Spike protein cleavage-activation in the context of the SARS-CoV-2 P681R mutation: an analysis from its first appearance in lineage A.23.1 identified in Uganda

Bailey Lubinski [1,2 #], Laura E. Frazier [1 #], My V.T. Phan [3], Daniel L. Bugembe [3], Jessie L. Cunningham [1], 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

[+] Lead corresponding

618 Tower Rd., Ithaca NY 14853, USA

jaj246@cornell.edu; grw7@cornell.edu

Abstract

Based on its predicted ability to affect transmissibility and pathogenesis, surveillance studies have highlighted the role of a specific mutation (P681R) in the S1/S2 furin cleavage site of the SARS-CoV-2 spike protein. Here we analyzed A.23.1, first identified in Uganda, as a P681R-containing virus several months prior to the emergence of B.1.617.2 (Delta variant). We performed assays using peptides mimicking the S1/S2 from A.23.1 and B.1.617 and observed significantly increased cleavability with furin compared to both an original B lineage (Wuhan-Hu1) and B.1.1.7 (Alpha variant). We also performed cell-cell fusion and functional infectivity assays using pseudotyped particles and observed an increase in activity for A.23.1 compared to an original B lineage spike. However, these changes in activity were not reproduced in the B lineage spike bearing only the P681R substitution. Our findings suggest that while A.23.1 has increased furin-mediated cleavage linked to the P681R substitution, this substitution needs to occur on the background of other spike protein changes to enable its 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 emerged as two main lineages, A and B, and now multiple sub-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

49 (Beta), B.1.1.28.1/P.1 (Gamma), B.1.617.2 (Delta), most of lineage A viruses remained at fairly lower
50 frequency and were more prevalent at the beginning of the pandemic in Asia. However, the A.23.1 viral
51 lineage is one of a few lineage A viruses that, due to local circumstances, became abundant in Uganda (5),
52 Rwanda (6) and South Sudan (7). A.23.1 evolved from the A.23 virus variant first identified in Uganda in
53 July 2020 and is characterized by three spike mutations F157L, V367F and Q613H (5). Subsequently, the
54 evolving A.23.1 lineage acquired additional spike substitutions (P681R), as well as in nsp6, ORF8 and ORF9
55 and with the acquisition of the E484K substitution A.23.1 was designated a variant under investigation
56 (VUI). By July 2021, the A.23.1 lineage has been observed with 1110 genomes reported from 47 countries
57 (GISAID, Pango Lineage report; https://cov-lineages.org/global_report_A.23.1.html).
58 Among several mutations in the A.23.1 lineage, the P681R mutation is of interest as it is part of a proteolytic
59 cleavage site for furin and furin-like proteases at the junction of the spike protein receptor-binding (S1) and
60 fusion (S2) domains (8). The S1/S2 junction of the SARS-CoV-2 S gene has a distinct indel compared to all
61 other known SARS-like viruses (Sarbecoviruses in *Betacoronavirus* lineage B)—the amino acid sequence of
62 SARS-CoV-2 S protein is $_{681}$-P-R-R-A-R|S-$_{686}$ with proteolytic cleavage (|) predicted to occur between the
63 arginine and serine residues depicted. Based on nomenclature established for proteolytic events (9), the R|S
64 residues are defined as the P1|P1' residues for enzymatic cleavage, with residue 681 of A.23.1 spike being
65 the P5 cleavage position. The ubiquitously-expressed cellular serine protease furin is highly specific and
66 cleaves at a distinct multi-basic motif containing paired arginine residues; furin requires a minimal motif of
67 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 (10). For SARS-
68 CoV-2, the presence of the S1/S2 "furin site" enhances virus transmissibility (11, 12). For the A.23.1 S,
69 P681R provides an additional basic residue at P5 and may modulate S1/S2 cleavability by furin, and hence
70 virus infection properties (13). Notably, the P681R substitution appears in several other lineages, most
71 notably B.1.617.2 (Delta) (N = 617590 genomes) but also the AY.X sub-lineages of B.1.617.2, B.1.617.1
72 (N= 6138), B.1.466.2 (N=2208), the B.1.1.7 sub-lineage Q.4 (2067), B.1.551 (N=722), AU.2 (N=302),
73 B.1.1.25 (N=509), B.1.466.2 (N=538), and other lineages (updated 22 Oct 21, https://outbreak.info;

74 https://outbreak.info/situation-reports?pango&muts=S%3AP681R) suggesting that the substitution may
75 provide an advantage for viruses encoding the substitution.
76 We previously studied the role of proteolytic activation of the spike protein of the lineage B SARS-CoV-2
77 isolates Wuhan-Hu1 and B.1.1.7 (14). Here, we used a similar approach to study the role of the proteolytic
78 activation of the spike protein in the context of the A.23.1 lineage virus, with a focus on the P681R
79 substitution to better understand the role of this notable change at the S1/S2 (furin) cleavage site.
80

81 **Results**

82 **Emergence and analysis of SARS-CoV-2 variants in Uganda, and evolution of the P681R mutation**
83 **and its role in the transmissibility and emergence of SARS-CoV-2**

84 A summary of the daily reported SARS-CoV-2 infections in Uganda is shown in Figure 1A, along with a
85 summary of SARS-CoV-2 lineage data in samples from Uganda (Figure 1B). The peak of infections in
86 December 2020-January 2021 corresponded to the circulation of the A.23.1 variant, which subsided, but
87 was replaced by a second larger peak of infections beginning in July 2021, primarily due to the emergence
88 of the B.1.617.2 variant (Delta), with additional variants being present over time. The circulation of the
89 A.23.1 variant appears to be fully displaced by the B.1.617.2 variant, which by July 2021 became the prevalent
90 variant in this country.
91 To further understand the evolution of the P681R substitution and its role in the transmissibility of SARS-
92 CoV-2, we monitored the frequency of substitutions at the S1/S2 (furin cleavage) site in the global
93 surveillance data and plotted these substitution data as fraction of total genomes. The initial B lineage virus
94 that spread out of Wuhan encoded a spike protein with P681 at the furin cleavage site, along with G614.
95 Fairly early in the epidemic, the D614G substitution appeared in the B.1 lineage and became prevalent in
96 May-December 2020 (Figure 2A). The B.1.1.7 (Alpha) lineage evolved from B.1, spread widely in the UK
97 and other regions, and encoded a P681H substitution in the G614D background. B.1.1.7 peaked globally in
98 March-April 2021 (Figure 2A). In most regions of the world, the B.1.617.2 (Delta) lineage encoding spike

99 D614G and P681R spread widely following its emergence in India in May-June 2021, and became the
100 dominant observed lineage globally (Figure 5A). In comparison, a distinct A lineage virus (A.23) containing
101 Q613H appeared in August-September 2020 and acquired the P681R mutation at a much earlier time
102 (December 2020-January 2021); however, it circulated only briefly (Figure 2B). These data suggest that while
103 P681R is important, it operates in the context of other viral mutations in the context of community spread,
104 with this functional context able to be addressed experimentally.
105

106 **Biochemical analysis of the SARS-CoV-2 A.23.1 S1/S2 cleavage site**

107 To gain insight into SARS-CoV-2 spike protein function and the proteolytic processing at the S1/S2 site,
108 we took a combined biochemical and cell-based strategy, with the rationale that along with other changes
109 in the spike protein, A.23.1, B.1.617.1 (Kappa) and B.1.617.2 (Delta) contain a P681R substitution at the
110 S1/S2 interface which may modulate spike protein function and that these mutations alter the furin cleavage
111 site—which can be monitored by analyzing downstream changes in the levels of cleaved products and in
112 virus-cell fusion and pseudoparticle activation. Sequences of representative S1/S2 sequences are summarized
113 in Figure 3A.
114 As an initial bioinformatic approach to assess biochemical function, we utilized the PiTou (15) and ProP
115 (16) protein cleavage prediction tools, comparing the spike proteins from A.23.1 to B.1.1.7 and the prototype
116 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
117 human respiratory betacoronaviruses (HCoV-HKU1 and HCoV-OC43) with identifiable furin cleavage sites
118 (Figure 3B). Both algorithms predicted an increase in the furin cleavage for the A.23.1 and B.1.617 lineages
119 compared to A.1/B.1, with PiTou also showing a marked increase compared to B.1.1.7. PiTou utilizes a
120 hidden Markov model specifically targeting 20 amino acid residues surrounding furin cleavage sites and is
121 expected to be a more accurate prediction tool. As expected, MERS-CoV showed a relatively low furin
122 cleavage score, with HCoV-HKU1 and HCoV-OC43 showing relatively high furin cleavage scores. Overall,

