SARS-COV-2 AND INFLUENZA SEQUENCING

GIHSN WEBINAR

LAURENCE JOSSET – HCL

ANTONIN BAL – HCL

QUENTIN SEMANAS - HCL

HADRIEN REGUE – HCL
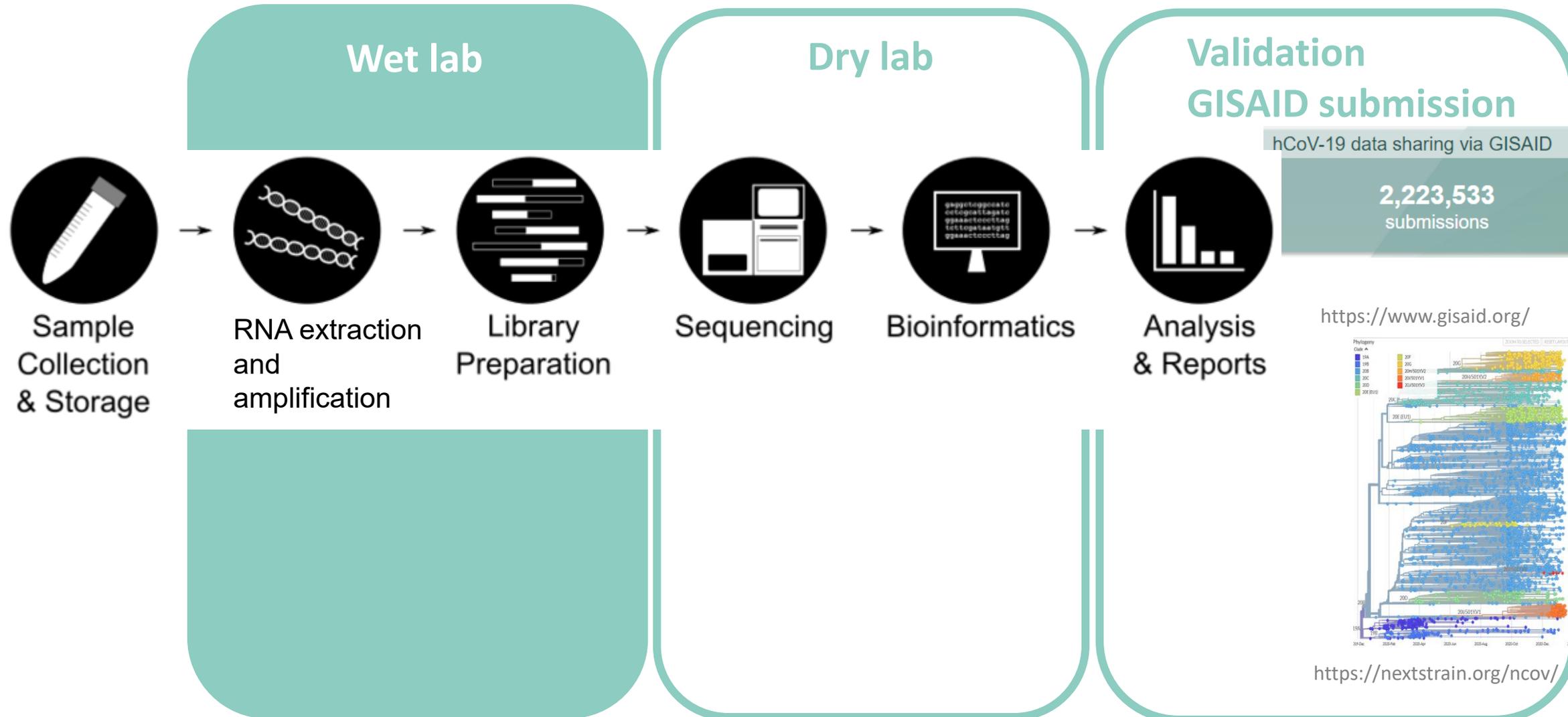
NATHALIE BERGAUD - HCL

CATHERINE COMMAILLE – IMPACT HEALTHCARE

2025/12/11   NGS TEAM- CNR VIRUS RESPIRATOIRES FRANCE SUD

HCL
HOSPICES CIVILS DE LYON

www.chu-lyon.fr

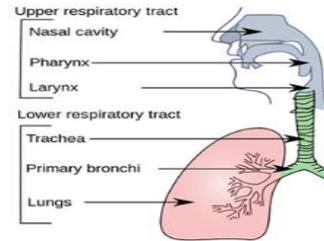# SARS-COV-2 AND INFLUENZA SEQUENCING

METAGENOMICS-NGS

METAGENOMICS-NGS

Detection of **known** and **unknown viruses**

Universal method for viral detection and whole genome sequencing

Influenza A &B, AdV, CMV, HHV-6, EBV, HBoV, HRV, RSV, PIV, MPV, CoV, Enterovirus, Measles, SARS-CoV-2

Bal et al., BMC Inf Dis, 2018

Upper respiratory tract
- Nasal cavity
- Pharynx
- Larynx

Lower respiratory tract
- Trachea
- Primary bronchi
- Lungs

Filtration → DNase → Extraction → **Nucleic acid precipitation** → **cDNA Amplification (WTA)** → **Lib Prep (COVIDSeq)** → Illumina Sequencing

February 2020

Molecular characterization of SARS-CoV-2 in the first COVID-19 cluster in France reveals an amino acid deletion in nsp2 (Asp268del)

A. Bal [1, 2, 3, 4, *], G. Destras [1, 2, 3, *], A. Gaymard [1, 2, 3], M. Bouscambert-Duchamp [1, 2], M. Valette [1, 2], V. Escuret [1, 2, 3], E. Frobert [1, 2, 3], G. Billaud [1, 2], S. Trouillet-Assant [3, 4], V. Cheynet [4], K. Brengel-Pesce [4], F. Morfin [1, 2, 3], B. Lina [1, 2, 3], L. Josset [1, 2, 3, *]

