Wastewater monitoring of human and avian influenza A viruses in Northern Ireland: a genomic surveillance study

Andrew J Lee, Stephen Carson, Marina I Reyne, Andrew Marshall, Daniel Moody, Danielle M Allen, Pearce Allingham, Ashley Levickas, Arthur Fitzgerald, Stephen H Bell, Jonathan Lock, Jonathon D Coey, Cormac McSparron, Behnam F Nejad, Evan P Troendle, David A Simpson, David G Courtney, Gisli G Einarsson, James P McKenna, Derek J Fairley, Tanya Curran, Jennifer M McKinley, Deirdre F Gilpin, Ken Lemon, John W McGrath, Connor G G Bamford

Summary

Background Influenza A viruses (IAVs) are significant pathogens of humans and other animals. Although endemic in humans and birds, novel IAV strains can emerge, jump species, and cause epidemics, like the latest variant of H5N1. Wastewater-based epidemiology (WBE) has been shown capable of detecting human IAVs. We aimed to assess whether whole-genome sequencing (WGS) of IAVs from wastewater is possible and can be used to discriminate between circulating strains of human and any non-human IAVs, such as those of avian origin.

Methods Using a pan-IAV RT-quantitative PCR assay, six wastewater treatment works (WWTWs) across Northern Ireland were screened from Aug 1 to Dec 5, 2022. A nanopore WGS approach was used to sequence RT-qPCR-positive samples. Phylogenetic analysis of sequences relative to currently circulating human and non-human IAVs was performed. For comparative purposes, clinical data (PCR test results) were supplied by The Regional Virus Laboratory, Belfast Health and Social Care Trust (Belfast, Northern Ireland, UK).

Findings We detected a dynamic IAV signal in wastewater from Sept 5, 2022, onwards across Northern Ireland, which did not show a clear positive relationship with the clinical data obtained for the region. Meta (mixed strain) whole-genome sequences were generated from wastewater samples displaying homology to only human and avian IAV strains. The relative proportion of IAV reads of human versus avian origin differed across time and sample site. A diversity in subtypes and lineages was detected (eg, H1N1, H3N2, and several avian). Avian segment 8 related to those found in recent H5N1 clade 2.3.4.4b was identified.

Interpretation WBE affords a means to monitor circulating human and avian IAV strains and provide crucial genetic information. As such, WBE can provide rapid, cost-effective, year-round One Health surveillance to help control IAV epidemic and pandemic-related threats. However, optimisation of WBE protocols are necessary to ensure observed wastewater signals not only correlate with clinical case data, but yield information on the wider environmental pan-influenz-ome.

Funding Department of Health for Northern Ireland.

Introduction

Influenza A virus (IAV; family *Orthomyxoviridae*; genus *Influenzavirus*) is an enveloped virus with a genome comprising eight segments (named 1–8) of single-stranded negative sense RNA.[1] The IAV genome encodes at least ten proteins, including the haemagglutinin (HA) and neuraminidase (NA) glycoproteins, and the innate immune antagonist non-structural protein 1 (NS1). IAV is a major pathogen of humans, livestock, and wildlife, and has been the causative agent of several pandemics over the last century following zoonosis from animal reservoirs such as pigs and birds. Clinical burden in humans and animals is extensive[1,2] but can be managed through prophylactic vaccination targeting circulating HA and NA sequences.[3]

IAV has many subtypes and although some are endemic in humans (eg, H1N1 and H3N2) and other mammals, wild birds belonging to the orders Anseriformes (eg, ducks) and Charadriiformes (eg, gulls[4]) are widely acknowledged as the major IAV genetic reservoirs. Most IAVs are considered low-pathogenicity avian influenza viruses (LPAIVs; eg, H13Nx and H16Nx, which infect gulls).[5,6] Others (eg, H5 and H7) can be high-pathogenicity avian influenza viruses (HPAIVs) due to the acquisition of a multibasic cleavage sequence in HA.[7] A novel strain of HPAIV H5Nx (clade 2.3.4.4b) was first detected in 2016 in western Siberia, Russia.[8] This variant spread globally, killing more than 50 million birds and infecting wild and farmed mammal species. Additional evidence suggesting mammal-to-mammal transmission (eg, mink) increases the pandemic potential of such circulating HPAIVs, emphasising the need for heightened, active surveillance in both humans and animals.[9]

Current human IAV monitoring usually entails clinical surveillance via primary care (sentinel general practitioner

Lancet Microbe 2024

Published Online https://doi.org/10.1016/S2666-5247(24)00175-7

School of Biological Sciences, Queen's University Belfast, Belfast, Northern Ireland, UK (A J Lee PhD, S Carson PhD, M I Reyne PhD, A Marshall PhD, D Moody PhD, D M Allen PhD, P Allingham MSc, A Levickas BSc, S H Bell PhD, J Lock MSc, J D Coey PhD, Prof J W McGrath PhD, C G G Bamford PhD); **Geography, School of Natural and Built Environment, Queen's University Belfast, Belfast, Northern Ireland, UK** (C McSparron PhD, B F Nejad PhD, Prof J M McKinley PhD); **Wellcome-Wolfson Institute for Experimental Medicine, School of Medicine, Dentistry and Biomedical Sciences, Queen's University Belfast, Belfast, Northern Ireland, UK** (A Fitzgerald MSc, E P Troendle PhD, Prof D A Simpson PhD, D G Courtney PhD); **School of Pharmacy, Queen's University Belfast, Belfast, Northern Ireland, UK** (G G Einarsson PhD, D F Gilpin PhD); **Regional Virology Laboratory, Belfast Health and Social Care Trust, Royal Victoria Hospital, Belfast, Northern Ireland, UK** (J P McKenna PhD, D J Fairley PhD, T Curran PhD); **Veterinary Sciences Division, Agri-Food and Biosciences Institute, Stormont, Belfast, Northern Ireland, UK** (K Lemon PhD); **Institute for Global Food Security, Queen's University Belfast, Belfast, Northern Ireland, UK** (Prof J W McGrath, C G G Bamford)

Correspondence to: Dr Andrew J Lee, School of Biological Sciences, Queen's University Belfast, Belfast BT9 5DL, Northern Ireland, UK **a.j.lee@qub.ac.uk**

Research in context

Evidence before this study

During the COVID-19 pandemic, wastewater-based epidemiology (WBE) gained popularity as a public health surveillance system. However, although WBE has been used for the surveillance of influenza A virus (IAV), so far it has had a human-centric focus, with genomic information derived from targeted sequencing of select portions of the IAV genome, namely the polymerase, haemagglutinin, and matrix segments. We searched PubMed, bioRxiv, and medRxiv for studies published in English between database inception and Jan 1, 2024, using the terms "wastewater-based epidemiology, WBE, wastewater surveillance, influenza A, IAV, sequencing, RT-PCR". We included studies if they used WBE to monitor IAV in human communities or the environment. The search yielded several WBE studies focusing on IAV, which showed that molecular detection and partial sequencing of this virus was possible from wastewater. However, noticeable limitations regarding sequencing approaches for influenza were evident, and no definitive evidence was presented regarding the detection of non-human IAV from wastewater. Our search identified no studies that confirmed whether whole-genome sequencing (WGS) of IAV from wastewater is possible and whether it can be used to discriminate between circulating strains of human IAV and IAV of avian origin.

Added value of this study

In this study, a One Health, integrated, genomic WBE approach was adopted. Using a pan-IAV RT-quantitative PCR (RT-qPCR) assay for screening, followed by long-read WGS to investigate positive samples, IAV concentrations were monitored before and during the start of a typical human influenza season, and WGS used to discriminate any circulating strains of IAV found in the wastewater. Previous studies have applied targeted RT-qPCR to monitor IAV concentrations with sequencing focused towards the antigenically important haemagglutinin and neuraminidase glycoproteins. To our knowledge, this is the first study to use a combined host-agnostic screening approach and WGS for IAV monitoring in wastewater. Using these methods, we are the first to show avian influenza in wastewater alongside human influenza, and the first to perform WGS of IAV from wastewater.

