Virologica Sinica 37 (2022) 198–207

Contents lists available at ScienceDirect

Virologica Sinica

journal homepage: www.keaipublishing.com/en/journals/virologica-sinica
www.virosin.org

Research Article

The mutational dynamics of the SARS-CoV-2 virus in serial passages *in vitro*

Sissy Therese Sonnleitner [a,b,*], Stefanie Sonnleitner [a], Eva Hinterbichler [a], Hannah Halbfurter [a], Dominik B.C. Kopecky [a], Stephan Koblmüller [c], Christian Sturmbauer [c], Wilfried Posch [b], Gernot Walder [a]

[a] Dr. Gernot Walder GmbH, Medical Laboratory, Department of Virology, Ausservillgraten, 9931, Austria
[b] Division of Hygiene and Medical Microbiology, Medical University of Innsbruck, Innsbruck, 6020, Austria
[c] Institute of Biology, University of Graz, Graz, 8010, Austria

ARTICLE INFO

Keywords:
SARS-CoV-2
Whole genome sequencing (WGS)
Mutational dynamics
Adaptation
Serial passage *in vitro*

ABSTRACT

Since its outbreak in 2019, Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) keeps surprising the medical community by evolving diverse immune escape mutations in a rapid and effective manner. To gain deeper insight into mutation frequency and dynamics, we isolated ten ancestral strains of SARS-CoV-2 and performed consecutive serial incubation in ten replications in a suitable and common cell line and subsequently analysed them using RT-qPCR and whole genome sequencing. Along those lines we hoped to gain fundamental insights into the evolutionary capacity of SARS-CoV-2 *in vitro*. Our results identified a series of adaptive genetic changes, ranging from unique convergent substitutional mutations and hitherto undescribed insertions. The region coding for spike proved to be a mutational hotspot, evolving a number of mutational changes including the already known substitutions at positions S:484 and S:501. We discussed the evolution of all specific adaptations as well as possible reasons for the seemingly inhomogeneous potential of SARS-CoV-2 in the adaptation to cell culture. The combination of serial passage *in vitro* with whole genome sequencing uncovers the immense mutational potential of some SARS-CoV-2 strains. The observed genetic changes of SARS-CoV-2 *in vitro* could not be explained solely by selectively neutral mutations but possibly resulted from the action of directional selection accumulating favourable genetic changes in the evolving variants, along the path of increasing potency of the strain. Competition among a high number of quasi-species in the SARS-CoV-2 *in vitro* population gene pool may reinforce directional selection and boost the speed of evolutionary change.

1. Introduction

Severe acute respiratory syndrome coronavirus (SARS-CoV-2) was first described in December 2019 (Zheng, 2020). To date it is responsible for (10th of January 2022) 5.49 million people's deaths (https://covid19-surveillance-report.ecdc.europa.eu/). Over the course of two years, its genetic structure has changed dramatically. The scientific community worldwide is currently witnessing the amazing evolutionary potential of SARS-CoV-2 in the rapid emergence of immune escape variants. We are gaining insight into the viral adaptational capacity to respond to new environmental influences. Quickly, immune escape variants that increased transmissibility or escape from specific antibodies gained traction and ground due to increasing immunity (https://www.ecdc.europa.eu/en/publications-data/covid-19-infographic-mutations-current-variants-concern).

Genetic changes take place particularly in the genetic region coding for spike as the main docking point for cell entry and in particular at the receptor binding domaine (RBD) as the key to gaining entry into the host cell (Andreano et al., 2020; Koenig et al., 2021; Verkhivker et al., 2021). Therefore, the spike protein represents the main target for neutralizing antibodies and thus is exposed to strong selective pressure (Brouwer et al., 2020; Cao et al., 2020; Ju et al., 2020; Liu et al., 2020; Rogers et al., 2020; Seydoux et al., 2020; Wec et al., 2020; Wu et al., 2020; Zost et al., 2020; Verkhivker et al., 2021; Elena and Sanjuán, 2005). Investigating mutations in the spike region simultaneously provides indications of the functionality and importance of this genomic region.

The reason for the high mutational potential of SARS-CoV-2 lies in the unique properties of its RNA virus polymerase with its intrinsic error rate and characteristically low fidelity (Elena and Sanjuán, 2005; Lauring and Andino, 2010). As in most RNA viruses, the lack of proofreading ability of

* Corresponding author. Dr. Gernot Walder GmbH, Medical Laboratory, Department of Virology, Ausservillgraten, 9931, Austria.
E-mail address: sissy.sonnleitner@infektiologie.tirol (S.T. Sonnleitner).

https://doi.org/10.1016/j.virs.2022.01.029
Received 26 September 2021; Accepted 21 January 2022
Available online 29 January 2022

the RNA virus polymerase leads to higher mutation rates, rapid replication kinetics and large population sizes compared to DNA viruses (Holland et al., 1982; Lauring and Andino, 2010). There are different types of mutations, with point nucleotide substitution (= single nucleotide polymorphisms, SNPs) being the most common. SNPs have been proposed to be one of the most important and most common mechanisms of viral evolution in nature (Lauring and Andino, 2010; Badua et al., 2021; Franco-Muñoz et al., 2020). Other forms of mutations are insertions and deletions (indels). These in particular illustrate in a fascinating way how RNA viruses manage loss and gain of genetic material and, together with SNP, provide valuable information for tracing origin and dispersion patterns of SARS-CoV-2 (Panzera et al., 2021). Due to the great importance of insertions and deletions, all three mutational events (substitutions, deletions and insertions) should be considered when calculating phylogenetic relationships to draw an authentic picture of the most likely genetic formation (Chrisman et al., 2021).

The current pandemic situation affords us all with the opportunity to witness the high mutational potential of SARS-CoV-2 under the selective pressure of rising immunity due to infection and vaccination. Still, we know very little about the baseline, meaning the natural evolutionary rate of SARS-CoV-2 under largely constant conditions as well as its response to a host change. Also, the functionality of genomic segments others than spike is still to a large extant unknown. A further question awaiting an answer is whether the potential of an immune escape variant lies dormant in every SARS-CoV-2 strain This study provides first insights into the mutational strategies of SARS-CoV-2 when adapting to cell culture and gives insight into the mechanism of successful adaptation and maximum fitness development in a single strain. We investigated the emergence, establishment and loss of mutations in ten different regional strains of SARS-CoV-2 in the course of ten serial passages in Vero B4 cells by means of RT-PCR and whole genome sequencing (WGS).

2. Materials and methods

2.1. Isolation of SARS-CoV-2

Over the course of four weeks in April 2020, every sample positive for SARS-CoV-2 in qPCR with a Ct < 27 was taken for isolation in cell culture. Isolation of SARS-CoV-2 was performed from naso-pharyngeal samples after positive specific qPCR (Gietl et al., 2021) by inoculation on Vero B4 (no. ACC-33, DSMZ) in T25 tissue culture flasks (Sarstedt, Germany, Cat. No. 83.3910.300) for 1 h at 36 °C. After incubation, the sample was removed and Medium199 (Lonza, Switzerland, Cat. No. BE12-117F) with 2.5% fetal calf serum (FCS, Lonza, Cat. No. 77227) and a mixture of antibiotics was added (streptomycin, vancomycin, penicillin, each 1 μg/mL). Among the ten strains, there were also two that could be isolated from the same patient at different times-once during the initial infection (A607) and once during recurrence (A4707). We monitored virus cultures daily for cytopathic effects and tested for specific viral RNA every three days. Isolation was considered successful when cytopathic effect was 80%–100% in passage 0 and the Ct value in RT-qPCR was lower than 20. Highly positive supernatants were harvested, centrifugated at 28,000×*g* for 5 min and stored at −80 °C in 10% FCS. All work involving infectious SARS-CoV-2 was conducted in a BSL3 facility, following the institutional guidelines and regulations.

2.2. Serial passage

Confluent Vero B4 cells (no. ACC-33, DSMZ) were cultured in Medium199 (Lonza). Ten 25 cm^2 cell culture flasks were infected with ten different isolates of the first regional outbreak in passage 0. Another one 25 cm^2 cell culture flask was treated exactly the same way but with medium instead of viral infection and served as a negative control. The initial concentration of the diverse strains of SARS-CoV-2 was ascertained to be Ct 15 in RT-qPCR at the beginning.

Vero B4 cells were cultured to a 100% confluent monolayer in Medium199 including 5% FCS (Lonza) in T25 tissue culture flasks (Sarstedt). The T25 flasks were infected with SARS-CoV-2 isolates of passage 0 after 24 h. After 1 h of infection, the monolayers were overlayed with 5 mL of Medium199 including 2% FCS and incubated at 35 °C. Every 14 days, supernatants were checked for viral replication via cytopathic effect and RT-qPCR and were harvested. Backups were stored at −80 °C in 10% FCS, whereas 1 mL of the supernatant was used to infect the next T25 flask with a new Vero B4 monolayer. This procedure started in September 2020 and was repeated over 125 days and 9 times until passage 10 in the end of January 2020. Quantitation of the viral load was performed by RT-qPCR. The conversion to plaque-forming units (PFU)/mL was done according to the specifications of Ramakrishnan, 2016, as described elsewhere in detail (Sonnleitner et al., 2021).

