The SARS-CoV-2 B.1.351 lineage (VOC $\beta$) is outgrowing the B.1.1.7 lineage (VOC $\alpha$) in some French regions in April 2021
Bénédicte Roquebert, Sabine Trombert-Paolantoni, Stéphanie Haim-Boukobza, Emmanuel Lecorche, Laura Verdurme, Vincent Foulongne, Mircea Sofonea, Samuel Alizon

To cite this version:
Bénédicte Roquebert, Sabine Trombert-Paolantoni, Stéphanie Haim-Boukobza, Emmanuel Lecorche, Laura Verdurme, et al.. The SARS-CoV-2 B.1.351 lineage (VOC $\beta$) is outgrowing the B.1.1.7 lineage (VOC $\alpha$) in some French regions in April 2021. Eurosurveillance, European Centre for Disease Prevention and Control, 2021, 26 (23), 10.2807/1560-7917.es.2021.26.23.2100447. hal-03289627

HAL Id: hal-03289627
https://hal.archives-ouvertes.fr/hal-03289627
Submitted on 17 Jul 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
To assess SARS-CoV-2 variants spread, we analysed 36,590 variant-specific reverse-transcription-PCR tests performed on samples from 12 April–7 May 2021 in France. In this period, contrarily to January–March 2021, variants of concern (VOC) β (B.1.351 lineage) and/or γ (P.1 lineage) had a significant transmission advantage over VOC α (B.1.1.7 lineage) in Île-de-France (15.8%; 95% confidence interval (CI): 15.5–16.2) and Hauts-de-France (17.3%; 95% CI: 15.9–18.7) regions. This is consistent with VOC β’s immune evasion abilities and high proportions of prior-SARS-CoV-2-infected persons in these regions.

‘Variants of concern’ (VOC) are severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) phenotypically distinct lineages that are associated with major epidemiological or clinical shifts. To date, four have been classified as such by the World Health Organization (WHO) [1]. The first, VOC α, which corresponds to Pango lineage B.1.1.7, nextstrain clade 20I/501Y.V1, and GISAID clade/lineage GRY, is currently causing the majority of infections in Europe and North America [2], whereas, the second, VOC β (Pango lineage: B.1.351; nextstrain clade: 20H/501Y.V2; GISAID clade/lineage: GH/501Y.V2) is the most common variant in South Africa [3]. The third variant, VOC γ (Pango lineage: P.1; nextstrain clade: 20I/501Y.V3; GISAID clade/lineage: GR/501Y.V3) dominates in Brazil and South America [4] and the fourth VOC δ (Pango lineage: B.1.617.2; nextstrain clade: 21A/S:478K; GISAID clade/lineage: G/452R.V3) caused a major epidemic wave in India [5]. The outcome of the (indirect) competition between variants is yet open. In France, the early introduction of VOC β in some regions makes it particularly important to monitor the spread of different variants [6].

PCR testing of SARS-CoV-2-positive clinical samples for variants

Since January 2021, the national guideline is to test all clinical samples that are positive for SARS-CoV-2 with an additional reverse-transcription (RT)-PCR to detect mutations indicative of certain variants [7,8]. Since April 2021, this variant-specific RT-PCR targets the N501Y mutation, which is shared by VOCs B.1.1.7, B.1.351 and P.1, and the E484K mutation, which is found in VOCs B.1.351 and P.1, as well as the variant of interest (VOI) B.1.525 (WHO: η; nextstrain clade: 20A/S484K; GISAID clade/lineage: G/484K.V3) [1], but not in B.1.1.7.