Implications of all the available evidence

Genomic surveillance of IAV using WBE approaches can complement established epidemiological methods, providing rapid and reliable additional information on the incidence and spread of IAV, a pathogen known for its pandemic potential. In the context of One Health routine pathogen monitoring, tracking emerging threats, or pandemic preparedness, integration of WBE alongside routine clinical or agricultural surveillance programmes might augment established veterinary and clinical testing structures.

services and in-hospital testing) with onward analysis of positive samples by PCR and sequencing. Avian IAV surveillance incorporates active (capture and swabbing of live birds, both wild and farmed) and passive (dead birds) monitoring (appendix p 10). Environmental sampling (eg, faeces or mud) has been explored but has not played a significant role in surveillance efforts globally,[10] despite research indicating that contaminated water was an important transmission mechanism.[11] In 2011, Heijnen and Medema,[12] who first illustrated the detection of IAV in sewage and surface water samples, postulated that avian reservoirs were a probable source of the IAV signal measured in the freshwater samples. However, they noted that confirmation was not possible as the available sequencing information was limited to small (95 bp), cloned quantitative RT-PCR (RT-qPCR) fragments.

See Online for appendix

Complementing established clinical surveillance methods, wastewater-based epidemiology (WBE) has recently proven an effective approach to track the prevalence, and identify variants, of SARS-CoV-2 circulating in the human population[13] and monitor enteric pathogens like adenoviruses[14] and other viruses, such as respiratory syncytial virus.[15,16] WBE studies have also shown the detection of human IAVs and influenza B virus in wastewater samples, using targeted PCR but with limited sequencing.[17–19] Since IAV is a recognised major pathogen of humans and other animals, responsible for seasonal epidemics and devastating pandemics, strengthening global surveillance systems for IAV (and other zoonotic diseases) should be a priority. Co-detection of human and non-human IAV in wastewater would show that WBE could act as an early warning surveillance system for not only human infections, but emerging, zoonotic IAVs of pandemic potential.

Here, we describe an integrated pan-IAV (ie, host-agnostic) genomic wastewater surveillance approach that uses RT-qPCR and nanopore whole-genome sequencing (WGS) to rapidly detect and sequence IAV-positive wastewater samples.

Methods

Study design

For this genomic surveillance study, we developed a One Health method to survey pan-IAV levels in wastewater via probe-based RT-qPCR and genetically characterise IAV-positive samples through WGS. This method was applied to clarified wastewater supernatant, from primary influent samples. These samples were collected from six wastewater treatment works (WWTWs) covering key population centres across each health trust of Northern Ireland and capturing 28·7% of the total population: Ballymena (42 969 residents), Belfast (228 939 residents), Craigavon (78 899 residents), Derry–Londonderry (94 655 residents), Enniskillen (15 115 residents), and North Down (73 384 residents; appendix pp 7–9, 11–12).

These select samples were provided for routine IAV surveillance by Northern Ireland Water and the Northern Ireland Environment Agency. The Regional Virus Laboratory, Belfast Health and Social Care Trust (Belfast, Northern Ireland, UK), which performs routine clinical testing for all of Northern Ireland (appendix p 6), supplied clinical data pre, peri, and post study period, in the form of total individual IAV RT-qPCR tests performed and IAV positive tests per epidemiological week. This study did not involve human participants or patient identifiable samples and so was exempt from ethical oversight.

Procedures

In-depth information on the method used is provided in the appendix (pp 1–7). Composite wastewater samples, comprising primary untreated influent were collected from the six WWTWs over 24 h, once per week, from Aug 1 to Dec 5, 2022, using an Isco Glacier autosampler (Isco, Lincoln, NE, USA). Primary influent wastewater samples (50 mL) were clarified by centrifugation at 4000 rpm (3434 rcf) for 10 min (4°C) before concentration using a CP-Select Concentrating Pipette with hollow fibre polysulfone high-flow pipette ultrafilter tips with a cutoff of 150 kDa (InnovaPrep, Drexel, MO, USA).

Total nucleic acids were extracted and purified from 200 µL of concentrated wastewater sample, on a Roche MagNA Pure 96 Instrument, using the DNA and Viral NA Small Volume Kit (Roche Diagnostics, Burgess Hill, West Sussex, UK) and eluted in a final volume of 50 µL. Wastewater was screened for the presence of any IAV using the host-agnostic SVIP-MPv2 RT-qPCR assay,[20] which targets a highly conserved region of the matrix protein segment of the IAV genome. IAV MP segment concentrations were normalised using wastewater flow rate and expressed as gene copies per 100 000 population equivalents per day.

A multisegment RT-PCR approach to simultaneously amplify all eight segments of the IAV genome from IAV MP-positive wastewater samples within a single PCR reaction was used.[21] Only IAV-positive samples with a RT-qPCR cycle threshold (Ct) value of 38 or less were selected for amplicon generation. Multisegment RT-PCR amplicons were prepared using the SQK-LSK109 kit and sequenced on FLO-MIN106D R9 flow cells using a GridION (Oxford Nanopore Technologies, Oxford, UK) with reads base called using Guppy (version 6.4.2). Utilising Centrifuge, a taxonomic classification first approach was used, before using Minimap2 to map IAV classified reads against key reference genomes (H1N1, H3N2, H5N1, and H13N6). Coverage calculations and the generation of consensus sequences used Samtools and BCFtools, respectively, with raw reads interrogated using BLAST+ (version 2.13.0; appendix p 5). Generated IAV consensus sequences were used for phylogenetic analysis, with focus on sequences where the consensus generated by each reference was similar to each other. Consensus sequence incorporation differed depending on the analysis used, with only homogeneous consensus sequences used for interrogation of databases and phylogenetics of avian segments 5, 7, and 8, whereas all consensus sequences available were used for human segment 4 and 6 analysis, using Nextclade to identify sequences of low quality, which were subsequently removed from analysis (appendix pp 5–6).