HCL
HOSPICES CIVILS DE LYON

PROS AND CONS FOR mNGS

REFERENCE METHOD

- Non-targeted method adapted for new emerging variant detection

- Low sensitivity: Ct<20
- Expensive
- Low throughput
- Time-consuming process

Charre, *Virus Evol*, 2020

HCL
HOSPICES CIVILS
DE LYON

AMPLICON-BASED TECHNICS
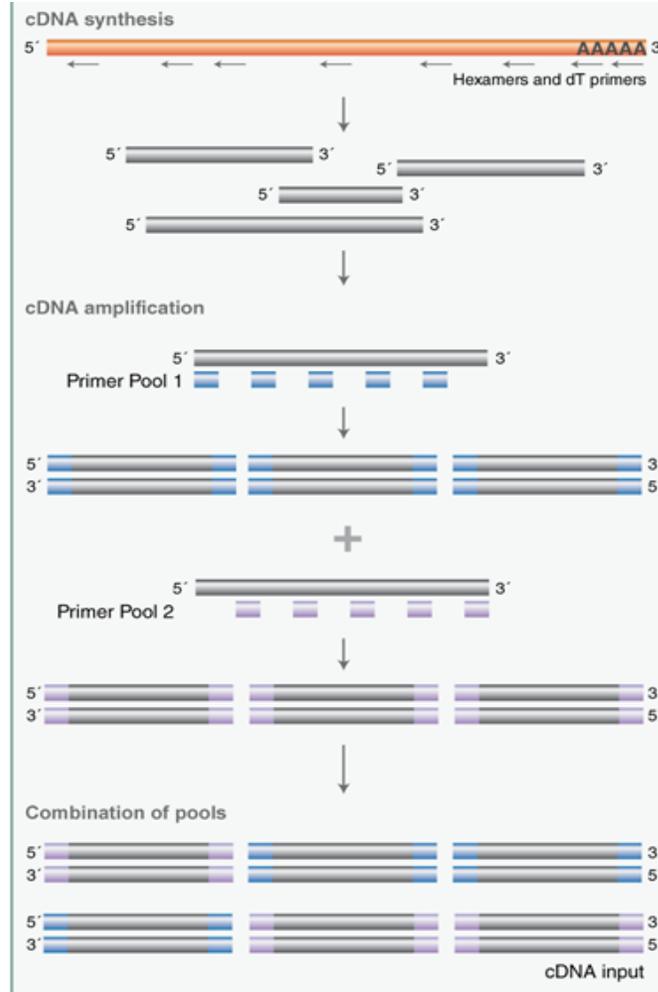
AMPLICON-BASED APPROACH – SARS-COV-2

ARTIC V5.4.2 98 AMPLICONS ~400 BP

1) Nucleic acid extraction

96-Well plate

2) Miniaturized and automatized cDNA synthesis and amplification

384-Well plate

3) Miniaturized and automatized Lib preparation
- Illumina : COVIDSeq
- ONT: sequencing of 400nt amplicons

4) Illumina / ONT Sequencing

AMPLICON-BASED APPROACH – INFLUENZA

UNIVERSAL PRIMERS

1) Nucleic acid extraction

96-Well plate

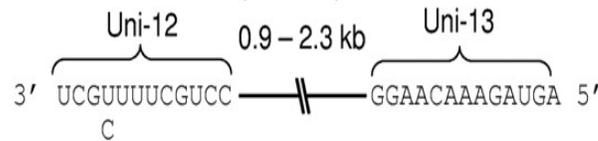2) Miniaturized and automatized cDNA synthesis and amplification

384-Well plate

Single-reaction genomic amplification accelerates sequencing and vaccine production for classical and Swine origin human influenza a viruses

Bin Zhou [1], Matthew E Donnelly, Derek T Scholes, Kirsten St George, Masato Hatta, Yoshihiro Kawaoka, David E Wentworth

```
        Uni-12          0.9 – 2.3 kb        Uni-13
3′   UCGUUUUCGUCC    ─||─   GGAACAAAGAUGA  5′
              C
```

Universal Influenza B Virus Genomic Amplification Facilitates Sequencing, Diagnostics, and Reverse Genetics

Bin Zhou,[a] Xudong Lin,[a] Wei Wang,[a] Rebecca A. Halpin,[a] Jayati Bera,[a] Timothy B. Stockwell,[a] Ian G. Barr,[b] and David E. Wentworth[⊠a]
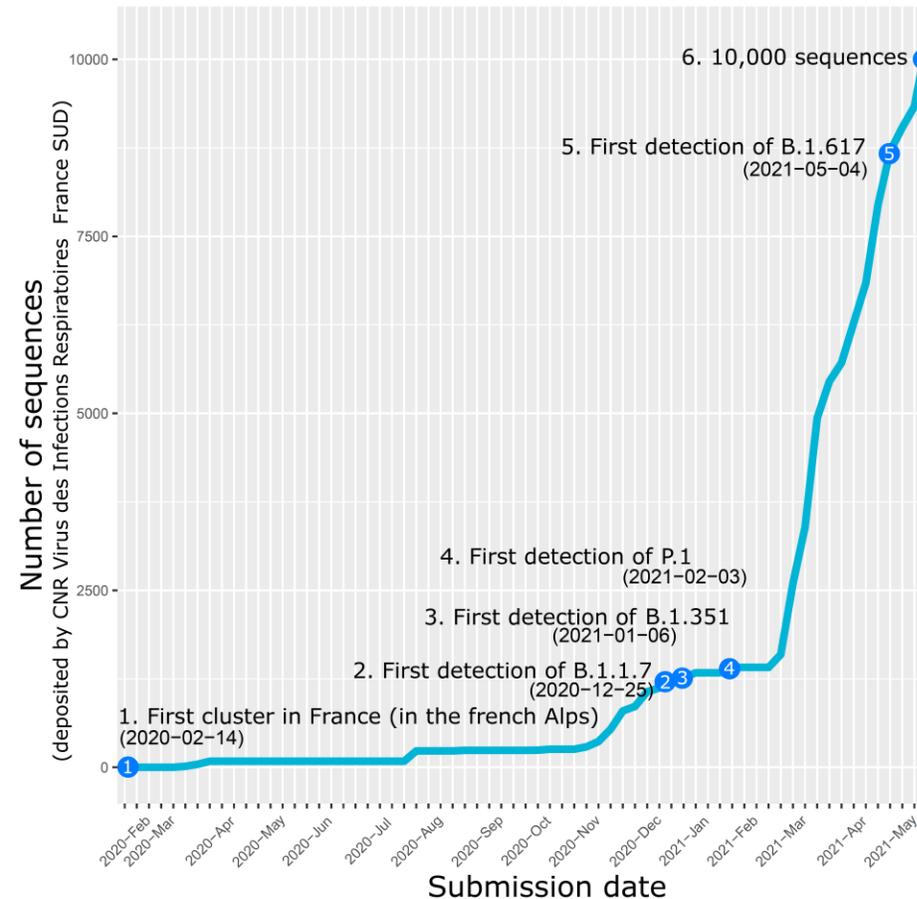
3) Miniaturized and automatized Lib preparation
- Illumina : COVIDSeq
- ONT: sequencing of 400nt amplicons