123 these analyses predict a distinct increase of furin cleavability for the spike protein of A.23.1 and B.1.617
124 lineage viruses compared to A.1/B.1 and B.1.1.7. lineage viruses.
125 To directly examine the activity of furin on the SARS-CoV-2 A.23.1 S1/S2 site, we used a biochemical
126 peptide cleavage assay to directly measure furin cleavage activity *in vitro* (17). The specific peptide sequences
127 used here were SARS-CoV-2 S1/S2 B.1.1.7 (TNSHRRARSVA), TNSPRRARSVA (Wuhan-Hu-1 S1/S2)
128 and TNSRRRARSVA (A.23.1 S1/S2). As predicted, furin effectively cleaved both the Wuhan-Hu-1 (WT)
129 and B.1.1.7 peptides, but with no significant differences (Figure 3C). Interestingly, and agreeing with the
130 PiTou prediction, we observed a significant increase in furin cleavage for the A.23.1 S1/S2 peptide (Figures
131 2B and C) compared to both Wuhan-Hu-1 (WT) and B.1.1.7. This comparative data with SARS-CoV S1/S2
132 sites reveals that the P681R substitution substantially increases cleavability by furin, beyond the small
133 increase noted previously for P681H (11).
134
135 **Cell-to-cell fusion assays of A.23.1 spike**
136 In order to assess the functional properties of the spike protein and to see if the P681R substitution provided
137 any advantage for cell-to-cell transmission or syncytia formation, we performed a cell-to-cell fusion assay in
138 which VeroE6 or Vero-TMPRSS2 cells were transfected with either the WT, A.23.1 or P681R spike gene.
139 We then evaluated syncytia formation as a read-out of membrane fusion. We observed an increase in the
140 syncytia formation following spike protein expression for either A.23.1 or Wuhan-Hu-1 harboring a P681R
141 mutation (P681R), compared to Wuhan-Hu-1 (WT) (Figure 4A). Vero-TMPRSS2 cells generally formed
142 more extensive syncytia than VeroE6 cells. This increase was evident by observation through fluorescence
143 microscopy, as well as by quantification of the syncytia and cell-to-cell fusion ratio (Figure 4A, B and C). An
144 increase in the number of nuclei involved in syncytia was observed in cells transfected with A.23.1 and
145 P681R S genes in both cell lines (Figure 4B). However, the increase was higher in Vero-TMPRSS2 cells in
146 all the three studied spike proteins. Membrane expressed spike cleavage also assessed using western blot
147 (Figure 4C). An increased cleavage ratio was observed in the A.23.1 and P681R membrane expressed spikes,

148 compared to WT. The cleavage ratio was similar in both cell lines. Band intensity was normalized to the
149 GLUT4 protein (housekeeping protein) band intensity. These data provide evidence that the P681R
150 mutation increases membrane fusion activity of the SARS-CoV-2 spike protein under the conditions tested.
151

152 **Functional analysis of virus entry using viral pseudoparticles**

153 To assess the functional importance of the S1/S2 site for SARS-CoV-2 entry, we utilized viral
154 pseudoparticles consisting of a murine leukemia virus (MLV) core displaying a heterologous viral envelope
155 protein to partially recapitulate the entry process of the native coronavirus. The pseudoparticles also contain
156 a luciferase reporter gene as well as the integrase activity to allow that integration into the host cell genome
157 to drive expression of quantifiable luciferase (MLVpp-SARS-CoV-2 S) (18). Using the HEK-293T cell line
158 for particle production, MLV pseudoparticles containing the spike proteins of A.23.1, Wuhan-Hu-1 SARS-
159 CoV-2 (WT), and a P681R point mutant of Wuhan-Hu-1 (P681R) were prepared. Positive-control particles
160 containing the vesicular stomatitis virus (VSV) G protein and negative-control particles (Δenvpp) lacking
161 envelope proteins were also prepared (not shown). Pseudoparticles were probed for their S content via
162 western blot (Figure 4A). Because SARS-CoV-2 S has an efficiently cleaved spike we also produced particles
163 under furin inhibition using dec-RVKR-CMK. This allowed changes in cleavage patterns between different
164 spike proteins to be visualized. We also treated the particles with exogenous furin (+ or – furin) to examine
165 the spike cleavage on both partially cleaved and uncleaved spikes, to further study the differences in furin
166 processing in the studied spikes. For both A.23.1 and P681R particles, we detected increased spike protein
167 cleavage compared to WT in the harvested particles (Figure 5A). Interestingly, we observed markedly
168 increased cleavage ratio for both A.23.1 and P681R spikes in the harvested pseudoparticles under furin-
169 inhibition conditions (Figure 5A), presumably as the enhanced furin cleavage motif site produced by the
170 P681R mutation rescued what may be a modest effect of dec-RVKR-CMK.

171 For SARS-CoV-2, furin is predicted to cleave during virus assembly and "prime" the spike protein at the
172 S1/S2 site for subsequent events during cell entry. However, a subsequent cleavage priming at a secondary

173 site (known as S2' site) is also needed to activate the spike's fusion machinery (1). SARS-CoV-2 is predicted
174 to enter Vero E6 cells using cathepsin L for activation during endosomal trafficking, in what is known as
175 the "late" pathway, whereas in Vero-TMPRSS2 is predicted to use a "early" pathway, with spike activated
176 by TMPRSS2 or other transmembrane serine proteases (TTSPs) at the cellular membrane (1). Here, we used
177 the Vero-TMPRSS2 and the Vero E6 cell lines, which are predicted to activate the SARS-CoV-2 using
178 TMPRSS2 and cathepsin L respectively. Considering that furin priming at the S1/S2 site normally occurs
179 during viral assembly, we used pseudoparticles that were produced without furin inhibitor, yielding cleaved
180 spike proteins. Vero-TMPRSS2 cells gave overall significantly higher luciferase signals indicative of more
181 efficient entry. In contrast, Vero E6 cells showed generally lowered infection levels. As expected, VSVpp
182 (positive control) pseudoparticles infected both cell lines with several orders of magnitude higher luciferase
183 units than the values reported with Δenvpp infection (data not shown). In Vero E6 cells, entry of A.23.1.
184 and P681R was lowered compared to wild type (Figure 5B). However, Vero-TMPRSS2 cells pseudoparticles
185 bearing the A.23.1 spike showed a significantly higher level of infection, indicating more efficient virus entry;
186 This was not reproduced for a P681R point mutant of Wuhan-Hu-1 (P681R), a result in line with previous
187 results indicating that other mutations in spike are needed for the increased cleavability imparted by the
188 P681R mutation to mediate enhanced virus infection.
189 As a further way to assess the entry mediated by A.23.1. and P681R-containing spike proteins, we tested
190 pseudoparticles in human lung A459 cells expressing ACE2 and TMPRSS2 (Figure 5C). These cells showed
191 a highly significant increase in infection by A.23.1 compared to Wuhan-Hu1. The point mutant of P681R
192 on the Wuhan-Hu1 background showed a decrease in infectivity, confirming that P681R (similarly to other
193 point mutants (19)) only has its functional consequence on the appropriate genetic background.
194

195 **Discussion**

196 Since late 2020, the evolution of the SARS-CoV-2 pandemic has been characterized by the emergence of
197 viruses bearing sets of substitutions/deletions, designated "variants of concern" (VOCs) and "variants under

198 investigation" (VUIs). These variants appear to have expanded following the selection for substitution or
199 deletions in the spike protein, such as D614G and Q613H, along with mutations in other viral proteins. The
200 substitutions encoded by such variants may alter virus characteristics including enhanced transmissibility and
201 antigenicity, some provide a direct advantage to avoid the changing developing immune responses in the
202 population due to prior exposure or vaccination as well as the social dynamics of the human population (4,
203 20-23). The specific case of the D614G is interesting, as this mutation have been shown to improve the
204 spike's open conformation for receptor binding, demonstrating an evolutionary advantage for the 614G
205 carrier virus (24). In fact, the explosive spread of COVID-19 cases can be tracked to the emergence of this
206 mutation, which provided the context for further evolution of the SARS-CoV-2 virus and the rising number
207 of new variants. The first notable SARS-CoV-2 VOC of 2021 was B.1.1.7 (Alpha), which among other
208 changes, encoded a P681H substitution in the spike S1/S2 furin cleavage site and has been linked to
209 increased transmissibility due to the presence of the additional basic amino acid, histidine (H). However,
210 histidine is unusual in that it has an ionizable side chain with a p*K*a near neutrality (25), and so is not
211 conventionally considered a basic amino acid. Most recently, the VOC (B.1.617.2, or Delta) has replaced
212 B.1.1.7(Alpha) as the dominant circulating virus globally, which like A.23.1 and sub-lineages B.1.617.1,
213 B.1.617.2 and B.1.617.3 encodes a P681R substitution, and is more conventionally "polybasic" in the S1/S2
214 cleavage motif, than the P681H of B.1.1.7 (Alpha) and is suggested to affect transmissibility and pathogenesis
215 (26). For the Delta variant (B.1.617.2), Saito *et al.* showed enhanced fusogenicity and viral entry in cells
216 expressing TMPRSS2 (Vero-TMPRSS2 and Calu-3) but lowered fusogenicity in Vero cells (27), with
217 equivalent results also shown by Peacock et al. in a range of TMPRSS2-expressing cells (28). Our data with
218 Vero cells in particular differ from those reported by Saito *et al.*, reinforcing the concept that cell-cell fusion
219 can be affected by many factors, including the specific growth conditions of the cells, and also suggesting
220 that other mutations in A.23.1 spike specifically affect fusion in non-TMPRSS2-expressing cells.
221 The A.23.1 variant pre-dated the B.1.617 lineage as a P681R-containing VOC/VOI by several months. It is
222 interesting to note that B.1.617.2 (Delta) has been shown to be a genetic outlier compared to other VOCs

