THE LANCET
Infectious Diseases

Supplementary appendix 2

This appendix formed part of the original submission and has been peer reviewed. We post it as supplied by the authors.

Supplement to: Delaugerre C, Foissac F, Abdoul H, et al. Prevention of SARS-CoV-2 transmission during a large, live, indoor gathering (SPRING): a non-inferiority, randomised, controlled trial. *Lancet Infect Dis* 2021; published online Nov 26. https://doi.org/10.1016/S1473-3099(21)00673-3.

SPRING trial appendix

1) Methods

Inclusion/non–inclusion criteria

Inclusion criteria

- Adults 18–45 years old
- No symptoms or contact with people with COVID-19 for 2 weeks
- Negative rapid antigen diagnostic test (RADT)-screening within 3 days before the gathering
- People who declared not having any serious risk factor [1]
- People who declared not living under the same roof as a person with such these factors
- Residing in the Paris region

Non-inclusion criteria

- Positive RADT test within 3 days before the gathering
- People with clinical signs suggestive of an infectious respiratory disease
- People with a risk factor for severe COVID-19
- People living with someone with risk factors for severe COVID-19
- Persons not covered by French National Health Insurance
- Someone who cannot stand for the duration of the experiment (about 5 hours from entry line to exit)
- Person under legal guardianship
- Pregnant woman or woman orally declaring nonuse of effective contraception
- Breastfeeding woman

Virology methods

Nasopharyngeal RADT

Nasopharyngeal antigen-testing used the Standard Q COVID-19 Ag test (SD Biosensor®, Chuncheongbuk-do, Republic of Korea, Roche Diagnostics), a chromatographic immunoassay for the detection of SARS-CoV-2 nucleocapsid (N) antigen. The result was read 15–30 minutes later, according to the manufacturer's instructions.

Saliva reverse-transcriptase–polymerase-chain-reaction (RT-PCR)

Participants were asked to self-collect a saliva sample after salivating for 30 seconds and spitting into the tube. The SARS-CoV-2 genome in fresh saliva was amplified, as previously described. [2]

Briefly, a 300-µl aliquot of saliva was mixed with 300 µl of NucliSENS® lysis buffer (Biomerieux, Marcy l'Etoile, France). Nucleic acids were extracted with the MGIEasy Nucleic Acid Extraction Kit (MGI Tech Co, Shenzhen, China) using an MGISP-960 instrument (MGI Tech Co). SARS-Cov-2 RNA was amplified using TaqPath™ COVID 19 CE IVD RT PCR Kit (Thermo Fisher Scientific, Coutaboeuf, France). The technique provides results expressed as a cycle threshold (Ct) for each targeted gene (*ORF1ab*, *N* and *S* genes). The cutoff value for RT-PCR Ct-positivity was <37 with the TaqPath™ COVID-19 CE IVD RT PCR Kit.

Whole-genome sequencing

The whole SARS-Cov-2 genome was sequenced on RT-PCR–positive saliva with at least two amplified targets to determine the variant(s) for both participant groups and transmission cluster(s) for the attendees. QIAseq SARS-Cov-2 Primer Panel (Qiagen, Hilden, Germany) was used to create tiled amplicons across the SARS-Cov-2 genome. DNA libraries were prepared using Nextera XT and

sequenced using Illumina Miseq with a 300 v2 cartridge (Illumina, San Diego, CA, USA). Consensus genomes were generated using Geneious prime [3] using NC_045512 as the reference genome. Genomes with <90% reference coverage were removed from the analysis. Consensus genomes were aligned using MAFFT [4] with 316 sequences from GISAID collected in the Paris region between 31/05/2021 and 08/06/2021. Best-fit nucleotide substitution GTR+G+I was applied to the datasets using model selection in IQ-Tree2, [5] followed by maximum likelihood phylogenetic tree construction with 1,000-bootstrap replicates. The ggtree R package was used for tree visualization. Lineages were assigned by NextClade, using pangolin and Clade, and confirmed with visual inspection.

Mask-wearing compliance

Mask-wearing compliance was estimated by processing video recordings with a computer algorithm developed by Datakalab® (https://www.datakalab.com/).

Video recordings

Throughout the concert, video recordings were captured by five different cameras, located in the arena along the path from the main entrance to arena floor.

