Spike protein cleavage-activation mediated by the SARS-CoV-2 P681R mutation: a case-study from its first appearance in variant of interest (VOI) A.23.1 identified in Uganda

Bailey Lubinski^{1,2 #}, Laura E. Frazier^{1,2 #}, My V.T. Phan³, Daniel L. Bugembe³, Tiffany Tang⁴, Susan Daniel⁴, Matthew Cotten^{3,5}, Javier A. Jaimes^{2 *} and Gary R. Whittaker^{2, 6 *}

¹ Graduate Program in Biological & Biomedical Sciences, Cornell University, Ithaca, NY, 14853, USA.

² Department of Microbiology & Immunology, College of Veterinary Medicine, Cornell University, Ithaca, NY, 14853, USA.

³ MRC/UVRI & London School of Hygiene and Tropical Medicine – Uganda Research Unit, Entebbe, Uganda.

⁴ Robert Frederick Smith School of Chemical & Biomolecular Engineering, Cornell University, Ithaca, NY, 14853, USA.

⁵ MRC Centre of Virus Research, University of Glasgow, Glasgow, United Kingdom.

⁶ Master of Public Health Program, Cornell University, Ithaca, NY, 14853, USA.

[#] Contributed equally

^{*} Corresponding authors

618 Tower Rd., Ithaca NY 14853, USA

jaj246@cornell.edu; grw7@cornell.edu

Graphical abstract

Abstract

The African continent like all other parts of the world with high infection/low vaccination rates can, and will, be a source of novel SARS-CoV-2 variants. The A.23 viral lineage, characterized by three spike mutations F157L, V367F and Q613H, was first identified in COVID-19 cases from a Ugandan prison in July 2020, and then was identified in the general population with the additional spike mutation P681R at the S1/S2 cleavage site to comprise lineage A.23.1 by September 2020 with subsequent spread to 26 other countries. The P681R spike substitution of A.23.1 is of note as it increases the number of basic residues in the sub-optimal SARS-CoV-2 spike protein furin cleavage site; as such, this substitution may affect viral replication, transmissibility, or pathogenic properties. The same P681R substitution has also subsequently appeared in B.1.617 variants, including B.1.617.2 (Delta). Here, we performed assays using fluorogenic

peptides mimicking the S1/S2 from A.23.1 and B.1.617 and observed significantly increased cleavability with furin, compared to sequences derived from the original Wuhan-Hu1 S1/S2. We performed cell-cell fusion and functional infectivity assays using pseudotyped particles harboring SARS-CoV-2 spike proteins and observed an increase in transduction for A.23.1-pseudotyped particles compared to Wuhan-Hu-1. However, these changes in activity were not reproduced in the original Wuhan-Hu-1 spike bearing only the P681R substitution. Our findings suggest that while A.23.1 has increased furin-mediated cleavage linked to the P681R substitution—which may affect viral infection and transmissibility—this substitution alone needs to occur on the background of other spike protein changes to enable its full functional consequences.

Key words:

SARS-CoV-2, A.23.1 variant, Uganda, spike, P681R mutation

Introduction

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is the agent causing the current COVID-19 pandemic (1). SARS-CoV-2 was first identified in late 2019 and has since spread rapidly throughout the world. The virus has evolved into two main lineages, A and B and multiple additional lineages. While the B.1 lineage became the dominant virus following its introduction into Northern Italy and spread through Europe/UK in February 2020, both A and B lineages remain in circulation globally (2). Both lineages have undergone significant diversification as they expanded; this expansion is apparently linked to a key S gene mutation—D614G in lineage B.1 and all sublineages, which has been linked to modest increase in virus transmissibility (3) and with Q613H found in lineage A.23/A.23.1. As Q613H is adjacent to D614G it may represent an example of convergent evolution that resulted in a more stabilized spike protein (4). D614G has now become established in circulating B and derived lineages. Compared with the lineage B.1 viruses that have successfully evolved into multiple variants of concern (VOC), including B.1.1.7 (Alpha) B.1.351 (Beta), B.1.1.28.1/P.1 (Gamma), B.1.617.2 (Delta), most of lineage A viruses remained at fairly lower

frequency and were more prevalent at the beginning of the pandemic in Asia. However, the A.23.1 viral lineage is one of a few lineage A viruses that, due to local circumstances, became abundant in Uganda (5), Rwanda (6) and South Sudan (7). A.23.1 evolved from the A.23 virus variant first identified in Uganda in July 2020 and is characterized by three spike mutations F157L, V367F and Q613H (8). Subsequently, the evolving A.23.1 lineage acquired additional spike substitutions (P681R), as well as in nsp6, ORF8 and ORF9 and with the acquisition of the E484K substitution A.23.1 was designated a variant under investigation (VUI). By July 2021, the A.23.1 lineage has been observed with 1110 genomes reported from 47 countries (GISAID, Pango Lineage report; https://cov-lineages.org/global_report_A.23.1.html).

Among several mutations in the A.23.1 lineage, the P681R mutation is of interest as it is part of a proteolytic cleavage site for furin and furin-like proteases at the junction of the spike protein receptor-binding (S1) and fusion (S2) domains (9). The S1/S2 junction of the SARS-CoV-2 S gene has a distinct indel compared to all other known SARS-like viruses (Sarbecoviruses in *Betacoronavirus* lineage B)—the amino acid sequence of SARS-CoV-2 S protein is ₆₈₁-P-R-R-A-R|S-₆₈₆ with proteolytic cleavage (|) predicted to occur between the arginine and serine residues depicted. Based on nomenclature established for proteolytic events (10), the R|S residues are defined as the P1|P1' residues for enzymatic cleavage, with residue 681 of A.23.1 spike being the P5 cleavage position. The ubiquitously-expressed cellular serine protease furin is highly specific and cleaves at a distinct multi-basic motif containing paired arginine residues; furin requires a minimal motif of R-x-x-R (P4-x-x-P1), with a preference for an additional basic (B) residue at P2; i.e., R-x-B-R (11). For SARS-CoV-2, the presence of the S1/S2 “furin site” enhances virus transmissibility (12, 13). For the A.23.1 S, P681R provides an additional basic residue at P5 and may modulate S1/S2 cleavability by furin, and hence virus infection properties (14). Notably, the P681R substitution appears in several other lineages, most notably B.1.617.2 (Delta)(N = 617590 genomes) but also the AY.X sublineages of B.1.617.2, B.1.617.1 (N= 6138), B.1.466.2 (N=2208), the B.1.1.7 sublineage Q.4 (2067), B.1.551 (N=722), AU.2 (N=302), B.1.1.25 (N=509), B.1.466.2 (N=538), and other lineages (updated 22 Oct 21, <https://outbreak.info>;

<https://outbreak.info/situation-reports?pango&mut=s%3AP681R>) suggesting that the substitution may provide an advantage for viruses encoding the substitution.