223 (29), raising the question of whether P681R (found in A.23.1 and B.1.617) ultimately results in a more
224 successful viral variant compared to P681H, found in B.1.1.7 (Alpha) and B.1.529 (Omicron BA.1/BA.2).
225 It is important to note that both lineages that have temporarily dominated Uganda have encoded the spike
226 P681R substitution, but in combination with distinct changes in the spike protein. In all cases, the position
227 681 change occurred after a change of position 613/614 (B to B.1 to B.1.1.7 (Alpha), B to B.1 to B.1.617.2
228 (Delta), A to A.23 to A.23.1, B to B.1 to B.1.1.7 to Q.4), and most recently B to B.1 to B.1.1.529 (Omicron).
229 This timing and linkage can be seen in the lineage prevalence charts (Figure 5A and B) where for each major
230 lineage the position 613/614 changes predate the position 681 changes.
231 The analyses reported here show that the substitution influences furin-mediated cleavage at the *in vitro* level,
232 with these results being consistent with other studies (22, 23). However, P681R may not be the sole driver
233 of spike protein function in vivo—a finding reinforced by the molecular studies described here. It would be
234 of interest to understand the additional spike-associated changes that cooperate with P681R. The
235 introduction of P681R alone into the WT Wuhan-Hu-1 spike did not reproduce the full activity of the A.23.1
236 spike (Figures 3 and 4). One limitation of this study is that isolated viruses of the A.23.1 lineage are not
237 available for infectivity assays, and so our work relies on the use of epidemiological tools and the
238 reconstruction of virus infection in biochemical and cell-based assays. Another limitation is that our "wild-
239 type" B lineage virus contains D614 and not 614G. Despite these limitations, we consider that our data
240 support our conclusion that the spike mutation P681R—by itself—is a not a primary driver of virus
241 transmissibility in the population, with A.23.1 giving unique insight into these aspects of the ongoing
242 COVID-19 pandemic, but requires the full context of additional spike and other viral changes seen in A.23.1,
243 B.1.617.2 and Q.4 for transmission success.
244 A.23.1 was a key SARS-CoV-2 variant spreading within Africa during the early part of 2021, and has been
245 defined (along with C.11) as an African VOI (30) having have multiple mutations on the spike glycoprotein
246 and evolving in a clocklike manner along with other variants. Epidemiological data from Uganda support
247 the importance of the P681R substitution in A.23.1 for community-wide transmission. The subsequent

248 decline of the P681R lineage A.23.1 in Uganda, combined with the in vitro analyses reported here clearly
249 showed that the P681R alone is not sufficient to drive such dominance. The P681R lineage B.1.617.2 (Delta)
250 likely benefited from additional S and other substitutions and eventually dominated the Uganda epidemic by
251 June 2021 (Figure 5), similar to patterns globally.
252 While P681R does make the S1/S2 cleavage site more basic in nature, such variant cleavage sites are still not
253 "ideal" for furin—as originally found in the prototype furin-cleaved virus mouse hepatitis virus (MHV)
254 (RRARR|S) (26, 31). The introduction of an arginine residue did appear to be making S1/S2 more
255 "polybasic" as the pandemic continued and transmissibility increased. While we should not over-simply the
256 complex process of spike protein activation, it will be interesting to see whether this progression of basic
257 residue addition continues with future variants, towards that seen in established community-acquired
258 respiratory coronaviruses such as HCoV-HKU-1 or HCoV-OC43, with S1/S2 sequences of RRKRR|S and
259 RRSRR|A, respectively (26). The recent emergence of B.1.1.529 (Omicron), without P681R but containing
260 distinct changes in its S1/S2 cleavage site (N579K, P681H) and apparently distinct properties in regard to
261 spike protein antigenicity, protease activation and fusion (for example see ref. (32)), has reaffirmed the notion
262 that the coronavirus spike protein is highly adaptable.
263
264 **Materials and Methods**
265 *Cells*
266 All cell lines were grown at 37°C with 5% CO_2. Vero E6 (ATCC CRL-1586) and Hek293T (ATCC CRL-
267 3216) cells were grown in Dulbeco's Modified Eagle's Medium (DMEM)(Corning) with 10% HyClone™
268 FetalClone® II (Cytiva), and 1% HEPES (Corning). Vero TMPRSS2 cells (JCRB Cell Bank JCRB1819) were
269 grown in DMEM, 10% Hyclone, 1% HEPES, and 1% Geneticin (Gibco). A549 ACE2 TMPRSS2 cells
270 (Invivogen a549-hace2tpsa) were maintained in DMEM with 10% FBS, 0.5 µg/ml of Puromycin (Invivogen
271 ant-pr-1) and 300 µg/ml of Hygromycin Gold (Invivogen ant-hg-1).
272

273 *Furin prediction calculations*

274 Prop: CoV sequences were analyzed using the ProP 1.0 Server hosted at:
275 https://services.healthtech.dtu.dk/service.php?ProP-1.0. PiTou: CoV sequences were analyzed using the
276 PiTou V3 software hosted at: http://www.nuolan.net/reference.html.

277

278 *Fluorogenic peptide assays*

279 Fluorogenic peptide assays were performed as described previously with minor modifications (33). Each
280 reaction was performed in a 100 µl volume consisting of buffer, protease, and SARS-CoV-2 S1/S2 Wuhan-
281 Hu-1 (WT) (TNSPRRARSVA), SARS-CoV-2 S1/S2 B.1.1.7 (TNSHRRARSVA) or SARS-CoV-2 S1/S2
282 A.23.1 (TNSRRRARSVA) fluorogenic peptide in an opaque 96-well plate. For trypsin catalyzed reactions,
283 0.8 nM/well TPCK trypsin was diluted in PBS buffer. For furin catalyzed reactions, 1 U/well recombinant
284 furin was diluted in a buffer consisting of 20 mM HEPES, 0.2 mM CaCl2, and 0.2 mM β-mercaptoethanol,
285 at pH 7.5. Fluorescence emission was measured once per minute for 60 continuous minutes using a
286 SpectraMax fluorometer (Molecular Devices) at 30°C with an excitation wavelength of 330 nm and an
287 emission wavelength of 390 nm. Vmax was calculated by fitting the linear rise in fluorescence to the equation
288 of a line.

