Article type: Research article

Full-length title:

Genome sequence analysis enabled deciphering the atypical evolution of COVID-19 epidemics in Marseille, France

Short title: Evolution of SARS-CoV-2 genotypes

Author list: Pierre-Edouard FOURNIER[1,2], Philippe COLSON[1,3], Anthony LEVASSEUR[1,3], Philippe GAUTRET[1,2], Marielle BEDOTTO[1,3], Véronique FILOSA[1], Jeremy DELERCE[1,3], Ludivine BRECHARD[1,2], Jean-Christophe LAGIER[1,3], Hervé CHAUDET[1,4], Florence FENOLLAR[1,2], Didier RAOULT[1,3]*

Affiliations: [1] IHU Méditerranée Infection; [2] Vecteurs - Infections Tropicales et Méditerranéennes (VITROME), Marseille, France ; [3] Aix-Marseille Univ, Microbes Evolution Phylogeny and Infections (MEPHI); [4] French Armed Forces Center for Epidemiology and Public Health (CESPA).

[¥] These authors contributed equally to this article

*** Address for correspondence:** 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

Word count: abstract, 222; text, 1,511

Figures: 4; **Table:** 1; **References:** 10

Abstract

Background. In Marseille, France, the COVID-19 outbreak exhibited a highly unusual kinetic for a viral respiratory disease, with the occurrence of two successive epidemic outbreaks.

Methods. By sequencing the genomes from 1038 SARS-CoV-2 strains from March through November 2020 and comparing them to those available in international databases, we confirmed that these outbreaks were caused by three main distinct variants intertwined with another 8 minor variants.

Results. The initial outbreak, from March to April 2020 was caused by the 20A variant. In the second outbreak, a first phase from July to August involved a milder variant that we named Marseille-1 and was possibly imported from sub-Saharan Africa through Maghreb. It was followed by a second, initially overlapping phase, from August 2020, in which the Marseille-4 variant, possibly originating from Northern Europe, and causing more severe infections, played the major role. Viral genotypes in the second outbreak exhibited a dramatically higher genetic diversity (mean ± standard deviation for genetic distance: $7.6 \times 10^{-4} \pm 3.8 \times 10^{-4}$ vs. $2.3 \times 10^{-4} \pm 1.1 \times 10^{-4}$, respectively; $p < 2.2 \times 10^{-16}$). Mutations observed in these 10 new viral variants are located in most SARS-CoV-2 genes including structural and non-structural genes.

Conclusions. Our results demonstrate that SARS-CoV-2 is able to genetically evolve rapidly into distinct variants that may cause new outbreaks, which questions the efficiency of monovalent vaccines.

Keywords**:** SARS-CoV-2; Covid-19; genome; variant; mutations; deletions; diversity

Introduction

The SARS CoV-2 outbreak that started in Wuhan, China, in December 2019, has rapidly spread around the world (https://coronavirus.jhu.edu/map.html). In Europe, the COVID-19 epidemic exhibited a very unusual profile for a respiratory viral disease, particularly in France. Initially, it exhibited, as does the vast majority of viral respiratory infection outbreaks, a banal bell shape (herein referred to as episode 1, which ran from February to May 2020, and practically disappeared in June). Surprisingly, at the end of June (episode 2), following the reopening of international borders, new cases appeared whose origin was clearly ferry boat travelers and sailors from North Africa. The number of cases rapidly increased but the disease developed a very atypical epidemic curve which led us to assume that the successive waves could be linked to distinct viral variants. Since the beginning of the epidemic, we have diagnosed 6,855 patients during episode 1 and 18,823 during episode 2. In an attempt to decipher this conundrum, we have sequenced the complete genomes from 1038 SARS-CoV-2 strains from patients during these two episodes.

Succession of SARS-CoV-2 variants

At Mediterranee Infection institute (IHU) in Marseille, France, we have carried out SARS-CoV-2 PCR since the end of January 2020 [1, 2], and have been monitoring the daily number of tests and positive cases since the first patient was diagnosed on 02/27/2020 [3, 4] (https://www.mediterranee-infection.com/covid-19/) (**Figure 1**). As of 22/12/2020, we tested 410,456 specimens for 293,683 individuals, of whom 25,678 were positive. Next-generation sequencing (NGS) of SARS-CoV-2 genomes was carried out using both Illumina (San Diego, CA, USA) and Oxford Nanopore (Oxford, UK) technologies, as previously described [3]. The mutational landscape of 1038 coronavirus genomes was tracked in terms of SNPs and insertions/deletions (**Table 1**). These included 487 genomes from episode 1 and 551 from episode 2. Means of 8 mutations (+/- 1.9 SD) and 16.3 mutations (+/- 6.8 SD) were reported per viral isolate for the first and second episodes, respectively. Among mutational events, we detected 78 non synonymous mutations in the spike protein including the previously reported D614G (**Table 1**). In addition, 49 and 22 nonsynonymous mutations targeted the nucleocapsid N

