2 sequence database, which is the primary SARS-CoV-2 sequence database resource with thousands of sequences linked to geographical distribution, includes 14 positions of interest in which the S-protein mutations occur, for instance, the RBD. The sampling dates are important, as they would offer indications for positive selection of a mutation by showing shifts (i.e. increases) in its frequency over time. A number of mutations have been detected in various domains of the protein, but the one that has attracted most of the attention, and which has important documented effects on the function of the protein is the D614G mutation: it was first sampled in Germany on January 28, 2020, and 4 other cases were sampled at the same time, or soon thereafter, in China: these early Chinese sequences were highly related to the German sequence, thus, the mutation may have originated either in China or Europe, since it was present in both areas at the end of January. The striking aspect of the mutation was its dramatic and rapid frequency increase with time: from 7 out of 183 sequences in the data base at the beginning of March, it increased to 29% in mid-March and proceeded to an alarmingly rapid increase in the next sampling [22]. At the end of January nearly all the global deposited sequences were the "wild type" with a D in position 614, but three months later the global frequency of the mutated variant had exceeded 50%, and was even higher in Europe (Fig. 4): its rapid increase indeed occurred predominantly in Europe (in two Countries, Italy and France, virtually all sequences, already in the early sampling, had the mutation) and then spread to the Americas and Oceania: wherever the mutation entered a region it rapidly became the dominant variant. By contrast, the frequency of the non-mutated variant remained relatively low in China and East Asia [22] (Fig. 4).

The D614G amino acid change is caused by an A-G nucleotide mutation at position 23,403 of the S- sequence, and it is almost always accompanied by three other mutations: a C-T mutation at position 241, a silent C-T mutation at position 3,037, and a C-T mutation at position 14,408 that results in a P323L change in the RNA polymerase.

The D614G mutation is associated with increased transmission, an effect that could have been inferred from the structural findings (Fig. 1) [19] that have shown that D614 is located on the surface of the trimeric protein. S1 could form contacts with a vicinal protomer through the side chain of D614 that could form a hydrogen bond with T859 of S2 of the neighbouring protomer. These, and other, structural considerations have been used to speculate on possible longer range effects of the mutation that could explain the increased infectivity. One point of discussion is whether the mutation would alter the release or shedding of the S1 domain after cleavage at the S1/S2 junction [22,23], or whether the shedding would instead be limited by the stabilization of the interaction between the S1 and S2 domains [24]. It has been speculated [22] that the D614G mutation would promote shedding of the S1 domain due to the loss of the proposed hydrogen between that D614 in subunit S1 could form with T 859 in subunit S2. Alternatively [24] (Fig. 5), it has been suggested that Q 613 forms a hydrogen bond with T859 in S2. The introduction of a glycine in the vicinal 614 position would increase the backbone flexibility allowing a more favorable orientation of Q613. D614 could also form an intra-domain salt bridge with R646, promoting a local conformation of S1 that would be unfavorable to the association with S2. A G in position 614 would impede this unfavorable configuration. As said, the questioni s debated, however, the binding of ACE2 to model cells increases when the cells express the mutated S-protein, i.e., there is in them more S1 available for the association with ACE2.

One question that is still open is the reason why the higher transmissibility of viruses carrying the D614G mutation does not appear to be associated with a difference in disease severity. A number of possible factors related to the virus have been considered, but it should not be forgotten that in the interplay virus-host the latter is also an actor: who could be influenced by epidemiological and evolutionary pressures. The introduction of a founder into a highly mobile and well connected population [23] may play a significant role in the rapid spread of the infection without effects on its severity. After all, factors of this type (see above) have certainly contributed to the fact the several small outbreaks of COVID-19 infection in the last decades of last century have remained limited and controlled until a complicating event -in this case the "*amplification*" factor of Wuhan's wet market-came into play [1].

4.2. COVID-19: spillover from the bat reservoir to humans?

The matter of the intermediate host in the transmission of coronaviruses from bats to humans is not settled [1]. During the SARS-CoV 2002–2003 epidemic, civet cats (*Paguma larvata)* had been considered the reservoir of the virus, as it had been isolated from clinically normal animals [25,26]. Therefore, the Guangzhou

