BMC Infectious Diseases

RESEARCH

Open Access

An epidemiologic surveillance study based on wastewater and respiratory specimens reveals influenza a virus prevalence and mutations in Taiyuan, China during 2023–2024

Lifeng Zhao[1†], Jihong Xu[1†], Jiane Guo[1], Ping Zhang[1], Xiaofang Guo[1], Zhihong Zuo[1], Li Gao[1], Zhao Jia[1,2], Puna Xue[1,2] and Jitao Wang[1,2*]

Abstract

Background Influenza is a major cause of morbidity and mortality. Influenza A virus (IAV) is one of the most important pathogens causing influenza and often causes global pandemics due to its tendency to mutate. We aim to use epidemiology based on wastewater and respiratory specimens to understand the occurrence of influenza A virus infections in Taiyuan City.

Methods A retrospective epidemiology surveillance was carried out at the First Hospital of Shanxi Medical University (FHSMU) and five wastewater treatment plants (WTPs) in Taiyuan city from 2023 to 2024. Reverse transcription real-time fluorescence quantitative polymerase chain reaction (RT-qPCR) was used to detect influenza A viruses in wastewater and respiratory specimens. High-throughput whole genome sequencing was performed on 17 strains obtained in this study, and subsequent analyses included characterization, phylogenetic construction, amino acid mutation analysis, and antigenic structural variability assessment.

Results 520 wastewater samples and 1,203 throat swab samples were collected. We detected RNA concentration from pH1N1 and H3N2 viruses in wastewater and got 17 genome sequences (5 of pH1N1 and 12 of H3N2) in respiratory specimens. Whole-genome sequencing showed co-prevalence of pH1N1 viruses in the branches of 6B.1 A.5a.2a.1 and H3N2 viruses in the branches of 3 C.2a1b.2a.2a.3a.a in Taiyuan from 2023 to 2024. Moreover, a HA mutation (N138D), predicted to be of high phenotypic consequence, was found in 8 Taiyuan H3N2 sequences.

Conclusion This study highlights the predominant presence of pH1N1 and H3N2 strains in Taiyuan. The analysis also identified amino acid site variations in the HA antigenic epitopes in H3N2 strains, which may contribute to immune escape.

[†]Lifeng Zhao and Jihong Xu contributed equally to this work.

*Correspondence:
Jitao Wang
wangjitao99@126.com

Full list of author information is available at the end of the article

Keywords Influenza a virus, Wastewater, Whole-genome sequencing, Mutation, Genome characterization

Background

Infections with influenza viruses cause acute respiratory infections with serious health burdens and economic losses [1]. There are approximately a billion cases of seasonal influenza related to influenza virus infection annually, including 3–5 million cases of severe illness [2]. All influenza viruses are enveloped negative-sense single-strand RNA viruses that can be further divided into four types-type A, type B, type C, and type D-based on the antigenicity of the nucleoprotein and matrix protein [2]. Influenza A viruses (IAVs) are the only known influenza viruses that can cause a pandemic [1, 3].

Influenza A virus contains eight segments, including hemagglutinin (HA), neuraminidase (NA), matrix protein (MP), nucleoprotein (NP), nonstructural (NS), polymerase acidic (PA), polymerase basic 1 (PB1), and polymerase basic 2 (PB2) [2]. HA and NA are the most antigenically variable, they are classified into antigenically diverse subtypes. To date, more than 16 antigenically different HA and 9 antigenically different NA serotypes or subtypes have been identified among the different avian strains [2]. In humans, current subtypes that routinely circulate include pH1N1 viruses (since 2009) and H3N2 viruses (since 1968) [2]. Both lineages are known to have caused pandemics over the past decades [1, 4].

IAVs can cause fever, cough, sore throat, pneumonia, and even death; all age groups can be affected, but there are groups that are more at risk than others, such as children < 5, pregnant women, elderly individuals, and immunocompromised patients [4–6]. Vaccination is an effective method for preventing influenza, but IAVs are prone to mutations, undergo antigenic drift, and produce immune escape [1, 2]. When a new or different type appears, a pandemic occurs, which infects people and has the ability to effectively spread from person to person, while people have almost no immunity to this virus. In addition, NA and PA inhibitors are currently specific drugs for treating influenza; however, with continuous use, viruses may develop drug-resistant mutations, which reduces the therapeutic effectiveness of these drugs [2].

Influenza surveillance in China is passive and draws from many sources. Usually, it relies on regular nucleic acid testing of clinical specimens by designated hospitals and clustered epidemic monitoring reports to identify diseases [7]. The use of a passive influenza monitoring system is voluntary and tends to identify infections in individuals with comorbidities or severe symptoms who require clinical care. Since influenza is self-limiting, most patients with mild symptoms do not seek treatment at the hospital. Therefore, there are obvious limitations in using clinical samples to monitor virus transmission in the community. At the same time, epidemiological information based on wastewater can provide a deeper understanding of the incidence of influenza in contributing communities [8]. Wastewater is a composite biological sample containing bodily excreta (including urine, feces, sputum, and mucus). The excreta of infected individuals may contain biomarkers of infectious diseases (including live and dead organisms, proteins, and nucleic acids), and methods have been developed to detect and quantify these biomarkers in wastewater matrices [8].

In contrast to seasonal diarrhea, influenza is highly prevalent during the winter and spring months in northern China [7, 9, 10], and monitoring typically occurs from October to April of the following year. Taiyuan is the provincial capital city of Shanxi Province and is also one of the densely populated cities in northern China; thus, the threat of influenza outbreaks in the region should not be underestimated. In this study, we tracked the presence of IAV RNA in wastewater after the end of the COVID-19 epidemic to assess the persistence of viral RNA in field wastewater discharged from hospital and municipal sewage systems. In addition, we studied all available sequence data for pH1N1 viruses and H3N2 viruses collected since July 2023 and June 2024, across all 8 nucleic acid fragments and 11 available amino acid sequences. Our goal was to gain insight into the epidemic characteristics of IAVs to improve outbreak prevention strategies.

Materials and methods

Specimen collection

Wastewater specimens were obtained from five wastewater treatment plants (WTPs) in Taiyuan city, namely Beijiao, Chengnan, Fendong, Jinyang, and Yangjiabao (Table 1 and Figure S1). These plants are responsible for the collection and treatment of domestic wastewater for more than 2 million people (Table 1). For each WTP, an FC-24 Automatic Water Reclamation Facility (Grasp, Beijing, China) was used to collect 24-hour mixed wastewater samples, collecting 100 ml per hour, totaling 2400 ml. Specimen collection cycles were twice a week at least 2 days apart. After collection, the specimens were stored at a low temperature (4 °C) and sent to the Influenza Surveillance Network Laboratory of the Taiyuan Center for Disease Control and Prevention (Taiyuan CDC) for IAV testing within 2 h. The containers used for sample collection were sterilized and disposable.

