D7.5 Report on feasibility study for influenza virus characterization, molecular epidemiology studies related to influenza related disease in vaccinated and unvaccinated subjects

DRIVE - 777363

Development of robust and innovative vaccine effectiveness

WP7 – Influenza Vaccine
Effectiveness Pilot Studies

Lead contributor	Bruno Lina (10–UCBL)
Other contributors	Xavier López-Labrador (1-FISABIO) Margarita Riera (3-P95) (reviewer) Caterina Rizzo (17-OPBG) Hanna HN. Nohynek (6 – THL)

Due date	30 March 2019
Delivery date	16 September 2019
Deliverable type	OTHER[1]
Dissemination level	PU

Description of Work	Version	Date
	V2.0	12 Sept 2019

---

[1] **Use one of the following codes:**
R: Document, report (excluding the periodic and final reports)
DEM: Demonstrator, pilot, prototype, plan designs
DEC: Websites, patents filing, press & media actions, videos, etc.
OTHER: Software, technical diagram, etc.

Document History

Version	Date	Description
V1.0	01 June 2019	First outline
V1.1	28 June 2019	First draft
V1.2	16 August 2019	Comments from ISC
V2.0	12 Sept 2019	Final version

Contents

Publishable Summary

For optimal analysis of strain specific Vaccine Effectiveness studies, a detailed characterization of the viruses detected in both cases and controls is important. The co-circulation of virus types and sub-types, and sometimes of different genetic and antigenic clades within a given subtype (i.e. A(H3N2) viruses during the 2017-18 and 2018-19 seasons) may lead to different VE according to the virus clade responsible for the infection in enrolled patients, and to potential differences in measured VE when this analysis is performed stratified by virus clade. The full genome sequencing of these viruses will provide WHO with a valuable tool to increase the size of the sequence dataset used during the selection of vaccine candidates during Vaccine Composition Meetings (VCM).

Introduction

The Development of Robust and Innovative Vaccine Effectiveness (DRIVE) project is a public-private partnership aiming to build capacity in Europe for estimating brand-specific influenza vaccine effectiveness (IVE). The DRIVE Project, which is funded by the Innovative Medicines Initiative (IMI), was initiated as a response to the changes for licensing of influenza vaccines in Europe. The new guidance on influenza vaccines by the European Medicines Agency (EMA) came into effect in the beginning of 2017. This guidance states that the performance of influenza vaccines should no longer be assessed based on serological assays, but should be based on post-authorization effectiveness studies [1].

In 2017/18, a pilot study was performed for conducting Influenza Vaccine Effectiveness studies through four test-negative design studies (TND) and one register-based cohort study. The tools and processes developed during the pilot season 2017/18, were used and further improved in the 2018/19 season. The main objective of the 2018/19 season was to estimate brand-specific seasonal IVE in Europe by health care setting and age group. During 2018/19 season, nine test-negative design studies (TND), one register-based cohort study and two cohort studies participate in DRIVE.

However, subtype-specific VE estimates may vary by site according to the clade distribution observed by each site may explain these differences (Fig 1 from Iannery et al. J Infect Dis. 2016;214(7):1010-1019.) .

Fig 1: differences measured in IVE of the A(H3N2) vaccine component according to the infecting clade

Genetic Group, Age	Influenza A(H3N2)–Positive Cases, Proportion (%) Vaccinated	Influenza Virus–Negative Controls, Proportion (%) Vaccinated	VE, % (95% CI)[a]
Overall[b]			
All ages	939/1817 (51.7)	3866/7078 (54.6)	7 (–5 to 17)
Genetic group 3C.3b			
All ages	56/156 (35.9)	3866/7078 (54.6)	44 (16 to 63)
Genetic group 3C.2a			
All ages	597/1101 (54.2)	3866/7078 (54.6)	1 (–14 to 14)
Genetic group 3C.3a			
All ages	31/55 (56.4)	3866/7078 (54.6)	–48 (–169 to 19)
Genetic group 3C.3			
All ages	27/47 (57.5)	3866/7078 (54.6)	1 (–87 to 48)

The integration of virological data in these type of studies is essential to interpret IVE results, particularly when low VE against a specific clade is suspected or when several clades are circulating. The distribution of clades by study site and by influenza period may help to understand the differences between site-specific VE and the changes in VE over the season.

Influenza characterization can be done by sequencing and HIA (Hemagglutination Inhibition Assays). Both methods are used in combination to monitor the antigenic evolution of the influenza viruses. This antigenic evolution of the viruses is a result of the constant emergence of immune escape mutants driven by the immune pressure to which the influenza viruses are exposed, and their high level of evolutionary capacities as a result of the lack of RNA proofreading during replication.

The antigenic evolution is monitored by WHO through the surveillance carried out by the National Influenza Centres, and analyzed by the WHO coordinating centre (WHOcc) the vaccine composition meetings. This detailed surveillance is mandatory, because of the very high annual burden of influenza for which optimal influenza vaccines are required. However, we know that for optimal effectiveness, the HA of the Influenza

vaccines have to be identical to these of the circulating strains. When antigenic differences (mismatches) are observed between the vaccine strain and the circulating viruses, the vaccine effectiveness can be reduced, sometime very significantly. To overcome this problem, the anticipation of virus evolution using tools like antigenic cartography and inference of virus fitness have been used to select the best vaccine candidates and propose vaccine compositions that are more likely matching the circulating strains. Because the selection process has to take place at least 6 months before the vaccination campaign, and 9 months before the actual start of the epidemic, the permanent evolution of the viruses (especially the A(H3N2) component) may become a permanent problem.

A more precise picture of the circulating strains (ie antigenic data, sequence data, clinical description associated to the case can have a positive impact in the strain selection process. For example, the early identification of escape mutants detected from the surveillance or from vaccinated individuals when reduced vaccine effectiveness is observed may be of very important value.

For that purpose, the new generation of sequencing machines (NGS) can be used for the implementation of large-scale molecular characterization of viruses, with detailed information regarding the consensus sequences of the HA and potentially the other gene segments. This technique is of interest, and largely used for A(H3N2) viruse for the identification of escape mutants (in combination with the conventional HIA), and for the description of the trend of their emergence (fig:2 example for the 2019 VCM for the northern hemisphere).