Statistical analysis

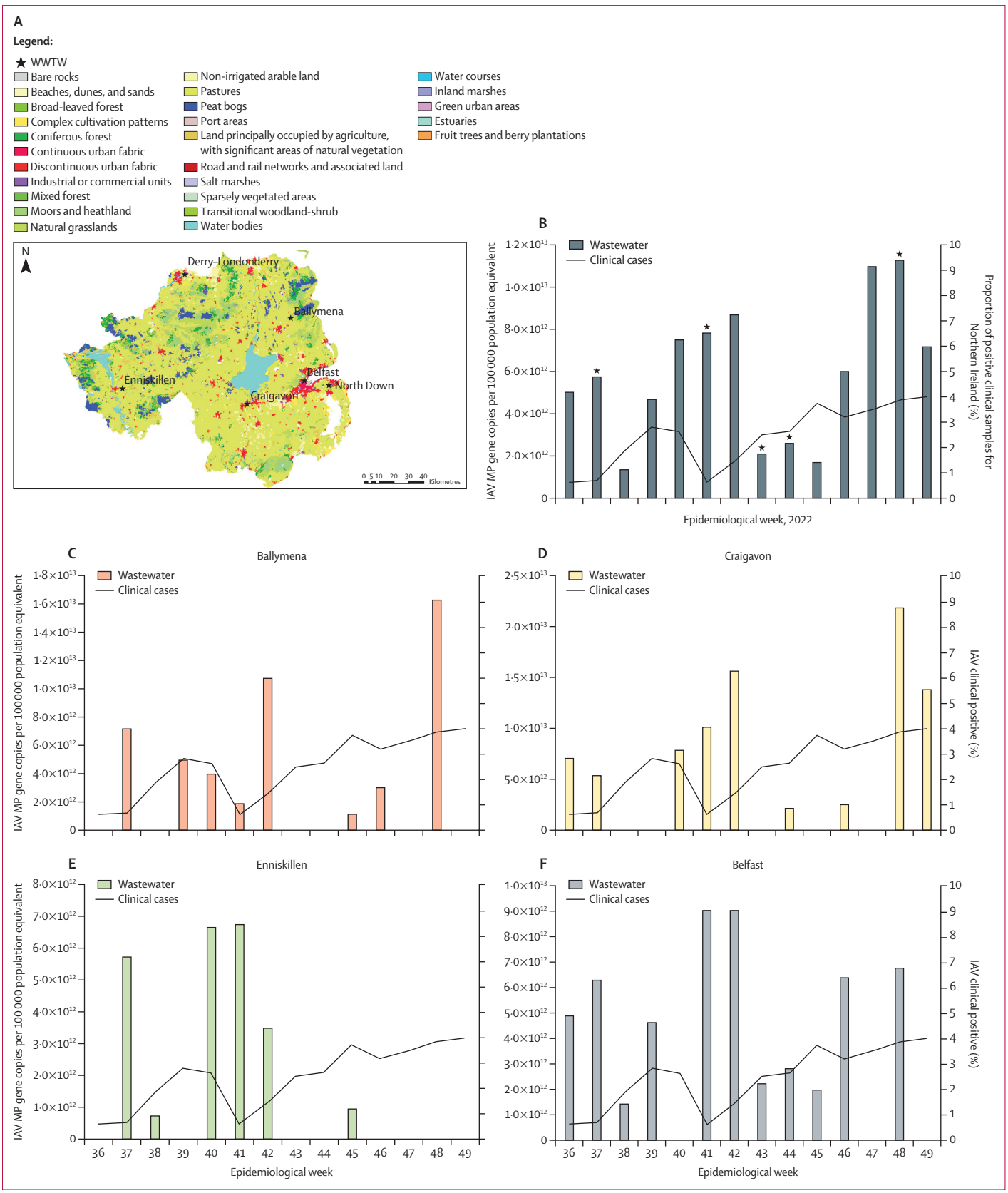
No sample size computation was performed because this investigation was carried out on samples delivered once per week as part of routine surveillance for IAV. The number of positive IAV clinical cases per week for Northern Ireland between Aug 1 and Dec 31, 2022, was provided by the Regional Virology Laboratory, Belfast Health and Social Care Trust, and percentage positivity per epidemiological week calculated from the total number of weekly tests for IAV carried out. For phylogenetic analysis, bootstrapping (n=1000 rapid bootstraps) was used to infer statistical support for phylogenies, using a support threshold of over 70% for a given node with the best substitution model selected using the Bayesian Information Criterion. Quantification of IAV RNA concentrations in wastewater for each sample collected was calculated using the following formulae. Gene copies per reaction: $G=10^{(\Delta Ct/m)}$, where G is IAV concentration in nucleic acid extract (gene copies per µL), ΔCt is the difference between the Ct value (sample nucleic acid) and the standard curve intercept (equivalent to one gene copy per µL), and m is slope of the standard curve. Gene copies per nucleic acid extract per sample: $G_S=V_e \times G_M/S_P$, where G_S is gene copy per sample, V_e is nucleic acid elution volume (µL), G_M is arithmetic mean of the replicate gene copy per µL (G) results, and S_P is proportion of concentrated virus included in nucleic acid extraction. Gene copies per L of wastewater: $G_L=1000 \times (G_S/V_S)$, where G_L is gene copies per L, G_S is gene copies per sample, and V_S is volume of wastewater processed (mL).

Role of the funding source

The funder of the study had no role in the study design, data collection, data analysis, data interpretation, or writing of the report.

Results

We screened 144 wastewater samples collected from six WWTWs across Northern Ireland (figure 1) between Aug 1 and Dec 5, 2022 (epidemiological weeks 31–49) for IAV. An IAV signal above the limit of detection was first identified from Sept 5, 2022 (epidemiological week 36) onwards, with a mean weekly concentration of $5{\cdot}03 \times 10^{12}$ gene copies per 100 000 population equivalents across the WWTWs screened (SD $1{\cdot}5 \times 10^{12}$; figure 1). IAV positivity (48 [H3·3%] of 144 samples) and its concentration in wastewater varied across the sites sampled during the period investigated, with three peaks observed during weeks 37 (mean gene copies per 100 000 population equivalents $5{\cdot}74 \times 10^{12}$ [SD $1{\cdot}31 \times 10^{12}$]), 42 ($8{\cdot}67 \times 10^{12}$ [$5{\cdot}08 \times 10^{12}$]), and 48 ($1{\cdot}13 \times 10^{13}$ [$2{\cdot}95 \times 10^{12}$]). Clinical percentage positivity remained below 1% until week 35 (eight [1·4%] of 577 samples) and did not sustain an upwards trajectory until



(Figure 1 continues on next page)

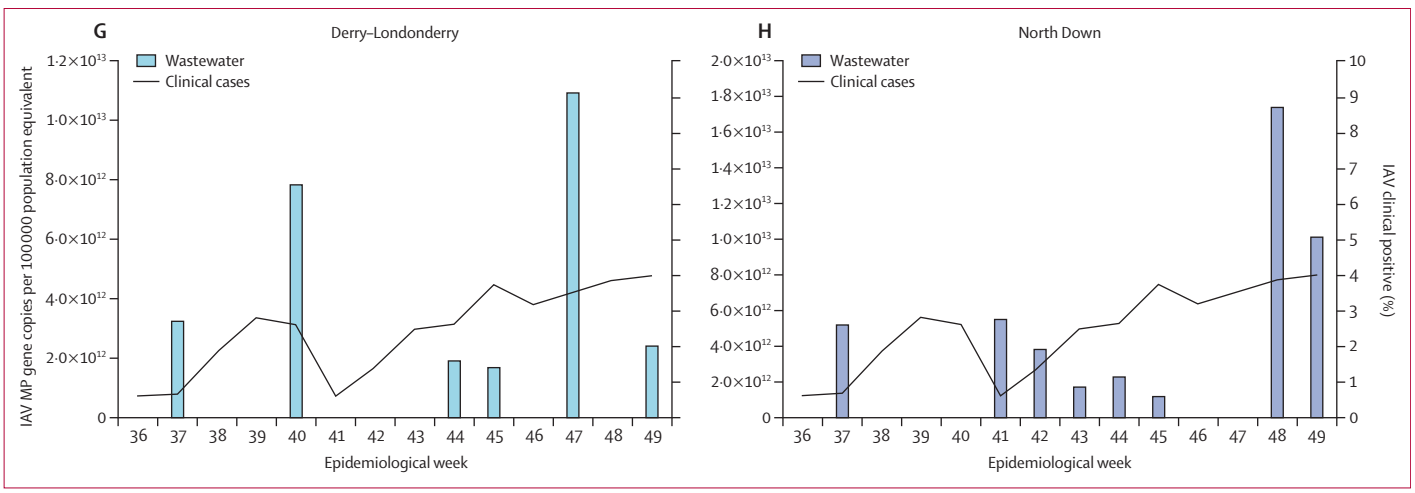


Figure 1: Dynamic IAV wastewater signal across Northern Ireland

(A) Corine Land Cover map showing geographical locations of WWTWs across Northern Ireland screened for IAV. Normalised and mean concentrations of IAV MP gene detected (gene copies per 100 000 population equivalent) per epidemiological week across the six WWTWs sampled (B) and for each WWTW: Ballymena (C), Craigavon (D), Enniskillen (E), Belfast (F), Derry–Londonderry (G), and North Down (H). Clinical data (black line) represents percentage IAV-positive PCR and point-of-care tests done by the Regional Virus Laboratory for Northern Ireland. Stars in (B) indicate epidemiological weeks in which IAV sequencing was successful. IAV RT-qPCR-positive results for each WWTW per epidemiological week is also included. IAV=influenza A virus. MP=matrix protein. WWTWs=wastewater treatment works.

week 42 (37 [1·4%] of 2577 samples), steadily increasing to 4·0% (107 of 2679 samples) during week 49 (189 [6·4%] of 2943 samples; figure 1). From week 50 onwards the clinical percentage positivity rose markedly, peaking during week 52 (335 [14·42%] of 2323 samples [data not shown]).

