SARS-CoV-2 variants combining spike mutations and the absence of ORF8 may be more transmissible and require close monitoring

Filipe Pereira [a], [b], [*]

[a] Centre for Functional Ecology, Department of Life Sciences, University of Coimbra, Calçada Martim de Freitas, 3000-456, Coimbra, Portugal
[b] IDENTIFICA Genetic Testing, Rua Simão Bolívar 259 3° Dir Tras, 4470-214, Maia, Portugal

ARTICLE INFO

ABSTRACT

The SARS-CoV-2 Variant of Concern 202012/01 (VOC-202012/01) emerged in southeast England and rapidly spread worldwide. This variant is believed to be more transmissible, with all attention being given to its spike mutations. However, VOC-202012/01 has also a mutation (Q27stop) that truncates the ORF8, a likely immune evasion protein. Removal of ORF8 changes the clinical outset of the disease, which may affect the virus transmissibility. Here I provide a detailed analysis of all reported ORF8-deficient lineages found in the background of relevant spike mutations, identified among 231,433 SARS-CoV-2 genomes. I found 19 ORF8 nonsense mutations, most of them occurring in the 5' half of the gene. The ORF8-deficient lineages were rare, representing 0.67% of sequenced genomes. Nevertheless, I identified two clusters of related sequences that emerged recently and spread in different countries. The widespread D614G spike mutation was found in most ORF-deficient lineages. Although less frequent, HV69-70del and L5F spike mutations occurred in the background of six different ORF8 nonsense mutations. I also confirmed that VOC-202012/01 is the ORF8-deficient variant with more spike mutations reported to date, although other variants could have up to six spike mutations, some of putative biological relevance. Overall, these results suggest that monitoring ORF8-deficient lineages is important for the progression of the COVID-19 pandemic, particularly when associated with relevant spike mutations.

1. Introduction

The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) triggered a major health crisis worldwide by causing the coronavirus disease 2019 (COVID-19). The virus spread rapidly from the original outbreak in Hubei province, China [1]. Several mutations have been described along the SARS-CoV-2 genome, as expected in a fast expanding population of a rapidly evolving RNA virus [2]. In November 2020, a variant named Variant of Concern 202012/01 (VOC-202012/01), originally termed VUI-202012/01, emerged in southeast England and rapidly spread locally and to other countries [3–5]. This variant belongs to lineage B.1.1.7 and gained particular interest by its fast spread and by having 14 non-synonymous mutations and 3 deletions. Focus has been placed on the mutations located in the spike protein due to their likely phenotypic effect. In particular, mutation N501Y has been shown to

provide a more favourable interaction with the angiotensin-converting enzyme 2 (ACE2) in mice [6]. P681H is located immediately adjacent to the furin cleavage site of importance for infection and transmission and the HV 69−70 deletion has been described in the context of evasion to the human immune response [7].

However, VOC-202012/01 has also a mutation that truncates the ORF8 accessory gene. It is now well recognised that SARS-CoV-2 can persist without a functional ORF8 protein, either by the accumulation of nonsense mutations or large deletions that remove or significantly change the ORF8 protein [8−10]. Moreover, substantial evidence has been accumulated over the last months to suggest that removal of ORF8 changes the clinical outset of the disease, with likely consequences on the transmissibility of the virus. A SARS-CoV-2 variant with a 382-nucleotide deletion in ORF8 results in a milder infection with less systemic release of proinflammatory cytokines, with patients having a longer duration of symptoms [10].

There is currently no data showing if VOC-202012/01 results in greater or lesser severity of disease in infected patients. However, preliminary data seems to indicate that this lineage is more

* Centre for Functional Ecology, Department of Life Sciences, University of Coimbra, Calçada Martim de Freitas, 3000-456, Coimbra, Portugal.
  E-mail address: fpereirapt@gmail.com.

transmissible than pre-existing variants of SARS-CoV-2 [4,11]. The co-occurrence of relevant spike mutations with the truncation of ORF8 raises the possibility that both loci are contributing to the increased transmissibility. The spike mutations may be affecting the receptor binding affinity of the spike protein enhancing the transmissibility of the virus, while the longer duration of symptoms resulting from the absence of ORF8 may increase the opportunity for transmission. It is therefore important to identify other SARS-CoV-2 variants that combine the absence of a functional ORF8 with spike mutation of putative effect on transmissibility and infection. If still in circulation, such linages may pose a risk of increased transmissibility, as observed with VOC-202012/01. Here I provide a detailed characterization of such variants, facilitating their future monitoring.

2. Materials and methods

2.1. SARS-CoV-2 sequences

SARS-CoV-2 sequences were obtained from the GISAID Initiative (https://www.gisaid.org/) using the filters "complete", "high coverage" and "low coverage excl" (all together), accessed on January 4, 2021. These filters excluded sequences with >1% Ns and with insertions/deletions not verified by the submitter. A total of 231,433 genomes were downloaded and analyzed for ORF8 and spike mutations. The sequences were aligned using the MAFFT version 7 online service [12], with the light-weight option for MSA of full-length SARS-CoV-2 genomes. The SARS-CoV-2 genome with accession number NC_045512.2 was used as reference.