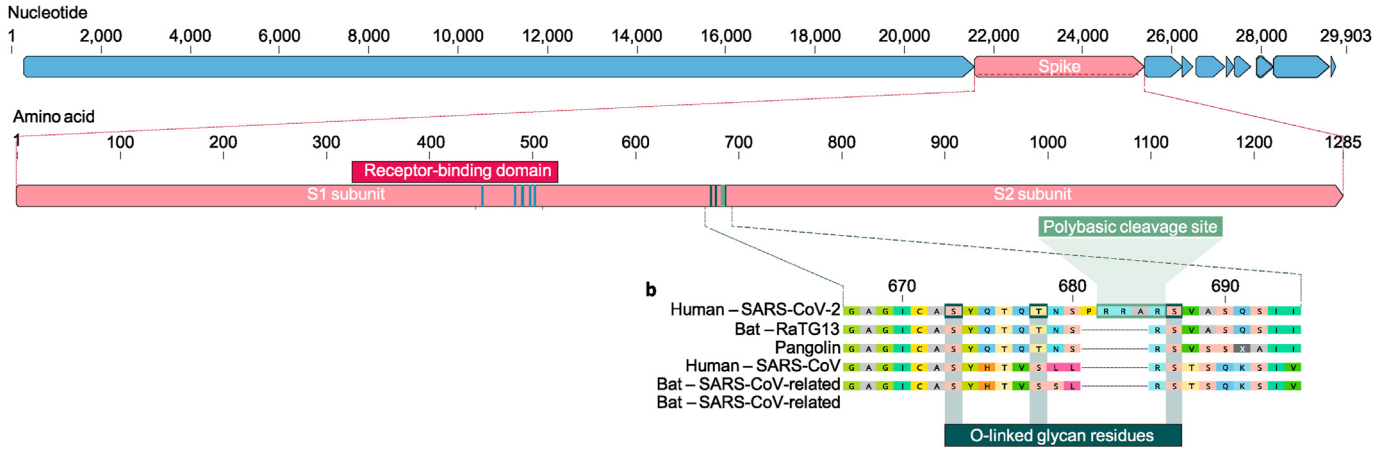

Fig. 2. Amino acid sequence of the Spike protein of human SARS-CoV-2 and related coronaviruses.The amino acid sequence of the Spike protein (red bar in A) is expanded in B, and shows a polybasic site (RRAR) at the junction of the two subunits S1and S2 of the protein. The alignment against related coronaviruses (C) shows that the polybasic site RRAR is unique to SARS-CoV-2. The three adjacent predicted O-linked glycans are also unique to SARS-CoV-2 Sequences shown are from NCBI GenBank, accession codes MN908947, MN996532, AY278741, KY417146and MK211376. The pangolin coronavirus sequences are a consensus generated from SRR10168377and SRR10168378(NCBI BioProject PRJNA573298. Modified from K. G. Andersen et al., 2020 [18] where additional details are found. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

Government ordered the destruction of a large number of civet cats in the wet-markets, an action that at the time was actually believed to have contained the re-emergence of SARS-CoV, as no further cases of the infection were detected. However, subsequent extensive epidemiology studies had failed to find SARS-CoV in farmed and wild-caught civet cats, indicating that other animal(s) might have been the natural virus reservoir of SARS-CoV [26,27]. It was also found that strains of SARS-CoV isolated from palm civets in 2002−2003 and 2005, had low affinity for human ACE2 receptors and low infectivity in human cells, but had high affinity for civet cat ACE2 receptors and high infectivity in civet cat cells [25],Therefore, there is a strong possibility that these animals were only incidental hosts for the SARS-CoV [1,3]. In fact, no intermediate host has yet been established for SARS-CoV [28].

In the case of SARS-CoV-2, Malayan pangolins (*Manis javanica*) were initially highly favored as the intermediate hosts. The genome of a new coronavirus in three sick pangolins from the illegal wildlife trade market was found to be highly related to that of SARS-CoV-2 [29], however, a later phylogenetic analysis and amino acid sequence on the S- protein of SARS-CoV-2 cast doubts on the hypothesis of pangolins as SARS-CoV-2 intermediate hosts [30].The doubts were reinforced by an analysis of Damas and coworkers [31] on the ACE2 receptors of 410 species of vertebrates (fishes, amphibians, birds, reptiles, and mammals) to assess their ability to bind SARS-CoV-2 spike proteins. The vertebrates were classified as having low, medium, and high score depending on the ability of their ACE2 receptors to bind SARS-CoV-2 Spike proteins [31]. The species that scored very low included all pangolin species such as Chinese pangolin (*Manis pentadactyla*), Sunda pangolin (*Manis javanica*), whitebellied pangolin (*Phataginus tricuspis*), and masked palm civet (*Paguma larvata*) [31] Another study on throat and rectal swabs of 334 pangolins (*Manis javanica* rescued from the wild in Peninsular Malaysia further exonerated the pangolin as reservoir or intermediate host for SARS-CoV-2, as no sample yielded a positive PCR result for coronaviruses [32]. Therefore, it seems very likely that the previous detection of CoV in pangolins might have been due to their exposure to infected humans, or other wildlife in the wildlife trade network [32]. Thus, also in the case of SARS-CoV-2 no intermediate host has thus so far been identified, and the possibility of a direct spillover of the virus from bats to humans seems a reasonable option. A recent report [1] has discussed the matter in some detail, also considering the possibility of “indirect” spillover through the food chain, i.e., via the contamination of food products, not just live animals. A study that supports the direct spillover is particularly interesting: Andersen and his coworkers [17] have hypothesized that the virus may have acquired the current genomic features through adaptation during undetected human-to-human transmissions. If the adaptation had occurred in a different animal, it would be logical to expect recurrences of the outbreaks, but if it had occurred in humans, the zoonotic transfers would demand the same series of adaptive mutations [17]. A time interval between different outbreaks would necessarily occur, to reflect the time needed by the virus to adapt from the reservoir to humans (17). This seems very unlikely, since SARS-CoV-2 adapted already very well to human hosts, with a very successful human-to-human transmission that allows the virus to be maintained within the population. Therefore, in the reasoning of Andersen and coworkers humans could be the new reservoirs of SARS-CoV-2 for their own species, eliminating the need of an additional spillover from animal to human to keep the infection going [1,17].

The considerations of the study by Andersen and coworkers are convincing, however, the matter of the intermediate host between bats and humans in the transfer of SARS-CoV-2 cannot be considered settled. Going back to the analysis of Damas and coworkers mentioned above on the ACE2 receptor in vertebrates [31] it can be noticed that it had found that the 18 species classified as “very high” were all Old World monkeys, among them macaques, which are the only known species that can be infected by SARS-CoV-2, with symptoms similar to those of humans. Macaques are widely traded for consumption in South East Asia, and are frequently found in markets where they are kept together with other animals, including bats. Considering macaques’ close phylogenetic relationship to humans, the possibility of the transfer of SARS-CoV-2 from bats to them as intermediate hosts for the spillover to humans seems plausible [3].

Finally, a recent report from Associated Press could be mentioned, as it is somehow related to the transmission of SARS-CoV-2: an outbreak of COVID-19 infection has killed at least 10,000 minks in Utah mink farms (https://www.cdc.gov/coronavirus/2019-ncov/daily-life-coping/animals.html). The susceptibility of minks to coronavirus was known, as the virus had been discovered in mink farms in the Nertherlands, and similar outbreaks had