4) Illumina / ONT Sequencing

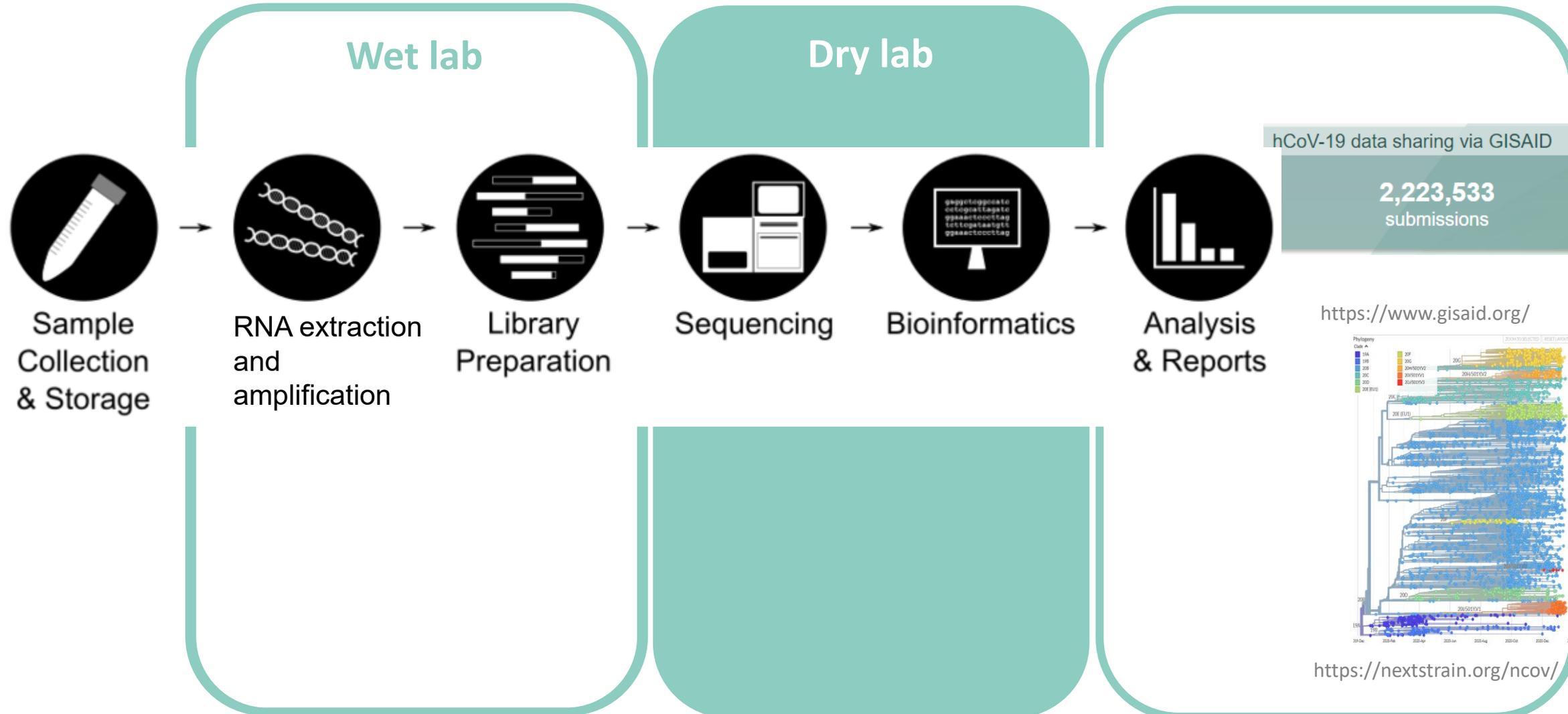# PROS AND CONS FOR AMPLICONS APPROACHES

- Sensitivity
- Cost-effective
- High throughput
- Short-time process

- Impacted by little genome variations (SNPs or indels)

Charre, *Virus Evol*, 2020

SARS-COV-2 AND INFLUENZA SEQUENCING

3 STEPS

Wet lab

Dry lab

Sample Collection & Storage

RNA extraction and amplification

Library Preparation

Sequencing

Bioinformatics

Analysis & Reports

hCoV-19 data sharing via GISAID

2,223,533 submissions

https://www.gisaid.org/

https://nextstrain.org/ncov/

BIOINFORMATICS

FROM RAW DATA TO THE CONSENSUS SEQUENCE: ILLUMINA SEQUENCING

Base call file → **Demultiplexing → Fastq** → **QC and Trimming** → **Alignment** → **Variant Calling** → **Consensus building → Fasta**

Read filtering based on:
-quality
-length

Aligning remaining reads on the reference

Detecting variants in the aligned reads compared to the reference

Reconstruct consensus sequence with retained variants, minimum depth of coverage: 10X

HCL
HOSPICES CIVILS
DE LYON

BIOINFORMATICS

OPEN SOURCE AND COMMERCIAL SOLUTIONS

For Illumina users...

« SEQMET », our « in house » pipeline
freely available on request:

Input viruses:
-SARS-CoV-2
-Influenza A/B (avian analysis are in
reworking)
-Respiratory Syncytial Viruses A/B

... and nanopore adepts

We haven't pipeline dedicated to ONT data analysis
to address, however we can provide help with your
data as well!

Several analysis workflows are proposed by ONT:

Influenza workflow :
https://github.com/epi2me-labs/wf-flu
SARS-CoV-2 workflow:
https://github.com/epi2me-labs/wf-artic

https://labs.epi2me.io/ provide a large
analysis panel (and tutorials) to process
online your data

BIOINFORMATICS

CONTACTS

- We can provide help in your setup process
- Feel free to contact us

NRC's Lyon Bioinformatician:
-Hadrien Regue: hadrien.regue@chu-lyon.fr

SARS-COV-2 AND INFLUENZA SEQUENCING

3 STEPS

Wet lab

Sample Collection & Storage → RNA extraction and amplification → Library Preparation