We used the ID SARS-CoV-2/N501Y/E484K Quadruplex assay (ID Solution, Grabels, France) to test 53,687 SARS-CoV-2 positive samples collected between 12 April and 7 May 2021 in 13 French regions, with the majority of samples coming from the Île-de-France region (Table 1). Some of the total samples (7–8%, Table 1) originated from hospitals (mostly hospitalised patients), the rest from the general population. We only analysed data from individuals aged from 5 to 80 years to minimise in the analysis the weight of preschool children and elderly persons in long-term care facilities. 17.3% of the tests could not be interpreted; this was mainly because the cycle threshold (Ct) value was too high to ensure an equal sensitivity for the N501Y and E484K targets. To avoid biasing the variant screening, all tests with Ct values strictly above 30, including those where a lineage could be assigned, were ignored (31.8%; 17,097/53,687). Overall, we analysed 68.2% (36,590/53,687) of all the samples tested (Table 1).
The specificity of the variant-specific RT-PCR we used is limited, since this PCR only targets two mutations. To gain additional insights regarding the type of variants circulating in the country, we sequenced 15% the samples collected on 30 March 2021 in France in this dataset for which the Ct was equal or lower than 28 using Twist Libraries and Illumina sequencing; the GISAID accession numbers are in Supplement S1. These samples were constituted for the most part (45%; 215/478) by samples from the Île-de-France region and showed a majority of viruses of the B.1.1.7 lineage (79.1%, 378/478; Supplementary Table S1). The other prevalent lineages were B.1.351 (7.9%, 38/478), B.1.525 (4.4%, 21/478), and B.1.214 (2.3%, 11/478), a lineage characterised by a variant not classified as a VOC, but which is under monitoring [9]. There were also lineages represented by less than 2% of the samples, such as P.1 (0.6%, 3/478). In the subsets of samples from the Île-de-France (n=215) and Hauts-de-France (n=48) regions, the order of prevalence of the VOCs was the same as for the overall samples (Supplementary Table S1). Results from Santé Publique France, the French National Public Health Agency, for the Île-de-France region in April (n=476 samples) also generally agreed with these findings (Supplementary Table S1). Only a few samples from April were sequenced from the Hauts-de-France (n=11) and Île-de-France (n=13) regions and for each of these regions, more than half of the samples were of the B.1.1.7 lineage.

Therefore, hereafter, samples with only the N501Y mutation detected are assumed to contain virus of B.1.1.7 lineage, samples with both N501Y and E484K mutations, mainly virus of the B.1.351 lineage with possibly a minority of P.1, samples with only the E484K mutation, virus of the B.1.525 lineage, and samples with no mutation, wild type SARS-CoV-2 (although these samples may contain also viruses of B.1.214 lineage, which also lack the two mutations).

Analysis of reverse-transcription PCR results

Raw proportions of each SARS-CoV-2 lineage deduced by RT-PCR are shown in Figure 1 and raw numbers are shown in Supplementary Figure S1. As shown in both Figures, the sampling intensity in the dataset varies strongly across regions, which explains that some weeks have extreme values (e.g. in Occitanie). Overall, we see that lineage B.1.1.7 is dominant is most regions, and that the Île-de-France is the region where the B.1.351 and/or P.1 lineages are the most frequently detected.

Ethical statement

This study has been approved by the Institutional Review Board of the CHU of Montpellier and is registered at ClinicalTrials.gov with identifier NCT04738331.

Lineage spreading in France

We used a multinomial log-linear model (the multinom function from the nnet R package [10]) to identify factors associated with the detection of certain lineages (B.1.1.7 being the variant of reference). The explanatory variables were the individual’s age (type of variable: integer), origin of the sample (hospital or community;
**Figure 1**

Raw daily cumulative frequencies of variant-specific reverse-transcription PCR test results for SARS-CoV-2 in eight French regions, 12 April–7 May 2021 (n = 33,583)