Respiratory specimens (throat swabs) were collected at the First Hospital of Shanxi Medical University (FHSMU), which is one of the national influenza surveillance sentinel hospitals in Shanxi Provinces, receiving at least 100,000 patients annually. A minimum of 20 swab

Table 1 Five wastewater treatment plants in this study

Serial number	Wastewater plants	Location (district, latitude/longitude)	Relevant population (thousand persons)	Wastewater treatment capacity (kiloton/per day)*
14,010,801	Beijiao (BJ)	Jiancaoping (N: 37°56′ E: 112°31′)	346.7	402. 46
14,010,502	Chengnan (CN)	Xiaodian (N: 37°45′ E: 112°34′)	172.4	200
14,010,503	Fendong (FD)	Xiaodian N: 37°37′ E: 112°28′	285.3	331
14,011,001	Jinyang (JY)	Jinyuan (N: 37°38′E: 112°24′)	275.7	320
14,010,501	Yangjiabao (YJB)	Xiaodian (N: 37°48′E: 112°33′)	1 347.4	1,562. 98

Note: * Average

specimens were collected from outpatients with influenza-like illness (ILI) (body temperature≥38 °C, with either cough or sore throat) each week between July 1, 2023 and June 30, 2024. Specimens were immediately stored at -20 °C and transferred to the Influenza Surveillance Network Laboratory of the Taiyuan CDC within 24 h for influenza virus nucleic acid testing.

Influenza a virus quantification in wastewater

Based on Lee's method, the quantification of influenza A virus was performed by constructing a standard curve using a dilution series of a qualified reference material on a PCR platform and evaluating linearity against corresponding threshold cycle (Ct) values [11, 12]. Specifically, 5-fold serial dilutions of reference material of inactivated influenza A viruses (pH1N1 and H3N2) (China National Sharing Platform for Reference Materials, Beijing, China) were used to generate standard curves (Supplementary Tables S1 and Figure S2). The series contained 10^4 to 10^1 copies per reaction, and each level ran at least three times. The HA gene viral load per reaction was calculated by interpolation to the respective standard curve (Supplementary Figure S2). In this process, total RNA from two reference materials (pH1N1 and H3N2) was extracted using a MGIEasy RNA Extraction Kit (MGI, Shenzhen, China) in an MGISP-960 Fully Automatic Nucleic Acid Extraction Workstation (MGI, Shenzhen, China), according to the manufacturer's instructions. HA gene detection and quantification were conducted using the Influenza A Virus pH1N1 (2009) and H3N2 Nucleic Acid Detection Kit (BioPerfectus, Taizhou, China). The PCR cycling parameters were set up according to the instructions: 50 °C for 30 min and 95 °C for 5 min, followed by 45 cycles of 94 °C for 10 s and 55 °C for 40 s in a CFX96 Real-time Thermal Cycler (Bio-Rad, Hercules, CA). The Ct values were determined using the default algorithm in CFX Manager Software 3.1 (Bio-Rad, Hercules, CA, USA). Positive internal controls (reagent kit) and non-template controls (H_2O) were included in each run.; Each sample undergoes at least 3 repeated experiments to ensure accuracy. Standard curves are constructed in Microsoft Excel (Version: Microsoft Office Professional 2016), and only the standard curve with an R^2-value greater than 0.99 is valid; otherwise, it is invalid (Supplementary Figure S2).

Polyethylene glycol precipitation was used for total RNA extraction from wastewater specimens according to Torii's method [13]. Specifically, 80 ml of wastewater were first centrifuged (2,500 g at 4 ℃ for 1 min) to remove solid impurities, and then 8 g of polyethylene glycol 8000 (PEG 8000) and 1.8 g of sodium chloride (NaCl) were added, followed by vortexing and centrifugation (12,000 g at 4 °C for 2 h). After removing the supernatant, the precipitate was washed with 1 ml of TE buffer, and the resulting solution was used to concentrate viruses. The concentrates were stored at -80 °C until downstream analysis, unless analyzed on the same day; The experimental process of nucleic acid extraction and PCR amplification was identical to that of standard curves (pH1N1 and H3N2). The quality of the extracted RNA was checked with a Qubit 4 fluorometer (Invitrogen, Singapore). An in-house control (positive internal control) and non-template controls were included at all stages of sample processing (virus isolation, nucleic acid extraction, and amplification) to assess potential inhibition and cross-contamination. Following the instructions of the reagent kit, a positive result was defined as a Ct value<42, and the in-house control was defined as a Ct value within 2 times the standard deviation to prove the effectiveness of the experiment, otherwise, the result was considered invalid. Each sample was tested at least three times and treated without any pasteurization to inactivate viruses in the wastewater to reduce virus loss. In order to maintain consistency and minimize human error, most of the sample processing steps were performed by a liquid handling robot (MGISP-960).

Influenza a virus detection in respiratory specimens

Original throat swab samples were collected from outpatients with ILI and immediately placed in non-inactivated Hank's medium tubes. The RNA extraction and nucleic acid detection were carried out using the same reagent kit as those used in wastewater experiments, and the program settings were also the same as previously mentioned. PCR cycling parameters were set up according to the instructions in a CFX96 Real-time Thermal Cycler

(Bio-Rad, Hercules, CA). A positive result was defined as a threshold cycle (Ct) value < 42, and a positive internal control was defined as a Ct value < 30.

Influenza a virus whole-genome sequencing

Among the positive respiratory samples, those with Ct values less than 30 were used for whole genome sequencing. Viral RNA was extracted by the same procedure as previously mentioned, and the quality of the extracted RNA was checked through a Qubit 4 Fluorometer (Invitrogen, Singapore). IAVs' gene capture was prepared using an ULSEN Ultra-Sensitive Influenza Virus Whole-Genome Capture Kit (MicroFuture, Beijing, China). The capture PCR cycling parameters were set up according to the instructions: 42 °C for 60 min and 94 °C for 2 min, followed by 45 cycles of 94 °C for 30 s, 57 °C for 30 s and 68 °C for 3 min in a T100 Thermal Cycler (Bio-Rad, Hercules, CA). Paired-end libraries for the MiniSeq platform were prepared using a Nextera XT DNA Library Prep Kit (Illumina, San Diego, California, USA), the libraries cycling parameters were set up according to the instructions: 72 °C for 3 min and 95 °C for 30 s, followed by 12 cycles of 95 °C for 10 s, 55 °C for 30 s and 72 °C for 30 s in a T100 Thermal Cycler (Bio-Rad, Hercules, CA). Genome sequencing was performed using a 300-cycle (2×150-bp paired-end) MiniSeq V2 Reagent Kit via the MiniSeq platform (Illumina, San Diego, California, USA) according to the manufacturer's protocols.

Genome assembly and analysis

Genome segments were assembled using MicrofluAnalyzer High-throughput Sequencing Data Analysis software (Version: 1.0.4.9) (MicroFuture, Beijing, China). Low-quality reads in demultiplexed data were removed. The genome sequences of 2023–2024 vaccine representative strains (A/Wisconsin/67/2022 and A/Darwin/9/2021) and various subtype reference strains were used to construct evolutionary tree by using the neighbor-joining method (kimura two-parameter model) of the MEGA 11 software (Version: 11.0.13). Multiple nucleotide, and amino acid sequence alignments were constructed using the ClustalW accessory application in the BioEdit sequence alignment editor (Version 7.2.5.27). A/Wisconsin/67/2022 (pH1N1) and A/Darwin/9/2021 (H3N2) were identified via alignment with MEGA software to analyze the amino acid difference sites. The data presented in the study were deposited in the GISAID repository (https://gisaid.org/), the accession numbers are listed in Supplementary Table S2, and all reference sequences were downloaded from the GISAID database (accession number list in Table S3).