Identified IAV reads in these RT-qPCR-positive samples ranged from 65 to 1440, with 7044 sequences detected overall. Reads were generally biased towards the smaller segments (ie, segments 7 and 8; table 1). No single sample produced reads for every segment. However, across the samples successfully sequenced, reads mapping to all eight segments were generated and are referred to here as meta-WGS to reflect sequencing of all segments and to distinguish from true WGS from clinical samples containing single strains. On average, percentage coverage of segments varied considerably from 25·7% (PB1) up to 99·6% (NP). Sequences mapping to segments 4 and 6, encoding HA and NA, were generated for nine (75%) of 12 samples. Mean read number and coverage were 32·3% (SD 61·1) and 29·5% (33·2) for HA segments and 10·2% (12·1) and 57·9% (44·1) for NA segments, respectively.

Initial read interrogation using a taxonomic classification first approach suggested the presence of diverse IAV strains, possibly consisting of both human-like and avian-like (figure 2). The relative abundance of these strains differed by sampling time and site. At first, heterogeneous mixes predominated, but the relative abundance of human-like IAV reads increased in samples from Belfast and Ballymena as the clinical influenza season commenced (figure 2). However, samples from North Down showed a more stable mix of human-like and avian-like sequences. Furthermore, we generated consensus sequences by mapping, focusing our downstream analysis on sequences where the consensus generated by each reference was similar to each other. This approach resulted in 25 consensus sequences spanning all IAV segments (table 2; appendix pp 23–59). Derived consensus sequences were compared with known IAV sequences by BLAST, revealing diverse origins corroborating the original analysis by our taxonomic classification first approach, including human subtypes (H1N1 and H3N2) but also avian subtypes, which are predominantly circulating in gull populations (H13N6), and even H5N1 clade 2.3.4.4b (Belfast; Oct 10, 2022) from a Eurasian curlew (a member of the Charadriiformes; table 2). As indicated by our initial taxonomic classification, we detect here that the proportion of avian-like sequences, compared with human-like sequences, decreases at later sampling dates.

To explore the avian IAV sequences further, we determined the phylogenetic relatedness of our derived consensus sequences relative to the two gull-associated subtypes (H13 and H16 Nx) as well as recent (2022) H5N1 clade 2.3.4.4b viruses found across Europe. Phylogenetic analysis of segment 8 (encoding the virulence factor NS1, and the essential nuclear export protein [NEP]) confirmed clear avian association but identified two distinct clades in our samples (figure 3). One clade (gull 2; sequences found in Belfast, Craigavon, and North Down on Sept 12, 2022) was closely related to, but distinct from, H16Nx viruses identified in Europe in 2013–15 (the Netherlands). The other clade (gull 3; detected in Belfast; Oct 10, 2023) was more closely related to recent European H5N1 viruses (clade 2.3.4.4b), which include viruses sequenced from birds and mammals (mink). However, phylogenetically, our H5N1-related sequence sat at a basal position to the H5N1 subclade along with other gull-associated sequences (found in Europe and China in 2020–21), suggesting that it is unlikely to be from H5N1 viruses but from the ancestor of the lineage that donated its segment 8 to H5N1 clade 2.3.4.4b. No other

	Segment 1 PB2 (2·3 kb)		Segment 2 PB1 (2·3 kb)		Segment 3 PA (2·2 kb)		Segment 4 HA (1·8 kb)		Segment 5 NP (1·6 kb)		Segment 6 NA (1·4 kb)		Segment 7 MP (1·0 kb)		Segment 8 NS (0·9 kb)		Total reads per sample (N)
	Reads	Coverage	Reads	Coverage	Reads	Coverage	Reads	Coverage	Reads	Coverage	Reads	Coverage	Reads	Coverage	Reads	Coverage	
Sept 12, 2022																	
Ballymena	3	50·5%	0	0	0	0	0	0	25	32·3%	0	0	67	90·7%	74	84·9%	169
Belfast	2	58·0%	0	0	0	0	6	76·5%	26	87·0%	1	96·8%	216	91·7%	326	89·9%	577
Craigavon	2	91·6%	35	25·7%	0	0	112	33·3%	13	88·2%	4	76·4%	277	92·7%	167	91·1%	595
Enniskillen	0	0	0	0	0	0	0	0	18	78·2%	31	95·7%	347	96·1%	228	93·8%	624
North Down	0	0	0	0	0	0	3	66·0%	14	82·6%	3	96·0%	120	89·2%	508	90·1%	648
Oct 10, 2022																	
Belfast	0	0	0	0	0	0	0	0	9	48·4%	27	58·3%	47	77·7%	48	76·8%	131
North Down	0	0	0	0	0	0	0	0	5	86·7%	0	0	60	88·7%	0	0	65
Oct 24, 2022																	
Belfast	0	0	0	0	0	0	0	0	0	0	12	87·1%	507	96·7%	676	92·1%	1195
Oct 31, 2022																	
Belfast	0	0	0	0	0	0	9	60·4%	1	99·6%	26	89·4%	98	93·9%	1306	90·4%	1440
North Down	0	0	0	0	0	0	0	0	0	0	0	0	96	93·4%	0	0	96
Nov 28, 2022																	
Ballymena	3	39·8%	0	0	2	68·9%	67	77·1%	85	88·3%	18	95·1%	279	89·4%	364	91·3%	818
Belfast	0	0	0	0	0	0	191	40·2%	12	87·6%	0	0	186	95·4%	282	91·2%	671
Per segment total reads and average coverage	10	60·0%	35	25·7%	2	68·9%	388	58·9%	208	77·9%	122	87·0%	2300	91·3%	3976	89·2%	..

Read number and average percentage coverage generated per segment, per sample. Identity and alignment length of raw reads investigated using Basic Local Alignment Search Tool against National Center for Biotechnology Information's available precompiled nucleotide database. IAV genome comprises eight RNA segments: PB1, PB2, PA, HA, NP, NA, MP, and NS. HA=haemagglutinin. IAV=influenza A virus. MP=matrix protein. NA=neuraminidase. NP=nucleoprotein. NS=non-structural protein. PA=polymerase acidic. PB1=polymerase basic 1. PB2=polymerase basic 2. WGS=whole-genome sequencing. WWTWs=wastewater treatment works.