<table>
<thead>
<tr>
<th>Region</th>
<th>Test Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Occitanie</td>
<td>(n = 412)</td>
</tr>
<tr>
<td>Provence-Alpes–Côte d'Azur</td>
<td>(n = 3,714)</td>
</tr>
<tr>
<td>Normandie</td>
<td>(n = 6,243)</td>
</tr>
<tr>
<td>Nouvelle–Aquitaine</td>
<td>(n = 1,351)</td>
</tr>
<tr>
<td>Hauts-de–France</td>
<td>(n = 4,593)</td>
</tr>
<tr>
<td>Île-de-France</td>
<td>(n = 14,953)</td>
</tr>
<tr>
<td>Bourgogne–Franche–Comté</td>
<td>(n = 455)</td>
</tr>
<tr>
<td>Centre–Val de Loire</td>
<td>(n = 1,862)</td>
</tr>
</tbody>
</table>

SARS-CoV-2: severe acute respiratory syndrome coronavirus 2; VOC: variant of concern; VOI: variant of interest.

* Only regions with more than 400 respective tests are shown.

* Characterisation as wild type SARS-CoV-2 is based on the absence of both N501Y and E484K mutations.

* Characterisation as B.1.351 and/or P.1 lineage (VOC β and/or γ) is based on the presence of both N501Y and E484K mutations.

* Characterisation as B.1.525 lineage (VOI η) is based on the simultaneous absence of N501Y and presence of E484K mutation.

* Characterisation as B.1.1.7 lineage (VOC α) is based on the simultaneous presence of N501Y and absence of E484K mutation.

The number of tests performed is indicated in each panel. For each day, the different colours indicate the proportion of tests belonging to each of the four screening categories (these sum to 1.0 every day). Regions with few tests can exhibit strong variations in frequencies for some days (e.g. Occitanie for days with only B.1.525 detected). See Supplementary Figure S1 for the raw numbers instead of the proportions.
The multinomial model revealed differences between lineages (Table 2). In terms of age, we found that older patients had a lower risk of being infected by B.1.351/P.1 and B.1.525 than by B.1.1.7 (our reference). In hospital settings, we found an over-representation of B.1.351/P.1 compared with B.1.1.7. When analysing region-specific temporal trends, we found that, for all regions, the risks of being infected by a wild type or a B.1.525 virus were either identical or lower than the risk of being infected by B.1.1.7. Conversely, we found that the risk of being infected by B.1.351/P.1 instead of B.1.1.7 significantly increased with time in Île-de-France, and to a lesser extent in Hauts-de-France and Nouvelle-Aquitaine.

**Transmission advantage of B.1.351/P.1 vs B.1.1.7**

Trends from the multinomial model should be treated with caution because of autocorrelation issues. Therefore, to investigate the temporal trends, we used the method described in [11] and, for each region of interest, fitted a logistic growth model to the fitted values of a generalised linear model (GLM) with three factors on the data sampled outside hospitals. In addition to the sampling date and the individual age, we also added the department (i.e. a within-region administrative unit), where the sample was performed. For simplicity, we tested the transmission advantage of B.1.351/P.1 compared with B.1.1.7 and neglected the other lineages in the analysis. We performed the analysis only in the Île-de-France, Hauts-de-France, and Nouvelle-Aquitaine regions. With few data regarding the coronavirus disease (COVID-19) epidemic serial interval in France, i.e. the time between the onset of the symptoms in an individual and that in a person he/she infects, we used the one from [12].

We found a transmission advantage of 15.8% (95% confidence interval: 15.5–16.2%) in Île-de-France and 17.3% (95%CI: 15.9–18.7%) in Hauts-de-France (Figure 2). In Nouvelle-Aquitaine, the logistic growth model was not significant, which could be due to the fact that this region was less affected by the third epidemic wave than the other two [13].