2.2. ORF8 and spike mutations

The list of ORF8 nonsense mutations was obtained from a previous work [13] which used data from the China National Center for Bioinformation (CNCB) [14] and CoV-GLUE database [15].The selection of spike mutations was based on the changes observed in the VOC-202012/01 variant and in two publications: 1) Korber et al. [16], which proposed an analysis pipeline to facilitate real-time tracking of spike mutation that may confer selective advantages in transmission or resistance to interventions and 2) Li et al. [17] that investigated SARS-CoV-2 spike mutation variants for the infectivity and reactivity to a panel of neutralizing antibodies and sera from convalescent patients.

2.3. Clade assignment

All sequences have been classified into clades using the Nextclade tool (https://clades.nextstrain.org/). The tool recognizes 11 clades defined by specific signature mutations. Clades 19A and 19B emerged in Wuhan and dominated the early outbreak. Clade 20A is derived from 19A and dominated the European outbreak in March 2020. From 20A, two major distinct subclades are derived (20B and 20C), in addition to subclade 20E (EU1). More recently, 20D and 20I have emerged over the summer of 2020 and include two "variants of concern" (VOC) with signature spike mutation N501Y.

3. Results and discussion

3.1. Nonsense mutations occur preferentially in the 5' half of the ORF8 gene

I identified 19 different types of nonsense mutations in the SARS-CoV-2 ORF8 gene (Table 1). Most of these mutations (n = 14) occurred in the 5' half of the ORF8 gene, resulting in short truncated proteins that are most likely non-functional (Fig. 1A). Overall,

premature ORF8 stop codons were found in 1,540 out of 231,433 genomes (0.67%). The number of observed ORF8 nonsense mutations (n = 1,788) was higher than the number of genomes due to the co-occurrence of more than one mutation in the same genome (248 cases, 16.1% of all ORF8-deficient genomes; Fig. 1B). I found one sequence (EPI_ISL_700264) with three (G8stop, E106stop and E110stop) and 247 sequences with two ORF8 premature stop codons. The most frequent combination of ORF8 mutations included Q18stop and E110stop, found in 241 sequences. Interestingly, 98% (243 out of 248) of the cases with more than one nonsense mutation included the two most ORF8 C-terminal premature stop codons (E106stop and E110stop). The only exception were the five sequences from the VOC-202012/01 variant that had both Q27stop and K68stop mutations (Fig. 1B).

Overall, these results support the previous claim that ORF8 has the highest coefficient of premature stop codons among all SARS-CoV-2 genes [18]. Moreover, the data suggests that more than one premature stop codon can occur in ORF8, perhaps due to the relaxation of selective pressures on a non-functional protein already with a premature stop codon. Nevertheless, E106stop and E110stop were the most prevalent mutations in variants with more than one premature stop. These mutations result in an almost complete ORF8 protein, and it remains to be determined if the resulting shorter protein is functional during the viral infection. If E106stop and E110stop occurred first in those lineages, there may not be any selective relaxation for the occurrence of the second premature stop codon. Although rare among all SARS-CoV-2 sequenced genomes, the rate of premature termination at ORF8 proves that the virus can persist without a functional protein, as suggested by the ORF8-deleted variants [19,20] and the now widespread VOC-202012/01 [4,11].

3.2. ORF8 nonsense mutations are recurrent in the SARS-CoV-2 phylogeny, but two main clusters emerged recently

When considered independently, the most frequent ORF8 nonsense mutations were Q18stop (n = 534), E64stop (n = 348), E110stop (n = 306), Q27stop (n = 158) and E106stop (n = 123) (Table 1). In all cases, the same mutation was found occurring independently in different clades of the SARS-CoV-2 phylogeny (Supplementary Material S1). The same pattern was observed even for rare mutations, including the five mutations observed in less than 10 sequences each (L7stop, Q29stop, K44stop, L60stop and C61stop). Variants with Q27stop and E106stop did not form any clear cluster, despite being frequent in our dataset.

The Q18stop/E110stop cluster of variants explained near half of the Q18stop cases and more than two thirds of the E110stop cases. The Q18stop/E110stop ORF8 variants formed a large cluster in clade 20B (Fig. 1C). A single Q18stop/E110stop ORF8 sequence was classified in clade 20A. The cluster observed in 20B included sequences from England (n = 233), Luxemburg (n = 6) and Belgium (n = 1). The close geographical proximity suggests a local origin, even though GISAID sequences are biases towards UK due to their remarkable sequencing efforts [21]. The positioning of the cluster within the recent 20B clade suggests a recent origin. According to the Nextclade tool, all sequences in this cluster had 19 or more mutations in relation to the root of the tree build with reference sequences. The sequence with more mutations within this cluster was EPI_ISL_741554, collected in England by 2020-12-20. Interestingly, this cluster did not include the VOC-202012/01 variant (defined by having Q27stop), but had some spike mutations of interest (Fig. 1C).