We previously studied the role of proteolytic activation of the spike protein of the lineage B SARS-CoV-2 isolates Wuhan-Hu1 and B.1.1.7 (15). Here, we used a similar approach to study the role of the proteolytic activation of the spike protein in the context of the A.23.1 lineage virus, with a focus on the P681R substitution to better understand the role of this notable change at the S1/S2 (furin) cleavage site.

Results

Emergence and analysis of SARS-CoV-2 variants in Uganda

A summary of the daily reported SARS-CoV-2 infections in Uganda is shown in Figure 1A, along with a summary of SARS-CoV-2 lineage data in samples from Uganda (Figure 1B). The peak of infections in December 2020-January 2021 corresponded to the circulation of the A.23.1 variant, which subsided, but was replaced by a second larger peak of infections beginning in July 2021, primarily due to the emergence of the B.1.617.2 variant (Delta), with additional variants being present over time. The circulation of the A.23.1 variant appears to be fully displaced by the B.1.617.2 variant, which by July 2021 became the prevalent variant in this country.

Bioinformatic and biochemical analysis of the SARS-CoV-2 A.23.1 S1/S2 cleavage site

Along with other changes in the spike protein, A.23.1, B.1.617.1 (Kappa) and B.1.617.2 (Delta) contain a P681R substitution at the S1/S2 interface (Figure 2A), which may modulate spike protein function based on changes to the furin cleavage site and downstream changes in the levels of cleaved products and in virus-cell fusion activation. To gain insight into proteolytic processing at the S1/S2 site of the SARS-CoV-2 spike protein, we first took a bioinformatic approach utilizing the PiTou (16) and ProP (17) cleavage prediction tools, comparing the spike proteins from A.23.1 to B.1.1.7 and the prototype SARS-CoV-2 from the A.1 and B.1 lineages (e.g. Wuhan-Hu-1), as well as to MERS-CoV and selected other human respiratory

betacoronaviruses (HCoV-HKU1 and HCoV-OC43) with identifiable furin cleavage sites (Figure 2B). Both algorithms predicted an increase in the furin cleavage for the A.23.1 and B.1.617 lineages compared to A.1/B.1, with PiTou also showing a marked increase compared to B.1.1.7. PiTou utilizes a hidden Markov model specifically targeting 20 amino acid residues surrounding furin cleavage sites and is expected to be more accurate prediction tool. As expected, MERS-CoV showed a relatively low furin cleavage score, with HCoV-HKU1 and HCoV-OC43 showing relatively high furin cleavage scores. Overall, these analyses predict a distinct increase of furin cleavability for the spike protein of A.23.1 and B.1.617 lineage viruses compared to A.1/B.1 and B.1.1.7. lineage viruses.

To directly examine the activity of furin on the SARS-CoV-2 A.23.1 S1/S2 site, we used a biochemical peptide cleavage assay to directly measure furin cleavage activity *in vitro* (18). The specific peptide sequences used here were SARS-CoV-2 S1/S2 B.1.1.7 (TNSHRRARSVA), TNSPRRARSVA (Wuhan-Hu-1 S1/S2) and TNSRRRARSVA (A.23.1 S1/S2). As predicted, furin effectively cleaved both the Wuhan-Hu-1 (WT) and B.1.1.7 peptides, but with no significant differences (Figure 2C). Interestingly, and agreeing with the PiTou prediction, we observed a significant increase in furin cleavage for the A.23.1 S1/S2 peptide (Figures 2B and C) compared to both Wuhan-Hu-1 (WT) and B.1.1.7. This comparative data with SARS-CoV S1/S2 sites reveals that the P618R substitution substantially increases cleavability by furin, beyond the small increase noted previously for P681H (12).

Cell-to-cell fusion assays of A.23.1 spike

In order to assess the functional properties of the spike protein and to see if the P681R substitution provided any advantage for cell-to-cell transmission or syncytia formation, we performed a cell-to-cell fusion assay in which VeroE6 or Vero-TMPRSS2 cells were transfected with either the WT, A.23.1 or P681R spike gene. We then evaluated syncytia formation as a read-out of membrane fusion. We observed an increase in the syncytia formation following spike protein expression for either A.23.1 or Wuhan-Hu-1 harboring a P681R mutation (P681R), compared to Wuhan-Hu-1 (WT) (Figure 3A). Vero-TMPRSS2 cells generally formed

more extensive syncytia than VeroE6 cells, with similar trends observed. This increase was evident by observation through fluorescence microscopy, as well as by quantification of the syncytia and cell-to-cell fusion ratio (Figure 3A, B and C). An increase in the number of nuclei involved in syncytia was observed in cells transfected with A.23.1 and P681R S genes in both cell lines (Figure 3B). However, the increase was higher in Vero-TMPRSS2 cells. We also observed an increase in the cell-to-cell fusion ratio (quantified by fluorescence intensity) in A.23.1 and P681R spike transfected cells, compared to WT (Figure 3C) and as with the previous analysis, the ratio was higher in Vero-TMPRSS2 cells. Membrane expressed spike cleavage also assessed using western blot (Figure 3D). An increased cleavage ratio was observed in the A.23.1 and P681R membrane expressed spikes, compared to WT. Interestingly, the cleavage ratio was similar in both cell lines. Band intensity was normalized to the GLUT4 protein (housekeeping protein) band intensity. These data provide evidence that the P681R mutation increases membrane fusion activity of the SARS-CoV-2 spike protein under the conditions tested.

