Emergence of Novel Combinations of SARS-CoV-2 Spike Receptor Binding Domain Variants in Senegal

Ambroise D Ahouidi (ambroise.ahouidi@iressef.org)
Epidemiological Surveillance and Training (IRESSEF)

Mary A Rodgers
Abbott Laboratories

Abdou Padane
Epidemiological Surveillance and Training (IRESSEF)

Nafissatou Leye
Epidemiological Surveillance and Training (IRESSEF)

Ana Olivo
Abbott Laboratories

Moustapha Mbow
Epidemiological Surveillance and Training (IRESSEF)

Aminata Mboup
Epidemiological Surveillance and Training (IRESSEF)

Papa Alassane Diaw
Epidemiological Surveillance and Training (IRESSEF)

Aminata Dia
Epidemiological Surveillance and Training (IRESSEF)

Barbara Harris
Abbott Laboratories

Yacine Amet Dia Padane
Epidemiological Surveillance and Training (IRESSEF)

Gora Lo
Epidemiological Surveillance and Training (IRESSEF)

Todd V Meyer
Abbott Laboratories

Cyrille K. Diedhiou
Epidemiological Surveillance and Training (IRESSEF)

Diabou Diagne
Epidemiological Surveillance and Training (IRESSEF)

Ndeye Coumba Toure Kane
Epidemiological Surveillance and Training (IRESSEF)

Gavin Cloherty

Abbott Laboratories

Souleymane Mboup

Epidemiological Surveillance and Training (IRESSEF)

Research Article

Keywords: SARS-CoV-2, COVID-19, whole genome sequencing, spike escape variant

Posted Date: June 14th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-567887/v1

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

Ambroise D Ahouidi[*1], Mary A Rodgers[*2], 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]

1 Institute for Health Research, Epidemiological Surveillance and Training (IRESSEF), Dakar, Senegal

2 Abbott Global Surveillance Program, Abbott Laboratories, Abbott Park, IL, USA

* Equal contribution and Correspondence: (ADA) ambroise.ahouidi@iressef.org; (MAR) mary.rodgers@abbott.com

Abstract: 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.

Keywords: SARS-CoV-2, COVID-19, whole genome sequencing, spike escape variant

Introduction

Ongoing viral evolution of 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, B.1.351, P.1, B.1.427, and B.1.429, immediately preceded their rise in prevalence and global spread [1-3]. 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, E417K/N, L452R, and E484K. Notably, the E484K, L452R, and S477N mutations in RBD had previously been demonstrated to confer immune escape in cell culture selection experiments [4], which is consistent with their increasing prevalence [5, 6], possibly due to increased viral fitness. 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 two waves occurring in March-November of 2020 (wave 1) and December 2020-March 2021 (wave 2). 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 [7]. To date, other variants of concern have not been reported in Senegal and the second wave has waned.

Results and Discussion

Genome coverage of >60% was achieved for N=213 specimens (N=96 first wave, N=117 second wave), with an average coverage depth of 43,006x (GISAID accession numbers EPI_ISL_1630259-1630270, 1633465-7, 1827859-950). 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 [10]. In Pangolin nomenclature [11], nine lineages were present in the first wave, which was predominated by B.1.416 (57/96, 59.4%, Figure 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 (Figure 1a). 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, Table 1), confirming a widespread distribution in western Senegal. The earliest B.1.1.7 infection in this study was diagnosed in mid-December in Thies, which predates the first case previously identified Senegal [7]. This patient was an individual 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 12% (14/117) of all wave two infections (Table 1). Several lineages carried either L452R or S477N, suggesting that these mutations likely arose independently in each lineage. When classified by clade, all of the L452R mutations were exclusively found in 19B clade genomes, whereas all S477N mutations were present in 20A sequences. 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 is present in several variants of concern and has been suggested to confer enhanced transmissibility [12]. The combination of both an escape mutation (L452R) and a mutation causing increased transmissibility (N501Y) is of concern for potential rapid spread of an immune escape variant. All three of the double mutants were present in sequences belonging to the A.27 lineage (clade 19B) and did not encode the D614G mutation that predominates most global infections today. We have provisionally named this lineage as A.27.N501Y.V4 (Table 1). 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.N501Y.V4 infections were diagnosed in the Almadie district of Dakar in mid-December, 2020 (Table 1).