The phylogenetic inference also revealed a large cluster of related ORF8-deficient variants with E64stop in clade 20B (Fig. 2). The 20B clade includes 265 (76.1%) of the 348 E64stop sequences.

Table 1
SARS-CoV-1 variants without a functional ORF8 due to nonsense mutations harbouring spike mutations of interest.

ORF8		Spike mutations									N° of spike sites with mutations
Nonsense mutation	N° sequences	L5F	H49Y	HV69-70del[a]	Y145H/del[a]	V367F	N501Y[a]	D614G	P681H[a]	P1263L	
L7stop	1	0	0	0	0	0	0	1	0	0	1
G8stop	72	0	2	0	0	0	0	21	0	1	3
Q18stop	534	11	1	25	15	1	3	527	1	0	8
E19stop	70	1	0	0	0	0	0	70	0	0	2
Q23stop[a]	18	0	0	0	0	0	0	16	0	0	1
Q27stop[a]	158	4	2	12	14	0	17	156	17	0	7
Q29stop	9	0	0	0	0	0	0	9	0	0	1
Y31stop	11	0	0	0	0	0	0	10	0	0	1
K44stop	3	0	0	0	0	0	0	3	0	0	1
W45stop	18	0	0	0	0	0	0	16	0	0	1
G50stop	18	0	0	0	0	0	0	14	0	0	1
K53stop	11	0	0	0	0	0	0	11	0	0	1
E59stop	67	2	0	2	0	0	0	65	0	0	3
L60stop	2	0	0	0	0	0	0	2	0	0	1
C61stop	3	0	0	0	0	0	0	3	0	0	1
E64stop	348	3	0	14	0	0	0	345	0	0	3
K68stop	16	0	0	5	4	0	5	16	5	0	5
E106stop	123	0	1	4	0	0	0	117	0	0	3
E110stop	306	11	0	0	15	2	0	302	0	1	5
Total	1788	32	6	62	48	3	25	1704	23	2	
N° of ORF8 nonsense mutations with spike mutations		6	4	6	4	2	3	19	3	2	

[a] Mutations found in the VOC-202012/01 variant.

Overall, the sequences were mainly obtained from samples in Northern Ireland ($n = 187$), Latvia ($n = 33$) and England ($n = 17$), although a few cases were reported in other countries. Two main branches occurred within this cluster, with one of the branches including all samples from Latvia ($n = 33$) and Finland ($n = 1$), one from Italy and a few from England ($n = 6$). This cluster represents 14.4% (33 out of 229) of all sequences from Latvia reported in the GISAID database. The other main branch within this cluster included all samples from Northern Ireland and most from England associated with this ORF8 mutation (Fig. 2). This cluster represents 12.1% (187 out of 1,537) of all sequences from Northern Ireland reported in the GISAID database. This pattern may indicate that both regions (Latvia and Northern Ireland/England) shared an early variant with an E64stop, which then spread locally in both areas. Whatever the case, the large number of related sequences deposited in GISAID suggests a fast and recent spread, which should require a careful monitoring.

3.3. SARS-CoV-2 variants with ORF8 Q27stop and E106stop are found in more than 20 countries, despite not being the most frequent ones

I searched for the presence of 20 spike mutations of interest [3,16,17] in all ORF8-deficient variants (Fig. 1A). The list includes two deletions (HV69-70del and Y145H/del) that remove amino acids from the spike protein. I found that 95.3% of ORF8-deficient variants had the spike D614G mutation. This mutation is associated with the B.1 lineage and now dominates the global distribution of SARS-CoV-2 lineages [22]. When not considering D614G, all variants with ORF8 nonsense mutations were found in at least two countries (Supplementary Fig. S2). The ORF8-deficient variants detected in more countries were E64stop (25 countries), E106stop (24 countries), Q18stop (23 countries) and Q27stop (20 countries), as shown in Fig. 3A and Supplementary Fig. S1. Q27stop and E106stop stand out by being present in more than 20 countries, despite being represented here by less than 200 sequences each (Fig. 3A). The Q27stop is present in the VOC-202012/01 variant, which lacks the widespread D614G mutation. The current number of VOC-202012/01 sequences reported so far is much higher than those analysed here because most sequences deposited in GISAID

do not comply with our quality criteria. Currently, only 13 out of 7,876 VOC-202012/01 sequences available in GISAID are complete and with high coverage according to the website's quality criteria (data accessed on January 4, 2021).

