TITLE PAGE

Full-length title:

Emergence in Southern France of a new SARS-CoV-2 variant of probably Cameroonian origin harbouring both substitutions N501Y and E484K in the spike protein

Short title (for the running head):

A new SARS-CoV-2 variant with spike substitutions N501Y and E484K

Author list: Philippe COLSON[1,2,3], Jérémy DELERCE[1], Emilie BUREL[1], Jordan DAHAN[4], Agnès JOUFFRET[5], Florence FENOLLAR[1,2,6], Nouara YAHI[7], Jacques FANTINI[7], Bernard LA SCOLA[1,2,3], Didier RAOULT[1,2] *

Affiliations: [1] IHU Méditerranée Infection, 19-21 boulevard Jean Moulin, 13005 Marseille, France; [2] Aix-Marseille Univ., Institut de Recherche pour le Développement (IRD), Microbes Evolution Phylogeny and Infections (MEPHI), 27 boulevard Jean Moulin, 13005 Marseille, France; [3] Assistance Publique-Hôpitaux de Marseille (AP-HM), 264 rue Saint-Pierre, 13005 Marseille, France; [4] Laboratoire de Biologie Médicale SYNLAB Provence Marseille, 25 rue Rabattu, 13015 Marseille, France; [5] Laboratoire de Biologie Médicale SYNLAB Provence Forcalquier, rue du Souvenir Français, 04300 Forcalquier, France; [6] Aix-Marseille Univ., Institut de Recherche pour le Développement (IRD), Vecteurs – Infections Tropicales et Méditerranéennes (VITROME), 27 boulevard Jean Moulin, 13005 Marseille, France; [7] Aix-Marseille Université, INSERM UMR S 1072, 51 boulevard Pierre Dramard, 13015 Marseille, France.

*** Corresponding author:** Didier Raoult, IHU Méditerranée Infection, 19-21 boulevard Jean Moulin, 13005 Marseille, France. Tel.: +33 413 732 401, Fax: +33 413 732 402; email:

didier.raoult@gmail.com

Keywords: SARS-CoV-2, variant, emergence, travel, southern France

Word counts: abstract, 198; text, 1,778

Figure: 1; **Table:** 1 **References:** 23

Supplementary Material: Figure: 1; **Tables:** 2; **References:** 2

ABSTRACT

SARS-CoV-2 variants have become a major virological, epidemiological and clinical concern, particularly with regard to the risk of escape from vaccine-induced immunity. Here we describe the emergence of a new variant. For twelve SARS-CoV-positive patients living in the same geographical area of southeastern France, qPCR testing that screen for variant-associated mutations showed an atypical combination. The index case returned from a travel in Cameroon. The genomes were obtained by next-generation sequencing with Oxford Nanopore Technologies on GridION instruments within ≈8 h. Their analysis revealed 46 mutations and 37 deletions resulting in 30 amino acid substitutions and 12 deletions. Fourteen amino acid substitutions, including N501Y and E484K, and 9 deletions are located in the spike protein. This genotype pattern led to create a new Pangolin lineage named B.1.640.2, which is a phylogenetic sister group to the old B.1.640 lineage renamed B.1.640.1. Both lineages differ by 25 nucleotide substitutions and 33 deletions. The mutation set and phylogenetic position of the genomes obtained here indicate based on our previous definition a new variant we named “IHU”. These data are another example of the unpredictability of the emergence of SARS-CoV-2 variants, and of their introduction in a given geographical area from abroad.

TEXT

SARS-CoV-2 has emerged in China in December 2019 and has been declared a pandemic 21 months ago [1]. We have shown since the summer of 2020 that several SARS-CoV-2 variants have emerged in our geographical area and caused distinct epidemics, either successive or superimposed [2,3]. In addition, we described that the origin of these variants was often their introduction from abroad but could also be mink. This was observed by genotyping, as of 09/12/2021, SARS-CoV-2 from almost 40,000 patients using next-generation sequencing (NGS) of complete genomes for more than 22,000 patients and implementing multiple qPCR specific of each variant for a more exhaustive assessment of their spread. Since then and with the emergence of the Alpha variant at the end of 2020, SARS-CoV-2 variants have become a major virological, epidemiological, and clinical concern, particularly with regard to the risk of escape from vaccine-induced immunity [4-7]. Here we describe the emergence in south-eastern France of a new variant of probably Cameroonian origin.