- One camera in the lobby during the security check, where half of the attendees were analyzed, with the other half being much farther away from this camera and, therefore, not taken into account.
- One camera on top of the main staircases leading to the arena floor.
- After the main stairway, attendees were split into two groups going through different corridors. A camera was set to record one of these corridors, during attendee entrance and exit.
- Two cameras on top of the stage recorded two different locations on the arena floor (about 400 attendees on each camera).

Three cameras had a resolution of 1920 × 1080 pixels, two others, located above the stage, had resolution of 3840 × 2160 pixels to obtain better image quality and mitigate the impact of lack of or changing lighting on analysis sensitivity.

Algorithm

The algorithm, based on the use of convolutional neural networks, [6-9] first detected faces in each video frame, then predicted whether these faces were wearing masks, yielding three predictions:

- “no mask”: the face is not wearing a mask, or wears it under the chin
- “inadequate”: the face wears a mask that does not cover the nose
- “adequate”: the face wears a mask that covers the mouth and nose.

Actually, a fourth prediction, i.e., “unassigned”, can be designated when the model is unsure of the prediction to be emitted for a specific person. For instance, if the person wears their mask on the nose, somewhere between “adequate” and “inadequate”, then the model may be unsure about which prediction to assign, and will label that face as “unassigned”. Unassigned predictions were excluded from the analysis of the corridors and the arena floor.

Algorithm performances

The algorithm’s sensitivity and specificity were calculated on a 10-minute sample from two different cameras: two in the lobby/staircase leading to the arena floor, the others above the stage recording the attendees. Each camera provides ~400 faces. Algorithm performances were estimated by counting the number of faces and manually labeling them, then comparing the latter to algorithm predictions. A confusion matrix was then established and used for the estimation.

Algorithm sensitivity was estimated at 98.3% for videos from cameras located on the way to the arena floor, and 97.7% for cameras recording the attendees during the concert. The slight decrease of the sensitivity is due to the latter's camera conditions made more difficult by lack of suddenly changing lighting.
The specificity was estimated at 99.8%, regardless of the camera used for the recording. In other words, almost every single person not wearing a mask (or wearing it inadequately) was correctly predicted as a "no mask".

Sample size calculation

In September 2020, the French government defined a SARS-CoV-2 contamination threshold of 250 per 100000 inhabitants over a sliding week as the maximum alert threshold per department. An incidence above the threshold leading to an enforcement of barrier and social distancing measures, such as: establishment of a curfew, closure of bars, restaurants and non-essential shops and generalization of remote working. We have therefore defined a threshold of 200 per 100000 inhabitants as the maximum incidence to maintain the holding of concerts. Considering an incidence of 200 per 100000 inhabitants in the control group and a maximum arena floor capacity of 5000 participants, and in comparison with superspreading events ranging from 10 to 200 cases (manuscript references 3-11), we stated that the non-inferiority should be ruled out if more than 5 additional cases are observed in the experimental arm compared to the control group. In this context, the non-inferiority margin was settled to a 0.35%.
Sample sizes of 4500 in the experimental group and 2250 in the control group achieve 85% power to detect a non-inferiority margin difference between the group proportions of 0.0035. The control group SARS-CoV-2 proportion is 0.002. The experimental group proportion is assumed to be 0.0055 under the null hypothesis of inferiority. The power was computed assuming that the experimental group proportion is 0.002. The test statistic used is the one-sided Score test (Miettinen & Nurminen). The significance level of the test was targeted at 0.025. The sample size calculation was performed using PASS statistical software (NCSS LLC., Kaysville, UT, USA). By anticipating 10% loss-to-follow-up rate for the primary endpoint, it was planned to randomise 5000 attendees and 2500 non-attendees."

Additional statistical analyses

Intention-to-treat analysis
Analysis of the primary endpoint was carried out on the ITT population, which included all randomized participants with or without protocol deviations. Multiple imputations were obtained to handle missing data using the "MICE" package. Non-inferiority analyses were repeated on five imputed datasets (**Table S3**).

Sensitivity analyses

Sensitivity analyses on different population sets

First, sensitivity analyses were done to assess robustness of the results on four different population sets (**table S4**):

- *Modified intention-to-treat (mITT)*
  This population included all randomized participants for whom the D7-saliva RT-PCR results

were available. All participants were analyzed in their assigned arm. Randomized experimental arm participants who did not come to the live concert were kept in the analysis as part of the attendees.