When considering all mutations together, the countries with more ORF8-deficient variants were England ($n = 458$), Northern Ireland ($n = 188$) and USA ($n = 171$). When excluding variants with D614G, the top countries were England ($n = 106$), Denmark ($n = 40$), Netherlands ($n = 16$), Scotland ($n = 12$) and USA ($n = 8$), as shown in Supplementary Fig. S2. With exception of Denmark, all these countries are in the top six of those submitting sequences to GISAID [21]. Therefore, the distribution of lineages is highly influenced by the sequencing efforts of these countries, and many cases certainly remain undetected. Nevertheless, some ORF8-deficient variants are found in several countries, which is not correlated with the number of reported sequences.

3.4. VOC-202012/01 is the ORF8-deficient variant with more spike mutations

All SARS-CoV-2 variants without a functional ORF8 had at least one relevant spike mutation (Table 1). D614G was the only spike mutation associated with all ORF8 mutations, being found in 1704 genomes. Apart from D614G, the most frequent mutations were the two spike deletions, HV69-70del ($n = 62$) and Y145H/del ($n = 48$). The HV69-70del mutation was found in six different ORF8 nonsense mutations, including the Q27stop observed in the VOC-202012/01 variant. The spike L5F mutation was also observed in the background of six different ORF8 nonsense mutations. The ORF8 Q18stop mutation was the most frequent in our dataset ($n = 534$) and was associated with the largest number of different spike mutations ($n = 8$; Fig. 3B). Despite having a lower number of sequence ($n = 158$), Q27stop stands out for being associated with seven different spike mutations, explained by the inclusion of the VOC-202012/01. Although with much lower numbers, the mutation K68stop ($n = 16$) was found in the background of five spike mutations. Eleven spike mutations were not found associated with ORF8 variants: L8V/W, N234Q, Q239K, L452R, A475V, G476S, V483A, F490L, V615I/F, A831V and D839Y/N/E.

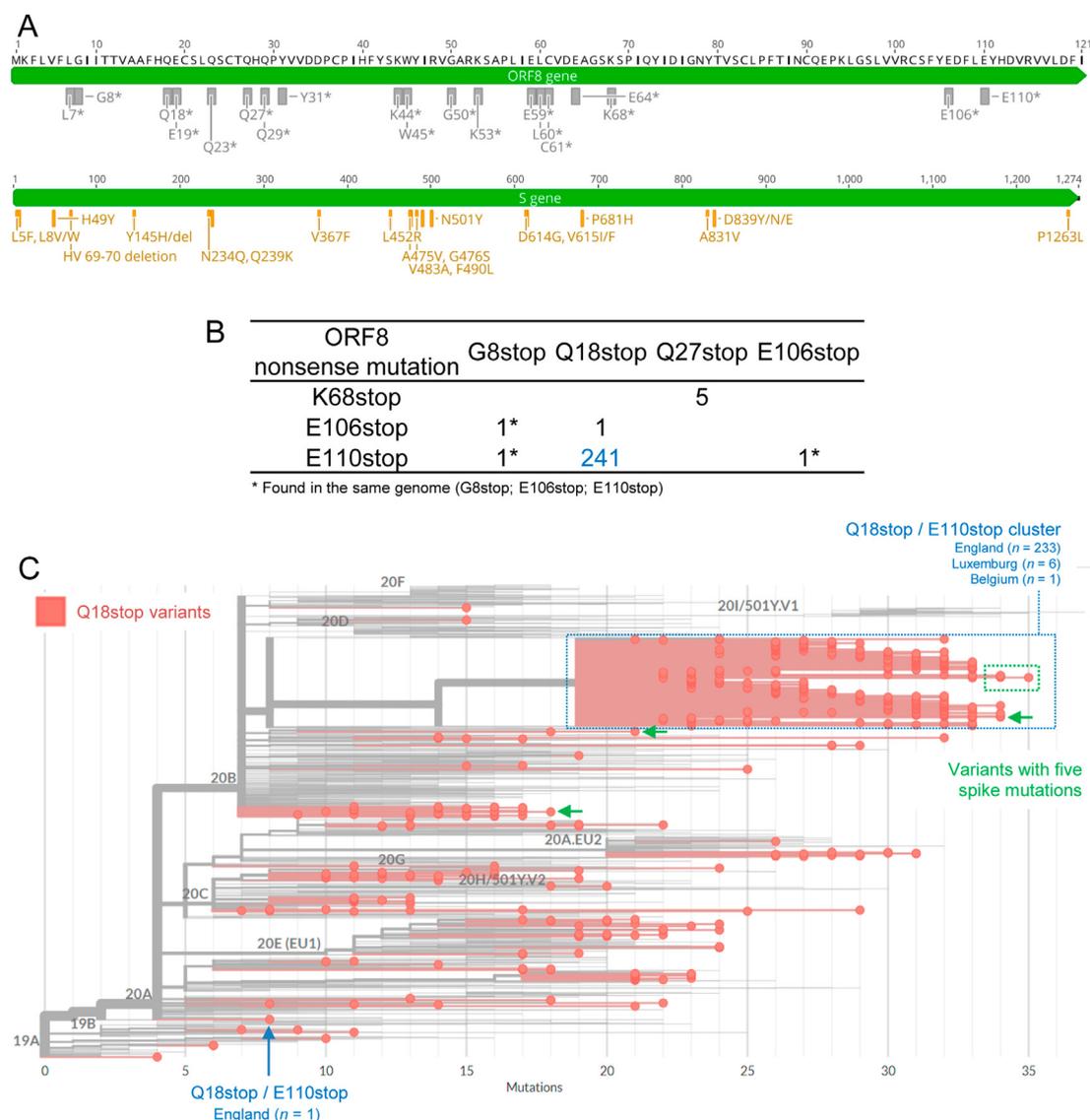The same SARS-CoV-2 variant can hold several spike mutations