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 (Figure 1b). The patient who was infected with this variant strain was diagnosed in December, 2020 in Diamniadio (Table 1). While E484K confers escape from neutralizing antibodies [13, 14], the N501T mutation enhances the spike receptor binding domain (RBD) affinity for ACE2 *in vitro* and is predicted to enhance transmissibility, similar to N501Y [15].

Strains harboring N501T first emerged in August of 2020 in Northern Italy [5] and the N501T mutation has been found recently in an emerging Brazilian lineage that differs from B.1.501T.V1 [16]. Alarmingly, N=2122 N501T strains were posted to GISAID from specimens collected in January-April 2021 from countries in Africa, Europe, Asia, North America, and South America (GISAID, date of accession April 18th, 2021) [5]. 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.

Materials and Methods

To compare the SARS-CoV-2 strains circulating during both waves of the pandemic in Senegal, a panel of 150 first wave and 150 second 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 [8]. 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 [9].

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.

Funding: This research was funded by Abbott Laboratories.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Ethical Committee of the Ministry of Health of Senegal (000129/MSAS/CNERS, date of approval July 2020).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Sequences generated in this study are available on the GISAID website under accession numbers: EPI_ISL_1630259-1630270, 1633465-7, 1827859-950.

Acknowledgments: Data, including metadata and sequence information was obtained from the GISAID database as a result of the work of data contributors, i.e., the Authors, the Originating laboratories responsible for obtaining the specimens, and the Submitting laboratories for generating the genetic sequence and metadata and sharing via the GISAID Initiative, on which this research is based. GISAID data are subject to GISAID's Terms and Conditions that can be accessed via www.gisaid.org. We thank Dr Gregory Orf (Abbott Laboratories) for NextClade support and Kenn Forberg and Dr Michael G Berg (Abbott Laboratories) for development of the SARS-CoV-2 xGen sequencing method. We thank Abdoul Aziz Diallo and Mame Khardiata Ndiaye (IRESSEF) for helping in samples collections.

Conflicts of Interest: MAR, AO, BH, TVM, and GC are employees of Abbott Laboratories.

References

1. Erik Volz; Swapnil Mishra MC, Jeffrey C. Barrett, Robert Johnson, Lily Geidelberg, Wes R Hinsley, Daniel J Laydon, Gavin Dabrera, Áine O'Toole, Roberto Amato, Manon Ragonnet-Cronin, Ian Harrison, Ben Jackson, Cristina V. Ariani, Olivia Boyd, Nicholas J Loman, John T McCrone, Sónia Gonçalves, David Jorgensen, Richard Myers, Verity Hill, David K. Jackson, Katy Gaythorpe, Natalie Groves, John Sillitoe, Dominic P. Kwiatkowski, The COVID-19 Genomics UK (COG-UK) consortium, Seth Flaxman, Oliver Ratmann, Samir Bhatt, Susan Hopkins, Axel Gandy, Andrew Rambaut, Neil M Ferguson. Transmission of SARS-CoV-2 Lineage B.1.1.7 in England: Insights from linking epidemiological and genetic data. MedRxiv 2021.

2. Tegally H, Wilkinson E, Giovanetti M, et al. Emergence of a SARS-CoV-2 variant of concern with mutations in spike glycoprotein. Nature 2021.

3. Faria NR, Mellan TA, Whittaker C, et al. Genomics and epidemiology of a novel SARS-CoV-2 lineage in Manaus, Brazil. medRxiv 2021.

4. Greaney AJ, Starr TN, Gilchuk P, et al. Complete Mapping of Mutations to the SARS-CoV-2 Spike Receptor-Binding Domain that Escape Antibody Recognition. Cell Host Microbe 2021; 29(1): 44-57 e9.

5. Elbe S, Buckland-Merrett G. Data, disease and diplomacy: GISAID's innovative contribution to global health. Glob Chall 2017; 1(1): 33-46.

6. Hodcroft E. "CoVariants: SARS-CoV-2 Mutations and Variants of Interest." https://covariantsorg/ 2021.

7. Padane A, Kanteh A, Leye N, et al. First detection of the British variant of SARS-CoV-2 in Senegal. New Microbes New Infect 2021: 100877.

