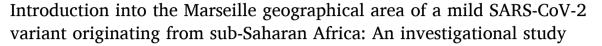Contents lists available at ScienceDirect

Travel Medicine and Infectious Disease

journal homepage: www.elsevier.com/locate/tmaid

Original article

Introduction into the Marseille geographical area of a mild SARS-CoV-2 variant originating from sub-Saharan Africa: An investigational study

Philippe Colson [a,b], Anthony Levasseur [a,b], Philippe Gautret [a,b], Florence Fenollar [a,c],
Van Thuan Hoang [a,c,d], Jeremy Delerce [a,b], Idir Bitam [c,e], Rachid Saile [f], Mossaab Maaloum [f],
Abdou Padane [g], Marielle Bedotto [a], Ludivine Brechard [a], Vincent Bossi [a],
Mariem Ben Khedher [a,b], Hervé Chaudet [a,c,h], Matthieu Million [a,b], Hervé Tissot-Dupont [a,b],
Jean-Christophe Lagier [a,b], Souleymane Mboup [g], Pierre-Edouard Fournier [a,c,**],
Didier Raoult [a,b,*]

[a] IHU Méditerranée Infection, Marseille, France
[b] Aix-Marseille Univ, Microbes Evolution Phylogeny and Infections (MEPHI), Marseille, France
[c] Aix-Marseille Univ, Vecteurs - Infections Tropicales et Méditerranéennes (VITROME), Marseille, France
[d] Thai Binh University of Medicine and Pharmacy, Thai Binh, Viet Nam
[e] Ecole Nationale Supérieure en Sciences de L'Aliment et des Industries Agroalimentaire, Alger, Algeria
[f] Faculty of Sciences Ben M'sik, Hassan II University of Casablanca, Morocco
[g] Institut de Recherche en Santé, de Surveillance Epidémiologique et de Formation (IRESSEF), Rufisque, Senegal
[h] French Armed Forces Center for Epidemiology and Public Health (CESPA), Marseille, France

ABSTRACT

Background: In Marseille, France, the COVID-19 incidence evolved unusually with several successive epidemic phases. The second outbreak started in July, was associated with North Africa, and involved travelers and an outbreak on passenger ships. This suggested the involvement of a new viral variant.
Methods: We sequenced the genomes from 916 SARS-CoV-2 strains from COVID-19 patients in our institute. The patients' demographic and clinical features were compared according to the infecting viral variant.
Results: From June 26th to August 14th, we identified a new viral variant (Marseille-1). Based on genome sequences (n = 89) or specific qPCR (n = 53), 142 patients infected with this variant were detected. It is characterized by a combination of 10 mutations located in the *nsp2, nsp3, nsp12, S, ORF3a, ORF8* and *N/ORF14* genes. We identified Senegal and Gambia, where the virus had been transferred from China and Europe in February–April as the sources of the Marseille-1 variant, which then most likely reached Marseille through Maghreb when French borders reopened. In France, this variant apparently remained almost limited to Marseille. In addition, it was significantly associated with a milder disease compared to clade 20A ancestor strains, in univariate analysis.
Conclusion: Our results demonstrate that SARS-CoV-2 can genetically diversify rapidly, its variants can diffuse internationally and cause successive outbreaks.

1. Introduction

The SARS-CoV-2 virus emerged in humans in Wuhan, China, in December 2019, prior to speading worldwide. In China and Europe, the epidemic had a bell shape typical of a respiratory virus (https://covid19-country-overviews.ecdc.europa.eu/) [1,2]. Surprisingly, in the

countries that closed their borders, the evolution varied: in some countries, no other epidemic was detected whereas in others, new epidemic waves occurred, caused by new variants [3]. Among the new sources that may explain the occurrence of different epidemics according to geographical zones, the role of intensive animal breeding like mink farming in Denmark [4] remains a mystery. In Marseille, the

bell-shaped curve ended in May, but new cases and then an atypically-shaped epidemic reappeared upon the border reopening. The reopening of borders with Maghreb occurred despite the fact that a very active COVID-19 outbreak was ongoing in Algeria. Interestingly, the first cases of the July epidemic had direct or indirect contacts with passengers from ferries coming from Tunisia or Algeria, which led us to suspect that this variant had an African origin. In our institute (Méditerranée Infection Institute [IHU]) in Marseille, Southern France, we investigated the viral genotypes from patients diagnosed in Marseille using genomic sequencing and genotype-specific PCR. Then, we also tested patients' specimens from Algerian, Moroccan and Senegalese residents for the presence of a new variant that we named Marseille-1.

