Phylodynamics of SARS-CoV-2 transmission in France, 2020

CROI – Conference on Retroviruses and Opportunistic Infections – February 12-16, 2022

Romain Coppée[1], François Blanquart[1], Aude Jary[3], Valentin Leducq[3], Anna Maria Franco Yusti[1,2], Lena Daniel[1,2], Charlotte Charpentier[1,2], Karen Zafilaza[3,4], Vincent Calvez[3,4], Diane Descamps[1,2], Anne-Geneviève Marcellin[3,4], Benoit Visseaux[1,2], Antoine Bridier-Nahmias[1]

[1] Université de Paris, Infection Modelisation Antimicrobial Evolution (IAME), Inserm UMR1137, Paris, France; [2] Service de Virologie, Hôpital Bichat Claude Bernard, AP-HP, Paris, France; [3] Sorbonne Université, Institut Pierre Louis d'Epidémiologie et de Santé Publique (iPLESP), AP-HP, Hôpital Pitié-Salpêtrière, Paris, France; [4] Service de Virologie, Hôpital de la Pitié-Salpêtrière, AP-HP, Paris, France

BACKGROUND

Story of COVID-19 epidemic waves in France during 2020 (Figure 1)

December 1, 2019	Outbreak of severe respiratory disease was identified in the city of Wuhan : severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)
January 24, 2020	First SARS-Cov-2 cases from Chinese travelers in France
Mid-March to mid-May	First French epidemic wave; Government applied a complete lockdown across the country
Mid-May to early-October	Measures were progressively relaxed; Restart of some business activities
Mid-October to mid-December	Second French epidemic wave; Second complete lockdown across the country
December 27, 2020	Vaccination campaign against SARS-CoV-2 started in France

Objective ➔ Elucidating the SARS-CoV-2 dynamic throughout the two French waves of the pandemic

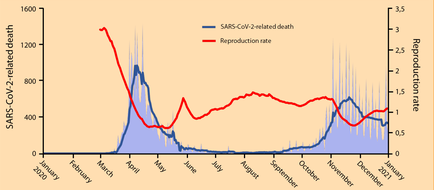


Figure 1 – Timeline of SARS-CoV-2-related deaths and reproduction rate in France, 2020. The two first French epidemic waves are respectively dated of March to late end-June 2020, and September to December 2020. SARS-CoV-2-related deaths are displayed as the daily number of deaths (light blue area) and as the mean of deaths in a week (curve).

METHODS

SARS-CoV-2 sequences from GISAID were included when meeting the following criteria:

1) complete viral genomes from infected individuals
2) maximal proportion of undetermined nucleotide bases = 5%
3) sequences with complete sampling locations
4) sampling dates comprised between January and December 2020

N = 604,568 sequences

Three geographic scales were investigated:
1) Continents + France
2) Europe + France
3) French administrative regions

} Independently investigated at the two French epidemic

For each investigated geographic scale and wave:

Analysis independently repeated 500 times

- ~1,000 sequences representative of each territory (using the number of SARS-CoV-2-related deaths per region and week. Sequences were drawn randomly from the GISAID extraction
- When a set was constituted, all the sequences were replaced in the whole dataset for the random selection of the next one
- Multiple sequence alignments and phylogenetic inference using IQTree; time to the most recent common ancestor estimated by computing the root-to-tip temporal regression using LSD2
- Ancestral state reconstructions of discrete, geographical data were performed with the ace function of ape R package. For each node was assigned the territory supported by ≥ 50% of the state assignment. Inter- and intra-transmissions were counted using breadth first search through a home-made R script (**Figure 2**).

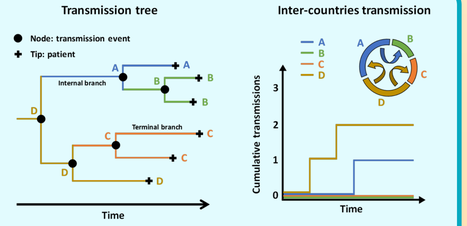


Figure 2 – Conceptual overview of phylodynamics analysis. Each tip of a phylogenetic tree corresponds to a patient infected by SARS-CoV-2 in a given geographical location and at a given date. Through maximum likelihood, each node may be assigned with a geographical location and each branch from this node correspond to a transmission event within this location, if the same location is also observed at the next node, or, in the contrary, from the ancestral node location to the following node location; the sum of branch lengths to root may be used to calculate the date of transmission.

RESULTS

Validation of the datasets:

- Each set, generated at different geographic scales and epidemic waves, was positively correlated to the theoretical SARS-CoV-2-related death ($p < 0.001$; Spearman's rank correlation)
- The number of sequences per territory was properly, temporally distributed within each wave
- The genomic diversity of SARS-CoV-2 in circulation in the different territories was found to be globally uniform

Which continents exchanged SARS-CoV-2 with France?