Dry lab

Sequencing → Bioinformatics

Validation

Analysis & Reports

hCoV-19 data sharing via GISAID

2,223,533 submissions

https://www.gisaid.org/

1. Criteria
2. Tools
3. GISAID
4. Indications

https://nextstrain.org/ncov/
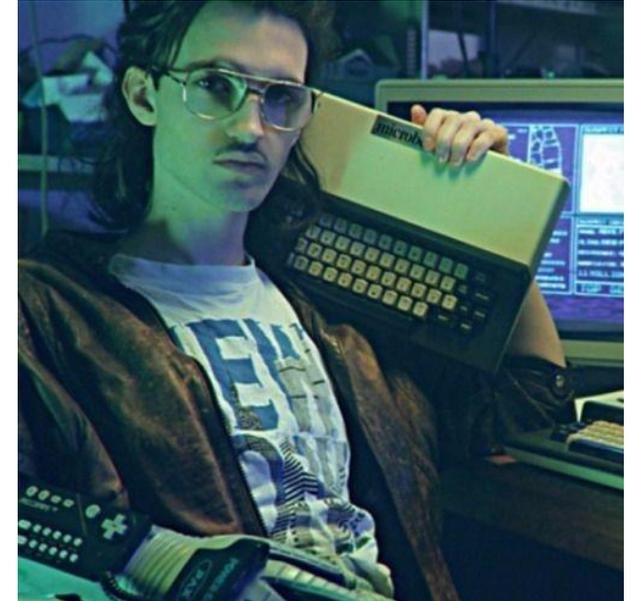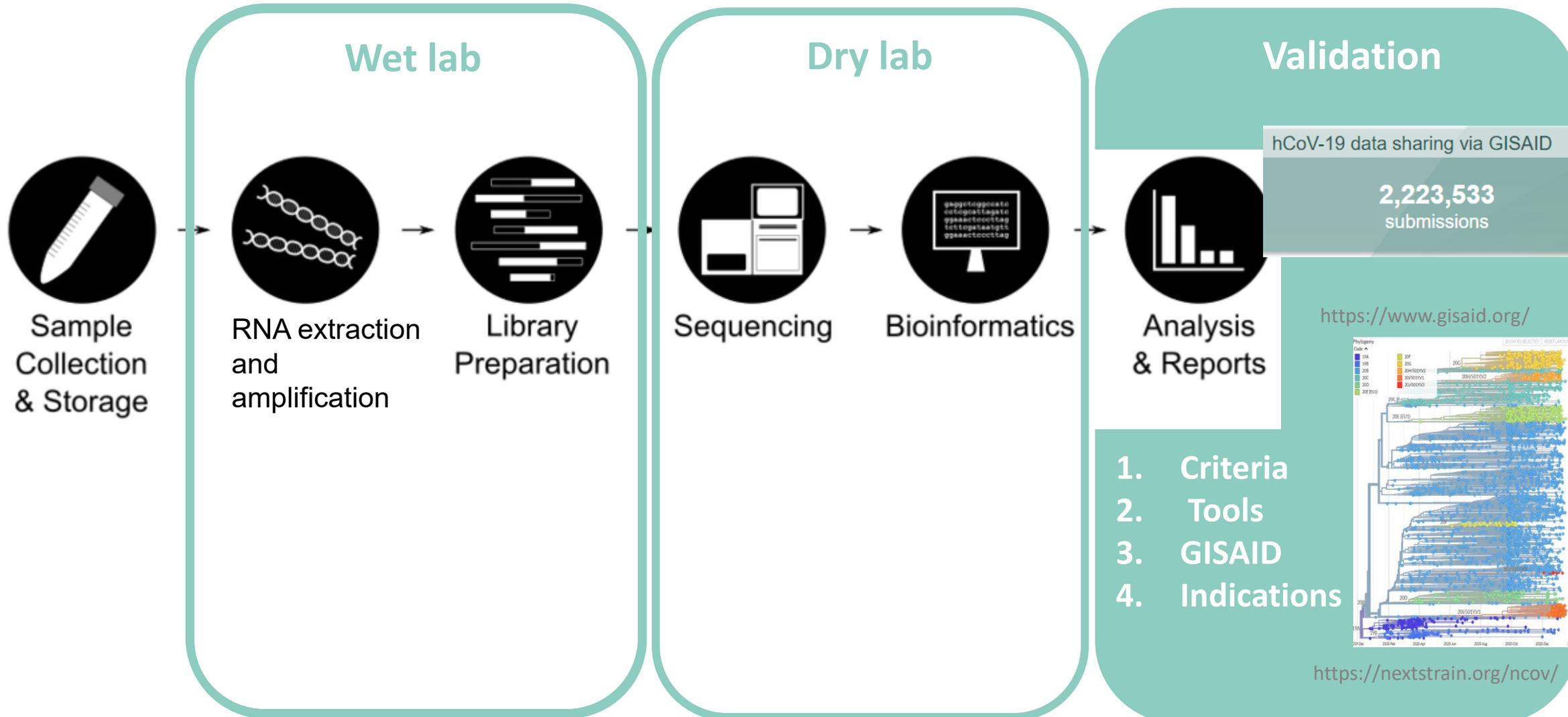
VALIDATION CRITERIA

VALIDATION CRITERIA

NEGATIVE CONTROLS - CONTAMINATION

- **3 No-Template-Controls** (**NTC**) per 96-well plate
- Nuclease free-water **processed through the whole** workflow (including extraction step)

Viral genome coverage for NTCs < 5%

HCL
HOSPICES CIVILS
DE LYON

VALIDATION CRITERIA

POSITIVE CONTROLS - CONTAMINATION

- Viral isolate belonging **to an historical clade**
- **Known mutations**
- **One positive control** / 96-well plate

Viral genome coverage for positive control > 95%

VALIDATION CRITERIA

COVERAGE

- Reminder: minimal depth of coverage / base for consensus sequence generation: 10X for Illumina sequencing
- Otherwise : **N**

The sequence is validated if the genome coverage is > 95 % with a mean depth of coverage > 200X
- SARS-CoV-2 : WGS with >95% coverage
- Influenza virus : HA and NA segments with >95% coverage

VALIDATION CRITERIA

NUMBER AND TYPE OF MUTATIONS

- Divergence (number and type of mutations):
  - Nextclade Mut and phylogenetic tree
  - Check unknown indel
  - compare mutation profile with outbreak.info (for SARS-CoV-2)
- Frameshift mutation: QC nextclade FrameShift
- Atypical set of mutations : QC nextclade Private Mutation
- Minor variants / ambiguous bases : QC Mixed sites

**Repeat extraction
+/- mNGS**

HCL
HOSPICES CIVILS
DE LYON

ID	Sequence name	QC	Clade	Mut.	non-ACGTN	Ns	Gaps
23	✅ Switzerland/AG-ETHZ-430474/2020	N M P C F S					
42	⚠️ USA/CA-CDC-LC0027675/2021	N M P C F S					
29	✅ BHR/30005866/2021	N M P C F S					
35	✅ USA/CA-CDC-LC0024684/2021	N M P C F S					
34	⚠️ USA/NJ-CDC-LC0019972/2021	N M P C F S					
22	✅ Switzerland/un-ETHZ-420516/2020	N M P C F S					
9	✅ HongKong/VM20006541/2020	N M P C F S					
67	✅ USA/FL-CDC-ASC210057011/2021	N M P C F S					
28	⚠️ USA/VA-DCLS-3814/2021	N M P C F S					
84	✅ Switzerland/100064/2020	N M P C F S					
30	⚠️ USA/TX-CDC-STM-000011246/2021	N M P C F S					
50	✅ USA/GA-CDC-ASC210019884/2021	N M P C F S					
73	✅ USA/IL-CDC-LC0051439/2021	N M P C F S					
57	✅ USA/VA-DCLS-5091/2021	N M P C F S					
21	✅ USA/JGGF/2020	N M P C F S					
19	✅ USA/WA-S3074/2020	N M P C F S	20G	20	0	1046	0

Overall QC score: 158
Overall QC status: bad
Detailed QC assessment:

N — Missing Data: mediocre
Missing data found. Total Ns: 2942 (3000 allowed). QC score: 98

M — Mixed Sites: good
No issues

P — Private Mutations: good
No issues

C — Mutation Clusters: good
No issues

F — Frame shifts: good
No issues

S — Stop codons: mediocre
1 misplaced stop codon(s) detected. Affected gene(s): ORF8. QC score: 75