protein and ORF8, respectively. Time-scaled phylogeny based on the 1038 full sequences enabled to differentiate at least ten clusters, assigned as variants Marseille-1 to Marseille-10 (**Figure 2**). The Marseille-4 cluster represented the majority of the identified variants, with 589 strains, followed by variants 1, 2 and 5 (114, 91 and 66 isolates, respectively). The most distant genomes from the Wuhan-1 isolate were variants Marseille-5 and -4, with 29 and 21 mutations, respectively. The Marseille-2 and -4 variants correspond to the Nexstrain clades 20A.EU1 and 20A.EU2, respectively (https://clades.nextstrain.org/) and the Marseille-1, 3, 5, 6, 7, 8, 9 and 10 variants are classified as Pangolin lineages B.1.5.12, B.1.5, B.1., B.1., B.1.7, B.1.1.10, B.1.1.55 and B.1.78, respectively.

The analysis of the genetic distance between the sequences from the first and second episodes showed a large difference (mean ± SD for genetic distance: $2.3 \times 10^{-4} \pm 1.1 \times 10^{-4}$ vs $7.6 \times 10^{-4} \pm 3.8 \times 10^{-4}$, respectively; $p < 2.2 \times 10^{-16}$). This indicates that the mean viral diversity expanded by 3.3 times between the two outbreaks. In addition, using Rao's quadratic entropy, the diversity between variants from both episodes was 1.7 times higher during episode 2 (1542.2) than that in episode 1 (911.6).

Viral epidemics

During episode 2, genome sequence analysis enabled us to identify epidemics caused by 10 distinct SARS-CoV-2 variants that occurred successively or overlapped between July and November, and were different from those in episode 1 (**Figure 3**). The emergence of these 10 variants, which were detected as early as weeks 25-33, followed the reopening of the French borders to tourists on June 15th. The episode 2 formed a jagged curve with two peaks during weeks 36 and 42 (**Figure 1**). However, the comparisons of the 10 variants to sequences available in GISAID (https://www.gisaid.org/) suggested that each of them followed its own agenda. Variants Marseille-1, -3, -4, -6, and -7 belong to GISAID clade GH and Nextstrain clade 20A, whereas variants -2 and -5 belong to Nextstrain clades 20D and 20C, respectively. Sequences with hallmark mutations from variants Marseille-2, -3, -4, -5 and -6 were present but considered as emerging sequences in the GISAID database.

Emergence of the Marseille-1 variant

The Marseille-1 variant emerged on week 27 and predominated in July before diseappearing at the end of August. Its onset coincided with the resumption on July 1st of maritime connections between Marseille and Maghreb. As 16 of the first 20 diagnosed cases (80%) were detected in boat travelers as well as sailors, we assume that this variant was brought to Marseille from Maghreb.

Emergence of the Marseille-4 variant

In weeks 29 and 30, five other variants, Marseille-2 to -5, emerged (**Figure 3**). Of those, Marseille-4 was the main variant from week 31, and is still present in France to date. Sequences from the Marseille-4 variant clustered with sequences in GISAID originating from Northern Europe, notably Denmark, UK, Switzerland, Belgium and the Netherlands. Given the prevalence of early positive samples in Denmark and the phylogenetic distance between this variant and other SARS-CoV-2 genotypes, we believe that it may have been transferred only recently to humans in this country. Since June 2020, Denmark has faced a SARS-CoV-2 epidemic in farmed minks [5]. Minks, and a wide array of animals including, among others, ferrets, dogs, cats, and primates, exhibit ACE2 proteins whose structure is highly similar to that of humans and are thus susceptible to SARS-CoV-2 [6]. As Denmark is the most important mink producer in Europe, with 17 million animals, it is possible that propagation of the virus in such a large animal population facilitated its genetic diversification. Then, as previously demonstrated in other intensive animal breeding countries [7], the virus was transferred from minks to humans [5]. Thus, we assume that the Marseille-4 variant was subsequently spread to the rest of Europe, notably France, by tourists from Northern Europe (Figure 4).

Other SARS-CoV-2 variants

The Marseille-2 and -5 variants were the second and third most prevalent variants from week 31. The best matches for Marseille-2 genomes in GISAID are from UK (90% (England, 51%)) and Spain. The Marseille-2 variant corresponds to the recently described 20A.EU1 cluster whose most likely source was described as Spain, in which it represents two thirds of released genomes. Then, this variant spread through Europe during the summer [8]. Together with Marseille-4, the Marseille-2

variant accounted for >80% of the genomes since week 40, Marseille-4 being 3.0-5.6 times more prevalent than Marseille-2. Homologs to the Marseille-5 variant were also identified in Switzerland and Norway whereas Marseille-3 sequences were obtained from France.

The Marseille-6 to -10 variants were first detected in weeks 30-33 and accounted for less than 10% of the genomes sequenced since then. The closest matches for Marseille-5, -6 and -7 genomes in GISAID are also sequences from European countries. Overall, we observed the disappearance (two successive weeks without case) of all variants but Marseille-4 6 to 11 weeks after their first detection.