2. Materials and methods

2.1. Virological diagnosis of SARS-CoV-2 infections

In IHU in Marseille, France, we have carried out SARS-CoV-2 RNA testing using real-time reverse transcription-PCR (qPCR) since the end of January 2020, as previously described [1,5]. The numbers of tests and cases were daily monitored since the first positive diagnosis on 02/27/2020 [2] (https://www.mediterranee-infection.com/covid-19/).

2.2. Study period and clinical samples

Whole genome sequencing of SARS-CoV-2 genome was performed from nasopharyngeal samples tested between June and August 2020 at IHU. Specimens with a cycle threshold value (Ct) lower than 20 were selected in priority, and those with a Ct between 20 and 30 were included secondarily to cover the study period more comprehensively.

2.3. Genome sequencing and analysis

Viral genomes were obtained using next-generation sequencing (NGS) and the Illumina technology (Illumina Inc., San Diego, CA, USA), as previously described [2,6]. Viral RNA was extracted from 200 μL of nasopharyngeal swab fluid using the EZ1 Virus Mini Kit v2.0, and was reverse transcribed using SuperScript IV (ThermoFisher Scientific, Waltham, MA, USA) prior to cDNA second strand synthesis with Klenow Fragment DNA polymerase (New England Biolabs, Beverly, MA, USA). The generated DNA was purified using Agencourt AMPure XP beads (Beckman Coulter, Villepinte, France) and sequenced using the Illumina Nextera XT Paired end strategy on a MiSeq instrument. Genome consensus sequences were generated with the CLC Genomics workbench v.7 by mapping on the SARS-CoV-2 genome GenBank accession no. NC_045512.2 (Wuhan-Hu-1 isolate) with the following thresholds: 0.8 for coverage and 0.9 for similarity. SARS-CoV-2 sequences obtained in our institute have been submitted to the GISAID database.

Sequences from complete genomes were analyzed using the Next-clade web-tool (https://clades.nextstrain.org/) [7]. Clades were defined based on the occurrence of at least five genomes sharing the same pattern of mutations. These genome sequences were compared to those available in the GISAID database (https://www.gisaid.org/). Phylogenetic trees were reconstructed by using Nextclade and visualized with iTOL (https://itol.embl.de/).

2.4. Marseille-1 variant specific RT-PCR

For the specimens with Ct values > 30 or those with Ct values < 30 but from which genome sequences were not obtained, we attempted to identify those harboring the Marseille-1 variant using RT-PCR targeting a fragment of the nucleocapsid-encoding gene harboring two mutations separated by 17 nucleotides concurrently present in the Marseille-1 variant. RT-PCR was performed with the PrimerF1 (forward, 5′-TCTACGCAGAAGGGAGCAGA-3′) and PrimerR1 (reverse, 5′-GGA-GAAGTTCCCCTACTGCTG-3′) primers, and the QuantiNova SYBR Green

RT-PCR kit (Qiagen, Hilden, Germany). In order to evaluate whether the Marseille-1 variant was also prevalent in these countries, this PCR system was also applied to 97 SARS-Cov-2 specimens from COVID-19-positive residents from Senegal, 278 from Algeria and 94 from Morocco. All specimens had been sampled in October and November 2020.

2.5. Comparisons of epidemiological and clinical features of patients diagnosed during phase 1 and phase 2

From February 29th to August 31st, 2020, the demographic and clinical features of the patients infected with the Marseille-1 variant were compared to those of the patients infected during phase 1 with 20A variants.