The index case was an adult first diagnosed as infected with SARS-CoV-2 by real-time reverse transcription PCR (qPCR) performed in a private medical biology laboratory on a nasopharyngeal sample collected mid-November 2021 (Table 1). He was vaccinated against SARS-CoV-2 and returned from a travel to Cameroon three days before. He developed mild respiratory symptoms the day before diagnosis. He lives in a small town of southeastern France. Subsequent detection by qPCR of three mutations in the spike gene to screen for variants, as systematically performed in France in case of SARS-CoV-2 positivity, revealed an atypical combination with L452R-negativity, E484K-positivity, and E484Q-negativity (Pentaplex assay, ID Solution, France) that did not correspond to the pattern of the Delta

variant involved in almost all SARS-CoV-2 infections at that time (Table 1). Respiratory samples collected from seven other SARS-CoV-2-positive patients living in the same geographical area exhibited the same combination of mutations screened by qPCR. They were two adults and five children (<15 years of age) (Table 1). The respiratory samples from these eight patients were sent to university hospital institute Méditerranée Infection for SARS-CoV-2 genome sequencing as recommended by French public health authorities. A rapid NGS procedure was launched overnight. It allowed obtaining SARS-CoV-2 genotype identification in ≈8 hours. Briefly, viral RNA was extracted from 200 µL of nasopharyngeal swab fluid using the KingFisher Flex system (Thermo Fisher Scientific, Waltham, MA, USA) following the manufacturer's instructions. Extracted RNA was reverse-transcribed using SuperScript IV (Thermo Fisher Scientific) and cDNA second strand was synthesized with LunaScript RT SuperMix kit (New England Biolabs) then amplified using a multiplex PCR protocol according to the ARTIC procedure (https://artic.network/) with ARTIC nCoV-2019 V3 panel of primers (IDT, Coralville, IA, USA). Finally, NGS was performed with the ligation sequencing kit and a GridION instrument of Oxford Nanopore Technologies (Oxford, UK) following manufacturer's instructions. Subsequently, fastq files were processed using the ARTIC field bioinformatics pipeline (https://github.com/artic-network/fieldbioinformatics). NGS reads were basecalled using Guppy (4.0.14) and aligned to the Wuhan-Hu-1 reference genome GenBank accession no. MN908947.3 using minimap2 (v2.17-r941) (https://github.com/lh3/minimap2) [8]. The ARTIC tool align_trim was used to softmask primers from read alignment and to cap sequencing depth at a maximum of 400. The identification of consensus-level variant candidates was performed using the Medaka (0.11.5) workflow developed by ARTIC (https://github.com/artic-network/artic-ncov2019). This strategy allowed assemblying the complete genome from NGS reads obtained within 30 min of run for cycle threshold values (Ct) of qPCR comprised between 15 and 27. SARS-CoV-2

genomes were classified into Nextclade and Pangolin lineages using web applications (https://clades.nextstrain.org/;https://cov-lineages.org/pangolin.html) [10,11,13]. They were deposited in the GISAID sequence database (https://www.gisaid.org/) [14] (Table 1). Phylogenies were reconstructed with the nextstrain/ncov tool (https://github.com/nextstrain/ncov) then visualized with Auspice (https://docs.nextstrain.org/projects/auspice/en/stable/).

The analysis of viral genomes revealed the presence of 46 nucleotide substitutions and 37 deletions, resulting in 30 amino acid substitutions and 12 deletions (Figure 1a; Supplementary Tables S1 and S2). Fourteen amino acid substitutions and 9 amino acid deletions are located in the spike protein. Substitutions N501Y and E484K are combined as in the Beta, Gamma, Theta and Omicron variants [5,15]. Substitution F490S is present as in the Lambda variant, and substitution P681H is present as in the Lambda and Omicron variants. In other structural proteins than the spike, amino acid changes include two substitutions in the nucleocapsid protein and one in the membrane protein. In non-structural proteins, amino acid changes include one substitution in proteins Nsp2, Nsp3, Nsp4, Nsp6, Nsp12 (RNA-dependent RNA polymerase), and Nsp13 (helicase); two substitutions in Nsp14 (3'-5'exonuclease); four substitutions in Nsp8 (which is part of the replication complex with Nsp7 and Nsp12); and three deletions in Nsp6. Finally, in regulatory proteins, amino acid changes include four substitutions in ORF3a, one in ORF9b and one in ORF8. In addition, codon 27 of ORF8 gene is changed into a stop codon, as in the Alpha variant [16]; some members of the Marseille-4 variant (B.1.160) that predominated in our geographical area between August 2020 and February 2021 also exhibit a stop codon in ORF8 gene but at another position [3].

