



HAL
open science

An amplicon-based nanopore sequencing workflow for rapid tracking of avian influenza outbreaks, France, 2020-2022

Guillaume Croville, Mathilda Walch, Aurélie Sécula, Laetitia Lèbre, Sonia Silva, Fabien Filaire, Jean-Luc Guerin

► **To cite this version:**

Guillaume Croville, Mathilda Walch, Aurélie Sécula, Laetitia Lèbre, Sonia Silva, et al.. An amplicon-based nanopore sequencing workflow for rapid tracking of avian influenza outbreaks, France, 2020-2022. *Frontiers in Cellular and Infection Microbiology*, 2024, 14, pp.1257586. 10.3389/fcimb.2024.1257586 . hal-04434201

HAL Id: hal-04434201

<https://hal.science/hal-04434201>

Submitted on 2 Feb 2024

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

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



OPEN ACCESS

EDITED BY

Nicolae Corcionivoschi,
Agri-Food and Biosciences Institute,
United Kingdom

REVIEWED BY

Yuebang Yin,
Erasmus Medical Center, Netherlands
Gloria Consuelo Ramirez-Nieto,
National University of Colombia, Colombia

*CORRESPONDENCE

Guillaume Croville

✉ guillaume.croville@envt.fr

†These authors share first authorship

RECEIVED 12 July 2023

ACCEPTED 04 January 2024

PUBLISHED 22 January 2024

CITATION

Croville G, Walch M, Sécula A, Lèbre L, Silva S,
Filaire F and Guérin J-L (2024) An amplicon-
based nanopore sequencing workflow for
rapid tracking of avian influenza outbreaks,
France, 2020-2022.
Front. Cell. Infect. Microbiol. 14:1257586.
doi: 10.3389/fcimb.2024.1257586

COPYRIGHT

© 2024 Croville, Walch, Sécula, Lèbre, Silva,
Filaire and Guérin. This is an open-access
article distributed under the terms of the
[Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/).
The use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

An amplicon-based nanopore sequencing workflow for rapid tracking of avian influenza outbreaks, France, 2020-2022

Guillaume Croville^{*†}, Mathilda Walch[†], Aurélie Sécula,
Laetitia Lèbre, Sonia Silva, Fabien Filaire and Jean-Luc Guérin

IHAP, Université de Toulouse, ENVT, INRAe, Toulouse, France

During the recent avian influenza epizootics that occurred in France in 2020/21 and 2021/22, the virus was so contagiousness that it was impossible to control its spread between farms. The preventive slaughter of millions of birds consequently was the only solution available. In an effort to better understand the spread of avian influenza viruses (AIVs) in a rapid and innovative manner, we established an amplicon-based MinION sequencing workflow for the rapid genetic typing of circulating AIV strains. An amplicon-based MinION sequencing workflow based on a set of PCR primers targeting primarily the hemagglutinin gene but also the entire influenza virus genome was developed. Thirty field samples from H5 HPAIV outbreaks in France, including environmental samples, were sequenced using the MinION MK1C. A real-time alignment of the sequences with MinKNOW software allowed the sequencing run to be stopped as soon as enough data were generated. The consensus sequences were then generated and a phylogenetic analysis was conducted to establish links between the outbreaks. The whole sequence of the hemagglutinin gene was obtained for the 30 clinical samples of H5Nx HPAIV belonging to clade 2.3.4.4b. The consensus sequences comparison and the phylogenetic analysis demonstrated links between some outbreaks. While several studies have shown the advantages of MinION for avian influenza virus sequencing, this workflow has been applied exclusively to clinical field samples, without any amplification step on cell cultures or embryonated eggs. As this type of testing pipeline requires only a short amount of time to link outbreaks or demonstrate a new introduction, it could be applied to the real-time management of viral epizootics.

KEYWORDS

influenza, virus, Oxford nanopore technologies, NGS, bioinformatics

Introduction

Avian influenza is a highly contagious infectious disease caused by Influenza A viruses belonging to the *Orthomyxoviridae* family (International Committee on Taxonomy of Viruses (ICTV)). Aquatic wild birds represent the natural reservoir of the virus, contributing to its spread and the generation of occasional epidemics characterized by the deaths of both wild and domestic birds, generating significant economic losses (Olsen et al., 2006; Causey and Edwards, 2008).

The highly pathogenic avian influenza viruses (HPAIVs) responsible for the 2020/21 and 2021/22 European epizootics belong to the H5N8 and H5N1 subtypes, both belonging to the A/Goose/Guangdong/1/1996 (GSGd) lineage. The HPAIVs actually circulating in Europe descended from the H5N8 variant (GsGd lineage of clade 2.3.4.4b) that emerged in Europe in 2014 after an intercontinental spread which started in Southeast Asia in early 2014 (Liu et al., 2005; Ku et al., 2014; Lee et al., 2014; Wu et al., 2014; Poen et al., 2018).

As of the 1st of March 2023, the 2020/21 epizootics led to 3,555 reported HPAI detections and affected around 22,400,000 poultry birds in 28 European countries (European Centre for Disease Prevention and Control, 2021). In 2021/22, the joint EFSA, ECDC and EU reference laboratory report listed a total of 2,467 outbreaks in poultry, 48 million birds culled in the affected establishments, 187 detections in captive birds, and 3,573 HPAI events in wild birds, affecting 37 European countries (European Centre for Disease Prevention and Control, 2022). The 2021/22 outbreak is considered the largest and most devastating HPAI outbreak ever to occur in Europe and, more recently, in the United States, where 46 American states were affected, resulting in 49 million dead birds (death as a result of AIV infection or culling) (CDC, 2022).

Because avian influenza viruses are highly contagious, and given the potential risk of long-distance spread of HPAI viruses from infected barns, it is crucial to conduct real-time genetic investigations to track circulating AIV strains. This real-time genetic monitoring of circulating strains aims to identify possible links between outbreaks in order to ultimately break the chain of infection.