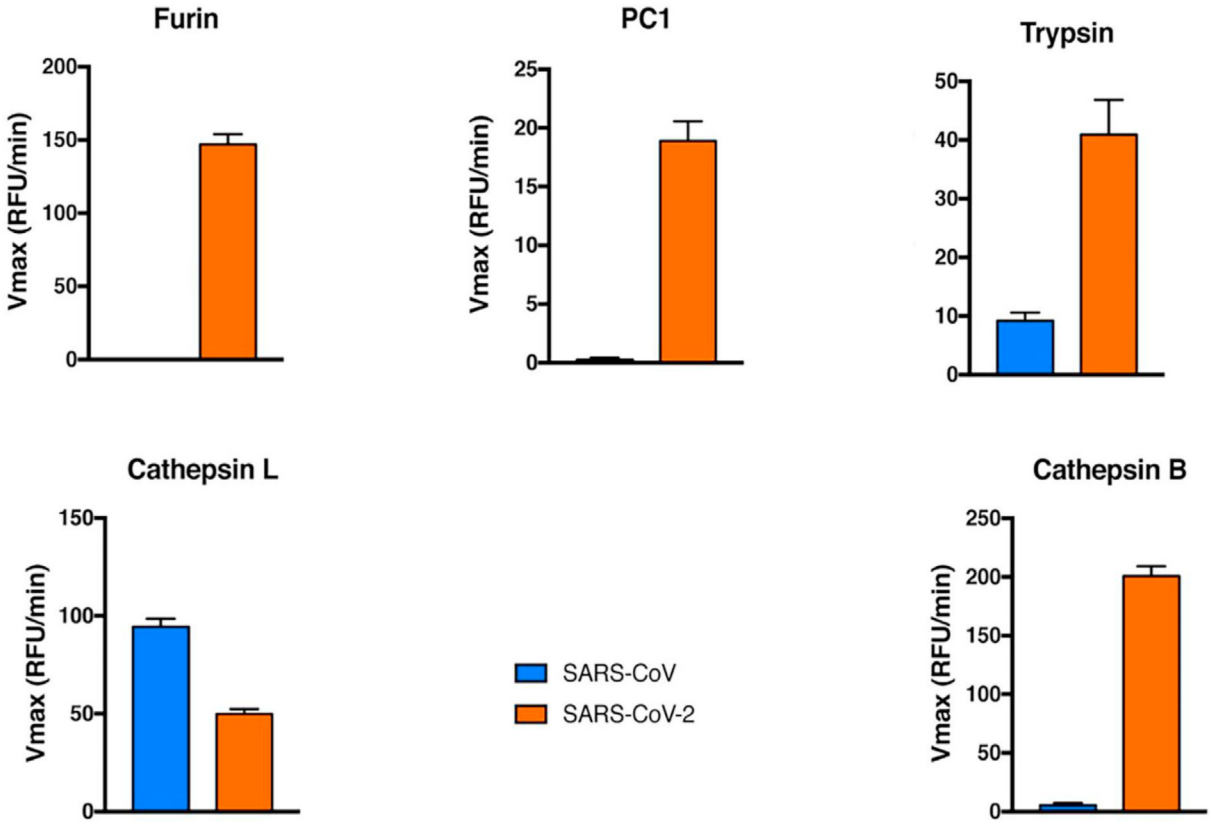

Fig. 3. Proteolytic cleavage of peptides containing the S1/S2 cleavage site of SARS-CoV-2 and SARS-CoV. A number of proteases are likely to be involved in Spike protein processing, furin and PC1, trypsin, type II transmembrane serine protease matripase (the transmembrane serine protease TMPRSS2 has not been tested), cathepsins L and B. Furin cleaves SARS-CoV-2, but not SARS-CoV. In addition to furin, other proteases cleave SARS-CoV-2 much more efficiently than SARS-CoV. The only protease that cleaves SARS-CoV more efficiently than SARS-CoV-2 is cathepsin L. Clearly, the acquisition of the polybasic site insert PRRA by COVID-17 broadens the repertoire of activating proteases. Modified from Jaimes et al. (21).

occurred elsewhere in the USA and in Europe. The USA Department of Agriculture said that there was no evidence of the transfer of the virus to humans, actually, people with COVID 19 could have spreaded the infection to the minks. However, the affected farms were quarantined to stop the spread, and stringent biosecurity measures were implemented.The United States Department of Agriculture has recommended that people with suspected or confirmed COVID 19 should avoid contact with pets or other animals to protect them from possible infection.

4.3. COVID-19: the pre-Wuhan period

When, years from now, the history of the COVID-19 will be written, three distinct phases will have to be distinguished: the origins, the localized explosion, and the final worldwide spreading. And there will be corollaries: the first, and most important, is to understand whether the transitions from one phase to the next was accidental or predictable. As mentioned at the beginning of this contribution, the idea that it all started in Wuhan at the beginning of this year is still widely repeated. Wuhan has certainly been the turning point in the history of the pandemic, however, only as the predictable and probably mandatory superspreader of what had been quietly boiling unnoticed, in China and elsewhere, long before,. At variance with the pre-Wuhan transition, the post Wuhan transition to pandemic wasn't instead mandatory nor was it predictable: it could as well not have occurred.

4.3.1. On the origins

After the SARS CoV epidemic of 2002–2003, a number of novel coronaviruses, which were termed SARS-like-CoV, were found in bats of the genus *Rhinolophus* from a number of Chinese provinces. Initially, they were unable to use the ACE2 receptors, and were thus not considered able to infect humans [30,33]. In 2013, various SARS-like-CoVs were discovered in a single colony of *Rhinolophus* bats in the Chinese province of Yunnan. Three of them, W1V1, W1V16, and RaTG13, had the ability to use ACE2 to enter human cells. Some individuals who had been in contact with bats, or had been exposed to bat secretions in their villages, had developed positivity to a SARS-like-CoV antibodies. However, these individuals did not recall particular symptoms in the months prior to the analysis. The three SARS-like-CoVs had about 95% identity to the SARS-CoV, and had adapted themselves to ACE2 receptors. Even if at that time SARS-CoV-2 was not yet known, ominous signs were evidently there that the potential for the emergence of outbreaks was present. Very likely, the still unknown SARS-CoV-2 was already active, but in a very limited way. All it could do was to initiate small localized outbreaks that did not (yet) have the strength to initiate an epidemic. As mentioned briefly above, and in more detail in Ref. [1], no transition of an initial episode to the epidemic phase can only occur if the virus is not permitted to spread to an adequately large number of individuals, as had happened whenever the infection occurred in remote Chinese villages close to bat caves, or when an infected individual is immediately hospitalized. Other factors which are less well characterized also play roles in limiting the outbreaks, e.g. the viral load itself, and the mobility and general