Functional analysis of virus entry using viral pseudoparticles

To assess the functional importance of the S1/S2 site for SARS-CoV-2 entry, we utilized viral pseudoparticles consisting of a murine leukemia virus (MLV) core displaying a heterologous viral envelope protein to partially recapitulate the entry process of the native coronavirus. The pseudoparticles also contain a luciferase reporter gene as well as the integrase activity to allow that integration into the host cell genome to drive expression of quantifiable luciferase (MLVpp-SARS-CoV-2 S) (19). Using the HEK-293T cell line for particle production, MLV pseudoparticles containing the spike proteins of A.23.1, Wuhan-Hu-1 SARS-CoV-2 (WT), and a P681R point mutant of Wuhan-Hu-1 (P681R) were prepared. Positive-control particles containing the vesicular stomatitis virus (VSV) G protein and negative-control particles ($\Delta envpp$) lacking envelope proteins were also prepared (not shown). The pseudoparticles were probed for their S content via western blot (Figure 4A). Because SARS-CoV-S has an efficiently cleaved spike we also produced particles under furin inhibition, using dec-RVKR-CMK which has previously shown to be effective inhibiting furin

and other proprotein convertases. This allows changes in cleavage patterns between different spike proteins to be visualized. We also treated the particles with exogenous furin (+ or – furin) to examine the spike cleavage on both partially cleaved and uncleaved spikes, to further study the differences in furin processing in the studied spikes. For both A.23.1 and P681R particles, we detected increased spike protein cleavage compared to WT in the harvested particles (Figure 4A). Interestingly, we observed markedly increased cleavage ratio for both A.23.1 and P681R spikes in the harvested pseudoparticles under furin-inhibition conditions (Figure 4A).

We next examined the infection properties of SARS-CoV-2 pseudoparticles in cell lines representative of both the “early” and “late” cell entry pathways (Figure 4B). Furin is predicted to cleave during virus assembly and “prime” the spike protein at the S1/S2 site for subsequent events during cell entry. However, a subsequent cleavage priming at a secondary site (known as S2’ site) is also needed to activate the spike’s fusion machinery (1). SARS-CoV-2 is predicted to enter Vero E6 cells using cathepsin L for S2’ priming during endosomal trafficking, in what is known as the “late” pathway, whereas in Vero-TMPRSS2 is predicted to use a “early” pathway, with the spike’s S2’ site being activated by TMPRSS2 or other transmembrane serine proteases (TSPs) at the cellular membrane. (1). Specifically, we used the Vero-TMPRSS2 (“early pathway”) and the Vero E6 (“late pathway”) cell lines, which are predicted to activate the SARS-CoV-2 S2’ using TMPRSS2 and cathepsin L respectively. Considering that furin priming at the S1/S2 site normally occurs during viral assembly, we used pseudoparticles that were produced without furin inhibitor, yielding cleaved spike proteins. Vero-TMPRSS2 cells gave overall significantly higher luciferase signal indicative of more efficient entry. In contrast, we observed lowered infection in Vero E6 cells. As expected, VSVpp (positive control) pseudoparticles infected both cell lines with several orders of magnitude higher luciferase units than the values reported with Δ envpp infection (data not shown).

Evolution of the P681R mutation and its role in the transmissibility and emergence of SARS-CoV-

To further understand the evolution of the P681R substitution and its role in the transmissibility of SARS-CoV-2, we monitored the frequency of substitutions at the S1/S2 (furin cleavage) site in the global surveillance data and plotted these substitution data as fraction of total genomes. The initial B lineage virus that spread out of Wuhan encoded a spike protein with P681 at the furin cleavage site, along with G614. Fairly early in the epidemic, the D614G substitution appeared in the B.1 lineage and became prevalent in May-December 2020 (Figure 5A). The B.1.1.7 (Alpha) lineage evolved from B.1, spread widely in the UK and other regions and encoded a P681H substitution in the G614D background. B.1.1.7 peaked globally in March-April 2021 (Figure 5A). In most regions of the world, the B.1.617.2 (Delta) lineage encoding spike D614G and P681R spread widely following its emergence in India in May-June 2021, and became the dominant observed lineage globally (Figure 5A). In comparison, a distinct A lineage virus (A.23) containing Q613H appeared in August-September 2020 and acquired the P681R mutation at a much earlier time (December 2020-January 2021); however, it circulated only briefly (Figure 5B).

Discussion

Since late 2020, the evolution of the SARS-CoV-2 pandemic has been characterized by the emergence of viruses bearing sets of substitutions/deletions, designated “variants of concern” (VOCs) and “variants under investigation” (VUIs). These variants appear to have expanded following the selection for substitution or deletions in the spike protein, such as D614G and Q613H, along with mutations in other viral proteins. The substitutions encoded by such variants may alter virus characteristics including enhanced transmissibility and antigenicity, some provide a direct advantage to avoid the changing developing immune responses in the population due to prior exposure or vaccination as well as the social dynamics of the human population (4, 20-23). The first notable SARS-CoV-2 VOC of 2021 was B.1.1.7 (Alpha), which among other changes, encoded a P681H substitution in the spike S1/S2 furin cleavage site and has been linked to increased transmissibility due to the presence of the additional basic amino acid, histidine (H). However, histidine is unusual in that it has an ionizable side chain with a pKa near neutrality (24), and so is not conventionally

considered a basic amino acid. Most recently, the VOC (B.1.617.2, or Delta) has replaced B.1.1.7(Alpha) as the dominant circulating virus globally, which like A.23.1 and (sub-lineages B.1.617.1 and B.1.617.3) encodes a P681R substitution, and is more conventionally “polybasic” in the S1/S2 cleavage motif, than the P681H of B.1.1.7 (Alpha) and is suggested to enhance transmissibility and pathogenesis (25, 26). However, the A.23.1 variant pre-dated B.1.617 as a P681R-containing VOC/VOI by several months. It is important to note that both lineages that have dominated Uganda have encoded the spike P681R substitution. A second important detail may be the timing of position 681 changes (either to H or R) relative to other changes in the spike. In all cases the position 681 change occurs after a change of position 613/614 (B to B.1 to B.1.1.7, B to B.1 to B.1.617.2, and A to A.23. A.23.1 and now B to B.1 to B.1.1.7 to Q.4)

The analyses reported here show that the substitution strongly influences furin-mediated cleavage at the *in vitro* level, with these results being consistent with other studies (23, 24). However, it is very clear that P681R is not the sole driver of spike protein function in vivo, a finding reinforced by the molecular studies described here. It would be of interest to understand associated spike changes that cooperate with P681R. The introduction of P681R alone into the WT Wuhan-1 spike did not reproduce the full activity of the A.23.1 spike (Figures 3 and 4). One limitation of this study is that isolated viruses of the A.23.1 lineage are not available for infectivity assays, and so our work relies on the use of epidemiological tools and the reconstruction of virus infection in biochemical and cell-based assays. Another limitation is that our “wild-type” virus contains D614 and not 614G. Despite these limitations, we consider that our data support our conclusion that the spike mutation P681R—by itself—is a not a primary driver of virus transmissibility in the population, with A.23.1 giving unique insight into these aspects of the ongoing COVID-19 pandemic, but requires the full context of additional spike and other viral changes seen in A.23.1, B.1.617.2 and Q.4 for transmission success.

