scientific reports

Check for updates

OPEN

Emergence of novel combinations of SARS-CoV-2 spike receptor binding domain variants in Senegal

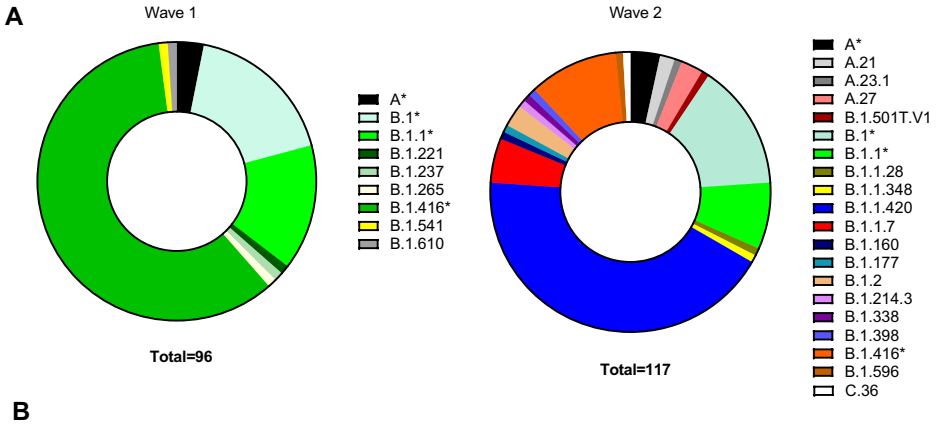
Ambroise D. Ahouidi[1,3]✉, Mary A. Rodgers[2,3]✉, Abdou Padane[1], Nafissatou Leye[1], Ana Olivo[2], Moustapha Mbow[1], Aminata Mboup[1], Papa Alassane Diaw[1], Aminata Dia[1], Barbara Harris[2], Yacine Amet Dia Padane[1], Gora Lo[1], Todd V. Meyer[2], Cyrille K. Diedhiou[1], Diabou Diagne[1], Ndeye Coumba Toure Kane[1], Gavin Cloherty[2] & Souleymane Mboup[1]

The emergence of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) lineages that carry mutations in the spike gene are of concern for potential impact to treatment and prevention efforts. To monitor for new SARS-CoV-2 mutations, a panel of specimens were sequenced from both wave one (N = 96), and wave two (N = 117) of the pandemic in Senegal by whole genome next generation sequencing. Amongst these genomes, new combinations of SARS-CoV-2 spike mutations were identified, with E484K + N501T, L452R + N501Y, and L452M + S477N exclusively found in second wave specimens. These sequences are evidence of local diversification over the course of the pandemic and parallel evolution of escape mutations in different lineages.

Ongoing viral evolution of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) threatens the efficacy of our strongest defenses against coronavirus disease 19 (COVID-19): vaccines, therapeutics, and diagnostics. To keep pace with continual viral diversification, molecular surveillance serves as a critical alert system for identifying new strains to evaluate for potential immune or diagnostic escape. Most recently, the identification of SARS-CoV-2 lineages of concern, B.1.1.7 (alpha), B.1.351 (beta), P.1 (gamma), and B.1.617.2 (delta), immediately preceded their rise in prevalence and global spread[1–4]. Subsequent reports have demonstrated that increased transmissibility and immune escape are linked to these lineages, which are defined by spike receptor binding domain (RBD) mutations, including N501Y, K417N/T, L452R, and E484K. Notably, the E484K and L452R mutations in RBD had previously been demonstrated to confer immune escape in cell culture selection experiments[5], which is consistent with their increasing prevalence[6,7], possibly due to increased viral fitness[8,9]. Therefore, vigilant monitoring of circulating strains for these mutations is of critical importance for potentially preventing their spread.