HCL
HOSPICES CIVILS DE LYON

Click "Tree" tab to load this tree view

https://nextstrain.org/ncov/open/global/6m
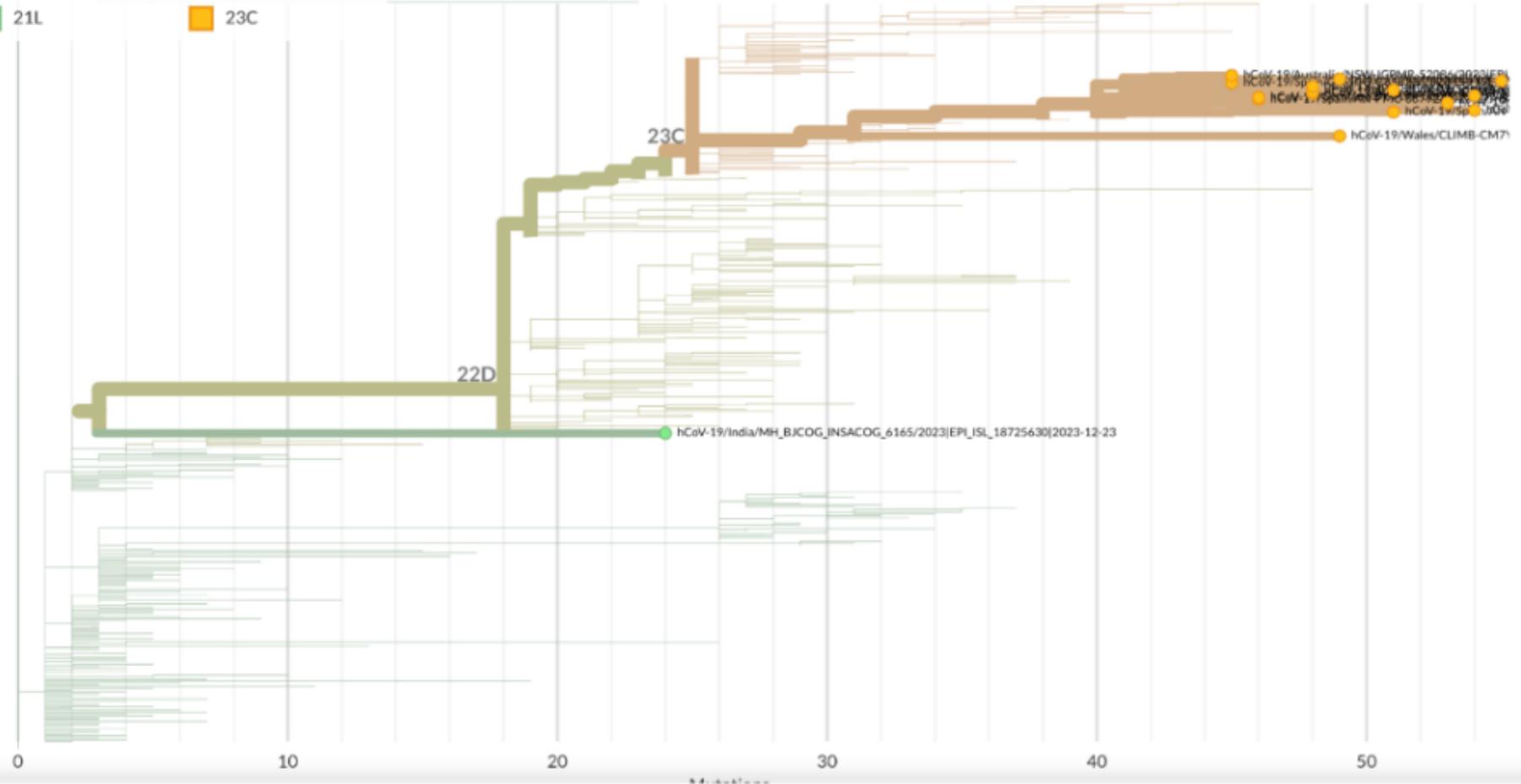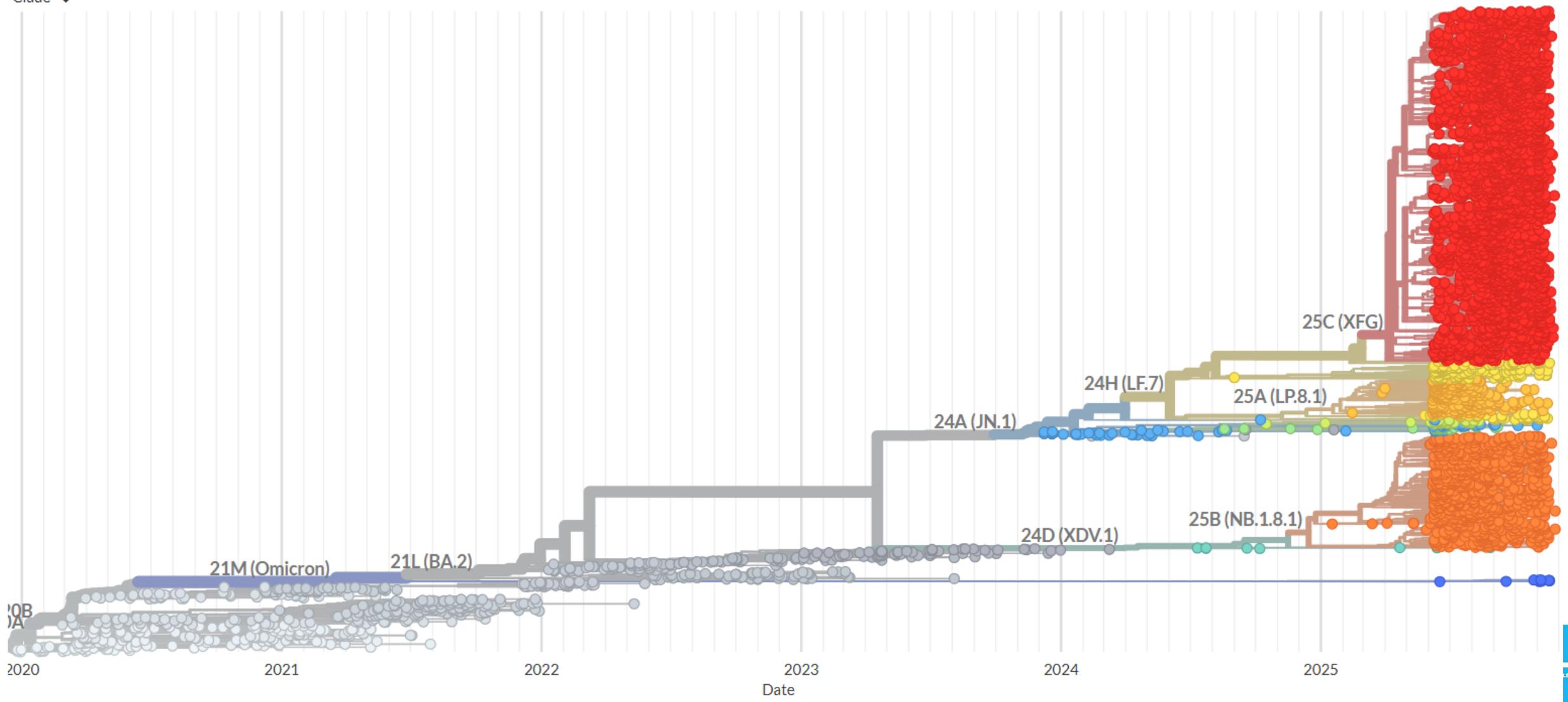
VALIDATION CRITERIA

MOLECULAR EPIDEMIOLOGY

- Different clades within the same cluster
- Clade different from the background
- Clade not circulating in a given area (return from travel)
- Former clades no longer circulating

