



Coding-Complete Genome Sequences of a Delta Subvariant (AY.33) of SARS-CoV-2 Obtained from Moroccan COVID-19 Patients

Mouhssine Hemlali,^{a,b} Taha Chouati,^{a,b} Hamza Ghammaz,^{a,d} Marouane Melloul,^{a,d} Sanaâ Alaoui Amine,^{a,b} Safaa Rhoulam,^a Nadia Touil,^{c,i} Khalid Ennibi,^h Hicham Oumzil,^g Rhajaoui Mohamed,^f Aguenou Hassan,^e  Elmostafa El Fahime^{a,b}

^aMolecular Biology and Functional Genomics Platform, National Center for Scientific and Technical Research (CNRST), Rabat, Morocco

^bNeuroscience and Neurogenetics Research Team, Faculty of Medicine and Pharmacy, University Mohammed V, Rabat, Morocco

^cCell Culture Unit, Center of Virology, Infectious, and Tropical Diseases, Med V Military Hospital, Rabat, Morocco

^dMicrobiology and Molecular Biology Team, Center of Plant and Microbial Biotechnology, Biodiversity, and Environment, Faculty of Sciences, Mohammed V University of Rabat, Rabat, Morocco

^eIbn Tofail University-CNESTEN, Joint Research Unit in Nutrition and Food, RDC-Nutrition AFRA/IAEA, Rabat-Kénitra, Morocco

^fNational Institute of Hygiene, Rabat, Morocco

^gNational Influenza Center, Virology Department, Institut National d'Hygiène, Rabat, Morocco

^hVirology and Infectious Tropical Diseases Center, Med V Military Hospital, Rabat, Morocco

ⁱGenomic Center for Human Pathologies (GENOPATH), Faculty of Medicine and Pharmacy, University Mohammed V, Rabat, Morocco

ABSTRACT We report here the complete genome sequences of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) strains obtained from Moroccan patients with COVID-19. The analysis of these sequences indicates that the identified strains belong to the AY.33 sublineage of the Delta variant.

In December 2019, the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), first described in the city of Wuhan in the Republic of China, and was identified as the etiological agent of the coronavirus disease 19 (COVID-19) (1). SARS-CoV-2 is a member of the *Coronaviridae* family and *Betacoronavirus* genus (2). In Morocco, as of 3 October 2021, more than 935,000 COVID cases had been reported, with a comparatively high death rate (14,339 deaths) (www.covidmaroc.ma). Here, we report the complete genome sequences of two SARS-CoV-2 strains recovered from Moroccan patients.

As part of the genomic surveillance against SARS-CoV-2, nasopharyngeal swabs from two Moroccan patients living in Casablanca were received on 7 and 8 July 2021, respectively. The viral RNA was extracted using the MagPurix instrument (Zinexts Life Science, Taiwan). Confirmation of the infection of the patients with SARS-CoV-2 was done by reverse transcription-quantitative PCR (qRT-PCR) using the GeneFinder COVID-19 Plus RealAmp kit at the National Center for Scientific and Technical Research, Rabat, Morocco. The quantity and quality of RNA were measured with a Qubit 4 fluorometer (Invitrogen, Thermo Fisher Scientific, USA) using the Qubit RNA high-sensitivity (HS) assay kit (Invitrogen, Thermo Fisher Scientific).

The viral RNA was reverse-transcribed using a SuperScript VILO cDNA synthesis kit (Invitrogen, Thermo Fisher Scientific, USA) and processed for library preparation using the Ion AmpliSeq SARS-CoV-2 research panel (Thermo Fisher, USA). The DNA libraries were then used for template preparation and chip loading on the Ion Chef instrument. Sequencing was carried out on the Ion S5 sequencer. Then, using Unipro UGENE version 38, a consensus sequence was generated by mapping the reads against the Wuhan-Hu-1 reference sequence (GenBank accession number [NC_045512.2](https://www.ncbi.nlm.nih.gov/nuccore/NC_045512.2)). All tools were run with default parameters unless otherwise specified.