Fig. 1. SARS-CoV-2 ORF8 and spike protein mutations. A) Location of mutations in ORF8 and spike proteins analyzed in this study. Nineteen nonsense mutations result in truncated or absent ORF8 proteins in circulating SARS-CoV-2 variants (top scheme). Twenty mutations in the spike protein were selected due to their putative influence in the virus infection and spread capacities (bottom scheme). B) Number of ORF8 nonsense mutations found simultaneously in the same SARS-CoV-2 genome. C) SARS-CoV-2 phylogeny highlighting the ORF8-deficient variants resulting from the nonsense mutation Q18stop. The cluster with variants harbouring both Q18stop and E110stop is identified by a blue box. Variants with five associated spike mutations are shown with green arrows and a box.

in relation to the reference genome, as observed in VOC-202012/01. The highest mean number of spike mutations per genome was observed in ORF8 variants K68stop (mean = 4; SD = 4.43), K53stop (mean = 3; SD = 0), Q27stop (mean = 2.51; SD = 2.75) and G50stop (mean = 2.40; SD = 1.30), as shown in Supplementary Fig. S3. The maximum number of spike mutations in a genome was found associated with Q27stop and K68stop (11 mutations), explained by including the VOC-202012/01. The 11 spike mutations found in both Q27stop and K68stop (I68del, HV69-70del, Y144del, Y145H/del, N501Y, A570D, D614G, P681H, T716I, S982A and D1118H) were found in eight sequences from England and two from the Netherlands (Table 1).

There is a clear gap between VOC-202012/01 and other variants in terms of the number of spike mutations (Supplementary Fig. S3). The next ORF8-deficient variant (E64stop) had six spike mutations (I68del, HV69-70del, L189F, N439K, D614G and V772I), two of them of putative biological relevance (HV69-70del and D614G). The sequences form a small cluster in clade 20A (Fig. 2), being found in

Denmark (n = 8), Germany (n = 2), Switzerland (n = 2) and USA (n = 1). The branch where these sequences occurred is predominantly found in central Europe, suggesting an origin in this area.

Several ORF8-deficient variants were found with five spike mutations. For example, a variant with five spike mutation, three of relevance (L5F, HV69-70del and D614G), was found associated with E59stop in a sample from Poland (EPI_ISL_732828). Ten sequences with five spike mutations were found associated with Q18stop, in three different combinations and branches of the SARS-CoV-2 phylogeny (Fig. 1C). The most divergent sequences within the Q18stop/E110stop cluster had five spike mutations and were detected in England.