Fig 2: monitoring of the respective proportions of the A(H3N2) clades presented during the feb 2019 VCM.
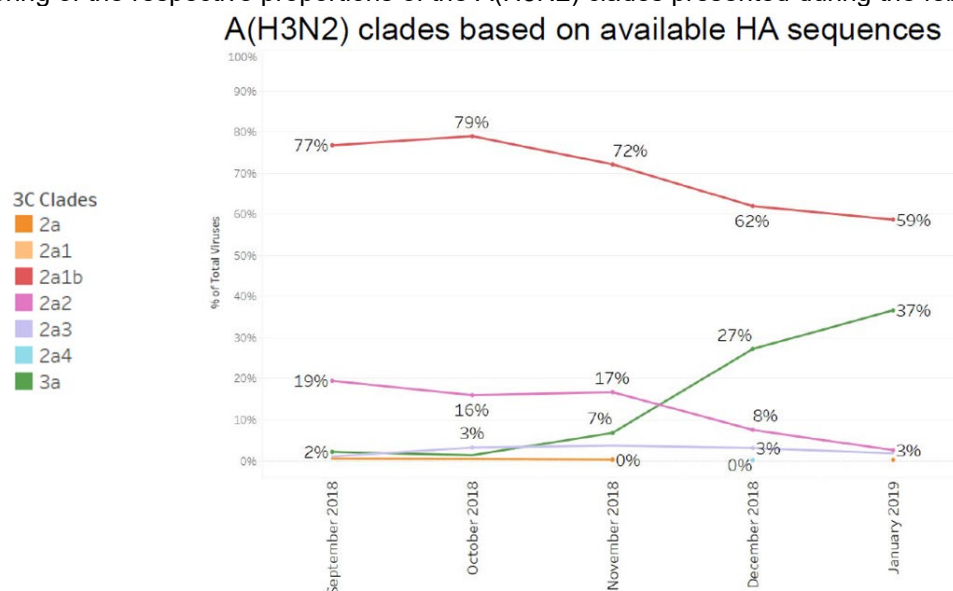
Last, in cases of vaccine failure, it is theoretically possible to identify in these patients subgroups of minority variants sequences that would reflect, by anticipation, some of the potentially emerging escape mutants that would drive the antigenic evolution of the currently circulating viruses.

A retrospective preliminary study performed on a limited number of hospital cases by the research group of UCBL (Simon B et al, Viruses, 2019) has demonstrated the capacity for NGS and a full genome sequencing strategy to identify a differential genetic sequence between viruses detected from severe and non-severe cases, as well as from vaccinated and non-vaccinated patients. This preliminary needs to be confirmed by larger studies with a robust design In addition, this study provided some technical information on the threshold to be used when analyzing minority variants (% of variant sequence as compared to the consensus), and the depth of sequencing required to be informative (minimum of reads for each nucleotide position).

Therefore, our aims are (1) to put in place a shared protocol for influenza virus NGS sequencing, and (2) to define how to select a representative sample of DRIVE specimens, to be analysed by NGS during the 2020-2021 studies, in order to identify/describe the viruses sequences, and to identify amino acid changes that can influence the effect of the influenza vaccines. This work should not modify the sampling strategy used for virological surveillance.

Objectives

Our main Objective is to describe, using NGS techniques, the potential genetic heterogeneity of viruses detected during the surveillance conducted by the DRIVE sites, and to identify key influenza virus genotypic evolutions that could affect vaccine effectiveness.

Secondary objectives are:
- to provide real-time data to support the decision for the strain selections carried out during the Vaccine Composition Meeting (VCM);
- to use the high throughput sequencing system available in some site (NGS) to provide sequence data allowing to estimate clade specific VE with pooled influenza data collected among sites participating in DRIVE;
- to provide clues for the antibody-induced antigenic evolution of circulating strains, and to detect if vaccine failure can be either a consequence or a reflect of genetic/antigenic evolution of viruses.

To be informative for the VCM, this analysis has to be done almost in real time, requiring a technical and human resource to perform this analysis (sequencing and big data management), and a system by which these specimens could be analyzed in a time frame adapted for the VCM and analysis of the virus evolution for the relevant stakeholders involved in the strain selection process.

Methods

Both the French National Influenza Reference Laboratory of Lyon, France and the Virology Laboratory in the Genomics & Health Area from FISABIO-Public Health in Valencia, Spain, have a strong experience in virus genetic characterisation by standard (Sanger) and NGS platforms. Before the start of the 2018-19 season, both laboratories prepared a common shared protocol for influenza virus NGS (Annex). Harmonization included materials required, laboratory techniques used between both partners, and the same sequencing platforms. As a proof-of-concept study, both laboratories applied the shared protocol to a limited number of influenza A(H3N2) and A(H1N1pdm09) samples during the 2018-19 season, obtaining satisfactory and comparable results (sufficient sequencing coverage for all the 8 segments of influenza viruses), thus consolidating the feasibility of extending whole-genome NGS sequencing of influenza to DRIVE.

For that purpose, preliminary exploring studies will be carried out by the UCBL lab on some specimens collected during the 2018-19 and 2019-2020 seasons. To meet the objectives of the study we will collect specimens from both vaccinated and non-vaccinated cases. These specimens will be analyzed by NGS sequencing and retrospectively compared to check for the variability of their genetic content. Instead of limiting the analysis to the classical characterization of the hemagglutinin and neuraminidase viral genes, a comprehensive sequencing of the full viral genome will be performed.

In addition, for the 2019-2020 call for IVE studies performed in the context of DRIVE will ask study sites to test the feasibility for doing NGS analysis (to be performed by the UCBL lab) of either HA and NA or WGS in a subset of the ILI/SARI cases recruited in the study (see case definition paragraph). This second, prospective step is required to test the logistics and the technical hurdles for a larger scale of analysis, and to identify sites with all the required technical capacities and know-how. Eventually, these sites will be asked to share their data within the consortium, to allow a multisite integration of the dataset generated during these studies.

Case definition

The ILI/SARI case definition used, and the procedures to identify influenza cases and controls are the ones defined in the DRIVE generic protocols.