2.6. Statistical analyses

Statistical tests were done using R 4.0.2 (https://cran.r-project.org/bin/windows/base/): Chi2 or Fisher's exact test for qualitative variables, and Student's t-test for quantitative variables. A $p < 0.05$ was considered statistically significant.

2.7. Ethics statement

The study was approved by the ethical committee of the Méditerranée Infection Institute (N°: 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-Hôpitaux de Marseille (APHM) and was recorded in the European General Data Protection Regulation registry under number RGPD/APHM 2019–73.

3. Results

3.1. Epidemiological evidence of an african source

The Marseille-1 variant emerged in week 27, when it accounted for 100% of sequenced genomes (Fig. 1). Between June 24th and July 15th, information on recent travel abroad by SARS-CoV-2-positive patients was documented at our institute for 939 of 1108 cases. This period coincided with an increase in the number of positive SARS-CoV-2 tests from patients returning from Tunisia, with 14 cases, following the reopening of borders, and the resumption of passenger boat traffic between France and this country (Supplementary Figure 1, Supplementary Table 1). Other patients who traveled by ferry between North Africa (Tunisia and Algeria) and Marseille were diagnosed between July 8th and 20th. In addition, two patients diagnosed on July 3rd and July 7th, who had no travel history, reported a contact with SARS-CoV-2-positive Algerian travelers as source of their COVID-19. Another 35 positive patients worked or traveled on ships from the same passenger ferry company that sailed between Maghreb and Marseille. Such an evidence of an African origin for the Marseille-1 variant motivated a genotypic study to evaluate the incidence and history of this variant.

3.2. Outcome of the Marseille-1 variant

As of December 17th, 2020, 405,070 tests were performed for SARS-CoV-2 infection for 289,689 patients, of whom 25,446 (8.8%) were positive (Fig. 2). We have diagnosed 6855 patients during phase 1 and 18,591 during phase 2. We obtained SARS-CoV-2 full-length genome sequences from 916 patients (submitted to the GISAID database). These included 429 genomes from phase 2, which were added to 487 genomes from phase 1 [11]. Time-scaled phylogeny enabled differentiating ten clusters, identified as variants Marseille-1 to Marseille-10, encompassing at least 5 genomes each [3]. Between July and August, the Marseille-1 variant predominated, with 38 complete genome sequences
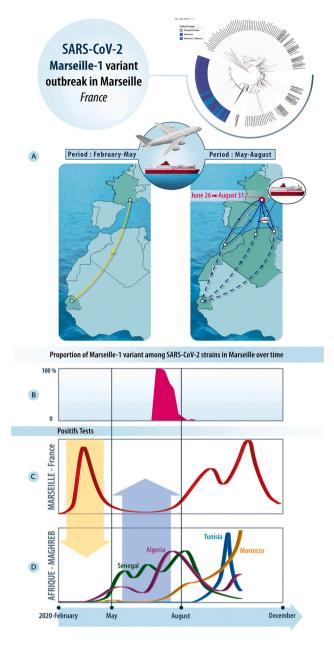
Fig. 1. The SARS-CoV-2 Marseille-1 variant outbreak. A. Left: transfer of the 20A variant from France to Senegal and Gambia during phase 1; right: proposed transfer route of the Marseille-1 variant from sub-Saharan Africa to Maghreb (dashed lines), and then to France.B. Evolution of the proportion of Marseille-1 variant genomes identified in Marseille over time. The time scale is featured at the bottom of figure. C. Epidemic curve of the COVID-19 in Marseille, France. D. Epidemic curves of the COVID-19 in Senegal, Morocco, Algeria and Senegal. The yellow arrow shows the transfer of the 20A variant from France to Senegal whereas the blue arrow shows the transfer of the Marseille-1 variant to France. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

being obtained. In addition, 51 draft genomes were identified as of the Marseille-1 variant, and this genotype was identified from an additional 53 samples by specific RT-PCR. Overall, 142 patients were detected as infected with the Marseille-1 variant. In addition, seven specimens/97 from Senegal, 44/278 from Algeria, and 15/94 from Morocco, all sampled in October and November 2020, were also RT-PCR-positive for the Marseille-1 variant (Supplementary Figure 1).

