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. Phan [3], Daniel L. Bugembe [3], Tiffany Tang [4], Susan Daniel [4], 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

grw7@cornell.edu; jaj246@cornell.edu

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 additional spike mutations (R102I, L141F, E484K and P681R) to comprise lineage A.23.1 by September 2020—with this virus being designated a variant of interest (VOI) in Africa and with subsequent spread to 26 other countries. The P681R spike substitution of the A.23.1 VOI 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 appeared in B.1.617 variants, including B.1.617.2 (Delta). Here, we performed assays using fluorogenic peptides mimicking the S1/S2 sequence from A.23.1 and B.1.617.2 and observed significantly increased cleavability with furin, compared to sequences derived from the original Wuhan-Hu1 S1/S2. We performed functional infectivity assays using pseudotyped MLV particles harboring SARS-CoV-2 spike proteins and observed an increase in transduction for A.23.1-pseudotyped particles compared to Wuhan-Hu-1 in Vero-TMPRSS2 and Calu-3 cells (with a presumed "early" entry pathway), although lowered infection in Vero E6 cells (with a presumed "late" entry pathway). However, these changes in infectivity 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 is not sufficient and needs to occur on the background of other spike protein changes to enable its full functional consequences.

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 sublineages. While the B 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 and Q613H in lineage A—an apparent example of convergent evolution that resulted in a more stabilized spike protein, with D614G linked to modest increase in virus transmissibility (3). The D614G/Q613H have now become established in circulating B/A lineage.

Compared with the lineage B 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 remain 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 has spread over 46 countries and was listed as variant of interest (VOI). 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 (4). Subsequently, the evolving A.23.1 lineage acquired additional spike substitutions (R102I, L141F, E484K and P681R), as well as in nsp6, ORF8 and ORF9. By July 2021, the A.23.1 lineage has been observed in 1102 genomes reported from 46 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 great importance 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 (5). The S1/S2 junction of the SARS-CoV-2 S gene has a distinct indel compared to all other 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 (6), 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 (7). For SARS-CoV-2, the presence of the S1/S2 "furin site" enhances virus transmissibility (8, 9). 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 (10). Notably, the P681R substitution appears in several other lineages, most notably B.1.627.2 (Delta) but also AY.2 (N=115), B.1.551 (N=156), AU.2 (N=302), B.1.1.25 (N=509), B.1.466.2 (N=538), 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 (11). 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 change at the S1/S2 (furin) cleavage site.

Methods

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 (12). Each reaction was performed in a 100 μl volume consisting of buffer, protease, and SARS-CoV-2 S1/S2 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 $CaCl_2$, 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 Veesler, 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.

Pseudoparticle Generation

Pseudoparticle generation was carried out using a murine leukemia virus (MLV)-based system as previously described with minor modification (13). HEK293T cells were seeded at $2.5x10^5$ cells/ml in a 6-well plate the day before transfection. Transfections were performed using polyethylenimine (PEI) and 1X Opti-Mem from Gibco. Cells were transfected with 800 ng of pCMV-MLV-gagpol, 600 ng of pTG-luc, and 600 ng of a plasmid containing the viral envelope protein of choice. Viral envelope plasmids included pcDNA-SARS-CoV-2 Wuhan-Hu1 S as the WT, pcDNA-SARS-CoV-2 Wuhan-Hu-1 P681R S, and pcDNA-SARS-CoV-2 A.23.1 S. pCAGGS-VSV G was used as a positive control and pCAGGS was used for a Δ-envelope negative control. At

48 hours post-transfection, the supernatant containing the pseudoparticles was removed, centrifuged to remove cell debris, filtered, and stored at -80°C.