Conclusion

Here we show that two epidemic phenomena occurred in France. The first, from March to May 2020, looked typical of a respiratory viral infection and was similar to the one observed in Eastern Asia, whereas the second evolved in two successive waves after an almost total disappearance. The onset of these waves paralleled the appearance of 10 significantly different viral variants that most likely resulted from multiple mutations. Most of these variants seem to have a lower epidemic potential than the first SARS-Cov-2 imported from China. In any case, the occurrence of several epidemics with different variants requires careful consideration of the management of future vaccines, including multi-variant composition.

We believe that the evolution of this virus corresponds exactly to a Darwinian model. Indeed, during episode 1, following the global spread of SARS-COV-2, two concomitant phenomena occurred. On the one hand, political responses have consisted in the interruption of travels and border closure, which resulted in the isolation of distinct ecosystems for several months. On the other hand, large concentrations of farmed minks were infected by human SARS-CoV-2. Under these conditions, speciation may occur in isolated ecosystems [9]. In the present case, the re-connection of isolated ecosystems (either countries and/or farmed animals) where different variants had developed generated new outbreaks in countries that were the most exposed to incoming populations such as travelers. This was in particular the case for Mediterranean countries and, in France, for Marseille that received an

elevated number of tourists in Summer 2020. In contrast, China did not experiment any “second wave”. We believe that the segregation of viral strains in isolated geographical areas and possibly in animals, may indeed explain the differences observed among epidemic curves around the world. This helps to explain the second episode that developed in Marseille, initially caused by an African variant that disappeared and then by emerging new variants linked to different areas of Europe, including Spain for variant 2, Northern Europe for variants 4 and 5, following the high touristic season in July-August with many visitors from Europe.

Acknowledgments

We are grateful to Olivia Ardizzoni, Madeleine Carrera, Vera Esteves-Vieira, Laurence Thomas, Priscilla Jardot and Raphael Tola for their technical help, and to Audrey Giraud-Gatineau and Lea Luciani for their help with data analysis. This manuscript has been edited by a native English speaker.

Ethical approval

The study was approved by the ethical committee of the Mediterranee Infection institute under reference 2020-016-2. Access to the patients' biological and registry data issued from the hospital information system was approved by the data protection committee of Assistance Publique-Hopitaux de Marseille (APHM) and was recorded in the European General Data Protection Regulation registry under number RGPD/APHM 2019-73.

Funding

This work was supported by the French Government under the "Investments for the Future" program managed by the National Agency for Research (ANR), Mediterranee-Infection 10-IAHU-03 and was also supported by Region Provence Alpes Cote d'Azur and European funding FEDER PRIMMI (Fonds Europeen de Developpement Regional-Plateformes de Recherche et d'Innovation Mutualisees Mediterranee Infection), FEDER PA 0000320 PRIMMI.

Transparency declaration

The 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.

References

(1) Lagier JC, Million M, Gautret P, et al. Outcomes of 3,737 COVID-19 patients treated with hydroxychloroquine/azithromycin and other regimens in Marseille, France: A retrospective analysis. Travel Med Infect Dis **2020 Jun 25**;101791.

(2) Colson P, Esteves-Vieira V, Giraud-Gatineau A, et al. Temporal and age distributions of SARS-CoV-2 and other coronaviruses, southeastern France. Int J Infect Dis **2020 Sep 23**; 101:121-5.

(3) Colson P, Lagier JC, Baudoin JP, Bou KJ, La SB, Raoult D. Ultrarapid diagnosis, microscope imaging, genome sequencing, and culture isolation of SARS-CoV-2. Eur J Clin Microbiol Infect Dis **2020 Aug**; 39(8):1601-3.

(4) Brahim BD, Fontanini A, Baudoin JP, et al. The Strengths of Scanning Electron Microscopy in Deciphering SARS-CoV-2 Infectious Cycle. Front Microbiol **2020**; 11:2014.

(5) Hammer AS, Quaade ML, Rasmussen TB, et al. SARS-CoV-2 Transmission between Mink (Neovison vison) and Humans, Denmark. Emerg Infect Dis **2020 Nov 18**; 27(2).

(6) Lam SD, Bordin N, Waman VP, et al. SARS-CoV-2 spike protein predicted to form complexes with host receptor protein orthologues from a broad range of mammals. Sci Rep **2020 Oct 5**; 10(1):16471.

(7) Koeijer AA, Hagenaars TJ, Leuken JPGV, Swart AN, Boender GJ. Spatial transmission risk during the 2007-2010 Q fever epidemic in The Netherlands: Analysis of the farm-to-farm and farm-to-resident transmission. PLoS One **2020**; 15(2):e0227491.

(8) hodcroft EB, Zuber M, Nadeau S, et al. Emergence and spread of a SARS-CoV-2 variant through Europe in the summer of 2020. medRxiv [Preprint] **2020**; https://doi.org/10.1101/2020.10.25.20219063.

(9) Darwin C. On the origin of species. London: John Murray, **1859**.

(10) Hadfield J, Megill C, Bell SM, et al. Nextstrain: real-time tracking of pathogen evolution. Bioinformatics **2018 Dec 1**; 34(23):4121-3.