Study period

The proportion of strains characterized will vary according to the phase in the influenza season: the proportion will be higher in the early and late phase and lower in the peak of the season. Each site will define the start/end of each influenza phase according to the surveillance indicators used in their site. During this preliminary phase, each site will define the proportion of strains to be characterized (sampling fraction) for each of the phases. Each site will:

- define the influenza phases;
- define the sampling fraction for each phase;
- document the sampling fraction for each phase and for each subtype.

Sample size

It is difficult to define the minimum sample size to meet the primary objective of the study. This will depend on several factor, but mainly, on the vaccination coverage among cases and controls. In the first phase of the study, sites will be required to provide a convenient sample of the positive specimens detected in their premises. The study sites will try to characterise the maximum number of strains taking into account the existing resources. Each site will define the proportion (representative of the whole collection if possible) of samples to be characterised.

We recommend to have a higher sampling fraction in the early phase (**if possible, characterise all strains**) of the influenza season to better understand what is circulating and to be able to provide results for the WHO vaccine strain selection committee that meets around end February each year.

We also recommend a higher sampling fraction in the late phase of the influenza season to detect possible changes in predominant clades.

Subsequently, WP7 will revise the number of strains characterised and included in the analysis. Based on the epidemiological situation, the early VE estimates and the available resources, WP7 will discuss the optimum sampling fraction that would be needed in each study site for the next 2019/20 season.

The random selection should be used to select the viruses included in the study sites (this does not apply to routine virological surveillance). At the end of each time period defined for the selection of strains, study sites will select viruses using the Bernoulli sampling method. This method ensures that each strain has the same probability of being selected. Specific guidelines will be provided to each study site.

Analysis

The number (and proportion) of viruses selected, eligible but excluded because they could not be characterized and the reasons for exclusion will be presented, according to the resources of the sites. Cases eligible, selected but not characterized will be described (age, vaccination status, date of onset). Cases excluded due above described reasons will be compared with cases with strains characterized in terms of date of onset, age, vaccination status, chronic diseases.

For each subtype/lineage, the following indicators will be computed in a centralized lab (UCBL):
• number and proportion of viruses selected for characterization;
• number and proportion of viruses sequenced with NGS techniques;
• number and proportion of viruses by clades taking into account the sampling fraction;
• distribution of clades by study site and for the pooled dataset;
• distribution of clades by time (weeks or months).

Discussion

The implementation of the first-year study will provide evidence for conducting future real-time monitoring of genetic sequences of circulating viruses detected in the IVE DRIVE studies. It will also be an additional tool for WHO during the vaccine composition meeting.
The second step will be the implementation of this characterization for all DRIVE IVE studies.

Conclusion

This work plan will be implemented with the help of the two laboratory structures already in capacity to generate influenza WGS data by NGS within the DRIVE network.

Annexes

INTRODUCTION
Version: *DRAFT*   ***Date :*** *27/06/2019*

There is a high interest to better understand and capture actual and prospective influenza genetic drifts and to characterize virus variants to support new influenza vaccines development. WHO and ECDC highly recommend and push forward for genetic characterization of influenza strains during seasons to better inform the vaccine recommendation committees and to enhance surveillance for new strains and/or potential pandemic viruses. Fine-tuned antigenic characterisation by monoclonal antibody techniques is used by WHOCC to characterize influenza strains yearly, but these techniques cannot be shared.

Both the French National Influenza Reference Laboratory of Lyon, France and the Virology Laboratory in the Genomics & Health Area from FISABIO-Public Health in Valencia, Spain have a strong experience in virus genetic characterization by standard (Sanger) and Next-Generation sequencing platforms, and will bring-up laboratory support to DRIVE . Harmonization of laboratory techniques used is necessary and possible within the network, according to the proposed following protocol.

Objectives of this study
The main objective is to sequence a significantly relevant sample of strains from the sites during the season, representing the diversity of circulating strains, to capture drifts. The study will also allow for the identification by comparative genomics of differences between circulating viruses and the A(H3N2), A(H1N1)pdm09 and B vaccine strains. Influenza A(H3N2) viruses are of particular interests because they are evolving rapidly and more genetic sequencing data is needed. Instead of limiting the analysis to the classical characterization of the hemagglutinin and neuraminidase viral genes, a comprehensive sequencing of the full viral genome will be performed. This strategy will add value to conventional characterization carried out in most IVE studies, and will help to understand drop in vaccine effectiveness, when antigenic/ genetic variants emerge before or during the Influenza season.

Materials required, procedures, and laboratory protocols.
The protocols following in annex detail the materials required and the laboratory procedures. These protocols are derived from UCB Lyon and FISABIO current procedures. Total nucleic acid from the samples to be included in the study can be extracted by the lab sites using the commercial nucleic acid extraction protocol 1 in annex, or alternatively an equivalent nucleic acid extraction protocol after approval by FISABIO/Lyon. This strategy would eventually facilitate easy shipping of the extracted nucleic acids to FISABIO/Lyon. Another option is to centralize nucleic acid extraction at FISABIO and Lyon: nasopharyngeal swabs in virus transport medium should be then shipped frozen to FISABIO/Lyon by specialized international courier.

PROTOCOL 1. TOTAL NUCLEIC ACID EXTRACTION
(Biomerieux NucliSENS easyMAG)
Version: DRAFT   *Date : 24/11/2017*

I. Introduction

The NucliSENS easyMAG is an automated platform for the isolation (purification and concentration) of total nucleic acids (RNA/DNA) from biological specimens. Used for batch processing of 1-24 samples.

The nucleic extraction method is based on Boom Chemistry using magnetic silica particles. Briefly, under high salt conditions, nucleic acid will bind to the silica particles. These silica particles act as solid phase and non-nucleic acid components are removed by several washing steps performed in the NucliSENS easyMAG instrument. Next, nucleic acids are eluded from the silica particles and the silica particles are removed from the extracted specimens by the NucliSENS easyMAG instrument. The resulting eluate contains purified and concentrated total nucleic acids.

II. Specimens Types

Viral transport media (VTM) from swabs collected in VTM. Use a minumun of 200 mL of VTM, up to 1 mL maximum, for final elution in 50 uL of elution buffer. Vortex on high for 10 seconds and brief spin before extraction.