4. Conclusion

The ORF8 accessory gene is not crucial for the replication of SARS-CoV-2, as proven by the several variants lacking this gene reported worldwide and studied here. The gene is most likely
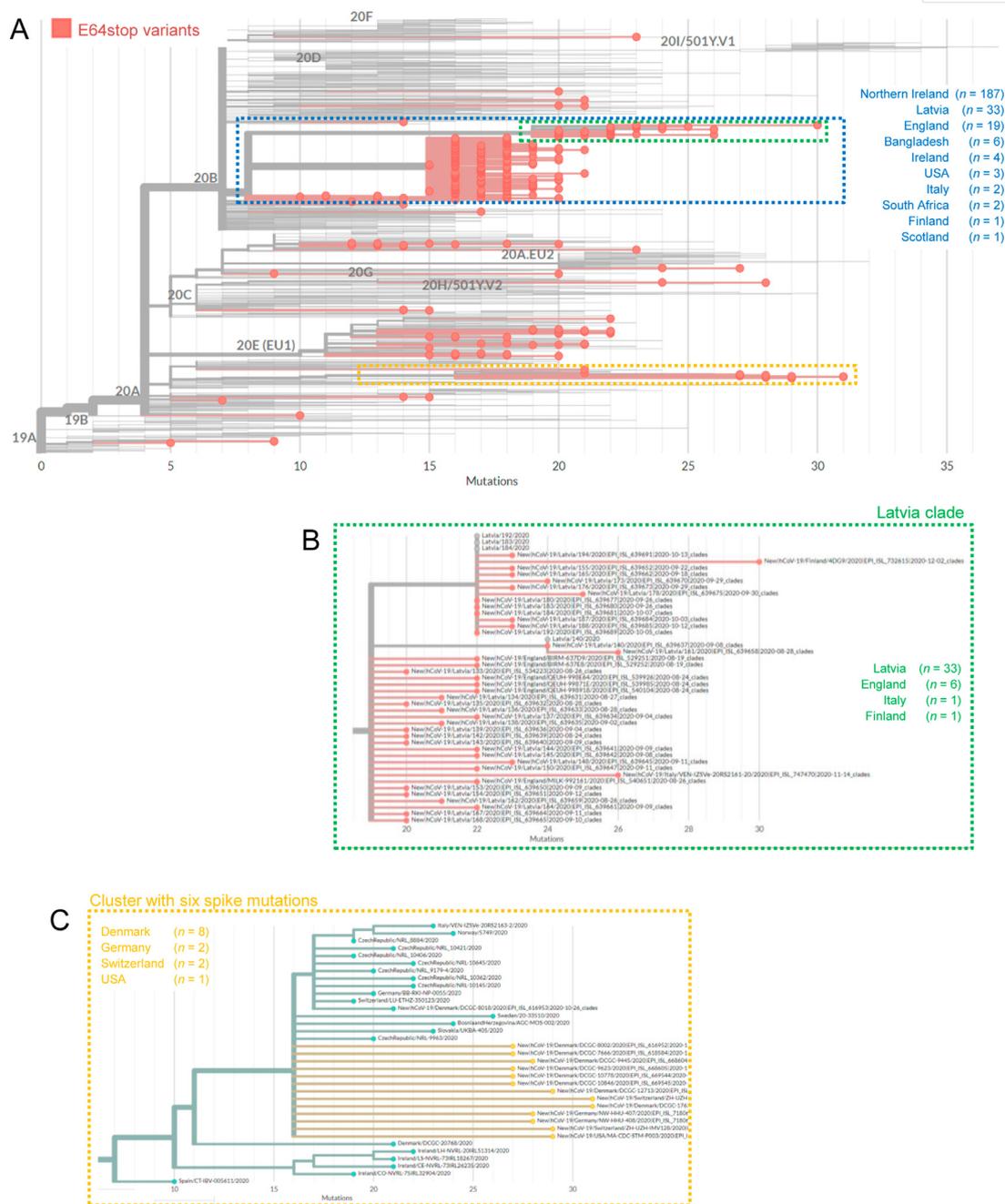
Fig. 2. SARS-CoV-2 phylogeny highlighting the ORF8-deficient variants resulting from the nonsense mutation E64stop. A) The main cluster with E64stop variants in clade 20B is identified by a blue box. B) The branch within this clade prevalent in Latvia is highlighted by a green box. C) Variants with six associated spike mutations located in clade 20A are highlighted by a yellow box.

involved in modulation of the host infected cell metabolism and innate immunity evasion, although the subject is still under investigation [8,9,23]. Removal of ORF8 results in less severe disease with a likely prolonged infection period [10], which may increase the opportunities for contagious and virus transmissibility. Curiously, middle and late phases of the SARS epidemic were characterized by the spread of SARS-CoV with either partial or complete deletions of the ORF8 gene [24,25]. The same may be happening in SARS-CoV-2, as more transmissible lineages may become dominant as the pandemic progresses.

Overall, I found that ORF8 nonsense mutations occur recurrently in the SARS-CoV-2 pandemic. It is difficult to estimate if such variants had any increased transmissibility, as genomes deposited in GISAID are a biased sample of circulating variants. Moreover, spread of a variant depends on many factors beyond the viral genetics, such as restrictions on host mobility, rates of diagnostic testing, clinical management, etc. In any case, I identified two clusters of ORF8-deficient variants that emerged recently in the 20B clade, which did not include the emerging VOC-202012/01 variant. The association of these ORF8-deficient variants with spike mutations of interest (as in VOC-202012/01) should raise concern, as they may be more transmissible. The recent origin of these clusters and the high number of variants detected in several countries should be further investigated.
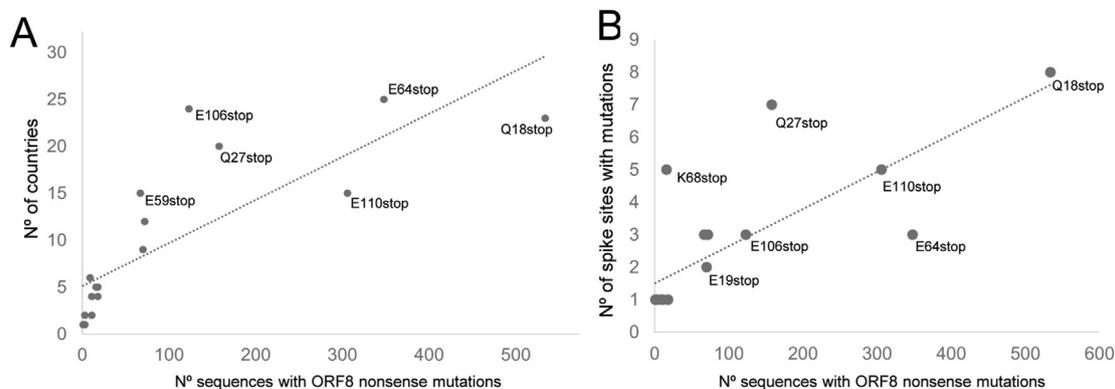
Fig. 3. Distribution of SARS-CoV-2 variants lacking ORF8 and with spike mutations of potential biological relevance. Relationship between the number of sequences with ORF8 nonsense mutations and A) the countries where they have been reported and B) the number of relevant spike mutations they include (a linear trendline is shown for both plots).