Nextclade identified a 20A lineage. Pangolin identified a B.1.640 lineage in primary analysis but a B.1 lineage with the -usher (Ultrafast Sample placement on Existing tRee;

https://genome.ucsc.edu/cgi-bin/hgPhyloPlace) option, which showed the phylogenetic placement of the genomes we obtained as an outgroup of the B.1.640 lineage and their clustering with a genome obtained late October in France (Ile-de-France) (EPI_ISL_5926666). The B.1.640 lineage corresponds to a variant first identified in France in April 2021, in Indonesia in August 2021, and in Republic of the Congo (Brazzaville) in September 2021, and it was involved in a cluster of cases in Brittany, France around mid-October 2021 [17]. As of 09/12/2021, 157 genomes were available from the GISAID database including 92 from France and 36 from the Republic of the Congo. The sets of spike mutations of the B.1.640 lineage and of genomes obtained here are similar, with 11 common nucleotide substitutions and 1 common deletion of 9 codons (Supplementary Figure S1, Tables S1-2). However, spike genes of both lineages differ by 7 mutations. In addition, 25 nucleotide substitutions and 33 nucleotide deletions located elsewhere in the genomes differ between the two genotypes. The pattern of mutations of present genomes hence indicates a new variant, which we named “IHU” (in reference to our institute), based on our previous definition [3]. Phylogeny performed with nextstrain/ncov tool (https://github.com/nextstrain/ncov) also showed that B.1.640 and IHU variants were most closely related between each other but comprised two divergent branches (Figure 1b). Their last common ancestor is estimated to date from January 2021 but no genome is currently available from GISAID that corresponds to it. Accordingly, a new Pangolin clade corresponding to the IHU variant was created on 07/12/2021 that was named B.1.640.2, the old B.1.640 clade being renamed B.1.640.1 (https://github.com/cov-lineages/pango-designation/issues/362). It encompasses present genomes and three other genomes comprising a sister group including the one recovered late October 2021 in France (Ile-de-France) (EPI_ISL_5926666) and two additional genomes obtained from samples collected late November in England (EPI_ISL_7181977) and Wales (EPI_ISL_7402094). As the index case was probably infected with the IHU variant during his

stay in Cameroon, we sought for this variant in GISAID among genomes from this country but as of 09/12/2021 none of the 556 available genomes belong to the B.1.640.1 or B.1.640.2 lineages.

We analyzed a complete structure of the spike protein of the IHU variant generated by incorporating its specific mutational profile to the original 20B SARS-CoV-2 (Wuhan-Hu-1 isolate with D614G substitution) [18] and fixing all gaps in the pdb file by incorporating the missing amino acids with the Robetta protein structure prediction tool [https://robetta.bakerlab.org/], followed by energy minimization with the Polak-Ribière algorithm as previously reported (Figure 1c) [19]. In the N-terminal domain (NTD), the 134-145 amino acid deletion is predicted to significantly affect the neutralizing epitope. Other changes involve amino acids at positions 96 and 190: in Wuhan-Hu-1 isolate, E96 and R190 induce a turn in NTD secondary structure through electrostatic interactions between each other. This interaction is conserved between substituted amino acids 96Q and 190S, which suggests the co-evolution of these changes. In the receptor binding domain (RBD), aside the well-known substitutions N501Y and E484K, several changes were predicted to significantly affect the neutralizing epitopes. Particularly, P681H is located in the cleavage site of S1-S2 subunits of the spike and is observed in other variants including the recently emerging Omicron [15]. Besides, D1139H substitution implies an amino acid involved in the fusion between the virus and the infected cell.Also, D614G is combined with T859N in the IHU variant. Interestingly, in the Wuhan-Hu-1 isolate, amino acids D614 and T859 from two subunits of the trimeric spike are face to face and lock the trimer in a closed conformation. Substitution D614G allows unlocking the trimer conformation, but this is predicted to be still easier in case of additional presence of substitution T859N.

Respiratory samples collected until end of November 2021 from four other SARS-CoV-2 positive patients living in the same city or borough than the index case could be

identified as containing the IHU variant by NGS within 24 hours after their reception (Table 1). All 12 IHU variant-positive samples showed the same combination of spike mutations as screened by real-time qPCR techniques: negativity for 452R and 484Q; positivity for 484K, 501Y [20], and 681H [3]. We also used the TaqPath COVID-19 kit (Thermo Fisher Scientific, Waltham, USA) that provided positive signals for all three genes targeted (ORF1, S, and N). Thus, the IHU variant can be distinguished by screening with qPCR assays from the Delta (L452R-positive) and Omicron (L452R-negative and negative for S gene detection by the TaqPath COVID-19 assay) variants that currently co-circulate in our geographical area. Finally, scanning electron microscopy using a SUV 5000 microscope (Hitachi High-Technologies Corporation, Tokyo, Japan) [21] allowed a quick visualization of the virus from a respiratory sample (Figure 1d).

Overall, these observations show once again the unpredictability of the emergence of new SARS-CoV-2 variants and their introduction from abroad, and they exemplify the difficulty to control such introduction and subsequent spread. They also warrant the implementation of genomic surveillance of SARS-CoV-2 that we started from the very beginning of the pandemic in our geographical area as soon as we diagnosed the first SARS-CoV-2 infection [21] and that we expanded during summer 2020 [2,3]. This surveillance has been implemented at the country scale in 2021 through the French Emergen consortium (https://www.santepubliquefrance.fr/dossiers/coronavirus-covid-19/consortium-emergen). It is too early to speculate on virological, epidemiological or clinical features of this IHU variant based on these 12 cases. For this purpose, respiratory samples from infected patients were inoculated on Vero E6 cells as previously described [22] to be able assessing the sensibility to neutralization by anti-spike antibodies elicited by vaccine immunization, or by prior infection [23].