Protein structure modeling

HA gene sequences obtained in this study were analyzed using the FluSurver (v1) mutations tool (https://gisaid.org/) that predicts mutation effects. HA mutations predicted to be of high phenotypic consequence (levels 2 and 3) were mapped to the influenza HA protein crystal structures (PDB:7KNA of pH1N1, 4WE8 of H3N2). The HA antigenic sites, as defined in Liu's paper [14], were overlaid onto protein structures. Structure models were visualized and manipulated using SWISS-MODEL (https://swissmodel.expasy.org/).

Results

Influenza in wastewater

A total of 520 samples were collected from 5 WTPs during the period of July 1, 2023, to June 30, 2024. The pH1N1 virus was detected from week 11 to 22 of 2024, with a concentration ranging from 0.63 copies/ml to 8.17 copies/ml (Fig. 1A). The epidemic of pH1N1 was characterized by a sporadic distribution, not accompanied by a markedly clustered outbreak. The H3N2 virus was detected continuously from week 41, 2023, to week 16, 2024, with a virus concentration ranging from 0.35 copies/mL to 51.81 copies/mL (Fig. 1B). Unlike pH1N1, the epidemic of H3N2 started early, peaked at week 49 in 2023, and began to decline at week 50, while pH1N1 was not detected (Fig. 1B). During the same period, although pH1N1 and H3N2 viruses coexisted, the content of the pH1N1 virus in wastewater was significantly less than that of H3N2 (Fig. 1A and B).

Influenza in outpatients with influenza-like illness

During the study period, a total of 1,203 outpatients provided respiratory specimens for influenza virus testing at the FHSMU. In detail, 135 (11.2%) tested positive for IAVs, including 16 for pH1N1 and 119 for H3N2, with an overall positive rate of 1.3% and 9.9%, respectively (Fig. 1C and D). For pH1N1, positive cases were mainly detected in week 11 to week 22 of 2024 with a characteristic of dispersed distribution (Fig. 1C). The earliest occurrence of H3N2 infection occurred in week 11, peaked at week 49, and began to decline in the following weeks (Fig. 1D). The positive rates of influenza A viruses in different months were statistically significant ($\chi2=279.489$, $p<0.001$) (Table 2). The positive rates of IAVs in males and females were 10.6% and 11.8%, respectively ($\chi2=0.438$, $p=0.508$) (Table 2). The positive rates of IAVs in different age groups were 8.9% for 0–20 years; 12.3% for 21–40 years; 10.9% for 41–60 years, and 9.6% for those above 60 years, respectively ($\chi2=2.238$, $p=0.524$) (Table 2).

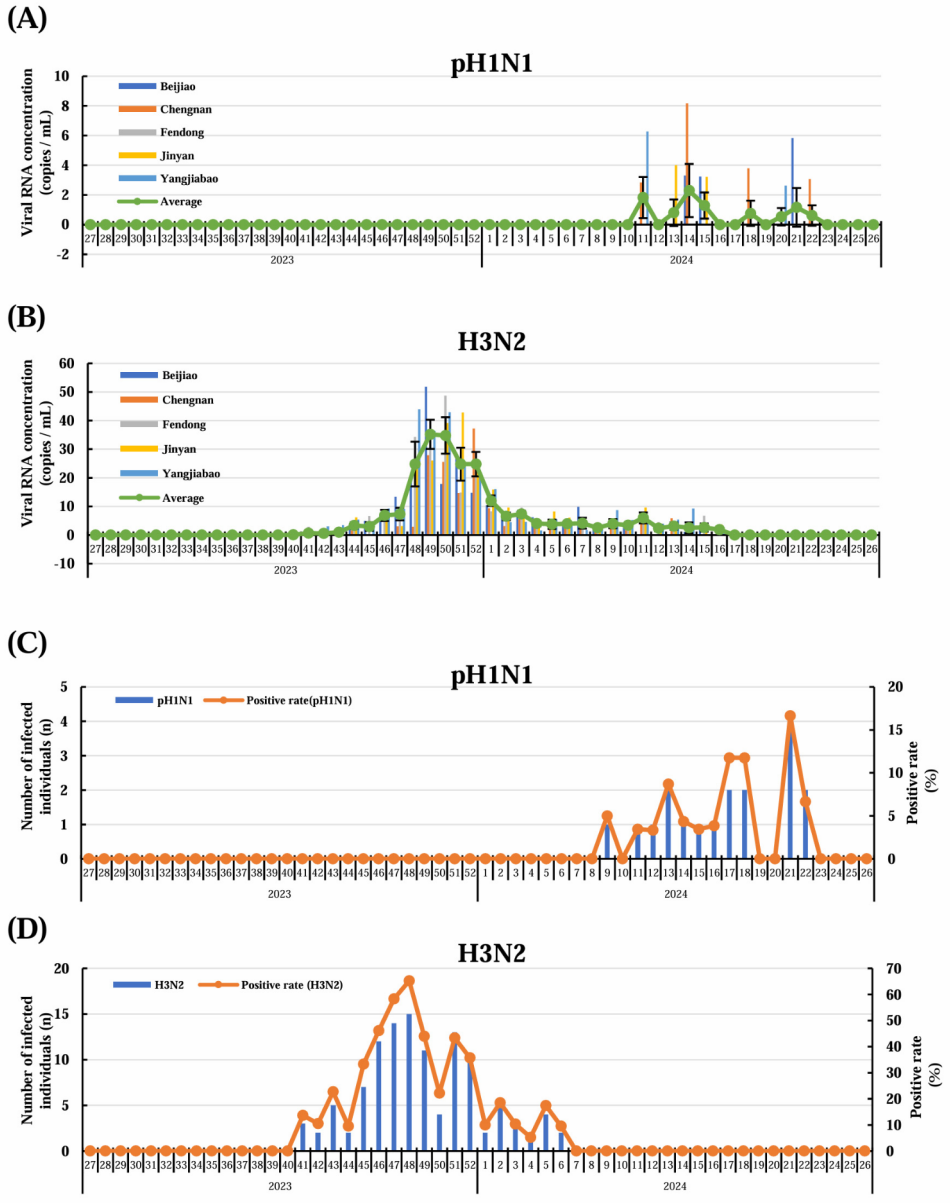

Fig. 1 Concentrations of IAV RNA in wastewater and comparison to designated hospital percent positivity data in Taiyuan city between 2023 and 2024. (**A**) pH1N1 RNA concentrations; (**B**) H3N2 RNA concentrations; (**C**) Laboratory-recognized infections and detection rate of pH1N1 viruses; (**D**) Laboratory-recognized infections and detection rate of H3N2 viruses. Average, weighted average of viral concentrations of five WTPs; Error bars: standard deviation

Genetic characterization of influenza viruses

In whole genome sequencing experiments, 17 strains were successfully sequenced (all 8 gene fragments), including 5 of pH1N1 and 12 of H3N2. In pH1N1, the nucleotide similarity of 8 gene fragments from 5 Taiyuan strains ranged from 99.2 to 100% (Table 3, pH1N1). All of the 5 pH1N1 strains' HA gene segments belonged to genetic clade 6B.1 A.5a.2a.1 (Fig. 2, HA). Compared with the WHO-recommended pH1N1 vaccine strain A/Wisconsin/67/2022 (2023–2024), the nucleotide similarity of the 8 gene segments varied from 97.9 to 100% (Table 3, pH1N1). Additionally, the evolutionary distances of HA, NA, MP, PA, NP, NS, PA, PB1, and PB2 genes in the 5 sequences were 0.015–0.017, 0.009–0.014, 0-0.003, 0.010–0.012, 0.006–0.009, 0.024–0.027, 0.011–0.013, and 0.017–0.018, respectively, compared with A/Wisconsin/67/2022 (Fig. 2).

