Independent acquisition of short insertions at the RIR1 site in the spike N-terminal domain of the SARS-CoV-2 BA.2 lineage

Samuele Greco[1], Marco Gerdol[1*]

[1]Department of Life Sciences, University of Trieste

*to whom correspondence should be addressed: mgerdol@units.it

Summary

Although the major SARS-CoV-2 omicron lineages share over 30 non-synonymous substitutions in the spike glycoprotein, they show several unique mutations that were acquired after their ancestral split. One of the most intriguing mutations associated with BA.1 is the presence of the inserted tripeptide Glu-Pro-Glu within the N-terminal domain, at a site that had previously independently acquired short insertions in several other SARS-CoV-2 lineages. Although the functional implications of the small nucleotide sequences found at this insertion hotspot, named RIR1, are still unclear, we have previously hypothesized that they may play a compensatory role in counterbalancing minor fitness deficits associated with other co-occurring spike non-synonymous mutations.

Here we show that similar insertion events have independently occurred at RIR1 at least 20 times in early 2022 within the BA.2 lineage, being occasionally associated with significant community transmission. One of these omicron sublineages, characterized by a Ser-Gly-Arg insertion in position 212, has been responsible of over 4,000 documented covid-19 cases worldwide between January and July 2022, for the most part concentrated in Denmark, where it reached a national prevalence close to 4% (10% in the Nordjylland region) in mid-May. Although the concurrent spread of the BA.2.12.1, BA.4 and BA.5 lineages led to the rapid decline of this BA.2 sublineage, the independent acquisition of several other RIR1 insertions on a BA.2 genomic background suggests that these events may provide a slight fitness advantage. Therefore, we they should be carefully monitored in the upcoming months in other emerging omicron-related lineages, including BA.5.

Keywords

Spike glycoprotein; viral evolution; indels; omicron.

Introduction

Since its first emergence in human during the Wuhan outbreak at the end of 2019, the genome of SARS-CoV-2 has been acquiring new mutations compared with the Wuhan-Hu-1 reference sequence

at a rate close to 9×10^{-4} substitutions/site/year, based on GISAID data (Shu and McCauley, 2017). Nevertheless, in line with observations previously collected in other betacoronaviruses, such mutations do not display a uniform distribution, being disproportionately located within the spike glycoprotein, which has a key role in regulating viral entry through the interaction with receptors expressed on host cell membranes (Boni et al., 2020; Guo et al., 2020). Most specifically, the overwhelming majority of S gene mutations target the S1 subunit, which includes a N-terminal domain (NTD) and the C-terminal receptor-binding domain (RBD), responsible of ACE-binding (Li et al., 2003; Hoffmann et al., 2020). As a major target of neutralizing antibodies (Ju et al., 2020; Walls et al., 2020), S1 has been characterized by strong positive selection and by a very high rate of non-synonymous substitutions during the first two years of the pandemics (i.e. nearly 2×10^{-2} amino acid/year, based on GISAID data (Shu and McCauley, 2017)). Starting from late 2020, several variants of interest (VOIs) and variants of concern (VOCs), characterized by the presence of non-synonymous spike mutations, emerged in different geographical locations, quickly outcompeting pre-existing viral lineages. Some of these mutations, such as N501Y, most likely increased the intrinsic transmissibility of the early VOCs alpha, beta and gamma, by increasing ACE2 binding affinity (Teruel et al., 2021; Zhu et al., 2021). However, other mutations associated with these VOCs targeted important antibody epitopes, highlighting the role played by growing population immunity in selecting variants characterized by immune escape properties (Harvey et al., 2021; Sabino et al., 2021). A few months later, the global spread of delta rapidly brought the other variants to the verge of extinction, revealing that the combination between mutations conferring antibody escape (e.g. L452R) (Planas et al., 2021) and enhanced intrinsic transmissibility (e.g. P681R) (Saito et al., 2021) could result in the generation of novel viral variants with highly enhanced fitness.

In late November 2021 the fifth VOC, named omicron, started to quickly spread in South Africa, showing a concerning association with reinfections (Pulliam et al., 2022). Omicron replaced delta with unprecedented speed, becoming the dominant SARS-CoV-2 lineage worldwide by early 2022. Unlike previous variants, which never showed the ability to fully escape the neutralizing activity of vaccine-elicited or convalescent sera, omicron displayed significant antigenic divergence compared with previous variants, suggesting its classification as a distinct SARS-CoV-2 serotype (Simon-Loriere and Schwartz, 2022). These immunological properties derive from the presence of a very high number of S1 mutations, several of which have been shown to mediate ACE2 recognition and/or immune escape (Cameroni et al., 2022; Mannar et al., 2022). Curiously, many of these mutations occur at sites subjected to strong purifying selection both in previous SARS-CoV-2 variants and in bat coronaviruses, suggesting that they may cooperatively interact to mitigate slight fitness deficits associated with individual substitution, with a significant structural and functional impact (Martin et

al., 2022a). For example they may, to some extent, also explain the preferential use of the TMPRSS2-independent and cathepsin-mediated endosomal cell entry route (Pia and Rowland-Jones, 2022), as well as the lower fusogenity and altered cell tropism displayed by this VOC (Meng et al., 2022).

Omicron includes different major lineages, which may derive from a chronic infection (Tegally et al., 2022) and rapidly outcompeted each other, gaining regional or global prevalence during early 2022. While BA.1 and BA.2 were the first ones to cause major outbreaks in human, at the time of the writing of this manuscript BA.5, endowed with multiple immune escape mutations, is quickly emerging as the dominant omicron lineage in multiple geographic locations (Cao et al., 2022; Hachmann et al., 2022; Wang et al., 2022). Although the major omicron lineages share several S1 mutations, they display a remarkable sequence divergence, which is roughly comparable with the divergence observed among earlier VOCs. One of the most puzzling mutations only found in BA.1 is a short insertion of three codons, which encodes the tripeptide Glu-Pro-Glu at position 215, within the spike NTD. Nearly all VOCs and VOIs carry small NTD deletions compared with Wuhan-Hu-1, which may either confer enhanced antibody escape or act as permissive mutations, counterbalancing the fitness cost of otherwise deleterious RBD mutations (Meng et al., 2021). These small deletions independently occurred on multiple occasions during SARS-CoV-2 evolution at Recurrent Deletion Regions (RDRs), marking one of the most significant signatures of convergent evolution documented in SARS-CoV-2. On the other hand, the occurrence of spike insertions has been much rarer and consequently subjected to far less study. We have previously shown that insertions of short sequence stretches, usually encoding 3 or 4 amino acids, had been independently acquired over 50 times at the very same NTD site, named Recurrent Insertion Region 1 (RIR1). Before the emergence of BA.1, RIR1 insertions have been documented in alpha, gamma and delta, where they did not lead to significant community spread, as well as in two other lineages (A.2.5 and B.1.214.2) that had a considerable diffusion in 2021 (Gerdol et al., 2022). Although the functional role of RIR1 insertions is presently unclear, their convergent acquisition by multiple viral lineages most certainly suggests the need for increased monitoring, due to the possibility of associated fitness advantages.

While BA.1 has been rapidly replaced by other omicron lineages characterized by higher intrinsic transmissibility or immune escape during early 2022, we here show that a few BA.2 sublineages have independently acquired RIR1 insertions. One of these, carrying a Ser-Arg-Gly insertion at position 212 and hereby named BA.2+ins(L), marks an interesting case due to the slight growth advantage it displayed over the parental BA.2 lineage in Denmark, before the rise of BA.5 led to its decline.

Materials and Methods

All SARS-CoV-2 genome sequence data deposited in GISAID (Shu and McCauley, 2017) up to July 16th 2022, was screened, looking for sequences belonging to the BA.2 lineage and showing insertions between positions 210 and 220 in the S gene. Genomes displaying unidentified amino acids (i.e. X) in such insertions were flagged as the likely product of mis-assembly and discarded. The quality of all retrieved genomes was evaluated with NextClade CLI 2.3.0, removing and all sequences marked as "bad" or "mediocre". Moreover, genomes assigned to the BA.2 lineage but carrying an "EPE" insertion in position 215, previously described as one of the lineage-defining mutations of the sister lineage BA.1, were flagged as likely misclassified and were therefore discarded. All resulting genomes were grouped based on the inserted nucleotide sequence, whose exact position and phase of insertion were defined based on the alignment with the reference SARS-CoV-2 genome sequence (Wuhan-Hu-1, NCBI accession ID: NC_045512.2) and a reference BA.2 genome (EPI_ISL_8128502). To remove possible sequencing and bioinformatics artefacts, only the insertions independently obtained by at least two different laboratories (based on GISAID metadata) were considered for further analysis. Following the nomenclature scheme proposed in our previous work (Gerdol et al., 2022), each of these groups was labeled with progressive Roman numerals, following their chronological order of identification, starting from insertion L.