2.3. Whole genome sequencing and mutation analysis

Backups of all passages were thawed and nucleic acids were isolated using the MagMAX™-96 Total RNA Isolation Kit (Thermo Fisher Scientific, Waltham, Massachusetts, USA, Cat. No. AM1830). Briefly, 200 μL PBS were taken from patient swab sample and mixed with 265 μL Binding Buffer, 5 μL proteinase K and 5 μL extraction control (Thermo Fisher Scientific, Cat. No. AM1830) according to the KingFisher™ extraction protocol for 200 μL sample volume (Thermo Fisher Scientific). After incubation at room temperature for at least 15 min, samples were transferred from tubes into 96-well KingFisher Deep well plates (Thermo Fisher Scientific) containing 280 μL isopropanol and 2 μL Mag-Bind Particles per well, using a KingFisher™ Flex Purification System (Thermo Scientific™, Cat. No. 5400620).

RT-qPCR was performed to evaluate initial viral load as a quantitative starting point for the creation of the library (Gietl et al., 2021). Libraries were prepared according to the COVID-19 Ion AmpliSeq library construction and sequencing protocol with the Library Kit Plus (Thermo Fisher Scientific, Cat. No. 4488990). Amplicons were cleaned up with AMPure XP beads (Beckman Coulter, Germany) with a 1:1 ratio. The libraries were quantified using the Ion Library TaqMan™ Quantitation Kit (Ion Torrent™, Cat. No. 4468802). Normalizing, pooling and sequencing were performed using an Ion Torrent™ S5 Plus. Ion Torrent Suite software (v 5.12.2) of the Ion S5 sequencer was used to map the generated reads to a SARS-CoV-2 reference genome (Wuhan-Hu-1; GenBank accession numbers NC_045512.2), using TMAP software included in the Torrent Suite. For mutation calls, the additional Ion Torrent plugin Variant caller (v5.12.0.4) was used, where all single-nucleotide variants (SNVs) were called with "Generic-S5/S5XL (510/520/530)-Somatic-Low Stringency" default parameters.

FASTA files containing the raw reads were inspected for quality criteria (mapped, filtered and target reads, mean depth and uniformity) using Thermo Fisher Software SARS-CoV-2 coverage Analysis v.1.3.0.2. Multiple sequence alignments were performed using Unipro UGENE software (Okonechnikov et al., 2012) as well as MEGA X software (Kumar et al., 2018), using NC 045512.2-Wuhan-Hu-1 as reference. Detailed information on the quality level of each individual sequence is given in Supplementary Table S1.

Viral genome assembly and screening for distinct mutations was performed online using nextstrain.org (https://github.com/nextstrain/ncov/blob/master/defaults/clades.tsv; https://clades.nextstrain.org/). The identification of pangolin lineages was carried out using Pangolin software, v.2.4.2. (https://pangolin.cog-uk.io/). The full-genome sequences are available at GISAID EpiCoV (https://gisaid.org/; no.EPI_ISL_2106191-21061201; EPI-ISL-3128605, 627, 629, 630, 620). Genetic distances were calculated in comparison to the reference genome of Wuhan (Wuhan-Hu-1; GenBank accession numbers NC_045512 and MN908947.3).

Phylogenetic relationships among the distinct viral genomes were inferred based on maximum likelihood in the IQ-Tree plugin (Nguyen et al., 2015) of PhyloSuite (Zhang et al., 2020). Indels were coded as binary characters using the "simple indel coding" algorithm (Simmons and Ochoterena, 2000) as implemented in 2matrix (Salinas and Little,

2014). We employed the best-fit model of molecular evolution (HKY + F + I), as inferred under the Akaike information criterion using ModelFinder (Kalyaanamoorthy et al., 2017) in PhyloSuite, and statistical support for nodes in the tree was assessed via ultrafast bootstrapping with 5,000 replicates (Hoang et al., 2018). Pairwise distances (mean number of substitutions per site) among and within strains were inferred in MEGA X and translated into substitution rates (mean substitutions per site per year) using the info on sampling time. The substitution rate was calculated as divergence rate/2.

All mutations described here are those resulting in a change in the amino acid sequence and include non-synonymous substitutions (aa substitutions), deletions and insertions.

2.4. Statistical analysis

Dichotomous data were evaluated for statistical significance by a chi-squared test or Fisher's exacta in the case of small group size ($n < 60$) (Microsoft® Excel®, Microsoft 395 MSO, Windows 10). A two-sided significance level of $P = 0.05$ was used for determining statistical significance.

3. Results

3.1. Characteristics of the SARS-CoV-2 strains used in serial passages

The strains were isolated from human swabs in the first pandemic wave in a routine laboratory in East Tyrol/Austria with a within group mean distance of 2.6×10^{-4} (Ct 10) between the ten strains in the beginning of the experiment at passage 0. The overall mean distance between all strains of passage (P) 0 and the reference genome of Wuhan was calculated with 4.4×10^{-4}, which translated to a substitution rate of 6×10^{-4} substitutions per site per year. All strains were D614G variants.

The pairwise distances between the diverse strains and the reference genome Wuhan are shown in Table 1. The SARS-CoV-2 strain with the lowest distance (2.0×10^{-4}) was A893 at the beginning of the experiment and A6137 was the strain with the largest evolutionary distance of 5.4×10^{-4} in comparison to the reference genome Wuhan-Hu-1.

In the course of the serial passages of the ten strains, the overall mean distance between the strains in P10 and the reference genome increased to 6.8×10^{-4} during the study period, implying a substitution rate of 4.8×10^{-4} nucleotide substitutions per site per year. The pairwise distance between all strains at the time point P0 and the time point P10 (intermediate time = four months) amounted to a pairwise distance of 4.4×10^{-4} and a substitution rate of 6.6×10^{-4}. The within group mean distance in P0 was 2.5×10^{-4}, and increased to 5.6×10^{-4} in P10, which in turn corresponded to a substitution rate of 4.7×10^{-4}.

Table 1
Characteristics of the used strains and its human origins.

Sample name	Pairwise distance (P0) ($\times 10^{-4}$)	Sequences (n)	Age (y)	Sex
A893	2.0	6	57	Female
A818	4.0	7	55	Male
A607	4.7	6	59	Male
A7326	5.0	5	22	Male
A6301	4.4	6	90	Female
A6137	5.4	6	15	Female
A4707	4.7	6	59	Male
B5155	4.0	2	28	Female
B4926	4.7	2	52	Female
PB5845	4.7	3	60	Male

Genetic distances (mean pairwise distances, as mean number of substitutions per site) of all ten strains to the reference genome Wuhan–Hu-1 were calculated before starting the experiment (P0). High quality sequences were used for analysis and two to five full-genome sequences of different *in vitro* passages were received by whole genome sequencing (WGS). Sequences (n), number of high-quality sequences received by WGS.

3.2. Development of the SARS-CoV-2 strains after serial passages

The ten strains developed differently throughout the study period of four months and ten passages, respectively. The three strains B4926, B5155 and B5845 did not show clear cytopathic effects in the first two passages and stopped replicating after passage three, confirmed by negative PCR-results. The other seven strains managed to replicate in cell culture over the entire time and achieved very low Ct-values of 8 in RT-qPCR, corresponding to a very high viral load of about 1.1×10^{7} PFU/mL. A detailed overview of the passaged strains and the different substitution rates throughout the *in vitro* passages is given in Table 2.

The substitution rates per year within the strains were between 1.0×10^{-4} in B4926 and 6.3×10^{-4} in A6137. On the inter-strain level, the substitution rates ranged from 4.9×10^{-4} up to 1.0×10^{-3}. The between group mean distances of the strains after ten passages in cell culture are shown in Table 3. The between group mean distances were calculated in comparison to the reference genome of Wuhan. A818 and A6137 were the strains with the lowest and highest divergence rates, respectively, with a pairwise distance of 3.3×10^{-4}, corresponding to an evolutionary rate of 7.1×10^{-4} nucleotides per site per year.

3.3. Genetic evolution in vitro

Sequencing was carried out at passage 0, 3, 5, 7 and 10 respectively, exclusively high quality sequences were used for further analysis and resulted in 49 SARS-CoV-2 whole genome sequences with 29,896 nucleotide sites each and enabled to investigate the chronology of the emergence of the individual mutations. The phylogenetic relationships of the ten strains at their different stages in the course of the serial passages are shown in Fig. 1. A6137 has developed furthest from the reference genome Wuhan-Hu-1 with an astonishingly high evolutionary rate of 1.0×10^{-3} nucleotides per site per year with nine acquired non-synonymous substitutions and an interesting deletion at nucleotide site 9797–9813 (aa position ORF8:A65-).