The SARS-CoV-2 pandemic in Senegal has surged in several waves occurring in March–November of 2020 (wave 1), December 2020–March 2021 (wave 2) and July–September 2021 (Wave 3). The first variant of concern that was reported in Senegal was B.1.1.7, which was first identified in a patient who was diagnosed on December 30th, 2020 during the second wave[10]. To compare the SARS-CoV-2 strains circulating during the first two waves of the pandemic in Senegal, a panel of 150 first wave and 150 s wave leftover nasopharyngeal specimens in viral transport media (VTM) were collected in a study approved by the Ethical Committee of the Ministry of Health of Senegal (000129/MSAS/CNERS). VTM specimens were sequenced by next generation sequencing (NGS) using a metagenomic approach with probe enrichment (xGen) and analysis on an Illumina HiSeq[11]. Genomes were assembled using BLAST and sequence NC_045512 as a reference, followed by clade assignment and mutation analysis with the NextClade tool (clades.nextstrain.org) and lineage assignments with the Pangolin tool[12]. Genome coverage of > 60% was achieved for N = 213 specimens (N = 96 first wave, N = 117 s wave), with an average coverage depth of 43,006x (GISAID accession numbers EPI_ISL_1630259-1630270). The first wave genomes fell into 3 clades: 19B (N = 3), 20A (N = 78), and 20B (N = 15), similar to the composition of strains in other countries around the same time period[13]. In Pangolin nomenclature[14], nine lineages were present in the first wave, which was predominated by B.1.416 (57/96, 59.4%, Fig. 1A). Viral diversity increased greatly in wave two with genomes from 9 clades present: 19A (N = 1), 19B (N = 11), 20A (N = 108), 20B (N = 81), 20C (N = 3), 20D (N = 1), 20E (N = 1), 20G (N = 1), and 20I (N = 1). Increased diversity of Pangolin lineages was also observed in the second wave, with 20 lineages identified, the majority of which were not present in the first wave (Fig. 1A).

[1]Institute for Health Research, Epidemiological Surveillance and Training (IRESSEF), Dakar, Senegal. [2]Abbott Global Surveillance Program, Abbott Laboratories, Abbott Park, IL, USA. [3]These authors contributed equally: Ambroise D. Ahouidi and Mary A. Rodgers. ✉email: ambroise.ahouidi@iressef.org; mary.rodgers@abbott.com



B

Specimen ID	*date of sampling*	*Collection location*	*% genome coverage*	*Pangolin*	*clade*	*Variant Classification*
SN-IR4-000972	1/2021	Dakar-Ngor	100.00%	A.21	19B	A.21.L452R
SN-IR1-0013954	12/2020	Dakar-Ngor	99.87%	A.27	19B	A.27.501Y.V4
SN-IR1-0014258	12/2020	Dakar-Ngor	99.83%	A.27	19B	A.27.501Y.V4
SN-IR1-0014313	12/2020	Dakar-Ngor	99.81%	A.27	19B	A.27.501Y.V4
SN-IR2-008858	12/2020	Diamniadio	100.00%	B.1	20C	B.1.501T.V1
SN-IR1-0016423	1/2021	Dakar-Ngor	99.85%	B.1.1.7	20I/501Y.V1	B.1.1.7
SN-IR8-000056	1/2021	Dakar	99.84%	B.1.1.7	20I/501Y.V1	B.1.1.7
SN-TH-TIV-2174	1/2021	Tivaoune	99.83%	B.1.1.7	20I/501Y.V1	B.1.1.7
SN-IR2-0010667	1/2021	Diamniadio	99.83%	B.1.1.7	20I/501Y.V1	B.1.1.7
SN-THI-21-027	1/2021	Thies	99.79%	B.1.1.7	20I/501Y.V1	B.1.1.7
SN-THI-20-2088	12/2020	Thies	99.72%	B.1.1.7	20I/501Y.V1	B.1.1.7