III. Materials, Equipments

NucliSENS easyMAG
NucliSENS easyMAG extraction products:
- NucliSENS Lysis Buffer
- NucliSENS easyMAG Lysis Buffer
- NucliSENS easyMAG Magnetic Silica
- NucliSENS easyMAG Extraction Buffer 1
- NucliSENS easyMAG Extraction Buffer 2
- NucliSENS easyMAG Extraction Buffer 3
- NucliSENS easyMAG Disposables –vessels and pipettes

Multichannel pipettor
1000uL, 100ul, 20ul Pipettors
Aerosol Resistant Tips (ART) 1000ul, 200ul, 20ul
Powder-free gloves
Internal controls (for human DNA and/or viruses)

IV. Worksheet.
Fill the lab worksheet with the sample details (máx. 22 samples + one positive + one negative control per run).

GENERAL PRECAUTIONS:

Use Specimen Preparation area; work always in BSC cabinet with gown and frequent glove change. For respiratory viruses, use of a face mask until the sample is neutralizad in lysis buffer is highly recommended.

Do not allow the NucliSENS easyMAG Lysis Buffer, NucliSENS easyMAG Extraction Buffer 1 or waste from the instrument to come in contact with acidic materials.

Certain reagents contain guanidine thiocyanate.
Buffers containing guanidine thiocyanate MUST NOT be mixed with cleaning solutions containing bleach. Liquid waste from extractions procedures contain guanidine thiocynate MUST NOT be mixed with other laboratory waste. Proper waste disposal is required.

IV.   Procedure

1. Turn ON easyMAG. Wait for orange light to turn GREEN, then switch ON the easyMAG computer.
2. Perform and record easyMAG Daily Maintence: Inspect drip tray.
   Inspect reagent cap filters.
   Inspect carbon filter.
   Empty & rinse waste bottle with water.

3. Log-in easyMAG   Login name: _____
   Password: _____

4. Click on "reagent" icon

5. Scan reagent stand and scan reagent bottle for all four reagents (A, B, C, D).

6. Click on "Sample" icon

For VTM/Respiratory viruses
   Enter Protocol: UTM-Respvir-1ml (or 200ml)
   Enter Matrix: Other
   Enter Volume (mL): 1.000 (or 200ml)
   Eluate (uL): 50 ul (reduce if sample volume is lower)

7. Click on sample ID bar and scan specimen barcodes (or input sample data)

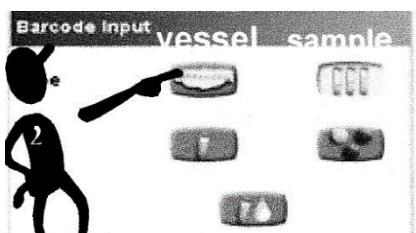8. Click on "organize run" icon

9. Click on "create run" icon

10. A box will appear with today's date and run #. Click on OK.

11. Click on to transfer all scanned items onto the worklist.

12. Click on "load run" icon

13. To identify vessels, click on vessel icon

14. Scan position A barcode and then vessel barcode for first 8 samples, continue with position B and C depending on how many samples you have.

15. Click on the silica icon

16. Scan in silica barcode from top of silica box.

17. Highlight all samples by touching screen or with mouse.

18. To apply silica barcode to all samples click on "downward arrow" icon

19. Click on the "progress tab" **>>>>** on third menu bar from top.

20. Print worklist. Select printer icon on right hand vertical menu bar. Pop-up appears, click OK.

21. Click on the "dispense lysis icon"

22. Once the lysis buffer has been dispensed, remove vessels form easyMag and put them in rack and take them to the BSC.

23. For respiratory viruses, add 1 mL of the VTM specimen to the lysis buffer.

24. Vortex internal control (IC). *(Not applicable for this study)*
    For Respiratory viruses, add 10uL of IC to the lysis buffer/specimen mix. *For CMV, BKV, HSV, VZV, PARVO, ENTERO, HCV, HBV, HIV, add 5 uL IC.* Mix up and down three times.

25. Incubate at room temperature for 10 minutes.

26. To make silica premix add 550 uL DEPC-PCR grade water to silica (program 1 of pipette)

27. Vortex silica premix before dispensing.

28. Dispense 125 uL silica into Elisa strip (program 2 of pipette)

29. After the 10 incubation period add 100 uL silica premix to vessel (program 3 of pipette)

30. Reload vessels into the easyMAG.

31. Scan position barcode and then vessel barcode (make sure the vessel icon is highlighted).

32. Click on start icon on right hand vertical menu bar.
    A pop-up box will appear asking if you have added the premix silica, click YES.

33. Once the samples have been eluted, transfer the eluate from the vessels to labeled 1.5 ml screw-cap o-ring microtube (use printed label with sample study number). Set pipette at 55uL to remove all of the eluate.

34. If elutes are to be tested same day store eluates at 4ºC in the fridge. If PCR / testing is to be performed at a later date, store eluate at -70ºC as soon as possible.

V.   Quality Control

Failed extraction runs, should be reviewed with senior and/or Charge technologist to establish causes and corrective actions; repeating extraction run maybe necessary.
If the easyMag system becomes inoperable call Biomerieux technical support phone number, provided on each easyMag.

VI.   References
NucliSENS easyMAG User Manual v2.0 ref: 280163

PROTOCOL 2.
Whole-genome amplification protocol
Influenza A (MBT 12/13 modified)
Version: DRAFT   *Date :* 27/06/2019

RT-PCR primers:

(MBTuni-12 [5'-ACGCGTGATCAGCAAAAGCAGG-3'] and MBTuni-13 [5'-ACGCGTGATCAGTAGAAACAAGG-3')

RT-PCR amplification set-up:

☐ Fill-in sample worksheet

PCR N°	Sample ID
Sample 1	*Xxxxxxx*
…	*…*
Sample N	*Xxxxxxx*

☐ Make RT-PCR mix (for 50 ul final reaction volume)

	for 1 sample	for N samples
Primer MBT12 10µM	2 µl	
Primer MBT13 10µM	2 µl	
2 X reaction mix	25µl	
Superscript III RT platinum Taq High Fidelity Enzyme Mix	1µl	
RNASin	0.5 µl	
PCR-grade water	9.5 µl	

☐ Add 10 µl of nucleic acid extract
☐ Vortex and spin.
☐ Place in thermocycler

Thermocycler program: (~3.5 hours):

Temperature	Duration	Cycle number
45°C	1 h	
94°C	2 min	
94°C	30 s	
45°C	30 s	X 5
68°C	3 min	
94°C	30 s	
57°C	30 s	X 31
68°C	3 min	

References:

- B Zhou; ME Donnelly; D T Scholes; K St. George; M Hatta ; Y Kawaoka and DE Wentworth. Single-Reaction Genomic Amplification Accelerates Sequencing and Vaccine Production for Classical and Swine Origin Human Influenza A Viruses. JOURNAL OF VIROLOGY, Oct. 2009, p. 10309–10313.