Acknowledgments

We are thankful to the Emergen French consortium (https://www.santepubliquefrance.fr/dossiers/coronavirus-covid-19/consortium-emergen). We are also grateful to Laurence Thomas, Claudia Andrieu, Ludivine Brechard, Mamadou Beye, Marielle Bedotto, Elsa Prudent, Sofiane Bakour, Jacques Bou Khalil, and Clio Grimaldier for their technical help.

Author contributions

Conceived and designed the experiments: PC, DR, JF, BLS. Contributed materials/analysis tools: PC, JDe, EB, JDa, AJ, FF, NY, JF. Analyzed the data: PC, DR, BLS, JD, EB, JF, NY. Wrote the paper: PC, JF, DR. All authors approved the last version of the manuscript.

Funding

This work was supported by the French Government under the “Investments for the Future” program managed by the National Agency for Research (ANR), Méditerranée-Infection 10-IAHU-03 and was also supported by Région Provence Alpes Côte d’Azur and European funding FEDER PRIMMI (Fonds Européen de Développement Régional-Plateformes de Recherche et d'Innovation Mutualisées Méditerranée Infection), FEDER PA 0000320 PRIMMI, and by Hitachi High-Technologies Corporation, Tokyo, Japan.

Conflicts of interest

DR has a conflict of interest being a consultant for Hitachi High-Technologies Corporation, Tokyo, Japan from 2018 to 2020. All other authors have no conflicts of interest to declare.

Funding sources had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of the manuscript.

Ethics

This study has been approved by the ethics committee of University Hospital Institute (IHU) Méditerranée Infection (N°2021-029). Access to the patients' biological and registry data issued from the hospital information system was approved by the data protection committee of Assistance Publique-Hôpitaux de Marseille (APHM) and was recorded in the European General Data Protection Regulation registry under number RGPD/APHM 2019-73.

REFERENCES

1. Cucinotta D, Vanelli M (2020) WHO Declares COVID-19 a Pandemic. Acta Biomed 91: 157-160.

2. Colson P, Levasseur A, Delerce J, Chaudet H, Bossi V, Ben Khedher M, Fournier PE, Lagier JC, Raoult D (2020) Dramatic increase in the SARS-CoV-2 mutation rate and low mortality rate during the second epidemic in summer in Marseille. IHU pre-prints https://doi.org/10.35088/68c3-ew82 (accessed 10 December 2021).

3. Colson P, Fournier PE, Chaudet H, Delerce J, Giraud-Gatineau A, Houhamdi L, Andrieu C, Brechard L, Bedotto M, Prudent E, Gazin C, Beye M, Burel E, Dudouet P, Tissot-Dupont H, Gautret P, Lagier JC, Million M, Brouqui P, Parola P, Drancourt M, La Scola B, Levasseur A, Raoult D (2021) Analysis of SARS-CoV-2 variants from 24,181 patients exemplifies the role of globalisation and zoonosis in pandemics. medRxiv doi: https://doi.org/10.1101/2021.09.10.21262922 (accessed 10 December 2021).

4. Hastie KM, Li H, Bedinger D, Schendel SL, Dennison SM, Li K, Rayaprolu V, Yu X, Mann C, Zandonatti M, Diaz Avalos R, Zyla D, Buck T, Hui S, Shaffer K, Hariharan C, Yin J, Olmedillas E, Enriquez A, Parekh D, Abraha M, Feeney E, Horn GQ; CoVIC-DB team1, Aldon Y, Ali H, Aracic S, Cobb RR, Federman RS, Fernandez JM, Glanville J, Green R, Grigoryan G, Lujan Hernandez AG, Ho DD, Huang KA, Ingraham J, Jiang W, Kellam P, Kim C, Kim M, Kim HM, Kong C, Krebs SJ, Lan F, Lang G, Lee S, Leung CL, Liu J, Lu Y, MacCamy A, McGuire AT, Palser AL, Rabbitts TH, Rikhtegaran Tehrani Z, Sajadi MM, Sanders RW, Sato AK, Schweizer L, Seo J, Shen B, Snitselaar JL, Stamatatos L, Tan Y, Tomic MT, van Gils MJ, Youssef S, Yu J, Yuan TZ, Zhang Q, Peters B, Tomaras GD, Germann T, Saphire EO (2021) Defining variant-

resistant epitopes targeted by SARS-CoV-2 antibodies: A global consortium study. Science 374: 472-478.

5. Harvey WT, Carabelli AM, Jackson B, Gupta RK, Thomson EC, Harrison EM, Ludden C, Reeve R, Rambaut A (2021) COVID-19 Genomics UK (COG-UK) Consortium, Peacock SJ, Robertson DL. SARS-CoV-2 variants, spike mutations and immune escape. Nat Rev Microbiol 19: 409-424.