C

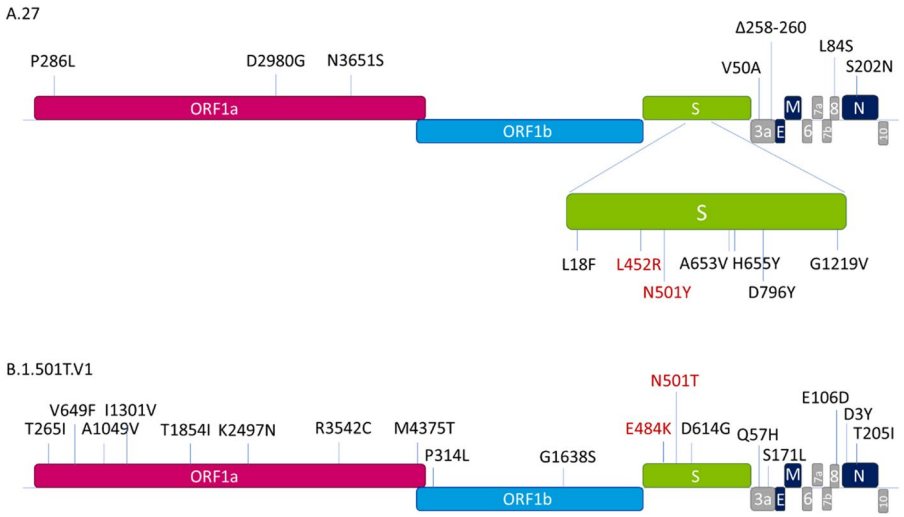


Figure 1. Molecular surveillance of SARS-CoV-2 in Senegal. In panel (**A**), the number of sequences classified in the indicated lineages present in waves one and two are shown proportionally to the total number of sequences generated with > 60% genome coverage from each wave as designated in the total numbers below each plot. An * indicates lineages that were present in both waves. In panel (**B**), the sequence and specimen metadata are listed for each strain carrying spike mutation of concern at position 501, 484, or 452. All sequences in this panel are from wave two. In panel (**C**), the lineage defining amino acid mutations (in comparison to the reference genome NC_045512) for the new strains identified in this study are shown. (**A**) Lineages identified, (**B**) Variant specimen summary, (**C**) Escape variant lineages.

Most notable amongst the new strains found exclusively in wave two, the B.1.1.7 variant accounted for 5% of all second wave infections (6/117) and was present in four different cities (Dakar, Tivaoune, Diamnadio, and Thies, Fig. 1B), confirming a widespread distribution in western Senegal. The earliest B.1.1.7 infection in this study was diagnosed on December 21st in Thies, which predates the first case previously identified Senegal[10]. The December 21st patient was a patient who was tested due to contact with an infected person, suggesting that B.1.1.7 was already circulating in Senegal in early December. The remaining 5 B.1.1.7 cases were all diagnosed in early January during the exponential phase of the second wave spike in cases.

Escape mutations in the spike protein were absent from wave one but were present in 4% (5/117) of all wave two infections (Fig. 1B, Supplemental Table 1). Additional details for all genomes with mutations of concern at position 501, 484, and/or 452 in the spike RBD are shown in Fig. 1B. When classified by clade, all of the L452R mutations were exclusively found in 19B clade genomes whereas the L452M mutation appears to have emerged in wave two in the 20A clade (Supplemental Table 1). In addition to strains carrying L452R individually, variant strains carrying a combination of L452R + N501Y (3/117, 2.6%) were also identified. The N501Y mutation confers higher affinity for the ACE2 receptor and is present in several variants of concern (alpha, beta, gamma) while L452R is a signature escape mutation found in the delta and epsilon lineages that also increases infectivity[4,8,9,15,16]. The combination of both of these mutations in one strain is of concern for potential rapid spread of an immune escape variant. All three of the genomes carrying the L452R/N501Y combination belonged to the A.27 lineage (clade 19B) and did not encode the D614G mutation that predominates most global infections today. Likewise, the other lineage defining mutations for variants of concern were absent in the A.27 genomes, with the exception of L18F and H655Y, which are both present in the gamma lineage (Fig. 1C). While 13 common single nucleotide polymorphisms (SNPs) were identified for this lineage, each individual genome had unique SNPs as well, suggesting they were not transmission linked cases. The three patients who had A.27 infections were diagnosed in the Almadie district of Dakar in December 2020 and ranged in age from 36 to 55 (Fig. 1B).