Pseudoparticle Infection Assay

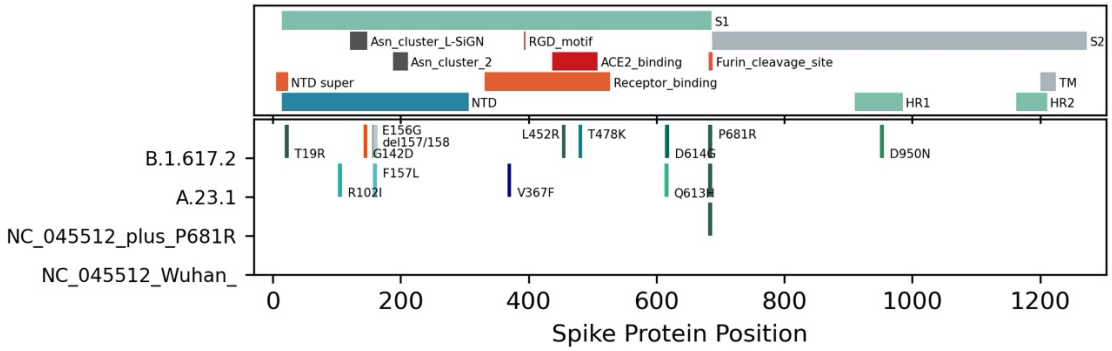
Infection assays were performed as previously described with minor adjustments (13). Vero E6 and Vero-TMPRSS2 cells were seeded at $4.5x10^5$ cells/ml in a 24-well plate the day before infection. Calu-3 cells were seeded at 10 $x10^5$ cells/ml in a 24-well plate three days before infection. Cells were washed three times with DPBS and infected with 200 µl of either VSV G, SARS-CoV-2 S, SARS-CoV-2 P681R S, SARS-CoV-2 A.23.1 S, or Δ-envelope pseudoparticles. Infected cells incubated on a rocker for 1.5 hours at 37°C, and then 300 µl of complete media was added and cells were left at 37°C. At 72 hours post-infection, cells were lysed and the level of infection was assessed using the Promega Luciferase Assay System. The manufacturer's protocol was modified by putting the cells through 3 freeze/thaw cycles after the addition of 100 µl of the lysis reagent. 10 µl of the cell lysate was added to 20 µl of luciferin, and then luciferase activity was measured using the Glomax 20/20 luminometer (Promega). Vero E6 and Vero-TMPRSS2 infection assays were replicated four times. The Calu-3 infection assays were replicated three times. Each assay was performed as technical replicates, with the same batch of pseudoparticles used for all of data.

Western blot analysis of pseudoparticles. A 3 ml volume of pseudoparticles was pelleted using a TLA-55 rotor with an Optima-MAX-E ultracentrifuge (Beckman Coulter) for 2 hours at 42,000 rpm at 4°C. Particles were resuspended in 30 µl DPBS buffer. Sodium dodecyl sulfate (SDS) loading buffer and DTT were added to samples and heated at 65°C for 20 minutes. Samples were separated on NuPAGE Bis-Tris gel (Invitrogen) and transferred on polyvinylidene difluoride membranes (GE). SARS-CoV-2 S was detected using a rabbit polyclonal antibody against the S2 domain (Cat: 40590-T62, Sinobiological) and an AlexaFluor 488 goat anti-rabbit antibody. Bands were detected using the ChemiDoc Imaging software (Bio-Rad) and band intensity was calculated using the analysis tools on Biorad Image Lab 6.1 software to determine the uncleaved to cleaved S ratios.

Results

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 and B.1.617.2 (Delta) contain a P681R substitution at the S1/S2 interface **(Figure 1A),** 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 (14) and ProP (15) cleavage prediction tools, comparing the spike proteins from A.23.1 to B.1.1.7 and the prototype SARS-CoV-2 Wuhan-Hu-1, as well as to SARS-CoV, MERS-CoV and selected other human respiratory betacoronaviruses (HCoV-HKU1 and HCoV-OC43); **Figure 1B.** Both algorithms predicted a significant increase in the furin cleavage for A.23.1 and B.1.617.2 (Delta) compared to Wuhan-Hu1 and B.1.1.7; in comparison, SARS-CoV is not predicted to be furin-cleaved; as expected MERS-CoV showed a relatively low furin cleavage score, with HCoV-HKU1 and HCoV-OC43 showing relatively high furin cleavage scores.