3.3. Phylogenetic study of the Marseille-1 variant

Phylogeny recontruction using genomic sequences available in GISAID showed that Marseille-1 variants belonged to a cluster that comprised almost only sequences from sub-Saharan Africa including from Senegal and Gambia as well as from the Marseille area (Fig. 3, Supplementary Figure 2). However, RT-PCR results demonstrated that Marseille-1 variants were also detected in travelers from Tunisia to Marseille as well as in Algerian and Moroccan residents. According to the date when sequenced strains were sampled, the earliest ancestors of Marseille-1 strains detected in the GISAID database were from samples collected in Senegal on March 31st, 2020, and in Gambia on June 1st, 2020 (Fig. 1).

Several studies have described that clade 20A strains were imported to sub-Saharan Africa, particularly Senegal and Gambia, mostly by French travelers during phase 1 [8–10]. We assume that the ancestors of Marseille-1 variants originated in Senegal or Gambia, through progressive accumulation of mutations by clade 20A strains, and that their offspring were then brought to Maghreb and imported to Marseille by travelers as well as ferry sailors from Algeria and Tunisia (Fig. 1). Such an hypothesis is supported by several arguments. First, a recent study of SARS-CoV-2 genomes suggested that sub-Saharan viral strains may have followed human routes through Mali, Tunisia or Egypt to Europe [10]. In addition, the emergence of Marseille-1 variants in Marseille followed the resumption on June 29th, 2020 of maritime and air connections between Maghreb and Marseille (Supplementary Table 1). However, despite the diffusion of the Marseille-1 epidemic in the population from Marseille, it seemingly barely spread outside the city, its prevalence rapidly decreased, being <10% for week 31 and it disappeared at the end of August.

3.4. Genotypic and amino acid patterns of Marseille-1 variants

Marseille-1 variants (that we propose to name 20A.MAR1) have a backbone of 10 mutations when compared to the Nextrain clade 20A strain, including C241U (*nsp2* gene [unknown function], synonymous), G1181U (*nsp2* gene, amino acid substitution V126F), C1625U (*nsp2* gene, amino acid substitution L274F), C3037U (*nsp3* [phosphoesterase], synonymous), C14408U (*nsp12* [RNA-dependent RNA polymerase, RdRp], amino acid substitution P314L), G22894A (spike-coding [*S*] gene, synonymous), A23403G (*S* gene, D614G), C25886U (ORF3a [predicted phosphoesterase], amino acid substitution S165F), G28198U (ORF8 [unknown function], amino acid substitution C102F) (Supplementary Figure 3) and G28851U (nucleocapsid-coding [*N*] gene, amino acid substitution S1931I). By comparison with the Wuhan-Hu-1 strain, two additional mutations were noted, including A20268G (*nsp 15*, synonymous) and C28833U (*N* gene, S187L; and ORF14, H34Y) (Table 1). Only amino acid substitutions V126F and L274F in the Nsp2 protein were found in the CoV-GLUE replacements database (http://cov-glue.cvr.gla.ac.uk/#/replacement) [11], in 123 and 308 GISAID genomes, respectively. The two additional mutations C22088U, corresponding to substitution L176F in the S protein and G5378A, corresponding to substitution G887S in the Nsp3 protein, were found in 612 and 9 GISAID genomes, respectively. In addition, we observed the successive occurrence of two additional mutations, C22088U (in Marseille-1A variants) then G5378A (in Marseille-1A1 variants) in 12 and 6, respectively, of the 38 full-length genomes obtained in Marseille. Moreover, 28 other mutations were found in between 1 and 6 of these Marseille-1 variant genomes, raising the total number of mutated nucleotide positions to 42 (Table 1). Interestingly, amino acid substitution C102F in ORF8 disrupts a disulfide bond close to the end of this 121 amino acid-long protein, which might alter its function, though as-yet undetermined.