In H3N2, the nucleotide similarity of 8 gene fragments from 12 Taiyuan strains ranged from 96.2 to 100% (Table 3, H3N2). All of the 12 Taiyuan H3N2 strains belonged to the genetic clade 3 C.2a1b.2a.2a.3a.a (Fig. 3, HA). Compared with the WHO-recommended H3N2 vaccine strain A/Darwin/9/2021 (Northern Hemisphere, 2023–2024), the nucleotide similarity of the 8 gene

Table 2 The positive rate of nucleic acid testing and composition characteristics of influenza in Taiyuan between 2023–2024

Variables	Specimens number (*n*)	pH1N1		H3N2		IAVs (pH1N1 + H3N2)	
		Positives (*n*)	Positive rate (%)	Positives(*n*)	Positive rate (%)	Positives(*n*)	Positive rate (%)
Month							
Jul-Sep 2023	267	0	0	0	0	0	0
Oct 2023	88	0	0	10	11.4	10	11.4
Nov 2023	101	0	0	47	46.5	47	46.5
Dec 2023	110	0	0	43	39.1	43	39.1
Jan 2024	109	0	0	16	14.7	16	14.7
Feb 2024	96	1	1.0	3	3.1	4	4.2
Mar 2024	110	4	3.6	0	0	4	3.6
Apr 2024	110	3	2.7	0	0	3	2.7
May 2024	113	8	7.1	0	0	8	7.1
Jun 2024	99	0	0	0	0	0	0.0
Gender							
Male	549	5	0.9	53	9.7	58	10.6
Female	654	11	1.7	66	10.1	77	11.8
Age							
0–20	179	1	0.6	15	8.4	16	8.9
21–40	682	10	1.5	74	10.9	84	12.3
40–60	165	4	2.4	14	8.5	18	10.9
61-	177	1	0.6	16	9.0	17	9.6

Table 3 Similarity analysis of the whole genome of influenza a virus in this study

Gene	Protein	Sequence similarity between Taiyuan strains (pH1N1 or H3N2) (%)		Similarity compared to vaccine strains* (%)	
		Nucleotides	Amino acid	Nucleotides	Amino acid
pH1N1					
HA	HA	99.6–100	99.6–100	97.9–98.2	97.9–98.1
NA	NA	99.2–100	99.3–100	98.5–99.0	98.5–98.7
MP	M1	99.7–100	100	99.4–99.7	100
	M2	100	100	99.6	100
NP	NP	99.7–100	100	97.4–97.5	98.9
NS	NS1	99.5–100	99.5–100	99.0-99.3	99.0-99.5
	NEP	99.7–100	100	98.9–100	98.3
PA	PA	99.7–100	99.5–100	98.7–100	98.8–99.1
	PA-X	99.5–100	99.5–100	98.2–98.7	98.2–98.6
PB1	PB1	99.6–100	99.8–100	98.7–98.9	99.6–99.7
PB2	PB2	99.7–100	100	98.2	99.6
H3N2					
HA	HA	98.7–100	98.4–100	97.9–98.2	97.3–98.0
NA	NA	99.0-100	98.9–100	98.9–99.2	98.9–99.3
MP	M1	98.4–100	100	98.5–99.8	99.6
	M2	96.2–100	94.8–100	96.9–99.3	94.8–98.9
NP	NP	99.1–100	99.7–100	97.7–98.1	98.9–99.1
NS	NS1	97.2–100	95.2–100	97.4–99.2	95.6–99.5
	NEP	97.1–100	94.8–100	97.2–99.5	95.2–99.6
PA	PA	99.1–100	99.4–100	97.9–98.1	98.4–98.8
	PA-X	99.2–100	99.2–100	97.7–98.0	96.8–98.0
PB1	PB1	99.0-99.9	99.7–100	98.8–99.1	99.4–99.6
PB2	PB2	98.9–100	99.6–100	97.5–98.0	99.2–99.4

Note: [*] Vaccine strain: A/Wisconsin/67/2022 (pH1N1) and A/Darwin/9/2021 (H3N2)