PROTOCOL 3.
Whole-genome amplification protocol
Influenza B (Zhou et al. modified)

Version: DRAFT   *Date : 27/06/2019*

RT-PCR primer cocktail:

Primer	Sequence	Volume (Concentration : 10 µM)
B-PBs-UniF	5'-GGGGGGAGCAGAAGCGGAGC-3'	10 µL
B-PBs-UniR	5'-CCGGGTTATTAGTAGAAACACGAGC-3'	10 µL
B-PA-UniF	5'-GGGGGGAGCAGAAGCGGTGC-3'	5 µL
B-PA-UniR	5'-CCGGGTTATTAGTAGAAACACGTGC-3'	5 µL
B-HANA-UniF	5'-GGGGGGAGCAGAAGCAGAGC-3'	10 µL
B-HANA-UniR	5'-CCGGGTTATTAGTAGTAACAAGAGC-3'	10 µL
B-NP-UniF	5'-GGGGGGAGCAGAAGCACAGC-3'	6 µL
B-NP-UniR	5'-CCGGGTTATTAGTAGAAACAACAGC-3'	6 µL
B-M-Uni3F	5'-GGGGGGAGCAGAAGCACGCACTT-3	3 µL
B-Mg-Uni3F	5'-GGGGGGAGCAGAAGCAGGCACTT-3'	3 µL
B-M-Uni3R	5'-CCGGGTTATTAGTAGAAACAACGCACTT-3'	6 µL
B-NS-Uni3F	5'-GGGGGGAGCAGAAGCAGAGGATT-3'	5 µL
B-NS-Uni3R	5'-CCGGGTTATTAGTAGTAACAAGAGGATT-3'	5 µL
	Total	**84 µL**

RT-PCR amplification set-up:

☐   Fill-in sample worksheet

PCR N°	Sample ID
Sample 1	*Xxxxxxxx*
…	*…*
Sample N	*Xxxxxxx*

RT-PCR Amplification mix

Reagent	For 1 sample	For n samples
RNAsin	0.2 µL	
2X RT-PCR Buffer	12,5 µL	
IBV-GA primer cocktail	2 µL	
RT/HiFi enzyme mix	0,75 µL	
Total	15.45 µL	

☐   Vortex and Spin

☐   Mix 15 µL of reagent mix in each tube and add 10 µl of nucleic acid extract

☐   Mix gently and spin.

Thermocycler program: (~5h17)

Temperature	Duration	Cycle number
45°C	60 min	1
55°C	30 min	
94°C	2 min	
94°C	20 sec	5
40°C	30 sec	
68°C	3 min 30	
94°C	20 sec	45
58 °C	30 sec	
68 °C	3 min 30	
68°C	10 min	1
15°C	∞	1

References

- B Zhou; ME Donnelly; D T Scholes; K St. George; M Hatta ; Y Kawaoka and DE Wentworth. Single-Reaction Genomic Amplification Accelerates Sequencing and Vaccine Production for Classical and Swine Origin Human Influenza A Viruses. JOURNAL OF VIROLOGY, Oct. 2009, p. 10309–10313.

PROTOCOL 4.
Agarose gel verification of whole genome
amplicons (Influenza A or B)

Version: DRAFT   *Date : 27/06/2019*

☐   Mix 200 ml TBE 5X with 800 ml of ultra-pure water.

NB : Maximum conservation : 1 month

Gel preparation X%

☐   Mix x g of agarose with 100 ml TBE 1X.

☐   After heating (600W for 2 min), add 10 µl Syber®Safe DNA gel strain 10 000 X and mix.

☐   Polymerization for 30 min (1 hour for concentration > 2%).

Amplicons verification

☐   Add and mix on parafilm 6µl of loading buffer 2X with 6 µl of DNA ladder or 6 µl of amplicons.

☐   Gently mix.

☐   Load 10 µl in the wells.

☐   Do these steps for each samples and the DNA ladder.

☐   Migration
  - o   180V for 50-60 min (4% gel)
  - o   200V for 30-35 min (3% gel)
  - o   180V for 40-45 min (2% gel)
  - o   200V for 20-25 min (1% gel)

☐   Annotate gel results

	PM											PM
Operator#1												
Operator#2												

	PM											PM
Operator#1												
Operator#2												

<div style="border:1px solid #000; text-align:center">

PROTOCOL 5.
Magnetic bead purification of whole genome
amplicons (Influenza A or B)

</div>

Version: DRAFT   *Date : 27/06/2019*

- ☐ Vortex the magnetic beads for 30 sec (Macherey Nagel) before use
- ☐ Transfer the amplicons in a 1.5ml LoBind tube
- ☐ Add 1.8Xµl beads/µl of Amplicons (or 0.6Xµl if size selection needed)
- ☐ Vortex 5sec and pulse spin
- ☐ Contact : 5 min at room temperature
- ☐ Put-on magnet for 3-5min
- ☐ Discard the supernatant
- ☐ Add 200µl of 70% ethanol
- ☐ Put on the magnet and discard the supernatant
- ☐ Add and discard ethanol 2 twice
- ☐ Completely dry the sample after washing.
- ☐ Eluate dried beads in appropriate volume of LowTe
- ☐ Mix gently and leave for contact for 5 min
- ☐ Put on magnet for 3 min
- ☐ Collect the supernatant
- ☐ Dose amplicons on a Nanodrop and measure the purity of the extract
- ☐ Dilute it if neede before final dosage on a Fluorometer
  - Normalize all concentration at 0.2 ng/µl

PROTOCOL 6.
Library Preparation Protocol of whole genome
amplicons (Influenza A or B).
NEXTERA XT – Illumina