By comparison with the original Wuhan-Hu-1 isolate, the Marseille-1 variant and its descendants (genotypes Marseille-1A and Marseille-1A1) differed by 12, 13 and 14 mutations, respectively. All three are classified
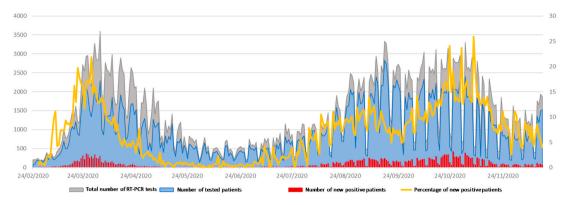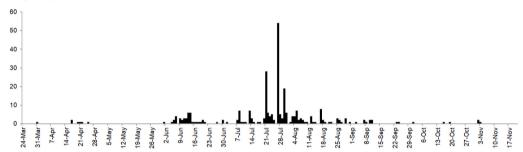
Fig. 2. Number of RT-PCR tests, tested patients, positive patients and percentage of positive patients from February 24th to December 17th, 2020.

a. Number of Marseille-1 genomes per day

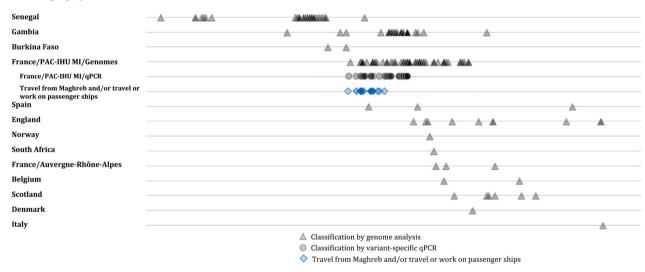**b. Time and geographical distribution of Marseille-1 variants**

Fig. 3. Evolution of the Marseille-1 variant epidemic, from Sub-Saharan Africa to Marseille. a) Numbers of patients with identified Marseille-1 strains detected over time in our institute; b) numbers of Marseille-1 genomes detected in our institute and in the GISAID database according to time of deposit and country. Triangles and circles indicate Marseille-1 strains identified using genome sequences or variant-specific PCR, respectively. Travel from Maghreb and/or travel or work on passenger ships is indicated by blue diamonds. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

in GISAID clade 20A and correspond to lineages B.1.5.12 and E.1 (an alias of B.1.5.12.1 lineage) in the Pangolin classification [12]. Their most recent identified ancestors are genomes from Senegal and Gambia that either harbored none of the 10 hallmark mutations, or C1625U, or C1625U associated to C25886U (Fig. 1). This strongly suggests the evolution of Marseille-1 ancestors in these countries through the successive occurrence of these mutations.

3.5. Demographic and clinical features of patients infected with the Marseille-1 variant

We compared the characteristics of 336 patients infected between March and April 2020 with clade 20A strains and 81 patients infected with the Marseille-1 variant (Table 2). The patients infected with the Marseille-1 variant were more frequently male and younger than those infected with clade 20A strains from phase 1. We acknowledge the fact that this specific demographic profile may reflect the traveler's profile of patients infected with the Marseille-1 variant during summer. Of the 417 patients, 56 were hospitalized. The hospitalization rate was lower in

Table 1
Nucleotide mutations and amino acid substitutions in the genomes of Marseille-1 variants.