3.3.1. Substitutions

It is noteworthy that the majority of the non-synonymous mutations occurred after passage four. While the average substitution rate for P1 to P4 was 8.1×10^{-4}, this rate rose sharply to 1.6×10^{-3} nucleotides per site per year in P5 to P10. A summary of all substitutions is given in Fig. 2.

3.3.2. Quantitative evaluation of the ability of reproduction in vitro

The initial load in all ten strains was adjusted to a Ct of 10 in *E*-gene screening PCR, which was equivalent to 1×10^{7} PFU/mL according to

Table 2
Overview of the different within mean group evolutionary rates of ten different strains after serial passage *in vitro* over a four-month study period under identical environmental conditions.

Strain	Within group mean pairwise distance ($\times 10^{-4}$)	Substitution rate per year ($\times 10^{-4}$)
B4926	0.67	1.0
B5155	0.67	1.0
A6301	0.85	1.3
A607	1.4	2.1
A7326	1.4	2.1
A893	1.6	2.4
B5845	1.8	2.7
A818	2.4	3.6
A4707	2.5	3.8
A6137	4.2	6.3

The within group mean distance reflects the mean number of substitutions per site within a group and thus gives an indication of the heterogeneity or homogeneity of a group at the beginning of the experiment. The substitution rate indicates the mean number of substitutions per site per year and reflects the tempo of evolution of the various strains since the separation of the reference strain Wuhan-Hu-1.

Table 3
Divergence of the isolated strains in serial passages *in vitro* and the Wuhan reference genome in ascending order.

Strain	Between group mean pairwise distance ($\times 10^{-4}$)	Substitution rate per year (8-month study period) ($\times 10^{-4}$)
A818	3.3	4.9
B5155	3.7	5.5
B5845	4.0	6.0
A893	4.3	6.4
B4926	4.7	7.1
A6301	4.9	7.3
A607	5.0	7.6
A7326	6.0	9.0
A4707	6.2	9.2
A6137	7.1	10.0

The between group mean distance gives the mean number of substitutions per site between the various strain and the reference genome Wuhan-Hu-1 at the end of the serial passages. .

our virus titration (Sonnleitner et al., 2021). After ten passages, three strains were negative in RT-qPCR and had already stopped replication within the first four passages. Four strains retained their viral loads without significant titer change whereas three strains could increase their abilities to reproduce in cell culture significantly and yielded high copy numbers, up to an increase of 4×10^3% up to 3×10^4%. The detailed quantitative development of the ten strains is shown in Table 4 and Fig. 2.

The increased replication success *in vitro* seems to be associated with an increased number of substitutions. Whereas the unsuccessful strains did not evolve aa substitutions, the most successful strains showed the largest number of acquired permanent substitutions and temporary mutational changes, respectively (2/3). The strains B5155, B4926 and B5845 did not manage to replicate in cell culture and did not evolve any aa substitutions or deletions. The strains A607, A818, A893 and A6301 managed to stay in culture during the three-month study period with a notable decrease or increase of viral load and showed a moderate number of aa substitutions. The three strains with a significant increase in viral load (A7326, A6137, A4707) showed larger numbers of mutations including temporary mutations, which occurred once or twice but had not been preserved. The four strains with low replication success evolved between two and six aa substitutions [mean = 4; standard deviation (SD) = 1.6], whereas the highly successful strains developed six to seven substitutional mutations (mean = 6.7, SD = 0.6), not including temporary mutations, which would increase to seven to twelve mutations during the same investigation period (mean = 9.3; SD = 2.5). Most of the genetic changes occurred in the genomic regions of spike and ORF1a. Table 5 gives a clear list of all acquired mutations during the experiment. Almost all genetic changes (5/7, 71.4%) that took place in A607 strain were recorded as temporary mutations in the closest related strain A4707. In contrast to strain A607, however, additional seven mutations developed in A707 (ORF1a: A1352V; S:H66R; S:E96G; S:N81T; S:D215A; ORF3a: M1I; E:S6L). A607 developed one substitution (E:V58A) that was not present in A4707 and one temporary insertion (S: 22116-AAAAAGTCTACT -22117). Not all strains developed substitutions in cell culture, but some evolved convergently in three or even four strains. S:E484D was the substitution which occurred four times independently in the strains A818, A893, A6301 and A6137. E:S6L and S:S247R developed three times convergently in the most successful strains A7326, A6137 and A4707. Other substitutions developed twice in diverse strains, as shown in Table 5 and Fig. 3. A number of substitutions (n = 20) were already included in some or all of the strains at P0 and were conserved over the full distance of the ten passages (Fig. 3).

Not all strains developed aa substitutions, but we could prove substitutions that developed convergently in two or even three strains in the course of the serial passage (Table 3, Fig. 4). The two aa substitutions S:S247R and E:S6L evolved convergently in all three most successful strains. One additional convergent substitution is found in the region ORF1a, seven in the spike-coding region and one each in the regions ORF3a, M and ORF8. Interestingly, two of these independent mutational changes occurred at the prominent positions Spike 484 and Spike 501, which are well-known in connection with immune-escape variants. In our case, however, the changes were E484D (from glutamic acid E to aspartic acid D instead of lysine K in variants of concern) and N501T (from asparagine N to threonine T instead of tyrosine Y) (Fig. 5).

S:E484D developed independently three times in strains A818, A893 and A6137. Also the mutations S:S247R and E:S6L were observed to

Fig. 1. Maximum likelihood tree based on the WGS sequences of ten SARS-CoV-2 strains in the course of serial passages *in vitro*. Sequencing was carried out at passages 0, 3, 5, 7 and 10 respectively, and only high-quality sequences were used for further analysis, as shown here in 49 SARS-CoV-2 with 29,896 nucleotide sites each. Numbers at nodes indicate bootstrap support values (only values > 50 are shown). Bootstrap is also given by branch colour.

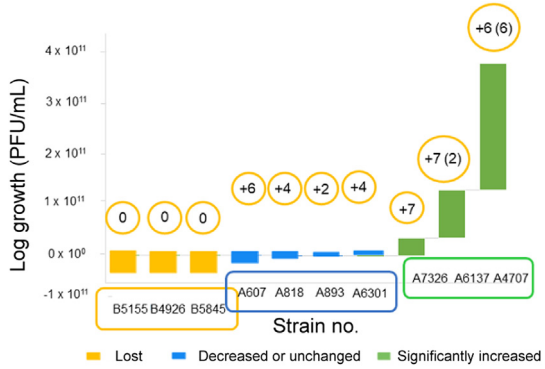

Fig. 2. Replication success of ten ancestral strains of SARS-CoV-2 over ten serial passages. The replication success is defined as the development of the viral load in the course of the experiment. All strains started with viral loads of 7.6×10^8. Three strains did not manage to replicate and got lost within the first three passages (yellow bars), four strains maintained in culture with a decreased or nearly unchanged viral load (blue bars), three strains showed significantly higher replication success (green bars). The bars show the growth factor of the diverse strains, calculated as the percentage growth of the strain during the experiment (PFU/mL at P10 × 100)/(PFU/mL at P0). The numbers above the bars represent the acquired aa substitutions during the serial passages. Numbers in parenthesis show additional temporary substitutions which did occur once in a sequence throughout the experiment but were not preserved. The first three strains, here marked with the first letter B, did not manage to replicate successfully in cell culture and were lost after at least four passages. Actual increases, thus successful adaptation to the cell cultures, were recorded in the six strains A818 to A4707, with the last three strains A7326, A6137 and A4707 achieving the highest increase in PFU/mL with growth of 4×10^3% up to 3×10^4%.

Table 4
The quantitative development of ten strains of SARS-CoV-2 after ten serial passages in Vero B4 cell culture flasks.

Strain	Ct P0	PFU/mL at P0 ($\times 10^8$)	Ct P10	PFU/mL at P10	Growth factor (%)
B5155	10	7.6	–	0	0
B4926	10	7.6	–	0	0
B5845	10	7.6	–	0	0
A607	10	7.6	24	1.3×10^3	2×10^{-4}
A818	10	7.6	16	2.6×10^6	3×10^{-1}
A893	10	7.6	14	1.7×10^7	2
A6301	10	7.6	12	2.6×10^6	15
A7326	10	7.6	6	3.4×10^{10}	4×10^3
A6137	10	7.6	5	8.7×10^{10}	1×10^4
A4707	10	7.6	4	2.3×10^{11}	3×10^4

Ct P0 and Ct P10 show the Ct values of the strains at passage 0 and 10, respectively, measured by RT-qPCR. The growth factor shows the percentage growth of the strain during the experiment (PFU/mL at P10 × 100)/(PFU/mL at P0).

develop convergently threefold in the highly successful strains A7326, A4707 and A6137. The following convergent substitutions were observed in two strains: S:N81T S:N501T, ORF1a:A1679V, M:H125Y and ORF8:P36L in strains A607 and A4707 and S:H66R, S:E96G, S:R1091C and ORF3a:M1I in strains A7326 and A4707.