This work aims to combine a multiplex PCR amplification of H5 avian influenza viruses with nanopore sequencing to evaluate the pathotypes and clades of circulating viruses, and identify the genetic links between 30 outbreaks that occurred in France between 2020 and 2022.

Materials and methods

Ethics statement

The procedures with animals implemented in the investigation of this study have been conducted in strict compliance with the French and European regulations on animal experiment and notifiable diseases: all samples were taken in the framework of the official surveillance of Highly Pathogenic Avian Influenza, under the supervision of French official veterinary services and in compliance with French legislation on veterinary practices, since they correspond strictly to sampling procedures routinely performed for official surveillance. They were also performed under the supervision of the Ethical Committee “Sciences & Santé Animales” n°115, Toulouse, France, according to EU law (Directive 2010/63/UE).

Sampling and nucleic acids extraction

In a first attempt, four AIV isolates of highly pathogenic (H5N8) and low pathogenicity avian influenza viruses (H5N3, H6N1, H7N9 and H9N2) (Table 1) were included in a validation assay and were processed under the same conditions described below for the clinical samples.

We conducted our study in seven administrative divisions of southwestern and western France (map available as a supplementary file, Map_France), accounting for more than 60% of French duck production. We sampled 30 flocks and collected tracheal swabs, feathers and dust, the latter two being known to be replication sites (Gaide et al., 2022) or carriers (Filaire et al., 2022) of HPAIV, respectively. The sampling information can be found in the supplementary file Data Sheet 1. On eleven farms, dust was collected using wipes as previously described (Filaire et al., 2022) and the swabs and immature feathers were taken from 20 animals per farm and analyzed individually with the Influenza A Real-time RT-PCR assay described below. Then, the swabs, feathers or dust samples with the lowest Ct values were selected and submitted to the NGS pipeline.

All samples were processed in a BSL3 laboratory until lysis in strict compliance with biosafety procedures. Swabs, feather pulps and wipes were placed in 1X PBS and vigorously vortexed for 30 s before viral RNA extraction using the Nucleospin RNA virus kit (Macherey Nagel). Viral RNA was then stored at -80°C until further use.

TABLE 1 Size (bp) of the consensus sequence obtained by nanopore sequencing on different AIV subtypes.

AIV subtype	NCBI taxon ID	Ct value (M gene)	PB2	PB1	PA	HA	NP	NA	M	NS
H5N3	437394	32	290	1109	498	1711	–	263	943	845
H6N1	1096539	32,41	457	1338	2158	1711	312	1364	989	845
H7N1	437402	31,84	2121	2110	2158	1693	1501	1354	989	845
H9N2	OQ179925 (access. number)	30,23	638	1101	998	161	1501	291	989	845

Influenza A Real-time RT-PCR assay

RT-qPCR was performed on 2 μ L of viral RNA using the ID GeneTM Influenza A Duplex PCR kit (IDVet, Grabels, France) for the detection of all type A influenza viruses. The PCR amplifications were run on a LC96 thermocycler (Roche, Basel, Switzerland) with the following parameters: a reverse transcription step of 10 min at 45°C followed by 10 min at 95°C and 40 cycles of 15 s at 95°C coupled with 60 s at 60°C.

Multiplex-PCR amplification

Hemagglutinin and whole genome amplification were performed using PCR primers pools made of oligonucleotides selected in the literature (Table 2). The selection criteria were (i) the hybridization sites with each primers pair encompassing the cleavage site, (ii) specificity for the H5 hemagglutinin gene and (iii) size diversity among the PCR products.

The PCR amplifications were performed as follows: extracted RNA was reverse transcribed from 10 μ L of RNA using a RevertAid First Strand cDNA Synthesis Kit (ThermoFisher, Waltham, USA) with 0.5 μ M of specific-primer (MBTuni-12 [5'-ACGCGTGATCAGCAAAAAGCAGG-3']) (Zhou et al., 2009). Two PCR assays were run separately and then mixed. First, the variable genomic regions encompassing the CS region were

amplified by a primer set designed to anneal the Hemagglutinin belonging to clade 2.3.4.4 (H5HP). Second, the eight genomic segments of influenza A virus were amplified in a single reaction inspired by Zhou and Wentworth with the following universal primers: MBTuni-12 [5'-ACGCGTGATCAGCAAAAAGCAGG] and MBTuni-13 [5'-ACGCGTGATCAGTAGAAACAAGG] (Zhou et al., 2009). The PCR reactions were performed in a final volume of 20 μ L using the PhusionTM High-Fidelity DNA Polymerase (ThermoFisher) with 8 μ L of primer mix H5HP (composed of 1 μ M of each primer) or 0.5 μ M of universal primers. The temperature cycle parameters were 98°C for 30 s, followed by 35 cycles (98°C for 10 s, 58°C for 20 s, and 72°C for 3 min) with the final extension at 72°C for 10 min. To assess the PCR amplification products, 5 μ L of each PCR product was checked on an agarose gel to roughly estimate the amplicons' sizes. Each sample produced satisfying amplicons for submission to nanopore sequencing. Such information is afterward critical for the library molarity calculation, which is based on amplicon size and the dsDNA concentration of the PCR products.

Nanopore sequencing

PCR products were pooled and purified with a 1:1 ratio of AMPure XP beads. DNA libraries were prepared from 200 fmol of purified PCR

TABLE 2 Names, positions, sequences and references of the PCR primers used in this study.