A.23.1 was a key SARS-CoV-2 variant spreading within Africa during the early part of 2021, and has been defined (along with C.11) as an African VOI (27) having have multiple mutations on the spike glycoprotein and evolving in a clocklike manner along with other variants. Epidemiological data from Uganda support

the importance of the P681R substitution in A.23.1 for community-wide transmission. The subsequent decline of the P681R lineage A.23.1 in Uganda, combined with the in vitro analyses reported here clearly showed that the P681R alone is not sufficient to drive such dominance. The P681R lineage B.1.617.2 (Delta) likely benefited from additional S and other substitutions and eventually dominated the Uganda epidemic by June 2021 (Figure 5), similar to patterns globally.

While P681R does make the S1/S2 cleavage site more basic in nature, such variant cleavage sites are still not “ideal” for furin—as originally found in the prototype furin-cleaved virus mouse hepatitis virus (MHV) (RRARR|S) (28, 29). The introduction of an arginine residue does appear to be making S1/S2 more “polybasic” as the pandemic continues and transmissibility increases. While we should not over-simplify the complex process of spike protein activation, it will be interesting to see whether this progression of basic residue addition continues with new variants, towards that seen in established community-acquired respiratory (CAR) coronaviruses such as HCoV-HKU-1 or HCoV-OC43, with S1/S2 sequences of RRKRR|S and RRSRR|A respectively.

Acknowledgements

This work was funded in part by the National Institute of Health research grant R01AI35270 (to GW and SD). We thank the global SARS-CoV-2 sequencing groups for their open and rapid sharing of sequence data and GISAID for providing an effective platform to make these data available. DLB, MVTP and MC were funded by the UK Medical Research Council (MRC/UK Research and Innovation) and the UK Department for International Development (DFID) under the MRC/DFID Concordat agreement (grant agreement no. NC_PC_19060) and Wellcome Trust, UK FCDO—Wellcome Epidemic Preparedness—Coronavirus (grant agreement no. 220977/Z/20/Z). TT was supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE-1650441 and the Samuel C. Fleming Family Graduate Fellowship.

Author contributions

Conceptualization: BL, LEF, TT, SD, JAJ and GRW. Methodology: BL, LEF, MVTP, DLB, TT, SD, MC, JAJ and GRW. Investigation: BL, LEF, MVTP, DLB, TT, JAJ. Writing - Original Draft: BL, TT, JAJ. and GRW. Writing - Review & Editing, BL, TT, SD, MC, JAJ and GRW. Visualization: BL, LEF, MVTP, TT and JAJ. Supervision: SD, MC, JAJ and G.R.W. Funding acquisition: SD, MC and GRW.

Declaration of Interests

The authors manifest no conflict of interest.

References

1. Whittaker GR, Daniel S, Millet JK. 2021. Coronavirus entry: how we arrived at SARS-CoV-2. *Curr Opin Virol* 47:113-120.
2. Rambaut A, Holmes EC, O'Toole A, Hill V, McCrone JT, Ruis C, du Plessis L, Pybus OG. 2020. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. *Nat Microbiol* 5:1403-1407.
3. Zhou B, Thao TTN, Hoffmann D, Taddeo A, Ebert N, Labroussaa F, Pohlmann A, King J, Steiner S, Kelly JN, Portmann J, Halwe NJ, Ulrich L, Trueb BS, Fan X, Hoffmann B, Wang L, Thomann L, Lin X, Stalder H, Pozzi B, de Brot S, Jiang N, Cui D, Hossain J, Wilson MM, Keller MW, Stark TJ, Barnes JR, Dijkman R, Jores J, Benarafa C, Wentworth DE, Thiel V, Beer M. 2021. SARS-CoV-2 spike D614G change enhances replication and transmission. *Nature* 592:122-127.
4. Harvey WT, Carabelli AM, Jackson B, Gupta RK, Thomson EC, Harrison EM, Ludden C, Reeve R, Rambaut A, Consortium C-GU, Peacock SJ, Robertson DL. 2021. SARS-CoV-2 variants, spike mutations and immune escape. *Nat Rev Microbiol* doi:10.1038/s41579-021-00573-0.
5. Bugembe DL, Phan MVT, Ssewanyana I, Semanda P, Nansumba H, Dhaala B, Nabadda S, O'Toole AN, Rambaut A, Kaleebu P, Cotten M. 2021. Emergence and spread of a SARS-CoV-2 lineage A variant (A.23.1) with altered spike protein in Uganda. *Nat Microbiol* doi:10.1038/s41564-021-00933-9.
6. Butera Y, Mukantwari E, Artesi M, D'Arc Umuringa J, O'Toole AN, Hill V, Rooke S, Hong SL, Dellicour S, Majyambere O, Bontems S, Boujemla B, Quick J, Resende PC, Loman N, Umumararungu E, Kabanda A, Murindahabi MM, Tuyisenge P, Gashegu M, Rwabihama JP, Sindayiheba R, Gikic D, Souopgui J, Ndifon W, Rutayisire R, Gatara S, Mpunga T, Ngamije D, Bours V, Rambaut A, Nsanzimana S, Baele G, Durkin K, Mutesa L, Rujeni N. 2021. Genomic Sequencing of SARS-CoV-2 in Rwanda: evolution and regional dynamics. *medRxiv* doi:10.1101/2021.04.02.21254839:2021.04.02.21254839.
7. Bugembe DL, Phan MVT, Abias A, Ayei J, Deng LL, Lako RLL, Rumunu J, Kaleebu P, Wamala JF, Hm JJ, Lodiongo DK, Bunga S, Cotten M. 2021. SARS-CoV-2 Variants, South Sudan, January–March 2021. *Emerging Infectious Disease journal* 27.
8. Bugembe DL, V.T.Phan M, Ssewanyana I, Semanda P, Nansumba H, Dhaala B, Nabadda S, O'Toole AN, Rambaut A, Kaleebu P, Cotten M. 2021. A SARS-CoV-2 lineage A variant (A.23.1)

