Emerging Microbes & Infections
2026, VOL. 15, 2755643 (5 pages)
https://doi.org/10.1080/22221751.2026.2755643

OPEN ACCESS Check for updates

Worldwide emergence of SARS-CoV-2 NB.1.8.1 sublineage with spike 420N and 478T dual substitution

Hoi-Wah Tsoi[a], Wan-Mui Chan[a], Allen Wing-Ho Chu[a,b], Cassie Yu-Sen Yip[a], Xiaojuan Zhang[a,b], Jonathan Daniel Ip[a,c], Yanling Luo[a], Hin Chu[a,b,d], Kwok-Yung Yuen[a,b,c,d,e] and Kelvin Kai-Wang To [a,b,c,d,e]

[a]State Key Laboratory of Emerging Infectious Diseases, Carol Yu Centre for Infection, Department of Microbiology, Li Ka Shing Faculty of Medicine, School of Clinical Medicine, the University of Hong Kong, Pokfulam, Hong Kong Special Administrative Region, People's Republic of China; [b]Centre for Virology, Vaccinology and Therapeutics, Hong Kong Science and Technology Park, Shatin, Hong Kong Special Administrative Region, People's Republic of China; [c]Department of Microbiology, Queen Mary Hospital, Pokfulam, Hong Kong Special Administrative Region, People's Republic of China; [d]Pandemic Research Alliance Unit at the University of Hong Kong, Pokfulam, Hong Kong Special Administrative Region, People's Republic of China; [e]Department of Infectious Disease and Microbiology, The University of Hong Kong-Shenzhen Hospital, Shenzhen, People's Republic of China

ABSTRACT

Background: After a year of quiescence, SARS-CoV-2 activity increased rapidly in Hong Kong since June 2026. We sought to characterize the genomic profile of SARS-CoV-2 driving this COVID-19 surge in Hong Kong and to examine the global dissemination of the identified circulating lineages.

Methods: We performed SARS-CoV-2 whole-genome sequencing on archived respiratory specimens collected at Queen Mary Hospital, Hong Kong, between January 2025 and July 2026. Viral lineages were assigned via Nextclade. Global SARS-CoV-2 metadata downloaded from GISAID as of July 2026 were analysed to track global evolutionary trends.

Results: Among 615 successfully sequenced strains, the XDV descendant NB.1.8.1 was the predominant lineage. Between April and July 2026, NB.1.8.1 descendants accounted for 95.6% (108/113) of sequenced strains, dominated by PQ.16.1.1 (81.5%). Spike substitutions 420N and 478T emerged rapidly. While 478T increased gradually throughout 2025, 420N emerged in April 2026 and reached 98.0% by June–July 2026. Strains harbouring 420N-478T dual substitution arose independently across distinct NB.1.8.1 lineages, accounting for 94.4% (102/108) of NB.1.8.1 descendants in Hong Kong during this period. Globally, the prevalence of NB.1.8.1 strains with 420N-478T dual substitution increased from <1% before April 2026 to 50.7% by mid-July 2026, with the highest prevalence observed in Southeast Asia and Australia.

Conclusion: The 2026 COVID-19 surge in Hong Kong was driven by NB.1.8.1 descendants harbouring spike 420N and 478T substitutions. The rapid worldwide expansion of these dual-substitution lineages highlights their competitive growth advantage and underscores the need for continued global genomic surveillance.

ARTICLE HISTORY Received 5 August 2026; Revised 18 September 2026; Accepted 22 October 2026

KEYWORDS SARS-CoV-2; D420N; K478T or I478T; amino acid substitutions; evolution

Introduction

Since emerging in late 2019, SARS-CoV-2 has evolved into different lineages, resulting in successive waves of infection. In Hong Kong, the first large SARS-CoV-2 outbreak occurred when the Omicron variant spread rapidly in early 2022, primarily driven by BA.2 [1,2]. Subsequent SARS-CoV-2 waves from 2022 to 2024 were characterized by BA.5, BA.2.3.20, XBB, EG.5.1 and JN.1 [3].

There is a noticeable shift in SARS-CoV-2 epidemiology in 2025 and 2026. Unlike the multiple discrete waves observed in preceding years, there was a single prolonged surge from spring to autumn in 2025, peaking in May, followed by a long period of quiescence from October 2025 to early spring 2026 [4]. However, SARS-CoV-2 activity rose sharply in June 2026. By mid-July 2026, SARS-CoV-2 was found in 10% of respiratory specimens tested at the Public Health Laboratory Services Branch in Hong Kong [4].