The genomes for samples 5858 and 5900 were obtained with an average read depth of 3,675 \times and 5,445 \times and showed lengths of 29,818 bp and 29,822 bp, respectively, and a GC

Editor Simon Roux, DOE Joint Genome Institute

Copyright © 2022 Hemlali et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Elmostafa El Fahime, m.elfahime@gmail.com.

The authors declare no conflict of interest.

Received 20 November 2021

Accepted 6 January 2022

Published 3 February 2022

TABLE 1 Amino acid variants of the sequences of hCoV-2019/5858 and hCoV2019-5900 in comparison to the reference strain (GenBank accession number [NC_045512.2](https://www.ncbi.nlm.nih.gov/nuccore/NC_045512.2)) using the web application CoV-GLUE

Open reading frame	Variants for:	
	hCoV-2019/5858	hCoV-2019/5900
ORF 1a	A1306S	A1306S
	P2046L	P2046L
	P2287S	P2287S
	V2930L	V2930L
	T3255I	T3255I
	L3606F	T3646A
	T3646A	
	I4049V	
ORF 1b	P314L	P314L
	G662S	G662S
	P1000L	P1000L
	A1918V	A1918V
S	T19R	T19R
	T29A	T29A
	G142D	G142D
	R158G	R158G
	T250I	T250I
	T299I	T299I
	L452R	L452R
	T478K	T478K
	Q613H	Q613H
	D614G	D614G
	P681R	P681R
D950N	D950N	
E157-F158-	E156-F157-	
ORF 3a	S26L	S26L
M		I82T
ORF 7a	V82A	V82A
	T120I	T120I
ORF 7b	T40I	T40I
ORF 8	D119- F120-	
ORF 9B	T60A	T60A
N	D63G	D63G
	R203M	R203M
	G215C	G215C
	N377Y	N377Y

content of 37.9%. The quality of the sequences was monitored using the Nextclade web tool version 1.7.2 (<https://clades.nextstrain.org>), although the determination of the mutations was carried out using the web application Cov-GLUE (<http://cov-glue.cvr.gla.ac.uk/>), and the lineage and sublineage were obtained using the web application Pangolin (<https://pangolin.cog-uk.io/>).

Both samples belonged to clade GK, sublineage AY.33 of Delta variant B.1.617.2 (Delta 2). The mutations within the viral genome of sample 5858 (31 mutations) and sample 5900 (30 mutations) are summarized in Table 1, and their localization within global SARS-CoV-2 diversity are illustrated in Fig. 1.

The genomes of both samples accumulated 4 new amino acid changes within the Spike protein (T29A, T50I, T299I, and Q613) in addition to the 10 modifications that characterize the Delta 2 variant.

T487K, which results in a significant modification in the conformation of the receptor binding domain (RBD), was linked to an increase in RBD binding affinity to the cell surface