- p>with altered spike has emerged and is dominating the current Uganda epidemic. medRxiv
-
- doi:10.1101/2021.02.08.21251393:2021.02.08.21251393.
9. Jaimes JA, Andre NM, Chappie JS, Millet JK, Whittaker GR. 2020. Phylogenetic Analysis and Structural Modeling of SARS-CoV-2 Spike Protein Reveals an Evolutionary Distinct and Proteolytically Sensitive Activation Loop. *J Mol Biol* 432:3309-3325.
10. Polgár L. 1989. General Aspects of Proteases, p 43-86, Mechanisms of Protease Action. CRC press, Boca Raton, FL.
11. Seidah NG, Prat A. 2012. The biology and therapeutic targeting of the proprotein convertases. *Nat Rev Drug Discov* 11:367-83.
12. Johnson BA, Xie X, Bailey AL, Kalveram B, Lokugamage KG, Muruato A, Zou J, Zhang X, Juelich T, Smith JK, Zhang L, Bopp N, Schindewolf C, Vu M, Vanderheiden A, Winkler ES, Swetnam D, Plante JA, Aguilar P, Plante KS, Popov V, Lee B, Weaver SC, Suthar MS, Routh AL, Ren P, Ku Z, An Z, Debbink K, Diamond MS, Shi PY, Freiberg AN, Menachery VD. 2021. Loss of furin cleavage site attenuates SARS-CoV-2 pathogenesis. *Nature* 591:293-299.
13. Peacock TP, Goldhill DH, Zhou J, Baillon L, Frise R, Swann OC, Kugathasan R, Penn R, Brown JC, Sanchez-David RY, Braga L, Williamson MK, Hassard JA, Staller E, Hanley B, Osborn M, Giacca M, Davidson AD, Matthews DA, Barclay WS. 2020. The furin cleavage site of SARS-CoV-2 spike protein is a key determinant for transmission due to enhanced replication in airway cells. bioRxiv doi:10.1101/2020.09.30.318311:2020.09.30.318311.
14. Hoffmann M, Kleine-Weber H, Schroeder S, Kruger N, Herrler T, Erichsen S, Schiergens TS, Herrler G, Wu NH, Nitsche A, Muller MA, Drosten C, Pohlmann S. 2020. SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* 181:271-280 e8.
15. Lubinski B, Tang T, Daniel S, Jaimes JA, Whittaker GR. 2021. Functional evaluation of proteolytic activation for the SARS-CoV-2 variant B.1.1.7: role of the P681H mutation. bioRxiv doi:10.1101/2021.04.06.438731.
16. Tian S, Huajun W, Wu J. 2012. Computational prediction of furin cleavage sites by a hybrid method and understanding mechanism underlying diseases. *Sci Rep* 2:261.
17. Duckert P, Brunak S, Blom N. 2004. Prediction of proprotein convertase cleavage sites. *Protein Eng Des Sel* 17:107-12.
18. Jaimes JA, Millet JK, Whittaker GR. 2020. Proteolytic Cleavage of the SARS-CoV-2 Spike Protein and the Role of the Novel S1/S2 Site. *iScience* 23:101212.
19. Millet JK, Tang T, Nathan L, Jaimes JA, Hsu HL, Daniel S, Whittaker GR. 2019. Production of Pseudotyped Particles to Study Highly Pathogenic Coronaviruses in a Biosafety Level 2 Setting. *J Vis Exp* doi:10.3791/59010.
20. Muller NF, Wagner C, Frazar CD, Roychoudhury P, Lee J, Moncla LH, Pelle B, Richardson M, Ryke E, Xie H, Shrestha L, Addetia A, Rachleff VM, Lieberman NAP, Huang ML, Gautom R, Melly G, Hiatt B, Dykema P, Adler A, Brandstetter E, Han PD, Fay K, Ilcisin M, Lacombe K, Sibley TR, Truong M, Wolf CR, Boeckh M, Englund JA, Famulare M, Lutz BR, Rieder MJ, Thompson M, Duchin JS, Starita LM, Chu HY, Shendure J, Jerome KR, Lindquist S, Greninger AL, Nickerson DA, Bedford T. 2021. Viral genomes reveal patterns of the SARS-CoV-2 outbreak in Washington State. *Sci Transl Med* 13.

21. Richard D, Shaw LP, Lanfear R, Acman M, Owen CJ, Tan CC, van Dorp L, Balloux F. 2021. A phylogeny-based metric for estimating changes in transmissibility from recurrent mutations in SARS-CoV-2. *bioRxiv* doi:10.1101/2021.05.06.442903:2021.05.06.442903.
22. Grubaugh ND, Hodcroft EB, Fauver JR, Phelan AL, Cevik M. 2021. Public health actions to control new SARS-CoV-2 variants. *Cell* 184:1127-1132.
23. Peacock TP, Penrice-Randal R, Hiscox JA, Barclay WS. 2021. SARS-CoV-2 one year on: evidence for ongoing viral adaptation. *J Gen Virol* 102.
24. Nelson DL, Cox MM. 2000. *Lehninger Principles of Biochemistry*. Worth Publishers.
25. Peacock TP, Sheppard CM, Brown JC, Goonawardane N, Zhou J, Whiteley M, de Silva TI, Barclay WS. 2021. The SARS-CoV-2 variants associated with infections in India, B.1.617, show enhanced spike cleavage by furin. *bioRxiv* doi:10.1101/2021.05.28.446163:2021.05.28.446163.
26. Saito A, Nasser H, Uriu K, Kosugi Y, Irie T, Shirakawa K, Sadamasu K, Kimura I, Ito J, Wu J, Ozono S, Tokunaga K, Butlertanaka EP, Tanaka YL, Shimizu R, Shimizu K, Fukuhara T, Kawabata R, Sakaguchi T, Yoshida I, Asakura H, Nagashima M, Yoshimura K, Kazuma Y, Nomura R, Horisawa Y, Takaori-Kondo A, Nakagawa S, Ikeda T, Sato K. 2021. SARS-CoV-2 spike P681R mutation enhances and accelerates viral fusion. *bioRxiv* doi:10.1101/2021.06.17.448820:2021.06.17.448820.
27. Wilkinson E, Giovanetti M, Tegally H, San JE, Lessels R, Cuadros D, Martin DP, Zekri A-RN, Sangare AK, Ouedraogo A-S, Sesay AK, Hammami A, Amuri AA, Sayed A, Rebai A, Elargoubi A, Trotter AJ, Keita AK, Sall AA, Kone A, Souissi A, Gutierrez AV, Page AJ, Iranzadeh A, Lambisia A, Sylverken A, Ibrahimi A, Dhaala B, Kouriba B, Kleinhans B, Brook C, Williamson C, Pratt CB, Akoua-Koffi CG, Agoti CN, Morang'a CM, Nokes DJ, Bridges DJ, Bugembe DL, Baker D, Doolabh D, Ssemwanga D, Tshabuila D, Bassirou D, Amuzu DSY, Goedhals D, Maruapula D, Foster-Nyarko E, Lusamaki EK, Simulundu E, et al. 2021. A year of genomic surveillance reveals how the SARS-CoV-2 pandemic unfolded in Africa. *medRxiv* doi:10.1101/2021.05.12.21257080:2021.05.12.21257080.
28. Sturman LS, Ricard CS, Holmes KV. 1985. Proteolytic cleavage of the E2 glycoprotein of murine coronavirus: activation of cell-fusing activity of virions by trypsin and separation of two different 90K cleavage fragments. *J Virol* 56:904-11.
29. Whittaker GR. 2021. SARS-CoV-2 spike and its adaptable furin cleavage site. *Lancet Microbe* 2:e488-e489.
30. Sun X, Tse LV, Ferguson AD, Whittaker GR. 2010. Modifications to the hemagglutinin cleavage site control the virulence of a neurotropic H1N1 influenza virus. *J Virol* 84:8683-90.