- Intra-transmissions counted for approximately 80% of the transmission events. This indicator was positively correlated to the theoretical number of deaths related to SARS-CoV-2 in each continent ($p < 0.001$ in both waves; Spearman's rank correlation), validating the model.

- For both French epidemic waves, Europe highly transmitted the virus to France and in the world.
- The second continent that transmitted SARS-CoV-2 to France was North America. South America, where the number of cases increased only after the first wave, lowly transmitted the virus.
- France was a main donor of SARS-CoV-2 to Europe at both waves, and highly transmitted the virus.

How the virus spread in Europe?

- <u>During the first wave</u>, we estimated fewer viral import events than export events for France, Ireland, Sweden and United Kingdom.
- France was the country that transmitted the most the virus to United Kingdom and Germany.
- United Kingdom was the European country that transmitted the most SARS-CoV-2 to France (and the main donor of SARS-CoV-2 in Europe), followed by Spain and Italy (**Figure 3A**).
- Germany was the lowest contributor for SARS-CoV-2 transmission to France (and to other European countries) (**Figure 3A**).

- <u>During the second wave</u>, we found less viral import events than export events for Greece, Russia and United Kingdom.
- United Kingdom remained the major source of SARS-CoV-2 transmissions in Europe, followed by Russia and Germany. Those countries, along with Italy, were the main SARS-CoV-2 transmitters to France.
- France actively participated in SARS-CoV-2 transmissions, being the second largest source of virus introductions in Belgium and the Netherlands.

The first and second French epidemic waves: two patterns of transmissions

- <u>During the first wave</u>, Ile-de-France (IDF; Paris area) highly transmitted SARS-CoV-2 to other French administrative regions, followed in this matter by Auvergne-Rhône-Alpes (ARA; Lyon area).
- Provence-Alpes-Côte d'Azur (PACA; Marseille area) also participated in virus dissemination although at lower rates. SARS-CoV-2 inter-transmission events were more limited for other regions (**Figure 3B**).

- <u>During the second wave</u>, SARS-CoV-2 inter-transmission events were not mostly dictated by IDF, ARA and PACA, that both participated concomitantly in virus spread within France.
- The other French regions participated only lowly to the French inter-region SARS-CoV-2 transmissions, compared to IDF, PACA and ARA. (**Figure 3B**).

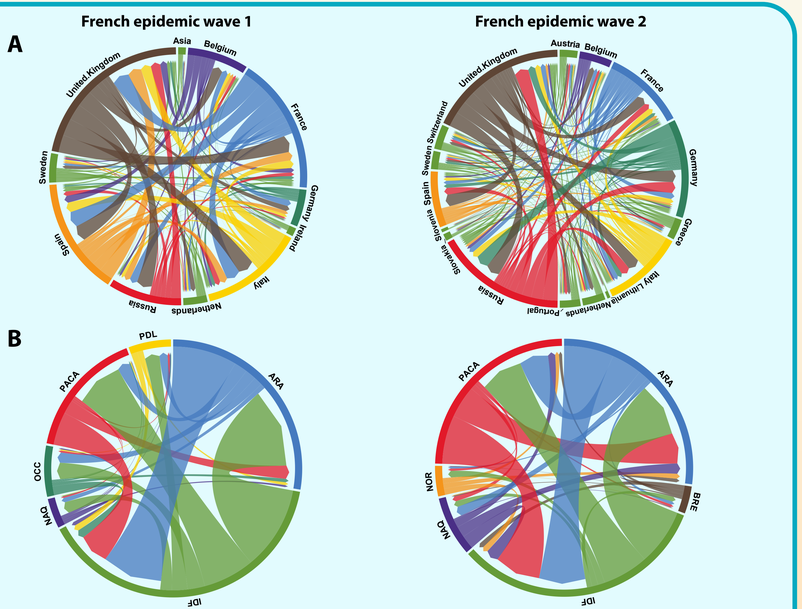


Figure 3 – Circular migration flow plot for each French epidemic wave, at the European (**A**) and French administrative regions (**B**) scales. In these plots, migration flow out of a particular location starts close to the outer ring and ends with an arrowhead more distant from the destination location. Each territory is represented with a specific color and labeled. Migration flow out correspond to the mean of the 500 independent sets of ~1,000 sequences

CONCLUSION

- It is impossible to perform such phylogenetic reconstruction with more than 600 000 sequences. On the other hand, using only 1 to 4 000 sequences can allow to accurately depict the viral diversity but is not enough for performing accurate phylodynamic reconstructions.
- The method proposed here has been specifically designed for tackling such big data-associated challenge.

- France was one of the largest contributor to cross-countries SARS-CoV-2 transmissions at the European scale.
- Germany lowly transmitted the virus during the first wave, suggesting that sanitary measures taken by the German government limited inter-transmission of the virus, at least within Europe.

- At the French administrative regions scale, SARS-CoV-2 inter-region transmission events showed two distinct patterns between the first and second French epidemic waves. Likely, this may be attributable to journey limitation within Europe during summer vacations.