Virus	S1/S2 Sequence	Furin Score PiTou	Furin Score ProP
SARS-CoV-2 (A.1/B.1)	672-ASYQTQTNSP**RR**A**R**|SVASQS-691	+9.196	0.626
SARS-CoV-2 (B.1.1.7)	669-ASYQTQTNS**H**R**R**A**R**|SVASQS-688	+9.907	0.757
SARS-CoV-2 (A.23.1)	669-ASYQTQTNS**RRR**A**R**|SVASQS-688	**+12.209**	**0.704**
SARS-CoV-2 (B.1.617)	669-ASYQTQTNS**RRR**A**R**|SVASQS-688	+12.209	0.704
MERS-CoV	738-LPDTPSTLTP**R**SV**R**|SVPGEM-757	+5.155	0.563
HCoV-HKU1	747-YNSPSSSSS**RRKRR**|SISASY-766	+14.634	0.918
HCoV-OC43 (clinical)	750-GYCVDYFKN**RR**S**RR**|AITTGY-769	+10.10	0.736

Figure 1 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 ProP[1] 1.0 and PiTou[2] 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 resides, 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.

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 cleavage activity in vitro (12). The specific peptide sequences used here were TNSRRRARSVA (A.23.1 S1/S2) and TNSPRRARSVA (Wuhan-Hu-1 S1/S2). We tested furin as the protease, along with trypsin as a control **(Figure 2).** As predicted, furin effectively cleaved the Wuhan-Hu-1 (WT) peptide but with a significant increase for the A.23.1 S1/S2 peptide. As expected, trypsin also effectively cleaved A.23.1 S1/S2 (results not shown?) . This comparative data with SARS-CoV S1/S2 sites reveals that the P618R substitution significantly increased cleavability by furin, with increases beyond that seen for B.1.1.7 (P681H) (12).

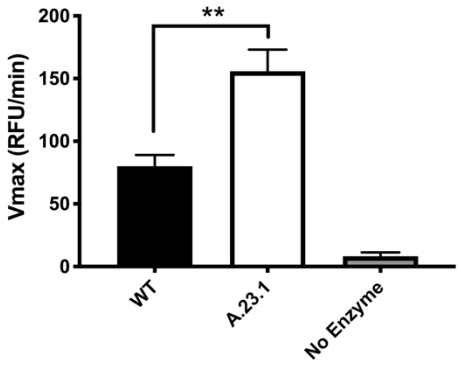


Figure 2. Fluorogenic peptide cleavage assays of the SARS-CoV-2 S1/S2 cleavage site. Peptides mimicking the S1/S2 site of the SARS-CoV-2 Wuhan-Hu-1 (WT) and A.23.1 variants were evaluated for *in vitro* cleavage by furin, compared to a control without protease.

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 luciferase, which is quantifiable (13). Using the HEK293T cell line for particle production, MLV pseudoparticles containing the A.23.1, Wuhan-Hu-1 SARS-CoV-2 S protein (WT), a P681R point mutant of Wuhan-Hu-1,

positive-control particles containing the vesicular stomatitis virus (VSV) G protein, and negative-control particles (Δenvpp) lacking envelope proteins (not shown) were prepared.

We examined infection properties of these SARS-CoV-2 pseudoparticles in cell lines representative of both the "early" and "late" cell entry pathways **(Figure 3).** Furin is predicted to cleave during virus assembly and "prime" the spike protein for subsequent entry. SARS-CoV-2 is predicted to enter cells using cathepsin L and the "early" pathway, whereas Vero-TMPRSS2 and Calu-3 cells are predicted to use a "late" pathway, with spike being activated for fusion by TMPRSS2 or other TTSPs—as the entry mechanisms of SARS-CoV-2 can be highly cell-type dependent (1). Specifically, we used the Vero-TMPRSS2 and Calu-3 ("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. While Vero-TMPRSS2 and Calu-3 cells gave overall significantly higher luciferase signal indicative of more efficient entry, 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 Δenvpp infection (negative control). We observed no difference in infection between pseudoparticles displaying spike protein with the P681R substitution in Calu-3 cells and lowered infection on Vero E6 and Vero-TMPRSS2 cells.