In any case, VOC-202012/01 is clearly the ORF8-deficient variant with more associated spike mutations detected so far. It has been speculated that this variant may have emerged in immunodeficient or immunosuppressed patients who are chronically infected with SARS-CoV-2 [5]. Because the ORF8 protein is believed to suppress immune responses, it is possible that the pressure to retain this protein is reduced in immunosuppressed patients. This hypothesis is supported by the recent observation of a new ORF8-deficient lineage (Q18 stop) in a lymphoma patient with long-term COVID-19, also associated with several spike mutations [26]. Further investigation of SARS-CoV-2 in long-term infected patients should be done to better understand the emergence of these variants. By all these reasons, monitoring of SARS-CoV-2 variants with mutations in ORF8 and spike proteins should be a priority in the ongoing COVID-19 pandemic, as they may become dominant. Even though the absence of ORF8 results in milder infections, the increase of transmissibility may result in an overall higher mortality.

Funding

Declaration of competing interestCOI

None to declare.

Acknowledgments

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bbrc.2021.02.080.

References

[1] M. Worobey, J. Pekar, B.B. Larsen, M.I. Nelson, V. Hill, J.B. Joy, A. Rambaut, M.A. Suchard, J.O. Wertheim, P. Lemey, The emergence of SARS-CoV-2 in Europe and North America, Science 370 (2020) 564–570.

[2] N.D. Grubaugh, M.E. Petrone, E.C. Holmes, We shouldn't worry when a virus mutates during disease outbreaks, Nat. Microbiol. 5 (2020) 529–530.

[3] M. Chand, S. Hopkins, G. Dabrera, C. Achison, W. Barclay, N. Ferguson, E. Volz, N. Loman, A. Rambaut, J. Barrett, Investigation of Novel SARS-COV-2 Variant: Variant of Concern 202012/01, Public Health England, 2020.

[4] N.G. Davies, R.C. Barnard, C.I. Jarvis, A.J. Kucharski, J. Munday, C.A. Pearson, T.W. Russell, D.C. Tully, S. Abbott, A. Gimma, Estimated Transmissibility and Severity of Novel SARS-CoV-2 Variant of Concern 202012/01 in England, medRxiv, 2020.

[5] A. Rambaut, N. Loman, O. Pybus, W. Barclay, J. Barrett, A. Carabelli, T. Connor, T. Peacock, D.L. Robertson, E. Volz, Preliminary Genomic Characterisation of an Emergent SARS-CoV-2 Lineage in the UK Defined by a Novel Set of Spike Mutations, COVID-19 Genomics Consortium UK (CoG-UK), 2020.

[6] H. Gu, Q. Chen, G. Yang, L. He, H. Fan, Y.-Q. Deng, Y. Wang, Y. Teng, Z. Zhao, Y. Cui, Adaptation of SARS-CoV-2 in BALB/c mice for testing vaccine efficacy, Science 369 (2020) 1603–1607.

[7] K.R. McCarthy, L.J. Rennick, S. Nambulli, L.R. Robinson-McCarthy, W.G. Bain, G. Haidar, W.P. Duprex, Recurrent deletions in the SARS-CoV-2 spike glycoprotein drive antibody escape, bioRxiv (2020), https://doi.org/10.1101/2020.11.19.389916.

[8] L. Zinzula, Lost in Deletion: the Enigmatic ORF8 Protein of SARS-CoV-2, Biochem. Biophys. Res. Commun. 538 (2021) 116–124.

[9] F. Pereira, Evolutionary dynamics of the SARS-CoV-2 ORF8 accessory gene, Infect. Genet. Evol. 85 (2020) 104525.

[10] B.E. Young, S.-W. Fong, Y.-H. Chan, T.-M. Mak, L.W. Ang, D.E. Anderson, C.Y.-P. Lee, S.N. Amrun, B. Lee, Y.S. Goh, Effects of a major deletion in the SARS-CoV-2 genome on the severity of infection and the inflammatory response: an observational cohort study, Lancet 396 (2020) 603–611.