- *Per-protocol 2*
  This population included all randomized participants meeting the eligibility criteria and for whom D7-saliva RT-PCR results were available (saliva swab collected between D6 and D15). Randomized experimental arm participants who did not come to the live concert were assigned to the control arm.

- *Per-protocol 3*
  This population included all randomized participants meeting the eligibility criteria, and for whom D0- and D7-saliva RT-PCR results were available (D0 saliva RT-PCR results available before D7; D7 saliva swab collected between D6 and D15). Randomized attendees who did not attend the live concert were assigned to the control arm. It was planned to exclude from this population all participants (attendees or nonattendees) with a positive D0-saliva RT-PCR.

- *Per-protocol 4*
  This population included all randomized participants who met the eligibility criteria and for whom D7-saliva RT-PCR results were available and within a restricted window of D6–D10 for collecting the saliva sample.

Sensitivity analyses on the definition of RT-PCR–positivity

Finally, three additional sensitivity analyses were run on the complete case population (per-protocol) using a modified definition of a positive RT-PCR result (**table S5**).

Number of targets
A sensitivity analysis was run considering participants with a RT-PCR saliva test positive for at least 2 of the 3 (*N*, *S* or *ORF1ab*) gene targets.

SARS-CoV-2–load threshold
A sensitivity analysis was run selecting a cycle threshold (Ct) of 28 for distinguishing a target (one or two genes) as positive or not:
Ct≥28: negative target (moderate or low load)
Ct<28: positive target (high load).

Supplemental references

ICH E9 statistical principles for clinical trials (https://www.ema.europa.eu/en/ich-e9-statistical-principles-clinical-trials)

Snapinn SM. Noninferiority trials. Curr Control Trials Cardiovasc Med. 2000 (PMID: 11714400) FDA Non-inferiority clinical trials to establish effectiveness: guidance for industry (https://www.fda.gov/regulatory-information/search-fda-guidance-documents/non-inferiority-clinical-trials

Rabe et al. Missing data handling in non-inferiority and equivalence trials: A systematic review. Pharm Stat. 2018 (PMID: 29797777).

2) Results

Table S1. Baseline characteristics of the intention-to-treat population according to randomisation arm.

Characteristic	Attendees (N=4451)	Non-attendees (N=2225)	All Participants (N=6676)
Age			
Median (IQR) — years	28 (23, 34)	27 (23, 34)	27 (23, 34)
Sex — no. (%)			
Female	2579 (57.9)	1289 (57.9)	3868 (57.9)
Male	1872 (42.1)	936 (42.1)	2808 (42.1)
Declared history of COVID-19 — no. (%)			
No	3673 (82.5)	1824 (82)	5497 (82.3)
Yes	778 (17.5)	401 (18.0)	1179 (17.7)
Vaccinated — no. (%)			
No vaccine	2212 (49.7)	1119 (50.3)	3331 (49.9)
First dose	1906 (42.8)	958 (43.1)	2864 (42.9)
Two doses	333 (7.5)	148 (6.7)	481 (7.2)
Median (IQR) inclusion-to-first-dose interval — days	14 (8, 27)	14 (7, 25)	14 (7, 26)
Median (IQR) inclusion-to-second-dose interval — days	26 (16, 43)	27 (18, 55)	26 (16, 46)

IQR, interquartile range

Table S2. Clinical characteristics and test results of participants* with positive RT-PCR at any time