Version: DRAFT   *Date :* 27/06/2019

DNA Tagmentation

- Vortex and spin all reagents.
- In each mix well:
  - Add 10 µl of Tagment DNA buffer
  - Add 5µL of amplicons (ADN 0.2ng/uL)
- Homogenize and seal the plate.
- Centrifuge at 280xG for 1 min at room temperature then transfer to a thermal cycler
- Thermocycler program:
  - 55°C for 5 min
  - 10°C ∞
- Add 5µL Neutralization buffer in each well
- Homogenize and seal the plate.
- Centrifuge at 280g for 1 min at room temperature then transfer to a thermal cycler
- Incubation for 5 min at room temperature

Amplification (library recovery PCR)

- Centrifuge i5 and i7 tubes before use
- Put i7 tubes in horizontal range and i5 in vertical
- In each well mix the tagmented sample with
  - 15µL of Nextera PCR Master Mix
  - 5µL of i5 index (white cap)
- Seal the plate
- Centrifuge at 280xG for 1 min at room temperature then transfer to a thermal cycler
- Thermocycler program:
  - 73°C for 3 min
  - 12 cycles of :
    - 95°C for 10 secondes
    - 55°C for 30 secondes
    - 72°C for 30 secondes
  - 72°C pour 5 minutes, then 10°C ∞
- STOPPING POINT (10°C ∞ thermos cycler or -4°C)

Library recovery PCR Clean up

- Vortex magnetic bead for 30 secs
- Prepare Ethanol 80%
- Centrifuge at 280xG for 1 min at room temperature then transfer to a thermal cycler
- Mix beads with tagmented DNA according to the following table:

Amplicon maximum length	Beads	Volumes
<300bp	Not recommended	Not recommended

300-500 bp	1 X vol	90uL
>500bp	0.6X vol	30uL
>500bp (for Miseq)	0.5X (for Miseq) vol	25uL

- ☐ Vortex (1800 rpm for 2 min)
- ☐ Incubate for 5 min at room temperature
- ☐ Put on magnet for 3 min
- ☐ Discard the supernatant
- ☐ Add 200ul of 80% ethanol
- ☐ Put on the magnet and discard the supernatant
- ☐ Add and discard ethanol 2 other times
- ☐ Completely dry the sample after the final washing step (15 min).
- ☐ Eluate dried beads in 52.5 of ReSuspension Buffer
- ☐ Mix gently and contact for 2 min
- ☐ Put on magnet for 2 min
- ☐ Collect 50µL of the supernatant
- ☐ STOPPING POINT seal the plate and stock -25°C / -15°C for 1 week maximum

Normalization

- ☐ Prepare 4 mL of 0.1N NaOH
- ☐ Prepare 4.4mL of LNA1 (for 96 samples) adding 800uL of LNB1.
- ☐ Homogeneise
- ☐ Add 45µL of LNA/B1 mix to each well
- ☐ Seal the plate
- ☐ Centrifuge at 1800 rpm for 30 minutes at room temperature then transfer to a thermal cycler
- ☐ Put on magnet for 2 min
- ☐ Discard the supernatant
- ☐ Add 45uL of LNW1 to each well
- ☐ Seal the plate
- ☐ Centrifuge at 1800 rpm for 5 minutes at room temperature then transfer to a thermal cycler
- ☐ Put on magnet for 2 min
- ☐ Discard the supernatant
- ☐ Add 45µL of LNW1 to each well (a 2nd time)
- ☐ Seal the plate
- ☐ Centrifuge at 1800 rpm for 5 minutes at room temperature then transfer to a thermal cycler
- ☐ Put on magnet for 2 min
- ☐ Discard the supernatant
- ☐ Add 30uL of NaOH 0.1N (for elution)
- ☐ Seal the plate
- ☐ Centrifuge at 1800 rpm for 5 minutes at room temperature then transfer to a thermal cycler
- ☐ Add 30µL of LNS1 in each well
- ☐ Put on magnet for 2 min
- ☐ Collect 30µL of supernatant
- ☐ Seal the plate
- ☐ Vortex 1000xG for 1 minute
- ☐ Pool 5µl of each well in a unique tube
- ☐ STOPPING POINT seal the plate and stock -25°C / -15°C for 1 week maximum

<div style="border:1px solid">

PROTOCOL 7.
Quantification of library preps (Kapa qPCR kit)

</div>

Version: DRAFT   *Date : 27/06/2019*

For the first use

- ☐ Mix 1 mL of 10X Primer Premix in 5 mL of 2X KAPA SYBR FAST qPCR Master Mix
- ☐ Vortex
- ☐ Stock in dark (-20°C)

Library dilutions

- ☐ Prepare 2 independent dilutions in duplicates, in low bind tubes

at 1 /200 :

Final dilution	1/10	1/100	1/200
PCR grade water	45 µl	45 µl	5 µl
	5 µl final library	5 µL 1/10	5µL 1/100

- ☐ at 1 /500 :

Final dilution	1/10	1/100	1/500
PCR grade water	45 µl	45 µl	20 µl
	5 µl final library	5 µL 1/10	5µL 1/100

Mix preparation

For 6 standards:
1 NTC
4 wells per dilution
12 wells for the standards

	1 tube (µL)	1 library (µL) = 18 wells	X library (µL); X = 13 wells + 4X
X KAPA SYBR FAST qPCR Master Mix + 10X Primer Premix	6	108	
H2O	2	36	
Total	8	144	

For 6 standards:
1 NTC
4 wells per dilution
8 wells for the standard

	1 tube (µL)	1 library (µL) = 14 wells	X library (µL); X = 10 wells + 4X
X KAPA SYBR FAST qPCR Master Mix + 10X Primer Premix	6	84	
H2O	2	28	
Total	8	112	

Run preparation

Plan

	1	2	3	4	5	6	7	8	9	10	11	12
A												
B		S1	S2	S3	S4	S5	S6					
C		S1	S2	S3	S4	S5	S6					
D		1-1/200	1-1/500	Blanc								
E		1-1/200	1-1/500									
F												
G												
H												

Mark at lower-

- ☐ Add
  - o 8µL de mix par well
  - o 2µL per standard
  - o 2µL per library dilution
  - o 2µL per no template control
- ☐ Seal the plate
- ☐ Centrifuge