Strains A607, A7326, A4707 and A6137 were those with the highest numbers of convergent substitutions and those with the highest evolutionary rate. None of the convergent substitutions arose in one of the unsuccessful strains which did not manage replication throughout the full length of the ten *in vitro* passages. All strains with convergent substitutions survived in all ten passages. The strains A7326, A4707 and A6137 were also those with the highest growth factor and could increase the starting number of PFL/mL by 79,000, 170,000 and 350,000%, respectively (Fig. 2).

3.3.3. Deletions

Only one aa deletion was found and concerned the lack of the amino acid at position 65 in the segment coding for ORF8 (ORF8:A65-). This deletion was detected in two strains (A818 and A6137).

3.3.4. Insertions

A noticeable change in the genetic structure of SARS-CoV-2 during serial passage *in vitro* came in the form of insertions. These additional nucleotide sequences developed in four of the ten strains in the course of the evolutionary adaptation to the cell culture. A detailed list of the strains with their individual insertions as well as the position and translation into the amino acid code is given in Table 6.

4. Discussion

In this study, we cultivated ten patient-derived Wuhan-near strains over ten *in vitro* passages. We inoculated SARS-CoV-2 onto a primate cell culture. To afford the virus with ample opportunity for adaptation, we consequently performed ten serial passages under constant conditions. Subsequently, we investigated the development of all ten strains using RT-qPCR and whole genome sequencing. The questions we were looking to answer were 1) whether all strains develop similarly 2) whether and how the environmental change was manifested genetically and 3) the genetic locations of the acquired mutations. Also, we were trying to further our understanding about each SARS-CoV-2 strain's potential to evolve into a variant of concern.

Interestingly, the development of the ten comparable ancestral strains was very different. Three have not been able to maintain replication over the first four passages and have diminished, four have been able to acclimate themselves more or less well and three have literally exploded *in vitro* both in the abundance of their genetic changes and in the viral load. That leaves us with the question of what characteristic defines the difference between the successful and unsuccessful strains over ten passages.

A comparison of the respective substitution rate per year *in vivo* an *in vitro* showed, on average, a slightly lower evolutionary rate *in vitro*. Nonetheless, the three most successful strains were able to catch up with the wildtype virus *in vivo* and achieved rates of 8×10^{-4} to 9×10^{-4} substitutions per site per year. The temporary mutations of the strains A6137 and A4707 were not included in this calculation of the substitution rates, which would increase the rate of evolution slightly. The twenty SNPs in comparison to the reference genome Wuhan-Hu-1, which the strains already had accumulated at the time of isolation, were maintained. A variety of substitutions occurred, and most of them in the three most successful strains, and none earlier than after four passages. Not all the strains developed substitutions. But there were two aa substitutions which occurred in all three highly successful strains. Other SNPs developed convergently in two or even three strains in the course of the serial passages. One of these convergent substitutions was found in the region ORF1a, seven in the spike-coding region and one each in the regions ORF3a, E, M and ORF8.

The SNPs which occurred convergently in three different strains were located in the genetic region coding for spike at amino acid positions 247 and 484, as well as position 6 of the region coding for the envelope proteins. SNP 247 is located in the N-terminal domaine (NTD) of spike 1 (S1) (Franco-Muñoz et al., 2020; Rohaim et al., 2021) and leads to the conversion of the original amino acid S serine to R arginine. The mutation on the position S:484 is a substitution in the RBD, which was described to impair neutralization resistance (Greaney et al., 2021). It is known to be an immune escape mutation when conversing from glutamine acid E to lysine K. This adaptation even developed four times. None of those were in any relation to each other. In our study, all four conversions occurred to aspartic acid D. They did however show their immune escape capability by resisting each of the four convalescent sera tested (Liu et al., 2021; Harvey et al., 2021). Another convergent point

Table 5
Overview of all acquired mutational changes during ten serial passages, consisting of aa substitutions, deletions and insertions.

	A607	A818	A893	A6301	A7326	A6137	A4707
ORF1a	ORF1a:A1679V	ORF1a:D1036E		ORF1a:E176G ORF1a:F924X*	ORF1a:A1352V	ORF1a:E673A ORF1a:S966F ORF1a:A1397V ORF1a:P1472S	ORF1a:A1352V ORF1a:A1679V*
ORF1b				ORF1b:E2072D		ORF1b:K82*	
Spike S	S:N81T **S:N501T**	S:L118F **S:E484D**	S:E180G **S:E484D**	**S:E484D**	S:H66R S:E96G S:S247R*** S:R1091C	S:H146R S:S247R*** **S:E484D**	S:H66R S:E96G S:S247R*** S:N81T* **S:N501T*** S:D215A*
	S: 22116-TAA...TAC-22117* S:214-RQSR-215	S:22206-GAG…AAT-22207		S:22199-TAC…GTT-22200			S:214-RQSR-215
ORF3a					ORF3a:M1I		ORF3a:M1I
E	E:V58A				E:S6L***	E:S6L***	E:S6L***
M	M:H125Y			M:S4F			M:H125Y*
ORF7a							
ORF8	ORF8:P36L	ORF8:E64D ORF8:A65-*				ORF8:A65-*	ORF8:P36L*
N							

Temporary mutations are written in grey with a *. Mutations in bold showed mutational hotspots which also occurred in the context of variants of concerns. The three strains A7326, A6137 and A4707 showed significantly higher replication success. Mutations which evolved convergently in all three most successful strains are marked with "***".

mutation occurred in the genomic region coding for the envelope E and was found in all three most successful strains of our experiment, A7326, A6137 and A4707. Five different substitutions that evolved twice in different strains were found in the region coding for spike: S:H66R, S:N81T, S:E96G, S:R1091C and notably, S:N501T. At exactly this position, an immune escape mutation has already been described in no less than four variants of concern. Those four are the Alpha (B.1.1.7), the Beta (B.1.135.1.), the Gamma variant (P1) and the Omicron variant (B.1.1.529). Interestingly again, in our case, the form of substitution was a different one. Asparagine N did not, as in the described cases, substitute to threonine T but to tyrosine Y. Further convergent SNPs were found at the positions ORF3a:M1I, membrane M:H125Y, ORF8:P36L and ORF1a:A1679V. Overall, the substitutions occurred in the regions coding for Spike, ORF1ab, E, M, ORF3a and ORF8.

In our experiment, ten out of 25 substitutions (40%), which were acquired after host change into cell culture, occurred in the genetic region coding for spike. The spike glycoprotein mediates viral entry into the host cell via the membrane-bound angiotensin-converting enzyme 2 (ACE2) and therefore, plays an important role in SARS-CoV-2 infectivity and transmissibility (Walls et al., 2020; Majumdar and Niyogi, 2021; Walls et al., 2020; Majumdar and Niyogi, 2021). This suggests that even in our artificial setting *in vitro*, where no large accumulation of active specific antibodies is to be expected, the spike region represents a mutational hotspot. Modifications in the amino acid structure can also present a decisive fitness advantage in this environment. The second most common substitutions were found in ORF1ab in our study (n = 9; 7 of it in ORF1a). ORF1ab (open reading frame) is subdivided into ORF1a and ORF1b and encodes for nine nonstructural proteins (Franco-Muñoz et al., 2020). The regions with the highest frequency of mutations were found in the sections coding for the non-structural protein (nsp) 2 and nsp3. They are substantive components for viral replication (Kim et al., 2020; Rohaim et al., 2021; Majumdar and Niyogi, 2021). The substitution at position ORF1a:A1679V is located in nsp3 and evolved three times independently in our experiment. Nsp3 is a large, multi-domain transmembrane protein with a number of activities. Those include the aforementioned participation in the viral replication, interaction with the nucleocapsid N protein, promotion of cytokine expression, cleavage of viral polyprotein and blocking of host innate immune response (Rohaim et al., 2021; Majumdar and Niyogi, 2021).

Several unique modifications were found in the ORF1a region of the most successful strain A4707, namely at positions E637A, S966F, A1397V and P1472S, all of which are located in the genetic regions coding for nsp2 and nsp3 (Majumdar and Niyogi, 2021). The remaining substitutions were located at the genomic regions of ORF3a, E, M and ORF8 in one or two strains in our experiment, respectively. E and M are structural proteins and encode the proteins for the viral envelope and membrane (Franco-Muñoz et al., 2020; Walls et al., 2020; Majumdar and Niyogi, 2021; Franco-Munoz et al., 2020; Walls et al., 2020, 2020; Majumdar and Niyogi, 2021, 2021). ORF3a and ORF8 are accessary proteins (Kim et al., 2020).

ORF3a is a sodium or calcium ion channel protein encoded by the region between S and E (Lu et al., 2006; Nieto-Torres et al., 2015). It is a lineage-specific accessory protein with three transmembrane domains and in combination with E it is required for viral replication and virulence (Castaño-Rodriguez et al., 2018; Siu et al., 2019). ORF8a encodes a membrane-associated structural protein with a single transmembrane domain (Guan et al., 2003). ORF8a has been proposed to enhance viral replication and induce apoptosis through a mitochondria-dependent pathway (Guan et al., 2003; Chen et al., 2011). Its topology and biological function, however, remain unknown. No mutational changes were observed in the regions coding for ORF3b, ORF6, ORF7ab, N, ORF9ab and ORF10, suggesting that there was no selective pressure enforcing modifications in these sections. To summarize, the single nucleotide substitutions took place mainly in the genetic region coding for spike, which was directly required for the entry into the target cell. These substitutions may therefore be considered directed mutations.