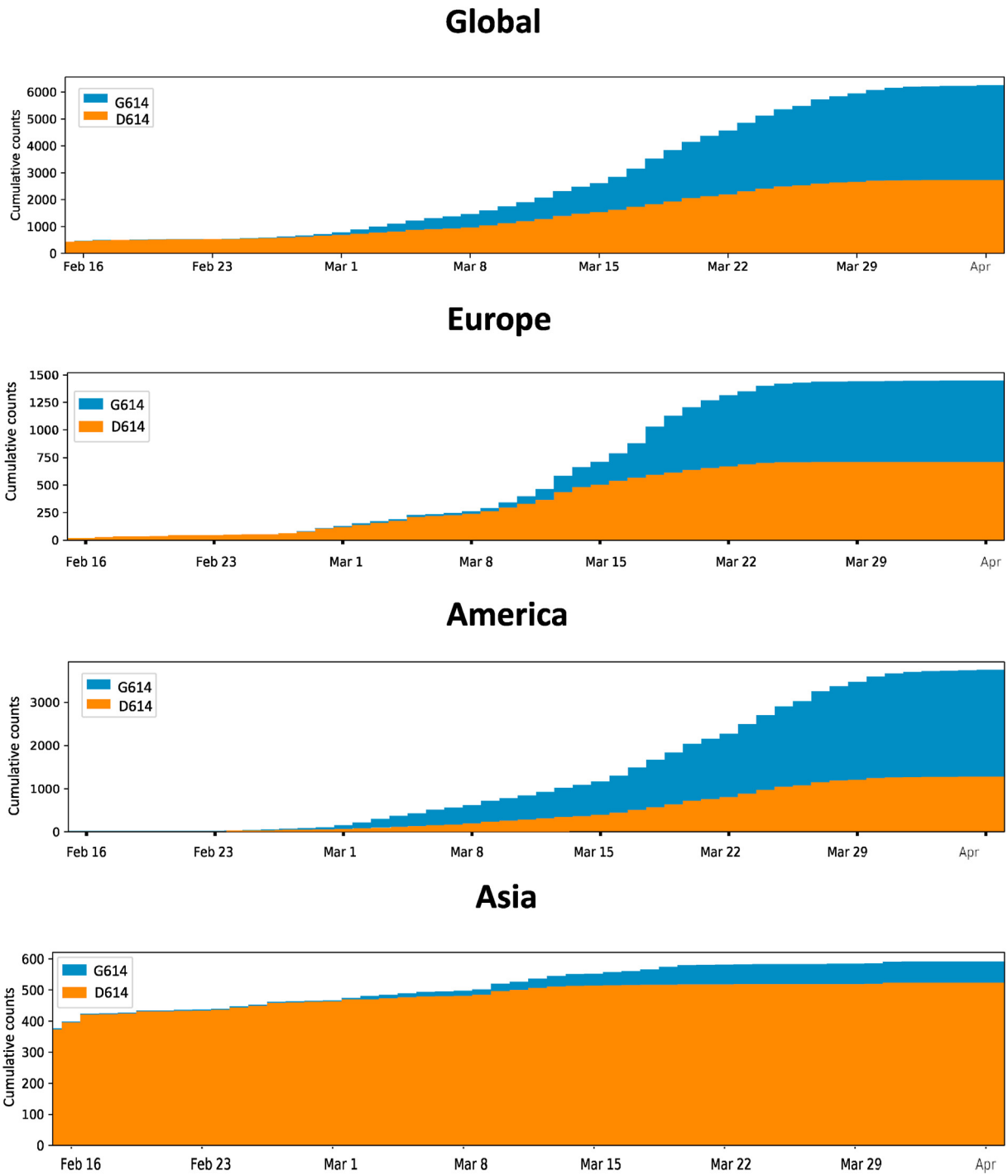

Fig. 4. Daily cumulative count of the relative amount of the wild type Spike protein of SARS-CoV-2 with the wild type D614 and of the mutant G614 in different world regions. After the mutant protein enters a region, it soon becomes the dominating variant. Modified from B. Korber et al. [22].

attitude of the founder.

In hindsight, indirect indications on the activity of the still unknown SARS-CoV-2 in the pre- Wuhan period are plentiful. In the first half of 2019, including the first Summer months, the cases of flu in China had exceeded the total in the preceding four years, with about 270 deaths (https://www.jiemian.com/article/3416522.htm). Long queues were recorded at pediatric hospitals for flu-related symptoms outside the influenza peak season (https://www.thepaper.cn/newsDetail_forward_5263340) (normally, people who visit Chinese hospitals claiming flu symptoms are not submitted to laboratory analysis, but are simply sent home with medications). A report from Harvard University, Office of Scholarly Communication, quotes a study by E.O. Nsoesie et al. (http://nrs.harvard.edu/urn-3:HUL.InstRepos:42669767) in which satellite imagery of hospital parking lots and Baidu search queries of disease related terms show an upward trend in hospital traffic and search volume beginning in late Summer and early Fall 2019. While queries of the respiratory symptom “cough” show fluctuations that coincide with the flu season, “diarrhea” is a more COVID-19 specific symptom. The increase of both signals precedes the documented