DNA Standard 1	20 pM
DNA Standard 2	2 pM
DNA Standard 3	0,2 pM
DNA Standard 4	0,02 pM
DNA Standard 5	0,002 pM
DNA Standard 6	0,0002 pM

Run

- ☐ New experiment
- ☐ Open Templates and choose the appropriate template file
- ☐ Program:

Step	Temperature	Duration	Cycles
Initial denaturation	95°C	5 min	1
Denaturation	95°C	30 sec	35
Hybridation/Elongation	60°C	45 sec	
Melt curve analysis	95 - 65 - 95°C		
Cooling	40°C	10min	1

Data analysis

- ☐ Report values on spreadsheet «model_input_KAPA.xls »

PROTOCOL 8.
Sequencing Run Illumina NextSeq V2

Version: DRAFT   *Date : 27/06/2019*

Run planning

☐ Connect NextSeq apparatus

Reagent Preparation

☐ Prepare 1mL of NaOH 0.2M
☐ De-freeze the cartridge with gentle agitation.
☐ Prepare the flowcell and dry it if needed
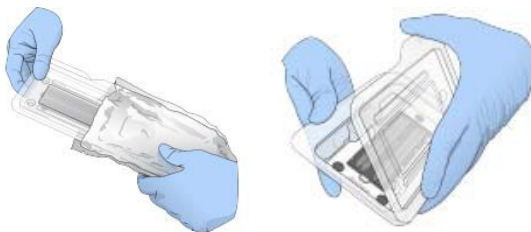
Librairies final denaturation and dilution

☐ In a Lowbind tube mix
  o 30µL of Pooled Librairies
  o 30µL of NaOH 0.2M
☐ Vortex and spin
☐ Incubate 5 min at room temperature
☐ Add 30uL of TrisHCL (200 mM, pH7)
☐ Vortex and spin
☐ Dilution of the denaturated library by adding Xul of HT1
☐ Vortex and spin
☐ Stock on ice

Denaturation and Dilution of PhiX control (if needed)

☐ Dilute PhiX to 4 nM : mix
  o 10uL of PhiX (10nM)
  o 15uL of ReSuspension Buffer
☐ Vortex and spin

☐ Stock : 3 months between -25°C and -15°C

Denaturation of the mixture

- ☐ In a Lowbind tube mix
  - o 5 µl of diluted PhiX
  - o 5uL of NaOH 0.2M
- ☐ Vortex and spin
- ☐ Incubate for 5 minutes at room temperature
- ☐ Add 5µL of Tris HCL (0.2M, pH7.0)
- ☐ Vortex and spin

Dilution of denaturated PhiX for final mix

- ☐ Add 985µL of HT1 (total volume: 1mL à 20 pM)
- ☐ Final dilution (1.8pM in 1.3mL)
  - o 117µL of denaturated PhiX (20pM)
  - o 1183µL of HT1
- ☐ Vortex and spin
- ☐ Stock : 2 weeks between -25°C and -15°C

Mix PhiX and Librairies

- ☐ In a Lowbind tube mix
  - o 1287 µL of the final librairies
  - o 13uL of the final PhiX
- ☐ Put on ice until sequencing

Cartridge loading

- ☐ Clean the n°10 lid
- ☐ Pierce the lid using a P1000 cone
- ☐ Transfer the whole Phix and Librairies dilution
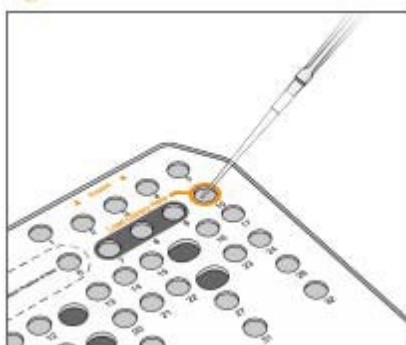
Sequencing start

☐ Put on the new flow cell after discarding the used flow cell

☐ Put on loading then next
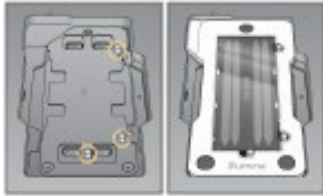
☐ Discard the used reagent and put the empty garbage cartridge

☐ Discard the used reagent cartridge and insert a new one

☐ Close and put Next

☐ Discard the sample reagent cartridge and insert a new one

☐ Select « load »

☐ Cartridge check by the system

☐ Click on next to launch the sequencing

☐ select « next »
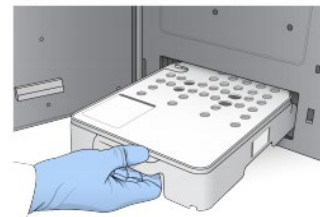
Run monitoring

☐ Check on screen :
  o Run progress :
  o Q-score distribution
  o Intensity
  o Cluster densité (K/mm3)
  o Cluster passing filter (%)
  o Estimated Yield (Gb)
  o Flow cell image
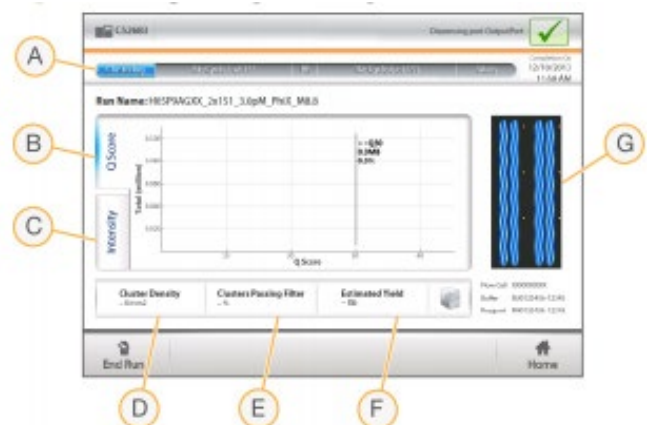
| PROTOCOL 10. |
| **Sequence data processing and** |
| **Bioinformatic analyses.** |

Version: DRAFT   *Date : 27/06/2019*

Following is an example of a sequence data processing and bioinformatic analysis pipeline currently in place at FISABIO-Public Health. This pipiline is used here as a mere reference for a minimum protocol, but can be eventually changed before implementation after consultation/agreement with the Lyon site. The pipieline can be also modified to introduce improvements through the duration of the study.