The molecular surveillance data from Denmark was subjected to further scrutiny, due to the local spread of one of the BA.2 sublineages carrying an insertion at RIR1, i.e. BA.2+ins(L). All genome data and associated metadata, deposited up to July 16th 2022, were downloaded from GISAID. This allowed to calculate the share of BA.2 genomes carrying insertion L, both at a national and at a regional scale (i.e. by separately taking into account Syddanmark, Sjælland, Nordjylland, Midtjylland and Hovestaden), up to July 7th 2022. The frequencies of observation of BA.2-ins(L) relative to all SARS-CoV-2 sequenced genomes were reported the 7-day moving average, with 95% confidence intervals.

Results and discussion

In total, we identified 20 independent events of insertion at RIR1 associated with the BA.2 lineage. These were chronologically ordered, from the least to the most recent one, and assigned Roman numerals, from insertion L to insertion LXVI, based on the naming scheme we have previously proposed (Gerdol et al., 2022), (**Table 1**).

The RIR1 insertions associated with BA.2 involved the acquisition of either two (a single case), three (ten cases), four (eight cases), or five (a single case) codons. However, their position was somewhat

different compared with those that arose in other pre-omicron variants throughout 2020 and 2021 (**Figure 1A**). Indeed, one insertion (i.e. LXVII) occurred within codon 211 at phase I, six (i.e. LI, LV, LXIII, LXV, LXVIII and LXIX) within codon 212 at phase II and eight others (i.e. L, LIII, LVI, LVII, LVIII, LIX, LXI and LXVI) within codon 213 at phase 0. None of the 49 RIR1 insertions described in our previous work was observed in these locations (Gerdol et al., 2022). The exact placement of the four remaining BA.2-associated RIR1 insertions could not be unambiguously ascertained due to the possibility of multiple alternative sequence alignments with the BA.2 reference.

The inserted di-, tri-, tetra- or penta-peptides in BA.2 were mostly placed between Leu212 and Val213 (albeit this residue is replaced by Gly in BA.2), and never observed between Arg214 and Asp215, the most common placement for pre-omicron RIR1 insertions (**Figure 1B**). Because of the out-of-frame nature of several insertions, in five cases (insertions LI, LV, LXV, LXVIII and LXIX) Leu212 was replaced with Phe and in a single case (insertion LXVII) Asn211 was replaced with Thr (**Figure 1B**). As of note, due to of the presence of the Val213Gly substitution in BA.2, the placement of most of these insertions by GISAID was often offset by a single amino acid (e.g. insertion L is reported as ins213GRG instead of ins212SGR). Consequently, a non-synonymous substitution was often incorrectly called in position 213 (e.g. Val213Ser instead of Val213Gly for insertion L).

As previously mentioned, only three out of the 49 RIR1 insertions described in our previous work were associated with a significant community spread, either locally, or globally: insertion III (lineage A.2.5), insertion IV (lineage B.1.214.2) and insertion XLI (lineage BA.1). The other 46 RIR1 insertions remained confined to small clusters counting no more than a few dozen sequenced cases, pointing out that not all RIR1 insertions were associated with the acquisition of increased fitness over previous variants. Due to the lack of structural overlap between RIR1 and important NTD antibody epitopes, we suggested that RIR1 insertions could mitigate slight negative fitness costs associated with the presence of non-synonymous RBD mutations affecting immune escape (Gerdol et al., 2022), as previously hypothesized for the H69/V70 deletion found in many VOCs and VOIs (Meng et al., 2021). The limited community spread of most SARS-CoV-2 variants bearing RIR1 insertions, together with their previously documented acquisition during *in vitro* experiments, (Committee for Medicinal Products for Human Use, 2021; Shiliaev et al., 2021), may suggest a relevant role of intra-host selection in their emergence.

Although BA.2 is currently rapidly declining worldwide due to the recent rise of BA.5, a few sublineages carrying insertions at RIR1 have been responsible of several hundred documented cases in early 2022 (**Table 1**). With the single exception of BA.2.41 (i.e. the sublineage carrying insertion LIII), all these lack of an official PANGO lineage designation. Five of these, whose continued

detection for several months indicates sustained community transmission (**Figure 2E**), will be briefly discussed below.

The most widespread BA-2 sublineage carrying an insertion at RIR1, with over 4,000 sequenced genomes worldwide, is BA.2+ins(L) is, first detected on January 7th 2022 and characterized by the presence of an in-frame TCCGGCAGA insertion between codon 212 and 213, which determines the acquisition of the tripeptide Ser-Gly-Arg in position 212 (**Figure 1**). The overwhelming majority of the genomes belonging to this sublineage derives from Denmark (79.5%), the country where it apparently first originated, followed by Germany (5.4%), Australia (4.3%), France (2.2%), USA (1.9%) and the United Kingdom (1.1%). As of July 16th, 2022, cases have however been detected in all continents, with reports from Austria, Belgium, Brazil, Canada, Croatia, Czech Republic, Estonia, Hong Kong, Israel, Italy, Japan, Luxembourg, Netherlands, New Zealand, Norway, Pakistan, Peru, Portugal, Singapore, Slovenia, South Africa, Spain, Sweden, Switzerland and Turkey (**Figure 2D**). The international spread of BA.2+ins(L) was likely the result of exportation from Denmark, where this this sublineage displayed a moderate (10-15%) weekly growth advantage over the parental BA.2 lineage from January to early April 2022. However, this trend was followed by a stationary phase (at a national prevalence comprised between 3 and 4%), which lasted approximately until mid-May (**Figure 2A**), coincident with the minimum in the daily number of nation-wide new cases, which fell below 1,000 for the first time since October 2021 (**Figure 2C**). Later, the sudden rise of other competing omicron sublineages (BA.2.12.1 and BA.4/BA.5 in particular) led the rapid decline of all BA.2 sublineages, including BA.2+ins(L) (**Figure 2B**).

The national spread of BA.2+ins(L) was initially driven by the Nordjylland region, where this sublineage was first detected, undergoing a significant expansion compared with BA.2 from mid-January up to its peak of prevalence, on May 11th, when this BA.2 sublineage accounted for approximately 9.9% of all cases recorded in the region. Midtjylland, the only Denmark region with land borders with Nordjylland, followed a similar trend: after the detection of the first cases in late January, the share of BA.2+ins(L) genomes progressively increased over time, reaching the peak at approximately 6.3% of all cases on May 13th. The first few BA.2+ins(L) cases were recorded in Syddanmark, the southernmost continental region in Denmark, in Hovestaden, the easternmost and capital region, in mid-January, whereas Sjælland, the largest insular region of Denmark, was the last Danish region to record this sublineage, on February 12th. Although not as widespread as in the two aforementioned regions, BA.2+ins(L) reached approximately 3.1% prevalence in Syddanmark (May 16th), 2.7% in Hovestaden (April 29th) and 2.6% in Sjælland (April 16th).

In the second half of May and during June 2022 the observation of BA.2+ins(L) became more and more sporadic in all Danish regions, to the point that only one out of over 5,000 genomes sequenced

in early July belongs to this lineage, even though a number of cases continue to be detected abroad (**Figure 2D**). Even though the emergence of BA.2.12.1, BA.4 and BA.5 in an epidemiological context with high population immunity and low incidence of infections led to the rapid demise of BA.2+ins(L), the moderate growth advantage displayed by this sublineage over the parental BA.2 for nearly three months in multiple Danish regions warrants further investigation. Indeed, it would be interesting to clarify whether this trend was driven by biological factors (e.g. an alteration of the structure and function of the spike protein) or rather by other epidemiological factors, such as founder effects, as previously observed for other SARS-CoV-2 lineages in the past (Hodcroft et al., 2021).