Figures

Figure 1. Number of qPCR tests performed at Mediterranee Infection institute from February to December 2020.

Grey: number of qPCR tests performed at IHU Mediterranee Infection; black curve: percentage of positive tests.

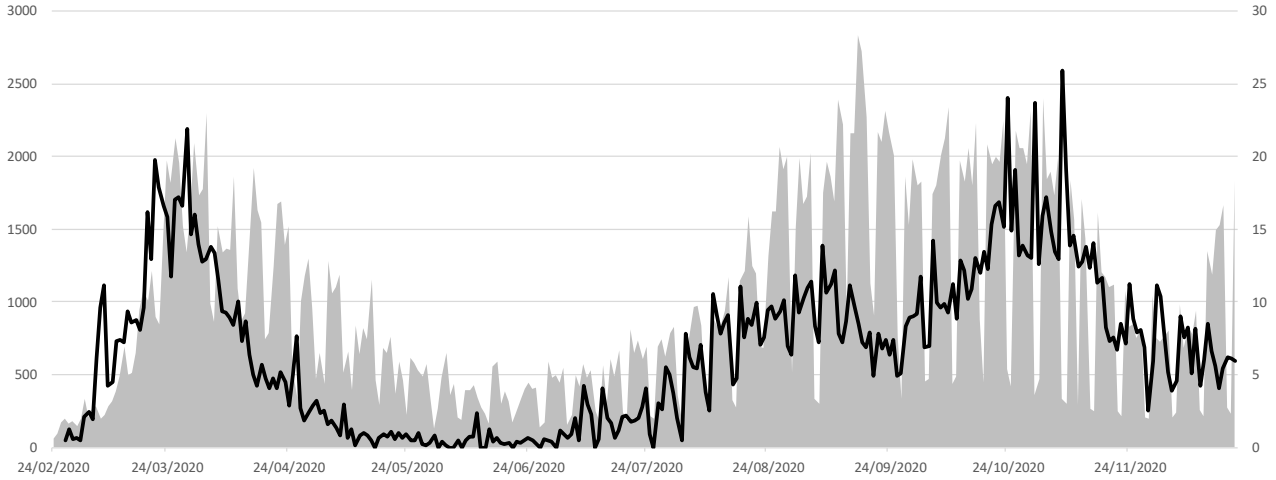


Figure 2. time-scaled phylogenetic tree based on SARS-CoV-2 full-length genomes obtained during the two outbreaks from February to May (episode 1) and from June to December (episode 2).