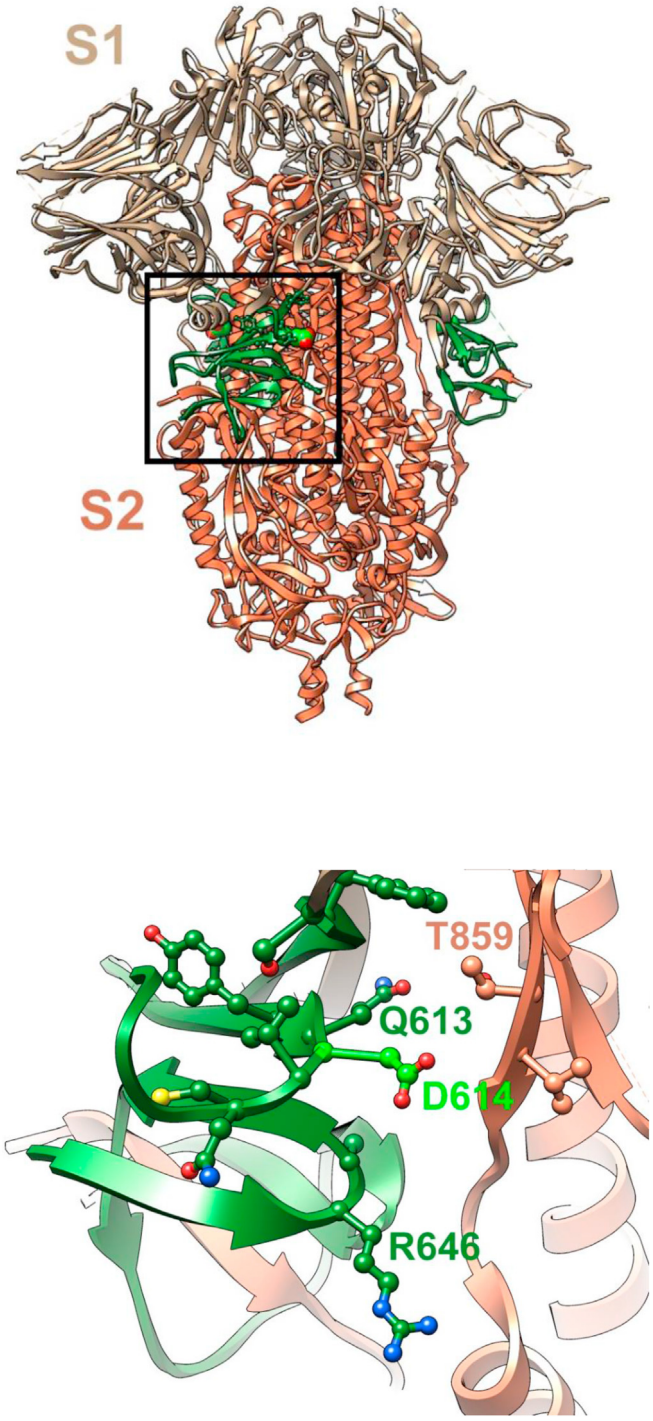

Fig. 5. The D614G mutation in the Spike protein of SARS-CoV-2. Top: cryo-EM structure of the S1 (brown) and S2 (orange) subunits of the Spike protein. Residues 581–676, a C-terminal region of the S1 domain involved in S2 interaction is colored green. Aspartic acid 614 is shown in light green. Bottom: the area indicated with a black square is shown magnified. Residues within 5.5 A of D614 are shown in a ball and stick representation (see text Modified from L.Zhang et al.) [24]. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

start of the COVID_19 epidemic in late December in Wuhan. Interestingly, the COVID-19 infection may have spread to other Countries as well in the final months of 2019. Unusually large numbers of anomalous interstitial pneumonias were observed in numerous Italian and French hospitals before the end of 2019. Retrospective scans of some of the patients admitted to hospitals in Colmar and Mulhouse had revealed cases with typical radiological features of COVID-19 pneumonia. The presence of COVID-19 infection in France before the end of 2019 was actually conclusively documented on a patient admitted to a Paris hospital at the end of December 2019 with severe pulmonary symptoms. Retrospective PCR tests on a stored respiratory sample demonstrated the presence of SARS-CoV-2.

One particularly important event in the pre-explosion phase have been the World Military games that took place in Wuhan from October 18 through October 27, 2019. The Games brought to Wuhan more tan 10,000 athletes from 110 Countries, including many from different Chinese provinces. The naturally brought to Wuhan the correspondingly large number of spectators expected of events of this importance. A number of athletes had developed symptoms that were recognized to be those of COVID-19 infection when they had returned to their home Countries, e.g., Italy, Sweden, France, the USA. The five USA athletes actually returned home with a diagnosis of malaria, ahead of the conclusion of the Games. No laboratory tests had been performed on the sick athletes in Wuhan: later on, some had described in vivid interviews their symptoms (at least one of them had transmitted the infection to his girl-friend), saying that many more athletes had been infected. The diseases of the athletes had naturally attracted larger attention, but they were, in a sense, the tip of the iceberg. It was common knowledge that numerous spectators of the Games had been affected as well: in hindsight, the Games had acted as a superspreaders of the infection, and their negative role was increased by the fact that a few weeks after their closing the New Year vacation took millions of Chinese to their home towns all over China. The timeline in Fig. 6 shows a visual rappresentation of the possible COVID-19 infection in China prior to the Wuan explosion.

4.3.2. On the Wuhan explosion

A still unknown virus was thus causing frequent outbreaks in the last months of 2019. The outbreaks had remained limited, as the conditions necessary to make them epidemic had not been meet. At end of 2019 they were - unfortunately-met with particular violence in Wuhan, specifically in the wet markets. Which are places frequently located close to residential areas, where food is stored a and sold, also in the form of all sorts of live animals that are slaughtered there, including civet cats, raccoon dogs, pigs and -yes- different species of bats. In the wet markets, live animals are promiscuously crowded in cages in non-existing hygienic conditions with dispersal of dangerous excrement [1]. These are conditions that obviously abolish the *dilution effect* of the biodiversity, and creating instead an *amplification zone,* where the high density of pathogen's reservoirs and of susceptible hosts can interact optimally, greatly increasing the chances of outbreaks [1,3]. Humans entering in contact with this *amplified zone* also represent susceptible hosts, who can be infected indirectly through an intermediate host or directly by the animal reservoir. As already briefly discussed above, viruses (SARS-CoV-2 in this case) could in this *amplified environment* modify their genome, generating variants, for instance, that are better adapted to humans, thus increasing their spreading power, first from animals to humans, and then from human to human [1]. The entire picture was further complicated by the fact that the Chinese New Year vacation is the highest season for the sale of the wet market animals, and by the fact that ii unfortunately coincides with the winter period, in which infective respiratory tract ailments occur with higher incidence.