Primer name	Position in the genome	Sequence (5'-3')	Ref
H5-1F	266	TGCCGGAATGGTCTTACATAGTG	(Ng et al., 2005)
fehH5_F	783	RAAYTTYATTGCTCCAGAAWATGCATAC	(Leijon et al., 2011)
H5_HP_EA_F1	984	CCTTGCDACTGGRCTCAG	(Naguib et al., 2017)
H5_HA_FP	985	CTTGCGACTGGGCTCAGAAAT	(Zhang et al., 2017)
H5_F_HA	127	ACACATGCYCARGACATACT	(Lee et al., 2001)
Kha_1_F	799	CCTCCAGARTATGCMTAYAAAATTGTC	(Slomka et al., 2007)
1057.1 F	967	GGRGAATGCCCAAATAYGT	(Gall et al., 2008)
1057.2 F	967	GGRARATGCCCCAGRTATGT	(Gall et al., 2008)
1057.3 F	967	GGRGAATGCCCAAARTAYAT	(Gall et al., 2008)
Bm_HA_1_F	1	TATTCGTCTCAGGGAGCAAAAAGCAGGGG	(Hoffmann et al., 2001)
H5-2R	1692	TCTGCATTGTAACGACCCATTG	(Ng et al., 2005)
rehH5_R	1100	TACCAACCGTCTACCATKCCYTG	(Leijon et al., 2011)
H5_HP_EA_R	1075	GTCTACCATTCCYTGCCA	(Naguib et al., 2017)
H5_3R	1130	CTCCCCTGCTCATGTGCTATG	(Jayawardena et al., 2007)
H5_B2a_R	1210	TTTTGTCAATGATTGAGTTGACCTTATTGG	(Pálmai et al., 2007)
1232.1_R	1142	CTGAGTCCGAACATTGAGTTGCTATGVTGRTAWCCATACCA	(Gall et al., 2008)
1232.2 R	1142	CTGAGTCCGAACATTGAGTTGATGYCTGAADCCRTACCA	(Gall et al., 2008)
Bm_NS_890R	1778	ATATCGTCTCGTATTAGTAGAAACAAGGGTGTTTT	(Hoffmann et al., 2001)
H5_HA_RP	1163	TTTGGGTGGATTCTTTGTCTGC	(Zhang et al., 2017)
kha 3	1100	TACCAACCGTCTACCATKCCYTG	(Slomka et al., 2007)

products using an SQK-LSK109 Ligation sequencing kit supplied by ONT (Oxford Nanopore Technologies, Oxford, United Kingdom) associated with an EXPNBD104 (ONT) native barcoding kit for the multiplexing of samples. For each run, 20 fmol DNA libraries were loaded on a FLO-FLG001 flongle and were run on a MinION Mk1C device (ONT) for 10 hours. High accuracy base-calling was performed in real-time with Guppy (v3.5) embedded in the MK1C software (v19.12.12) with the ‘Trim Barcode’ option on (ONT).

Sequencing data analysis

The quantity and quality of data were checked by looking at the sequencing report generated by MinKNOW 2.0 and using NanoPlot from the NanoPack tool set (De Coster et al., 2018). The fastq files were aligned using minimap2 (Li, 2018) and the samtools (Danecek et al., 2021) package. The HA and whole-genome consensus sequences were generated using the consensus command of the iVar (Grubaugh et al., 2019) pipeline that embeds the samtools mpileup tool (Li, 2011). The consensus sequences were then manually checked using Bioedit (v7.2.6) and IGV (v2.8.2) software. Nanoplot (Python v3.6.3) and BBmap (v.38.31) were used to assess the quality of the NGS reads and to visualize the coverage of the genome.

The command lines were used as follows:

```
minimap2 -ax map-ont reference.fa mysample.fastq
> mysample.sam
samtools sort mysample.sam > mysample.sort.bam
samtools index mysample.sort.bam
samtools flagstat mysample.sort.bam > mysample_flagstat
samtools depth mysample.sort.bam > mysample_samdepth
samtools mpileup -A -Q 0 mysample.sort.bam | ivar consensus
-p prefix -q 10 -t 0 -m 1
```

Phylogeny

Two phylogenetic trees were built with (i) a set of 282 HA sequences from the 2020/21 H5N8 epizootics (including the 11 sequences from this study) and (ii) a set of 91 HA sequences from the 2021/22 H5N1 epizootics (including the 19 sequences from this study). Both sets of sequences are available on the GISAID EpiFlu™ Database and are listed in the [Supplementary File Table 2](#).

The sequences were selected on the basis of their sampling locations and dates and were processed as follows: multiple sequences alignment with MAFFT version 7 (Kuraku et al., 2013; Katoh et al., 2019).

Maximum Likelihood trees construction with IQ-TREE (Nguyen et al., 2015; Trifinopoulos et al., 2016), the model of substitution selected by ModelFinder (Kalyaanamoorthy et al., 2017) was K3Pu+F+I+G4 for the 2020-2021 H5N8 tree and K3Pu+F+G4 for the 2021-2022 H5N1 tree. The phylogenetic bootstrapping (1000 replicates) was performed with UFBoot

(Hoang et al., 2018). Finally, the tree figures were generated using FigTree v1.4.4.

Statistical analysis

Statistical analyses were performed using R Studio software (<https://www.R-project.org>), with R version 4.1.1 (2021-08-10), to compare Ct value and percentage of data aligning with the HA segment as a function of sample type. Data for each variable were analyzed using non-parametric tests, Kruskal-Wallis and Wilcoxon-Mann-Whitney Test, with Holm’s p-value adjustment method.

Results

In silico validation of PCR-based enrichment

An *in-silico* analysis consisting in the alignment of 20 PCR primers (Table 2) selected in the literature with a panel of H5 hemagglutinin was first performed for the constitution of the multiplex PCR panel. The selection of the PCR primers was based on their hybridization locations across the HA segment to guarantee a whole or partial gene amplification including the cleavage site. We also included MBTuni-12 and MBTuni-13 universal primers (Zhou et al., 2009) for the full amplification of any AIV subtype. The amplicon sizes produced by the different combinations of PCR primers are available in Table 3.

Validation of the multiplex-PCR sequencing on different AIV subtypes

The evaluation of the PCR panel was then performed on several AIV subtypes of highly pathogenic (H5N8) and low pathogenicity avian influenza viruses (H5N3, H6N1, H7N1 and H9N2). These isolates were submitted to our nanopore protocol: RNA extraction, PCR amplification and ONT sequencing. These assays resulted in a complete coverage of the HA gene and a near-complete to whole sequencing of the seven other RNA segments (Table 1).