Methodology

Furin prediction calculations. Prop: CoV sequences were inputted into the ProP 1.0 Server hosted at: cbs.dtu.dk/services/ProP/ PiTou: CoV sequences were analyzed using the PiTou V3 software hosted at: <http://www.nuolan.net/reference.html>.

Fluorogenic peptide assays

Fluorogenic peptide assays were performed as described previously with minor modifications (18). Each reaction was performed in a 100 µl volume consisting of buffer, protease, and SARS-CoV-2 S1/S2 Wuhan-Hu-1 (WT) (TNSPRRARSVA), SARS-CoV-2 S1/S2 B.1.1.7 (TNSHRRARSVA) or SARS-CoV-2 S1/S2 A.23.1 (TNSRRRARSVA) fluorogenic peptide in an opaque 96-well plate. For trypsin catalyzed reactions, 0.8 nM/well TPCK trypsin was diluted in PBS buffer. For furin catalyzed reactions, 1 U/well recombinant furin was diluted in a buffer consisting of 20 mM HEPES, 0.2 mM CaCl₂, and 0.2 mM β-mercaptoethanol, at pH 7.5. Fluorescence emission was measured once per minute for 60 continued minutes using a SpectraMax fluorometer (Molecular Devices) at 30°C with an excitation wavelength of 330 nm and an emission wavelength of 390 nm. Vmax was calculated by fitting the linear rise in fluorescence to the equation of a line.

Synthesis and cloning of the A.23.1 spike protein

The sequence for the A.23.1 spike gene from isolate SARS-CoV-2 A.23.1 hCoV-19/Uganda/UG185/2020 (EPI_ISL_955136) was obtained from GISAID (<https://www.gisaid.org/>), codon-optimized, synthesized and cloned into a pcDNA 3.1+ vector for expression (GenScript). line.

Site-directed Mutagenesis

Primers (ACCTGGCTCTCCTTCGGGAGTTTGTCTGG/CCAGACAAACTCCCGAAGGAGAGCCAGGT) for mutagenesis were designed using the Agilent QuickChange Primer Design tool to create the P681R mutation (CCA->CGA). Mutagenesis was carried out on a pCDNA-SARs2 Wuhan-Hu-1 S plasmid using the Agilent QuickChange Lightning Mutagenesis kit (The original plasmid was generously provided by David Veessler, University of Washington USA). The mutated pcDNA-SARS-CoV-2 Wuhan-Hu-1 P681R S plasmid was used to transform XL-10 gold ultracompetent cells, which were grown up in small culture, and

then plasmid was extracted using the Qiagen QIAprep Spin Miniprep Kit. Sanger Sequencing was used to confirm incorporation of the mutation.

Cell-to-cell fusion assay

Vero E6 and Vero-TMPRSS2 cells were transfected with a plasmid harboring the spike gene of the SARS-CoV-2 Wuhan-Hu-1 S (WT), SARS-CoV-2 A.23.1 S, SARS-CoV-2 Wuhan-Hu-1 with a P to R mutation in the 681 amino acid position (P681R), or a delta-spike pCDNA3.1+ (S-) plasmid, and evaluated through an immunofluorescence assay (IFA), nuclei involved in syncytia counting, cell-to-cell fusion ratio and western blot. Transfection was performed on 8-well glass slides at 90% confluent cells using Lipofectamine® 3000 (Cat: L3000075, Invitrogen Co.), following the manufacturer's instructions and 250 ng of plasmid DNA per well were transfected. Cells were then incubated at 37°C with 5% of CO₂ for 28 hours. Syncytia was visualized through fluorescence microscopy using a previously described method (15). The spike expression was detected using a SARS-CoV-2 spike antibody (Cat: 40591-T62, Sino Biological Inc.) at 1/500 dilution for 1 hour. Secondary antibody labeling was performed using AlexaFluor™ 488 goat anti-rabbit IgG antibody (Cat: A32731, Invitrogen Co.) at a 1/500 dilution for 45 minutes. Representative images of each treatment group were taken, and these images were used to calculate the percent of nuclei involved in the formation of syncytia and for quantification of the cell-to-cell fusion ratio based in fluorescence intensity. Images were taken at 20X on the Echo Revolve fluorescent microscope (Model: RVL-100-M). The nuclei were counted manually using the Cell Counter plugin in ImageJ (<https://imagej.nih.gov/ij/>). Cells that expressed the spike protein and contained 4 or more nuclei were considered to be one syncytium. Cell-to-cell fusion ration was calculated using image J version 1.53 (National Institutes of Health). Expressed protein from Vero E6 and Vero-TMPRSS2 cells was analyzed through a cell surface biotinylation assay western blot as described previously. (30) Spike protein was detected using the antibodies described previously in this section. GLUT4 protein was used as a housekeeping expression control and labeled using a GLUT4 monoclonal antibody (Cat: MA5-17175, Invitrogen Co.).