Repeat extraction

TOOLS FOR QUALITY CONTROL AND VISUALIZATION

NEXTCLADE

NEXTCLADE

PANGOLIN

HTTPS://PANGOLIN.COG-UK.IO/

INSAFLU

https://insaflu.readthedocs.io/en/latest/routine_genomic_surveillance.html

INSAFLU

EXEMPLE WITH AVIAN SAMPLES

Sample Name	Uploaded Date	#Fastq Files	Technology	Classification	Data Set	Alerts	#Quality Seq. (Fastq1)-(Fastq2)
SAMPLE08_S8	09-12-2025 10:24	2	Illumina/IonTorrent	A-H5N1	Generic	0	ⓘ (874676/147.6)-(874676/144.7) (1749352)
SAMPLE09_S9	09-12-2025 10:21	2	Illumina/IonTorrent	A-H5N1	Generic	0	ⓘ (1169243/147.5)-(1169243/144.9) (2338486)
SAMPLE10_S10	09-12-2025 10:20	2	Illumina/IonTorrent	A-H5N1	Generic	0	ⓘ (998971/147.5)-(998971/144.8) (1997942)

Quick subtyping when uploading your data

Get all needed results, including coverage of each segments and consensus sequences

Coverage plot for the HA segment

VACCINE COMPOSITION MEETING

Timeliness sharing of relevant results for the WHO Vaccine Composition Meeting (scheduled in February for the Northern Hemisphere and in September for Southern Hemisphere).
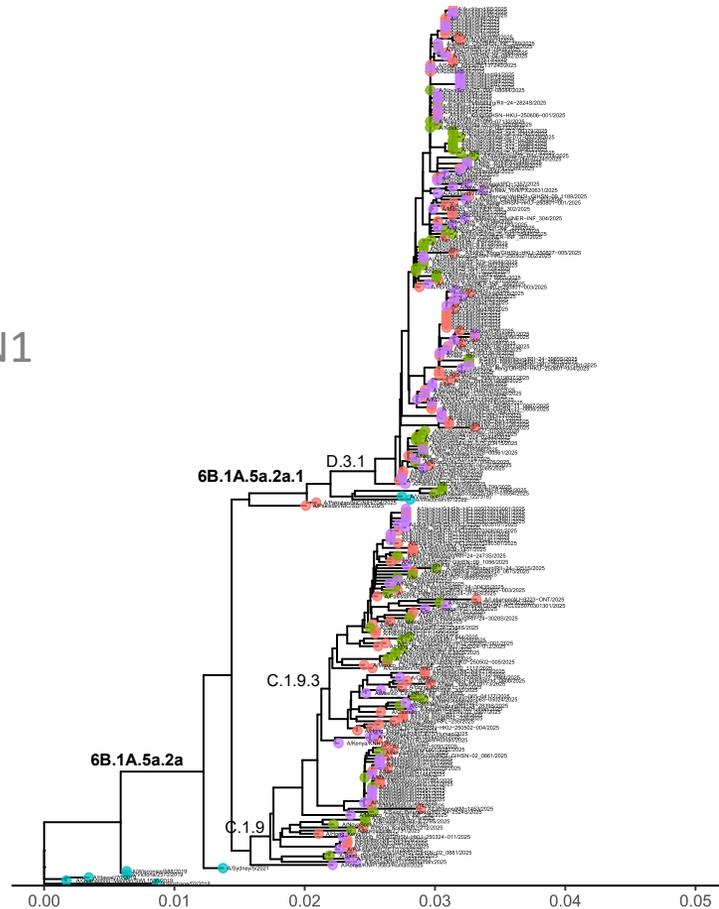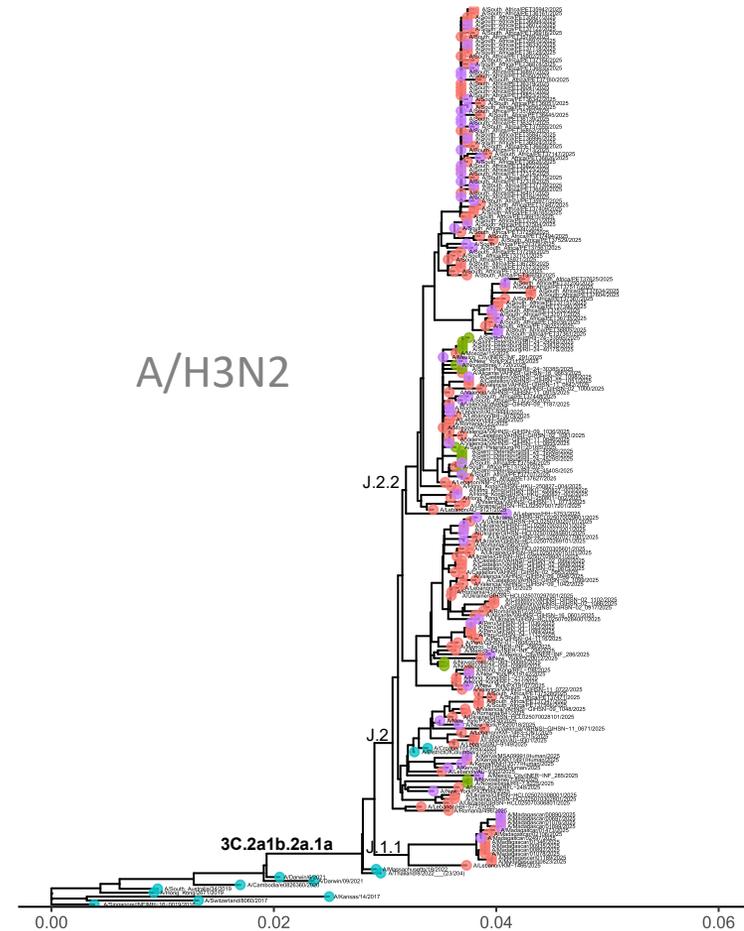
GISAID

GISAID

CREATING AN ACCOUNT

GISAID – INFLUENZA SEQUENCE SUBMISSION

SINGLE UPLOAD

A/[country]/GIHSN-[your Sample_ID]/2025

BATCH UPLOAD

There is no clear GIHSN option to select using influenza batch upload template.

The « GIHSN » tag have to be present in the Isolate Name.

As we working with GISAID for this issue, you can adress your GISHN sequences using the « Note » column in the batch upload template file.

NB: before submitting batch upload, you need to use at least single upload once by your own to get a lab ID.