Sequence processing, mapping and variant calling.

A pipeline mapping of R1 and R2 sequencing reads will be carried out by the bwa mem program (Li, et al., 2009; Li, et al., 2010), and recalibrated using the Picard program https://broadinstitute.github.io/picard/ to generate bam files.

Variants against reference influenza genomes (i.e. from those viral Straits included in the seasonal vaccine) will be called using the freebayes program (Garrison et al., 2009), to generate variant calling tables including frequencies, using following parameters:

Raw variants files: SampleName/var/SampleName.freebayes.raw.vcf

-F 0.2 = minimum alternate fraction 20%

-C 2 = minimum aalternate count 2

-m 0 = minimum mapping quality 0

-q 0 = minimum base quality 0

Filtered variants [-F 0.2 -C 5 -m 30 -q 20] files: SampleName/var/SampleName.freebayes.vcf

-F 0.2 = minimum alternate fraction 20%

-C 5 = minimum alternate count 5

-m 30 = minimum mapping quality 30

-q 20 = minimum base quality 20

Annotation of SNPs/InDels against reference influenza genomes (i.e. from those viral Straits included in the seasonal vaccine) will be carried out using the snpEff program (Cingolani, et al., 2012) to produce annotated files.

References

J. Allaire, et al. markdown: Markdown rendering for R. R package version 0.7.4. 2014. URL: http://CRAN.R-project.org/package=markdown.

S. Anders. "Visualization of genomic data with the Hilbert curve". In: Bioinformatics 25.10 (may. 2009). [PubMed Central:PMC2677744] [DOI:10.1093/bioinformatics/btp152] [PubMed:19297348] , pp. 1231-1235.

C. Boettiger. knitcitations: Citations for knitr markdown files. R package version 1.0.5. 2014. URL: http://CRAN.R-project.org/package=knitcitations.

P. Cingolani, et al. "A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3". In: Fly (Austin) 6.2 (2012). [PubMed Central:PMC3679285] [DOI:10.4161/fly.19695] [PubMed:22728672] , pp. 80-92.

L. Guy, et al. "genoPlotR: comparative gene and genome visualization in R". In: Bioinformatics 26.18 (sep. 2010). [PubMed Central:PMC2935412] [DOI:10.1093/bioinformatics/btq413] [PubMed:20624783] , pp. 2334-2335.

M. Lawrence, et al. "Software for computing and annotating genomic ranges". In: PLoS Comput. Biol. 9.8 (2013). [PubMed Central:PMC3738458] [DOI:10.1371/journal.pcbi.1003118] [PubMed:23950696] , p.

e1003118.

H. Li, et al. "Fast and accurate long-read alignment with Burrows-Wheeler transform". In: Bioinformatics 26.5 (mar. 2010). [PubMed Central:PMC2828108] [DOI:10.1093/bioinformatics/btp698] [PubMed:20080505] , pp. 589-595.

H. Li, et al. "The Sequence Alignment/Map format and SAMtools". In: Bioinformatics 25.16 (ago. 2009). [PubMed Central:PMC2723002] [DOI:10.1093/bioinformatics/btp352] [PubMed:19505943] , pp. 2078-2079.

M. Morgan, et al. "ShortRead: a bioconductor package for input, quality assessment and exploration of high-throughput sequence data". In: Bioinformatics 25.19 (oct. 2009). [PubMed Central:PMC2752612] [DOI:10.1093/bioinformatics/btp450] [PubMed:19654119] , pp. 2607-2608.

V. Obenchain, et al. "VariantAnnotation: a Bioconductor package for exploration and annotation of genetic variants". In: Bioinformatics 30.14 (jul. 2014). [PubMed Central:PMC4080743] [DOI:10.1093/bioinformatics/btu168] [PubMed:24681907] , pp. 2076-2078.

D. Sarkar. Lattice: Multivariate Data Visualization with R. ISBN 978-0-387-75968-5. New York: Springer, 2008. URL: http://lmdvr.r-forge.r-project.org.

DRIVE 2019 Annual Forum
Virological characterization, WGS pilot

Task leaders:

F. Xavier López-Labrador
FISABIO – Public Health, Generalitat Valenciana, Valencia, Spain
CIBER in Epidemiology and Public Health, Instituto de Salud Carlos III, Spain

Bruno Lina
CNR des virus Influenza, CHU Hospices Civils de Lyon
Virpath, CIRI, U1111, UMR 5308, ENS, UCBL, Lyon, France

Influenza viruses

Influenza epidemiology

- Viruses are circulating worldwide
- Evolving constantly
- Type A viruses multiple hosts: human, flock, swine: complex evolution
- The epidemiology is different in the different regions of the world
- The impact is correlated to:
  - the type and subtype
  - the escape to pre-existing immunity
- Disease burden is often underestimated, but much better monitored these last years
- The diagnostic tools are powerful

Types, sub-types and lineages of seasonal influenza viruses

Implementation of a pilot Influenza WGS study in Valencia, Spain

From the past 2017-18 season: WGS analyses for identifying strain variation and evolution

Table 2. Example of automatically-generated variant report (vcf) for a complete influenza virus genome, filtered by missense mutations only

DRIVE Influenza WGS pilot

- Common WGS protocol in place, prepared by UCBL-Lyon and FISABIO-Public Health, Valencia.
- WP7, D7.5: "*Report on feasibility study for influenza virus characterization, molecular epidemiology studies related to influenza related disease in vaccinated and unvaccinated subjects*"
- Run on already collected samples from the 2018/19 season
- Using the collected specimens from the study sites we can detect:
  - Substitutions associated to vaccine failure (antigenic drift)
  - Trend of substitutions associated to genetic evolution (clades/subclades)
  - Minority variants detected prior to the emergence of sustained substitutions associated to antigenic drift
- Use pilot to evaluate feasibility for timely feeding GISAID database for WHO
- No funding available, UCBL-Lyon has funding for aprox. 100 samples
- FISABIO-Public Health has funding for 100 additional samples
- Option for sites with ongoing NGS assays to share sequencing data