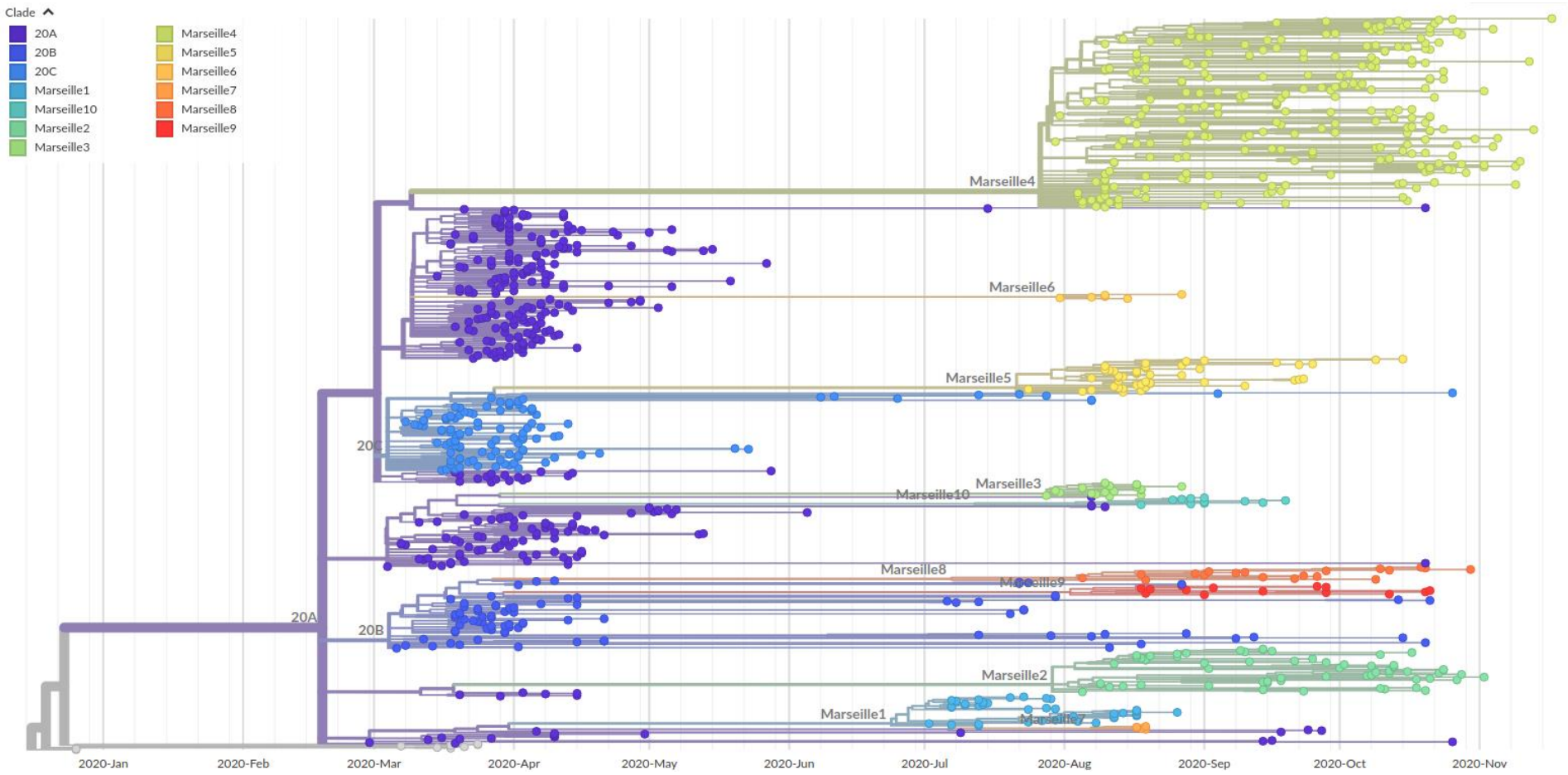


Figure 3. Proportion of SARS-CoV-2 genome sequences per ISO week that fall within each variant.

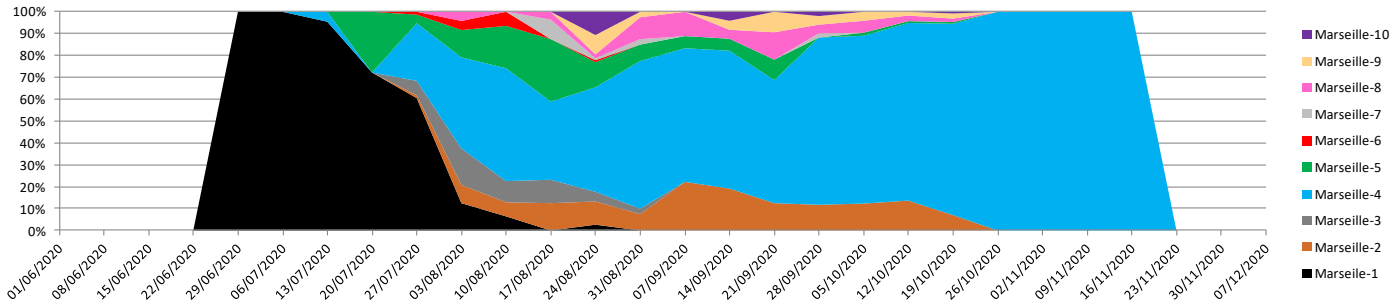
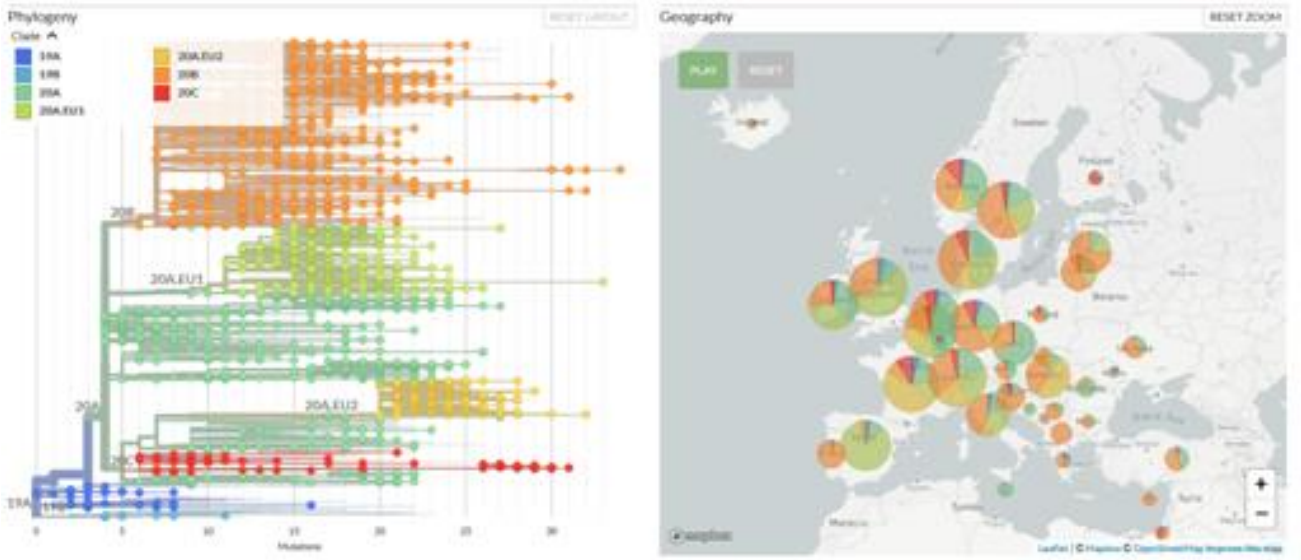


Figure 4. Europe (A) and France (B) maps showing the numbers and distribution into genetic variants of SARS-CoV-2 genome sequences obtained from nasopharyngeal specimens. The images were obtained from the Nextstrain webtool (https://clades.nextstrain.org/).