Table 1: **Description of IAV WGS in wastewater by WWTWs**

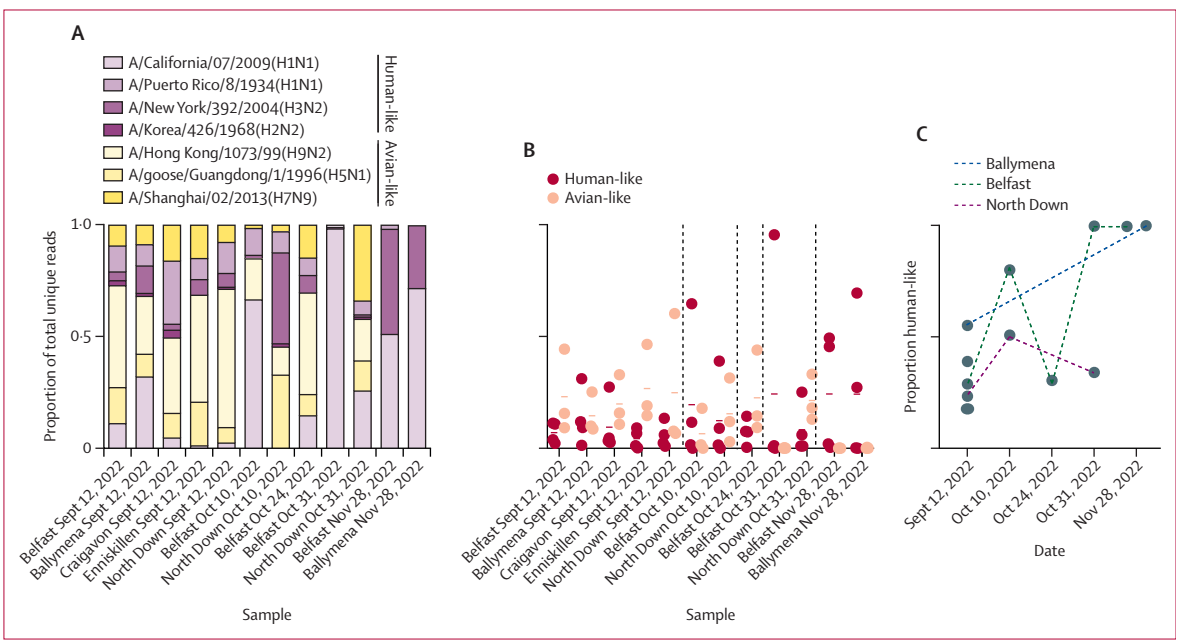


Figure 2: **Taxonomic classification-first interrogation of IAV reads from wastewater**

(A) Proportion of unique reads from samples were assigned a similarity to one of seven IAV strains (A/California/07/2009[H1N1], A/Puerto Rico/8/1934[H1N1], A/New York/392/2004[H3N2], A/Korea/426/1968[H2N2], A/Hong Kong/1073/99[H9N2], A/goose/Guangdong/1/1996[H5N1], or A/Shanghai/02/2013[H7N9]) in the Centrifuge database. These strains were grouped into human-like (H1N1, H3N2, and H2N2; purple shades) and avian-like (H9N2, H5N1, and H7N9; yellow shades). (B) Data from (A) shown aggregated by human-like and avian-like for each sample, with median and range illustrated. (C) Fractional proportion of aggregate human-like sequences shown sample date only for all sites with links for those for which multidate data are available. Graphs generated in Graphpad Prism, version 9.0. IAV=influenza A virus.

	Segment	Mapping	Subtype	Host species	GISAID ID	Name
Sept 12, 2022						
Ballymena	1	H1N1	H16N3	Gull	EPI2181372	A/gull/Shandong/W1359/2021 (A/H16N3) segment 1 (PB2)
Belfast	1	H1N1	H13N8	Black-headed gull	EPI1942887	A/Chroicocephalus_ridibundus/Belgium/13464/2020 (A/H13N8) segment 1 (PB2)
Belfast	5	H13N6	H13N2	Armenian gull	EPI617828	A/Armenian gull/Republic of Georgia/2/2012 (A/H13N2) segment 5 (NP)
Belfast	6	H13N6	H13N6	Common gull	EPI2195559	A/common_gull/Poland/MW241/2011 (A/H13N6) segment 6 (NA)
Belfast	8	H13N6	H13N6	Black-headed gull	EPI889316	A/black-headed gull/Netherlands/88/2012 (A/H13N6) segment 8 (NS)
Craigavon	1	H1N1	H16N3	Gull	EPI2181372	A/gull/Shandong/W1359/2021 (A/H16N3) segment 1 (PB2)
Craigavon	2	H13N6	H13N8	Common guillemot	EPI1884896	A/Common guillemot/Sweden/SVA210729SZ0317/FB002610/I-2021 (A/H13N8) segment 2 (PB1)
Craigavon	6	H1N1	H1N1	Human	EPI2214233	A/England/223980556/2022 (A/H1N1 pdm09) segment 6 (NA)
Craigavon	8	H13N6	H13N6	Black-headed gull	EPI889316	A/black-headed gull/Netherlands/88/2012 (A/H13N6) segment 8 (NS)
Enniskillen	5	H13N6	H13N8	Common guillemot	EPI1884899	A/Common guillemot/Sweden/SVA210729SZ0317/FB002610/I-2021 (A/H13N8) segment 5 (NP)
Enniskillen	6	H13N6	H13N6	Common gull	EPI2195559	A/common_gull/Poland/MW241/2011 (A/H13N6) segment 6 (NA)
North Down	4	H13N6	H13N6	Black-headed gull	EPI890024	A/black-headed gull/Netherlands/5/2014 (A/H13N6) segment 4 (HA)
North Down	5	H13N6	H13N2	Black-headed gull	EPI1014396	A/black-headed gull/Netherlands/9/2014 (A/H13N2) segment 5 (NP)
North Down	8	H13N6	H13N6	Black-headed gull	EPI889316	A/black-headed gull/Netherlands/88/2012 (A/H13N6) segment 8 (NS)
Oct 10, 2022						
Belfast	5	H1N1	H1N1	Human	EPI2569903	A/France/ARA-HCL023049801501/2023 (A/H1N1 pdm09) segment 5 (NP)
Belfast	6	H1N1	H1N1	Human	EPI2550192	A/Poland/PL86/2022 (A/H1N1 pdm09) segment 6 (NA)
Belfast	7	H1N1	H1N1	Human	EPI2295864	A/Catalonia/NSVH171694150/2022 (A/H1N1 pdm09) segment 7 (MP)
Belfast	8	H13N6	H13N8 and H5N1	Black-headed gull and Eurasian curlew	EPI1942894 and EPI2540344	A/Chroicocephalus_ridibundus/Belgium/13464/2020 (A/H13N8) segment 8 (NS) and A/Numenius_arquata/Belgium/01456_0003/2023 (A/H5N1) segment 8 (NS)
North Down	7	H13N6	H13N8	Black-headed gull	EPI1942893	A/Chroicocephalus_ridibundus/Belgium/13464/2020 (A/H13N8) segment 7 (MP)
Oct 24, 2022						
Belfast	6	H1N1	H1N1	Human	EPI2550184	A/Ireland/67945/2022 (A/H1N1 pdm09) segment 6 (NA)
Belfast	6	H1N1	H1N1	Human	EPI2414712	A/Czech Republic/5176/2022 (A/H1N1 pdm09) segment 6 (NA)
Belfast	7	H1N1	H1N1	Human	EPI2571961	A/Jamaica/7605/2023 (A/H1N1 pdm09) segment 7 (MP)
Nov 28, 2022						
Ballymena	1	H3N2	H3N2	Human	EPI2213118	A/Denmark/3021/2022 (A/H3N2) segment 1 (PB2)
Ballymena	3	H3N2	H3N2	Human	EPI2193946	A/Barry/3792/2022 (A/H3N2) segment 3 (PA)
Belfast	5	H1N1	H1N1	Human	EPI2571981	A/Costa Rica/6859/2023 (A/H1N1 pdm09) segment 5 (NP)

Table describing BLAST (GISAID) results for consensus sequences from samples sites and dates. Sequences spanned all segments and interrogation of GISAID database allowed low resolution subtyping by identifying known subtype of the top-ranked similarity hit. For each top hit, the subtype, host species, GISAID ID, and name for each sequence is shown. For Belfast, Oct 10, 2022, two hits including similarity to H5N1 are shown. HA=haemagglutinin. IAV=influenza A virus. MP=matrix proteins. NA=neuraminidase. NP=nucleoprotein. NS=non-structural proteins. PA=polymerase acidic. PB1=polymerase basic 1. PB2=polymerase basic 2.

Table 2: **Similarity of wastewater IAV consensus sequences to known sequences**

H5-like sequences were detected. A similar relationship to H5 viruses was observed with a single sequence of segment 5 (appendix p 20). Interrogation of the single segment 7 sequence identified similarity to sequences from Europe (2020–21), China (2021), and Alaska (2020; appendix p 21).