Furthermore, a large number of SNPs occurred in ORF1ab, a region that encodes for those non-structural proteins directly related to viral replication. The few other substitutions in our experiment occurred in genetic regions associated with viral replication.

Hitherto unknown insertions occurred in our experiment in four of the ten strains and comprised nine to twelve nucleotides. Two of the four insertions evolved divergently in two strains. These were strain A607, a primary infection and strain A4707, isolated from the same patient in the course of a febrile recurrence of the SARS-CoV-2 infections four weeks later.

Those two insertions were identified as codon-aligned insertions which have not been previously found (Singer et al., 2020). However, comparable insertions have been described at aa position 214/nucleotide position 22004 in wildtype SARS-CoV-2 strains: ins214AAG, ins214TDR and ins214KLGP (Garushyants et al., 2021) and were located at the N-terminal domain (NTD) of the spike protein. Due to the fact that these insertions evolved parallel in multiple strains at exactly the same position S:214 the authors suggest a role in infection. This goes hand in hand with the observation of multiple deletion variants in the same region, in particular

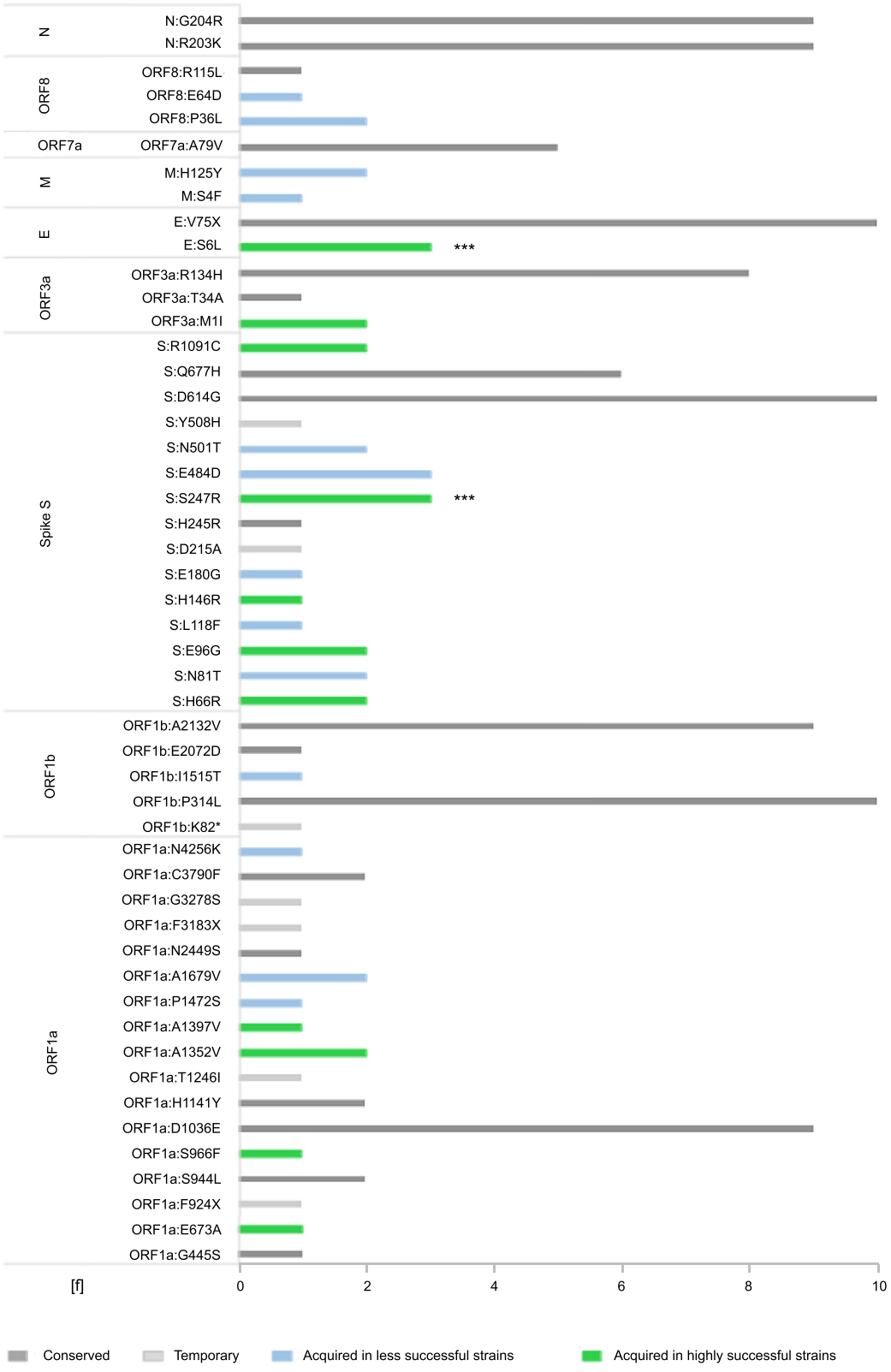

Fig. 3. A summary of all aa substitutions in SARS-CoV-2 in a study of the mutational dynamics of ten endemic strains of SARS-CoV-2 during serial passage. Conserved substitutions are those mutations acquired before isolation of the virus in cell culture and shown as dark-grey bars. All these conserved substitutions were sustained over the study period. Acquired substitutions evolved in the course of the serial passages. Acquired substitutions of less successful strains are shown in green, those of the highly successful strains in blue. Every temporary substitution could only be determined once in a sequence and disappeared in the course of the following passages. f, shows the frequency of each mutation (i.e., number of strains the mutation were found in). ***, convergent mutation in all three most successful strains.

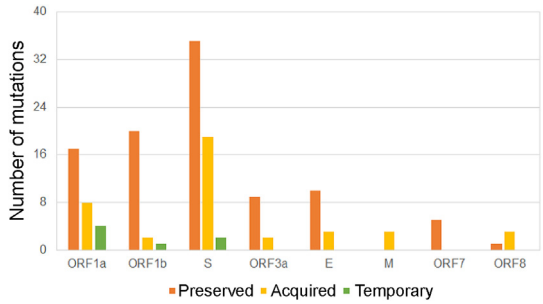

Fig. 4. Types of all acquired amino acid (aa) substitution in 11 coding sequences, acquired by SARS-CoV-2 strains in serial passages *in vitro*. This figure summarizes the distribution of aa substitutions on 11 coding sequences (ORF1a, ORF1b, S, ORF3a, E, M, ORF6, ORF7, ORF8, N, and ORF10). Orange bars indicate mutations acquired by the strains prior to isolation and preserved throughout the full range of serial passages, yellow bars indicate substitutions acquired by the strains throughout the *in vitro* serial passages, green bars show temporary mutations during one serial passage.

21971–22005 and does not rule out immune escape (McCarthy et al., 2021). Again, we found mutations of SARS-CoV-2 that were suspected to support immune escape but evolved *in vitro* in the absence of antibodies to force immune escape mutations. We therefore consider position S:214 to be a frequent adaptation of SARS-CoV-2 to Vero B4 cells. Furthermore, we assume that insertions in the genomic position S:214 can bring a fitness advantage in our experimental habitat, since the strains with the insertions are also the ones with the highest growth rate *in vitro*.

The different replication success of the strains A607 and A4707 is notable as these two strains were isolated coming from the same patient, once during initial infection and again during recurrence three weeks later. According to phylogenetic analysis based on whole genome sequences, both strains started with an apparently identical genetic framework. After ten *in vitro* passages, A4707 had developed into one of the most successful strains with an increase of replication by 3.5×10^{10}% (Table 4), while A607 was significantly weaker in replication. It was barely able to survive in the cell culture and showed no noticeable multiplication ($P < 0.05$). Only one genetic difference could be identified: A4707 had the insertion S:214-RQSR-215 at passage 1, whereas A607 started with a non-coding insertion in S:22116-AAAAAGTCTACT-22117, which was replaced by S:214-RQSR-215 not earlier than passage 7. In the course of the experiment, almost all genetic changes (5/7, 71.4%) that took place in the A607 strain were temporary mutations and were replaced by six novel substitutions in the strain A4707. Potentially, this indicates an association between the temporary fitness disadvantage of A607 *in vitro* and the two different spike insertions.