The second most widespread BA.2 sublineage characterized by an insertion at RIR1 is BA.2+ins(LIII)/BA.2.41, with nearly 1,000 sequenced genomes (**Figure 2E**, **Table 1**). First detected on February 5th, 2022 in England, BA.2.41 displays a longer insertion compared with BA.2+ins(L), i.e. ACAGTAGGAGGA, which is also found in-frame between codon 212 and 213, and encodes the tetrapeptide Thr-Val-Gly-Gly (**Figure 1**). About 60% of the 914 sequenced genomes carrying this insertion sequenced so far (as of July 16th, 2022) have been detected in England, where this sublineage is still occasionally sequenced, even though it never exceeded 1% prevalence over five months of community transmission. However, signs of international spread of BA.2.41 are marked by the identification of this sublineage in in Australia, Austria, Belgium, Canada, Czech Republic, Denmark, France, Georgia, Germany, Ireland, Israel, Italy, Netherlands, Norway, Portugal, Spain, Trinidad & Tobago, Turkey and USA.

Another notewhorty lineage is represented by BA.2+ins(LIV), characterized by a ACACCTACCTTA insertion, which occurs in an ambiguous site and could therefore either be interpreted either as ins:211LTPT or as ins:212:TPTL (**Figure 1**). This sublineage, first detected on February 19th 2022, mostly spread in Germany (where 63% of the genomes have been sequenced). However, it continued community transmission for at least four months (**Figure 2E**) led to its detection in seven other European countries, with two single cases detected outside the continent, in Israel and Canada.

Two additional relevant BA.2 sublineages carrying insertions at RIR1 mainly spread in the United States. BA.2+ins(LV), whose CATGGCGGAGCT out-of-frame insertion within codon 212 encodes the tetrapeptide Met-Ala-Glu-Leu (**Figure 1**), has been first identified on March 4th 2022, and has been detected other 67 times since, over a period of three months (**Table 1**, **Figure 2E**). Unlike the three aforementioned sublineages, insertion LV has been only spotted on four occasions abroad, pointing out limited exportation.

BA.2+ins(LX), carrying an in-frame GAAACAGCA insertion between codons 212 and 213, encoding the tripeptide Glu-Thr-Ala (**Figure 1**), was first identified on April 6th 2022 and has been so far responsible for 58 cases (**Table 1**, **Figure 2E**). Like insertion LV, insertion LX did not lead to significant clusters of infections abroad, with just three cases reported outside of the American continent.

Curiously, most BA.2-associated RIR1 insertions did not bear other relevant S1 non-synonymous substitutions targeting the RBD or known important epitopes recognized by neutralizing antibodies (**Table 1**), questioning their previously hypothesized "compensatory" role in stabilizing otherwise slightly deleterious mutations with significant effect on immune escape (Gerdol et al., 2022). For example, the 212:SGR spike insertion was the only lineage-defining non-synonymous mutation found in BA.2+ins(L). A notable exception was the presence of Y144del in BA.2+ins(LV), a deletion which occurs in RDR2 and is associated with antibody escape (McCarthy et al., 2021). BA.2+ins(LIII) also displayed the presence of Thr1213Ser, a S2 non-synonymous conservative substitution rarely observed in SARS-CoV-2 prior to the emergence of BA.2, which helped its designation as BA.2.41.

The independent spread and continued community transmission of these sublineages suggests that the acquisition of RIR1 insertions on a BA.2 genomic background may occasionally provide a moderate fitness increase, even though the precise mechanism by which this would occur is still unclear. It is worth mentioning that several the mutations shared by the different omicron lineages, if taken individually, would theoretically have a detrimental impact viral fitness. Nevertheless, they are thought to cooperatively interact by positive epistasis, mitigating the negative fitness costs associated with their presence, being possibly responsible of the remarkable structural and functional alterations of the omicron spike protein (Martin et al., 2022b). Insertion XLI (i.e. ins215EPE), which is included among the mutations found in BA.1 but not shared with BA.2 and BA.5, undoubtedly emerged after the split between these lineages and its presence may therefore provide a fitness advantage in combination with any of the other spike mutations associated with BA.1, but absent in BA.2 or BA.5. Nevertheless, none of these mutations was observed in any of the BA.2 sublineages carrying RIR1 insertions (**Table 1**). Interestingly, the only study carried out so far which has explicitly tested the functional role of ins214EPE in BA.1 suggests that this insertion may improve spike expression (Javanmardi et al., 2022), which may be consistent with the significant structural alterations determined by RIR1 insertions on the spike protein trimer (Gerdol et al., 2022). Omicron displays enhanced inter-domain and inter-subunit packing, as well as to a higher accessibility for ACE2 binding, driven by its higher predisposition towards an open configuration (Ye et al., 2022), and it would be of high interest to investigate whether the presence of RIR1 insertions have a significant impact on these properties. Undoubtedly, further structural studies would be required to clarify the effects of the acquisition of RIR1 insertions on a BA.2 genetic background in

terms of the impact on the structural conformation and stability of spike trimers. Moreover, continued monitoring of similar insertions would be important for other emerging omicron lineages, such as BA.5, which as of July 16th 2022 only displays a single sequenced genome with a RIR1 insertion (i.e. EPI_ISL_13773295, carrying an in-frame AGAACAACCTAC insertion between codons 212 and 213, encoding the tetrapeptide Arg-Thr-Thr-Tyr).

Table and figure captions

Table 1: Summary of the 20 independent RIR1 insertions identified in the SARS-CoV-2 BA.2 lineage. *only mutations associated with at least 50% of the sequenced genomes are reported; **Latest detection, as of April 16th, 2022.

Design ation	Insertion	Insertion type	GISAID entries	Other spike mutations not shared with BA.2*	Earliest detection	Latest detection**
L	212:SGR	in frame (codon 213 phase 0)	4002	/	Jan 7th, 2022	Jul 5th, 2022
LI	212:NNR	out-of-frame (codon 212 phase 2)	2	G75D, L212F	Feb 4th, 2022	Feb 7th, 2022
LIII	212:TVGG	in frame (codon 213 phase 0)	914	T1231S	Feb 5th, 2022	Jul 4th, 2022
LIV	211:LTPT/212:TPTL	ambiguous; in frame (codon 212 phase 0/codon 213 phase 0)	128	/	Feb 19th, 2022	Jun 16th, 2022
LV	212:MAEL	out-of-frame (codon 212 phase 2)	68	Y144del, L212F	Mar 4th, 2022	Jun 5th, 2022
LVI	212:TGNTL	in frame (codon 213 phase 0)	44	/	Mar 14th, 2022	Jun 20th, 2022
LVII	212:SEE	in frame (codon 213 phase 0)	3	/	Mar 25th, 2022	May 23rd, 2022
LVIII	212:QGK	in frame (codon 213 phase 0)	31	/	Mar 28th, 2022	May 24, 2022
LIX	212:QEG	in frame (codon 213 phase 0)	2	/	Apr 5th, 2022	Apr 8th, 2022
LX	212:ETA	ambiguous (codon 213 phase 0/phase I)	58	/	Apr 6th, 2022	Jun 20th, 2022
LXI	212:IRQ	in frame (codon 213 phase 0)	50	/	Apr 11th, 2022	Jun 11th, 2022
LXII	212:EIVS	ambiguous (codon 213 phase 0/phase I)	8	/	Apr 13th, 2022	Apr 28th, 2022
LXIII	212:AE	out-of-frame (codon 212 phase 2)	4	/	Apr 14th,	May 15th,

|						2022	2022
LXIV	211:FNTY/212:NTYL	ambiguous (codon 211 phase 2/codon 212 phase 0/1/2)	3	/		Apr 16th, 2022	May 19th, 2022
LXV	212:RDG	out-of-frame (codon 212 phase 2)	2	L212F		Apr 17th, 2022	Apr 20th, 2022
LXVI	212:TME	in frame (codon 213 phase 0)	2	P1143S		Apr 21st, 2022	Apr 25st, 2022
LXVII	211:REPD	out-of-frame (codon 211 phase 1)	3	N211T		May 5th, 2022	Jul 5th, 2022
LXVIII	212:SEAG	out-of-frame (codon 212 phase 2)	6	L212F		May 11th, 2022	Jun 11th, 2022
LXIX	212:SKI	out-of-frame (codon 212 phase 2)	2	L212F		May 16th, 2022	May 16th, 2022
LXX	212:GVER/213:VERG	ambiguous; out-of-frame (codon 212 phase 2)/in-frame (codon 213 phase 0)	2	/		May 14th, 2022	Jun 1st, 2022