The resulting PCR products obtained from two H5N8 sample were loaded on an agarose gel to demonstrate the amplification of the HA in several parts, as well as the whole genome amplification with the universal primers (Supplementary file image 1).

Application of the multiplex-PCR sequencing pipeline on two sets of field samples originating from 11 H5N8 and 19 H5N1 HPAIV outbreaks sampled in 2020-2021 and 2021-2022, respectively

Two sets of field samples were included in this study, both originating from French outbreaks: 11 from December 2020 to

TABLE 3 Hemagglutinin amplicons sizes (bp) according to primer combination.

Rev primer name →	H5-2R	REHH5 R	H5 HP EA R	H5 3R	H5 B2A R	1232.1 R	1232.2 R	BM NS 890R	HA RP	KHA 3
For primer name ↓										
I057.1 F	1426	834	809	864	944	876	876	1512	897	834
H5 HP EA F1	909	317	292	347	427	359	359	995	380	317
KHA 1 F	708	116	91	146	226	158	158	794	179	116
FEHH5 F	707	115	90	145	225	157	157	793	178	115
H5 F HA	1565	973	948	1003	1083	1015	1015	1651	1036	973
I057.3 F	893	301	276	331	411	343	343	979	364	301
H5 HA FP	725	133	108	163	243	175	175	811	196	133
H5-1F	725	133	108	163	243	175	175	811	196	133
I057.2 F	725	133	108	163	243	175	175	811	196	133
BM HA 1 F	1692	1100	1075	1130	1210	1142	1142	1778	1163	1100

February 2021, and 19 from late November 2021 to February 2022 (Supplementary File Data Sheet 1). The Ct values ranged between 13.03 and 28.5, which is representative of high and low viral loads, respectively. These field samples included mostly tracheal swabs, but also feather pulp, cloacal swabs and dust wipes, to cover the different types of samples available for AIV surveillance.

Each sample was submitted to HA and whole-genome amplification using the multiplex PCR pipeline described previously. The sequencing yields varied from 2.34 to 474.13 Mb per sample, while the N50 values were comprised between 469 and 1 818 bp. The amount of data corresponding to the AIV allowed a sequencing depth of 10x on the entire HA gene for 27 of the 30 samples Supplementary File Data Sheet 1.

The 30 figures corresponding to the 30 sequenced samples are displayed in Figure 1. Influenza genomes were not systematically completely sequenced, but HA was always amplified with a sequencing depth varying roughly from 10^2 to 10^5 X.

The sequencing metrics such as read length, N50 and number of reads for each sample are available in the Supplementary File Data Sheet 1. An example of cumulative output (bases) is shown in the Supplementary File Image 2A, corresponding to a run where four samples were multiplexed on a flongle flowcell. The Supplementary File Image 2B and 2C show a read length histogram and the bases count per barcode, respectively. Finally, the Supplementary File Image 2D shows the coverage and sequencing depth obtained for one sample.

The consensus sequences were obtained for each hemagglutinin, using iVar, and were deposited in the Genbank database.

Phylogeny

Two phylogenetic trees (Supplementary File Data Sheets 2 and 3) were built in order to determine whether the outbreaks were related to each other (e.g. farm network) or were due to new viral introductions (e.g. an unnoticed outbreak, introduction from wild birds). The tree of the Supplementary File Data Sheet 2 shows the

relationship between nine outbreaks (blue taxa) in southwest France (Gers, Hautes-Pyrénées, Landes) between 6 December 2020 and 20 February 2021. In addition, at a distance of about 300 km (in Vendée and Deux-Sèvres), two other outbreaks (green taxa) were sampled on 13 December 2020. The tree in the Supplementary File Data Sheet 2 shows a common recent ancestor to the 11 French sequences, but it is clear that the nine outbreaks in the southwest are not directly related to the outbreaks in Vendée and Deux-Sèvres. The tree of the Supplementary File Data Sheet 3 shows 19 outbreaks clustering in six different groups. A set of 11 sequences (pink taxa) from southwest France sampled over a 4-week period (18 December 2021 to 19 January 2022) illustrates the circulation of one viral strain within a restricted area, suggesting a link between these farms that could be the source of the virus circulation. In addition, in the same area and approximately at the same time, we analyzed six outbreaks (yellow, light red, dark red taxa) which group in three different clusters that were clearly not related to the 11 infected farms described above (pink taxa).

Statistical analysis

Statistical analysis established a correlation (i) between sample type and the Ct value of the M gene qPCR (Supplementary File Image 3) and (ii) between sample type and the proportion of sequencing data belonging to the HA gene (Supplementary File Image 4). In both cases, feathers emerged as the most relevant samples, with significantly lower Ct values and significantly higher HA data than for dust and tracheal swabs.

Discussion

In this study, we aimed to amplify in a single multiplex PCR reaction a large range of AIV strains. We focused first on those belonging to the H5 subtype. Sequencing the HA was a priority and