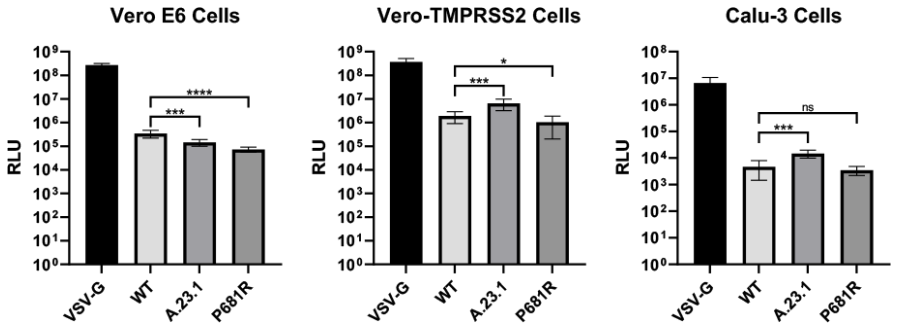


Figure 3. Pseudoparticle infectivity assays in Vero E6, Vero-TMPRSS2 or Calu-3 cells. Cells were infected with MLVpps harboring the VSV-G, SARS-CoV-2 S (WT), SARS-CoV-2 S A.23.1 variant, SARS-CoV-2 S WT with P681R mutation. Data represents the average luciferase activity of cells of three (Vero E6 and Vero-TMPRSS2) or two (Calu-3) biological replicates.

The pseudoparticles were also probed for their S content via western blot and densitometry. For both A.23.1 and Wuhan Hu-1 P681R particles, we detected increased cleavage of the spike protein compared to Wuhan-Hu1 (WT); see **Figure 4.**

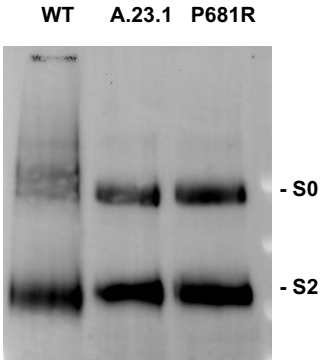


Figure 4. Western blot analysis of SARS-CoV-2 pseudoparticles generated with SARS-CoV-2 Wuhan Hu-1 S (WT), SARS-CoV-2 S A.23.1 variant, or SARS-CoV-2 S WT with P681R mutation. Uncleaved (S0) and cleaved (S2) bands are indicated. SARS-CoV-2 S was detected using a rabbit antibody against the SARS-CoV-2 S2 region; hence cleaved S1 is not visible. Equivalent MLV content was confirmed using a mouse antibody against MLVp30 (not shown).

Cell-cell fusion assays of A.23.1 spike

In order 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 the WT, A.23.1 and Wuhan-Hu-1 P681R spike gene. We then evaluated syncytia formation as a read-out of membrane fusion **(Figure 5).** We observed an increase in the syncytia formation following spike protein expression for either A.23.1 or Wuhan-Hu-1 P681H, compared to Wuhan-Hu-1 (WT). Vero-TMPRSS2 cells generally formed more extensive syncytia than VeroE6 cells, with similar trends observed. These data provide evidence that the P681R mutation increases membrane fusion activity of the SARS-CoV-2 spike protein under the conditions tested.