Figure 1: Panel A: Multiple sequence alignment of the nucleotide sequences of the S gene of viral genomes belonging to the BA.2 lineage bearing insertions at RIR1, compared with the reference sequences of Wuhan Hu-1 and BA.2. The multiple sequence alignment only shown a small region of the S gene (i.e. codon 210- codon 214). Red vertical bars highlight codon boundaries. Panel B: amino acid sequences encoded by the nucleotide sequences displayed in panel A. Codon/residue numbering refers to the Wuhan-Hu-1 reference sequence. Asterisks mark ambiguous insertion placements (see **Table 1** for details).
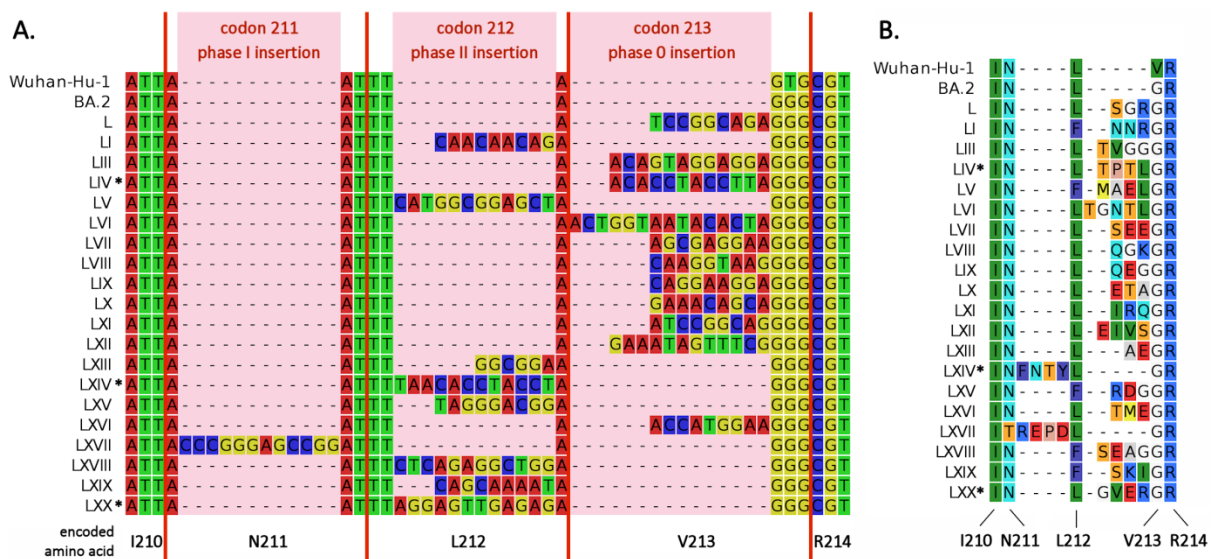
Figure 2: Panel A: Share of BA.2 genomes bearing the L insertion at RIR1 in Denmark. The graph reports the 7-day moving average frequencies of observations of BA.2+ins(L) genomes, relative to all

sequenced SARS-CoV-2 genomes, with 95% confidence intervals, up to July 7th 2022. Panel B: prevalence of the main viral lineages circulating in Denmark in 2022, up to July 7th 2022. These include BA.1 (and related sublineages), BA.2+ins212(L), BA.2.12.1, BA.2 (and related sublineages, except BA.2+ins212(L) and BA.2.12.1) and BA.4/BA.5. Panel C: 7-day moving average of daily new covid-19 cases observed in Denmark in 2022, up to July 7th 2022. Panel D: timing of the detection of sequenced BA.2 genomes carrying insertion L in five Denmark regions and other different countries, based on sampling date. Only countries with >= 10 sequenced genomes are reported, whereas the others were collapsed in geographic macroareas. Panel E: timing of the detection of sequenced BA.2 genomes carrying relevant spike insertions at RIR1, namely insertion L (ins212:SGR), LIII (ins212:TVGG), LIV (ins211:LTPT/212:TPTL), LV (ins212:MAEL) and LX (ins212:ETA), worldwide, based on sampling date.
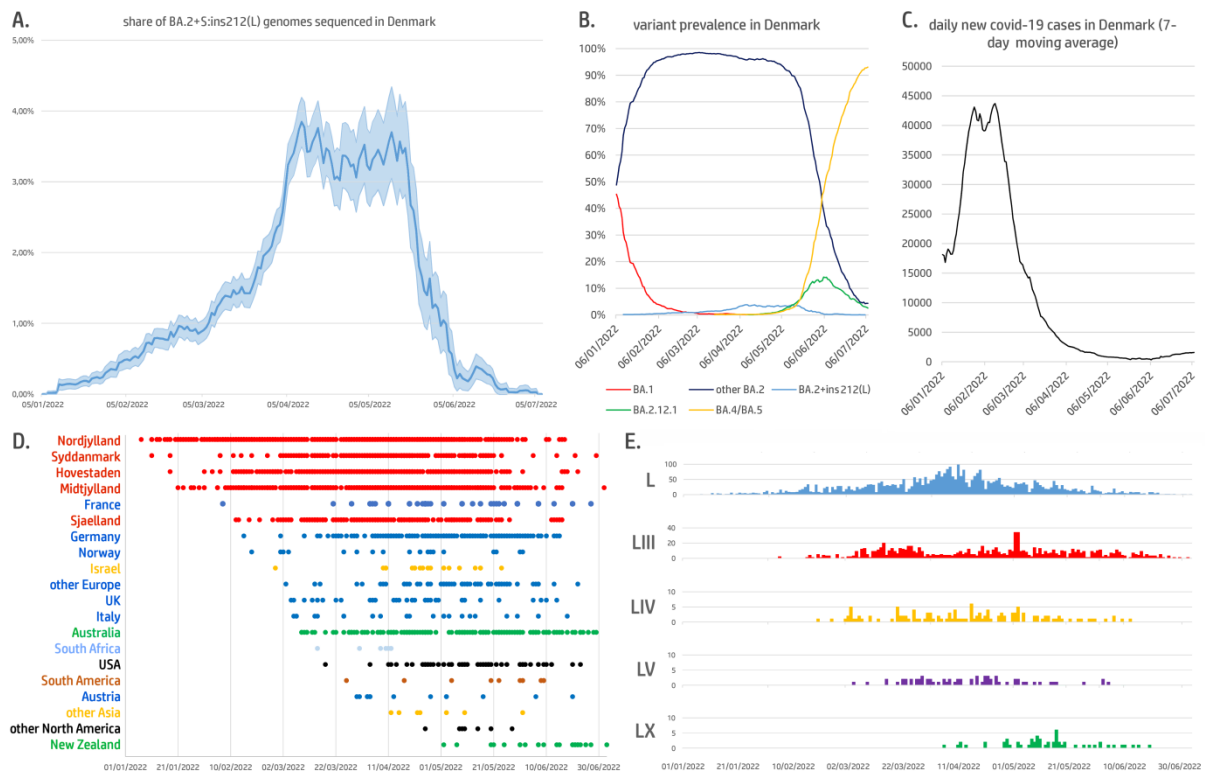
Conflict of Interest Statement

The authors declare they have no competing interests.

Ethics Statement

The authors confirm that the ethical policies of the journal, as noted on the journal's author guidelines page, have been adhered to. No ethical approval was required as this is a review article with no original research data.

References

Boni, M.F., P. Lemey, X. Jiang, T.T.-Y. Lam, B.W. Perry, T.A. Castoe, A. Rambaut, and D.L. Robertson, 2020: Evolutionary origins of the SARS-CoV-2 sarbecovirus lineage responsible for the COVID-19 pandemic. *Nature Microbiology* **5**, 1408–1417, DOI: 10.1038/s41564-020-0771-4.

Cameroni, E., J.E. Bowen, L.E. Rosen, C. Saliba, S.K. Zepeda, K. Culap, D. Pinto, L.A. VanBlargan, A. De Marco, J. di Iulio, F. Zatta, H. Kaiser, J. Noack, N. Farhat, N. Czudnochowski, C. Havenar-Daughton, K.R. Sprouse, J.R. Dillen, A.E. Powell, A. Chen, C. Maher, L. Yin, D. Sun, L. Soriaga, J. Bassi, C. Silacci-Fregni, C. Gustafsson, N.M. Franko, J. Logue, N.T. Iqbal, I. Mazzitelli, J. Geffner, R. Grifantini, H. Chu, A. Gori, A. Riva, O. Giannini, A. Ceschi, P. Ferrari, P.E. Cippà, A. Franzetti-Pellanda, C. Garzoni, P.J. Halfmann, Y. Kawaoka, C. Hebner, L.A. Purcell, L. Piccoli, M.S. Pizzuto, A.C. Walls, M.S. Diamond, A. Telenti, H.W. Virgin, A. Lanzavecchia, G. Snell, D. Veesler, and D. Corti, 2022: Broadly neutralizing antibodies overcome SARS-CoV-2 Omicron antigenic shift. *Nature* **602**, 664–670, DOI: 10.1038/s41586-021-04386-2.