In addition to the L452R + N501Y double mutant, a single genome was identified that carried a unique combination of E484K + N501T spike RBD mutations in a B.1 lineage genome (clade 20C) with D614G also present. This lineage has been provisionally named B.1.501T.V1 (Fig. 1C). The patient who was infected with this variant strain was a patient who was diagnosed in December 2020 in Diamniadio (Fig. 1B). While E484K confers escape from neutralizing antibodies[17,18], the N501T mutation enhances the spike receptor binding domain (RBD) affinity for ACE2 in vitro and is predicted to enhance transmissibility, similar to N501Y[19,20]. Strains harboring N501T first emerged in August of 2020 in Northern Italy[6] and the N501T mutation has been found recently in an emerging Brazilian lineage that differs from B.1.501T.V1[21]. Alarmingly, N = 2122 N501T strains were posted to GISAID from specimens collected in the months that followed the identification of this specimen in Senegal (January–April 2021) from countries in Africa, Europe, Asia, North America, and South America (GISAID, date of accession April 18th, 2021)[6]. Altogether, these trends suggest that convergent evolution around the world is leading to mutations at spike positions E484 and N501 in many lineages, suggesting a possible increased fitness for viruses carrying these mutations.

Received: 27 May 2021; Accepted: 16 November 2021
Published online: 08 December 2021

References

1. Volz, E. *et al.* The COVID-19 Genomics UK (COG-UK) Consortium, Flaxman, S., Ratmann, O., Bhatt, S., Hopkins, S., Gandy, A., Rambaut, A. & Ferguson, N. M. Transmission of SARS-CoV-2 Lineage B.1.1.7 in England: Insights from linking epidemiological and genetic data. *Nature* **593**(7858), 266–269. https://doi.org/10.1038/s41586-021-03470-x (2021).
2. Tegally, H., Wilkinson, E., Giovanetti, M. *et al.* Detection of a SARS-CoV-2 variant of concern in South Africa. *Nature* **592**(7854), 438-443.https://doi.org/10.1038/s41586-021-03402-9 (2021).
3. Faria, N. R., Mellan, T. A., Whittaker, C. *et al.* Genomics and epidemiology of a novel SARS-CoV-2 lineage in Manaus, Brazil. *medRxiv* (2021).
4. Cherian, S. *et al.* SARS-CoV-2 spike mutations, L452R, T478K, E484Q and P681R, in the second wave of COVID-19 in Maharashtra, India. *Microorganisms* **9**(7), 1542 (2021).
5. Greaney, A. J. *et al.* Complete mapping of mutations to the SARS-CoV-2 spike receptor-binding domain that escape antibody recognition. *Cell Host Microbe* **29**(1), 44–57 (2021).
6. Elbe, S. & Buckland-Merrett, G. Data, disease and diplomacy: GISAID's innovative contribution to global health. *Glob Chall.* **1**(1), 33–46 (2017).
7. Hodcroft, E. CoVariants: SARS-CoV-2 mutations and variants of interest. https://covariants.org/ (2021).
8. Motozono, C. *et al.* SARS-CoV-2 spike L452R variant evades cellular immunity and increases infectivity. *Cell Host Microbe* **29**(7), 1124–1136 (2021).
9. Vogel, M., Augusto, G., Chang, X. *et al.* Molecular definition of severe acute respiratory syndrome coronavirus 2 receptor-binding domain mutations: Receptor affinity versus neutralization of receptor interaction. *Allergy* https://doi.org/10.1111/all.15002 (2021).
10. Padane, A. *et al.* First detection of the British variant of SARS-CoV-2 in Senegal. *New Microbes New Infect.* **41**, 100877 (2021).
11. Forberg, K. O. G., Meyer, T. V., Mowerman, I., Mohaimani, A., Faron, M., Jennings, C., Landay, A. L., Goldstein, Y., Fox, A., Berg, M. G., Cloherty, G. A. SNP and phylogenetic characterization of low viral load SARS-CoV-2 specimens by target enrichment. *Front. Virol.* (Under Review) (2021).
12. O'Toole, Á., Scher, E., Underwood, A., Jackson, B., Hill, V., McCrone, J. T., Ruis, C., Abu-Dahab, K., Taylor, B., Yeats, C., du Plessis, L., Aanensen, D., Holmes, E., Pybus, O. & Rambaut, A. Pangolin: Lineage assignment in an emerging pandemic as an epidemiological tool. githubcom/cov-lineages/pangolin (2021).
13. Hadfield, J. *et al.* Nextstrain: Real-time tracking of pathogen evolution. *Bioinformatics* **34**(23), 4121–4123 (2018).
14. Rambaut, A. *et al.* A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. *Nat. Microbiol.* **5**(11), 1403–1407 (2020).
15. Kirby, T. New variant of SARS-CoV-2 in UK causes surge of COVID-19. *Lancet Respir. Med.* **9**, e20–e21 (2021).
16. Deng, X., Garcia-Knight, M. A., Khalid, M. M., *et al.* Transmission, infectivity, and neutralization of a spike L452R SARS-CoV-2 variant. *Cell* **184**(13), 3426–3437.e8. https://doi.org/10.1016/j.cell.2021.04.025 (2021).