289

290 *Synthesis and cloning of the A.23.1 spike protein*

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

294

295 *Site-directed mutagenesis*

296 Primers
297 (ACCTGGCTCTCCTTCGGGAGTTTGTCTGG/CCAGACAAACTCCCGAAGGAGAGCCAGGT)

298 for mutagenesis were designed using the Agilent QuickChange Primer Design tool to create the P681R
299 mutation (CCA->CGA). Mutagenesis was carried out on a pCDNA-SARs2 Wuhan-Hu-1 S plasmid using
300 the Agilent QuickChange Lightning Mutagenesis kit (The original plasmid was generously provided by David
301 Veesler, University of Washington USA). The mutated pcDNA-SARS-CoV-2 Wuhan-Hu-1 P681R S
302 plasmid was used to transform XL-10 gold ultracompetent cells, which were grown up in small culture, and
303 then plasmid was extracted using the Qiagen QIAprep Spin Miniprep Kit. Sanger sequencing was used to
304 confirm the incorporation of the mutation.
305
306 *Cell-cell fusion assay*
307 Vero E6 and Vero-TMPRSS2 cells were transfected with a plasmid harboring the spike gene of the SARS-
308 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
309 the 681 amino acid position (P681R), or an empty pCDNA3.1+ (S-) plasmid, and evaluated through an
310 immunofluorescence assay (IFA) to quantify nuclei involved in syncytia formation. Transfection was
311 performed on 8-well glass slides at 90% confluent cells using Lipofectamine® 3000 (Cat: L3000075,
312 Invitrogen Co.), following the manufacturer's instructions and 250 ng of plasmid DNA per well were
313 transfected. Cells were then incubated at 37°C with 5% of CO2 for 28 hours. Syncytia were visualized
314 through fluorescence microscopy using a previously described method (14). The spike expression was
315 detected using a SARS-CoV-2 spike antibody (Cat: 40591-T62, Sino Biological Inc.) at 1/500 dilution for 1
316 hour. Secondary antibody labeling was performed using AlexaFluorTM 488 goat anti-rabbit IgG antibody
317 (Cat: A32731, Invitrogen Co.) at a 1/500 dilution for 45 minutes. Representative images of each treatment
318 group were used to calculate the percent of nuclei involved in the formation of syncytia. Images were taken
319 at 20X on the Echo Revolve fluorescent microscope (Model: RVL-100-M). Nuclei were counted manually
320 using the Cell Counter plugin in ImageJ (https://imagej.nih.gov/ij/). Cells that expressed the spike protein
321 and contained 4 or more nuclei were considered to be one syncytium.
322

323 *Cell surface expression of spike protein*

324 For analysis of cell surface expression Vero E6 and Vero TMPRSS2 cells were seeded at $5x10^5$ cells/ml in a
325 6 well plate. The following day, each well was transfected using polyethylenimine (PEI) and 1X Gibco® Opti-
326 Mem with 2000 ng of one of the following plasmids: SARS-CoV-2 Wuhan-Hu-1 S (WT), SARS-CoV-2
327 A.23.1 S, SARS-CoV-2 Wuhan-Hu-1 P681R, or an empty pCDNA3.1+ (S-) plasmid. 24 hours post-
328 transfection, expressed protein on Vero E6 and Vero-TMPRSS2 cells was analyzed through a cell surface
329 biotinylation assay western blot as described previously (34). Spike protein was detected via Western Blot
330 using the antibodies described in the cell-cell fusion assay. GLUT4 protein was used as a housekeeping
331 expression control and labeled using a GLUT4 monoclonal antibody (Cat: MA5-17175, Invitrogen Co.).

332

333 *Pseudoparticle generation*

334 Pseudoparticle generation was carried out using a murine leukemia virus (MLV)-based system as previously
335 described with minor modification (18). HEK-293T cells were seeded at $2.5x10^5$ cells/ml in a 6-well plate
336 the day before transfection. Transfection was performed using polyethylenimine (PEI) and 1X Gibco® Opti-
337 Mem (Life Technologies Co.). Cells were transfected with 800ng of pCMV-MLV *gag-pol*, 600ng of pTG-Luc,
338 and 600 ng of a plasmid containing the viral envelope protein of choice. Viral envelope plasmids included
339 pcDNA-SARS-CoV-2 Wuhan-Hu1 S as the WT, pcDNA-SARS-CoV-2 Wuhan-Hu-1 P681R S, and
340 pcDNA-SARS-CoV-2 A.23.1 S. pCAGGS-VSV G was used as a positive control and pCAGGS3.1+ was
341 used for an empty plasmid negative control (S-). 48 hours post-transfection, the supernatant containing the
342 pseudoparticles was removed, centrifuged to remove cell debris, filtered, and stored at -80°C.

343

344 *Pseudoparticle Infection Assay*

345 Infection assays were performed as previously described with minor adjustments (18). Vero E6 and Vero-
346 TMPRSS2 cells were seeded at $4.5x10^5$ cells/ml, while A549 ACE2 TMPRSS2 cells were seeded at $4x10^5$
347 cells/ml in a 24-well plate the day before infection. Cells were washed three times with DPBS and infected

348 with 200 µl of either VSV G, SARS-CoV-2 S, SARS-CoV-2 P681R S, SARS-CoV-2 A.23.1 S, or empty
349 plasmid (S-) pseudoparticles. Infected cells were incubated on a rocker for 1.5 hours at 37°C, then 300 µl of
350 complete media were added and cells were left at 37°C. At 72 hours post-infection, cells were lysed and the
351 level of infection was assessed using the Luciferase Assay System (Cat: E1501, Promega Co.). The
352 manufacturer's protocol was modified by putting the cells through 3 freeze/thaw cycles after the addition
353 of 100 µl of the lysis reagent. 10 µl of the cell lysate was added to 20 µl of luciferin, and then luciferase
354 activity was measured using the Glomax 20/20 luminometer (Promega Co.). Vero E6 and Vero-TMPRSS2
355 infection assays were replicated four times. A549 ACE2 TMPRSS2 infection assays were repeated 3 times.
356 Each assay was performed with three technical replicates. Vero E6 and Vero-TMPRSS2 infection assays
357 were completed using the same batch of pseudoparticles, while the A549 ACE2 TMPRSS2 infection assays
358 were carried out using a newly made batch.
359
360 *Western blot analysis of pseudoparticles*
361 A 3 ml volume of pseudoparticles was pelleted using a TLA-55 rotor with an Optima-MAX-E
362 ultracentrifuge (Beckman Coulter) for 2 hours at 42,000 rpm at 4°C. untreated particles were resuspended
363 in 30 µl DPBS buffer. Pseudopartices were generated as described in the pseudoparticle generation section,
364 with the furin inhibitor dec-RVKR-CMK (Cat:35-011, Tocris) being added during transfection to select
365 wells. For the + furin treated MLVpps, particles were resuspended in 30 µL of furin buffer consistent in 20
366 mM HEPES, 0.2 mM CaCl2, and 0.2 mM β-mercaptoethanol (at pH 7.0). Furin-treated particles were later
367 incubated with 6 U of recombinant furin for 3 h at 37 °C. Sodium dodecyl sulfate (SDS) loading buffer and
368 DTT were added to all samples and heated at 95°C for 10 minutes. Samples were separated on NuPAGE
369 Bis-Tris gel (Invitrogen) and transferred on polyvinylidene difluoride membranes (GE). SARS-CoV-2 S was
370 detected using a rabbit polyclonal antibody against the S2 domain (Cat: 40590-T62, Sinobiological) and an
371 AlexaFluor 488 goat anti-rabbit antibody. Bands were detected using the ChemiDoc Imaging software (Bio-

372 Rad) and band intensity was calculated using the analysis tools on Biorad Image Lab 6.1 software to
373 determine the uncleaved to cleaved S ratios.
374
375 *Uganda cases vs. viral lineages over time*
376 Daily reported SARS-CoV-2 infections The Uganda daily SARS-CoV-2 positive samples numbers were
377 retrieved from Our World in Data (https://ourworldindata.org/coronavirus) and the 7 day average was
378 determined. Uganda SARS-CoV-2 lineage data were generated from the MRC Uganda genomic data
379 deposited in the GISAID database (https://www.gisaid.org/). SARS-CoV-2 Pango lineages (2) were
380 determined using the pangolin module pangoLEARN (https://github. com/cov-lineages/pangolin).
381
382 *Spike position 681 and 613/614 changes in global data*
383 All available spike protein sequences were obtained from the GISAID database. The frequency of P681,
384 P681R, P681H, D614 D614G and D613H were counted by string matching using Ack
385 (http://beyondgrep.com/) to the major variations of the 88 amino acid peptide sequence (aa 605 to 691)
386 spanning the two relevant sites (D614_P681 (Wuhan_B), D614G_P681 (B.1), D614_P681H,
387 D614G_P681H (B.1.1.7), D614_P681R, Q613_P681 (Wuhan_B), Q613H_P681 (A.23), Q613H_P681R
388 (A.23.1), D614G_P681R (B.1.617.2). Fractions of available total genomes for each month encoding each
389 peptide variant were visualized in a heatmap. Additional changes at position H655Y (present in the Gamma
390 lineage) were also included in the count and fraction calculation but the total numbers were minor.
391
392 *Quantification and Statistical Analysis*
393 All statistical analysis was performed using GraphPad Prism for Mac OS X, GraphPad Software, San Diego,
394 California USA, www.graphpad.com. Two sample T-tests were used to compare SARS-CoV-2 Wuhan-Hu
395 1 to SARS-CoV-2 A.23.1 or SARS-CoV-2 P681R, with significant P values reported in the figure legends.
396 Standard deviation was calculated and included in graphs when appropriate.