The evolutionary dynamics of RNA viruses in general are complex and may strongly be influenced by their high mutation rates, rapid replication kinetics, and large population sizes (Holland et al., 1982; Lauring and Andino, 2010). The low fidelity of the viral RNA polymerase ensures the constant development of point mutations, and theoretically, an immense number of different quasi-species is represented in a florid infection (Vignuzzi et al., 2005). Also, an RNA virus does not consist of a single strain but comprises a multitude of similar variants (Lauring and Andino, 2010). What becomes visible for us via whole genome sequencing is merely the master sequence which is surrounded by a

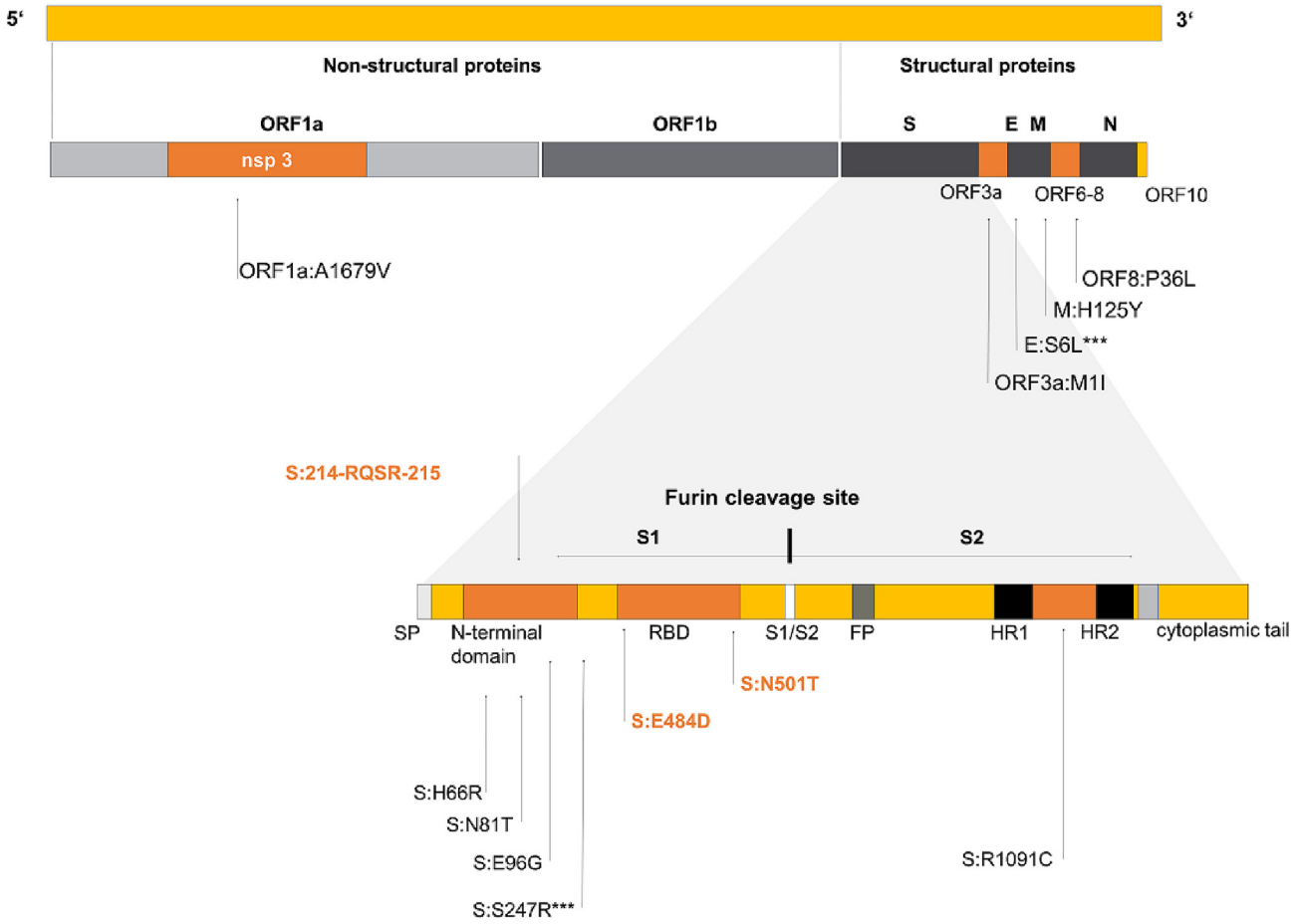

Fig. 5. Localization of the aa substitutions and insertions developed convergently from two or more strains during serial passage of ten SARS-CoV-2 strains *in vitro*. All convergently evolved mutations occurred in the regions coding for ORF1a/nsp3 (1); ORF3a (1), E (1), M (1), ORF8 (1) and spike (8). Substitutions marked with "***" evolved convergently in all three highly successful strains.

Table 6
Insertions that developed in the course of the experiment.

Strain	aa position	nc position	Sequence	Length	Type of insertion	Location
A607*	S:22116-AAA...ACT-22117	22216	AAAAAGTCTACT	12	Novel non-codon-aligned insertion	S1/NTD
A607	S:214-RQSR-215	22204	CGGCAGTCGAGA	12	Novel codon-aligned insertion	S1/NTD
A6301	S:22199-TAC…GTT-22200	22199	ACAGAAGAGGTT	12	Novel non-codon-aligned insertion	S1/NTD
A4707	S:214-RQSR-215	22204	CGGCAGTCGAGA	12	Novel codon-aligned insertion	S1/NTD
A818	S:22206-GAG…AAT-22207	22206	GAGAAGAAT	9	Novel non-codon-aligned insertion	S1/NTD

Four of ten endemic strains of SARS-CoV-2 developed 9 to 12mer nucleotide insertions. One strain (A607) developed two different insertions sequentially. The nucleotide sequences, positions and translations into amino acid codons are shown here.
The insertion labelled with a "*" is a temporary insertion, later replaced by S:214-RQSR-215.
nc, nucleotide; *length*, nucleotide length; *aa*, amino acid; *NTD*, N-terminal domaine; *S1*, spike 1.

collection of quasi-species with sometimes extensive genetic heterogeneity (Nowak, 1992; Biebricher and Eigen, 2006; Lauring and Andino, 2010). The comparable genetic evolution of A607 and A4707 *in vitro* is another hint towards the quasi-species theory. We assume that a high portion of the measured mutations have already been underrepresented quasi-species, arising in response to the predominant environmental conditions. We see further confirmation of this theory in the circumstance that the ten strains of our study evolved very differently. Not all strains in this experiment, in fact, only four of ten, had the potential to develop the necessary mutational adaptations to gain noteworthy fitness advantages.

5. Conclusions

Our experiment shows that SARS-CoV-2 is highly able to adapt to primate cell culture and changes in that environment. We were especially astonished by the rapidity of some strains' adaption as well as the noticeable increase in mutational change, especially in the coding region for spike, though not all strains apparently carried that same potential for intensive genetic adaption. Some even abandoned replication after a few passages. Others, isolated from the same patient at different times, developed surprisingly similarly and evolved several convergent mutations, until the end, their only difference was one insertion.

Since this study gives insights into the mutational adaptations of different strains of SARS-CoV-2 to a widely used type of cell culture, we believe it can provide a sound foundation for advanced studies concerning the evolutionary dynamics of SARS-CoV-2 and other RNA-viruses *in vitro*. Our experiments showed that high mutation rates occurred even in the absence of antibodies. These mutations can be substitutions as well as insertions and deletions. Furthermore, we found convergent mutations, indicating adaptation to the host change, in our case at the earliest at passage four. This should be kept in mind when cultivating SARS-CoV-2, e.g. in neutralization tests in BSL-3 laboratories. Thusly, we propose that very early passages or recently isolated strains of SARS-CoV-2 should be used in diagnostic applications. In summary, the observed genetic changes of SARS-CoV-2 *in vitro* cannot be explained solely by selectively neutral mutations but possibly result from the action of directional selection accumulating favourable genetic changes in the evolving variants, along the path of increasing potency of the strain. Competition among a high number of quasi-species in the SARS-CoV-2 *in vitro* population gene pool may reinforce directional selection and boost the speed of evolutionary change.

Data availability

The full-genome sequences obtained in the course of this study are available at GISAID EpiCoV (https://gisaid.org/; no.EPI_ISL_2106191-21061201; EPI-ISL-3128605, 627, 629, 630, 620).

Ethics statement

This article does not contain any studies with human or animal subjects performed by the author.

Author contributions

Sissy Therese Sonnleitner: conceptualization, data curation, formal analysis, funding acquisition, investigation, methodology; writing-original draft preparation. Stefanie Sonnleitner: methodology, writing-editing. Eva Hinterbichler: investigation, methodology. Hannah Halbfurter: project administration. Dominik B. C. Kopecky: visualization, organization. Stephan Koblmüller: conceptualization, writing-editing, funding acquisition. Christian Sturmbauer: conceptualization, writing-editing, funding acquisition. Wilfried Posch: validation, supervision. Gernot Walder: supervision.

Conflict of interest

The authors have no conflict of interest to declare.

Acknowledgements

We gratefully acknowledge the financial support of the Austrian Research Promotion Agency (FFG), Grant No. 35863961. The authors report no potential conflict of interest. We also thank Ramona Polster, MSc, for excellent technical assistance, DDI Martin Lamprecht for valuable suggestions and all the kind colleagues from the routine diagnostics team of the Dr. Gernot Walder laboratory for great support.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.virs.2022.01.029.