6. Tao K, Tzou PL, Nouhin J, Gupta RK, de Oliveira T, Kosakovsky Pond SL, Fera D, Shafer RW (2021) The biological and clinical significance of emerging SARS-CoV-2 variants. Nat Rev Genet 22: 757-773.

7. Wilder-Smith A (2021) What is the vaccine effect on reducing transmission in the context of the SARS-CoV-2 delta variant? Lancet Infect Dis S1473-3099(21)00690-3. doi: 10.1016/S1473-3099(21)00690-3. Epub ahead of print. PMID: 34756187; PMCID: PMC8554481.

8. Li H (2018). Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics 34: 3094-3100.

9. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics 25: 2078-2079.

10. Hadfield J, Megill C, Bell SM, Huddleston J, Potter B, Callender C, Sagulenko P, Bedford T, Neher RA (2018). Nextstrain: real-time tracking of pathogen evolution. Bioinformatics 34: 4121-4123.

11. Aksamentov I, Roemer C, Hodcroft EB, Neher RA (2021). Nextclade: clade assignment, mutation calling and quality control for viral genomes. Zenodo https://doi.org/10.5281/zenodo.5607694.

12. Garrison E, Marth G (2012). Haplotype-based variant detection from short-read sequencing. arXiv.org. https://arxiv.org/abs/1207.3907 (accessed 10 December 2021).
13. Rambaut A, Holmes EC, O'Toole Ã, Hill V, McCrone JT, Ruis C, du Plessis L, Pybus OG (2020). A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. Nat. Microbiol 5: 1403-1407.
14. Alm E, Broberg EK, Connor T, Hodcroft EB, Komissarov AB, Maurer-Stroh S, Melidou A, Neher RA, O'Toole A, Pereyaslov D, WHO European Region sequencing laboratories and GISAID EpiCoV group; WHO European Region sequencing laboratories and GISAID EpiCoV group (2020) Geographical and temporal distribution of SARS-CoV-2 clades in the WHO European Region, January to June 2020. Euro. Surveill 25: 2001410.
15. Karim SSA, Karim QA (2021) Omicron SARS-CoV-2 variant: a new chapter in the COVID-19 pandemic. Lancet 398: 2126-2128.
16. Rambaut A, Loman N, Pybus O, Barclay W, Barrett J, Carabelli A, Connor T, Peacock T, Robertson DL, Volz E, on behalf of COVID-19 Genomics Consortium UK (CoG-UK) (2020) Preliminary genomic characterisation of an emergent SARS-CoV-2 lineage in the UK defined by a novel set of spike mutations. Virological Pre-print. https://virological.org/t/preliminary-genomic-characterisation-of-an-emergent-sars-cov-2-lineage-in-the-uk-defined-by-a-novel-set-of-spike-mutations/563
17. Le Page M (2021) New variant gains ground. New Sci 252: 8.
18. Benton DJ, Wrobel AG, Roustan C, Borg A, Xu P, Martin SR, Rosenthal PB, Skehel JJ, Gamblin SJ (2021). The effect of the D614G substitution on the structure of the spike glycoprotein of SARS-CoV-2. Proc. Natl. Acad. Sci. U S A. 118:e2022586118.

19. Fantini J, Yahi N, Azzaz F, Chahinian H (2021) Structural dynamics of SARS-CoV-2 variants: A health monitoring strategy for anticipating Covid-19 outbreaks. J Infect 83: 197-206.

20. Bedotto M, Fournier PE, Houhamdi L, Colson P, Raoult D (2021) Implementation of an in-house real-time reverse transcription-PCR assay to detect the emerging SARS-CoV-2 N501Y variants. J Clin Virol 140: 104868.

21. Colson P, Lagier JC, Baudoin JP, Bou Khalil J, La Scola B, Raoult D (2020) Ultrarapid diagnosis, microscope imaging, genome sequencing, and culture isolation of SARS-CoV-2. Eur J Clin Microbiol Infect Dis 39: 1601-1603.

22. Wurtz N, Penant G, Jardot P, Duclos N, La Scola B (2021) Culture of SARS-CoV-2 in a panel of laboratory cell lines, permissivity, and differences in growth profile. Eur J Clin Microbiol Infect Dis 40: 477-484.

23. Jaafar R, Boschi C, Aherfi S, Bancod A, Le Bideau M, Edouard S, Colson P, Chahinian H, Raoult D, Yahi N, Fantini J, La Scola B (2021) High individual heterogeneity of neutralizing activities against the original strain and nine different variants of SARS-CoV-2. Viruses 13: 2177.