397

398 **Acknowledgements**

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

408

409 **Author contributions**

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

414

415 **Declaration of Interests**

416 The authors manifest no conflict of interest.

417

418 **References**

419 1. Whittaker GR, Daniel S, Millet JK. 2021. Coronavirus entry: how we arrived at SARS-CoV-2.
420 Current Opinion in Virology 47:113-120.

421 2. Rambaut A, Holmes EC, O'Toole Á, Hill V, McCrone JT, Ruis C, du Plessis L, Pybus OG. 2020. A
422 dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. Nature
423 Microbiology 5:1403-1407.
424 3. Zhou B, Thao TTN, Hoffmann D, Taddeo A, Ebert N, Labroussaa F, Pohlmann A, King J, Steiner
425 S, Kelly JN, Portmann J, Halwe NJ, Ulrich L, Trüeb BS, Fan X, Hoffmann B, Wang L, Thomann L,
426 Lin X, Stalder H, Pozzi B, de Brot S, Jiang N, Cui D, Hossain J, Wilson MM, Keller MW, Stark TJ,
427 Barnes JR, Dijkman R, Jores J, Benarafa C, Wentworth DE, Thiel V, Beer M. 2021. SARS-CoV-2
428 spike D614G change enhances replication and transmission. Nature 592:122-127.
429 4. Harvey WT, Carabelli AM, Jackson B, Gupta RK, Thomson EC, Harrison EM, Ludden C, Reeve R,
430 Rambaut A, Peacock SJ, Robertson DL, Consortium C-GU. 2021. SARS-CoV-2 variants, spike
431 mutations and immune escape. Nature Reviews Microbiology 19:409-424.
432 5. Bugembe DL, Phan MVT, Ssewanyana I, Semanda P, Nansumba H, Dhaala B, Nabadda S, O'Toole
433 ÁN, Rambaut A, Kaleebu P, Cotten M. 2021. Emergence and spread of a SARS-CoV-2 lineage A
434 variant (A.23.1) with altered spike protein in Uganda. Nature Microbiology 6:1094-1101.
435 6. Butera Y, Mukantwari E, Artesi M, Umuringa Jda, O'Toole ÁN, Hill V, Rooke S, Hong SL, Dellicour
436 S, Majyambere O, Bontems S, Boujemla B, Quick J, Resende PC, Loman N, Umumararungu E,
437 Kabanda A, Murindahabi MM, Tuyisenge P, Gashegu M, Rwabihama JP, Sindayiheba R, Gikic D,
438 Souopgui J, Ndifon W, Rutayisire R, Gatare S, Mpunga T, Ngamije D, Bours V, Rambaut A,
439 Nsanzimana S, Baele G, Durkin K, Mutesa L, Rujeni N. 2021. Genomic sequencing of SARS-CoV-
440 2 in Rwanda reveals the importance of incoming travelers on lineage diversity. Nature
441 Communications 12:5705.
442 7. Bugembe DL, Phan MVT, Abias AG, Ayei J, Deng LL, Lako RLL, Rumunu J, Kaleebu P, Wamala
443 JF, Hm JJ, Lodiongo DK, Bunga S, Cotten M. 2021. SARS-CoV-2 Variants, South Sudan, January-
444 March 2021. Emerg Infect Dis 27:3133-3136.

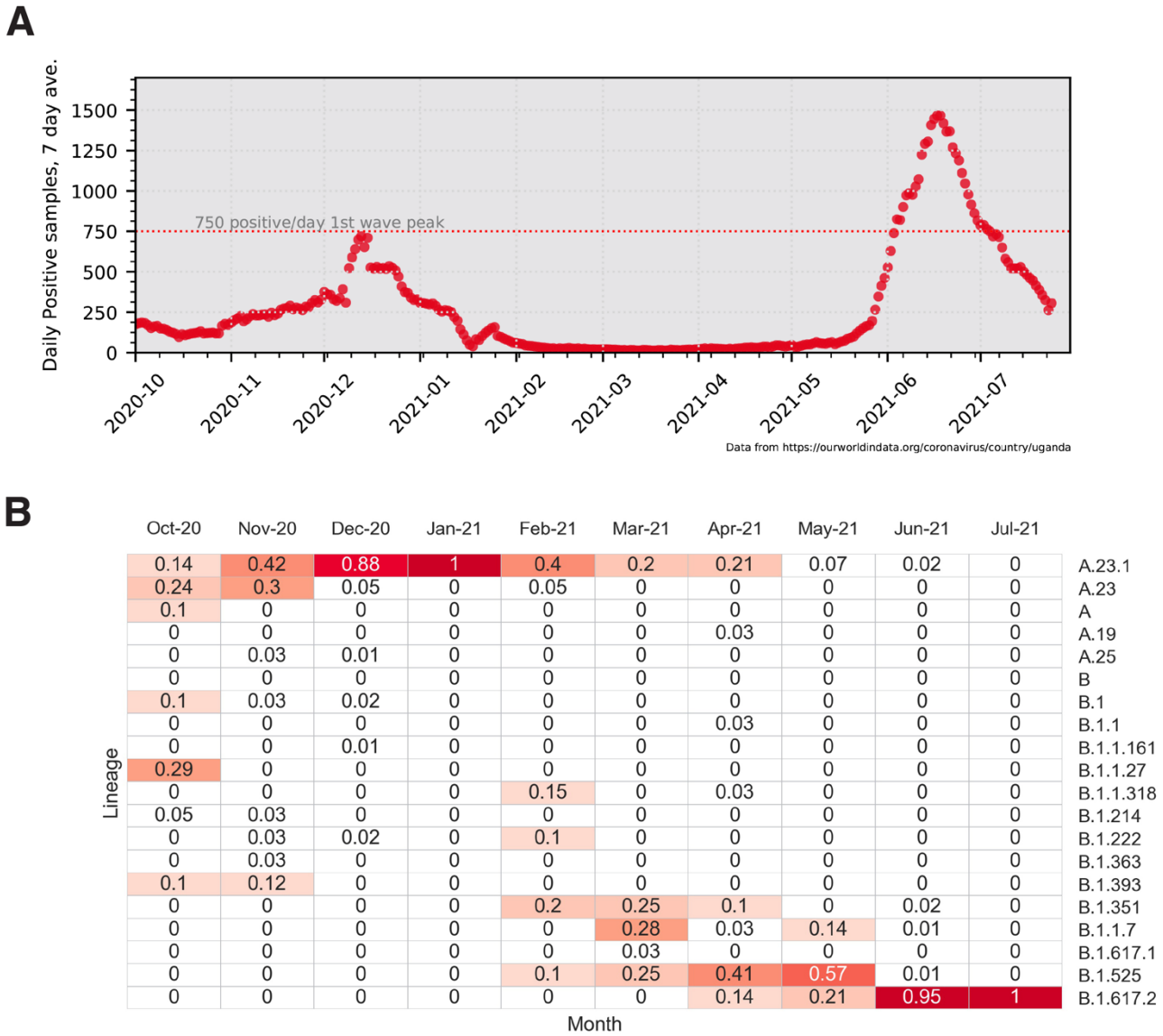
445 8. Jaimes JA, Andre NM, Chappie JS, Millet JK, Whittaker GR. 2020. Phylogenetic Analysis and
446 Structural Modeling of SARS-CoV-2 Spike Protein Reveals an Evolutionary Distinct and
447 Proteolytically Sensitive Activation Loop. J Mol Biol doi:10.1016/j.jmb.2020.04.009.
448 9. Polgár L. 1989. General Aspects of Proteases, p 43-86, Mechanisms of Protease Action. CRC Press.
449 10. Seidah NG, Prat A. 2012. The biology and therapeutic targeting of the proprotein convertases.
450 Nature Reviews Drug Discovery 11:367-383.
451 11. Johnson BA, Xie X, Bailey AL, Kalveram B, Lokugamage KG, Muruato A, Zou J, Zhang X, Juelich
452 T, Smith JK, Zhang L, Bopp N, Schindewolf C, Vu M, Vanderheiden A, Winkler ES, Swetnam D,
453 Plante JA, Aguilar P, Plante KS, Popov V, Lee B, Weaver SC, Suthar MS, Routh AL, Ren P, Ku Z,
454 An Z, Debbink K, Diamond MS, Shi P-Y, Freiberg AN, Menachery VD. 2021. Loss of furin cleavage
455 site attenuates SARS-CoV-2 pathogenesis. Nature 591:293-299.
456 12. Peacock TP, Goldhill DH, Zhou J, Baillon L, Frise R, Swann OC, Kugathasan R, Penn R, Brown
457 JC, Sanchez-David RY, Braga L, Williamson MK, Hassard JA, Staller E, Hanley B, Osborn M,
458 Giacca M, Davidson AD, Matthews DA, Barclay WS. 2020. The furin cleavage site of SARS-CoV-2
459 spike protein is a key determinant for transmission due to enhanced replication in airway cells.
460 bioRxiv doi:10.1101/2020.09.30.318311:2020.09.30.318311.
461 13. Hoffmann M, Kleine-Weber H, Schroeder S, Kruger N, Herrler T, Erichsen S, Schiergens TS,
462 Herrler G, Wu NH, Nitsche A, Muller MA, Drosten C, Pohlmann S. 2020. SARS-CoV-2 Cell Entry
463 Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. Cell
464 doi:10.1016/j.cell.2020.02.052.
465 14. Lubinski B, Fernandes MHV, Frazier L, Tang T, Daniel S, Diel DG, Jaimes JA, Whittaker GR. 2022.
466 Functional evaluation of the P681H mutation on the proteolytic activation of the SARS-CoV-2
467 variant B.1.1.7 (Alpha) spike. iScience 25:103589.
468 15. Tian S, Huajun W, Wu J. 2012. Computational prediction of furin cleavage sites by a hybrid method
469 and understanding mechanism underlying diseases. Sci Rep 2:261.