Number of mutated genomes	Proportion of mutated genomes (%)	Nucleotide position	Gene	Wild-type nucleotide	Mutated nucleotide	Codon change	Codon number	Amino acid substitution
1	3	236	5UTR	G	U	-	-	-
38	100	241	5UTR	C	U	-	-	-
1	3	745	nsp1	C	U	AAC>AAU	160	N
1	3	865	nsp2	C	U	UGC>UGU	20	C
38	100	1181	nsp2	G	U	GUU>UUU	126	V126F
38	100	1625	nsp2	C	U	CUC>UUC	274	L274F
38	100	3037	nsp3	C	U	UUC>UUU	106	F
1	3	3505	nsp3	C	U	AAC>AAU	262	N
6	16	5378	nsp3	G	A	GGU>AGU	887	G887S
1	3	6681	nsp3	C	U	GCU>GUU	1321	A1321V
1	3	12436	nsp8	U	A	GUU>GUA	115	V
38	100	14408	nsp12b	C	U	CCU>CUU	314	P314L
1	3	14562	nsp12b	U	C	UAU>UAC	365	Y
38	100	20268	nsp15	A	G	UUA>UUG	216	L
18	47	22088	S	C	U	CUU>UUU	176	L176F
3	8	22583	S	G	A	GUU>AUU	341	V341I
38	100	22894	S	G	A	AAG>AAA	444	K
38	100	23403	S	A	G	GAU>GGU	614	D614G
2	5	23583	S	A	.	UAU>-	674	-
2	5	23584	S	U	.	UAU>-	674	-
2	5	23585	S	C	.	CAG>-	675	-
2	5	23586	S	A	.	CAG>-	675	-
2	5	23587	S	G	.	CAG>-	675	-
2	5	23588	S	A	.	ACU>-	676	-
2	5	23589	S	C	.	ACU>-	676	-
2	5	23590	S	U	.	ACU>-	676	-
2	5	23591	S	C	.	CAG>-	677	-
2	5	23592	S	A	.	CAG>-	677	-
2	5	23593	S	G	.	CAG>-	677	-
2	5	23594	S	A	.	ACU>-	678	-
2	5	23595	S	C	.	ACU>-	678	-
2	5	23596	S	U	.	ACU>-	678	-
2	5	23597	S	A	.	AAU>-	679	-
1	3	24197	S	G	U	GCG>UCG	879	A879S
38	100	25886	ORF3a	C	U	UCU>UUU	165	S165F
1	3	27143	M	C	U	AAC>AAU	?	N
1	3	27222	ORF6	U	C	UUU>UUC	7	F
1	3	27371	ORF6	C	U	CCA>CUA	57	P57L
38	100	28198	ORF8	G	U	UGU>UUU	102	C102F
1	3	28703	N	G	U	GAU>UAU	144	D144Y
38	100	28833	N	C	U	UCA>UUA	187	S187L
38	100	28851	N	G	U	AGU>AUU	193	S193I

(Genome identity columns for individual isolates PACL-HU-0942 (2020-07-02) through PACL-HU-2281-A (2020-08-03) omitted.)

Table 2
Demographics, outcomes and clinical symptoms in patients infected with different SARS-CoV-2 variants.

Demographics and outcomes (N = 417)	20A (N = 336)		Marseille-1 (N = 81)		p-value*
	n	%	n	%	
Male gender	149	44.4	46	56.8	0.044
Age (mean ± SD)	50.2 ± 22.4		36.7 ± 17.5		<0.0001
Hospitalization	53	15.8	3	3.5	0.004
Transfer to intensive care unit	5	1.5	0	0	0.59
Death	10	3.0	0	0	0.22

Symptoms (N = 320)	20A (N = 251)		Marseille-1 (N = 69)		p-value[a]
	n	%	n	%	
Cough	121	48.2	30	43.5	0.49
Rhinitis	104	41.4	22	31.9	0.15
Anosmia	76	30.3	25	36.2	0.35
Ageusia	72	28.7	22	31.9	0.61
Dyspnea	71	28.3	9	13.0	0.01
SpO 2 <96%	49	19.5	2	2.9	0.001

[a] Chi2 or Fisher exact test for qualitative variables. Student's t-test for quantitative variables.

patients infected with the Marseille-1 variant (3.5 *vs* 15.8%, $p = 0.004$). Ten patients died and five were transferred to intensive care unit, all of whom were infected with 20A variants (mortality rates of 3.0 and 0 for the 20A and Marseille-1 variants, respectively). Clinical symptoms were available in 320 patients (Table 2). Patients infected with the Marseille-1 variant suffered less frequently from dyspnea and hypoxemia. In contrast, rhinitis, anosmia and ageusia were not significantly different between patients infected with either of the two variants. Overall, the Marseille-1 variant exhibited a milder phenotype [13].