To assess whether our derived sequences might also provide information on the evolution of the human epidemic, derived human segments 4 and 6 consensus sequences were analysed further. Following removal of low-quality sequences, this analysis included 11 sequences from seven samples. Phylogenetic analyses, focused on our segment 4 and 6 sequences, compared with those from recently circulating IAVs (globally and 2022 Northern Ireland) showed diversity in WBE sequences, from at least one H1N1 clade (6B.1A.5a.2a.1 [5a.2a.1]) and three H3N2 clades (all Bangladesh-like 3C.2a1b.2a.2a.1 [2a.1], 3C.2a1b.2a.2a.3a.1 [2a.3a.1], and 3C.2a1b.2a.2b [2b]), as suggested by our initial taxonomic classification (figure 4). Our sequences were similar to strains circulating in 2022–23 in Northern Ireland but phylogenetically distinct from HA sequences incorporated into the vaccines used in that season (ie, 5a.2 for H1N1 and 3C.2a1b.2a.2 for H3N2).

Discussion

Here we show that through WGS, WBE—a surveillance approach traditionally focused on monitoring human infections—can also be repurposed to distinguish between circulating strains of IAV and characterise animal pathogens, such as avian influenza.

We detected a clear and dynamic IAV signal in our wastewater samples, but comparing our wastewater viral load data to the clinical data for all of Northern Ireland did not show a clear positive relationship. Although the clinical positivity rate increased over time, our IAV signal showed several clear peaks and troughs. WGS suggested that earlier in the season more avian IAVs dominated before a later shift to human IAVs. Considering our genomics findings, a lack of correlation could be expected given the overwhelming non-human avian IAV signal detected, potentially reflecting dynamics of avian IAVs in wild avian reservoirs. Comparing

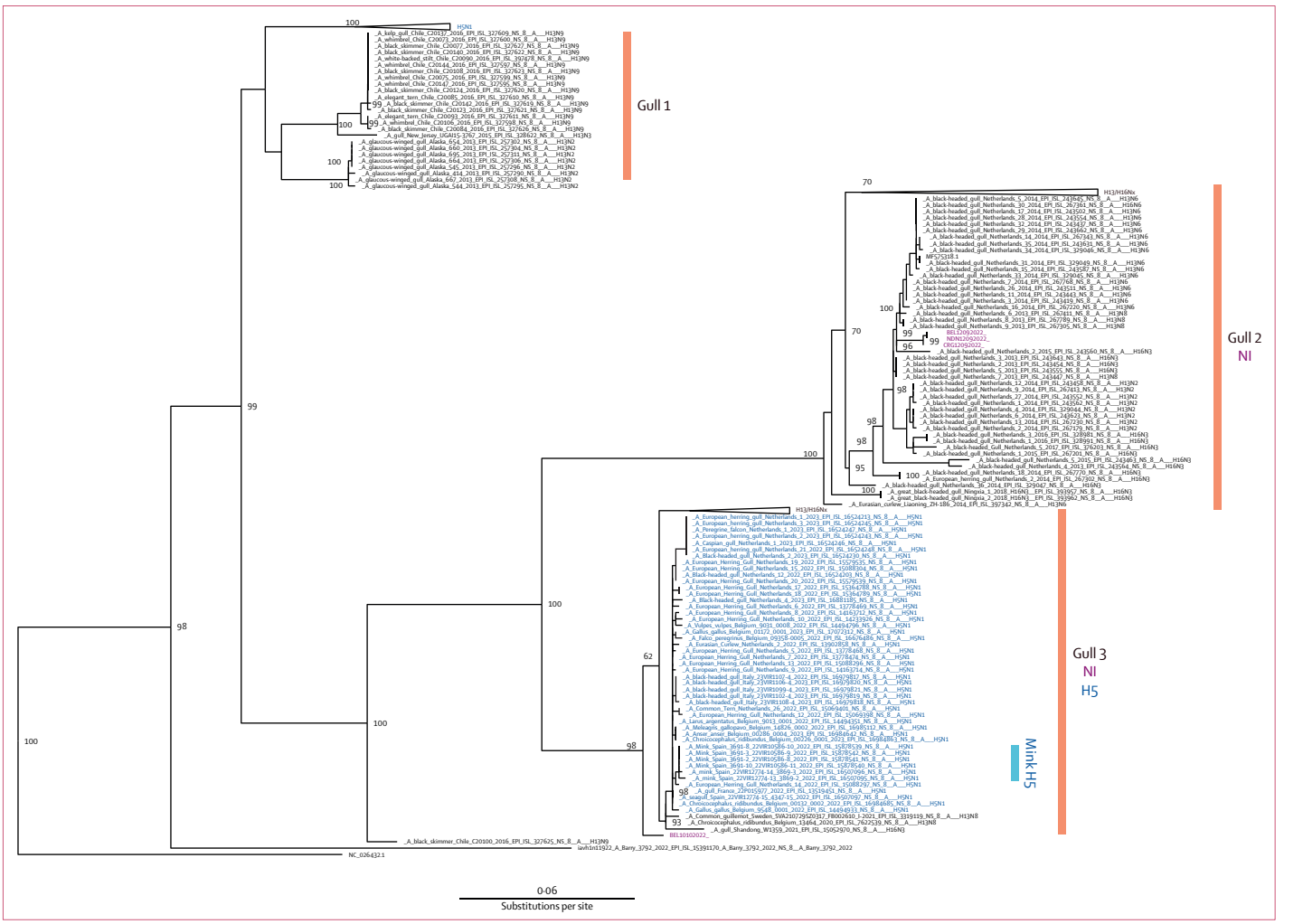


Figure 3: **Maximum-likelihood phylogeny of wastewater-derived avian-like segment 8 sequences**

Phylogenetic tree of the four avian-like segment 8 sequences (pink) alongside all GISAID gull H13-like and H16-like sequences (black), and 2022 H5N1 European sequences (blue; accessed Jan 9, 2023). Three clades are highlighted (gull 1–3) with presence of N1 (pink) and H5 (blue) annotated. Mink H5N1 sequences are highlighted in cyan. For illustrative purposes, several clades are shown collapsed in cartoon form. Relevant bootstrap replicate values are shown on the tree.

peaks in early September and early October, distinct avian IAV strains were identified.

Our data revealed diverse avian IAV lineages, notably recent H13-like and H16-like viruses, which are LPAIV subtypes infecting gulls and closely related species belonging to the order Charadriiformes, poultry, and mammals.[5,6,22] Remarkably, we identified one sequence (segment 8 BEL10102022) with close affinity to contemporaneous HPAI H5N1 viruses found across Europe, proven capable of spreading between mammals.[9] Although recently in Northern Ireland there have been confirmed cases of gulls infected with H5N1,[23] it is unlikely that we detected the HPAIV H5N1 clade 2.3.4.4b as H5N1 homology was identified only for one segment (segment 8; NS1 and NEP) and phylogenetic analysis suggested that the sequences we detected fell outside of the recent H5N1 group. Consistent with this interpretation, we only found non-H5 avian H13 segments, with no H5 identified. Notably, recent European H5N1 viruses have reassorted with local gull-associated viruses, acquiring segments 3, 5, and 8.[9] Gulls have been shown to be important in spreading HPAIV H5.[24]

Previous WBE studies of IAV have adopted a human-centric screening and sequencing approach, possibly explaining why convincing evidence of avian IAV has not been reported thus far. It is also probable that timing and geography are factors impacting avian IAV detection. The stable mixture of human and avian IAV seen throughout the study period at the Belfast and North Down WWTWs might reflect the coastal location of these sites: both are close to important seabird colonies in Belfast Lough, the Copeland Islands, and Strangford Lough. The other sampled WWTW sites (excepting Derry–Londonderry, from which no sequences were recovered) are located much further inland.