470 16. Duckert P, Brunak S, Blom N. 2004. Prediction of proprotein convertase cleavage sites. Protein
471 Engineering, Design and Selection 17:107-112.
472 17. Jaimes JA, Millet JK, Whittaker GR. 2020. Proteolytic Cleavage of the SARS-CoV-2 Spike Protein
473 and the Role of the Novel S1/S2 Site. iScience 23.
474 18. Millet JK, Tang T, Nathan L, Jaimes JA, Hsu HL, Daniel S, Whittaker GR. 2019. Production of
475 Pseudotyped Particles to Study Highly Pathogenic Coronaviruses in a Biosafety Level 2 Setting. J
476 Vis Exp doi:10.3791/59010.
477 19. Goodman LB, Whittaker GR. 2021. Public health surveillance of infectious diseases: beyond point
478 mutations. The Lancet Microbe 2:e53-e54.
479 20. Müller NF, Wagner C, Frazar CD, Roychoudhury P, Lee J, Moncla LH, Pelle B, Richardson M, Ryke
480 E, Xie H, Shrestha L, Addetia A, Rachleff VM, Lieberman NAP, Huang M-L, Gautom R, Melly G,
481 Hiatt B, Dykema P, Adler A, Brandstetter E, Han PD, Fay K, Ilcisin M, Lacombe K, Sibley TR,
482 Truong M, Wolf CR, Boeckh M, Englund JA, Famulare M, Lutz BR, Rieder MJ, Thompson M,
483 Duchin JS, Starita LM, Chu HY, Shendure J, Jerome KR, Lindquist S, Greninger AL, Nickerson
484 DA, Bedford T. 2021. Viral genomes reveal patterns of the SARS-CoV-2 outbreak in Washington
485 State. Science Translational Medicine 13:eabf0202.
486 21. Richard D, Shaw LP, Lanfear R, Corbett-Detig R, Hinrichs A, McBroome J, Turakhia Y, Acman M,
487 Owen CJ, Tan CCS, van Dorp L, Balloux F. 2021. A phylogeny-based metric for estimating changes
488 in transmissibility from recurrent mutations in SARS-CoV-2. bioRxiv
489 doi:10.1101/2021.05.06.442903:2021.05.06.442903.
490 22. Grubaugh ND, Hodcroft EB, Fauver JR, Phelan AL, Cevik M. 2021. Public health actions to control
491 new SARS-CoV-2 variants. Cell 184:1127-1132.
492 23. Peacock TP, Penrice-Randal R, Hiscox JA, Barclay WS. 2021. SARS-CoV-2 one year on: evidence
493 for ongoing viral adaptation. Journal of General Virology 102.

494 24. Yurkovetskiy L, Wang X, Pascal KE, Tomkins-Tinch C, Nyalile TP, Wang Y, Baum A, Diehl WE,
495 Dauphin A, Carbone C, Veinotte K, Egri SB, Schaffner SF, Lemieux JE, Munro JB, Rafique A,
496 Barve A, Sabeti PC, Kyratsous CA, Dudkina NV, Shen K, Luban J. 2020. Structural and Functional
497 Analysis of the D614G SARS-CoV-2 Spike Protein Variant. Cell 183:739-751.e8.
498 25. Nelson DL, Cox MM. 2000. Lehninger Principles of Biochemistry. Worth Publishers.
499 26. Whittaker GR. 2021. SARS-CoV-2 spike and its adaptable furin cleavage site. The Lancet Microbe
500 2:e488-e489.
501 27. Saito A, Irie T, Suzuki R, Maemura T, Nasser H, Uriu K, Kosugi Y, Shirakawa K, Sadamasu K,
502 Kimura I, Ito J, Wu J, Iwatsuki-Horimoto K, Ito M, Yamayoshi S, Loeber S, Tsuda M, Wang L,
503 Ozono S, Butlertanaka EP, Tanaka YL, Shimizu R, Shimizu K, Yoshimatsu K, Kawabata R,
504 Sakaguchi T, Tokunaga K, Yoshida I, Asakura H, Nagashima M, Kazuma Y, Nomura R, Horisawa
505 Y, Yoshimura K, Takaori-Kondo A, Imai M, Chiba M, Furihata H, Hasebe H, Kitazato K, Kubo H,
506 Misawa N, Morizako N, Noda K, Oide A, Suganami M, Takahashi M, Tsushima K, Yokoyama M,
507 Yuan Y, et al. 2022. Enhanced fusogenicity and pathogenicity of SARS-CoV-2 Delta P681R
508 mutation. Nature 602:300-306.
509 28. Peacock TP, Sheppard CM, Brown JC, Goonawardane N, Zhou J, Whiteley M, de Silva TI, Barclay
510 WS. 2021. The SARS-CoV-2 variants associated with infections in India, B.1.617, show enhanced
511 spike cleavage by furin. bioRxiv doi:10.1101/2021.05.28.446163:2021.05.28.446163.
512 29. Hill V, Du Plessis L, Peacock TP, Aggarwal D, Colquhoun R, Carabelli AM, Ellaby N, Gallagher E,
513 Groves N, Jackson B, McCrone JT, O'Toole Á, Price A, Sanderson T, Scher E, Southgate J, Volz
514 E, The C-gUKc, Barclay WS, Barrett JC, Chand M, Connor T, Goodfellow I, Gupta RK, Harrison
515 EM, Loman N, Myers R, Robertson DL, Pybus OG, Rambaut A. 2022. The origins and molecular
516 evolution of SARS-CoV-2 lineage B.1.1.7 in the UK. bioRxiv
517 doi:10.1101/2022.03.08.481609:2022.03.08.481609.

518 30. Wilkinson E, Giovanetti M, Tegally H, San JE, Lessells R, Cuadros D, Martin DP, Rasmussen DA,
519 Zekri A-RN, Sangare AK, Ouedraogo A-S, Sesay AK, Priscilla A, Kemi A-S, Olubusuyi AM,
520 Oluwapelumi AOO, Hammami A, Amuri AA, Sayed A, Ouma AEO, Elargoubi A, Ajayi NA,
521 Victoria AF, Kazeem A, George A, Trotter AJ, Yahaya AA, Keita AK, Diallo A, Kone A, Souissi A,
522 Chtourou A, Gutierrez AV, Page AJ, Vinze A, Iranzadeh A, Lambisia A, Ismail A, Rosemary A,
523 Sylverken A, Femi A, Ibrahimi A, Marycelin B, Oderinde BS, Bolajoko B, Dhaala B, Herring BL,
524 Njanpop-Lafourcade B-M, Kleinhans B, McInnis B, et al. 2021. A year of genomic surveillance
525 reveals how the SARS-CoV-2 pandemic unfolded in Africa. Science 374:423-431.
526 31. Sturman LS, Ricard CS, Holmes KV. 1985. Proteolytic cleavage of the E2 glycoprotein of murine
527 coronavirus: activation of cell-fusing activity of virions by trypsin and separation of two different
528 90K cleavage fragments. Journal of Virology 56:904-911.
529 32. Willett BJ, Grove J, MacLean OA, Wilkie C, Logan N, Lorenzo GD, Furnon W, Scott S, Manali M,
530 Szemiel A, Ashraf S, Vink E, Harvey WT, Davis C, Orton R, Hughes J, Holland P, Silva V, Pascall
531 D, Puxty K, da Silva Filipe A, Yebra G, Shaaban S, Holden MTG, Pinto RM, Gunson R, Templeton
532 K, Murcia PR, Patel AH, The C-GUKC, Haughney J, Robertson DL, Palmarini M, Ray S, Thomson
533 EC. 2022. The hyper-transmissible SARS-CoV-2 Omicron variant exhibits significant antigenic
534 change, vaccine escape and a switch in cell entry mechanism. medRxiv
535 doi:10.1101/2022.01.03.21268111:2022.01.03.21268111.
536 33. Jaimes JA, Millet JK, Goldstein ME, Whittaker GR, Straus MR. 2019. A Fluorogenic Peptide
537 Cleavage Assay to Screen for Proteolytic Activity: Applications for coronavirus spike protein
538 activation. J Vis Exp doi:10.3791/58892.
539 34. Sun X, Tse LV, Ferguson AD, Whittaker GR. 2010. Modifications to the Hemagglutinin Cleavage
540 Site Control the Virulence of a Neurotropic H1N1 Influenza Virus. Journal of Virology 84:8683-
541 8690.
542