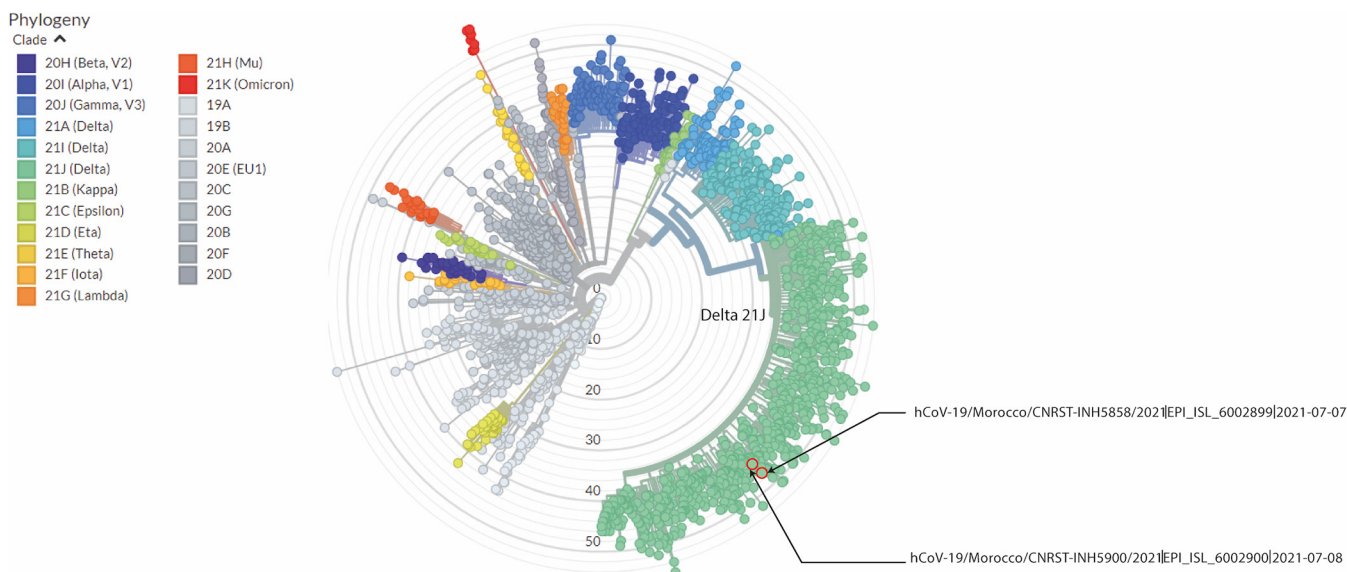


FIG 1 Phylogenetic tree of SARS-CoV-2 generated from Nextstrain using the Nextclade maximum likelihood module. The hCoV-19/5858 and hCoV-19/5900 strains are situated in the Delta lineage (green) and, more specifically, in the AY.33 sublineage.

angiotensin converting enzyme 2 (ACE2) receptor (3). In addition, Q613H was also reported as an amino acid modification of little-known impact; however, given its close proximity to the well-known D614G, it is possible that this amino acid change would enhance the role of the latter in terms of virus transmissibility (4, 5).

Furthermore, among amino acid modifications affecting the receptor binding motif (RBM) region, L452R is present in variants which exhibit high binding affinity to the ACE2 receptor and may also contribute to humoral immune escape (6). Another interesting amino acid change is P681R, located between subunits S1 and S2. It has been established that the P681R change increases S1/S2 cleavage, which induces a quicker fusion between virus and cell membrane, thus facilitating viral transmissibility (7).

Finally, the N-terminal domain (NTD) (an essential binding site for antibody 4A8 [8]) is altered through G142D. This mutation contributes to changing the structure of the NTD, resulting in the binding prevention of neutralization antibodies (NAbs) to this domain (9). The characteristic B.1.617.2 and B.1.617.3 deletions 157 to 158 can also be observed and constitute another mechanism by which the virus acquired vaccine escape abilities (10).

The emergence and spread of new variants derived from volatile organic compounds (VOCs) are a global concern, their escape from vaccination is not yet well studied, and the genomic surveillance of the genome of SARS-CoV-2 remains a primary approach to identify variants and prevent their spread.

Data availability. The SARS-CoV-2 genome sequences 5858 and 5900 were deposited into the GISAID database under the identifiers [EPI_ISL_6002899](https://gisaid.org/record/EPI_ISL_6002899) and [EPI_ISL_6002900](https://gisaid.org/record/EPI_ISL_6002900), respectively, and in NCBI GenBank under the accession numbers [OL375239](https://ncbi.nlm.nih.gov/nucl/OL375239) and [OL375240](https://ncbi.nlm.nih.gov/nucl/OL375240), respectively. Raw reads from next-generation sequencing (NGS) are available in the BioProject database under SRA number [PRJNA780922](https://bioinformatics.ncbi.nlm.nih.gov/sra/study/SRR13150847). The specific raw read library identifiers of samples 5858 and 5900 are [SRX13150847](https://bioinformatics.ncbi.nlm.nih.gov/sra/study/SRR13150847) and [SRX13150848](https://bioinformatics.ncbi.nlm.nih.gov/sra/study/SRR13150848), respectively.

ACKNOWLEDGMENT

This study was supported and financed by the Hassan II Academy of Science and Technology, Morocco.