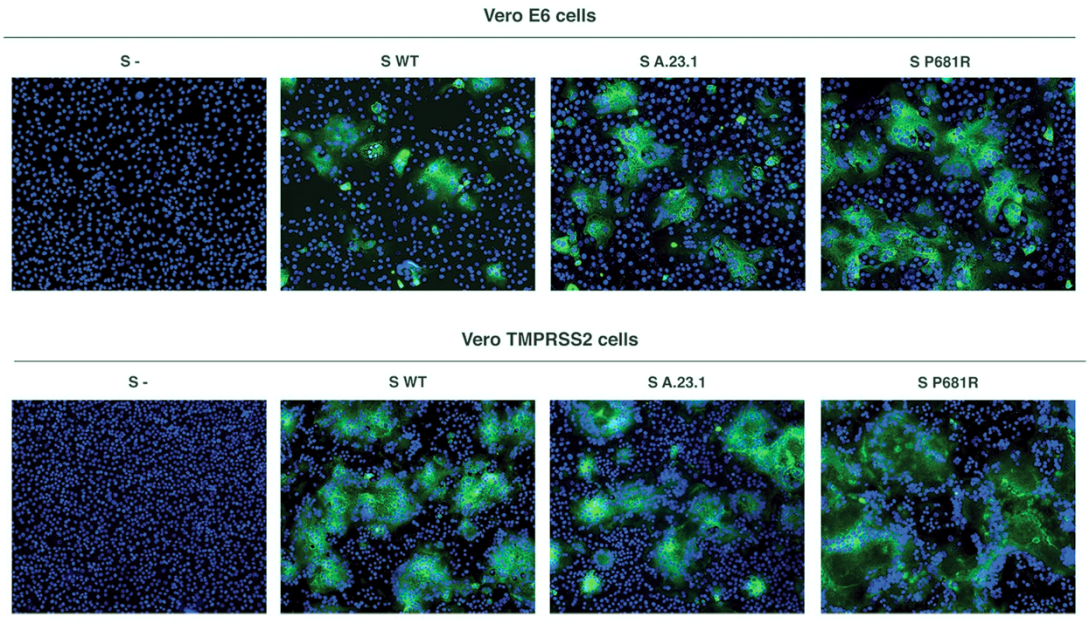


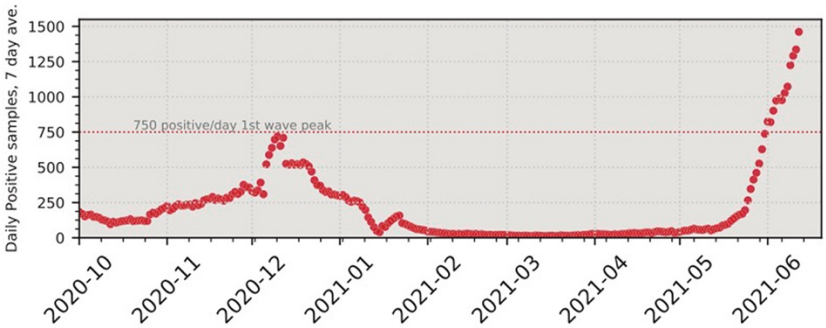
Figure 5. Cell-cell fusion assay of SARS-CoV-2 Wuhan-Hu-1 S (WT), SARS-CoV-2 S A.23.1 variant, or SARS-CoV-2 S WT with P681R mutation. S- = non-transfected cells. SARS-CoV-2 S was detected using a rabbit antibody against the SARS-CoV-2 S2 region.

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 of interest" (VOIs). These variants appear to have expanded following the selection for substitution or deletions in the spike protein, such as D614G and Q613H as well as other viral proteins. The substitutions encoded by VOCs and VOIs 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 (16-20).

To date, the most prevalent SARS-CoV-1 VOC of 2021 has been 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 p*K*a near neutrality (21), and so is not conventionally considered a basic amino acid.