FIGURE LEGENDS

Figure 1. Virological features and scanning electron microscopy image of the SARS-CoV-2 IHU variant

a: Map of the IHU variant genome showing amino acid substitutions and deletions.

b: Phylogeny reconstruction performed using the nextstrain/ncov tool (https://github.com/nextstrain/ncov) then visualized with Auspice (https://docs.nextstrain.org/projects/auspice/en/stable/). The genome of the original Wuhan-Hu-1 SARS-CoV-2 isolate (GenBank accession no. NC_045512.2) was added as outgroup, in addition to SARS-CoV-2 genomes of Pangolin lineages B.1.640.1 and B.1.640.2. X-axis shows time.

c: Representations of the spike of the IHU variant showing the location of all its amino acid substitutions. N-terminal domain (NTD) mutations are in blue; receptor binding domain (RBD) mutations are in red; mutations involved in ACE-2 unmasking are in yellow; mutations at S1-S2 cleavage site are in green; mutations at fusion region are in cyan.

d: Scanning electron microscopy image obtained using a SUV 5000 microscope from a respiratory sample positive for the SARS-CoV-2 IHU variant (Hitachi High-Technologies Corporation, Tokyo, Japan).

TABLES

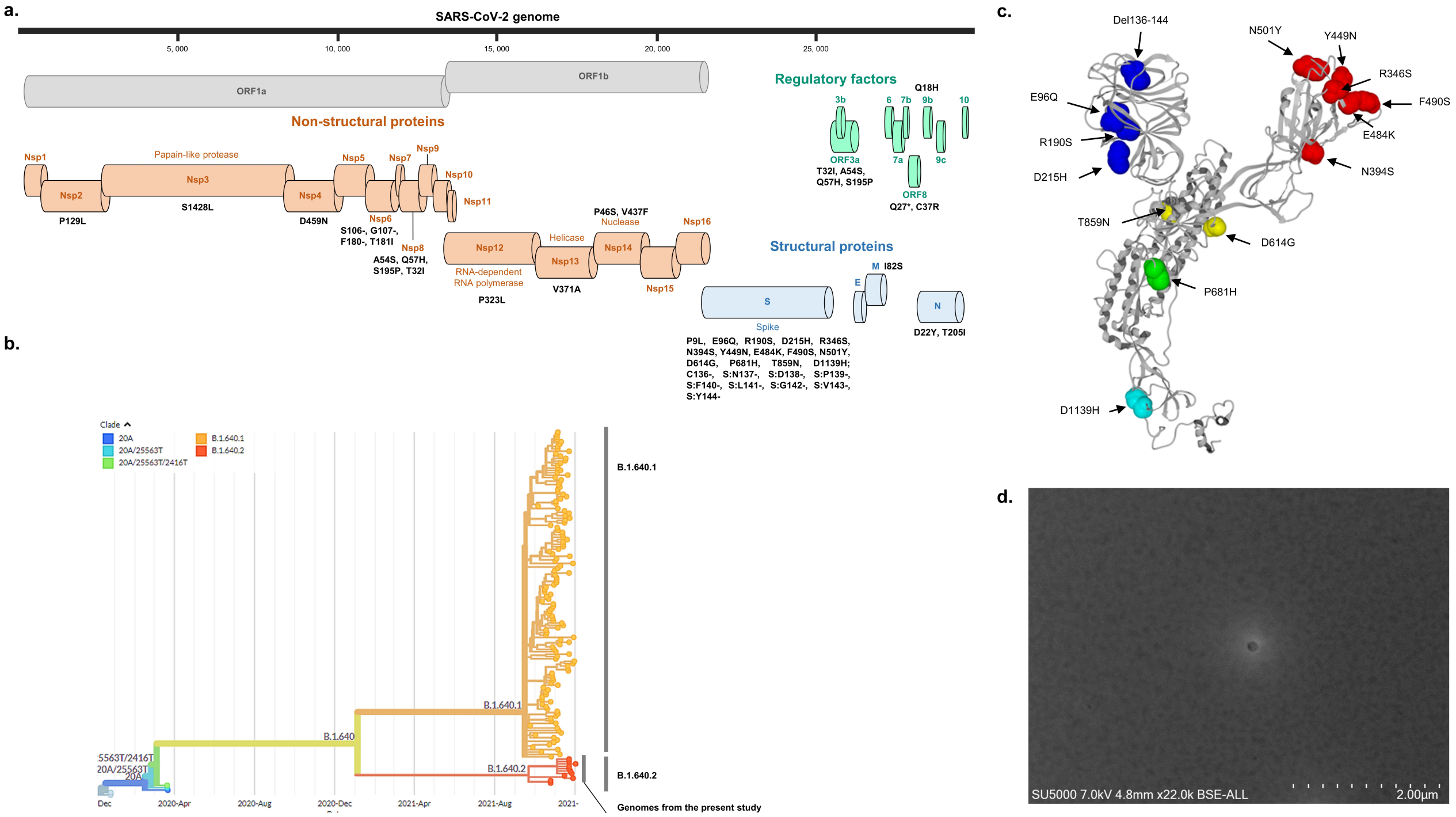
Table 1. Main epidemiological and virological features of cases identified with infection with the SARS-CoV-2 IHU variant