It should have become clear from the discussion above that the transition from the relative "dormancy" of the SARS-CoV-2 in the second half of 2019 to its violent explosion at the turn of the

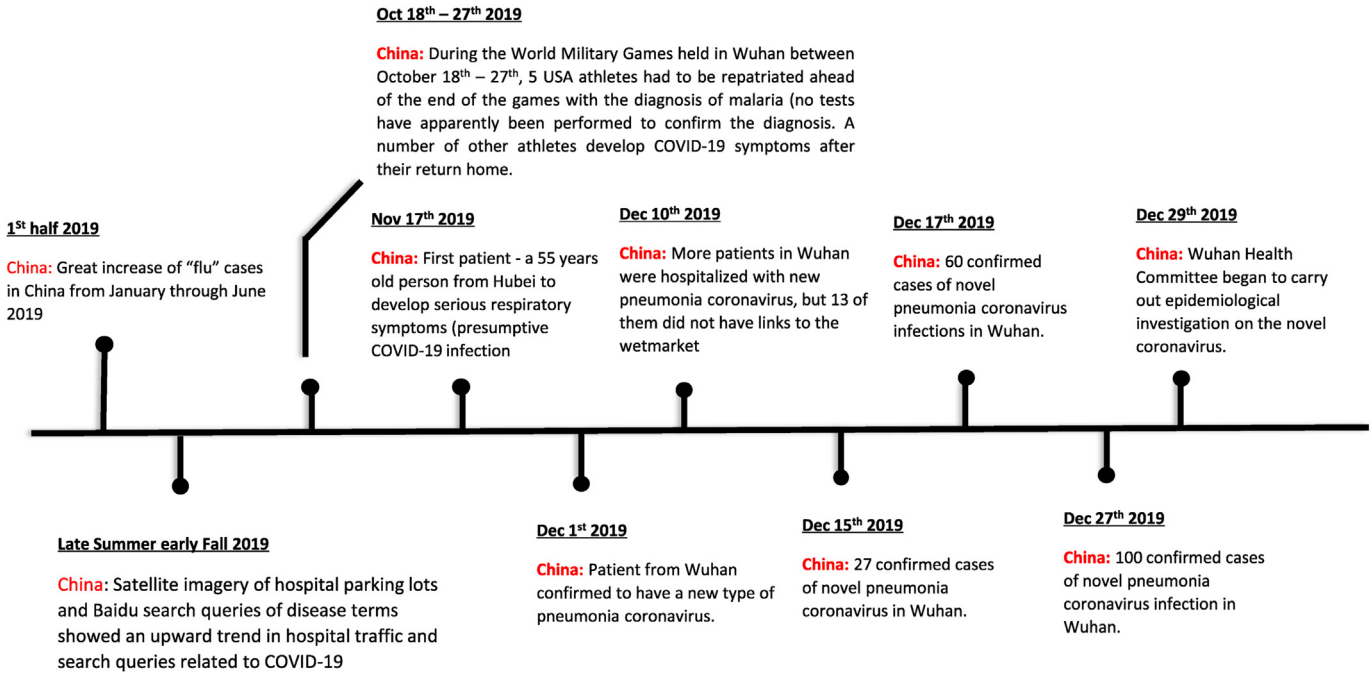

Fig. 6. A timeline of the of the possible COVID-19 infection in China prior to the Wuhan explosion.

century was in all likelihood bound to happen. Wuhan was the place where a number of unfavorable factors -the wet-markets, the World Military Games and their aftermath, the winter season-all coincided to make the transition possible. Knowing what we have now learned we can plausibly say tht the epidemic was predictable.

4.3.3. On the worldwide spreading

This last portion of the contribution is easier to deal with: the biosystem changes and the other factors described in the last few lines were predictable. If not in Wuhan, sooner or later the COVID-19 would in all likelihood have become an epidemic. Something, for instance akin to the 2002–2003 SARS epidemic. The difference, in the case of SARS-CoV-2, is its terribly efficient worldwide diffusion. Which. from all what we have learned, is due to the appearance of a mutation in the Spike protein of the virus that has tremendously increased its transmissibility. A mutation that could not have been predicted. This appears to be the key difference between the virus that initiated the epidemic in Wuhan in December–January, and the virus that beginning at the end of January of this year, has invaded the world in what probably is the most devastating pandemic in history. Clearly, all this was not predictable and, to make the details of the story even more difficult to understand, there is the finding that -see above-the mutation had most likely first occurred during the epidemic in Wuhan and/or China: true, the first case was reported in Germany, but it may have been indirectly connected with Wuhan. What is clear is that the mutation evidently had a positive selection character given its fantastically rapid spreading, that in a few weeks after its first appearance has made it by far the dominant variant of the virus all over the world. However, not In China, where, for reasons that are still not understood, it apparently did not spread with the same fantastic success.

Declaration of competing interest

The authors declare that they don't have conflicts of interest.

Aknowledgement

The authors are greatly indebted to Andrea Nicoletto for his competent work in the preparation of the Figures.