A



B

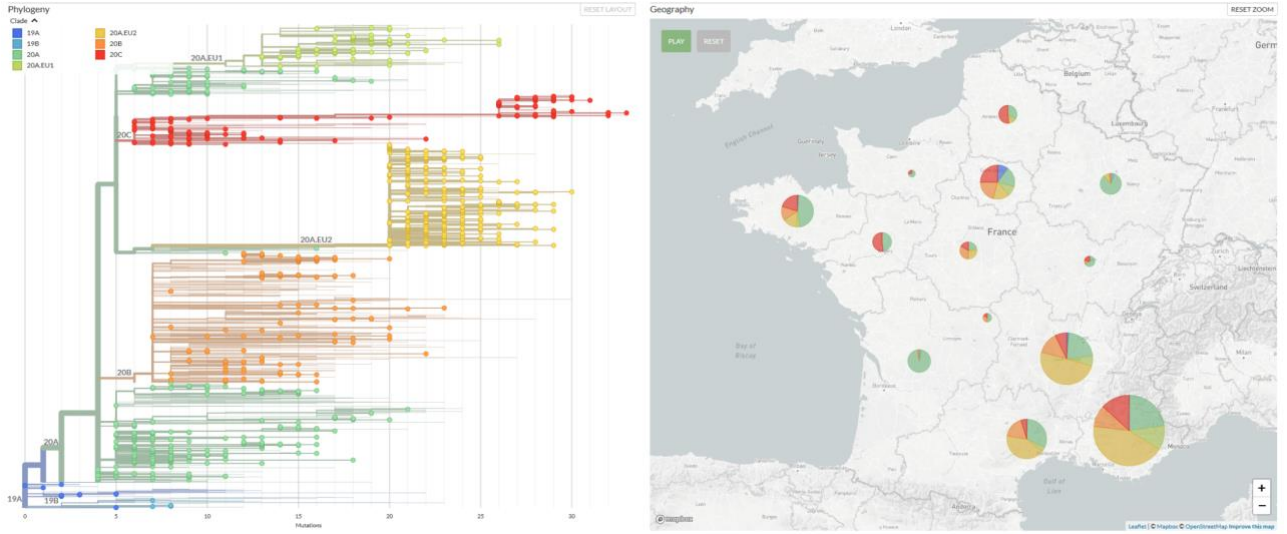


Table 1. Number of SARS-CoV-2 genomes per genotype and period (February-May (episode 1) and June–November (episode 2))

Period	Closest ancestor	Lineage	Number of mutations relative to the Wuhan-Hu-1 genome	February-May (period #1)		June-November (period #2)		Total	Hallmark nucleotide mutations
				Number	%	Number	%		
Periods #1 and #2	20A	**20A/15324T**	5	**44**	*89,8*	**5**	*10,2*	**49**	- - C15324T
	20A/15324T	**20A/15324T-3**	6	**9**	*90,0*	**1**	*10,0*	**10**	- - - G9479T
	20A/18877T	**20A/18877T-1**	7	**1**	*33,3*	**2**	*66,7*	**3**	- - - - C26735T
	20A	**20A/20268G**	5	**2**	*28,6*	**5**	*71,4*	**7**	- - A20268G
	20A	**20B**	7	**36**	*65,5*	**19**	*34,5*	**55**	- - G28881A, G28882A, G28883C
	20A/25563T	**20C**	6	**51**	*96,2*	**2**	*3,8*	**53**	- - - C1059T
	20C	**20C-2**	7	**4**	*80,0*	**1**	*20,0*	**5**	- - - - C10582T
Period #1 only	Wuhan	**19**	0	**7**	*100,0*	**0**	*0,0*	**7**	
	Wuhan	**20A**	4	**7**	*100,0*	**0**	*0,0*	**7**	- C241T, C3037T, C14408T, A23403G
	20A	**20A/25563T**	5	**16**	*100,0*	**0**	*0,0*	**16**	- - G25563T
	20A/25563T	**20A/18877T**	6	**1**	*100,0*	**0**	*0,0*	**1**	- - - C18877T
	20A/20268G	**20A/20268G-2**	6	**10**	*100,0*	**0**	*0,0*	**10**	- - - C18175T
	20A/15324T	**20A/15324T-1**	6	**7**	*100,0*	**0**	*0,0*	**7**	- - - C2189T
	20A/15324T-1	**20A/15324T-1a**	7	**15**	*100,0*	**0**	*0,0*	**15**	- - - - G4105T
	20A/15324T	**20A/15324T-4**	6	**8**	*100,0*	**0**	*0,0*	**8**	- - - C19602T
	20A/25563T	**20A/25563T-1**	6	**10**	*100,0*	**0**	*0,0*	**10**	- - - C2416T
	20A/25563T-1	**20A/25563T-1a**	8	**1**	*100,0*	**0**	*0,0*	**1**	- - - - C13458T, C20946T
	20A/25563T-1a	**20A/25563T-1a1**	9	**6**	*100,0*	**0**	*0,0*	**6**	- - - - - C6401T
	20A/25563T-1a1	**20A/25563T-1a1a**	10	**9**	*100,0*	**0**	*0,0*	**9**	- - - - - - A1198G
	20A/25563T-1	**20A/25563T-1b**	7	**121**	*100,0*	**0**	*0,0*	**121**	- - - - G8371T
	20A/25563T-1b	**20A/25563T-1b1**	8	**8**	*100,0*	**0**	*0,0*	**8**	- - - - - C26907A
	20A/25563T-1b	**20A/25563T-1b2**	9	**6**	*100,0*	**0**	*0,0*	**6**	- - - - - C9996T, G29747T
	20A/25563T-1b	**20A/25563T-1b3**	8	**5**	*100,0*	**0**	*0,0*	**5**	- - - - - G26718T
	20A/25563T-1b	**20A/25563T-1b4**	8	**6**	*100,0*	**0**	*0,0*	**6**	- - - - - G25909T
	20A/25563T-1b	**20A/25563T-1b5**	8	**5**	*100,0*	**0**	*0,0*	**5**	- - - - - G25767T
	20A/25563T-1b	**20A/25563T-1b6**	9	**6**	*100,0*	**0**	*0,0*	**6**	- - - - - A20294G, G29779C