Case no.	Age	Diagnostic qPCR Ct	Results of qPCR used to screen for the presence of SARS-CoV-2 spike substitutions	Results of the TaqPath COVID-19 qPCR assay (Targets: ORF1, S, and N genes)	Genome GISAID Id.
1 *	Adult	27	L452R-neg.; E484K-pos.; E484Q-neg.; N501Y-pos.; P681H-pos.	Pos. for all three genes	EPI_ISL_7156955
2	Child	21	L452R-neg.; E484K-pos.; E484Q-neg.; N501Y-pos.; P681H-pos.	Pos. for all three genes	EPI_ISL_7314302
3	Child	15	L452R-neg.; E484K-pos.; E484Q-neg.; N501Y-pos.; P681H-pos.	Pos. for all three genes	EPI_ISL_7381031
4	Child	18	L452R-neg.; E484K-pos.; E484Q-neg.; N501Y-pos.; P681H-pos.	Pos. for all three genes	EPI_ISL_7381062
5	Adult	15	L452R-neg.; E484K-pos.; E484Q-neg.; N501Y-pos.; P681H-pos.	Pos. for all three genes	EPI_ISL_7156959
6	Adult	17	L452R-neg.; E484K-pos.; E484Q-neg.; N501Y-pos.; P681H-pos.	Pos. for all three genes	EPI_ISL_7314417
7	Child	19	L452R-neg.; E484K-pos.; E484Q-neg.; N501Y-pos.; P681H-pos.	Pos. for all three genes	EPI_ISL_7314514
8	Child	26	L452R-neg.; E484K-pos.; E484Q-neg.; N501Y-pos.; P681H-pos.	Pos. for all three genes	EPI_ISL_7314471
9	Adult	15	L452R-neg.; E484K-pos.; E484Q-neg.; N501Y-n.t..; P681H- n.t.	Pos. for all three genes	EPI_ISL_7552465
10	Adult	16	L452R-neg.; E484K-pos.; E484Q-neg.; N501Y-n.t..; P681H- n.t.	Pos. for all three genes	EPI_ISL_7552470
11	Adult	22	L452R-neg.; E484K-pos.; E484Q-neg.; N501Y-n.t..; P681H- n.t.	Pos. for all three genes	EPI_ISL_7552483
12	Adult	15	L452R-neg.; E484K-pos.; E484Q-neg.; N501Y-n.t..; P681H- n.t.	Pos. for all three genes	EPI_ISL_7601710

* Index case; travelled to Cameroon

Ct, cycle threshold; Id., identifier; neg., negative; N, nucleocapsid; no., number; n.t., not tested; ORF1, open reading frame 1; pos., positive; qPCR, real-time reverse-transcription PCR; S, spike.

All 12 respiratory samples were collected between mid and end of November 2021.


a.
SARS-CoV-2 genome
5, 000
10, 000
15, 000
20, 000
25, 000
ORF1a
ORF1b
Non-structural proteins
Nsp1
P129L
Papain-like protease
Nsp2
Nsp3
S1428L
Nsp4
D459N
Nsp5
Nsp6
S106-, G107-, F180-, T181I
Nsp7
Nsp8
A54S, Q57H, S195P, T32I
Nsp9
Nsp10
Nsp11
Nsp12
RNA-dependent RNA polymerase
P323L
Helicase
Nsp13
V371A
P46S, V437F
Nuclease
Nsp14
Nsp15
Nsp16
Regulatory factors
3b
ORF3a
T32I, A54S, Q57H, S195P
6
7a
7b
ORF8
Q27*, C37R
9b
9c
Q18H
10
Structural proteins
S
Spike
P9L, E96Q, R190S, D215H, R346S, N394S, Y449N, E484K, F490S, N501Y, D614G, P681H, T859N, D1139H; C136-, S:N137-, S:D138-, S:P139-, S:F140-, S:L141-, S:G142-, S:V143-, S:Y144-
E
M
I82S
N
D22Y, T205I
b.
Clade
20A
20A/25563T
20A/25563T/2416T
B.1.640.1
B.1.640.2
B.1.640.1
B.1.640
B.1.640.2
B.1.640.1
B.1.640.2
Genomes from the present study
Dec
2020-Apr
2020-Aug
2020-Dec
2021-Apr
2021-Aug
2021-
c.
Del136-144
N501Y
Y449N
R346S
F490S
E96Q
E484K
R190S
N394S
D215H
T859N
D614G
P681H
D1139H
d.
SU5000 7.0kV 4.8mm x22.0k BSE-ALL
2.00µm

SUPPLEMENTARY MATERIAL

SUPPLEMENTARY FIGURE LEGENDS

Supplementary Figure S1. Microarray showing the distribution along the SARS-CoV-2 genome and in viral genes of nucleotide changes observed in comparison with the genome of the Wuhan-Hu-1 isolate for the Pangolin B.1.640.1 and B.1.640.2 (IHU variant) lineages.