GISAID – SARS-COV-2 SEQUENCE SUBMISSION

SINGLE UPLOAD

Virus detail

Virus name*	
	hCoV-19/Country/Identifier/2023
Accession ID	
Type	betacoronavirus
Passage details/history*	
	Example: Original, Vero

hCoV-19/[country]/GIHSN-[your Sample_ID]/2025

Institute information

Originating lab*	
	Where the clinical specimen or virus isolate was first obtained
Address*	
Sample ID given by the originating lab	
Submitting lab*	CNR Virus des Infections Respiratoires, France SUD
	Where sequence data have been generated and submitted to GISAID
Address*	103 Grande Rue de la Croix-Rousse 69004 Lyon France
Sample ID given by the Submitting lab	
Consortium	Global Influenza Hospital Surveillance Network (GIHSN)
Authors*	*Jane Doe, John Doe*

GISAID – SARS-COV-2 SEQUENCE SUBMISSION

BATCH UPLOAD

GISAID – SARS-COV-2 SEQUENCE SUBMISSION

BATCH UPLOAD

Column information		
Submitter	mandatory	enter your GISAID-Username
FASTA filename	mandatory	the filename that contains the sequence without path (e.g. all_sequences.fasta not c:\users\meier\docs\all_sequences.fasta)
Virus name	mandatory	e.g. hCoV-19/Netherlands/Gelderland-01/2020 (Must be FASTA-Header from the FASTA file all_sequences.fasta)
Type	mandatory	default must remain "betacoronavirus"
Passage details/history	mandatory	e.g. Original, Vero
Collection date	mandatory	Date in the format YYYY or YYYY-MM or YYYY-MM-DD
Location	mandatory	e.g. Europe / Germany / Bavaria / Munich
Additional location information		e.g. Cruise Ship, Convention, Live animal market
Host	mandatory	e.g. Human, Environment, Canine, Manis javanica, Rhinolophus affinis, etc
Additional host information		e.g. Patient infected while traveling in ….
Sampling Strategy		e.g. Sentinel surveillance (ILI), Sentinel surveillance (ARI), Sentinel surveillance (SARI), Non-sentinel-surveillance (hospital), Non-sentinel-surveillance (GP network), Longitudinal sampling on same patient(s), S gene dropout
Gender	mandatory	Male, Female, or unknown
Patient age	mandatory	e.g. 65 or 7 months, or unknown
Patient status	mandatory	e.g. Hospitalized, Released, Live, Deceased, or unknown
Specimen source		e.g. Sputum, Alveolar lavage fluid, Oro-pharyngeal swab, Blood, Tracheal swab, Urine, Stool, Cloakal swab, Organ, Feces, Other
Outbreak		Date, Location e.g. type of gathering, Family cluster, etc.
Last vaccinated		provide details if applicable
Treatment		Include drug name, dosage
Sequencing technology	mandatory	e.g. Illumina Miseq, Sanger, Nanopore MinION, Ion Torrent, etc.
Assembly method		e.g. CLC Genomics Workbench 12, Geneious 10.2.4, SPAdes/MEGAHIT v1.2.9, UGENE v. 33, etc.
Coverage		e.g. 70x, 1,000x, 10,000x (average)
Originating lab	mandatory	Where the clinical specimen or virus isolate was first obtained
Address	mandatory	
Sample ID given by the originating laboratory		
Submitting lab	mandatory	Where sequence data have been generated and submitted to GISAID
Address	mandatory	
Sample ID given by the submitting laboratory		
Sequencing Consortium		Sequencing consortium the submitting lab is affiliated to
Authors	mandatory	a comma separated list of Authors with complete First followed by Last Name
Comment	leave empty	do not use this column
Comment Icon	leave empty	do not use this column
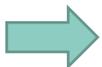
Fill this column with « Global Influenza Hospital Surveillance Network (GIHSN) »

GISAID: WHICH INFORMATION TO ADDRESS TO THE GIHSN

When sending your sample for sequencing to Lyon, results are provided
In a Excel sheet with your own sample number and the corresponding
GISAID Isolate ID to address to the GIHSN form.

	A	Z
1	Isolate_Id	Originating_Sample_Id
2	EPI_ISL_19256954	01-1489
3	EPI_ISL_19256955	01-1494
4	EPI_ISL_19256956	01-1496
5	EPI_ISL_19256957	01-1500
6	EPI_ISL_19256958	01-1504
7	EPI_ISL_19256959	01-1505
8	EPI_ISL_19256960	01-1508
9	EPI_ISL_19256961	01-1509
10	EPI_ISL_19256962	01-1510
11	EPI_ISL_19256963	01-1513
12	EPI_ISL_19256964	01-1518
13	EPI_ISL_19256965	01-1527
14	EPI_ISL_19256966	01-1529
15	EPI_ISL_19256967	01-1530
16	EPI_ISL_19256968	04-0431
17	EPI_ISL_19256969	04-0434

Influenza result table

	A	W
1	Isolate ID	provider_sample_id
2	EPI_ISL_19257028	B-9 01-1497 MAI
3	EPI_ISL_19257029	B-1 04-0358 MAI
4	EPI_ISL_19257030	B-5 04-0415 MAI

SARS-CoV-2 result table

HCL
HOSPICES CIVILS
DE LYON

GISAID: SEND THE SHIPMENT FORM CAREFULLY FILLED

When sending your sample for sequencing to Lyon, please fill information carefully. Sample registering and submission will be easier. Thank you☺

Name	Maiden Name	Surname	Gender	Date of birth	Age	Age_Unit	Collection number	type of samples	Collection_Date	Collection_Time	Country	Region	City	Host

Name PCR : covid	Ct PCR : COVID	Name PCR : Influenza	Ct PCR : Influenza	Name PCR : Influenza	Ct PCR : Influenza B	Influenza subtype	Ct subtype	02 Requireme	Health_Status	Originating lab	Address Origin	Authors

PROCEDURES

HCL
HOSPICES CIVILS
DE LYON

❖ All sites are expected to generate Whole genome sequencing (WGS) for **100 Influenza positive specimens (min. 50)***

❖ All sites must submit WGS to the GISAID EpiFlu™ database **on a regular basis**, and in a timeframe so that results are **available for the WHO Vaccine Composition Meeting (VCM)**

  ❖ In addition, WGS for SARS-CoV2 (if performed) are encouraged to be submitted to GISAID database to add to public knowledge and support WHO initiatives.

❖ **Each GISAID Accession Number (Isolate ID : "EPI_ISL...") must be reported** in the corresponding clinical questionnaire, allowing for the linkage of WGS data with clinical data.

❖ Storage (**preferably -70°C**, if not possible -20°C) of all influenza positive and negative study samples should be carried out for a minimum of one year. This will assure sample availability for additional retrospective investigations (e.g. SARS-CoV-2 or pathogen discovery initiatives) if necessary.

*Some sites are in capacity to do more (Spain, NYC US, Pakistan & NZ)

❖ If the site has no capacities to generate WGS data locally (on its own, or through the usual lab they cooperate with), the site may ship its specimens to the GIHSN sequencing platform at the NIC in Lyon, France. **This has to be discussed and agreed with the Foundation for Influenza Epidemiology before start of the season.**

❖ If specimens are to be sent to Lyon: Shipments are organized by the NIC in Lyon, but sites should be careful to group their shipments (<u>2 shipments max per year</u>).