17. Cele, S., Gazy, I., Jackson, L. *et al.* Escape of SARS-CoV-2 501Y.V2 from neutralization by convalescent plasma. *Nature* **593**(7857), 142–146. https://doi.org/10.1038/s41586-021-03471-w (2021).
18. Dejnirattisai, W., Zhou, D., Supasa, P. *et al.* Antibody evasion by the P.1 strain of SARS-CoV-2. *Cell* **184**(11), 2939–2954.e9. https://doi.org/10.1016/j.cell.2021.03.055 (2021).
19. Wan, Y., Shang, J., Graham, R., Baric, R. S. & Li, F. Receptor recognition by the novel coronavirus from Wuhan: An analysis based on decade-long structural studies of SARS coronavirus. *J. Virol.* **94**(7), e00127-20 (2020).
20. Starr, T. N. *et al.* Deep mutational scanning of SARS-CoV-2 receptor binding domain reveals constraints on folding and ACE2 binding. *Cell* **182**(5), 1295–1310 (2020).
21. Moreira, F. R. R., Bonfim, D. M., Geddes, V. E. V., Zauli, D. A. G., do Prado Silva, J., de Lima, A. B., Malta, F. S. V., de Souza Ferreira, A. C., Pardini, V. C., Queiroz, D. C., de Souza, R. M., de Araújo, J. L. F., Alves, H. J., Silva, A. V. F. G., Resende, G. G., de Menezes, A. L., de Oliveira, E. S., de Oliveira, J. S., Teixeira, M. M., Luiz, L. M., Gomez, R. S., Fonseca, P. L. C., Moreira, R. G., Tanuri, A., de Souza, W. M., Faria, N. R., Voloch, C. M., de Souza, R. P. & Aguiar, R. S. Increasing frequency of SARS-CoV-2 lineages B.1.1.7, P.1 and P.2 and identification of a novel lineage harboring E484Q and N501T spike mutations in Minas Gerais, Southeast Brazil. Virologicalorg (2021).

Acknowledgements

We gratefully acknowledge the authors, originating and submitting laboratories of the genetic sequence and metadata made available through GISAID. This Project was funded by Abbott.

Author contributions

Conceptualization, A.D.A., M.A.R., G.C., and S.M.; methodology, T.V.M.; investigation, A.O.; D.D.; C.K.D.; Y.A.D.P.; P.A.D.; and A.D.; resources, G.C.; data curation, B.H., A.O., A.D.A. and M.A.R.; writing—original draft preparation, M.A.R.; A.D.A.; writing—review and editing, A.D.A., M.A.R., B.H., G.C., A.M.; A.P.; N.L.; N.C.T.K.; M.M.; and G.L.; visualization, M.A.R..; supervision, G.C. and S.M.; All authors have read and agreed to the published version of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-021-02874-z.

Correspondence and requests for materials should be addressed to A.D.A. or M.A.R.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.