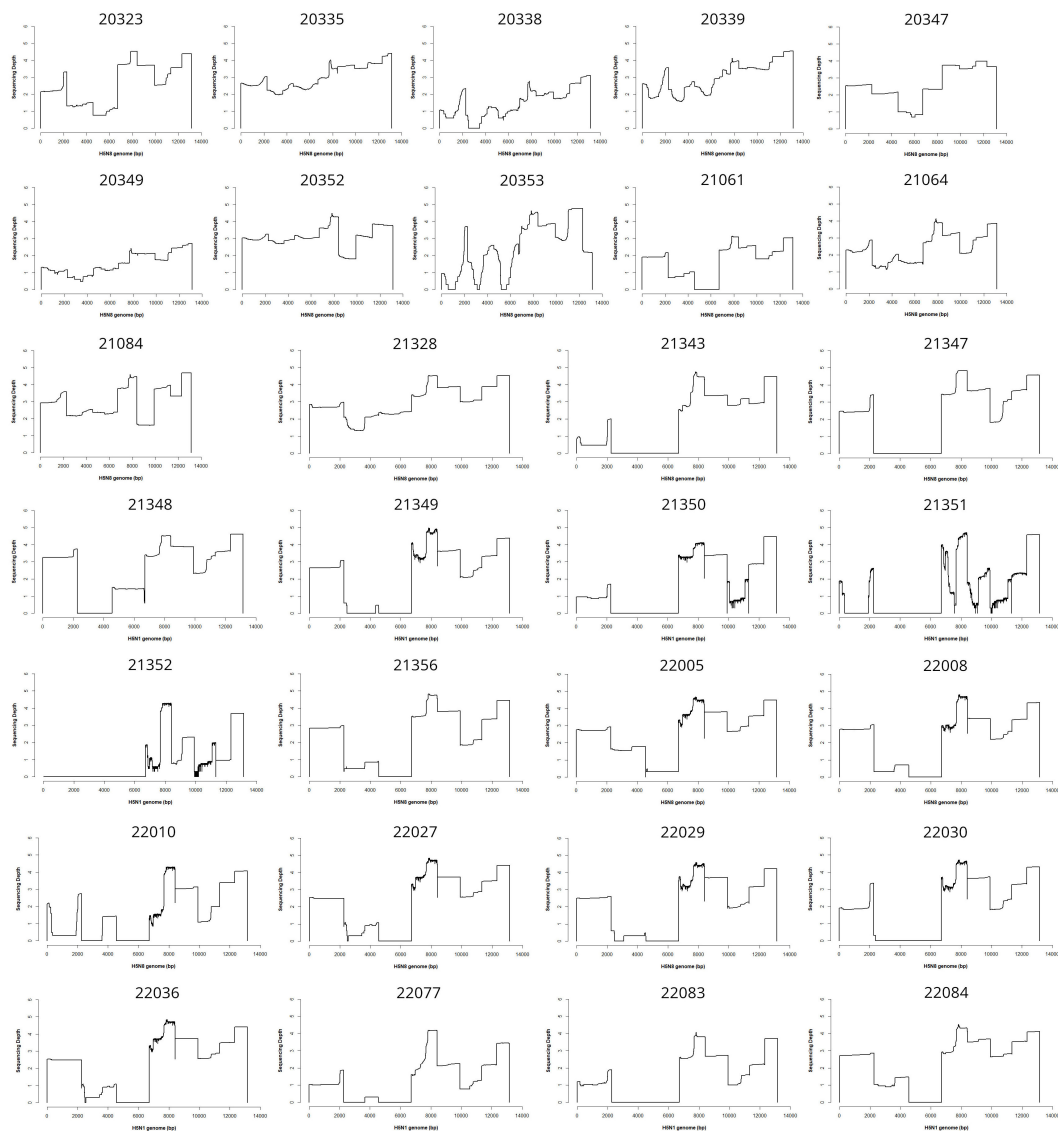


FIGURE 1

Coverage plots. This file shows the coverage and sequencing depth (logarithmic scale) plots obtained for the 30 sequenced samples all over the genomes.

the *in-silico* design was performed to guarantee, at the minimum, the amplification of the cleavage site of the HA gene for pathotyping. As the viral loads and the integrity of the RNA directly extracted from field samples could be seriously weak or damaged, the PCR panel was designed to produce small (i.e. < 250 bp) to large (>700 bp) amplicons that systematically encompassed the HA cleavage site. Moreover, universal AIV PCR primers were included for a broad-range detection of all type A influenza viruses. While official analyses require tracheal or cloacal swabs, alternative samples have shown promising performances for AIV detection, such as feathers (Gaide et al., 2022) and dust sampling in poultry farms (Filaire et al., 2022). The relevance of feathers was confirmed, as this type of samples provided more AIV-specific sequencing data. Based on an amplicon-based approach, the workflow described here is intended to focus on influenza viruses with an increased sensitivity, and is not designed for an unbiased metagenomic

analysis. This is intentional as AIVs, due to their pathological and epidemiological significance, require a fully dedicated tool that efficiently provides the data needed for surveillance. Due to the genetic evolution of AIVs, a constant update of the PCR panel will be required.

As summarized in previous studies (BioWatch PCR Assays: Building Confidence, Ensuring Reliability; Abbreviated Version, 2015), two main approaches can be used for viral detection in animals or their environment. PCR (either monoplex or multiplex) remains the mostly widely used as a surveillance or diagnostic tool due to (i) its specificity and low detection threshold, (ii) the inexpensive and simple application of the method and (iii) the robustness and simplicity of its interpretation. However, some limitations have been pointed out, such as nonspecific hybridization, the limited resolution of the information provided, and the difficulty in detecting highly diverse viruses and those for

which detailed genetic information (such as pathotype) is crucial. Moreover, PCR cannot handle the detection of minority variants, which can be relevant for viral surveillance.

For these reasons, monitoring protocols increasingly include the use of Next Generation Sequencing (NGS) or third generation sequencing methods that are known to rapidly provide very high-throughput data. Manufacturers in the molecular biology field are now developing tools enabling any laboratory operator to perform NGS in the laboratory or even in the field. However, although these assays are attractive and perform quite well, their running time and cost render them unfeasible for routine clinical applications. Furthermore, online user-friendly bioinformatics tools such as EPI2ME WIMP workflow do not allow the personalized configuration level that can be reached with open source algorithms which also can be run offline. Offline real-time classification pipelines based on open source tools have already been developed (Deshpande et al., 2019).

Another critical aspect of NGS is the amount of data required for the detection of scarce pathogens whose nucleic acids can be lost among non-viral sequences. For example, a study on Zika virus sequencing showed that 2 million reads were required for the detection of 223 viral reads (Quick et al., 2017). Since the beginning of metagenomics, the scientific community has done everything possible for the development of enrichment techniques, such as host nucleic acids depletion (Matranga et al., 2014) and probe capture (Li et al., 2020) followed by massive deep sequencing. These techniques are interesting for a global and exhaustive description of pathogen communities present in a clinical or environmental sample. Nonetheless, regardless of the enrichment method and sequencing technology, NGS cannot offer detection levels as sensitive as those obtained with PCR. Because PCR offers a fast, cheap and convenient target enrichment, several studies have already shown the advantages of combining PCR with NGS (Quick et al., 2017; Onda et al., 2018; Kim et al., 2018).