543 **Figures:**



544

545 **Figure 1: Uganda SARS-CoV-2 cases and lineages, October 2020 to July 2021. A.** 7-day average positive
546 cases numbers were plotted by day, the peak of 750 cases/ per day observed in the first wave of infections
547 in January 2021 is indicated with a dotted line. Case data were obtained from Our World in Data
548 (https://ourworldindata.org/). **B.** Monthly SARS-CoV-2 lineage data for Uganda. All Uganda full genome
549 sequences from GISAID (https://www.gisaid.org/) were retrieved, lineage types using the Pango tool
550 (https://cov-lineages.org/resources/pangolin.html), and the fraction of each month's total genomes were
551 plotted. Fractions were indicated in each cell and cells are colored (white to dark red) by increasing fraction.
552

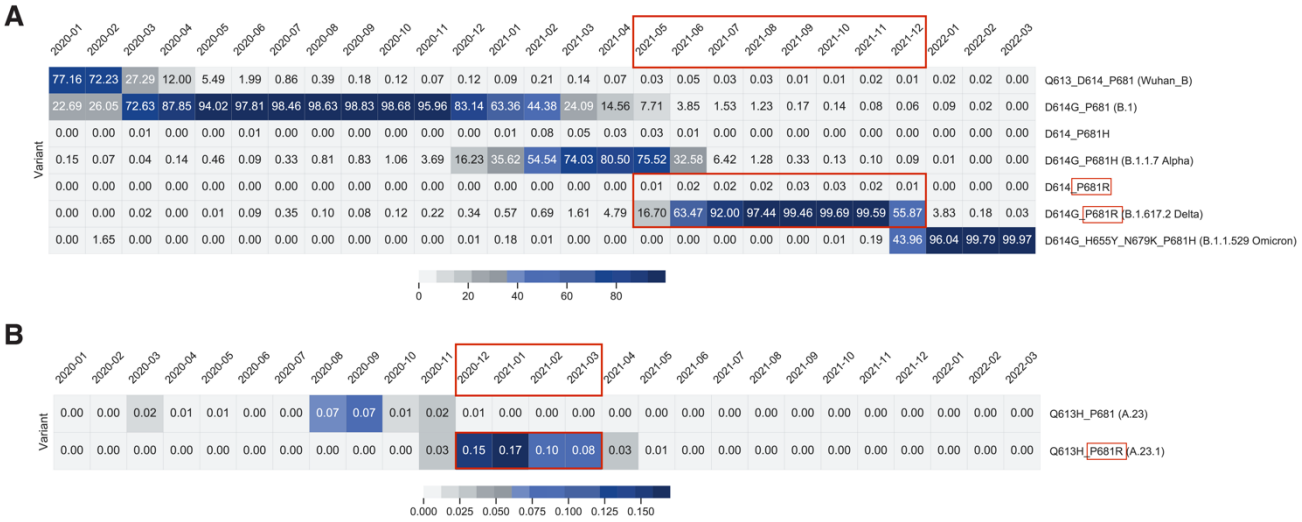
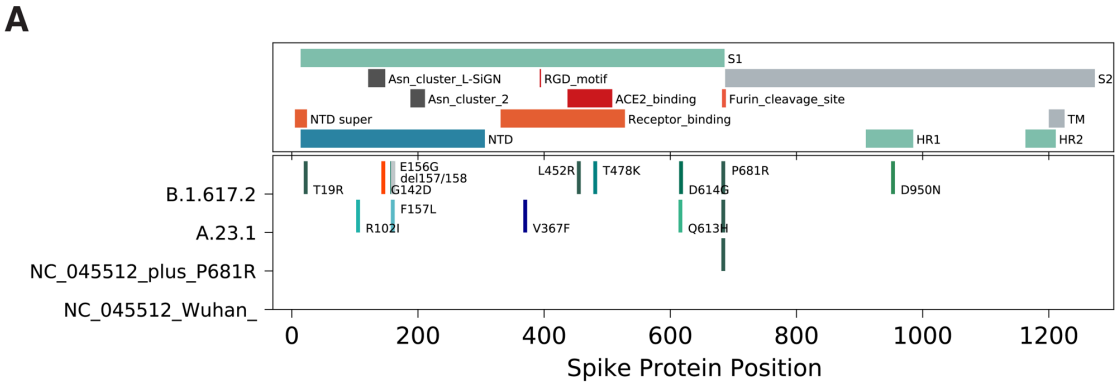


Figure 2. Frequency of P681, P681R, P681H, D614 D614G and D613H substitutions. The frequency of substitutions was counted by string matching to a peptide sequencing spanning the position 613 to 681) including the relevant sites at 613/614 and 681. (D614_P681 (Wuhan_B), D614G_P681 (B.1), D614_P681H, D614G_P681H (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. **A.** Lineage B relevant substitutions. **B.** Lineage A.23 and A.23.1 relevant substitutions. The time periods where P681R was dominant in each lineage are shown in red boxes.



B

Virus	S1/S2 Sequence	Furin Score PiTou	Furin Score ProP
SARS-CoV-2 (A.1/B.1)	672-ASYQTQTNSPRRAR|SVASQS-691	+9.196	0.626
SARS-CoV-2 (B.1.1.7)	669-ASYQTQTNSHRRAR|SVASQS-688	+9.907	0.706
SARS-CoV-2 (A.23.1)	669-ASYQTQTNSRRRAR|SVASQS-688	**+12.209**	**0.704**
SARS-CoV-2 (B.1.617.2)	669-ASYQTQTNSRRRAR|SVASQS-688	+12.209	0.704
MERS-CoV	738-LPDTPSTLTPRSVR|SVPGEM-757	+5.155	0.556
HCoV-HKU1	747-YNSPSSSSSRRKRR|SISASY-766	+14.634	0.868
HCoV-OC43 (clinical)	750-GYCVDYFKNRRSRR|AITTGY-769	+10.10	0.736