In this study, we conducted a longitudinal genomic surveillance using SARS-CoV-2 positive respiratory specimens collected during the 2025 and 2026 epidemic waves in Hong Kong. We demonstrate that JN.1-XDE recombinant lineage NB.1.8.1 predominated in both 2025 and 2026. Notably, the NB.1.8.1 in 2026 carried spike 420N and 478T dual substitution. Located in the receptor binding domain (RBD), these substitutions have been shown to confer

CONTACT Kelvin Kai-Wang To kelvinto@hku.hk Department of Microbiology, Li Ka Shing Faculty of Medicine, School of Clinical Medicine, Carol Yu Centre for Infection, Queen Mary Hospital, The University of Hong Kong, Pokfulam, Hong Kong Special Administrative Region, People's Republic of China
Supplemental data for this article can be accessed online at https://doi.org/10.1080/22221751.2026.2755643.

© 2026 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group, on behalf of Shanghai Shangyixun Cultural Communication Co., Ltd
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

immune escape or affect binding to the host receptor angiotensin-converting enzyme 2 (ACE2) [5–8]. Importantly, the dual substitution arose independently in different sublineages.

Materials and methods

Clinical specimens and nucleic acid extraction

Archived respiratory specimens collected between January 2025 and July 2026 and tested positive for SARS-CoV-2 by reverse transcriptase-polymerase chain reaction (RT-PCR) at Queen Mary Hospital in Hong Kong were retrieved. Total nucleic acid was extracted using the GeneRotex 96 Nucleic Acid Extractor using the Viral DNA and RNA Extraction Kit (T014H) (Tianlong).

Viral genome sequencing and bioinformatics analysis

Whole viral genome sequencing directly from respiratory specimens was performed using the Oxford Nanopore MinION device (Oxford Nanopore Technologies, Oxford, United Kingdom) with the Midnight protocol as we described previously [9]. For bioinformatics analysis, consensus sequences were generated using wf-artic workflow (v1.0.0) via the EPI2ME Desktop application. A normalized value of 999999 and super accuracy mode for basecalling were applied as described previously [9,10]. The sequence MN908947.3 obtained from NCBI was used as the reference. The consensus sequences were verified manually using Integrative Genomics Viewer (IGV) v2.17.2 [11]. SARS-CoV-2 lineage was assigned using the online Nextclade tool (v3.21.2; https://clades.nextstrain.org; accessed on 31st July 2026). All sequences were deposited onto the Global Initiative on Sharing All Influenza Data (GISAID) database (Supplementary Table S1). A whole viral genome was considered successfully sequenced if at least 85% of the genome achieved a sequencing depth of ≥20x.

Global SARS-CoV-2 sequence analysis

The analysis of global SARS-CoV-2 sequences was performed as we described previously [12]. We downloaded the SARS-CoV-2 sequences and metadata from GISAID on 29 July 2026. A SARS-CoV-2 sequence was included if it had been marked as "complete" and "betacoronavirus" in the metadata file and was directly obtained from a human clinical specimen without *in vitro* passage. A sequence was excluded if the collection date information was not complete, the sequence was originated from Hong Kong, both spike amino acid residue 420 and 478 could not be sequenced, or failed NextClade Cli analysis. The summary of the amino acid occurrence and monthly prevalence were calculated by an in-house python script using Python v3.9.11 with modules Pandas v3.0.0, and tqdm v4.67.1. The doubling time of strains with 478T only or 420N-478T dual substitution was determined as described previously [2].

In silico *analysis of ACE2 binding*

The *in silico* analysis of ACE2 binding was performed using NextClade, which is based on the estimator from Starr et al. [13].

Results

Genomic surveillance in Hong Kong from January 2025 to July 2026

Between January 2025 and July 2026, we have sequenced a total of 772 strains collected in Queen Mary Hospital, in which 615 strains had at least 85% of the genome achieving a sequencing depth of ≥20x (Supplementary Table S2). NB.1.8.1, a descendant of XDV, was the predominant lineage during this period (Supplementary Figure S1). The median age of individuals with NB.1.8.1 is similar to those with non-NB.1.8.1 lineage (72 vs 68 years).