- ❖ Contact <u>nathalie.bergaud01@chu-lyon.fr</u> with the following information : **address, 2 contacts** (mail+ tel), **nature and number of samples, number of tubes, volume, number and size of boxes**

- ❖ The carrier World Courier will provide box + dry-ice + temperature probe

- ❖ You will receive an **excel file to be completed comprehensively** with all the information regarding your samples. Please send it back asap to Lyon

- ❖ Send preferably **P0 (400 µl minimum)**, if not possible, RNA (50 µl minimum stored at -70°C)

GIHSN WGS REPORTS ARE PREPARED BY THE LAB IN LYON AND SENT TWICE A YEAR (FEBRUARY & SEPTEMBER) TO WHO PRIOR TO THE VACCINE COMPOSITION MEETING
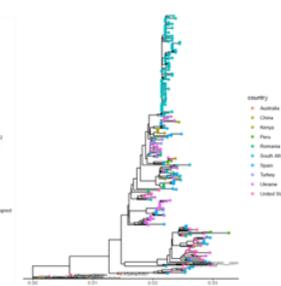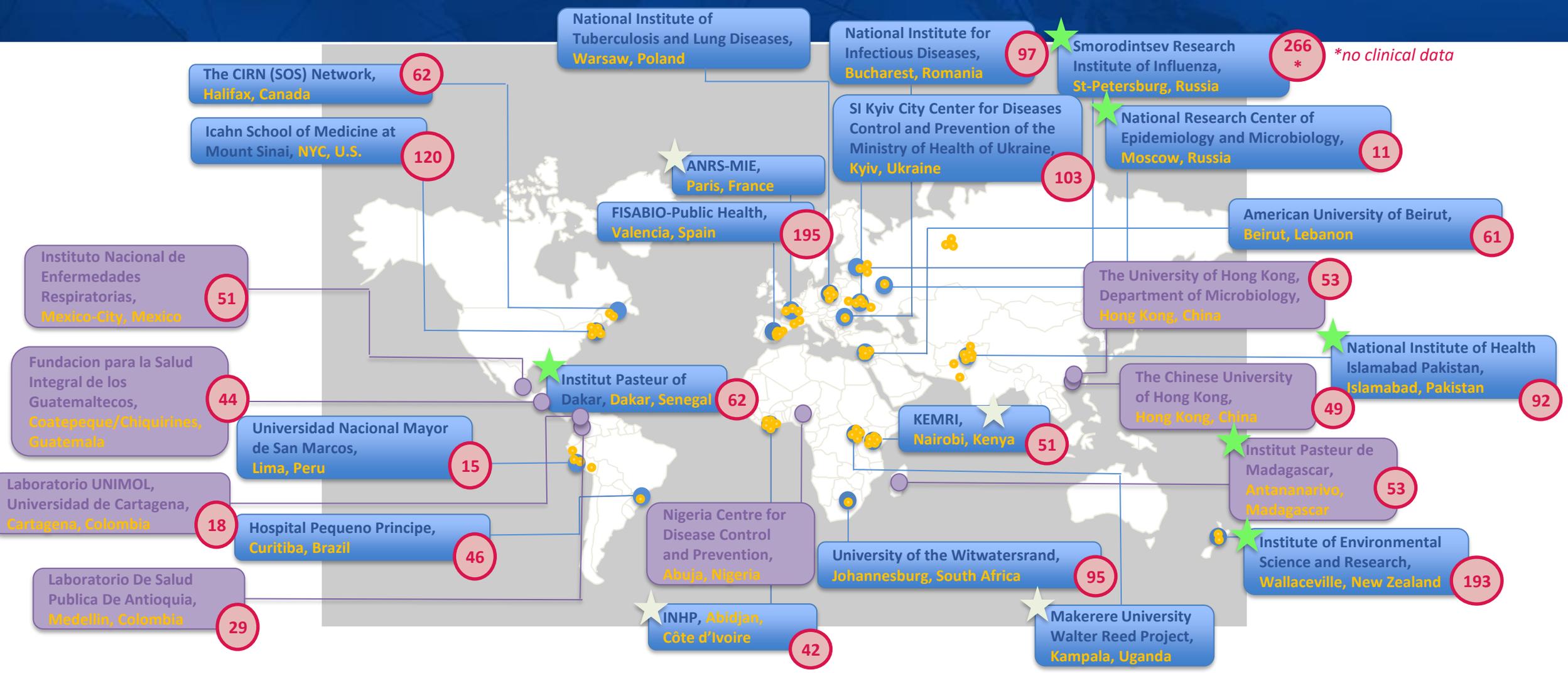

❖ GIHSN sites which proceed with <u>WGS locally</u> are expected to share their WGS results **end of January at the latest** for the VCM of February and **end of August** for the VCM of September

❖ GIHSN sites which send their <u>samples to Lyon for WGS</u> are expected to send all relevant samples **by mid-January at the latest** and **mid-August**.

WGS STATUS (INFLUENZA) TO DATE FOR 2024-2025 (1758 TO DATE)

National Institute of Tuberculosis and Lung Diseases, Warsaw, Poland

The CIRN (SOS) Network, Halifax, Canada — 62

Icahn School of Medicine at Mount Sinai, NYC, U.S. — 120

National Institute for Infectious Diseases, Bucharest, Romania — 97

Smorodintsev Research Institute of Influenza, St-Petersburg, Russia — 266 * *no clinical data*

SI Kyiv City Center for Diseases Control and Prevention of the Ministry of Health of Ukraine, Kyiv, Ukraine — 103

National Research Center of Epidemiology and Microbiology, Moscow, Russia — 11

ANRS-MIE, Paris, France

FISABIO-Public Health, Valencia, Spain — 195

American University of Beirut, Beirut, Lebanon — 61

Instituto Nacional de Enfermedades Respiratorias, Mexico-City, Mexico — 51

The University of Hong Kong, Department of Microbiology, Hong Kong, China — 53

Fundacion para la Salud Integral de los Guatemaltecos, Coatepeque/Chiquirines, Guatemala — 44

National Institute of Health Islamabad Pakistan, Islamabad, Pakistan — 92

Institut Pasteur of Dakar, Dakar, Senegal — 62

The Chinese University of Hong Kong, Hong Kong, China — 49

Universidad Nacional Mayor de San Marcos, Lima, Peru — 15

KEMRI, Nairobi, Kenya — 51

Laboratorio UNIMOL, Universidad de Cartagena, Cartagena, Colombia — 18

Institut Pasteur de Madagascar, Antananarivo, Madagascar — 53

Hospital Pequeno Principe, Curitiba, Brazil — 46

Nigeria Centre for Disease Control and Prevention, Abuja, Nigeria

University of the Witwatersrand, Johannesburg, South Africa — 95

Institute of Environmental Science and Research, Wallaceville, New Zealand — 193

Laboratorio De Salud Publica De Antioquia, Medellin, Colombia — 29

INHP, Abidjan, Côte d'Ivoire — 42

Makerere University Walter Reed Project, Kampala, Uganda

Legend

- 🔵 Sites already in GIHSN 2023-24
- 🟣 New sites (2024-25)
- ⭕ xx Number of WGS Flu per site
- 🟡 Hospital location
- ⭐ National Influenza Center (NIC)
- ☆ Close collaboration with NIC

Global Influenza Hospital Surveillance Network

WGS REPORTING FOR 2025-2026

❖ WGS reporting **to start ASAP**, in view of the next VVCM (February 2026)

❖ WGS results to be shared end of January 2026 **at the latest** (exact dates to be circulated by email)

THANK YOU!