Table 1 (continued)

Period	Closest ancestor	Lineage	Number of mutations relative to the Wuhan-Hu-1 genome	February-May (period #1)		June-November (period #2)		Total	Hallmark nucleotide mutations
				Number	%	Number	%		
	20A/25563T-1b	**20A/25563T-1b7**	8	**1**	*100,0*	**0**	*0,0*	**1**	- - - - - A12334T
	20A/25563T-1b7	**20A/25563T-1b7a**	9	**5**	*100,0*	**0**	*0,0*	**5**	- - - - - - C15810T
	20A/25563T-1b	**20A/25563T-1b8**	8	**5**	*100,0*	**0**	*0,0*	**5**	- - - - - C2910T
	20A/25563T-1b	**20A/25563T-1b9**	10	**5**	*100,0*	**0**	*0,0*	**5**	- - - - - G11083T, C19160T, G28027T
	20A/25563T-1	**20A/25563T-1c**	10	**6**	*100,0*	**0**	*0,0*	**6**	- - - - G3338T, G15438T, C21597T, C25731T
	20B	**20B-1**	8	**11**	*100,0*	**0**	*0,0*	**11**	- - - C313T
	20B-1	**20B-1a**	10	**6**	*100,0*	**0**	*0,0*	**6**	- - - - G19518T, G28845T
	20B-1	**20B-1b**	9	**5**	*100,0*	**0**	*0,0*	**5**	- - - - G24998T
	20C	**20C-1**	7	**5**	*100,0*	**0**	*0,0*	**5**	- - - - C9286T
	20C	**20C-3**	8	**4**	*100,0*	**0**	*0,0*	**4**	- - - - T13006C, C25688T
	20C-3	**20C-3a**	9	**6**	*100,0*	**0**	*0,0*	**6**	- - - - - C14391T
	20C	**20C-4**	7	**7**	*100,0*	**0**	*0,0*	**7**	- - - - G27996T
	20C	**20C-5**	7	**19**	*100,0*	**0**	*0,0*	**19**	- - - - G11083T
	20C	**20C-6**	7	**5**	*100,0*	**0**	*0,0*	**5**	- - - - G8039T
Period #2 only	20A/20268G-1	**20A/20268G-1a**	7	**0**	*0,0*	**2**	*100,0*	**2**	- - - - C25731T
		20B-2	9	**0**	*0,0*	**6**	*100,0*	**6**	- - - T27299C, T29148C
	20C-2	**20C-2a**	9	**0**	*0,0*	**7**	*100,0*	**7**	- - - - - C27804T, C28830A
	20A/20268G-1	**Marseille-1**	12	**0**	*0,0*	**24**	*100,0*	**24**	- - - - G1181T, C1625T, G22894A, C25886T, G28198T, G28851T
	Marseille-1	**Marseille-1-A**	13	**0**	*0,0*	**12**	*100,0*	**12**	- - - - - C22088T
	Marseille-1-A	**Marseille-1-A1**	14	**0**	*0,0*	**6**	*100,0*	**6**	- - - - - - G5378A
	20A	**Marseille-2**	11	**0**	*0,0*	**17**	*100,0*	**17**	- - T445C, C6286T, G21255C, C22227T, C26801G, C28932T, G29645T
	Marseille-2	**Marseille-2-A**	15	**0**	*0,0*	**1**	*100,0*	**1**	- - - G4006T, C5170T, G11132T, T26609C
	Marseille-2-A	**Marseille-2-A1**	16	**0**	*0,0*	**22**	*100,0*	**22**	- - - - G5504A
	Marseille-2-A	**Marseille-2-A2**	17	**0**	*0,0*	**7**	*100,0*	**7**	- - - - C10615T, C23191T
	Marseille-2	**Marseille-2-B**	12	**0**	*0,0*	**5**	*100,0*	**5**	- - - C15738T
	Marseille-2-B	**Marseille-2-B1**	15	**0**	*0,0*	**10**	*100,0*	**10**	- - - - G5950T, G14500T, G17427T
	20A/15324T	**Marseille-3**	12	**0**	*0,0*	**20**	*100,0*	**20**	- - - C1912T, G5210A, C17470T, C21191T, A23148G, T27125C, C28854T