NB.1.8.1 descendants accounted for 95.6% (108/113) of strains successfully sequenced between 19th April and 6th July 2026, followed by XFG descendants (2.7%; 3/113) and BA.3.2 descendants (1.8%; 2/113) (Supplementary Table S3). Among the NB.1.8.1 descendants, PQ.16.1.1 (designated by the World Health Organization as a variant under monitoring on July 27, 2026 [14]) or its descendants, SV.1 and SV.2, was most predominant (81.5%; 88/108), followed by PQ.2 descendants (16.7%; 18/108), PQ.10 descendant (RH.1) (0.9%; 1/108), and PQ.17 descendant (RK.1) (0.9%; 1/108) (Supplementary Figure S2 and Supplementary Table S3).

The lineage defining amino acid substitutions include the spike protein D253G, N417T, and D420N for PQ.16.1, and K478T reversion for PQ.16.1.1 [15]. Notably, 420N is also lineage defining substitutions for RK.1, while 478T is also a lineage defining substitution for PQ.2.1 and PQ.17. Hence, we further analysed the prevalence of 420N and 478T among our strains. 478T exceeded 10% of all strains since July 2025 and peaked at 56.7% (17/30) in August 2025 (Figure 1(A)). In contrast, 420N was not found in 2025 and early 2026, but was first detected in April 2026. The proportion of strains with D420N substitution increased rapidly from 33.3% (2/6) in April 2026 to 98.0% (97/99) in June to July 2026, and almost all strains with 420N also carry 478T. 420N and/or 478T were only present among the

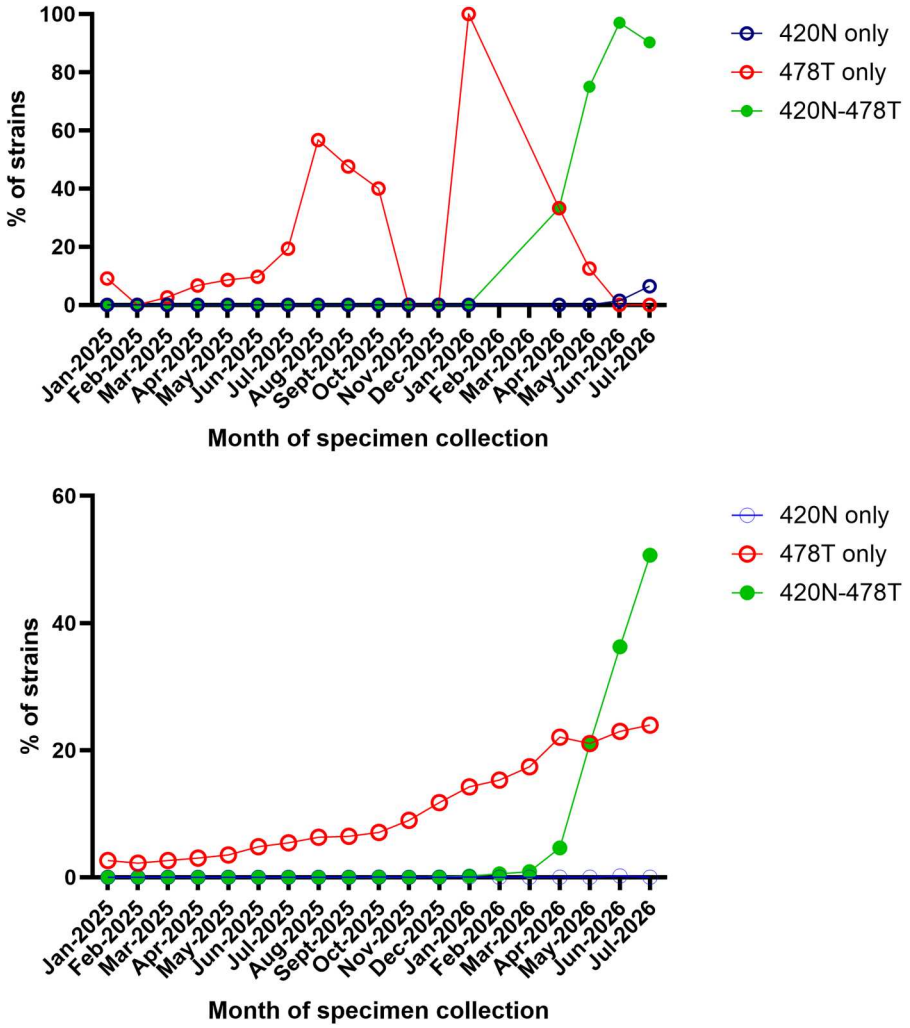


Figure 1. Prevalence of SARS-CoV-2 strains with spike 420N and/or 478T substitutions. (A) Our cohort. (B) Worldwide, excluding Hong Kong.