However, metagenomics should also be considered for its capacity to detect relevant coinfecting pathogens, as well as potential emerging or re-emerging epizootic microorganisms. A shotgun metagenomic workflow for the detection of a wide variety of microbial genomes would require hundreds of megabases per sample. This kind of protocol therefore cannot be routinely applied in veterinary surveillance, but it could be used for field cases of great clinical or epidemiological importance.

ONT technology is highly scalable, from the smallest consumable unit, namely “Flongle”, to high throughput platforms. The Flongle flow cell used in this study allowed a sequential analysis of the avian influenza outbreaks during the course of the epizootics. In addition to enabling low scale sequencing runs, the separate management of each case was an effective way to prevent cross-contamination. Moreover, while the accuracy of sequencing is often identified as a weakness of nanopore technology, it was demonstrated that a sequencing depth cutoff of 10x resulted in a consensus sequence accuracy of 99.95-100%, identical to Illumina sequences (Xu et al., 2021). As we obtained a sequencing depth of 10x over the whole HA gene for 27 of the 30 samples, confidence in

the quality of the consensus sequence is high. This type of pipeline, combined with a larger quantity of sequencing data, could therefore be adapted to the detection of minority variants. The bioinformatics pipeline described in this manuscript was deliberately based on open-source algorithms that are freely available on dedicated internet platforms. Based on our dataset, we were able to identify AIV HA subtypes, the sequence of the cleavage site and potential clusters of outbreaks. One of the main concerns voiced by government veterinary services is the need to unveil links between outbreaks. These links may suggest direct or indirect farm-to-farm contamination routes. In contrast, when an emerging strain is identified in a region, this may indicate a new introduction, for example from wild birds.

Beyond the application presented in this study, Oxford Nanopore sequencing of AIV whole genomes can contribute to the analysis of Defective Interfering Particles (DIPs). Since DIPs are generated through incorporation of highly truncated forms of genome segments, the long sequences generated represent a clear asset. As DIPs clearly play a role in the biology of influenza viruses (Świątoń et al., 2020; Mendes and Russell, 2021), their study through a routine sequencing protocol represents an advantage. The phylogenetic analysis carried out in this study also revealed greater HA variability in the 2021-22 epizootic compared with 2020-21. Within each cluster, however, related strains showed genetic evolution, suggesting local circulation of strains. The fact that the different outbreaks analyzed were divided into distinct clusters also suggests that there was no diffusion between the different groups of outbreaks, suggesting that the spread of the virus was restricted to adjacent areas. More generally, these phylogenetic trees show that this type of protocol can rapidly determine whether there is viral circulation between outbreaks or whether it is a new introduction. These data, coupled with epidemiological data, can help to understand virus circulation and, if necessary, break transmission chains.

The protocol presented in this study can be carried out for a set of 30 samples in less than 24 hours, starting from the reception of the samples to the production of the consensus sequences. This turnaround time could, however, be reduced by the use of (i) a simplified extraction process, (ii) an automatized PCR amplification and library preparation (eg: voltrax) and a built-in analysis pipeline such as the RAMPART model (RAMPART).

We propose here a MinION workflow for the rapid typing of avian influenza viruses, evaluated on clinical samples, without any amplification step on cell cultures or embryonated eggs. The objective of this proof-of-concept study was not to perform a comprehensive phylogenetic survey, but rather to demonstrate the possible contribution of this workflow to the real-time management of viral epizootics by unveiling very quickly possible links and specific patterns of transmission. The implementation of this kind of workflow – from the sample to the phylogenetic tree – in the framework of official surveillance will obviously need further validation by reference laboratories, but it addresses a substantial need for the surveillance of AIVs and, more broadly, emerging viruses.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/[Supplementary Material](#).

Ethics statement

The animal study was approved by Ethical Committee “Sciences & Santé Animales” n°115, Toulouse, France, according to EU law (Directive 2010/63/UE). The study was conducted in accordance with the local legislation and institutional requirements.

Author contributions

GC: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. MW: Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing – review & editing. AS: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. LL: Formal analysis, Investigation, Writing – review & editing. SS: Formal analysis, Investigation, Methodology, Writing – review & editing. FF: Formal analysis, Investigation, Methodology, Writing – review & editing. JG: Conceptualization, Funding acquisition, Methodology, Project administration, Resources, Supervision, Validation, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This study was funded by (i) the Direction Generale de l'Alimentation,

References

- BioWatch PCR Assays: Building Confidence, Ensuring Reliability; Abbreviated Version (2015) (Washington, D.C).
- Causey, D., and Edwards, S. V. (2008). Ecology of avian influenza virus in birds. *J. Infect. Dis.* 197, S29–S33. doi: 10.1086/524991
- CDC (2022) *U.S. Approaches Record Number of Avian Influenza Outbreaks* (Centers for Disease Control and Prevention). Available at: <https://www.cdc.gov/flu/avianflu/spotlights/2022-2023/nearing-record-number-avian-influenza.htm> (Accessed November 28, 2022).
- Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., et al. (2021). Twelve years of SAMtools and BCFtools. *Gigascience* 10, giab008. doi: 10.1093/gigascience/giab008
- De Coster, W., D'Hert, S., Schultz, D. T., Cruts, M., and Van Broeckhoven, C. (2018). NanoPack: visualizing and processing long-read sequencing data. *Bioinformatics* 34, 2666–2669. doi: 10.1093/bioinformatics/bty149
- Deshpande, S. V., Reed, T. M., Sullivan, R. F., Kerkhof, L. J., Beigel, K. M., and Wade, M. M. (2019). Offline next generation metagenomics sequence analysis using MinION Detection Software (MINDS). *Genes (Basel)* 10 (8), 578. doi: 10.3390/genes10080578
- European Centre for Disease Prevention and Control (2021) *Avian influenza overview February – May 2021*. Available at: <https://www.ecdc.europa.eu/en/publications-data/avian-influenza-overview-february-may-2021> (Accessed January 11, 2022).
- European Centre for Disease Prevention and Control (2022) *2021-2022 data show largest avian flu epidemic in Europe ever*. Available at: <https://www.ecdc.europa.eu/en/news-events/2021-2022-data-show-largest-avian-flu-epidemic-europe-ever> (Accessed November 28, 2022).
- Filaire, F., Lebre, L., Foret-Lucas, C., Vergne, T., Daniel, P., Lelièvre, A., et al. (2022). Highly pathogenic avian influenza A(H5N8) clade 2.3.4.4b virus in dust samples from poultry farms, France 2021. *Emerg. Infect. Dis.* 28, 1446–1450. doi: 10.3201/eid2807.212247