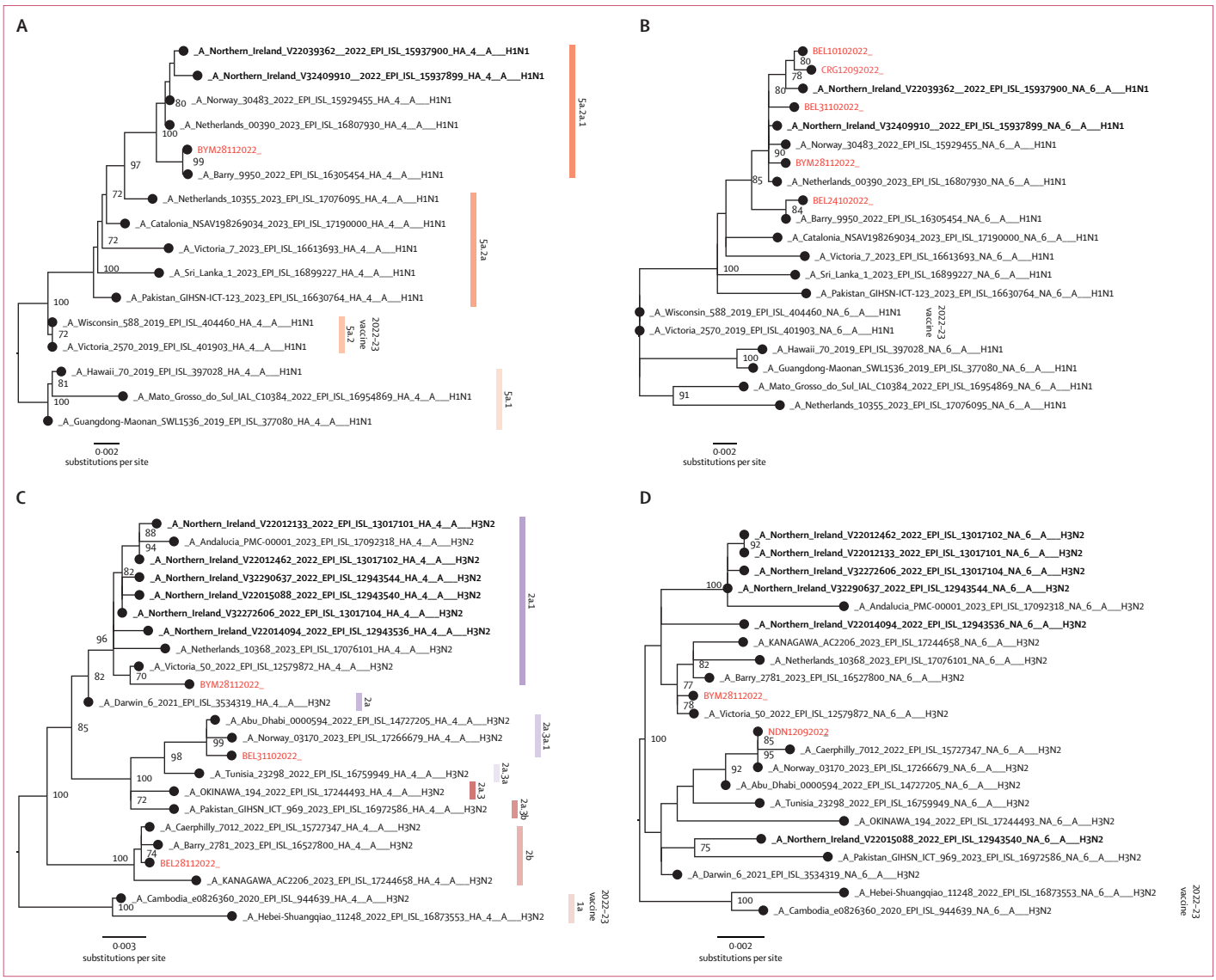


Figure 4: **Maximum-likelihood phylogeny of wastewater-derived human-like segments 4 (HA) and 6 (NA) sequences from H1N1 and H3N2**
Phylogenetic trees of the 11 sequences (red) of human-like segment H1N1 segments 4 (A) and 6 (B) and H3N2 segments 4 (C) and 6 (D) alongside select recent worldwide sequences and all available relevant Northern Ireland clinical sequences (bold black). H1N1 Northern Ireland clinical sequences were sampled during the 2022–23 season (September to December, 2022), whereas the H3N2 Northern Ireland clinical sequences were sampled in spring, 2022, during the end of the 2021–22 season (January, 2021 to April, 2022). Established HA sequence clades are shown alongside sequence groupings. Northern Hemisphere 2022–23 winter season vaccine strain sequences highlighted. Only bootstrap replicate values greater than 69 are shown on the tree.

Concerning the origin of our avian IAV sequences, WWTWs, urban environments, and neighbouring agricultural areas are recognised as important habitats for birdlife year-round. Consequently, direct deposition or runoff of avian faecal matter from fields or metropolitan areas entering sewage systems could be a source.[25–27] Nevertheless, other mechanisms should be considered, such as the role of mammal vectors[28] or that true human infection exists. However, we consider these scenarios less probable as no human infections with H13 or H16 subtypes were noted.

Our protocol generated sequence information from the early stages of the human 2022–23 IAV epidemic in Northern Ireland, in particular covering segments 4 (HA) and 6 (NA), which facilitate subtyping and antigenic comparison. Given the dearth of locally available clinical IAV sequences, an accurate comparison between wastewater and clinical data is challenging, although wastewater-derived sequences generally matched contemporary globally circulating viruses. Having access to provincial clinical IAV sequence data for correlation with our own would have aided spatial resolution. Comparing our sequences with those of the chosen vaccine strains for 2022–23 showed phylogenetic distinctness, which might impact antigenicity, although vaccine match was generally considered good for that year.[29] This finding suggests that sensitive and effective real-time WBE IAV

sequencing could be used as a tool to inform in-season estimates of vaccine effectiveness and clinical burden. However, this approach would require further in-depth molecular and serological analysis of derived HA and NA sequences, and development of more specific RT-qPCR assays.

Our study was limited by the number of IAV-positive wastewater samples detected and by the subsequent sequencing success rate for those positive samples—circumstances probably influenced by uncertainties around both the stability of viral RNA biomarkers in wastewater, and variations in viral titres present in the waste stream. Despite these limitations we show that meta-WGS of IAV from wastewater is possible, notwithstanding the unique challenges such complex sample matrices pose. Co-purified PCR-inhibiting substances and the fragmented state of shed virus might explain the inconsistency between our RT-qPCR results and the number of successfully sequenced samples. Although the pan-IAV assay requires only a small fragment of the segment to be present, the WGS approach requires intact viral RNA termini, and their degradation due to exonuclease activity will substantially impact amplicon generation. Alternatively, it is possible that post-PCR carryover of enduring contaminants damaged or blocked flow cell pores. Using newer, inhibitor-tolerant, cDNA synthesis systems, stringent clean-up steps, and updated sequencing reagent chemistries and technologies could notably improve sequencing outcomes. An explicit focus on capturing more virus, reducing inhibitors, and enhancing nucleic acid extraction and purification steps should also be prioritised.

Nanopore sequencing has already proven useful for rapid end-to-end, in-field characterisation of avian influenza.[30] Our workflow, tailored appropriately, could offer a cheaper alternative for genomic IAV wastewater surveillance, useful for low-income and middle-income countries, where access to fully equipped laboratories is not readily available, but environmental surveillance of contaminated surface waters or sewage is of concern.

In conclusion, we show that human-centric WBE can be refocused as a useful tool for combined avian and human IAV genomic surveillance. WBE might augment established veterinary and clinical testing structures, offering advance warning on the incidence and spread of IAV and providing a cost-effective means to develop a global, year-round One-Health sentinel service for IAV.