NB.1.8.1 descendants, but not the XFG or BA.3.2 descendants (Supplementary Table S3). Among strains collected from April to July 2026, 420N, 478T and 420N-478T dual substitution were present in 97.2% (105/108), 97.2% (105/108) and 94.4% (102/108) of NB.1.8.1 descendants, respectively (Supplementary Table S3).

To determine if strains with 420N-478T dual substitution were also increasing worldwide, we analysed the data from GISAID, excluding data from Hong Kong. Our analysis showed a substantial increase in 420N-478T dual substitution worldwide in 2026. The prevalence of 420N-478T has increased from <1% in 2025 to 36.3% in June 2026 and 50.7% in July 2026 (Figure 1(B)). The doubling time of strains with 420N-478T dual substitution (1.21 months [95% confidence interval: 1.03–1.42]) was shorter than that of strains with 478T substitution only (5.46 months [95% confidence interval: 4.87–6.16]). Among the SARS-CoV-2 strains with 420N-478T dual substitution and collected between January 2025 and July 2026, 97.5% (1272/1304) were NB.1.8.1 descendants (Supplementary Table S4). The proportion of 420N-478T among NB.1.8.1 descendants varies between different regions. Among regions with at least 50 strains deposited in GISAID in 2026, 420N-478T dual substitution was most prevalent in Singapore (43.5%; 741/1704), followed by Thailand (34.6%; 27/78) and Taiwan (32.3%; 20/62) (Supplementary Table S5). Outside Asia, 420N-478T dual substitution was most prevalent in Australia (10.2%, 294/2879).

To determine whether 420N-478T dual substitution arose through convergent evolution or from a single common ancestor, we constructed a phylogenetic tree of all NB.1.8.1 strains from Queen Mary Hospital collected between January 2025 and July 2026. The phylogenetic tree showed that NB.1.8.1 strains harbouring 420N-478T dual substitution emerged independently across distinct lineages instead of descending from a single common ancestor (Supplementary Figure S3).

To investigate whether the rapid increase in the prevalence of 420N-478T is related to ACE2 binding, we evaluated the ACE2 binding using NextClade, which is based on the work by Starr et al. [13]. Compared to the BA.2.86 baseline, the ACE2 binding

affinities of the PQ.16-420D-478I, PQ.16-420D-I478T, and PQ.16-D420N-I478T are −2.25882, −2.2775 and −2.31451 Δ-$\log_{10}(K_D)$, respectively, suggesting that 420N-478T dual substitution confers a slightly lower binding affinity to ACE2.

Discussion

In this genomic surveillance study, we identified the rapid emergence of NB.1.8.1 descendants carrying spike 420N-478T dual substitution during the 2026 summer COVID-19 season in Hong Kong. Notably, 420N-478T dual substitution has arisen independently across multiple genetically divergent NB.1.8.1 sublineages, including PQ.16.1.1, PQ.2.1.11, and RK.1, suggesting convergent evolution. NB.1.8.1 descendants with 420N-478T substitutions also increased substantially worldwide, especially in Southeast Asia and Australia.

The acquisition of the D420N substitution within the RBD appears to be a key molecular determinant of this ongoing SARS-CoV-2 surge. Witte *et al* found that the D420N substitution emerged during the passage of SARS-CoV-2 BA.2 variant with monoclonal neutralizing antibody [5]. Synthetic virus with D420N substitution showed a substantial reduction in the neutralizing activity of class 1/2 or class 1/4 antibodies. Furthermore, deep mutational scanning by Dadonaite *et al* highlighted D420N as a primary escape mutation against polyclonal serum antibodies derived from individuals vaccinated and infected during the XBB era [6]. Recently, He *et al* found that both PQ.16.1.1 and RK.1, both with D420N as clade defining substitutions, evade some class 1 monoclonal antibodies [16]. Notably, the impact of D420N is strain dependent. D420N affects the neutralizing activity of C3095 (a class 1/2 mAb) for BA.1 or BA.2 but not the ancestral strain [5].