Cao, Y., A. Yisimayi, F. Jian, W. Song, T. Xiao, L. Wang, S. Du, J. Wang, Q. Li, X. Chen, Y. Yu, P. Wang, Z. Zhang, P. Liu, R. An, X. Hao, Y. Wang, J. Wang, R. Feng, H. Sun, L. Zhao, W. Zhang, D. Zhao, J. Zheng, L. Yu, C. Li, N. Zhang, R. Wang, X. Niu, S. Yang, X. Song, Y. Chai, Y. Hu, Y. Shi, L. Zheng, Z. Li, Q. Gu, F. Shao, W. Huang, R. Jin, Z. Shen, Y. Wang, X. Wang, J. Xiao, and X.S. Xie, 2022: BA.2.12.1, BA.4 and BA.5 escape antibodies elicited by Omicron infection. *Nature*1–3, DOI: 10.1038/s41586-022-04980-y.

Committee for Medicinal Products for Human Use, 2021: GlaxoSmithKline use of sotrovimab (VIR-7831/GSK4182136) for the treatment of COVID-19 (No. EMA/304600/2021). p. 76. . European Medicines Agency.

Gerdol, M., K. Dishnica, and A. Giorgetti, 2022: Emergence of a recurrent insertion in the N-terminal domain of the SARS-CoV-2 spike glycoprotein. *Virus Research* **310**, 198674, DOI: 10.1016/j.virusres.2022.198674.

Guo, H., B.-J. Hu, X.-L. Yang, L.-P. Zeng, B. Li, S. Ouyang, and Z.-L. Shi, 2020: Evolutionary Arms Race between Virus and Host Drives Genetic Diversity in Bat Severe Acute Respiratory Syndrome-Related Coronavirus Spike Genes. *J Virol* **94**, e00902-20, DOI: 10.1128/JVI.00902-20.

Hachmann, N.P., J. Miller, A.Y. Collier, J.D. Ventura, J. Yu, M. Rowe, E.A. Bondzie, O. Powers, N. Surve, K. Hall, and D.H. Barouch, 2022: Neutralization Escape by SARS-CoV-2 Omicron Subvariants BA.2.12.1, BA.4, and BA.5. *N Engl J Med* **387**, 86–88, DOI: 10.1056/NEJMc2206576.

Harvey, W.T., A.M. Carabelli, B. Jackson, R.K. Gupta, E.C. Thomson, E.M. Harrison, C. Ludden, R. Reeve, A. Rambaut, S.J. Peacock, and D.L. Robertson, 2021: SARS-CoV-2 variants, spike mutations and immune escape. *Nat Rev Microbiol* **19**, 409–424, DOI: 10.1038/s41579-021-00573-0.

Hodcroft, E.B., M. Zuber, S. Nadeau, T.G. Vaughan, K.H.D. Crawford, C.L. Althaus, M.L. Reichmuth, J.E. Bowen, A.C. Walls, D. Corti, J.D. Bloom, D. Veesler, D. Mateo, A. Hernando, I. Comas, F.

González-Candelas, T. Stadler, and R.A. Neher, 2021: Spread of a SARS-CoV-2 variant through Europe in the summer of 2020. *Nature* **595**, 707–712, DOI: 10.1038/s41586-021-03677-y.

Hoffmann, M., H. Kleine-Weber, S. Schroeder, N. Krüger, T. Herrler, S. Erichsen, T.S. Schiergens, G. Herrler, N.-H. Wu, A. Nitsche, M.A. Müller, C. Drosten, and S. Pöhlmann, 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, DOI: 10.1016/j.cell.2020.02.052.

Javanmardi, K., T.H. Segall-Shapiro, C.-W. Chou, D.R. Boutz, R.J. Olsen, X. Xie, H. Xia, P.-Y. Shi, C.D. Johnson, A. Annapareddy, S. Weaver, J.M. Musser, A.D. Ellington, I.J. Finkelstein, and J.D. Gollihar, 2022: Antibody escape and cryptic cross-domain stabilization in the SARS-CoV-2 Omicron spike protein. *bioRxiv*2022.04.18.488614, DOI: 10.1101/2022.04.18.488614.

Ju, B., Q. Zhang, J. Ge, R. Wang, J. Sun, X. Ge, J. Yu, S. Shan, B. Zhou, S. Song, X. Tang, J. Yu, J. Lan, J. Yuan, H. Wang, J. Zhao, S. Zhang, Y. Wang, X. Shi, L. Liu, J. Zhao, X. Wang, Z. Zhang, and L. Zhang, 2020: Human neutralizing antibodies elicited by SARS-CoV-2 infection. *Nature* **584**, 115–119, DOI: 10.1038/s41586-020-2380-z.

Li, W., M.J. Moore, N. Vasilieva, J. Sui, S.K. Wong, M.A. Berne, M. Somasundaran, J.L. Sullivan, K. Luzuriaga, T.C. Greenough, H. Choe, and M. Farzan, 2003: Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. *Nature* **426**, 450–454, DOI: 10.1038/nature02145.

Mannar, D., J.W. Saville, X. Zhu, S.S. Srivastava, A.M. Berezuk, K.S. Tuttle, A.C. Marquez, I. Sekirov, and S. Subramaniam, 2022: SARS-CoV-2 Omicron variant: Antibody evasion and cryo-EM structure of spike protein-ACE2 complex. *Science* **375**, 760–764, DOI: 10.1126/science.abn7760.

Martin, D.P., S. Lytras, A.G. Lucaci, W. Maier, B. Grüning, S.D. Shank, S. Weaver, O.A. MacLean, R.J. Orton, P. Lemey, M.F. Boni, H. Tegally, G. Harkins, C. Scheepers, J.N. Bhiman, J. Everatt, D.G. Amoako, J.E. San, J. Giandhari, A. Sigal, NGS-SA, C. Williamson, N.-Y. Hsiao, A. von Gottberg, A. De Klerk, R.W. Shafer, D.L. Robertson, R.J. Wilkinson, B.T. Sewell, R. Lessells, A. Nekrutenko, A.J. Greaney, T.N. Starr, J.D. Bloom, B. Murrell, E. Wilkinson, R.K. Gupta, T. de Oliveira, and S.L. Kosakovsky Pond, 2022a: Selection analysis identifies unusual clustered mutational changes in Omicron lineage BA.1 that likely impact Spike function. *bioRxiv*2022.01.14.476382, DOI: 10.1101/2022.01.14.476382.

Martin, D.P., S. Lytras, A.G. Lucaci, W. Maier, B. Grüning, S.D. Shank, S. Weaver, O.A. MacLean, R.J. Orton, P. Lemey, M.F. Boni, H. Tegally, G.W. Harkins, C. Scheepers, J.N. Bhiman, J. Everatt, D.G. Amoako, J.E. San, J. Giandhari, A. Sigal, NGS-SA, C. Williamson, N. Hsiao, A. von Gottberg, A. De Klerk, R.W. Shafer, D.L. Robertson, R.J. Wilkinson, B.T. Sewell, R. Lessells, A. Nekrutenko, A.J. Greaney, T.N. Starr, J.D. Bloom, B. Murrell, E. Wilkinson, R.K. Gupta, T. de Oliveira, and S.L. Kosakovsky Pond, 2022b: Selection analysis identifies clusters of unusual mutational changes in Omicron lineage BA.1 that likely impact Spike function. *Mol Biol Evol*msac061, DOI: 10.1093/molbev/msac061.

McCarthy, K.R., L.J. Rennick, S. Nambulli, L.R. Robinson-McCarthy, W.G. Bain, G. Haidar, and W.P. Duprex, 2021: Recurrent deletions in the SARS-CoV-2 spike glycoprotein drive antibody escape. *Science* **371**, 1139–1142, DOI: 10.1126/science.abf6950.