4. Discussion

Here we demonstrate that two SARS-CoV-2 epidemic phenomena occurred in France. The first one, from March to May 2020, exhibited a usual evolution for a respiratory viral infection, and was similar to the one observed in China. In contrast, following an almost total disappearance of SARS-CoV-2 diagnoses, the second epidemic phase evolved in successive or overlapping waves. These waves resulted from the occurrence of 10 viral variants exhibiting substantial genetic diversity between each other. Among them, the Marseille-1 variant caused a short outbreak than ran from July to August 2020, and remained essentially restricted to the Marseille area in France. In addition, our study demonstrated that the Marseille-1 variant was present in sub-Saharan (Senegal, Gambia) and North Africa (Tunisia, Algeria, Morocco). This variant is most likely a descendant from clade 20A strains transferred to sub-Saharan Africa by French travelers during phase 1 [10,14], prior to genetically evolving onsite and being later brought back to Maghreb and then to Marseille by travelers. The Marseille-1 variant was associated with a milder clinical outcome and a lower epidemic potential, and no associated death were observed. Criteria for hospital admission and transfer to intensive care unit did not differ over the study period. Patients infected with the Marseille-1 variant were younger that those infected with the clade 20-A which may have played a role in the severity of the disease. In addition, several factors known to be associated with severity including lymphopenia, D-dimer counts, troponin level, among others, were not considered in this analysis. More in-depth multivariate analysis is needed before any firm conclusion can be drawn about the clinical severity of the Marseille-1 variant. In addition, its epidemic potential was lower, and no case of re-infection with this variant was detected, contrasting with what we observed with other variants [15].

The evolution of the Marseille-1 variant as we understand it corresponds to a Darwinian model. Indeed, during phase 1, following the global spread of clade 20A strains, public health responses have consisted in the interruption of travels and border closure. This resulted in the isolation of distinct populations and ecosystems for several months,

in which viral variants may have evolved individually. However, we cannot rule out that the mutations observed in the Marseille-1 variant by comparison to the Chinese variant occurred in an animal population in sub-Saharan countries, as observed recently in farmed minks in Northern Europe [4]. Another example of an animal reservoir in coronavirus transmission was demonstrated for the MERS-CoV and camels in the Middle East [16]. In Sub-Saharan Africa, notably Senegal, African green monkeys, which were recently reported to be susceptible to SARS-CoV-2 [17], can form large colonies and live in close contact with humans in rural areas, may potentially play such a reservoir role. Under these conditions, what generally occurs is speciation [18]. As a matter of fact, the existence of animal reservoirs, infected during the first phase, may explain the differences in epidemic curves observed among countries. However, the consequences of viral variant selection in massive animal groups, and subsequent human infections, remain to be determined, especially because the immunity acquired by patients during phase 1 may not be protective against a re-infection with another variant [15].

Then, when international borders reopened and travels resumed, the reconnection of these isolated ecosystems where different variants had developed generated new outbreaks in areas that were the most exposed to incoming populations. This was in particular the case for Marseille where daily passenger air and boat traffic exchanges with Maghreb occur. Marseille is alas familiar with boat importation of epidemics from the South, notably plague and cholera [19]. Since November 2020, several new SARS-CoV-2 variants emerged, including the highly transmissible variants 20I/501Y.V1 in United Kingdom and 20H/501Y.V2 in South Africa. Like Marseille-1, these two variants have spread internationally with infected travelers. However, in contrast with the 20I/501Y.V1 and 20H/501Y.V2 variants that keep spreading, the diffusion of the Marseille-1 variant stopped in Marseille where it disappeared concomitantly with the emergence of other variants including the one that still predominates to date. This difference is currently unexplained but may be related to the great number of mutations in the spike protein-coding gene from the 20I/501Y.V1 and 20H/501Y.V2 variants, notably the N501Y mutation that may make them more infective.

Therefore, our results confirm that SARS-CoV-2 is able to genetically diversify rapidly, its variants to diffuse internationally with travelers and cause successive outbreaks, even in populations beforehand exposed to the original virus.

Funding

Declaration of competing interests

The authors declare no competing interests. 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.