Table 1 (continued)

Period	Closest ancestor	Lineage	Number of mutations relative to the Wuhan-Hu-1 genome	February-May (period #1)		June-November (period #2)		Total	Hallmark nucleotide mutations
				Number	%	Number	%		
	20A/18877T-1	**Marseille-4**	20	**0**	*0,0*	**168**	*100,0*	**168**	- - - - - C4543T, G5629T, G9526T, C11497T, G13993T, G15766T, A16889G, G17019T, G22992A, C25710T, T26876C, G28975C, G29399A
	Marseille-4	**Marseille-4-A**	21	**0**	*0,0*	**19**	*100,0*	**19**	- - - - - - G28086T
	Marseille-4-A	**Marseille-4-A1**	21	**0**	*0,0*	**11**	*100,0*	**11**	- - - - - - - C27942T
	Marseille-4	**Marseille-4-B**	21	**0**	*0,0*	**4**	*100,0*	**4**	- - - - - - C8937T
	Marseille-4	**Marseille-4-C**	24	**0**	*0,0*	**16**	*100,0*	**16**	- - - - - - C222T, C503T, G26720T
	Marseille-4	**Marseille-4-D**	21	**0**	*0,0*	**6**	*100,0*	**6**	- - - - - - G18105T
	Marseille-4	**Marseille-4-E**	21	**0**	*0,0*	**6**	*100,0*	**6**	- - - - - - G27877T
	Marseille-4	**Marseille-4-F**	22	**0**	*0,0*	**6**	*100,0*	**6**	- - - - - - T26442C, G29511T
	Marseille-4	**Marseille-4-G**	21	**0**	*0,0*	**5**	*100,0*	**5**	- - - - - - C23191T
	Marseille-4	**Marseille-4-H**	21	**0**	*0,0*	**12**	*100,0*	**12**	- - - - - - G29701A
	Marseille-4	**Marseille-4-I**	21	**0**	*0,0*	**9**	*100,0*	**9**	- - - - - - G2600T
	Marseille-4	**Marseille-4-J**	22	**0**	*0,0*	**7**	*100,0*	**7**	- - - - - - A2647G, G25534T
	20C-2a	**Marseille-5**	26	**0**	*0,0*	**25**	*100,0*	**25**	- - - - - - C3099T, G4960T, C4965T, C6070T, C7303T, C7564T, C9246T, C10279T, C10301A, C10525T, G10688T, G11851T, C14230A, G21800T, G27632T, G29402T, G29779T
	Marseille-5	**Marseille-5-A**	29	**0**	*0,0*	**8**	*100,0*	**8**	- - - - - - - T8743C, G16377T, G16647A
	Marseille-5	**Marseille-5-B**	27	**0**	*0,0*	**15**	*100,0*	**15**	- - - - - - - C9142T
	20A/25563T-1b	**Marseille-6**	17	**0**	*0,0*	**7**	*100,0*	**7**	- - - - - C9430T, A15477T, C18395T, A20622T, G20623T, A20624T, C23730T, A26319G, C28854T, G29044A
	20A/20268G-1a	**Marseille-7**	9	**0**	*0,0*	**5**	*100,0*	**5**	- - - - C2706T, G27463C
	20B	**Marseille-8**	12	**0**	*0,0*	**1**	*100,0*	**1**	- - - C5055T, G11851T, A12755G, G24812T, C26895T
	Marseille-8	**Marseille-8-A**	13	**0**	*0,0*	**18**	*100,0*	**18**	- - - - C28887T
	20B	**Marseille-9**	12	**0**	*0,0*	**7**	*100,0*	**7**	- - - A11782G, T21570G, C21575T, T25473C, C28253T
	Marseille-9	**Marseille-9-A**	13	**0**	*0,0*	**1**	*100,0*	**1**	- - - - C24138T
	Marseille-9-A	**Marseille-9-A1**	15	**0**	*0,0*	**6**	*100,0*	**6**	- - - - - A20003C, C24083T
	20A/15324T	**Marseille-10**	13	**0**	*0,0*	**1**	*100,0*	**1**	- - - C3602T, C6941T, C21855T, A25505G, G25906C, G25996T, C28651T, C28869T
	Marseille-10	**Marseille-10-A**	14	**0**	*0,0*	**10**	*100,0*	**10**	- - - - A16044T
		Total	-	**491**	-	**547**	-	**1038**	

1