Most recently, a new 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 (22, 23). However, the A.23.1 variant pre-dated B.1.617 as a P681R-containing VOC/VOI by several months. The A.23.1 emerged in Uganda in October 2020, to become the dominant virus in the country (up to 100% prevalence on January 2021) **(Figure 6)**. However, during the early months of 2021, A.23.1 prevalence declined to <1% prevalence in May 2021 and other variants were observed in Uganda, including VOC B.1.351 and VOI B.1.525, both without the P681R substitution. Similar to other regions of the world, the B.1.617.2 (Delta) lineage encoding P681R appeared and by June became the dominant observed lineage. **(Figure 6)**



Lineage	Oct-20	Nov-20	Dec-20	Jan-21	Feb-21	Mar-21	Apr-21	May-21	Jun-21
A.23.1	0.14	0.42	0.88	1	0.4	0.2	0.21	0.08	0
A.23	0.24	0.3	0.05	0	0.05	0	0	0	0
A	0.1	0	0	0	0	0	0	0	0
A.19	0	0	0	0	0	0	0.03	0	0
A.25	0	0.03	0.01	0	0	0	0	0	0
B	0	0	0	0	0	0	0	0	0
B.1	0.1	0.03	0.02	0	0	0	0	0	0
B.1.1	0	0	0	0	0	0	0.03	0	0
B.1.1.161	0	0	0.01	0	0	0	0	0	0
B.1.1.27	0.29	0	0	0	0	0	0	0	0
B.1.1.318	0	0	0	0	0.15	0	0.03	0	0
B.1.214	0.05	0.03	0	0	0	0	0	0	0
B.1.222	0	0.03	0.02	0	0.1	0	0	0	0
B.1.363	0	0.03	0	0	0	0	0	0	0
B.1.393	0.1	0.12	0	0	0	0	0	0	0
B.1.351	0	0	0	0	0.2	0.25	0.1	0	0
B.1.1.7	0	0	0	0	0	0.28	0.03	0.17	0.03
B.1.617.1	0	0	0	0	0	0.03	0	0	0
B.1.525	0	0	0	0	0.1	0.25	0.41	0.67	0
B.1.617.2	0	0	0	0	0	0	0.14	0.08	0.97

Month

Figure 6. Uganda SARS-CoV-2 cases and lineages, October 2020 to June 2021. Panel 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/). **Panel 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, (2) and the fractions of the months total genomes were plotted by month. Fractions were indicated in each cell and cells are colored (white to dark red) by increasing fraction.

Overall, these epidemiological data support the importance of the P681R substitution for community wide transmission. Both SARS-CoV-2 lineages that became the dominant lineages in Uganda encode this substitution, however several VOC and VOI lineages lacking the substitution, e.g., B.1.1.7 (Alpha), B.1.351 (Beta) and B.1.525 (Eta), also entered the country but did not achieve dominance (Figure 6). The subsequent decline of the P681R lineage A.23.1, combined with the in vitro analyses reported here clearly showed that the P681R alone is not sufficient to drive such dominance. Furthermore, the *in vitro* analyses reported here show that the substitution strongly influences furin-mediated cleavage, with these results being consistent with other studies (23, 24). However, it is very clear that P681R is not the sole driver of virus transmissibility, 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 produce the full activity of the A.23.1 spike (Figure 3).

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) (24). 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-simply 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.

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. Bugembe DL, V.T.Phan M, Ssewanyana I, Semanda P, Nansumba H, Dhaala B, Nabadda S, O'Toole ÁN, Rambaut A, Kaleebu P, Cotten M. 2021. A SARS-CoV-2 lineage A variant (A.23.1) with altered spike has emerged and is dominating the current Uganda epidemic. medRxiv doi:10.1101/2021.02.08.21251393:2021.02.08.21251393.
5. 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.
6. Polgár L. 1989. General Aspects of Proteases, p 43-86, Mechanisms of Protease Action. CRC press, Boca Raton, FL.
7. Seidah NG, Prat A. 2012. The biology and therapeutic targeting of the proprotein convertases. Nat Rev Drug Discov 11:367-83.
8. 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.
9. 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.
10. 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.
11. 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.
12. 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.
13. 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.
14. 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.
15. Duckert P, Brunak S, Blom N. 2004. Prediction of proprotein convertase cleavage sites. Protein Eng Des Sel 17:107-12.
16. 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.

17. 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.

18. 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.

19. 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.

20. 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.

21. Nelson DL, Cox MM. 2000. Lehninger Principles of Biochemistry. Worth Publishers.

22. 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.

23. 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.

24. 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.