Almost all strains with spike 420N also carry 478T. While T478K substitution was the hallmark of the Delta variant, the initial NB.1.8.1 strains possessed 478I. Previous studies showed that different substitutions at residue 478 can affect ACE2 binding. While T478K enhances the binding, T478I does not significantly affect ACE2 binding [7,8]. Strains with 478T reversion rapidly increased since July 2025 to about 50%. However, strains with 478T without 420N did not further expand in 2025. Instead, the expansion of NB.1.8.1 only occurred in 2026 when 478T was accompanied by 420N, suggesting that the advantage in transmission or human infection of 420N-478T dual substitution over those with single 478T substitution. Our *in silico* results showed that 420N-478T dual substitution had a slightly lower binding affinity to ACE2 receptor. This result concurs with a recent study which showed that PQ.16.1.1 has lower binding affinity to the ACE2 receptor than NB.1.8.1 using a pseudovirus binding assay [16]. Further studies are required to confirm whether the 420N-478T dual substitution is responsible for the change in ACE2 binding affinity and whether the change in binding affinity affect viral fitness and transmission in humans.

There are several limitations in this study. First, we only included clinical specimens from hospital patients, and therefore the lineages from milder cases were excluded. Second, we only included clinical specimens from a single hospital, which may not be generalizable to the entire city. Third, there is substantial variation in the number of sequences uploaded to GISAID between countries or regions. Estimates of variant prevalence would be less accurate for places with fewer sequences. Furthermore, the estimated global prevalence may be disproportionately influenced by places with a higher number of sequences.

In summary, the 2026 COVID-19 surge in Hong Kong was predominated by NB.1.8.1 descendants with 420N-478T dual substitution, which arose independently in different lineages. Although strains with 420N-478T substitutions are spreading rapidly, there is currently no epidemiological evidence to suggest that these strains cause more severe disease. Further studies are required to determine how the dual substitution has competitive advantages over strains with single substitution, especially for immune escape and receptor binding.

Acknowledgements

We gratefully acknowledge the authors originating and submitting laboratories of the SARS-CoV-2 genetic sequences and metadata made available through GISAID on which this research is based (Supplementary Table S6).

Author contributions

CRediT: **Hoi-Wah Tsoi:** Conceptualization, Data curation, Formal analysis, Writing – review & editing; **Wan-Mui Chan:** Data curation, Methodology, Writing – review & editing; **Allen Wing-Ho Chu:** Data curation, Methodology, Writing – review & editing; **Cassie Yu-Sen Yip:** Formal analysis, Writing – review & editing; **Xiaojuan Zhang:** Data curation, Writing – review & editing; **Jonathan Daniel Ip:** Data curation, Formal analysis, Writing – review & editing; **Yanling Luo:** Formal analysis, Writing – review & editing; **Hin Chu:** Formal analysis, Writing – review & editing; **Kwok-Yung Yuen:** Funding acquisition, Supervision, Writing – review & editing; **Kelvin Kai-Wang To:** Conceptualization, Formal analysis, Funding acquisition, Supervision, Writing – original draft.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This study was supported by the Partnership Programme of Enhancing Laboratory Surveillance and Investigation of Emerging Infectious Diseases and Antimicrobial Resistance for the Department of Health, HKSAR, Theme-Based Research Scheme [T11/709/21-N] of the Research Grants Council, HKSAR; the Global Influenza Hospital Surveillance Network (GISHN); and funding received from private donors including Richard Yu and Carol Yu, Shaw Foundation Hong Kong, Michael Seak-Kan Tong, May Tam Mak Mei Yin, Lee Wan Keung Charity Foundation Limited, Providence Foundation Limited (in memory of the late Lui Hac-Minh), Hong Kong Sanatorium & Hospital, Respiratory Viral Research Foundation Limited, Hui Ming, Hui Hoy and Chow Sin Lan Charity Fund Limited, Chan Yin Chuen Memorial Charitable Foundation, Marina Man-Wai Lee and May-Loy Kao.