References

Andreano, E., Piccini, G., Licastro, D., Casalino, L., Johnson, N.V., Paciello, I., Dal Monego, S., Pantano, E., Manganaro, N., Manenti, A., Manna, R., Casa, E., Hyseni, I., Benincasa, L., Montomoli, E., Amaro, R.E., McLellan, J.S., Rappuoli, R., 2020. SARS-CoV-2 Escape bioRxiv.

Badua, C.L.D.C., Baldo, K.A.T., Medina, P.M.B., 2021. Genomic and proteomic mutation landscapes of SARS-CoV-2. J. Med. Virol. 93, 1702–1721.

Biebricher, C.K., Eigen, M., 2006. What is a quasispecies? Curr. Top. Microbiol. Immunol. 299, 1–31.

Brouwer, P.J.M., Caniels, T.G., van der Straten, K., Snitselaar, J.L., Aldon, Y., Bangaru, S., Torres, J.L., Okba, N.M.A., Claireaux, M., Kerster, G., Bentlage, A.E.H., van Haaren, M.M., Guerra, D., Burger, J.A., Schermer, E.E., Verheul, K.D., van der Velde, N., van der Kooi, A., van Schooten, J., van Breemen, M.J., Bijl, T.P.L., Sliepen, K., Aartse, A., Derking, R., Bontjer, I., Kootstra, N.A., Wiersinga, W.J., Vidarsson, G., Haagmans, B.L., Ward, A.B., de Bree, G.J., Sanders, R.W., van Gils, M.J., 2020. Potent neutralizing antibodies from COVID-19 patients define multiple targets of vulnerability. Science 369, 643–650.

Cao, Y., Su, B., Guo, X., Sun, W., Deng, Y., Bao, L., Zhu, Q., Zhang, X., Zheng, Y., Geng, C., Chai, X., He, R., Li, X., Lv, Q., Zhu, H., Deng, W., Xu, Y., Wang, Y., Qiao, L., Tan, Y., Song, L., Wang, G., Du, X., Gao, N., Liu, J., Xiao, J., Su, X.D., Du, Z., Feng, Y., Qin, C., Jin, R., Xie, X.S., 2020. Potent neutralizing antibodies against SARS-CoV-2 identified by high-throughput single-cell sequencing of convalescent patients' B cells. Cell 182, 73–84 e16.

Castaño-Rodriguez, C., Honrubia, J.M., Gutiérrez-Álvarez, J., DeDiego, M.L., Nieto-Torres, J.L., Jimenez-Guardeño, J.M., Regla-Nava, J.A., Fernandez-Delgado, R., Verdia-Báguena, C., Queralt-Martín, M., Kochan, G., Perlman, S., Aguilella, V.M., Sola, I., Enjuanes, L., 2018. Role of Severe acute respiratory syndrome coronavirus viroporins E, 3a, and 8a in replication and pathogenesis. mBio 9, e02325–17.