Ministère de l'Agriculture et de l'Alimentation, France and (ii) the “France Futur Elevage” Carnot of INRAE Institute in collaboration with the French Institut Carnot “Microbes-Santé”, in the framework of the FIELD project.

Acknowledgments

We are grateful to the genotoul bioinformatics platform Toulouse Occitanie (Bioinfo Genotoul, <https://doi.org/10.15454/1.5572369328961167E12>) for providing help and/or computing and/or storage resources. We would like to thank farmers for their collaboration with the sampling of their flocks. We would like to thank Theseo France, part of LanXess group, for the loan of the MinION MK1C and for funding F.F.'s PhD program.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcimb.2024.1257586/full#supplementary-material>

- Gaïde, N., Lucas, M.-N., Delpont, M., Croville, G., Bouwman, K. M., Papanikolaou, A., et al. (2022). Pathobiology of highly pathogenic H5 avian influenza viruses in naturally infected Galliformes and Anseriformes in France during winter 2015–2016. *Vet. Res.* 53, 11. doi: 10.1186/s13567-022-01028-x
- Gall, A., Hoffmann, B., Harder, T., Grund, C., and Beer, M. (2008). Universal primer set for amplification and sequencing of HAO cleavage sites of all influenza A viruses. *J. Clin. Microbiol.* 46, 2561–2567. doi: 10.1128/JCM.00466-08
- Grubaugh, N. D., Gangavarapu, K., Quick, J., Matteson, N. L., De Jesus, J. G., Main, B. J., et al. (2019). An amplicon-based sequencing framework for accurately measuring intrahost virus diversity using PrimalSeq and iVar. *Genome Biol.* 20, 8. doi: 10.1186/s13059-018-1618-7
- Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q., and Vinh, L. S. (2018). UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Mol. Biol. Evol.* 35 (2), 518–522. doi: 10.1093/molbev/msx281
- Hoffmann, E., Stech, J., Guan, Y., Webster, R. G., and Perez, D. R. (2001). Universal primer set for the full-length amplification of all influenza A viruses. *Arch. Virol.* 146, 2275–2289. doi: 10.1007/s007050170002
- International Committee on Taxonomy of Viruses (ICTV) *Orthomyxoviridae*. Available at: <https://talk.ictvonline.org/taxonomy/> (Accessed May 25, 2021).
- Jayawardena, S., Cheung, C. Y., Barr, I., Chan, K. H., Chen, H., Guan, Y., et al. (2007). Loop-mediated isothermal amplification for influenza A (H5N1) virus. *Emerging Infect. Dis. J. CDC* 13 (6), 899–901. doi: 10.3201/eid1306.061572
- Kalyanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A., and Jermini, L. S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* 14 (6), 587–589. doi: 10.1038/nmeth.4285
- Katoh, K., Rozewicki, J., and Yamada, K. D. (2019). MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Briefings Bioinf.* 20, 1160–1166. doi: 10.1093/bib/bbx108
- Kim, W.-K., No, J. S., Lee, S.-H., Song, D. H., Lee, D., Kim, J.-A., et al. (2018). Multiplex PCR-based next-generation sequencing and global diversity of Seoul virus in humans and rats. *Emerging Infect. Dis. J. CDC* 24 (2), 249–257. doi: 10.3201/eid2402.171216
- Ku, K. B., Park, E. H., Yum, J., Kim, J. A., Oh, S. K., and Seo, S. H. (2014). Highly pathogenic avian influenza A(H5N8) virus from waterfowl, South Korea. *Emerg. Infect. Dis.* 20, 1587–1588. doi: 10.3201/eid2009.140390
- Kuraku, S., Zmasek, C. M., Nishimura, O., and Katoh, K. (2013). aLeaves facilitates on-demand exploration of metazoan gene family trees on MAFFT sequence alignment server with enhanced interactivity. *Nucleic Acids Res.* 41, W22–W28. doi: 10.1093/nar/gkt389
- Lee, M. S., Chang, P. C., Shien, J. H., Cheng, M. C., and Shieh, H. K. (2001). Identification and subtyping of avian influenza viruses by reverse transcription-PCR. *J. Virol. Methods* 97, 13–22. doi: 10.1016/s0166-0934(01)00301-9
- Lee, Y.-J., Kang, H.-M., Lee, E.-K., Song, B.-M., Jeong, J., Kwon, Y.-K., et al. (2014). Novel reassortant influenza A(H5N8) viruses, South Korea. *Emerg. Infect. Dis.* 20, 1087–1089. doi: 10.3201/eid2006.140233
- Leijon, M., Ullman, K., Thyselius, S., Zohari, S., Pedersen, J. C., Hanna, A., et al. (2011). Rapid PCR-based molecular pathotyping of H5 and H7 avian influenza viruses. *J. Clin. Microbiol.* 49, 3860–3873. doi: 10.1128/JCM.01179-11
- Li, H. (2011). A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* 27, 2987–2993. doi: 10.1093/bioinformatics/btr509
- Li, H. (2018). Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* 34, 3094–3100. doi: 10.1093/bioinformatics/bty191
- Li, B., Si, H.-R., Zhu, Y., Yang, X.-L., Anderson, D. E., Shi, Z.-L., et al. (2020). Discovery of bat coronaviruses through surveillance and probe capture-based next-generation sequencing. *mSphere* 5 (1). doi: 10.1128/mSphere.00807-19