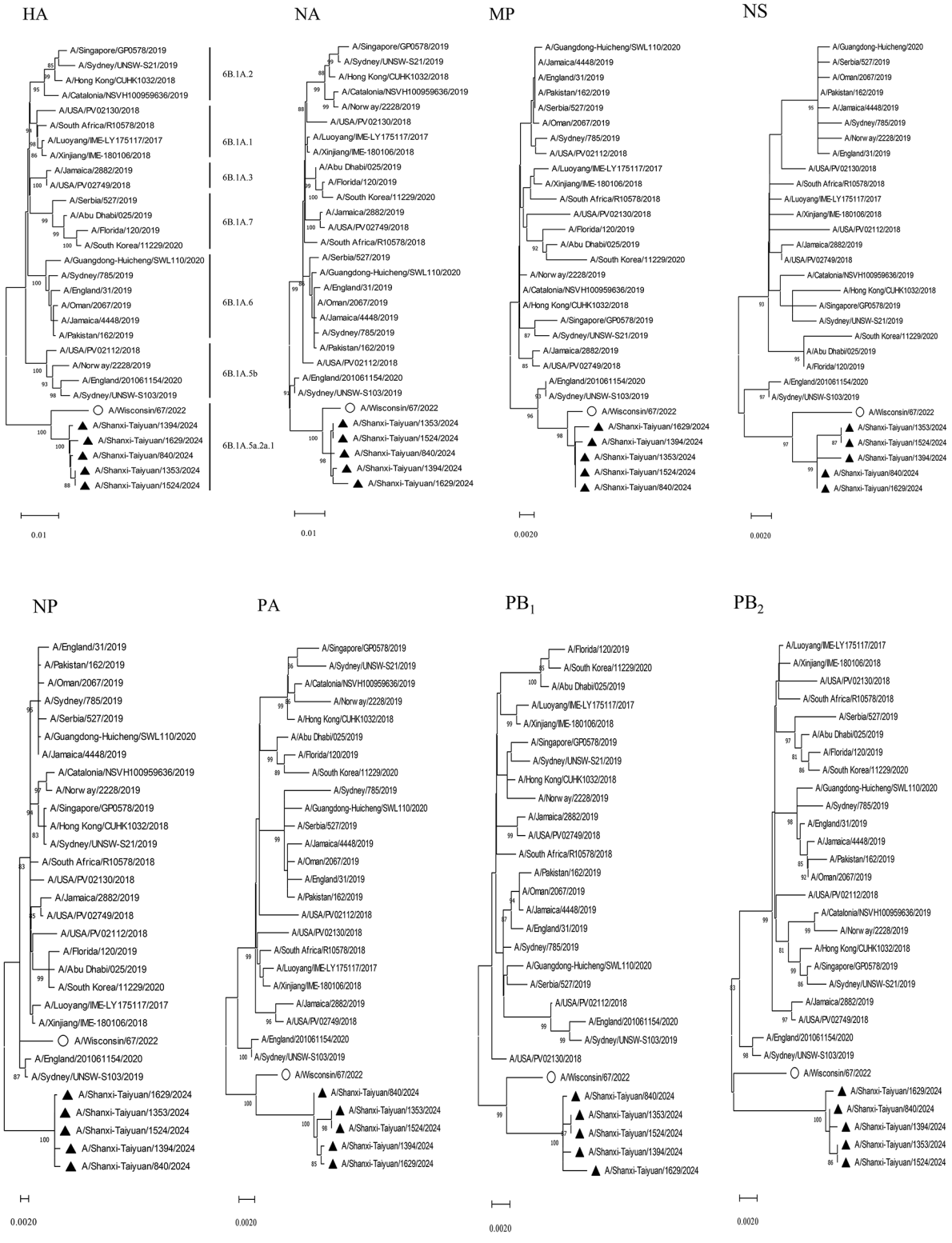

Fig. 2 Phylogenetic analysis of pH1N1 viruses circulating in Taiyuan city during 2023–2024 (black triangles) compared with the Northern Hemisphere vaccine strain A/Wisconsin/67/2022 (white circle) and reference sequences of other branches. Using the Kimura two-parameter model of the Neighbor-joining method of the MEGA 11.0.13 software to construct the evolutionary tree of eight gene segments of IAV. The reliability was run by bootstrap analysis with 1000 replications. The bootstrap value is hidden if it is less than 80. HA, hemagglutinin. NA, neuraminidase. MP, matrix protein. NP, nucleoprotein. NS, nonstructural. PA, polymerase acidic. PB1, polymerase basic 1. PB2, polymerase basic 2

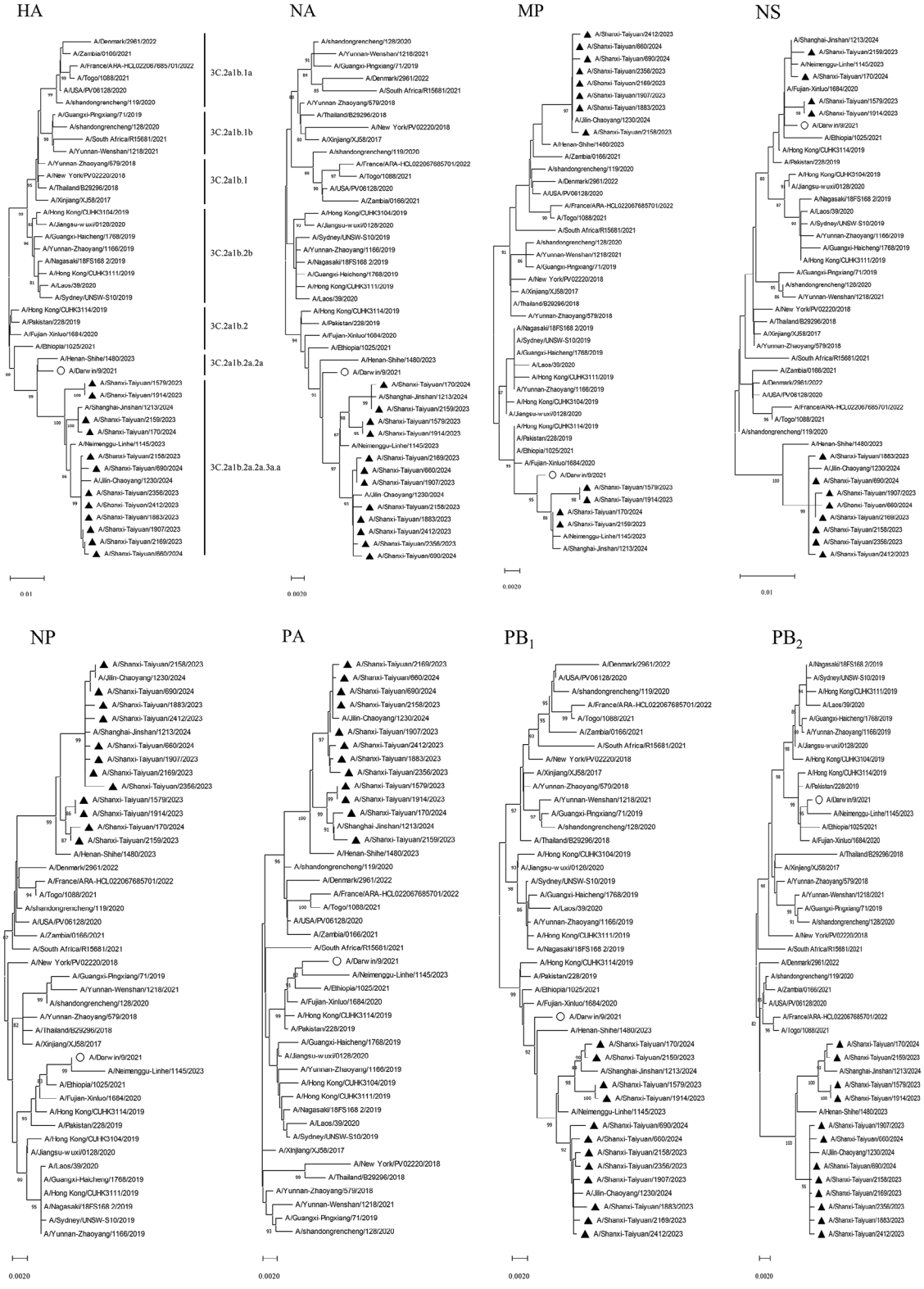

Fig. 3 Phylogenetic analysis of H3N2 viruses circulating in Taiyuan city during 2023–2024 (black triangles) compared with the Northern Hemisphere vaccine strain A/Darwin/9/2021 (white circle) and reference sequences of other branches. Using the Kimura two-parameter model of the Neighbor-joining method of the MEGA 11.0.13 software to construct the evolutionary tree of eight gene segments of IAV. The reliability was run by bootstrap analysis with 1000 replications. The bootstrap value is hidden if it is less than 80. HA, hemagglutinin. NA, neuraminidase. HA, hemagglutinin. NA, neuraminidase. MP, matrix protein. NP, nucleoprotein. NS, nonstructural. PA, polymerase acidic. PB1, polymerase basic 1. PB2, polymerase basic 2

segments varied from 96.9 to 99.8% (Table 3, H3N2). The amino acid similarity of the encoded protein was the lowest at 94.8% and the highest at 99.8% (Table 3, H3N2). Additionally, the evolutionary distances of the HA, NA, MP, NP, NS, PA, PB1, and PB2 genes in the 12 sequences were 0.018–0.021, 0.007–0.010, 0.001–0.014, 0.019–0.023, 0.007–0.026, 0.018–0.020, 0.008–0.010, and 0.020–0.025, respectively, when compared with A/Darwin/9/2021 (Fig. 3).