431

432 **Pseudoparticle Generation**

433 Pseudoparticle generation was carried out using a murine leukemia virus (MLV)-based system as previously
 434 described with minor modification (19). HEK293T cells were seeded at 2.5×10^5 cells/ml in a 6-well plate the
 435 day before transfection. Transfections were performed using polyethylenimine (PEI) and 1X Opti-Mem
 436 from Gibco. Cells were transfected with 800 ng of pCMV-MLV-gagpol, 600 ng of pTG-luc, and 600 ng of
 437 a plasmid containing the viral envelope protein of choice. Viral envelope plasmids included pcDNA-SARS-
 438 CoV-2 Wuhan-Hu1 S as the WT, pcDNA-SARS-CoV-2 Wuhan-Hu-1 P681R S, and pcDNA-SARS-CoV-2
 439 A.23.1 S. pCAGGS-VSV G was used as a positive control and pCAGGS was used for a Δ -envelope negative
 440 control. At 48 hours post-transfection, the supernatant containing the pseudoparticles was removed,
 441 centrifuged to remove cell debris, filtered, and stored at -80°C .

442

443 **Pseudoparticle Infection Assay**

444 Infection assays were performed as previously described with minor adjustments (19). Vero E6 and Vero-
 445 TMPRSS2 cells were seeded at 4.5×10^5 cells/ml in a 24-well plate the day before infection. Calu-3 cells were
 446 seeded at 10×10^5 cells/ml in a 24-well plate three days before infection. Cells were washed three times with
 447 DPBS and infected with 200 μl of either VSV G, SARS-CoV-2 S, SARS-CoV-2 P681R S, SARS-CoV-2
 448 A.23.1 S, or Δ -envelope pseudoparticles. Infected cells incubated on a rocker for 1.5 hours at 37°C , and then
 449 300 μl of complete media was added and cells were left at 37°C . At 72 hours post-infection, cells were lysed
 450 and the level of infection was assessed using the Promega Luciferase Assay System. The manufacturer's
 451 protocol was modified by putting the cells through 3 freeze/thaw cycles after the addition of 100 μl of the
 452 lysis reagent. 10 μl of the cell lysate was added to 20 μl of luciferin, and then luciferase activity was measured
 453 using the Glomax 20/20 luminometer (Promega). Vero E6 and Vero-TMPRSS2 infection assays were
 454 replicated four times. The Calu-3 infection assays were replicated three times. Each assay was performed as
 455 technical replicates, with the same batch of pseudoparticles used for all of data.

456

457 **Western blot analysis of pseudoparticles.** A 3 ml volume of pseudoparticles was pelleted using a TLA-
 458 55 rotor with an Optima-MAX-E ultracentrifuge (Beckman Coulter) for 2 hours at 42,000 rpm at 4°C.
 459 untreated particles were resuspended in 30 µl DPBS buffer. For the + furin treated MLVpps, particles were
 460 resuspended in 30 µL of furin buffer consistent in 20 mM HEPES, 0.2 mM CaCl₂, and 0.2 mM β-
 461 mercaptoethanol (at pH 7.0). Particles were later incubated with 6 U of recombinant furin for 3 h at 37 °C.
 462 Sodium dodecyl sulfate (SDS) loading buffer and DTT were added to samples and heated at 65°C for 20
 463 minutes. Samples were separated on NuPAGE Bis-Tris gel (Invitrogen) and transferred on polyvinylidene
 464 difluoride membranes (GE). SARS-CoV-2 S was detected using a rabbit polyclonal antibody against the S2
 465 domain (Cat: 40590-T62, Sinobiological) and an AlexaFluor 488 goat anti-rabbit antibody. Bands were
 466 detected using the ChemiDoc Imaging software (Bio-Rad) and band intensity was calculated using the
 467 analysis tools on Biorad Image Lab 6.1 software to determine the uncleaved to cleaved S ratios.

468

469 **Uganda cases vs. viral lineages over time.** Daily reported SARS-CoV-2 infections The Uganda daily
 470 SARS-CoV-2 positive samples numbers were retrieved from Our World in Data
 471 (<https://ourworldindata.org/coronavirus>) and the 7 day average was determined. Uganda SARS-CoV-2
 472 lineage data were generated from the MRC Uganda genomic data deposited in the GISAID database
 473 (<https://www.gisaid.org/>). SARS-CoV-2 Pango lineages (2) were determined using the pangolin module
 474 pangoLEARN (<https://github.com/cov-lineages/pangolin>).

475

476 **Spike position 681 and 613/614 changes in global data.** All available spike protein sequences were
 477 obtained from the GISAID database. The frequency of P681, P681R, P681H, D614 D614G and D613H
 478 were counted by string matching using Ack (<http://beyondgrep.com/>) to the major variations of the 88
 479 amino acid peptide sequence (aa 605 to 691) spanning the two relevant sites (D614_P681 (Wuhan_B),
 480 D614G_P681 (B.1), D614_P681H, D614G_P681H (B.1.1.7), D614_P681R, Q613_P681 (Wuhan_B),

481 Q613H_P681 (A.23), Q613H_P681R (A.23.1), D614G_P681R (B.1.617.2). Fractions of available total
 482 genomes for each month encoding each peptide variant were visualized in a heatmap. Additional changes at
 483 position H655Y (present in the Gamma lineage) were also included in the count and fraction calculation but
 484 the total numbers were minor.

485

486

487

488

489

490

491

492

493

A

B

	Oct-20	Nov-20	Dec-20	Jan-21	Feb-21	Mar-21	Apr-21	May-21	Jun-21	Jul-21	
Lineage	0.14	0.42	0.88	1	0.4	0.2	0.21	0.07	0.02	0	A.23.1
	0.24	0.3	0.05	0	0.05	0	0	0	0	0	A.23
	0.1	0	0	0	0	0	0	0	0	0	A
	0	0	0	0	0	0	0.03	0	0	0	A.19
	0	0.03	0.01	0	0	0	0	0	0	0	A.25
	0	0	0	0	0	0	0	0	0	0	B
	0.1	0.03	0.02	0	0	0	0	0	0	0	B.1
	0	0	0	0	0	0	0.03	0	0	0	B.1.1
	0	0	0.01	0	0	0	0	0	0	0	B.1.1.161
	0.29	0	0	0	0	0	0	0	0	0	B.1.1.27
	0	0	0	0	0.15	0	0.03	0	0	0	B.1.1.318
	0.05	0.03	0	0	0	0	0	0	0	0	B.1.214
	0	0.03	0.02	0	0.1	0	0	0	0	0	B.1.222
	0	0.03	0	0	0	0	0	0	0	0	B.1.363
	0.1	0.12	0	0	0	0	0	0	0	0	B.1.393
	0	0	0	0	0.2	0.25	0.1	0	0.02	0	B.1.351
	0	0	0	0	0	0.28	0.03	0.14	0.01	0	B.1.1.7
	0	0	0	0	0	0.03	0	0	0	0	B.1.617.1
	0	0	0	0	0.1	0.25	0.41	0.57	0.01	0	B.1.525
	0	0	0	0	0	0	0.14	0.21	0.95	1	B.1.617.2
	Month										