8. Kenn Forberg GO, Todd V. Meyer, Illya Mowerman, Aurash Mohaimani, Matthew Faron, Cheryl Jennings, Alan L. Landay, Yitz Goldstein, Amy Fox, Michael G. Berg, Gavin A. Cloherty. SARS-CoV-2 xGen target enrichment sequencing detects the early emergence of spike Q677 mutations. Frontiers in microbiology 2021; (Submitted).

9. Áine O'Toole ES, Anthony Underwood, Ben Jackson, Verity Hill, JT McCrone, Chris Ruis, Khali Abu-Dahab, Ben Taylor, Corin Yeats, Louis du Plessis, David Aanensen, Eddie Holmes, Oliver Pybus, Andrew Rambaut. pangolin: lineage assignment in an emerging pandemic as an epidemiological tool. githubcom/cov-lineages/pangolin](githubcom/cov-lineages/pangolin 2021.

10. Hadfield J, Megill C, Bell SM, et al. Nextstrain: real-time tracking of pathogen evolution. Bioinformatics 2018; 34(23): 4121-3.

11. Rambaut A, Holmes EC, O'Toole A, et al. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. Nat Microbiol 2020; 5(11): 1403-7.

12. Kirby T. New variant of SARS-CoV-2 in UK causes surge of COVID-19. Lancet Respir Med 2021; 9.

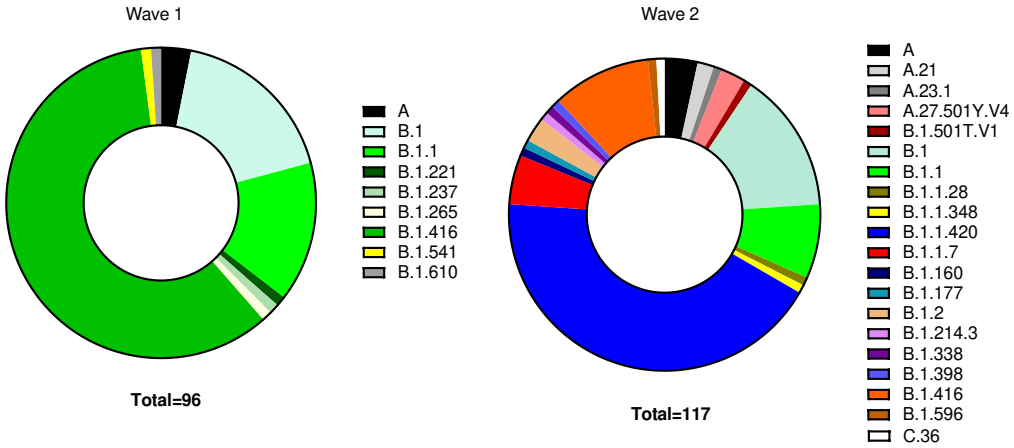
13. Cele S, Gazy I, Jackson L, et al. Escape of SARS-CoV-2 501Y.V2 from neutralization by convalescent plasma. Nature 2021.

14. Dejnirattisai W, Zhou D, Supasa P, et al. Antibody evasion by the P.1 strain of SARS-CoV-2. Cell 2021.

15. Wan Y, Shang J, Graham R, Baric RS, Li F. Receptor Recognition by the Novel Coronavirus from Wuhan: an Analysis Based on Decade-Long Structural Studies of SARS Coronavirus. Journal of virology 2020; 94(7).

16. Filipe Romero Rebello Moreira DMB, Victor Emmanuel Viana Geddes, Danielle Alves Gomes Zauli, Joice do Prado Silva, Aline Brito de Lima, Frederico Scott Varella Malta, Alessandro Clayton de Souza Ferreira, Victor Cavalcanti Pardini, Daniel Costa Queiroz, Rafael Marques de Souza, João Locke Ferreira de Araújo, Hugo José Alves, Ana Valesca Fernandes Gilson Silva, Gustavo Gomes Resende, André Luiz de Menezes, Eneida Santos de Oliveira, Jaqueline Silva de Oliveira, Mauro Martins Teixeira, Lucyene Miguita Luiz, Ricardo Santiago Gomez, Paula Luize Camargos Fonseca, Rennan Garcias Moreira, Amilcar Tanuri, William Marciel de Souza, Nuno Rodrigues Faria, Carolina Moreira Voloch, Renan Pedra de Souza, Renato Santana Aguiar. 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. Virological.org 2021.