Meng, B., A. Abdullahi, I.A.T.M. Ferreira, N. Goonawardane, A. Saito, I. Kimura, D. Yamasoba, P.P. Gerber, S. Fatihi, S. Rathore, S.K. Zepeda, G. Papa, S.A. Kemp, T. Ikeda, M. Toyoda, T.S. Tan, J. Kuramochi, S. Mitsunaga, T. Ueno, K. Shirakawa, A. Takaori-Kondo, T. Brevini, D.L. Mallery, O.J. Charles, CITIID-NIHR BioResource COVID-19 Collaboration, Genotype to Phenotype Japan (G2P-Japan) Consortium, Ecuador-COVID19 Consortium, J.E. Bowen, A. Joshi, A.C. Walls, L. Jackson, D. Martin, K.G.C. Smith, J. Bradley, J.A.G. Briggs, J. Choi, E. Madissoon, K.B. Meyer, P. Mlcochova, L. Ceron-Gutierrez, R. Doffinger, S.A. Teichmann, A.J. Fisher, M.S. Pizzuto, A. de Marco, D. Corti, M. Hosmillo, J.H. Lee, L.C. James, L. Thukral, D. Veesler, A. Sigal, F. Sampaziotis, I.G. Goodfellow, N.J. Matheson, K. Sato, and R.K. Gupta, 2022: Altered

TMPRSS2 usage by SARS-CoV-2 Omicron impacts infectivity and fusogenicity. *Nature* **603**, 706–714, DOI: 10.1038/s41586-022-04474-x.

Meng, B., S.A. Kemp, G. Papa, R. Datir, I.A.T.M. Ferreira, S. Marelli, W.T. Harvey, S. Lytras, A. Mohamed, G. Gallo, N. Thakur, D.A. Collier, P. Mlcochova, S.C. Robson, N.J. Loman, T.R. Connor, T. Golubchik, R.T. Martinez Nunez, C. Ludden, S. Corden, I. Johnston, D. Bonsall, C.P. Smith, A.R. Awan, G. Bucca, M.E. Torok, K. Saeed, J.A. Prieto, D.K. Jackson, W.L. Hamilton, L.B. Snell, C. Moore, E.M. Harrison, S. Goncalves, D.J. Fairley, M.W. Loose, J. Watkins, R. Livett, S. Moses, R. Amato, S. Nicholls, M. Bull, D.L. Smith, J. Barrett, D.M. Aanensen, M.D. Curran, S. Parmar, D. Aggarwal, J.G. Shepherd, M.D. Parker, S. Glaysher, M. Bashton, A.P. Underwood, N. Pacchiarini, K.F. Loveson, K.E. Templeton, C.F. Langford, J. Sillitoe, T.I. de Silva, D. Wang, D. Kwiatkowski, A. Rambaut, J. O'Grady, S. Cottrell, M.T.G. Holden, E.C. Thomson, H. Osman, M. Andersson, A.J. Chauhan, M.O. Hassan-Ibrahim, M. Lawniczak, A. Alderton, M. Chand, C. Constantinidou, M. Unnikrishnan, A.C. Darby, J.A. Hiscox, S. Paterson, I. Martincorena, E.M. Volz, A.J. Page, O.G. Pybus, A.R. Bassett, C.V. Ariani, M.H.S. Chapman, K.K. Li, R.N. Shah, N.G. Jesudason, Y. Taha, M.P. McHugh, R. Dewar, A.S. Jahun, C. McMurray, S. Pandey, J.P. McKenna, A. Nelson, G.R. Young, C.M. McCann, S. Elliott, H. Lowe, B. Temperton, S. Roy, A. Price, S. Rey, M. Wyles, S. Rooke, S. Shaaban, M. de Cesare, L. Letchford, S. Silveira, E. Pelosi, E. Wilson-Davies, M. Hosmillo, Á. O'Toole, A.R. Hesketh, R. Stark, L. du Plessis, C. Ruis, H. Adams, Y. Bourgeois, S.L. Michell, D. Gramatopoulos, J. Edgeworth, J. Breuer, J.A. Todd, C. Fraser, D. Buck, M. John, G.L. Kay, S. Palmer, S.J. Peacock, D. Heyburn, D. Weldon, E. Robinson, A. McNally, P. Muir, I.B. Vipond, J. Boyes, V. Sivaprakasam, T. Salluja, S. Dervisevic, E.J. Meader, N.R. Park, K. Oliver, A.R. Jeffries, S. Ott, A. da Silva Filipe, D.A. Simpson, C. Williams, J.A.H. Masoli, B.A. Knight, C.R. Jones, C. Koshy, A. Ash, A. Casey, A. Bosworth, L. Ratcliffe, L. Xu-McCrae, H.M. Pymont, S. Hutchings, L. Berry, K. Jones, F. Halstead, T. Davis, C. Holmes, M. Iturriza-Gomara, A.O. Lucaci, P.A. Randell, A. Cox, P. Madona, K.A. Harris, J.R. Brown, T.W. Mahungu, D. Irish-Tavares, T. Haque, J. Hart, E. Witele, M.L. Fenton, S. Liggett, C. Graham, E. Swindells, J. Collins, G. Eltringham, S. Campbell, P.C. McClure, G. Clark, T.J. Sloan, C. Jones, J. Lynch, B. Warne, S. Leonard, J. Durham, T. Williams, S.T. Haldenby, N. Storey, N.-F. Alikhan, N. Holmes, C. Moore, M. Carlile, M. Perry, N. Craine, R.A. Lyons, A.H. Beckett, S. Goudarzi, C. Fearn, K. Cook, H. Dent, H. Paul, R. Davies, B. Blane, S.T. Girgis, M.A. Beale, K.L. Bellis, M.J. Dorman, E. Drury, L. Kane, S. Kay, S. McGuigan, R. Nelson, L. Prestwood, S. Rajatileka, R. Batra, R.J. Williams, M. Kristiansen, A. Green, A. Justice, A.I.K. Mahanama, B. Samaraweera, N.F. Hadjirin, J. Quick, R. Poplawski, L.M. Kermack, N. Reynolds, G. Hall, Y. Chaudhry, M.L. Pinckert, I. Georgana, R.J. Moll, A. Thornton, R. Myers, J. Stockton, C.A. Williams, W.C. Yew, A.J. Trotter, A. Trebes, G. MacIntyre-Cockett, A. Birchley, A. Adams, A. Plimmer, B. Gatica-Wilcox, C. McKerr, E. Hilvers, H. Jones, H. Asad, J. Coombes, J.M. Evans, L. Fina, L. Gilbert, L. Graham, M. Cronin, S. Kumziene-Summerhayes, S. Taylor, S. Jones, D.C. Groves, P. Zhang, M. Gallis, S.F. Louka, I. Starinskij, C. Jackson, M. Gourtovaia, G. Tonkin-Hill, K. Lewis, J.M. Tovar-Corona, K. James, L. Baxter, M.T. Alam, R.J. Orton, J. Hughes, S. Vattipally, M. Ragonnet-Cronin, F.F. Nascimento, D. Jorgensen, O. Boyd, L. Geidelberg, A.E. Zarebski, J. Raghwani, M.U.G. Kraemer, J. Southgate, B.B. Lindsey, T.M. Freeman, J.-P. Keatley, J.B. Singer, L. de Oliveira Martins, C.A. Yeats, K. Abudahab, B.E.W. Taylor, M. Menegazzo, J. Danesh, W. Hogsden, S. Eldirdiri, A. Kenyon, J. Mason, T.I. Robinson, A. Holmes, J. Price, J.A. Hartley, T. Curran, A.E. Mather, G. Shankar, R. Jones, R. Howe, S. Morgan, E. Wastenge, M.R. Chapman, S. Mookerjee, R. Stanley, W. Smith, T. Peto, D. Eyre, D. Crook, G. Vernet, C. Kitchen, H. Gulliver, I. Merrick, M. Guest, R. Munn, D.T. Bradley, T. Wyatt, C. Beaver, L. Foulser, S. Palmer, C.M. Churcher, E. Brooks, K.S. Smith, K. Galai, G.M. McManus, F. Bolt, F. Coll, L. Meadows, S.W. Attwood, A. Davies, E. De Lacy, F. Downing, S. Edwards, G.P. Scarlett, S. Jeremiah, N. Smith, D. Leek, S. Sridhar, S. Forrest, C. Cormie, H.K. Gill, J. Dias, E.E. Higginson, M. Maes, J. Young, M. Wantoch, D. Jamrozy, S. Lo, M. Patel, V. Hill, C.M. Bewshea, S. Ellard, C. Auckland, I.