- Liu, J., Xiao, H., Lei, F., Zhu, Q., Qin, K., Zhang, X.-W., et al. (2005). Highly pathogenic H5N1 influenza virus infection in migratory birds. *Science* 309, 1206. doi: 10.1126/science.1115273
- Matranga, C. B., Andersen, K. G., Winnicki, S., Busby, M., Gladden, A. D., Tewhey, R., et al. (2014). Enhanced methods for unbiased deep sequencing of Lassa and Ebola RNA viruses from clinical and biological samples. *Genome Biol.* 15, 519. doi: 10.1186/s13059-014-0519-7
- Mendes, M., and Russell, A. B. (2021). Library-based analysis reveals segment and length dependent characteristics of defective influenza genomes. *PLoS Pathog.* 17, e1010125. doi: 10.1371/journal.ppat.1010125
- Naguib, M. M., Graaf, A., Fortin, A., Luttermann, C., Wernery, U., Amarín, N., et al. (2017). Novel real-time PCR-based patho- and phylotyping of potentially zoonotic avian influenza A subtype H5 viruses at risk of incursion into Europe in 2017. *Euro Surveill* 22, 30435. doi: 10.2807/1560-7917.ES.2017.22.1.30435
- Ng, E. K. O., Cheng, P. K. C., Ng, A. Y. Y., Hoang, T. L., and Lim, W. W. L. (2005). Influenza A H5N1 detection. *Emerg. Infect. Dis.* 11, 1303–1305. doi: 10.3201/eid1108.041317
- Nguyen, L. T., Schmidt, H. A., von Haeseler, A., and Minh, B. Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* 14 (6), 587–589. doi: 10.1038/nmeth.4285
- Olsen, B., Munster, V. J., Wallensten, A., Waldenström, J., Osterhaus, A. D. M. E., and Fouchier, R. A. M. (2006). Global patterns of influenza A virus in wild birds. *Science* 312, 384–388. doi: 10.1126/science.1122438
- Onda, Y., Takahagi, K., Shimizu, M., Inoue, K., and Mochida, K. (2018). Multiplex PCR targeted amplicon sequencing (MTA-seq): simple, flexible, and versatile SNP genotyping by highly multiplexed PCR amplicon sequencing. *Front. Plant Sci.* 9. doi: 10.3389/fpls.2018.00201
- Pálmai, N., Erdélyi, K., Bálint, Á., Márton, L., Dán, Á., Deim, Z., et al. (2007). Pathobiology of highly pathogenic avian influenza virus (H5N1) infection in mute swans (*Cygnus olor*). *Avian Pathol.* 36, 245–249. doi: 10.1080/03079450701341957
- Poen, M. J., Bestebroer, T. M., Vuong, O., Scheuer, R. D., Jeugd, H. P., Kleyheeg, E., et al. (2017). Local amplification of highly pathogenic avian influenza H5N8 viruses in wild birds in the Netherlands 2016 to 2017. *Eurosurveillance* 23, 17. doi: 10.2807/1560-7917.ES.2018.23.4.17-00449
- Quick, J., Grubaugh, N. D., Pullan, S. T., Claro, I. M., Smith, A. D., Gangavarapu, K., et al. (2017). Multiplex PCR method for MinION and Illumina sequencing of Zika and other virus genomes directly from clinical samples. *Nat. Protoc.* 12, 1261–1276. doi: 10.1038/nprot.2017.066
- RAMPART (rampart). Available at: <https://artic-network.github.io/rampart/> (Accessed January 11, 2022).
- Slomka, M. J., Coward, V. J., Banks, J., Löndt, B. Z., Brown, I. H., Voermans, J., et al. (2017). Identification of sensitive and specific avian influenza polymerase chain reaction methods through blind ring trials organized in the European Union. *Avian Dis.* 51, 227–234. doi: 10.1637/7674-063006R1.1
- Świątęto, E., Tarasiuk, K., and Śmietanka, K. (2020). Low pathogenic avian influenza virus isolates with different levels of defective genome segments vary in pathogenicity and transmission efficiency. *Veterinary Res.* 51, 108. doi: 10.1186/s13567-020-00833-6
- Trifinopoulos, J., Nguyen, L. T., von Haeseler, A., and Minh, B. Q. (2016). W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Res.* 44 (W1), W232–W235. doi: 10.1093/nar/gkw256
- Wu, H., Peng, X., Xu, L., Jin, C., Cheng, L., Lu, X., et al. (2014). Novel reassortant influenza A(H5N8) viruses in domestic ducks, eastern China. *Emerg. Infect. Dis.* 20, 1315–1318. doi: 10.3201/eid2008.140339
- Xu, Y., Lewandowski, K., Downs, L. O., Kavanagh, J., Hender, T., Lumley, S., et al. (2021). Nanopore metagenomic sequencing of influenza virus directly from respiratory samples: diagnosis, drug resistance and nosocomial transmission, United Kingdom 2018/19 influenza season. *Eurosurveillance* 26, 2000004. doi: 10.2807/1560-7917.ES.2021.26.27.2000004
- Zhang, Z., Liu, D., Sun, W., Liu, J., He, L., Hu, J., et al. (2017). Multiplex one-step Real-time PCR by Taqman-MGB method for rapid detection of pan and H5 subtype avian influenza viruses. *PLoS One* 12, e0178634. doi: 10.1371/journal.pone.0178634
- Zhou, B., Donnelly, M. E., Scholes, D. T., St. George, K., Hatta, M., Kawakoba, Y., et al. (2009). Single-reaction genomic amplification accelerates sequencing and vaccine production for classical and swine origin human influenza A viruses. *J. Virol.* 83, 10309–10313. doi: 10.1128/JVI.01109-09