(**a**)



(**b**)

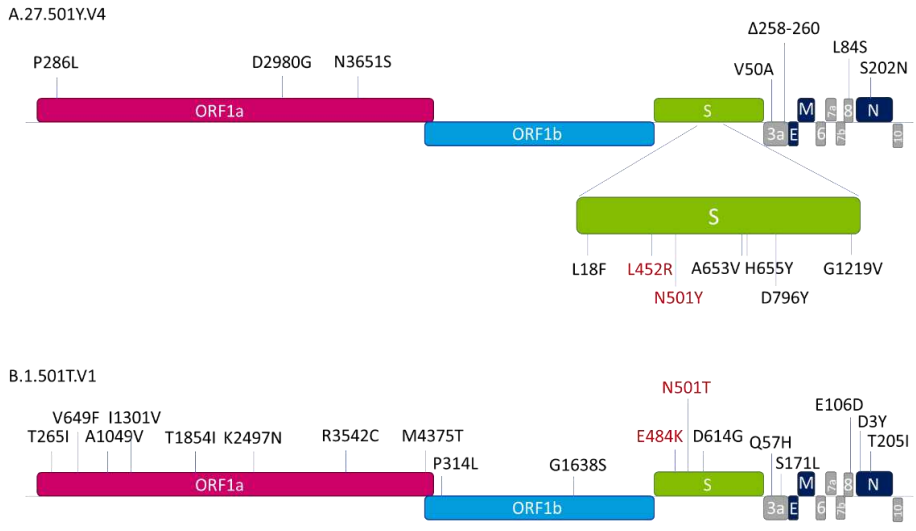


Figure 1. Molecular surveillance of SARS-CoV-2 in Senegal. In panel a, the number of sequences classified in the indicated lineages present in waves one and two are shown proportionally to the total number of sequences generated with >60% genome coverage from each wave as designated in the total numbers below each plot. In panel b, the lineage defining amino acid mutations (in comparison to the reference genome NC_045512) for the new strains identified in this study are shown.

236

237

238 **Table 1.**

Specimen ID	Collection Location	% genome coverage	Lineage	Clade	Variant Classification
SN-IR4-000972	Dakar-Ngor	100.00%	A.21	19B	A.21.L452R
SN-IR1-0013954	Dakar-Ngor	99.87%	A.27	19B	A.27.501Y.V4
SN-IR1-0014258	Dakar-Ngor	99.83%	A.27	19B	A.27.501Y.V4
SN-IR1-0014313	Dakar-Ngor	99.81%	A.27	19B	A.27.501Y.V4
SN-IR1-0018256	Dakar-Ngor	99.86%	B.1.160	20A.EU2	B.1.160.S477N
SN-MBO-20-1752	Mbour	78.71%	B.1.338	20A	B.1.338.S477N
SN-DKi12-002	Dakar	80.60%	B.1	20A	B.1.S477N
SN-THI-21-024	Thies	99.99%	B.1.416	20A	B.1.416.S477N
SN-THI-21-036	Thies	99.91%	B.1.416	20A	B.1.416.S477N
SN-CHRT-0500	Thies	99.90%	B.1.416	20A	B.1.416.S477N
SN-MBO-20-1743	Mbour	99.90%	B.1.416	20A	B.1.416.S477N
SN-THI-MEK-0547	Meckhe	99.75%	B.1.416	20A	B.1.416.S477N
SN-TH-TIV-2176	Tivaoune	99.61%	B.1.416	20A	B.1.416.S477N
SN-IR2-008858	Diamniadio	100.00%	B.1	20C	B.1.501T.V1
SN-IR1-0016423	Dakar-Ngor	99.85%	B.1.1.7	20I/501Y.V1	B.1.1.7
SN-IR8-000056	Dakar	99.84%	B.1.1.7	20I/501Y.V1	B.1.1.7
SN-TH-TIV-2174	Tivaoune	99.83%	B.1.1.7	20I/501Y.V1	B.1.1.7
SN-IR2-0010667	Diamniadio	99.83%	B.1.1.7	20I/501Y.V1	B.1.1.7
SN-THI-21-027	Thies	99.79%	B.1.1.7	20I/501Y.V1	B.1.1.7
SN-THI-20-2088	Thies	99.72%	B.1.1.7	20I/501Y.V1	B.1.1.7