**Discussion**

When analysing the results of variant-specific tests on samples obtained from January to March 2021, we found that the B.1.1.7’s (VOC α) transmission advantage relative to wild type lineages was larger than
there, because of the high proportion of individuals variants with a transmission advantage is occurring date by the epidemic [13], it is possible that a shift in being one of the French regions the most impacted to evasion properties [15,16]. Therefore, Île-de-France than B.1.1.7. The B.1.351 lineage has known immune β) and possibly P.1 (VOC γ) spreading more rapidly this trend appears to have shifted with B.1.351 (VOC β) lineages; based on the sequencing of a subset of samples in March and April however, the B.1.351 (VOC β) appears to dominate over the P.1 (VOC γ) lineage in France and in the regions considered. Therefore the results are considered as likely reflecting mostly the B.1.351 variant (VOC β) spread.

Each region is represented in a different colour, with the triangles indicating the GLM-fitted values and the line the output of the logistic growth model estimation. The shaded area in the same colour as the line represents the 95%CI. The text indicates the estimated transmission advantage of B.1.351 (VOC β) with respect to B.1.1.7 (VOC α) and the 95% CIs.

that of B.1.351 (VOC β) relative to wild type lineages [14]. During April 2021, in at least two French regions, this trend appears to have shifted with B.1.351 (VOC β) and possibly P.1 (VOC γ) spreading more rapidly than B.1.1.7. The B.1.351 lineage has known immune evasion properties [15,16]. Therefore, Île-de-France being one of the French regions the most impacted to date by the epidemic [13], it is possible that a shift in variants with a transmission advantage is occurring there, because of the high proportion of individuals with immunity acquired through prior-SARS-CoV-2-infections. Vaccination might favour immune escape mutants [17] but the coverage with COVID-19 vaccines is homogeneous among French regions. Our results call for more detailed analyses regarding the link between the transmission advantage of the B.1.351 variant and the proportion of the population with immunity (following infection or vaccination) in different French regions.

There are some limitations to this analysis. First, although we performed sequencing to distinguish between B.1.351 and P.1 lineages, in some samples collected in March and, to a lesser extent, April, further sequencing will be needed to validate our assumption that the transmission advantage belongs to B.1.351, to P.1, or to both. Second, France had entered a third national lockdown on 3 April, which means that most of the tests analysed here were performed in a declining epidemic [18]. If we what assume to be the B.1.351 lineage causes infections that have a shorter generation interval than the B.1.1.7 lineage, this could affect the transmission advantage estimates. While there are some data on generation intervals for COVID-19 epidemics in France [19], studies are so far limited. Furthermore, analyses performed on the detailed United Kingdom epidemic data found that the hypothesis of differences in generation interval between B.1.1.7 and wild type lineages was less likely than other hypotheses, especially differences in contagiousness [2]. Finally, it is unlikely that non-pharmaceutical interventions would affect differently the transmission of the variants.

In conclusion, given the progressive lifting of the control measures in June 2021 in France [18], these results call for particular care regarding vaccination rollout and the maintenance of non-pharmaceutical prevention until vaccine coverage reaches levels compatible with spontaneous regression of the epidemic.

Acknowledgements
We thank the ETE modelling team for discussion, as well as the CNRS, the IRD, the ANR, and the Région Occitanie for funding (PHYEPI grant). We thank the Infectious diseases department in Cerba laboratory, and the technical and IT teams for their technical assistance.

Funding statement: We thank the CNRS, the IRD, the ANR and the Région Occitanie (PHYEPI project) for funding.

Conflict of interest
None declared.

Authors’ Contributions
Bénédicte Roquebert, Sabine Trombert-Paolontani, Stéphanie Haim-Boukobza, Emmanuel Lecorche, Vincent Foulonoge, and Laura Verdurme collected the RT-PCR data. Mircea T. Sofonea and Samuel Alizon performed the statistical analysis. Samuel Alizon wrote the first version of the article. All the authors contributed to the final version of the manuscript.

References


License, supplementary material and copyright
This is an open-access article distributed under the terms of the Creative Commons Attribution (CC BY 4.0) Licence. You may share and adapt the material, but must give appropriate credit to the source, provide a link to the licence and indicate if changes were made.

Any supplementary material referenced in the article can be found in the online version.