Amino acid variant analysis

Using the recommended numbering scheme for influenza HA subtypes, compared with A/Wisconsin/67/2022 and A/Darwin/9/2021 (the vaccine strains), we identified nine non-synonymous mutations in the HA amino acid sequence of pH1N1, and 21 in H3N2 (Supplementary Tables S4–S5). In this study, in H3N2, a N138D mutation was found in HA 140-loop, which has removed a potential N-glycosylation site at position 138, and it may also affect the antigenic and other properties of this strain (Fig. 4). Of these mutations, one mutation (N138D) in H3N2 was predicted to be of high phenotypic consequence (FluSurver tool interest levels 2). We mapped this mutation of interest to the antigenic sites of influenza HA protein crystal structures (Fig. 4). For NA proteins, we identified 10 mutation sites in pH1N1 and 11 sites in H3N2 (Supplementary Tables S4–S5), and no mutations in catalytic active sites or drug resistance were found [15]. 28 mutations were found on other protein sites in pH1N1, including: 5 sites in NP, 2 sites in NS1, 2 sites in NEP, 8 sites in PA, 5 sites in PA-X, 3 sites in PB1, and 3 sites in PB2 (Supplementary Table S4). In H3N2, 88 amino acid sites were found to be mutated in the other proteins, including: 1 site in M1, 5 sites in M2, 6 sites in NP, 12 sites in NS1, 17 sites in NEP, 15 sites in PA, 9 sites in PA-X, 9 sites in PB1, 7 sites in PB1-F2, and 9 sites in PB2 (Supplementary Table S5).

Discussion

Influenza remains one of the major public health concerns because it causes annual epidemics and can potentially instigate a global pandemic. In this study, WTP-based and hospital-based epidemiology have all been used to determine human influenza A virus infection rates and epidemic characteristics in Taiyuan.

Wastewater-based epidemiology can be used to obtain information on the circulation of influenza viruses citywide without the need to test many people, as a single wastewater sample is representative of an entire community [8]. We were able to estimate trends in infection incidence and quantify the effective reproductive number of the pH1N1 and H3N2 viruses in five WTPs in Taiyuan city. In this study, the prevalence of influenza virus obtained from wastewater was generally similar to that in hospitals, and the correlation between viral RNA in wastewater and case positivity suggested that the relative concentration of IAVs in wastewater was associated with local disease incidence (Fig. 1). Therefore, more information about local circulating influenza viruses, obtained from wastewater specimens, can inform local clinical, public health, and individual decision making. Medical doctors can use information on circulating viruses to aid in differential diagnosis and making decisions about patient specimen testing that could influence the use of therapeutics.

Previously, it was shown that the delay between the first detection of IAV RNA and the detection of an epidemic in wastewater and clinical data was 9 days [16]. This is in line with our findings, which had a lower resolution, because of wastewater testing at 3-day intervals (min. 2 days, max. 4 days) and clustering of IAV cases on a weekly basis. The incidence of IAV positivity in the population may change rapidly over a period of days [17]; therefore, weekly wastewater collection is not optimal for accurate monitoring of viruses. In the initial stages of an IAV epidemic, the number of infected people is small, and even comprehensive sampling may miss the virus in the influent water of a WTP. This observation was obtained in two recently published articles, which showed that, in the initial stages of IAV epidemics, the number of positive and negative samples fluctuated daily in the tested wastewater samples [8, 16]. Similar observations can also be observed in several monitoring data sets of SARS-CoV-2 in wastewater [18]. Therefore, frequent composite wastewater sampling and virus detection are necessary for sensitive and accurate monitoring of IAV epidemics, but economic factors also need to be considered.

Hospital-based surveillance also has many advantages over wastewater, such as the possibility of obtaining epidemiological characteristics of ILI cases and isolating live virus strains for virological studies. In this study, from July 2023 to June 2024, Taiyuan City's influenza surveillance network laboratory reported a 11.2% positive rate of IAV among ILI-cases in the FHSMU, with H3N2 being the primary prevalent subtype during the winter and spring seasons (Table 2; Fig. 1). Interestingly, our results partially differ from a study conducted in the United States during the same period of time. In Arizona, pH1N1 and H3N2 were co-endemic between October 2023 and February 2024, but pH1N1 is the predominantly endemic strain and contributed the most positive cases [19]. This indicates that regional differences in influenza still exist in the context of increasingly frequent personnel travel and the global spread of the influenza virus.

Outbreaks of IAVs are generally seasonal and cause annual epidemics worldwide. Due to their frequent reassortment and evolution, annual surveillance is of