239

Figures

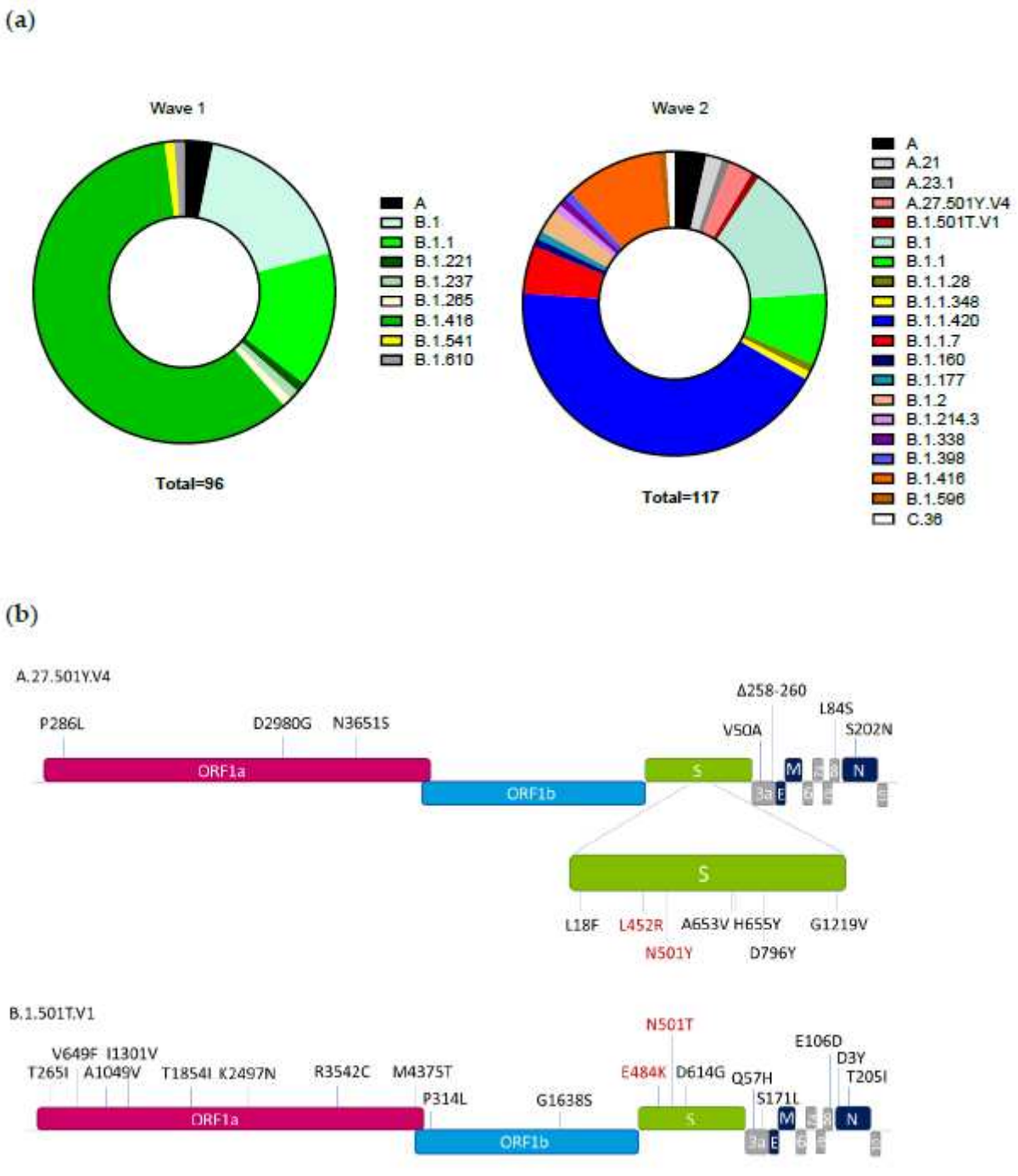


Figure 1

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. In panel b, the lineage defining amino acid mutations (in comparison to the reference genome NC_045512) for the new strains identified in this study are shown.