Ethics statement

This study was approved by the Institutional Review Board of the University of Hong Kong/Hospital Authority Hong Kong West Cluster (UW 25-141). Written informed consent was waived since archive anonymized specimens were used.

ORCID

Kelvin Kai-Wang To http://orcid.org/0000-0002-1921-5824

References

[1] Chen LL, Abdullah SMU, Chan WM, et al. Contribution of low population immunity to the severe Omicron BA.2 outbreak in Hong Kong. Nat Commun. 2022;13(1):3618. doi:10.1038/s41467-022-31395-0

[2] Cheng VCC, Ip JD, Chu AWH, et al. Rapid spread of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) Omicron subvariant BA.2 in a single-source community outbreak. Clin Infect Dis. 2022;75(1):e44–e49. doi:10.1093/cid/ciac203

[3] Zhang X, Lam SJ, Ip JD, et al. Characterizing fitness and immune escape of SARS-CoV-2 EG.5 sublineage using elderly serum and nasal organoid. iScience. 2024;27(5):109706. doi:10.1016/j.isci.2024.109706

[4] Centre for Health Protection. Local situation of COVID-19 activity (as of Jul 15, 2026). [cited 2026 Jul 22]. Available from: https://www.chp.gov.hk/files/pdf/covid_flux_week28_16_7_2026_eng.pdf

[5] Witte L, Baharani VA, Schmidt F, et al. Epistasis lowers the genetic barrier to SARS-CoV-2 neutralizing antibody escape. Nat Commun. 2023;14(1):302. doi:10.1038/s41467-023-35927-0

[6] Dadonaite B, Brown J, McMahon TE, et al. Spike deep mutational scanning helps predict success of SARS-CoV-2 clades. Nat. 2024;631(8021):617–6626. doi:10.1038/s41586-024-07636-1

[7] Ataya F, Alamro A, Alghamdi A, et al. SARS-CoV-2 spike mutations alter structure and energetics to modulate ACE2 binding immune evasion and viral adaptation. Sci Rep. 2025;15(1):37546. doi:10.1038/s41598-025-15979-6

[8] Rucker G, Qin H, Zhang L. Structure, dynamics and free energy studies on the effect of point mutations on SARS-CoV-2 spike protein binding with ACE2 receptor. PLoS One. 2023;18(10):e0289432. doi:10.1371/journal.pone.0289432

[9] Wong SC, Chen JH, Chau PH, et al. Tracking SARS-CoV-2 RNA in the air: lessons from a COVID-19 outbreak in an infirmary unit. Am J Infect Control. 2025;53(3):348–3356. doi:10.1016/j.ajic.2024.10.035

[10] Ip JD, Chu WM, Chan WM, et al. The significance of recurrent de novo amino acid substitutions that emerged during chronic SARS-CoV-2 infection: an observational study. EBioMedicine. 2024;107:105273. doi:10.1016/j.ebiom.2024.105273

[11] Robinson JT, Thorvaldsdottir H, Winckler W, et al. Integrative genomics viewer. Nat Biotechnol. 2011;29(1):24–226. doi:10.1038/nbt.1754

[12] Ip JD, Wing-Ho Chu A, Chan WM, et al. Global prevalence of SARS-CoV-2 3CL protease mutations associated with nirmatrelvir or ensitrelvir resistance. EBioMedicine. 2023;91:104559. doi:10.1016/j.ebiom.2023.104559

[13] Starr TN, Greaney AJ, Stewart CM, et al. Deep mutational scans for ACE2 binding, RBD expression, and antibody escape in the SARS-CoV-2 Omicron BA.1 and BA.2 receptor-binding domains. PLoS Pathog. 2022;18(11):e1010951. doi:10.1371/journal.ppat.1010951

[14] World Health Organization. Tracking SARS-CoV-2 variants. 2026 [cited 2026 Sep 6]. Available from: https://www.who.int/activities/tracking-SARS-CoV-2-variants

[15] Anonymous. Pango-designation. 2026 [cited 2026 Jul 28]. Available from: https://github.com/cov-lineages/pango-designation/blob/master/lineage_notes.txt

[16] He P, Song Y, Guo C, et al. Antibody evasion and receptor binding of SARS-CoV-2 variants PQ.16.1.1 and RK.1. Lancet Infect Dis. 2026;26(10):e386–e387. doi:10.1016/S1473-3099(26)00429-9