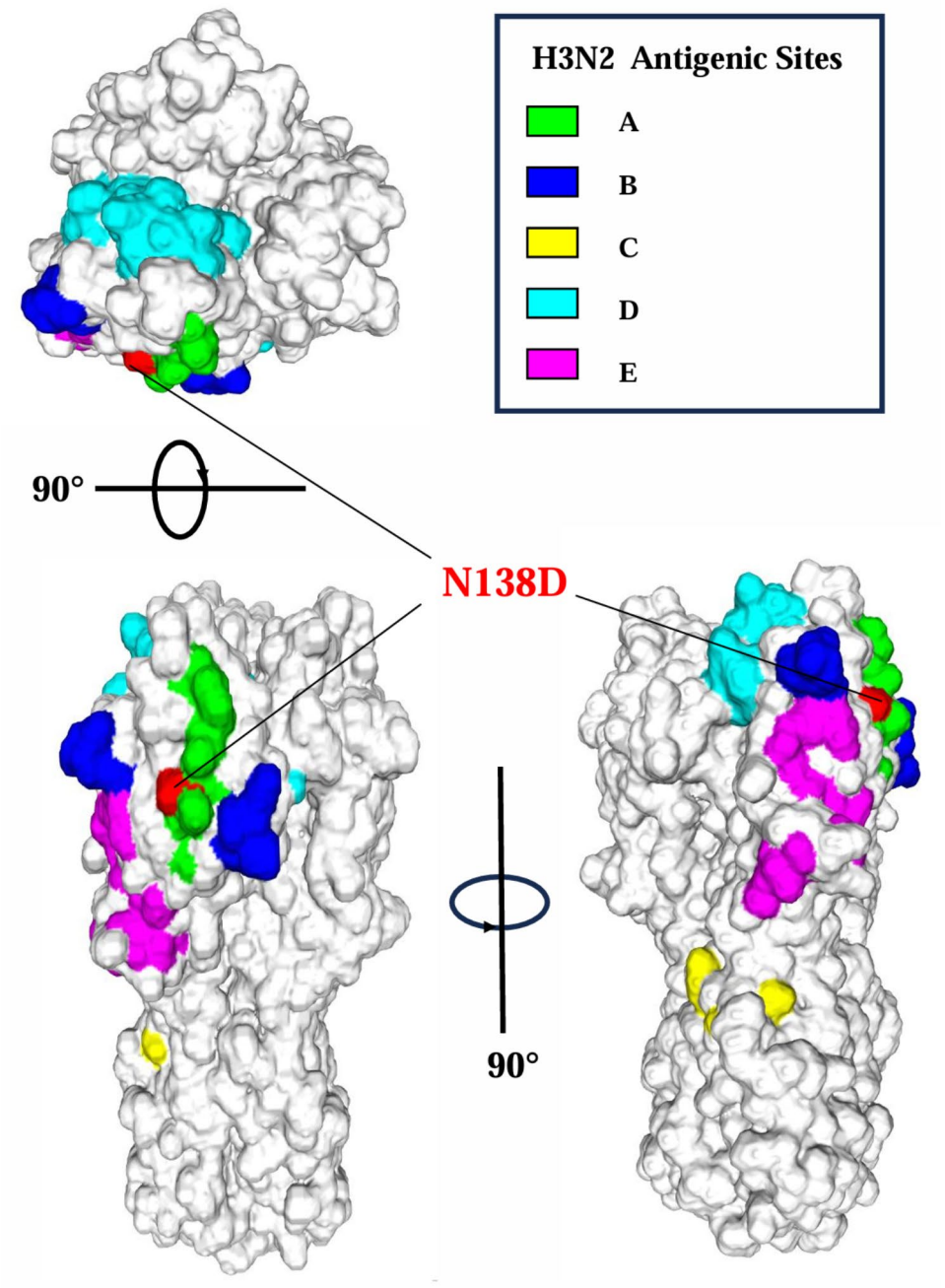

Fig. 4 Structure of influenza HA proteins and locations of important residues. Mutations predicted by FluSurver as interest levels 2 and 3 (red), and antigenic sites (shaded color regions) were mapped to representative crystal structures for H3N2 proteins (PDB: 4WE8)

paramount importance to guide vaccine strategies. Since the end of the COVID-19 pandemic, a global increase in influenza virus has been noted worldwide. In this study, we found that the evolution of the IAVs has not stopped locally, but the genetic diversity of the strains has significantly decreased. In pH1N1, phylogenetic analysis revealed that the strains identified in this study clustered with the corresponding vaccine strains (Fig. 2). Although there are certain genetic differences, they all belong to the 6B.1 A.5a.2a.1 branch, indicating that the local strains are relatively homogeneous in genotype and match the vaccine strain (Fig. 2). In the 2022–2023 influenza season, pH1N1 viruses in Taiyuan city were mainly located in the 6B.1 A.5a.2a branch. Now the pH1N1 viruses in Taiyuan city have evolved from the 6B.1 A.5a.2a branch to the 6B.1 A.5a.2a.1 evolutionary branches (Fig. 2). In H3N2, unlike the vaccine strain belonging to branch 3 C.2a1b.2a.2a, all the Taiyuan strains have evolved into a new branch belonging to 3 C.2a1b.2a.2a.3a.a (Fig. 3). In addition, two distinct branches were found in the MP

and NS genes for the first time in H3N2 strains (Fig. 3); These two branches have been confirmed to exist in other regions of China (Fig. 3), which suggests that a co-prevalence of two different genotypes has evolved among the H3N2 strains prevalent in the world. Therefore, although influenza occurs annually, continuous surveillance is needed for the early detection of variants that can cause human pandemics and to guide health authorities in the proper inclusion of viral lineages in seasonal vaccines.

In this study, the genome of the circulating IAV in Taiyuan from 2023 to 2024 was characterized. Whole genome analysis revealed amino acid substitutions across eight segmented genes (Table S4 and S5). HA is crucial for the antigenic variation of the influenza virus, and its heavy chain region contains both antigenic determinants and receptor binding sites [20]. In detail, the motif at positions 138–140 changed from NES (glyco) to DES (no glyco); and a mutation at the position equivalent to HA 138 has been reported in the literature to be related to antigenic drift / escape mutants and mild drug resistance [21]. In addition, the isolates accumulated mutations at several antigenic sites, which may reduce the protective efficacy of the vaccine. In the context of our study, this indicates that the protective efficacy of the 2023–2024 H3N2 vaccine is likely reduced due to mutations in the antigenic sites of the HA protein, particularly if the vaccination rate in China remained at the same level in 2023 as in prior years. Of course, more research is needed to verify this hypothesis.

Currently, NA, PA, PB1, and PB2 inhibitors are recognized as specific drugs for the treatment of IAVs. The function of NA is to cleave the glycosidic bonds between HA and host cell, allowing virus to be released from the surface. NA inhibitors (such as oseltamivir, zanamivir and peramivir) can specifically bind to the NA active site, inhibit NA activity, and inhibit the release of the IAV. NA plays a critical role in the release stage, and mutational analysis showed NA might be associated with viral entry via amino acids residues R118, E119, D151, R152, W178, I222, E227, E276, R292, and R371 [15]. In this study, we did not find any amino acid mutations occurring at the active site mentioned above, whether in the pH1N1 or H3N2 strains in Taiyuan (Supplementary Table S4 and S5). Assembly of the PB1, PA, and PB2 into the polymerase complex is a prerequisite for virus replication. Amino acid residues in the N-terminal (1-196 amino acids) of PA play a key role in endonuclease activity, protein stability and promoter binding. PA inhibitors are currently important specific drugs for treating influenza, and their sites of action are T20, F24, M34, N37, and I38 [22]. None of the Taiyuan pH1N1 and H3N2 strains examined in this study had mutations at these sites, which indicates that the relevant medications are still available for the treatment of influenza. In the previous work of Sugiyama and colleagues, the key residues determining the PB1-PB2 interface were identified, including L695, F699, V715, V719, A722, I746 andI750 located at the C-terminal end of PB1 [23]. Meanwhile, inhibition of the binding of PA endonuclease and PB2 cap is an effective means to develop influenza virus-targeted drugs. PB1 gene can also undergo mutations that enhance replication ability, such as D27V/N and N44Q, and PB2 binds to the cap structure through Q325, W359, and Y434 alleles [24]. In this study, none of the Taiyuan pH1N1 and H3N2 strains had mutations in these sites either. Although no amino acid mutations were found at the active site, we still found over 10 mutation sites at other positions whether on pH1N1 or H3N2 (Supplementary Table S4 and S5), and with the continuous mutation of the influenza virus genes, it may reduce the binding of the inhibitor to NA and PA, and reduce the virus's sensitivity to the inhibitor, resulting in drug-resistant mutants that make the treatment of influenza more challenging.

Our results demonstrate that although influenza is a respiratoryly transmitted disease, wastewater-based influenza surveillance is an effective method for analyzing the epidemic characteristics of influenza viruses in the environment. We analyzed the influenza epidemic and gene evolution variation in Taiyuan from 2023 to 2024, showing that the antigenic epitopes of the H3N2 influenza virus have partially mutated and formed new variants. Furthermore, the 17 nucleic acid and protein sequences in Taiyuan have undergone some variation, displaying differences in variation sites, homology, evolutionary characteristics, and genetic distances. This suggests that the pH1N1 and H3N2 strains are still evolving and mutating; thus, further strengthening of influenza surveillance is needed.