References

[1] S. Platto, T. Xue, E. Carafoli, COVID-19: an announced pandemic, Cell Death Dis. 11 (2020) 799–812.
[2] V. Berti Deslandes, J. Tandjaoui-Lambotte, C. Alloui, E. Carbonelle, J.R. Zahar, S. Brichler, Y. Cohen, SARS-CoV-2 was already spreading in France in late December 2019, Int. J. Antimicrob. Agents 55 (2020), https://doi.org/10.1016/j.ijantimicag.2020.106006. Epub 2020 May 3.
[3] S. Platto, J. Zhou, W. Yanqing, W. Huo, E. Carafoli, Biodiversity loss and COVID-19 pandemic: the role of bats in the origin and the spreading of the disease, Biochem. Biophys. Res. Commun. (2020), https://doi.org/10.1016/j.bbrc.2020.10.028. PMID: 33092787; PMCID: PMVC7566801.
[4] A.D. Luis, D.T.S. Hayman, T.J. O'Shea, P.M. Cryan, A.T. Gilbert, J.R.C. Pulliam, J.N. Mills, M.E. Timonin, C.K.R. Willis, A.A. Cunningham, A.R. Fooks, C.E. Rupprecht, J.L.N. Wood, C.T. Webb, A comparison of bats and rodents as reservoirs of zoonotic viruses: are bats special ? Proc. Royal Soc. B 280 (2013) https://doi.org/10.1098/rspb.2012.2753. Print 2013 Apr 7.
[5] P. Forster, L. Forster, C. Renfrew, M. Forster, Phylogenetic network analysis of SARS-CoV-2 genomes, Proc. Natl. Acad. Sci. U.S.A. 117 (2020) 9241–9243.
[6] P. Zhou, X.-L. Yang, X.-G. Wang, B. Hu, L. Zhang, W. Zhang, H.-R. Si, Y. Zhu, B. Lei, C.-L. Huang, H.-D. Chen, J. Chen, Y. Luo, H. Guo, R.-D. Jiang, M.-Q. Liu, Y. Chen, X.-R. Shen, X. Wang, X.-S. Zheng, K. Xhao, Q.-J. Chen, F. Deng, L.-L. Liu, B. Yan, F.-X. Zhzn, Y.-Y. Wang, G.-F. Xiao, Z.-L. Shi, A pneumonia outbreak associated with a new coronavirus of probable bat origin, Nature 579 (2020) 270–273.
[7] W. Li, M.J. Moore, N. Vasilieva, J. Sui, S.K. Wong, M.A. Berne, M. Somasundaran, J.L. Sullivan, K. Luzuriaga, T.C. Greenhough, H. Choe, M. Farzan, Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirys, Nature 426 (2003) 450–454.
[8] M. Letko, A. Marzi, V. Munster, Functional assessment of cell entry and receptor usage for SARS-CoV-2 and other lineage B betacoronaviruses, Nature Microbiol 5 (2020) 562–569.
[9] M. Hoffmann, H. Kleine-Weber, N. Krueger, M. Mueller, C. Drosten, S. Poehlmann, The novel coronavirus 2019 (2019-nCoV) uses the SARS coronavirus receptor ACE2 and the cellular protease TMPRSS2for entry into target cells, BioRxiv doi.org/10.1101/2020.01.31.929042.
[10] V.S. Raj, H. Mou, S.L. Smits, D.H.W. Dekkers, M.A. Mueller, R. Dijkman, D. Muth, J.A.A. Demmers, A. Zaki, R.A.M. Fouchier, F. Thiel, C. Drosten, P.J.M. Rottier, A.D.M.E. Osterhaus, B.J. Bosch, B.L. Haagmans, Dipeptidyl peptidase 4 is a functional receptor for the emerging coronavirus-EMC, Nature 495 (2013) 251–254.
[11] X. Zhang, K. Wu, D. Wang, X. Yue, D. Song, Y. Zhu, J. Wu, Nucleocapsid protein