[11] E. Volz, S. Mishra, M. Chand, J.C. Barrett, R. Johnson, L. Geidelberg, W.R. Hinsley, D.J. Laydon, G. Dabrera, Á. O'Toole, Transmission of SARS-CoV-2 Lineage B. 1. 1. 7 in England: insights from Linking Epidemiological and Genetic Data, medRxiv, 2021, 2020.2012. 2030.20249034.

[12] K. Katoh, J. Rozewicki, K.D. Yamada, MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization, Briefings Bioinf. 20 (2019) 1160–1166.

[13] F. Pereira, SARS-CoV-2 variants lacking a functional ORF8 may reduce accuracy of serological testing, J. Immunol. Methods 488 (2021) 112906.

[14] W.-M. Zhao, S.-H. Song, M.-L. Chen, D. Zou, L.-N. Ma, Y.-K. Ma, R.-J. Li, L.-L. Hao, C.-P. Li, D.-M. Tian, The 2019 Novel Coronavirus Resource, Yi chuan Hereditas 42, 2020, pp. 212–221.

[15] J. Singer, R. Gifford, M. Cotten, D. Robertson, CoV-GLUE: A Web Application for Tracking SARS-CoV-2 Genomic Variation, 2020.

[16] B. Korber, W.M Fischer, S. Gnanakaran, H. Yoon, J. Theiler, W. Abfalterer, N. Hengartner, E.E. Giorgi, T. Bhattacharya, B. Foley, Tracking changes in SARS-CoV-2 Spike: evidence that D614G increases infectivity of the COVID-19 virus, Cell 182 (2020) 812–827, e819.

[17] Q. Li, J. Wu, J. Nie, L. Zhang, H. Hao, S. Liu, C. Zhao, Q. Zhang, H. Liu, L. Nie, The impact of mutations in SARS-CoV-2 spike on viral infectivity and antigenicity, Cell 182 (2020) 1284–1294, e1289.

[18] K. Pancer, A. Milewska, K. Owczarek, A. Dabrowska, M. Kowalski, P.P. Łabaj, W. Branicki, M. Sanak, K. Pyrc, The SARS-CoV-2 ORF10 is not essential in vitro or in vivo in humans, PLoS Pathog. 16 (2020), e1008959.

[19] Y.-N. Gong, K.-C. Tsao, M.-J. Hsiao, C.-G. Huang, P.-N. Huang, P.-W. Huang, K.-M. Lee, Y.-C. Liu, S.-L. Yang, R.-L. Kuo, SARS-CoV-2 genomic surveillance in Taiwan revealed novel ORF8-deletion mutant and clade possibly associated with infections in Middle East, Emerg. Microb. Infect. (2020) 1–37.

[20] Y.C. Su, D.E. Anderson, B.E. Young, M. Linster, F. Zhu, J. Jayakumar, Y. Zhuang, S. Kalimuddin, J.G. Low, C.W. Tan, Discovery and genomic characterization of a 382-nucleotide deletion in ORF7b and ORF8 during the early evolution of SARS-CoV-2, mBio (2020) 11.

[21] Y. Furuse, Genomic sequencing effort for SARS-CoV-2 by country during the pandemic, Int. J. Infect. Dis. 103 (2020) 305–307.

[22] E. Volz, V. Hill, J.T. McCrone, A. Price, D. Jorgensen, Á. O'Toole, J. Southgate,

R. Johnson, B. Jackson, F.F. Nascimento, Evaluating the effects of SARS-CoV-2 Spike mutation D614G on transmissibility and pathogenicity, Cell 184 (2020) 64–75, e11.

[23] T.G. Flower, C.Z. Buffalo, R.M. Hooy, M. Allaire, X. Ren, J.H. Hurley, Structure of SARS-CoV-2 ORF8, a rapidly evolving immune evasion protein, Proc. Natl. Acad. Sci. Unit. States Am. 118 (2021).

[24] C.S.M.E. Consortium, Molecular evolution of the SARS coronavirus during the course of the SARS epidemic in China, Science 303 (2004) 1666–1669.

[25] Y. Guan, B. Zheng, Y. He, X. Liu, Z. Zhuang, C. Cheung, S. Luo, P. Li, L. Zhang, Y. Guan, Isolation and characterization of viruses related to the SARS coronavirus from animals in southern China, Science 302 (2003) 276–278.

[26] G. Bazykin, O. Stanevich, D. Danilenko, A. Fadeev, K. Komissarova, A. Ivanova, M. Sergeeva, K. Safina, E. Nabieva, G. Klink, Emergence of Y453F and Δ69-70HV mutations in a lymphoma patient with long-term COVID-19, virological.org, https://virological.org/t/emergence-of-y453f-and-69-70hv-mutations-in-a-lymphoma-patient-with-long-term-covid-19/580.