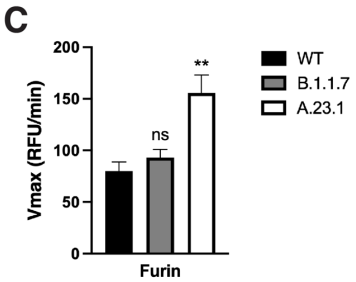
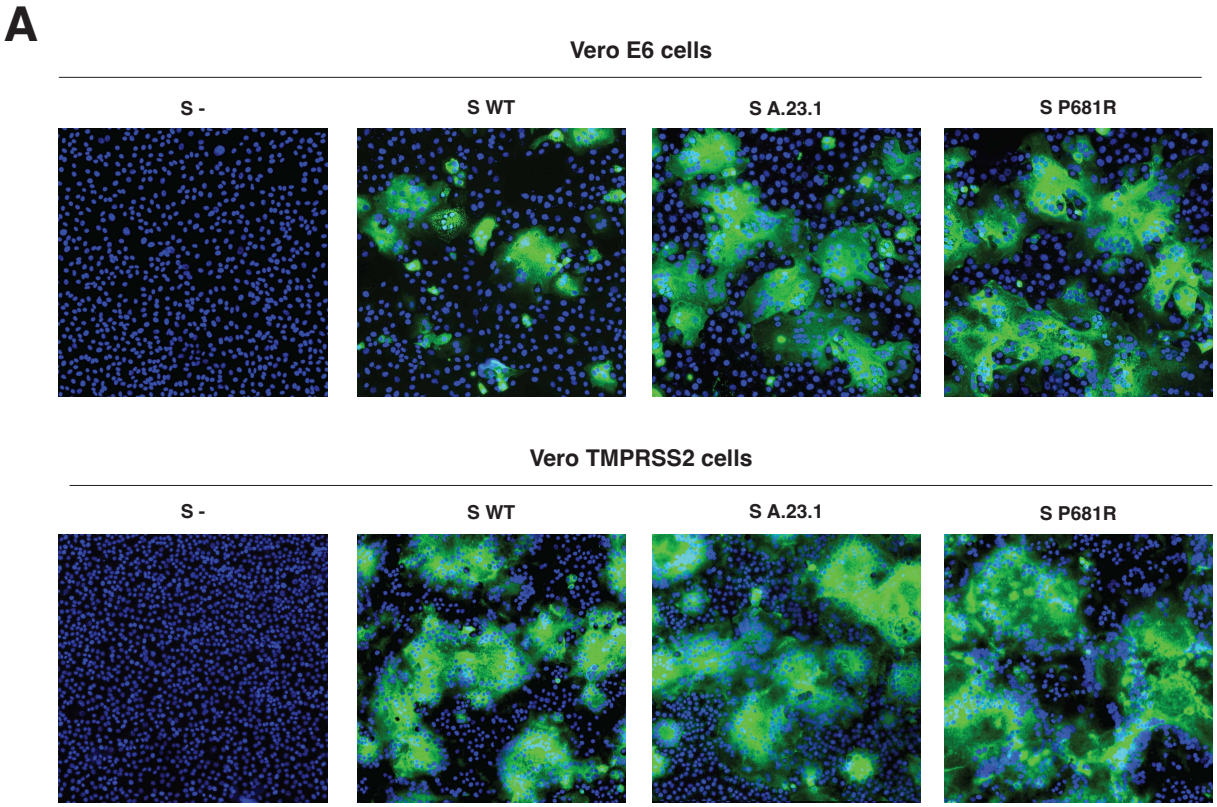


Figure 3: SARS-CoV-2 A.23.1 S sequence changes and S1/S2 furin cleavage. A. Summary of notable functional domains and sequence changes in the spike gene of A.23.1 compared to Wuhan-Hu-1 and B.1.617.2 (Delta). **B.** Furin cleavage score analysis of CoV S1/S2 cleavage sites. CoV S sequences were analyzed 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 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. **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. Error bars represent G standard errors (n = 9). Asterisks indicate statistical significance compared to the untreated control. Statistical analysis was performed using an unpaired Student's t test. ** $p < 0.01$.



B

Vero E6 cells			
	# Total nuclei	# Nuclei involved in syncytia	% Nuclei involved in syncytia
S -	6,824	0	0
S WT	4,866	345	7.09
S A.23.1	4,627	1,054	22.77
S P681R	4,703	955	20.30

Vero-TMPRSS2 cells			
	# Total nuclei	# Nuclei involved in syncytia	% Nuclei involved in syncytia
S -	13,051	0	0
S WT	10,441	7,718	73.92
S A.23.1	7,861	6,367	80.99
S P681R	10,747	8,033	74.74

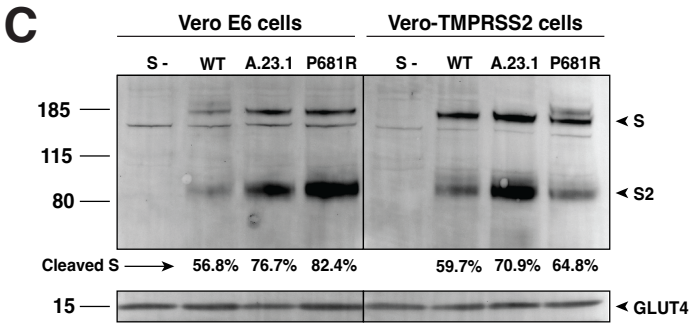


Figure 4: Cell-to-cell fusion in SARS-CoV-2 A.23.1 S expressing cells. A. Cell-to-cell fusion assay of SARS-CoV-2 Wuhan-Hu-1 S (WT), SARS-CoV-2 S A.23.1 variant, or SARS-CoV-2 S WT with P681R mutation. S- = non-transfected cells. SARS-CoV-2 S was detected using a rabbit antibody against the SARS-CoV-2 S2 region. **B.** Syncytia quantification by number of nuclei involved in syncytia. **C.** 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.

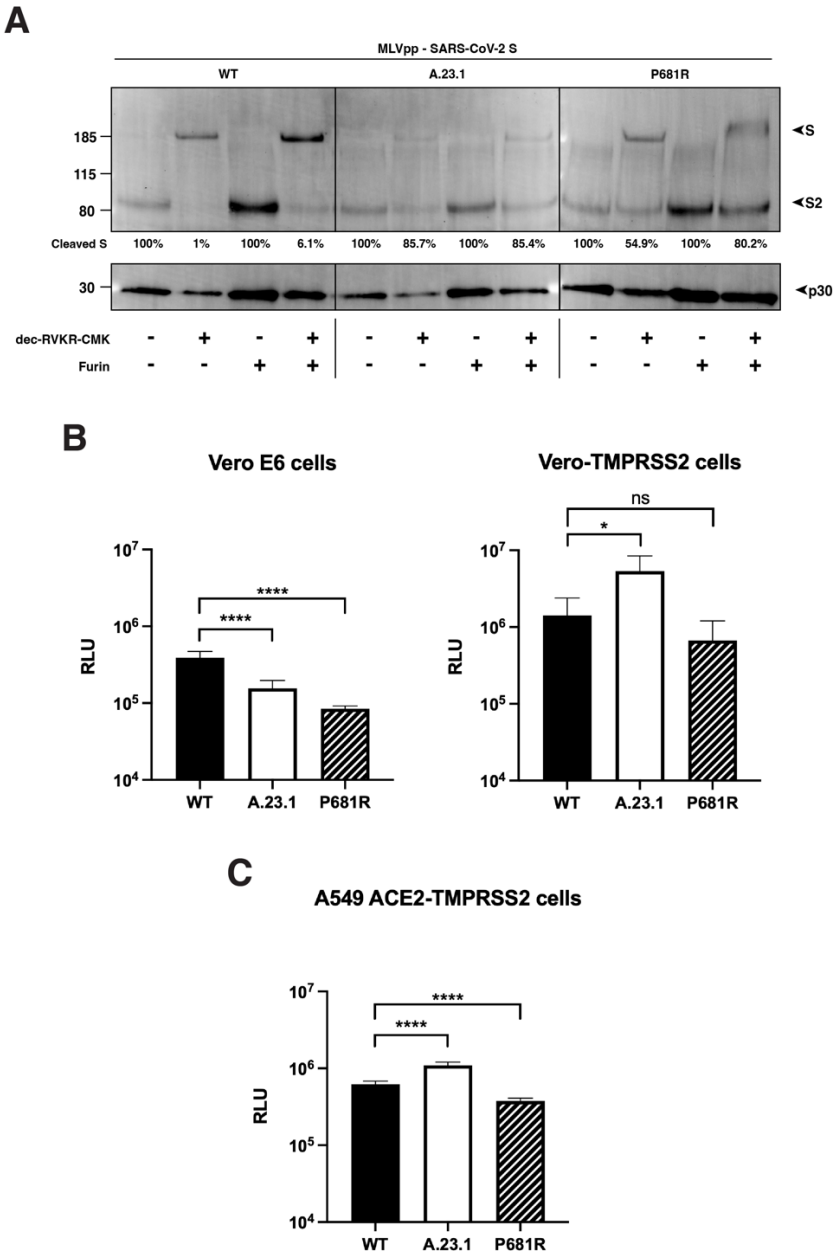


Figure 5: SARS-CoV-2 A.23.1 variant S1/S2 cleavage site activation and role in viral entry. A. Western blot analysis of MLVpp-SARS-CoV-2 S produced in ± dec-RVKR-CMK and treated with ± furin. S was detected using a rabbit antibody against the SARS-CoV-2 S2 subunit. MLV content was detected using a mouse antibody against MLV p30. **B.** Pseudoparticle infectivity assays in Vero E6 and Vero-TMPRSS2 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 four independent experiments (Vero E6 and Vero-TMPRSS2). Error bars represent G standard deviation (n = 3). Asterisks indicate statistical significance compared to the untreated control. Statistical analysis was performed using an unpaired Student's t test. * $p < 0.1$, **** $p < 0.0001$. **C.** Pseudoparticle infectivity assays in A549-ACE2-TMPRSS2 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 independent experiments. Error bars represent G standard deviation (n = 3). Asterisks indicate statistical significance compared to the untreated control. Statistical analysis was performed using an unpaired Student's t test. **** $p < 0.0001$.