Harrison, C. Bishop, V. Chalker, A. Richter, A. Beggs, A. Best, B. Percival, J. Mirza, O. Megram, M. Mayhew, L. Crawford, F. Ashcroft, E. Moles-Garcia, N. Cumley, R. Hopes, P. Asamaphan, M.O. Niebel, R.N. Gunson, A. Bradley, A. Maclean, G. Mollett, R. Blacow, P. Bird, T. Helmer, K. Fallon, J. Tang, A.D. Hale, L.R. Macfarlane-Smith, K.L. Harper, H. Carden, N.W. Machin, K.A. Jackson, S.S.Y. Ahmad, R.P. George, L. Turtle, E. O'Toole, J. Watts, C. Breen, A. Cowell, A. Alcolea-Medina, T. Charalampous, A. Patel, L.J. Levett, J. Heaney, A. Rowan, G.P. Taylor, D. Shah, L. Atkinson, J.C.D. Lee, A.P. Westhorpe, R. Jannoo, H.L. Lowe, A. Karamani, L. Ensell, W. Chatterton, M. Pusok, A. Dadrah, A. Symmonds, G. Sluga, Z. Molnar, P. Baker, S. Bonner, S. Essex, E. Barton, D. Padgett, G. Scott, J. Greenaway, B.A.I. Payne, S. Burton-Fanning, S. Waugh, V. Raviprakash, N. Sheriff, V. Blakey, L.-A. Williams, J. Moore, S. Stonehouse, L. Smith, R.K. Davidson, L. Bedford, L. Coupland, V. Wright, J.G. Chappell, T. Tsoleridis, J. Ball, M. Khakh, V.M. Fleming, M.M. Lister, H.C. Howson-Wells, L. Berry, T. Boswell, A. Joseph, I. Willingham, N. Duckworth, S. Walsh, E. Wise, N. Moore, M. Mori, N. Cortes, S. Kidd, R. Williams, L. Gifford, K. Bicknell, S. Wyllie, A. Lloyd, R. Impey, C.S. Malone, B.J. Cogger, N. Levene, L. Monaghan, A.J. Keeley, D.G. Partridge, M. Raza, C. Evans, K. Johnson, E. Betteridge, B.W. Farr, S. Goodwin, M.A. Quail, C. Scott, L. Shirley, S.A.J. Thurston, D. Rajan, I.F. Bronner, L. Aigrain, N.M. Redshaw, S.V. Lensing, S. McCarthy, A. Makunin, C.E. Balcazar, M.D. Gallagher, K.A. Williamson, T.D. Stanton, M.L. Michelsen, J. Warwick-Dugdale, R. Manley, A. Farbos, J.W. Harrison, C.M. Sambles, D.J. Studholme, A. Lackenby, T. Mbisa, S. Platt, S. Miah, D. Bibby, C. Manso, J. Hubb, G. Dabrera, M. Ramsay, D. Bradshaw, U. Schaefer, N. Groves, E. Gallagher, D. Lee, D. Williams, N. Ellaby, H. Hartman, N. Manesis, V. Patel, J. Ledesma, K.A. Twohig, E. Allara, C. Pearson, J.K.J. Cheng, H.E. Bridgewater, L.R. Frost, G. Taylor-Joyce, P.E. Brown, L. Tong, A. Broos, D. Mair, J. Nichols, S.N. Carmichael, K.L. Smollett, K. Nomikou, E. Aranday-Cortes, N. Johnson, S. Nickbakhsh, E.E. Vamos, M. Hughes, L. Rainbow, R. Eccles, C. Nelson, M. Whitehead, R. Gregory, M. Gemmell, C. Wierzbicki, H.J. Webster, C.L. Fisher, A.W. Signell, G. Betancor, H.D. Wilson, G. Nebbia, F. Flaviani, A.C. Cerda, T.V. Merrill, R.E. Wilson, M. Cotic, N. Bayzid, T. Thompson, E. Acheson, S. Rushton, S. O'Brien, D.J. Baker, S. Rudder, A. Aydin, F. Sang, J. Debebe, S. Francois, T.I. Vasylyeva, M.E. Zamudio, B. Gutierrez, A. Marchbank, J. Maksimovic, K. Spellman, K. McCluggage, M. Morgan, R. Beer, S. Afifi, T. Workman, W. Fuller, C. Bresner, A. Angyal, L.R. Green, P.J. Parsons, R.M. Tucker, R. Brown, M. Whiteley, J. Bonfield, C. Puethe, A. Whitwham, J. Liddle, W. Rowe, I. Siveroni, T. Le-Viet, A. Gaskin, R. Johnson, I. Abnizova, M. Ali, L. Allen, R. Anderson, C. Ariani, S. Austin-Guest, S. Bala, J. Barrett, A. Bassett, K. Battleday, J. Beal, M. Beale, S. Bellany, T. Bellerby, K. Bellis, D. Berger, M. Berriman, P. Bevan, S. Binley, J. Bishop, K. Blackburn, N. Boughton, S. Bowker, T. Brendler-Spaeth, I. Bronner, T. Brooklyn, S.K. Buddenborg, R. Bush, C. Caetano, A. Cagan, N. Carter, J. Cartwright, T.C. Monteiro, L. Chapman, T.-J. Chillingworth, P. Clapham, R. Clark, A. Clarke, C. Clarke, D. Cole, E. Cook, M. Coppola, L. Cornell, C. Cornwell, C. Corton, A. Crackett, A. Cranage, H. Craven, S. Craw, M. Crawford, T. Cutts, M. Dabrowska, M. Davies, J. Dawson, C. Day, A. Densem, T. Dibling, C. Dockree, D. Dodd, S. Dogga, M. Dorman, G. Dougan, M. Dougherty, A. Dove, L. Drummond, M. Dudek, L. Durrant, E. Easthope, S. Eckert, P. Ellis, B. Farr, M. Fenton, M. Ferrero, N. Flack, H. Fordham, G. Forsythe, M. Francis, A. Fraser, A. Freeman, A. Galvin, M. Garcia-Casado, A. Gedny, S. Girgis, J. Glover, O. Gould, A. Gray, E. Gray, C. Griffiths, Y. Gu, F. Guerin, W. Hamilton, H. Hanks, E. Harrison, A. Harrott, E. Harry, J. Harvison, P. Heath, A. Hernandez-Koutoucheva, R. Hobbs, D. Holland, S. Holmes, G. Hornett, N. Hough, L. Huckle, L. Hughes-Hallet, A. Hunter, S. Inglis, S. Iqbal, A. Jackson, D. Jackson, C.J. Verdejo, M. Jones, K. Kallepally, K. Kay, J. Keatley, A. Keith, A. King, L. Kitchin, M. Kleanthous, M. Klimekova, P. Korlevic, K. Krasheninnkova, G. Lane, C. Langford, A. Laverack, K. Law, S. Lensing, A. Lewis-Wade, J. Liddle, Q. Lin, S. Lindsay, S. Linsdell, R. Long, J. Lovell, J. Lovell, J. Mack, M. Maddison, A. Makunin, I. Mamun, J. Mansfield, N. Marriott, M. Martin, M. Mayho, J. McClintock, S. McHugh, L. MapcMinn, C. Meadows, E. Mobley, R. Moll, M. Morra, L.

Morrow, K. Murie, S. Nash, C. Nathwani, P. Naydenova, A. Neaverson, E. Nerou, J. Nicholson, T. Nimz, G.G. Noell, S. O'Meara, V. Ohan, C. Olney, D. Ormond, A. Oszlanczi, Y.F. Pang, B. Pardubska, N. Park, A. Parmar, G. Patel, M. Payne, S. Peacock, A. Petersen, D. Plowman, T. Preston, M. Quail, R. Rance, S. Rawlings, N. Redshaw, J. Reynolds, M. Reynolds, S. Rice, M. Richardson, C. Roberts, K. Robinson, M. Robinson, D. Robinson, H. Rogers, E.M. Rojo, D. Roopra, M. Rose, L. Rudd, R. Sadri, N. Salmon, D. Saul, F. Schwach, P. Seekings, A. Simms, M. Sinnott, S. Sivadasan, B. Siwek, D. Sizer, K. Skeldon, J. Skelton, J. Slater-Tunstill, L. Sloper, N. Smerdon, C. Smith, C. Smith, J. Smith, K. Smith, M. Smith, S. Smith, T. Smith, L. Sneade, C.D. Soria, C. Sousa, E. Souster, A. Sparkes, M. Spencer-Chapman, J. Squares, R. Stanley, C. Steed, T. Stickland, I. Still, M. Stratton, M. Strickland, A. Swann, A. Swiatkowska, N. Sycamore, E. Swift, E. Symons, S. Szluha, E. Taluy, N. Tao, K. Taylor, S. Taylor, S. Thompson, M. Thompson, M. Thomson, N. Thomson, S. Thurston, D. Toombs, B. Topping, J. Tovar-Corona, D. Ungureanu, J. Uphill, J. Urbanova, P.J. Van, V. Vancollie, P. Voak, D. Walker, M. Walker, M. Waller, G. Ward, C. Weatherhogg, N. Webb, A. Wells, E. Wells, L. Westwood, T. Whipp, T. Whiteley, G. Whitton, S. Widaa, M. Williams, M. Wilson, S. Wright, L.M. Duncan, A.M. Carabelli, J.C. Kenyon, A.M. Lever, A. De Marco, C. Saliba, K. Culap, E. Cameroni, N.J. Matheson, L. Piccoli, D. Corti, L.C. James, D.L. Robertson, D. Bailey, and R.K. Gupta, 2021: Recurrent emergence of SARS-CoV-2 spike deletion H69/V70 and its role in the Alpha variant B.1.1.7. *Cell Reports* **35**, 109292, DOI: 10.1016/j.celrep.2021.109292.