Figure 1: Uganda SARS-CoV-2 cases and lineages, October 2020 to June 2021. **A.** 7-day average positive cases numbers were plotted by day, the peak of 750 cases/ per day observed in the first wave of infections in January 2021 is indicated with a dotted line. Case data were obtained from Our World in Data (<https://ourworldindata.org/>). **B.** Monthly SARS-CoV-2 lineage data for Uganda. All Uganda full genome sequences from GISAID (<https://www.gisaid.org/>) were retrieved, lineage types using the Pango tool (<https://cov-lineages.org/resources/pangolin.html>), and the fraction of each month's total genomes were plotted. Fractions were indicated in each cell and cells are colored (white to dark red) by increasing fraction.

A

B

Virus	S1/S2 Sequence	Furin Score PiTou	Furin Score ProP
SARS-CoV-2 (A.1/B.1)	672-ASYQTQNSP RRAR SVASQS-691	+9.196	0.626
SARS-CoV-2 (B.1.1.7)	669-ASYQTQNS RRAR SVASQS-688	+9.907	0.706
SARS-CoV-2 (A.23.1)	669-ASYQTQNS RRAR SVASQS-688	+12.209	0.704
SARS-CoV-2 (B.1.617.2)	669-ASYQTQNS RRAR SVASQS-688	+12.209	0.704
MERS-CoV	738-LPDTPTSLTP RSVR SVPGEM-757	+5.155	0.556
HCoV-HKU1	747-YNPSSSSSS RRRR SISASY-766	+14.634	0.868
HCoV-OC43 (clinical)	750-GYCVDFK RRSR AITTG-769	+10.10	0.736

C

Figure 2: A. Summary of notable functional domains and sequence changes in the spike gene of A.23.1 compared to Wuhan-Hu-1 and B.1.617.2 (Delta). **B.** Furin cleavage score analysis of CoV S1/S2 cleavage sites. CoV S sequences were analyzed using the ProP1 1.0 and PiTou2 3.0 furin prediction algorithm, generating a score with bold numbers indicating predicted furin cleavage. (|) denotes the position of the predicted S1/S2 cleavage site. Basic residues, arginine (R) and lysine (K), are highlighted in blue, with histidine in purple. Sequences corresponding to the S1/S2 region of SARS-CoV-2 (QHD43416.1), SARS-CoV (AAT74874.1), MERS-CoV (AFS88936.1), HCoV-HKU1 (AAT98580.1), HCoV-OC43 (KY369907.1) were obtained from GenBank. Sequences corresponding to the S1/S2 region of SARS-CoV-2 B.1.1.7 (EPI_ISL_1374509) and SARS-CoV-2 A.23.1 hCoV-19/Uganda/UG185/2020 (EPI_ISL_955136), were obtained from GISAID. **C.** Fluorogenic peptide cleavage assays of the SARS-CoV-2 S1/S2 cleavage site. Peptides mimicking the S1/S2 site of the SARS-CoV-2 Wuhan-Hu-1 (WT – P681), B.1.1.7 (P681H) and A.23.1 (P681R) variants were evaluated for in vitro cleavage by furin, compared to trypsin control.

A

B

Vero E6 cells				Vero-TMPRSS2 cells			
	# Total nuclei	# Nuclei involved in syncytia	% Nuclei involved in syncytia		# Total nuclei	# Nuclei involved in syncytia	% Nuclei involved in syncytia
S -	6,824	0	0	S -	13,051	0	0
S WT	4,866	345	7.09	S WT	10,441	7718	73.92
S A.23.1	4,627	1054	22.77	S A.23.1	7,861	6367	80.99
S P681H	4,703	955	20.30	S P681H	10,747	8033	74.74

C

D

Figure 3: A. Cell-to-cell fusion assay of SARS-CoV-2 Wuhan-Hu-1 S (WT), SARS-CoV-2 S A.23.1 variant, or SARS-CoV-2 S WT with P681R mutation. S- = non-transfected cells. SARS-CoV-2 S was detected using a rabbit antibody against the SARS-CoV-2 S2 region. **B.** Syncytia quantification by number of nuclei involved in syncytia. **C.** Cell-to-cell fusion ratio measured by fluorescence intensity. **D.** Western blot analysis of membrane expressed S proteins and GLUT4 (housekeeping expression protein). All the experiments were performed on Vero E6 and Vero-TMPRSS2 cells.

535

A

B

Figure 4: A. Western blot analysis of MLVpp-SARS-CoV-2 S produced in $\pm$ dec-RVCR-CMK and treated with $\pm$ furin. S was detected using a rabbit antibody against the SARS-CoV-2 S2 subunit. MLV content was detected using a mouse antibody against MLV p30. **B.** Pseudoparticle infectivity assays in Vero E6, Vero-TMPRSS2 or Calu-3 cells. Cells were infected with MLVpps harboring the VSV-G, SARS-CoV-2 S (WT), SARS-CoV-2 S A.23.1 variant, SARS-CoV-2 S WT with P681R mutation. Data represents the average luciferase activity of cells of three (Vero E6 and Vero-TMPRSS2).

A

B

Figure 5: The frequency of P681, P681R, P681H, D614 D614G and D613H substitutions were counted by string matching to a peptide sequencing spanning the position 613 to 681) including the relevant sites at 613/614 and 681. (D614_P681 (Wuhan_B), D614G_P681 (B.1), D614_P681H, D614G_P681H (B.1.1.7), D614_P681R, Q613_P681 (Wuhan_B), Q613H_P681 (A.23), Q613H_P681R (A.23.1), D614G_P681R (B.1.617.2). Fractions of total genomes available for each month were plotted. Color bar at the bottom of each panel indicate fraction/color code. Panel A. Lineage B relevant substitutions. Panel B. Lineage A.23 and A.23.1 relevant substitutions.