Arm/Part.no	Attended,	Age	Sex	Vaccine	Last vaccine-	RADT-	Saliva RT-PCR results					Symptomatic	
	Yes/No			doses received	dose-to-D0, days	to-D0, days	Day	*N* gene (Ct)	*S* gene (Ct)	*ORFab* gene (Ct)	SARS-CoV-2 variant	Yes/No	1st date
RT-PCR–positive on D7													
Exp/Part.1	Yes	25	F	0	–	1	D0	29.59	NA	28.92	Alpha (B.1.1.7)	No	–
							D7	26.4	NA	26.6		No	–
Exp/Part.2	Yes	24	F	0	–	2	D0	NA	NA	33.93	NA	No	–
							D7	33.37	NA	33.99	NA	No	–
Exp/Part.4	Yes	25	M	0	–	2	D0	30.27	NA	31.38	NA	No	–
							D7	28.17	NA	30.75	NA	No	–
Exp/Part.5	Yes	26	M	0	–	2	D0	31.01	NA	32.71	NA	No	–
							D7	28.64	NA	29.88	NA	No	–
Exp/Part.7	Yes	43	M	0	–	2	D0	Negative				No	–
							D7	26.33	23.66	24.86	20A/440K (B.1.619)	Yes	05-june
Exp/Part.8	Yes	32	F	1	>15	3	D0	Negative				No	–
							D7	27.73	25.91	26.81	Beta (B.1.351)	Yes	02-june
Exp/Part.11	Yes	25	M	0	–	3	D0	Negative				No	–
							D7	24.67	NA	No	Alpha (B.1.1.7)	No	–
Exp/Part.13	Yes	20	F	1	<14	1	D0	29.21	NA	No		No	–
							D7	29.21	NA	28.99	Alpha (B.1.1.7)	Yes	30-may
Control/ Part.14	No	32	M	0	–	2	D0	30.97	NA	32.25	NA	Yes	
							D7	33.93	NA	33.34	NA	Yes	17-may
Control/Part.17	No	27	F	1	>15	3	D0	Negative					
							D7	22.58	NA	22.49	Alpha (B.1.1.7)	Yes	29-may
Control/ Part.20	No	27	M	1	<14	2	D0	Negative					
							D7	32.78	NA	31.97	NA	Yes	01-june
RT-PCR–positive only on D0													
Exp/Part.3	Yes	37	M	0	–	1	D0	NA	NA	35.84	NA	–	–
							D7	No sample					
Exp/Part.6	Yes	29	F	1	–	3	D0	33.25	NA	33.24	NA	Yes	21-may
							D7	Negative					
Exp/Part.9	Yes	18	M	0	–	2	D0	29.33	NA	33.27	NA	No	–
							D7	Negative					
Exp/Part.10	Yes	24	F	1	<14	3	D0	30.82	NA	32.37	NA	No	31-may
							D7	Negative					
Exp/Part.12	Yes	21	M	0	–	1	D0	29.98	NA	31.53	NA	No	07-may
							D7	Negative					
Control/ Part.15	No	35	M	0	–	3	D0	33.84	NA	38.66	NA	No	–
							D7	Negative					
Control/ Part.16	No	40	F	0	–	1	D0	NA	NA	36.17	NA	No	–
							D7	Negative					
Control/ Part.18	No	41	M	1	<14	3	D0	NA	NA	35.29	NA	No	–

							D7	Negative					
Control/ Part.19	No	19	F	1	>15	2	D0	32.86	NA	32.33	NA	No	–
							D7	Negative					

*All had negative, preconcert screening with a rapid antigen diagnostic test (RADT). ID, participant number; Exp, experimental; Ct, cycle threshold; D, day, NA, not amplified

Table S3 : Salivary carriage at D0 and D7

		Attend live concert		**Not attend live concert**		**Total**	
		PCR saliva result at D7		PCR saliva result at D7		PCR saliva result at D7	
		+	-	+	-	+	-
PCR saliva result at D0	+	5	4	1	4	6	8
	-	3	3819	2	1894	5	5713

Table S4. Intention-to-treat (ITT) non-inferiority analyses of five Imputed datasets

ITT imputed-population	Attendees (N=4451)	Non-attendees (N=2225*)	Rate difference	Incidence rate ratio
Infections	***No. (%) [95% CI]***		**[95% CI]**	**[95% CI]**
Dataset 1	9 (0.20%) [0.09 to 0.38]	4 (0.18%) [0.05 to 0.46]	0.02% [–0.27 to +0.23]	1.11 [0.37 to 3.44]
Dataset 2	9 (0.20%) [0.09 to 0.38]	3 (0.13%) [0.03 to 0.39]	0.07% [-–0.21 to +0.28]	1.54 [0.44 to 5.12]
Dataset 3	10 (0.22%) [0.11 to 0.41]	5 (0.22%) [0.07 to 0.52]	0.00% [–0.32 to +0.23]	1.00 [0.36 to 2.79]
Dataset 4	8 (0.18%) [0.08 to 0.35]	5 (0.22%) [0.07 to 0.52]	–0.05% [–0.36 to +0.17]	0.82 [0.36 to 2.32]
Dataset 5	8 (0.18%) [0.08 to 0.35]	3 (0.13%) [0.03 to 0.39]	0.05% [–0.23 to +0.24]	1.38 [0.38 to 4.63]