Pia, L., and S. Rowland-Jones, 2022: Omicron entry route. *Nat Rev Immunol* **22**, 144–144, DOI: 10.1038/s41577-022-00681-9.

Planas, D., D. Veyer, A. Baidaliuk, I. Staropoli, F. Guivel-Benhassine, M.M. Rajah, C. Planchais, F. Porrot, N. Robillard, J. Puech, M. Prot, F. Gallais, P. Gantner, A. Velay, J. Le Guen, N. Kassis-Chikhani, D. Edriss, L. Belec, A. Seve, L. Courtellemont, H. Péré, L. Hocqueloux, S. Fafi-Kremer, T. Prazuck, H. Mouquet, T. Bruel, E. Simon-Lorière, F.A. Rey, and O. Schwartz, 2021: Reduced sensitivity of SARS-CoV-2 variant Delta to antibody neutralization. *Nature* **596**, 276–280, DOI: 10.1038/s41586-021-03777-9.

Pulliam, J.R.C., C. van Schalkwyk, N. Govender, A. von Gottberg, C. Cohen, M.J. Groome, J. Dushoff, K. Mlisana, and H. Moultrie, 2022: Increased risk of SARS-CoV-2 reinfection associated with emergence of Omicron in South Africa. *Science* **0**, eabn4947, DOI: 10.1126/science.abn4947.

Sabino, E.C., L.F. Buss, M.P.S. Carvalho, C.A. Prete, M.A.E. Crispim, N.A. Fraiji, R.H.M. Pereira, K.V. Parag, P. da S. Peixoto, M.U.G. Kraemer, M.K. Oikawa, T. Salomon, Z.M. Cucunuba, M.C. Castro, A.A. de S. Santos, V.H. Nascimento, H.S. Pereira, N.M. Ferguson, O.G. Pybus, A. Kucharski, M.P. Busch, C. Dye, and N.R. Faria, 2021: Resurgence of COVID-19 in Manaus, Brazil, despite high seroprevalence. *The Lancet* **397**, 452–455, DOI: 10.1016/S0140-6736(21)00183-5.

Saito, A., T. Irie, R. Suzuki, T. Maemura, H. Nasser, K. Uriu, Y. Kosugi, K. Shirakawa, K. Sadamasu, I. Kimura, J. Ito, J. Wu, K. Iwatsuki-Horimoto, M. Ito, S. Yamayoshi, S. Loeber, M. Tsuda, L. Wang, S. Ozono, E.P. Butlertanaka, Y.L. Tanaka, R. Shimizu, K. Shimizu, K. Yoshimatsu, R. Kawabata, T. Sakaguchi, K. Tokunaga, I. Yoshida, H. Asakura, M. Nagashima, Y. Kazuma, R. Nomura, Y. Horisawa, K. Yoshimura, A. Takaori-Kondo, M. Imai, S. Tanaka, S. Nakagawa, T. Ikeda, T. Fukuhara, Y. Kawaoka, and K. Sato, 2021: Enhanced fusogenicity and pathogenicity of SARS-CoV-2 Delta P681R mutation. *Nature* 1–10, DOI: 10.1038/s41586-021-04266-9.

Shiliaev, N., T. Lukash, O. Palchevska, D.K. Crossman, T.J. Green, M.R. Crowley, E.I. Frolova, and I. Frolov, 2021: Natural and Recombinant SARS-CoV-2 Isolates Rapidly Evolve In Vitro to Higher Infectivity through More Efficient Binding to Heparan Sulfate and Reduced S1/S2 Cleavage. *Journal of Virology* **95**, e01357-21, DOI: 10.1128/JVI.01357-21.

Shu, Y., and J. McCauley, 2017: GISAID: Global initiative on sharing all influenza data – from vision to reality. *Euro Surveill* **22**, DOI: 10.2807/1560-7917.ES.2017.22.13.30494.

Simon-Loriere, E., and O. Schwartz, 2022: Towards SARS-CoV-2 serotypes? *Nat Rev Microbiol* **20**, 187–188, DOI: 10.1038/s41579-022-00708-x.

Tegally, H., M. Moir, J. Everatt, M. Giovanetti, C. Scheepers, E. Wilkinson, K. Subramoney, Z. Makatini, S. Moyo, D.G. Amoako, C. Baxter, C.L. Althaus, U.J. Anyaneji, D. Kekana, R. Viana, J. Giandhari, R.J. Lessells, T. Maponga, D. Maruapula, W. Choga, M. Matshaba, M.B. Mbulawa, N. Msomi, Y. Naidoo, S. Pillay, T.J. Sanko, J.E. San, L. Scott, L. Singh, N.A. Magini, P. Smith-Lawrence, W. Stevens, G. Dor, D. Tshiabuila, N. Wolter, W. Preiser, F.K. Treurnicht, M. Venter, G. Chiloane, C. McIntyre, A. O'Toole, C. Ruis, T.P. Peacock, C. Roemer, S.L.K. Pond, C. Williamson, O.G. Pybus, J.N. Bhiman, A. Glass, D.P. Martin, B. Jackson, A. Rambaut, O. Laguda-Akingba, S. Gaseitsiwe, A. von Gottberg, and T. de Oliveira, 2022: Emergence of SARS-CoV-2 Omicron lineages BA.4 and BA.5 in South Africa. *Nat Med*1–1, DOI: 10.1038/s41591-022-01911-2.

Teruel, N., O. Mailhot, and R.J. Najmanovich, 2021: Modelling conformational state dynamics and its role on infection for SARS-CoV-2 Spike protein variants. *PLOS Computational Biology* **17**, e1009286, DOI: 10.1371/journal.pcbi.1009286.

Walls, A.C., Y.-J. Park, M.A. Tortorici, A. Wall, A.T. McGuire, and D. Veesler, 2020: Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* **181**, 281-292.e6, DOI: 10.1016/j.cell.2020.02.058.

Wang, Q., Y. Guo, S. Iketani, M.S. Nair, Z. Li, H. Mohri, M. Wang, J. Yu, A.D. Bowen, J.Y. Chang, J.G. Shah, N. Nguyen, Z. Chen, K. Meyers, M.T. Yin, M.E. Sobieszczyk, Z. Sheng, Y. Huang, L. Liu, and D.D. Ho, 2022: Antibody evasion by SARS-CoV-2 Omicron subvariants BA.2.12.1, BA.4, & BA.5. *Nature*1–3, DOI: 10.1038/s41586-022-05053-w.

Ye, G., B. Liu, and F. Li, 2022: Cryo-EM structure of a SARS-CoV-2 omicron spike protein ectodomain. *Nat Commun* **13**, 1214, DOI: 10.1038/s41467-022-28882-9.

Zhu, X., D. Mannar, S.S. Srivastava, A.M. Berezuk, J.-P. Demers, J.W. Saville, K. Leopold, W. Li, D.S. Dimitrov, K.S. Tuttle, S. Zhou, S. Chittori, and S. Subramaniam, 2021: Cryo-electron microscopy structures of the N501Y SARS-CoV-2 spike protein in complex with ACE2 and 2 potent neutralizing antibodies. *PLOS Biology* **19**, e3001237, DOI: 10.1371/journal.pbio.3001237.