Contributors

AJL and CGGB contributed to the conceptualisation of the study, data curation, formal analysis, investigation, method, visualisation, and writing of the original draft manuscript. SC and MIR contributed to data curation, investigation, and visualisation. AM, DM, DMA, PA, AL, AF, SHB, JL, and JDC contributed to the investigation. CM, BFN, EPT, DAS, GGE, and JMM contributed to formal analysis, data curation, and visualisation. DGC, JPM, DJF, and TC contributed to data curation and resource provision. KL contributed to the conceptualisation of the study, formal analysis, and resource provision. DFG and JWM contributed to the conceptualisation of the study, funding acquisition, resources, and supervision. All coauthors had full access to all the data in the study, contributed to the revision and editing of subsequent drafts of the manuscript and approved the final version for submission. AJL, CGGB, JPM, and KL accessed and verified the underlying data of the study.

Declaration of interests

DFG and JWM report grants from the Northern Ireland Department of Health and Public Health Agency. All other authors declare no competing interests.

Data sharing

All relevant data are included in the manuscript and appendix. The generated consensus sequence data were submitted to GenBank (accession identifiers PP664069—PP664104) and are also available in the appendix (pp 23–59). All sequences not generated by us and used for the phylogenetic analysis shown were obtained from GISAID.

Acknowledgments

This work was supported by funding from the Department of Health for Northern Ireland through the Northern Ireland Wastewater Surveillance Programme. We are grateful to the staff members of Northern Ireland Water and the Northern Ireland Environment Agency for the provision of wastewater samples. We also thank Phil Atkinson from the British Trust for Ornithology for helpful discussions.

Editorial note: The Lancet Group takes a neutral position with respect to territorial claims in published maps and institutional affiliations.

References

1 Krammer F, Smith GJD, Fouchier RAM, et al. Influenza. *Nat Rev Dis Primers* 2018; **4:** 3.
2 Uyeki TM, Hui DS, Zambon M, Wentworth DE, Monto AS. Influenza. *Lancet* 2022; **400:** 693–706.
3 Wei CJ, Crank MC, Shiver J, Graham BS, Mascola JR, Nabel GJ. Next-generation influenza vaccines: opportunities and challenges. *Nat Rev Drug Discov* 2020; **19:** 239–52.
4 Yoon SW, Webby RJ, Webster RG. Evolution and ecology of influenza A viruses. *Curr Top Microbiol Immunol* 2014; **385:** 359–75.
5 Fouchier RA, Munster V, Wallensten A, et al. Characterization of a novel influenza A virus hemagglutinin subtype (H16) obtained from black-headed gulls. *J Virol* 2005; **79:** 2814–22.
6 Verhagen JH, Poen M, Stallknecht DE, et al. Phylogeography and antigenic diversity of low-pathogenic avian influenza H13 and H16 viruses. *J Virol* 2020; **94:** e00537–20.
7 Munster VJ, Schrauwen EJ, de Wit E, et al. Insertion of a multibasic cleavage motif into the hemagglutinin of a low-pathogenic avian influenza H6N1 virus induces a highly pathogenic phenotype. *J Virol* 2010; **84:** 7953–60.
8 Lee DH, Sharshov K, Swayne DE, et al. Novel Reassortant clade 2.3.4.4 avian influenza A(H5N8) virus in wild aquatic birds, Russia, 2016. *Emerg Infect Dis* 2017; **23:** 359–60.
9 Agüero M, Monne I, Sánchez A, et al. Highly pathogenic avian influenza A(H5N1) virus infection in farmed minks, Spain, October 2022. *Euro Surveill* 2023; **28:** 2300001.
10 Hood G, Roche X, Brioudes A, et al. A literature review of the use of environmental sampling in the surveillance of avian influenza viruses. *Transbound Emerg Dis* 2021; **68:** 110–26.
11 Pepin KM, Leach CB, Barrett NL, et al. Environmental transmission of influenza A virus in mallards. *MBio* 2023; **14:** e0086223.
12 Heijnen L, Medema G. Surveillance of influenza A and the pandemic influenza A (H1N1) 2009 in sewage and surface water in the Netherlands. *J Water Health* 2011; **9:** 434–42.
13 Li L, Uppal T, Hartley PD, et al. Detecting SARS-CoV-2 variants in wastewater and their correlation with circulating variants in the communities. *Sci Rep* 2022; **12:** 16141.
14 Reyne MI, Allen DM, Levickas A, et al. Detection of human adenovirus F41 in wastewater and its relationship to clinical cases of acute hepatitis of unknown aetiology. *Sci Total Environ* 2023; **857:** 159579.
15 Allen DM, Reyne MI, Allingham P, et al. Genomic analysis and surveillance of respiratory syncytial virus (RSV) using wastewater-based epidemiology (WBE). *J Infect Dis* 2024; **2024:** jiae205.

16 Boehm AB, Hughes B, Duong D, et al. Wastewater concentrations of human influenza, metapneumovirus, parainfluenza, respiratory syncytial virus, rhinovirus, and seasonal coronavirus nucleic-acids during the COVID-19 pandemic: a surveillance study. *Lancet Microbe* 2023; **4:** e340–48.

17 Mercier E, D'Aoust PM, Thakali O, et al. Municipal and neighbourhood level wastewater surveillance and subtyping of an influenza virus outbreak. *Sci Rep* 2022; **12:** 15777.

18 Markt R, Stillebacher F, Nägele F, et al. Expanding the pathogen panel in wastewater epidemiology to influenza and norovirus. *Viruses* 2023; **15:** 263.

19 Vo V, Harrington A, Chang CL, et al. Identification and genome sequencing of an influenza H3N2 variant in wastewater from elementary schools during a surge of influenza A cases in Las Vegas, Nevada. *Sci Total Environ* 2023; **872:** 162058.

20 Nagy A, Černíková L, Kunteová K, et al. A universal RT-qPCR assay for "One Health" detection of influenza A viruses. *PLoS One* 2021; **16:** e0244669.

21 Zhou B, Donnelly ME, Scholes DT, et al. Single-reaction genomic amplification accelerates sequencing and vaccine production for classical and Swine origin human influenza a viruses. *J Virol* 2009; **83:** 10309–13.

22 Yu Z, Cheng K, Gao Y. Poultry infection with influenza viruses of wild bird origin, China, 2016. *Emerg Infect Dis* 2018; **24:** 1375–77.

23 UK Department of Agriculture, Environment and Rural Affairs. Bird flu alert. April 26, 2024. https://www.daera-ni.gov.uk/articles/latest-situation (accessed Sept 6, 2024).

24 Hill NJ, Bishop MA, Trovão NS, et al. Ecological divergence of wild birds drives avian influenza spillover and global spread. *PLoS Pathog* 2022; **18:** e1010062.

25 Gough S, Gillings S, Vickery JA. The value and management of wastewater treatment works for breeding and wintering birds in lowland eastern England. BTO Research Report, number 333 2003.

26 Schwemmer P, Garthe S, Mundry R. Area utilization of gulls in a coastal farmland landscape: habitat mosaic supports niche segregation of opportunistic species. *Landsc Ecol* 2008; **23:** 355–67.

27 Spelt A, Williamson C, Shamoun-Baranes J, Shepard E, Rock P, Windsor S. Habitat use of urban-nesting lesser black-backed gulls during the breeding season. *Sci Rep* 2019; **9:** 10527.

28 Velkers FC, Blokhuis SJ, Veldhuis Kroeze EJB, Burt SA. The role of rodents in avian influenza outbreaks in poultry farms: a review. *Vet Q* 2017; **37:** 182–94.

29 McLean HQ, Petrie JG, Hanson KE, et al. interim estimates of 2022–23 seasonal influenza vaccine effectiveness, Wisconsin, October 2022–February 2023. *MMWR Morb Mortal Wkly Rep* 2023; **72:** 201–05.

30 de Vries EM, Cogan NOI, Gubala AJ, et al. Rapid, in-field deployable, avian influenza virus haemagglutinin characterisation tool using MinION technology. *Sci Rep* 2022; **12:** 11886.