*Among the 2227 randomized participants in the control group, 2 withdrew their consent form and data, and were consequently excluded from the non-inferiority intention-to-treat analyses

Table S5. Non-inferiority sensitivity analyses for the modified intention-to-treat, and per-protocol 2, 3 and 4 populations

Population	Attendees	Non-attendees	Rate difference
	N; no. (%) [95% CI]		**[95% CI]**
Modified intention-to-treat	4138; 8 (0.19%) [0.08 to 0.38]	1959; 3 (0.15%) [0.03 to 0.45]	0.04% [–0.27 to +0.25]
Per-protocol 2	3917; 8 (0.20%) [0.09 to 0.40]	2149; 3 (0.14%) [0.03 to 0.41]	0.06% [–0.22 to +0.29]
Per-protocol 3	3821; 3 (0.08%) [0.02 to 0.23]	1895; 2 (0.11%) [0.01 to 0.38]	–0.03% [-0.31 to +0.14]
Per-protocol 4	3850; 8 (0.21%) [0.09 to 0.41]	1911; 3 (0.16%) [0.03 to 0.46]	0.05% [–0.27 to +0.28]

Table S6. Non-inferiority sensitivity analyses according to the RT-PCR test-positive definition for the all participants (per-protocol)

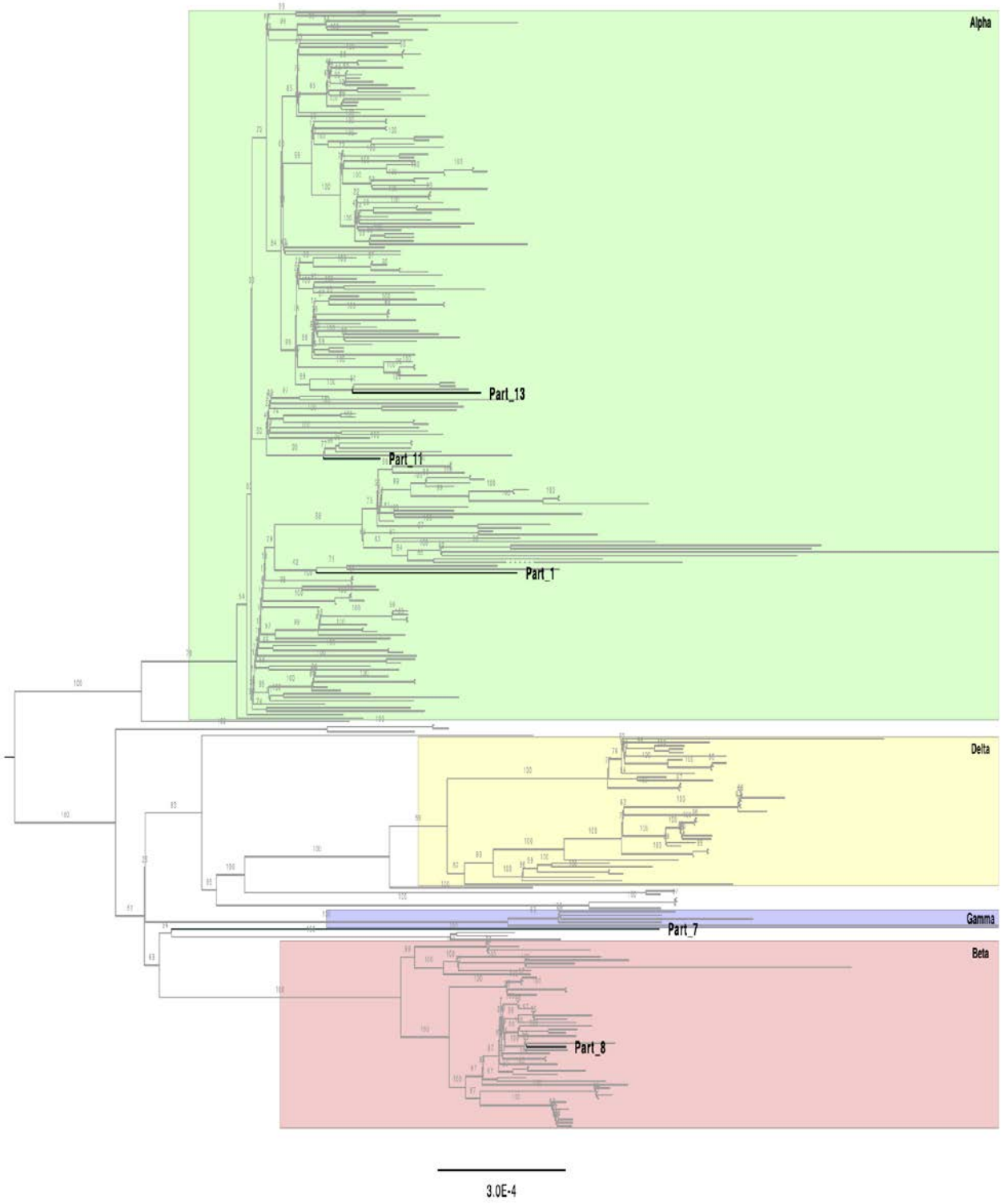
Criteria of positivity	Attendees (N=3917)	Non-attendees (N=1947)	Rate difference
	No. (%) [95% CI]		**[95% CI]**
Ct<37; 2 genes	8 (0.20%) [0.09 to 0.40]	3 (0.15%) [0.03 to 0.45]	0.05% [–0.26 to +0.28]
Ct<28; 1 gene	5 (0.13%) [0.04 to 0.30]	1 (0.05%) [0.00 to 0.29]	0.08% [–0.17 to +0.26]
Ct<28; 2 genes	5 (0.13%) [0.04 to 0.30]	1 (0.05%) [0.00 to 0.29]	0.08% [–0.17 to +0.26]