Chen, C.C., Krüger, J., Sramala, I., Hsu, H.J., Henklein, P., Chen, Y.M., Fischer, W.B., 2011. ORF8a of SARS-CoV forms an ion channel: experiments and molecular dynamics simulations. Biochim. Biophys. Acta 1808, 572–579.
Chrisman, B.S., Paskov, K., Stockham, N., Tabatabaei, K., Jung, J.Y., Washington, P., Varma, M., Sun, M.W., Maleki, S., Wall, D.P., 2021. Indels in SARS-CoV-2 occur at template-switching hotspots. BioData Min. 14, 20.
Elena, S.F., Sanjuán, R., 2005. Adaptive value of high mutation rates of RNA viruses: separating causes from consequences. J. Virol. 79, 11555–11558.
Franco-Muñoz, C., Álvarez-Díaz, D.A., Laiton-Donato, K., Wiesner, M., Escandón, P., Usme-Ciro, J.A., Franco-Sierra, N.D., Flórez-Sánchez, A.C., Gómez-Rangel, S., Rodríguez-Calderon, L.D., Barbosa-Ramirez, J., Ospitia-Baez, E., Walteros, D.M., Ospina-Martinez, M.L., Mercado-Reyes, M., 2020. Substitutions in spike and nucleocapsid proteins of SARS-CoV-2 circulating in South America. Infect. Genet. Evol. 85, 104557.
Garushyants, S.K., Rogozin, I.B., Koonin, E.V., 2021. Insertions in SARS-CoV-2 Genome Caused by Template Switch and Duplications Give Rise to New Variants that Merit Monitoring. bioRxiv.
Gietl, S., Schönegger, C.M., Falk, M., Weiler, S., Obererlacher, S., Jansen, B., Sonnleitner, S.T., Walder, G., 2021. Home quarantine in COVID-19: a study on 50 consecutive patients in Austria. Clin. Med. 21, e9–e13.
Greaney, A.J., Loes, A.N., Crawford, K.H.D., Starr, T.N., Malone, K.D., Chu, H.Y., Bloom, J.D., 2021. Comprehensive mapping of mutations in the SARS-CoV-2 receptor-binding domain that affect recognition by polyclonal human plasma antibodies. Cell Host Microbe 29, 463–476 e6.
Guan, Y., Zheng, B.J., He, Y.Q., Liu, X.L., Zhuang, Z.X., Cheung, C.L., Luo, S.W., Li, P.H., Zhang, L.J., Guan, Y.J., Butt, K.M., Wong, K.L., Chan, K.W., Lim, W., Shortridge, K.F., Yuen, K.Y., Peiris, J.S., Poon, L.L., 2003. Isolation and characterization of viruses related to the SARS coronavirus from animals in southern China. Science 302, 276–278.
Harvey, W.T., Carabelli, A.M., Jackson, B., Gupta, R.K., Thomson, E.C., Harrison, E.M., Ludden, C., Reeve, R., Rambaut, A., Peacock, S.J., Robertson, D.L., Consortium, C.-G.U.C.-U., 2021. SARS-CoV-2 variants, spike mutations and immune escape. Nat. Rev. Microbiol. 19, 409–424.
Hoang, D.T., Chernomor, O., von Haeseler, A., Minh, B.Q., Vinh, L.S., 2018. UFBoot2: improving the ultrafast bootstrap approximation. Mol. Biol. Evol. 35, 518–522.
Holland, J., Spindler, K., Horodyski, F., Grabau, E., Nichol, S., VandePol, S., 1982. Rapid evolution of RNA genomes. Science 215, 1577–1585.
Ju, B., Zhang, Q., Ge, J., Wang, R., Sun, J., Ge, X., Yu, J., Shan, S., Zhou, B., Song, S., Tang, X., Lan, J., Yuan, J., Wang, H., Zhao, J., Zhang, S., Wang, Y., Shi, X., Liu, L., Wang, X., Zhang, Z., Zhang, L., 2020. Human neutralizing antibodies elicited by SARS-CoV-2 infection. Nature 584, 115–119.
Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., von Haeseler, A., Jermiin, L.S., 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat. Methods 14, 587–589.
Kim, J.S., Jang, J.H., Kim, J.M., Chung, Y.S., Yoo, C.K., Han, M.G., 2020. Genome-wide identification and characterization of point mutations in the SARS-CoV-2 genome. Osong Public Health Res. Perspect. 11, 101–111.
Koenig, P.A., Das, H., Liu, H., Kümmerer, B.M., Gohr, F.N., Jenster, L.M., Schiffelers, L.D.J., Tesfamariam, Y.M., Uchima, M., Wuerth, J.D., Gatterdam, K., Ruetalo, N., Christensen, M.H., Fandrey, C.I., Normann, S., Tödtmann, J.M.P., Pritzl, S., Hanke, L., Boos, J., Yuan, M., Zhu, X., Schmid-Burgk, J.L., Kato, H., Schindler, M., Wilson, I.A., Geyer, M., Ludwig, K.U., Hällberg, B.M., Wu, N.C., Schmidt, F.I., 2021. Structure-guided multivalent nanobodies block SARS-CoV-2 infection and suppress mutational escape. Science 371, eabe6230.
Kumar, S., Stecher, G., Li, M., Knyaz, C., Tamura, K., 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol. Biol. Evol. 35, 1547–1549.
Lauring, A.S., Andino, R., 2010. Quasispecies theory and the behavior of RNA viruses. PLoS Pathog. 6, e1001005.
Liu, L., Wang, P., Nair, M.S., Yu, J., Rapp, M., Wang, Q., Luo, Y., Chan, J.F., Sahi, V., Figueroa, A., Guo, X.V., Cerutti, G., Bimela, J., Gorman, J., Zhou, T., Chen, Z., Yuen, K.Y., Kwong, P.D., Sodroski, J.G., Yin, M.T., Sheng, Z., Huang, Y., Shapiro, L., Ho, D.D., 2020. Potent neutralizing antibodies against multiple epitopes on SARS-CoV-2 spike. Nature 584, 450–456.
Liu, Z., VanBlargan, L.A., Bloyet, L.M., Rothlauf, P.W., Chen, R.E., Stumpf, S., Zhao, H., Errico, J.M., Theel, E.S., Liebeskind, M.J., Alford, B., Buchser, W.J., Ellebedy, A.H., Fremont, D.H., Diamond, M.S., Whelan, S.P.J., 2021. Identification of SARS-CoV-2 spike mutations that attenuate monoclonal and serum antibody neutralization. Cell Host Microbe 29, 477–488 e4.
Lu, W., Zheng, B.J., Xu, K., Schwarz, W., Du, L., Wong, C.K., Chen, J., Duan, S., Deubel, V., Sun, B., 2006. Severe acute respiratory syndrome-associated coronavirus 3a protein forms an ion channel and modulates virus release. Proc. Natl. Acad. Sci. U. S. A. 103, 12540–12545.
Majumdar, P., Niyogi, S., 2021. SARS-CoV-2 mutations: the biological trackway towards viral fitness. Epidemiol. Infect. 149, e110.
McCarthy, K.R., Rennick, L.J., Nambulli, S., Robinson-McCarthy, L.R., Bain, W.G., Haidar, G., Duprex, W.P., 2021. Recurrent deletions in the SARS-CoV-2 spike glycoprotein drive antibody escape. Science 371, 1139–1142.
Nguyen, L.T., Schmidt, H.A., von Haeseler, A., Minh, B.Q., 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol. Evol. 32, 268–274.
Nieto-Torres, J.L., Verdiá-Báguena, C., Jimenez-Guardeño, J.M., Regla-Nava, J.A., Castaño-Rodriguez, C., Fernandez-Delgado, R., Torres, J., Aguilella, V.M., Enjuanes, L., 2015. Severe acute respiratory syndrome coronavirus E protein transports calcium ions and activates the NLRP3 inflammasome. Virology 485, 330–339.
Nowak, M.A., 1992. What is a quasispecies? Trends Ecol. Evol. 7, 118–121. https://doi.org/10.1016/0169-5347(92)90145-2.
Okonechnikov, K., Golosova, O., Fursov, M., team, U., 2012. Unipro UGENE: a unified bioinformatics toolkit. Bioinformatics 28, 1166–1167.
Panzera, Y., Ramos, N., Frabasile, S., Calleros, L., Marandino, A., Tomás, G., Techera, C., Grecco, S., Fuques, E., Goñi, N., Ramas, V., Coppola, L., Chiparelli, H., Sorhouet, C., Mogdasy, C., Arbiza, J., Delfraro, A., Pérez, R., 2021. A deletion in SARS-CoV-2 ORF7 identified in COVID-19 outbreak in Uruguay. Transbound Emerg. Dis. 68, 3075–3082.
Rogers, T.F., Zhao, F., Huang, D., Beutler, N., Burns, A., He, W.T., Limbo, O., Smith, C., Song, G., Woehl, J., Yang, L., Abbott, R.K., Callaghan, S., Garcia, E., Hurtado, J., Parren, M., Peng, L., Ramirez, S., Ricketts, J., Ricciardi, M.J., Rawlings, S.A., Wu, N.C., Yuan, M., Smith, D.M., Nemazee, D., Teijaro, J.R., Voss, J.E., Wilson, I.A., Andrabi, R., Briney, B., Landais, E., Sok, D., Jardine, J.G., Burton, D.R., 2020. Isolation of potent SARS-CoV-2 neutralizing antibodies and protection from disease in a small animal model. Science 369, 956–963.
Rohaim, M.A., El Naggar, R.F., Clayton, E., Munir, M., 2021. Structural and functional insights into non-structural proteins of coronaviruses. Microb. Pathog. 150, 104641.
Salinas, N.R., Little, D.P., 2014. 2matrix: a utility for indel coding and phylogenetic matrix concatenation(1.). Appl. Plant Sci. 2, apps.1300083.
Seydoux, E., Homad, L.J., MacCamy, A.J., Parks, K.R., Hurlburt, N.K., Jennewein, M.F., Akins, N.R., Stuart, A.B., Wan, Y.H., Feng, J., Whaley, R.E., Singh, S., Boeckh, M., Cohen, K.W., McElrath, M.J., Englund, J.A., Chu, H.Y., Pancera, M., McGuire, A.T., Stamatatos, L., 2020. Analysis of a SARS-CoV-2-infected individual reveals development of potent neutralizing antibodies with limited somatic mutation. Immunity 53, 98–105 e5.
Simmons, M.P., Ochoterena, H., 2000. Gaps as characters in sequence-based phylogenetic analyses. Syst. Biol. 49, 369–381.
Singer, J., Gifford, R., Cotten, M., Robertson, D., 2020. CoV-GLUE: A Web Application for Tracking SARS-CoV-2 Genomic Variation. https://doi.org/10.20944/preprints202006.0225.v1.
Siu, K.L., Yuen, K.S., Castaño-Rodriguez, C., Ye, Z.W., Yeung, M.L., Fung, S.Y., Yuan, S., Chan, C.P., Yuen, K.Y., Enjuanes, L., Jin, D.Y., 2019. Severe acute respiratory syndrome coronavirus ORF3a protein activates the NLRP3 inflammasome by promoting TRAF3-dependent ubiquitination of ASC. Faseb. J. 33, 8865–8877.
Sonnleitner, S.T., Dorighi, J., Jansen, B., Schönegger, C., Gietl, S., Koblmüller, S., Sturmbauer, C., Posch, W., Walder, G., 2021. An in vitro model for assessment of SARS-CoV-2 infectivity by defining the correlation between virus isolation and quantitative PCR value: isolation success of SARS-CoV-2 from oropharyngeal swabs correlates negatively with Cq value. Virol. J. 18, 71.
Verkhivker, G.M., Agajanian, S., Oztas, D.Y., Gupta, G., 2021. Comparative perturbation-based modeling of the SARS-CoV-2 spike protein binding with host receptor and neutralizing antibodies: structurally adaptable Allosteric communication hotspots define spike sites targeted by global circulating mutations. Biochemistry 60, 1459–1484.
Vignuzzi, M., Stone, J.K., Andino, R., 2005. Ribavirin and lethal mutagenesis of poliovirus: molecular mechanisms, resistance and biological implications. Virus Res. 107, 173–181.
Walls, A.C., Park, Y.J., Tortorici, M.A., Wall, A., McGuire, A.T., Veesler, D., 2020. Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. Cell 183, 1735.
Wec, A.Z., Wrapp, D., Herbert, A.S., Maurer, D.P., Haslwanter, D., Sakharkar, M., Jangra, R.K., Dieterle, M.E., Lilov, A., Huang, D., Tse, L.V., Johnson, N.V., Hsieh, C.L., Wang, N., Nett, J.H., Champney, E., Burnina, I., Brown, M., Lin, S., Sinclair, M., Johnson, C., Pudi, S., Bortz, R., Wirchnianski, A.S., Laudermilch, E., Florez, C., Fels, J.M., O'Brien, C.M., Graham, B.S., Nemazee, D., Burton, D.R., Baric, R.S., Voss, J.E., Chandran, K., Dye, J.M., McLellan, J.S., Walker, L.M., 2020. Broad neutralization of SARS-related viruses by human monoclonal antibodies. Science 369, 731–736.
Wu, Y., Wang, F., Shen, C., Peng, W., Li, D., Zhao, C., Li, Z., Li, S., Bi, Y., Yang, Y., Gong, Y., Xiao, H., Fan, Z., Tan, S., Wu, G., Tan, W., Lu, X., Fan, C., Wang, Q., Liu, Y., Zhang, C., Qi, J., Gao, G.F., Gao, F., Liu, L., 2020. A noncompeting pair of human neutralizing antibodies block COVID-19 virus binding to its receptor ACE2. Science 368, 1274–1278.
Zhang, D., Gao, F., Jakovlić, I., Zou, H., Zhang, J., Li, W.X., Wang, G.T., 2020. PhyloSuite: an integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. Mol. Ecol. Resour. 20, 348–355.
Zheng, J., 2020. SARS-CoV-2: an emerging coronavirus that causes a global threat. Int. J. Biol. Sci. 16, 1678–1685.
Zost, S.J., Gilchuk, P., Case, J.B., Binshtein, E., Chen, R.E., Nkolola, J.P., Schäfer, A., Reidy, J.X., Trivette, A., Nargi, R.S., Sutton, R.E., Suryadevara, N., Martinez, D.R., Williamson, L.E., Chen, E.C., Jones, T., Day, S., Myers, L., Hassan, A.O., Kafai, N.M., Winkler, E.S., Fox, J.M., Shrihari, S., Mueller, B.K., Meiler, J., Chandrashekar, A., Mercado, N.B., Steinhardt, J.J., Ren, K., Loo, Y.M., Kallewaard, N.L., McCune, B.T., Keeler, S.P., Holtzman, M.J., Barouch, D.H., Gralinski, L.E., Baric, R.S., Thackray, L.B., Diamond, M.S., Carnahan, R.H., Crowe, J.E., 2020. Potently neutralizing and protective human antibodies against SARS-CoV-2. Nature 584, 443–449.