of SARS- CoV activates interleukin-6 expression through cellular transcription factor NF-kB, Virology 365 (2007) 324–335.
[12] C. Walls, M.A. Tortorici, B.-J. Bosch, B. Frenz, P.J.M. Rottier, F. DiMaio, F.A. Rey, D. Veesler, Cryo-electron microscopy structure of a coronavirus spike glycoprotein trimer, Nature 351 (2016) 114–119.
[13] D. Wrapp, N. Wang, K.S. Corbett, J.A. Goldsmith, C.-L. Hsieh, O. Abiona, B.S. Graham, J.S. McLellan, Cryo-EM structure of the 2019-nCoV spike in the perfusion conformation, Science 367 (2020) 1260–1263.
[14] M. Hoffmann, H. Kleine-Weber, S. Schroeder, N. Krueger, T. Herrler, S. Erichsen, T.S. Schiergens, G. Herrler, N.H. Wu, A. Nitsche, M.A. Mueller, C. Drosden, S. Poehlmann, SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor, Cell 181 (2020), 271-280.
[15] M. Hoffmann, H. Hofmann Winkler, S. Poehlmann, Priming time: how cellular proteases arm coronavirus spike proteins, in: E. Boettger-Frieberts- aeuser, W. Gaertner, H.D. Klenk (Eds.), Activation of Viruses by Host Proteases, Springer, Cham, 2018, pp. 71–98.
[16] J. Fantini, C. Di Scala, H. Chah16 inian, N. Yahi, Structural and molecular modelling studies reveal a new mechanism of action of Chloroquine and hydroxychloroquine against SARS-CoV-2 infection, Int. J. Antimicrob. Agents 55 (2020), https://doi.org/10.1016/j.ijantimicag.2020.105960. Epub 2020 Apr 3.
[17] Y. Peng Wu, B. Huang, X. Ding, X. Wang, P. Niu, J. Meng, Z. Zhu, Z. Zhang, J. Wang, J. Sheng, L. Quan, Z. Xia, W. Tan, G. Cheng, T. Jiang, Genome composition and divergence of the novel coronavirus (2019-nCoV)m originating in China, Cell Host Microbe 27 (2020) 325–328.
[18] K.G. Andersen, A. Rambaut, W. I. Lipkin, E.C. Holmes, R.F. Garry, The proximal origin of SARS-CoV-2, Nat. Med.. Doi. 10.1038/s41591-020-0820-9.
[19] M. Gheblawi, K. Wang, A. Viveiros, Q. Nguyen, J.-C. Zhong, A.J. Turner, m.K. Raizada, M.P. Grant, G.Y. Oudit, Angiotensin converiting enzyme 2: SARS-CoV-2 receptor and regulator of the renin-angiotensin system, Circ. Res. 126 (2020) 1457–1475.
[20] J. Shang, G. Ye, K. Shi, Y. Wan, C. Luo, H. Aihara, Q. Geng, A. Auerbach, F. Li, Structural basis of receptor recognition by SARS-CoV-2, Nature 581 (2020) 221–222.
[21] J.A. Jaimes, J.K. Milet, G.R. Whittaker, Proteolytic cleavage of the SARS-CoV-2 spike protein and the role of the novel S1/S2 site, iScience 182 (2020) 1–5, 812–82723.
[22] B. Korber , W.M Fischer, S. Gnanakaran, H. Yoon, J. Theiler, W. Abfalterer, B. Foley, E.E. Giorgi, T. Bhattakharya, M.D. Parker, D.G. Partridge, C.M. Evans, T.I. de Slva, On behalf of the Sheffield COVID-19 Genomics Group, C.C. La Branche,C. Montefiori, Spike mutation pipeline reveals the emergence of a morec transmissible form of SARS.CoV-2, BioRxiv,doi.10.1101/2020.04.29.099054.
[23] B. Korber, W.M. Fischer, S. Gnanakaran, H. Yoon, J. Theiler, W. Abfalterer, N. Hengartner, E.E. Giorgi, T. Bhattacharya, B. Foley, K.M. Hastie, M.D. Parker, D.G. Partridge, C.M. Evans, T.M. Freeman, T.I. de Silva, , on behalf of the Sheffield COVID-19 Genomics Group, C. McDanald, L.G. Perz, H. Tang, A. Moon-Walker, S.P. Whelan, C.C. LaBranche, E.O. Shaphire, D.C. Montefiori, Cell 182 (2020) 812–827.
[24] L. Zhang, C.B. Jackson, H. Mou, A. Oiha, S.S. Rangarajan, T. Izard, M. Farzan, H. Choe, The D614G mutation in the SARS-CoV-2 spike protein reduces S1 shedding and increases infectivity, bioRxiv, Doi. 10.1101/2020.06.12.148726.
[25] Y. Guan, B.J. Zheng, Y.Q. He, X.L. Liu, Z.X. Zhuang, C.L. Cheung, S.W. Luo, P.H. Li, L.J. Zhang, Y.J. Guan, K.M. Butt, K.L. Wong, K.W. Chan, W. Lim, K.F. Shortridge, K.Y. Yuen, J.S.M. Peiris, L.L.M. Poon, Isolation and characterization of viruses related to the SARS coronavirus from animals in Southern China, Science 302 (2003) 276–278.
[26] D. Wu, C. Tu, C. Xin, H. Xuan, Q. Meng, Y. Liu, Y. Yu, Y. Guan, Y. Jiang, X. Yin, G. Crameri, M. Wang, C. Li, S. Liu, M. Liao, L. Feng, H. Xiang, J. Sun, J. Chen, Y. Sun, S. Gu, N. Liu, D. Fu, B.T. Eaton, L.-F. Wang, X. Kong, Civets are equally susceptible to experimental infection by two different severe acute respiratory syndrome coronavirus isolates, J. Virol. 79 (2005) 2620–2625.
[27] B. Kan, M. Wang, H. Jing, H. Xu, X. Jiang, M. Yan, W. Liang, H. Zheng, K. Wan, Q. Liu, B. Cui, Y. Xu, E. Zhang, H. Wang, J. Ye, G. Li, M. Li, Z. Cui, X. Qi, K. Chen, L. Du, K. Gao, Y-t. Zhao, X-z. Zou, Y.-J. Feng, Y.-F. Gao, R. Hai, D. Yu, Molecular evolution analysis and geographic investigatiin of severe acute respiratory syndrome coronavirus-like virus in palm civets at an animal market and on farms, J. Virol. 79 (2005) 11892–11900.
[28] L.F. Wang, Z. Shi, S. Zhang, H. Field, P. Daszak, B.T. Eaton, Review of bats and SARS emerging infectious diseases. https://doi.org/10.3201/eid1212.060401.
[29] T.T.-Y. Lam, N. Jia, Y.-W. Zhang, M.H.- Shum, J.-F. Jiang, H.-C. Zhou, Y.-G. Tong, Y.-X. Shi, X.-B. Ni, Y.-S. Liao, W.-J. Li, B.-G. Jiang, W.W.T.-T. Juan, K. Zheng, X.-M. Cui, J. Li, G.-Q. Pei, X. Xiang, W.Y.-M. Cheung, L.-F. Li, F.-F. Sun, S. Qin, Ji-C. Huang, G.M. Leung, E.C. Holmes, Y.-L. Hu, Y.I. Guan, W.-C. Cao, Identifying SARS-Co V2-related coronaviryses in Malayan pangolins, Nature 583 (2020) 282–285.
[30] L. Ping, J.-Z. Jiang, Y.X.-F. Wan, Y. Hua, L. Li, J. Zhou, X. Wang, F. Hou, J. Chen, J. Zhou, J. Chen, Are pangolins the intermediate hosts of the 2019 novel coronavirus(SARS-CoV-2)? PLoS Pathog. 16 (2020) 1–13.
[31] J. Damas, G.M. Hughes, K.C. Keough, C.A. Painter, N.S. Persky, M. Corbp, M. Hiller, k.P. Koepfli, A.R. Pfenning, H. Zhao, G.P. Genereux, R. Swofford, K.S. Pollard, O.A. Ryder, M.T. Nweeia, K. Lindblad-Toh, E.C. Teeling, E.K. Karksson, H.A. Levin, Broad host range of SARS- CoV-2 analysis of ACE2 in vertebrates, by comparative and structural analysis of ACE2 in vertebrates, bioRxiv: https://doi.org/10.1101/2020.04.16.045302.
[32] J. Lee, M. T. Hughes, M-H. Lee, H. Field, J. J. Rovie-Ryan, F.T. Sitam, S. Sipangkui, S.K,S.S. Nathan, D. Ramirez, S.V. Kumar, H. Lasimbang, J.H. Epstein, P. Daszak, No evidence of coronaviruses or other potentially zoonotic viruses in Sunda pangolins (Manis javanica) entering the wildlife trtade via Malysia, bioRxiv: https://doi.org/10.1101/2020.06.19.158717.
[33] X.C. Tang, J.X. Zhang, S.Y. Zhang, P. Wang, X.H. Fan, L.F. Li, G. Li, B.Q. Dong, W. Liu, C.L. Cheung, K.M. Wu, W.J. Song, D. Vijaykrishna, L.L.M. Poon, J.S.M. Peiris, G.J.D. smith, H. Chen, Y. Guan, Prevalence and genetic diversity of coronaviruses in baths from China, J. Virol. 80 (2006) 7481–7490.