Genomes were analyzed using the Nextstrain web-tool (https://clades.nextstrain.org/) [1,2]. Representation is adapted from Nextclade sequence analysis web application output (https://clades.nextstrain.org/).

SUPPLEMENTARY TABLES

Supplementary Table S1. Comparison of nucleotide mutational patterns of the B.1.640.1 and B.1.640.2 lineages

B.1.640.1	B.1.640.2 (IHU variant)
C241T	C241T
C601T	-
C1191T	C1191T
A1620G	-
C2416T	C2416T
C2455T	C2455T
C3037T	C3037T
G6622T	-
G7328T	-
C9711T	-
G9756A	-
-	G9929A
C10029T	-
-	T10561C
-	Deletion 11288-11296
T11418C	-
-	C11514T
C11956T	C11956T
C14408T	C14408T
C16575T	-
C16869T	C16869T
-	C17004T
-	T17348C
-	A17916G
-	C18175T
-	C18804T
-	G19348T
-	T19680C
C20283T	-
-	A21258G
C21588T	C21588T
G21848C	G21848C
Deletion 21968-21994	Deletion 21968-21994
G22132T	G22132T
T22191C	-
-	G22205C
A22600C	A22600C
A22743G	A22743G
T22907A	T22907A
-	G23012A
T23030C	-
T23031**G**	T23031**C**
A23063T	A23063T
A23403G	A23403G
C23604A	C23604A
C24138A	C24138A
G24368C	-
-	G24977C
C25487T	C25487T
G25563T	G25563T
A26492T	-
T26767C	T26767G
C27513T	C27513T
C27807T	C27807T
-	T27833C
C27972T	C27972T
-	T28002C
Deletion 28271	Deletion 28271
T28297C	C28312T
G28337T	G28337T
C28887T	C28887T
T29377C	T29377C
G29405C	-
Deletion 29738-29758	-
G29779T	G29779T

Genomes were analyzed using the Nextstrain web-tool (https://clades.nextstrain.org/) [1,2].

Spike region is indicated by a grey background.

Supplementary Table S2. Comparison of amino acid mutational patterns of the B.1.640.1 and B.1.640.2 lineages

B.1.640.1	B.1.640.2 (IHU variant)
M:I82T	M:I82S
N:D22Y	N:D22Y
N:T205I	N:T205I
N:E378Q	-
ORF1a:P309L	ORF1a:P309L
ORF1a:E452G	-
ORF1a:L2119F	-
ORF1a:A2355S	-
ORF1a:S3149F	-
ORF1a:R3164H	-
-	ORF1a:D3222N
ORF1a:T3255I	-
-	Deletions ORF1a:S3675-, ORF1a:G3676-, ORF1a:F3677-
ORF1a:V3718A	-
-	ORF1a:T3750I
ORF1b:P314L	ORF1b:P314L
-	ORF1b:V1294A
-	ORF1b:P1570S
-	ORF1b:V1961F
ORF3a:T32I	ORF3a:T32I
ORF3a:Q57H	ORF3a:Q57H
ORF8:Q27*	ORF8:Q27*
ORF9b:I5T	-
ORF9b:Q18H	-
-	ORF8:C37R
-	ORF9b:P10L
-	ORF9b:Q18H
S:P9L	S:P9L
S:E96Q	S:E96Q
Deletions S:C136-, S:N137-, S:D138-, S:P139-, S:F140-, S:L141-, S:G142-, S:V143-, S:Y144-	Deletions S:C136-, S:N137-, S:D138-, S:P139-, S:F140-, S:L141-, S:G142-, S:V143-, S:Y144-
S:R190S	S:R190S
S:I210T	-
-	S:D215H
S:R346S	S:R346S
S:N394S	S:N394S
S:Y449N	S:Y449N
-	S:E484K
S:F490R	S:F490S
S:N501Y	S:N501Y
S:D614G	S:D614G
S:P681H	S:P681H
S:T859N	S:T859N
S:D936H	-
-	S:D1139H

Genomes were analyzed using the Nextstrain web-tool (https://clades.nextstrain.org/) [1,2].

Spike region is indicated by a grey background.

REFERENCES

1. Hadfield J, Megill C, Bell SM, Huddleston J, Potter B, Callender C, Sagulenko P, Bedford T, Neher RA (2018). Nextstrain: real-time tracking of pathogen evolution. Bioinformatics 34: 4121-4123.

2. Aksamentov I, Roemer C, Hodcroft EB, Neher RA (2021). Nextclade: clade assignment, mutation calling and quality control for viral genomes. Zenodo https://doi.org/10.5281/zenodo.5607694.