Viral subtyping and cluster transmission analysis

To identify the SARS-CoV-2 subtype, 20 positive samples with at least two amplified virus genes, regardless of cycle threshold (Ct), underwent whole-genome-sequencing. Sequencing was successful for six (5 attendees and 1 non-attendee) and failed for the 14 others (all Ct > 30). The SARS-CoV-2 variants were lineage and clade classified as four Alpha (B.1.1.7), one Beta (B.1.351) and one 20A/440K. The phylogenetic tree confirmed absence of transmission cluster during the gathering (appendix p10, figure S1).

Figure S1. Phylogenetic analysis of transmission clusters

Maximum likelihood phylogenetic tree of SARS-Cov-2 whole-genome sequences from participants along with 316 GISAID sequences from the Paris Region collected during the study period. Participants' sequences available from are highlighted in black bold. The phylogeny was estimated with Iqtree on 10,001,000-bootstrap replicates.


Alpha
Part_13
Part_11
Part_1
Delta
Gamma
Part_7
Beta
Part_8
3.0E-4

References

1. Avis provisoire Recommandations relatives à la prévention et à la prise en charge du COVID-19 chez les patients à risque de formes sévères [Internet]. [cited 2021 Aug 4];Available from: https://www.hcsp.fr/Explore.cgi/AvisRapportsDomaine?clefr=775

2. Kernéis S, Elie C, Fourgeaud J, et al. Accuracy of saliva and nasopharyngeal sampling for detection of SARS-CoV-2 in community screening: a multicentric cohort study. Eur J Clin Microbiol Infect Dis 2021;

3. Kearse M, Moir R, Wilson A, et al. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 2012;28(12):1647–9.

4. Katoh K, Misawa K, Kuma K, Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res 2002;30(14):3059–66.

5. Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol 2015;32(1):268–74.

6. Dapogny A, Bailly K, and Cord, M. DeCaFA. deep convolutional cascade for face alignment in the wild. In Proceedings of the IEEE/CVF International Conference on Computer Vision 2019 : 6893-6901.

7. Yvinec E, Dapogny A, and Bailly K. DeeSCo. Deep Heterogeneous Ensemble with Stochastic Combinatory Loss for Gaze Estimation. 15th IEEE International Conference on Automatic Face and Gesture Recognition 2020

8. Dapogny A, Bailly K, and Cord M. Deep Entwined Learning Head Pose and Face Alignment Inside an Attentional Cascade with Doubly-Conditional Fusion. IEEE International Conference on Automatic Face and Gesture Recognition 2020

9. Dapogny A, Cord M, and Pérez P. The Missing Data Encoder: Cross-Channel Image Completion with Hide-And-Seek Adversarial Network. AAAI Conference on Artificial Intelligence 2020