Author contributions

Conceived and designed the study: PC, DR. Designed and/or performed experiments: VTH, JD, MB, LB, VB, MBK. Analyzed patients' data: PG, MM, JCL. Analyzed and interpreted data: PC, AL, PG, FF, IB, RS, MM, AP, HC, MM, HTD, JCL, SM, PEF, DR. Wrote the manuscript: PC, AL, PG, FF, MM, HC, MM, HTD, JCL, PEF, DR. All authors read and approved the final manuscript.

Acknowledgments

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.tmaid.2021.101980.

References

[1] Colson P, Esteves-Vieira V, Giraud-Gatineau A, Zandotti C, Filosa V, Chaudet H, et al. Temporal and age distributions of SARS-CoV-2 and other coronaviruses, southeastern France. Int J Infect Dis 2020;101:121–5.

[2] 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;39(8):1601–3.

[3] Fournier PE, Colson P, Levasseur A, Gautret P, Luciani L, Bedotto M, et al. Genome sequence analysis enabled to decipher the atypical evolution of COVID-19 epidemics in Marseille. 2020.

[4] 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.

[5] Lagier JC, Million M, Gautret P, Colson P, Cortaredona S, Giraud-Gatineau A, et al. Outcomes of 3,737 COVID-19 patients treated with hydroxychloroquine/ azithromycin and other regimens in Marseille, France: a retrospective analysis. Trav Med Infect Dis 2020:101791.

[6] Levasseur A, Delerce J, Caputo A, Brechard L, Colson P, Lagier JC, et al. Genomic diversity and evolution of coronavirus (SARS-CoV-2) in France from 309 COVID-19-infected patients. 2020.

[7] Hadfield J, Megill C, Bell SM, Huddleston J, Potter B, Callender C, et al. Nextstrain: real-time tracking of pathogen evolution. Bioinformatics 2018;34(23):4121–3.

[8] Lalaoui R, Bakour S, Raoult D, Verger P, Sokhna C, Devaux C, et al. What could explain the late emergence of COVID-19 in Africa? New Microbes New Infect 2020; 38:100760.

[9] Dia N, Lakh NA, Diagne MM, Mbaye KD, Taieb F, Fall NM, et al. COVID-19 outbreak, Senegal, 2020. Emerg Infect Dis 2020;26(11):2772–4.

[10] Wruck W, Adjaye J. Transmission of SARS-COV-2 from China to Europe and west Africa: a detailed phylogenetic analysis. 2020.

[11] Singer J, Gifford R, Cotten M, Robertson D. CoV-GLUE: a web application for tracking SARS-CoV-2 genomic variation. 2020.

[12] Rambaut A, Holmes EC, O'Toole A, Hill V, McCrone JT, Ruis C, et al. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. Nat Microbiol 2020;5(11):1403–7.

[13] Gautret P, Colson P, Lagier JC, Camoin-Jau L, Giraud-Gatineau A, Boudjema S, et al. Different pattern of the second outbreak of COVID-19 in Marseille, France. Int J Infect Dis 2021;102:17–9.

[14] Sun H, Dickens BL, Cook AR, Clapham HE. Importations of COVID-19 into African countries and risk of onward spread. BMC Infect Dis 2020;20(1):598.

[15] Colson P, Finaud M, Levy N, Lagier JC, Raoult D. Evidence of SARS-CoV-2 re-infection with a different genotype. J Infect 2020.

[16] Dudas G, Carvalho LM, Rambaut A, Bedford T. MERS-CoV spillover at the camel-human interface. Elife 2018;7.

[17] Woolsey C, Borisevich V, Prasad AN, Agans KN, Deer DJ, Dobias NS, et al. Establishment of an African green monkey model for COVID-19 and protection against re-infection. Nat Immunol 2021;22(1):86–98.

[18] Darwin C. On the origin of species. London: John Murray; 1859.

[19] Barbieri R, Signoli M, Cheve D, Costedoat C, Tzortzis S, Aboudharam G, et al. Yersinia pestis: the natural history of plague. Clin Microbiol Rev 2020;34(1).