REFERENCES

- Wang H, Li X, Li T, Zhang S, Wang L, Wu X, Liu J. 2020. The genetic sequence, origin, and diagnosis of SARS-CoV-2. *Eur J Clin Microbiol Infect Dis* 39:1629–1635. <https://doi.org/10.1007/s10096-020-03899-4>.
- Wang C, Liu Z, Chen Z, Huang X, Xu M, He T, Zhang Z. 2020. The establishment of reference sequence for SARS-CoV-2 and variation analysis. *J Med Virol* 92:667–674. <https://doi.org/10.1002/jmv.25762>.
- Pascarella S, Ciccozzi M, Zella D, Bianchi M, Benedetti F, Benvenuto D, Broccolo F, Cauda R, Caruso A, Angeletti S, Giovanetti M, Cassone A. 2021.

- SARS-CoV-2 B.1.617 Indian variants: are electrostatic potential changes responsible for a higher transmission rate? *J Med Virol* 93:6551–6556. <https://doi.org/10.1002/jmv.27210>.
- Hou YJ, Chiba S, Halfmann P, Ehre C, Kuroda M, Dinnon KH, Leist SR, Schäfer A, Nakajima N, Takahashi K, Lee RE, Mascenik TM, Graham R, Edwards CE, Tse LV, Okuda K, Markmann AJ, Bartelt L, de Silva A, Margolis DM, Boucher RC, Randell SH, Suzuki T, Gralinski LE, Kawaoka Y, Baric RS. 2020. SARS-CoV-2 D614G variant exhibits efficient replication ex vivo and transmission in vivo. *Science* 370:1464–1468. <https://doi.org/10.1126/science.abe8499>.
 - Peacock TP, Penrice-Randal R, Hiscox JA, Barclay WS. 2021. SARS-CoV-2 one year on: evidence for ongoing viral adaptation. *J Gen Virol* 102:001584. <https://doi.org/10.1099/jgv.0.001584>.
 - Venkata-Viswanadh V, Lai L, Sahoo MK, Floyd F, Sibai M, Solis D, Flowers MW, Hussaini L, Ciric CR, Bechnack S, Stephens K, Mokhtari EB, Mudvari P, Creanga A, Pegu A, Derrien-Colemyn A, Henry AR, Gagne M, Graham BS, Wrammert J, Douek DC, Boritz E, Pinsky BA, Suthar MS. 2021. Infection and vaccine-induced neutralizing antibody responses to the SARS-CoV-2 B.1.617.1 variant. *bioRxiv* 2021.05.09.443299.
 - Garcia-Beltran WF, Lam EC, St Denis K, Nitido AD, Garcia ZH, Hauser BM, Feldman J, Pavlovic MN, Gregory DJ, Poznansky MC, Sigal A, Schmidt AG, lafrate AJ, Naranbhai V, Balazs AB. 2021. Multiple SARS-CoV-2 variants escape neutralization by vaccine-induced humoral immunity. *Cell* 184:2372–2383.e9. <https://doi.org/10.1016/j.cell.2021.03.013>.
 - Chi X, Yan R, Zhang J, Zhang G, Zhang Y, Hao M, Zhang Z, Fan P, Dong Y, Yang Y, Chen Z, Guo Y, Zhang J, Li Y, Song X, Chen Y, Xia L, Fu L, Hou L, Xu J, Yu C, Li J, Zhou Q, Chen W. 2020. A neutralizing human antibody binds to the N-terminal domain of the Spike protein of SARS-CoV-2. *Science* 369:650–655. <https://doi.org/10.1126/science.abc6952>.
 - Shen L, Triche TJ, Bard JD, Biegel JA, Judkins AR, Gai X. 2021. Spike protein NTD mutation G142D in SARS-CoV-2 Delta VOC lineages is associated with frequent back mutations, increased viral loads, and immune evasion. *medRxiv* 2021.09.12.21263475.
 - Mishra T, Joshi G, Kumar A, Dalavi R, Pandey P, Shukla S, Mishra RK, Chande A. 2021. B.1.617.3 SARS CoV-2 spike E156G/Δ157-158 mutations contribute to reduced neutralization sensitivity and increased infectivity. *bioRxiv* 2021.10.04.463028.