Abbreviations

IFV	Influenza virus
IAV	Influenza A virus
WTP	Wastewater treatment plant
CDC	Center for Disease Control and Prevention
qRT-PCR	Real-time quantitative reverse transcription polymerase chain reaction

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12879-024-10169-7.

Supplementary Material 1

Acknowledgements

The authors wish to thank all staff of Prof. Jianguo Xu's State Key Laboratory for Infectious Disease Prevention and Control, China CDC.

Author contributions

JW and JX developed the original idea and reviewed the manuscript. LZ contributed to the development of the protocol, analyzed the data, and wrote the manuscript. LZ, JG, PZ, ZJ, PX, XG, ZZ, and LG contributed to the sample collection. LZ analyzed and interpreted the patient data regarding

the influenza. JX, PZ and JG contributed to the development of the protocol and provided technical support. All authors read and approved the final manuscript.

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

Data availability
The datasets presented in this study can be found in GISAID online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary material.

Declarations

Ethics approval and consent to participate
This study was approved by the Institutional Review Board and Human Research Ethics Committee of the Taiyuan Center for Disease Prevention and Control, and written consent was obtained. The children were enrolled after informed and verbal consent was obtained from the parents or guardians. Although the option of written approval was also provided, all participants chose to give their verbal consent due to convenience. Tests were performed on existing respiratory specimens (throat swab), and no additional tests involving any invasive procedures, were performed on patients. These procedures were approved by the above ethics committee.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Department of Microbiology Test, Taiyuan Center for Disease Control and Prevention, No. 22 Huazhang West Street, Taiyuan, Shanxi Province 030032, China
[2]School of Public Health, Shanxi Medical University, NO. 56 Xinjian South Road, Taiyuan, Shanxi Province 030001, China

Received: 2 September 2024 / Accepted: 30 October 2024
Published online: 12 November 2024

References

1. Krammer F, Smith GJD, Fouchier RAM, Peiris M, Kedzierska K, Doherty PC, et al. Influenza Nat Reviews Disease Primers 2018. 2018;4:1.
2. Klingen TR, Loers J, Stanelle-Bertram S, Gabriel G, McHardy AC. Structures and functions linked to genome-wide adaptation of human influenza a viruses. Sci Rep 2019. 2019;9:1.
3. Ferrari A, Schiavetti I, Ogliastro M, Minet C, Sibilio R, Giberti I et al. Co-detection of respiratory pathogens among ILI patients: characterization of samples collected during the 2018/19 and 2019/20 pre-pandemic seasons. BMC Infectious Diseases 2024 24:1. 2024;24:1–10.
4. Nayak J, Hoy G, Gordon A. Influenza in children. Cold Spring Harb Perspect Med. 2021;11:1–17.
5. Gaitonde DY, Moore FC, Morgan MK. Influenza: diagnosis and treatment. Am Fam Physician. 2019;100:751–8.
6. Chao S, Wang Y, Wu B, Li R, Dong J, Ji L, et al. Characterization of viral infections in children with influenza-like-illness during December 2018–January 2019. Front Cell Infect Microbiol. 2023;13:1351814.
7. Li L, Liu T, Wang Q, Ding Y, Jiang Y, Wu Z et al. Genetic characterization and whole-genome sequencing-based genetic analysis of influenza virus in Jining City during 2021–2022. Front Microbiol. 2023;14.
8. Boehm AB, Hughes B, Duong D, Chan-Herur V, Buchman A, Wolfe MK, 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:e340–8.
9. Zhao L, Xiong Y, Meng D, Guo J, Li Y, Liang L, et al. An 11-year study of shigellosis and Shigella species in Taiyuan, China: active surveillance, epidemic characteristics, and molecular serotyping. J Infect Public Health. 2017;10:794–8.
10. Zhao L, Shi X, Meng D, Guo J, Li Y, Liang L, et al. Prevalence and genotype distribution of group a rotavirus circulating in Shanxi Province, China during 2015–2019. BMC Infect Dis. 2021;21:1–14.
11. Lee MJ, Quantifying. SARS-CoV-2 viral load: current status and future prospects. Expert Rev Mol Diagn. 2021;21:1.
12. Boehm AB, Hughes B, Duong D, Chan-Herur V, Buchman A, Wolfe MK, 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.
13. Torii S, Furumai H, Katayama H. Applicability of polyethylene glycol precipitation followed by acid guanidinium thiocyanate-phenol-chloroform extraction for the detection of SARS-CoV-2 RNA from municipal wastewater. Sci Total Environ. 2021;756:143067.
14. Stray SJ, Pittman LB. Subtype- and antigenic site-specific differences in biophysical influences on evolution of influenza virus hemagglutinin. Virol J. 2012;9:91.
15. Chen F, Liu T, Xu J, Huang Y, Liu S, Yang J. Key amino acid residues of neuraminidase involved in influenza a virus entry. Pathog Dis. 2019;77.
16. Toribio-Avedillo D, Gómez-Gómez C, Sala-Comorera L, Rodríguez-Rubio L, Carcereny A, García-Pedemonte D, et al. Monitoring influenza and respiratory syncytial virus in wastewater. Beyond COVID-19. Sci Total Environ. 2023;892:164495.
17. Shaw Stewart PD, Bach JL. Temperature dependent viral tropism: understanding viral seasonality and pathogenicity as applied to the avoidance and treatment of endemic viral respiratory illnesses. Rev Med Virol. 2022;32:e2241.
18. Ahmed W, Simpson SL, Bertsch PM, Bibby K, Bivins A, Blackall LL, et al. Minimizing errors in RT-PCR detection and quantification of SARS-CoV-2 RNA for wastewater surveillance. Sci Total Environ. 2022;805:149877.
19. Maqsood R, Smith MF, Holland LRA, Sullins RA, Holland SC, Tan M et al. Influenza Virus Genomic Surveillance, Arizona, USA, 2023–2024. Viruses. 2024;16.
20. Liu STH, Behzadi MA, Sun W, Freyn AW, Liu WC, Broecker F, et al. Antigenic sites in influenza H1 hemagglutinin display species-specific immunodominance. J Clin Invest. 2018;128:4992.
21. Philpott M, Hioe C, Sheerar M, Hinshaw VS. Hemagglutinin mutations related to attenuation and altered cell tropism of a virulent avian influenza a virus. J Virol. 1990;64:2941.
22. Takashita E. Influenza polymerase inhibitors: mechanisms of action and resistance. Cold Spring Harb Perspect Med. 2021;11.
23. Sugiyama K, Obayashi E, Kawaguchi A, Suzuki Y, Tame JRH, Nagata K, et al. Structural insight into the essential PB1-PB2 subunit contact of the influenza virus RNA polymerase. EMBO J. 2009;28:1803–11.
24. Yuan S, Wen L, Zhou J. Inhibitors of Influenza A Virus polymerase. ACS Infect Dis. 2018;4:218–23.

Publisher's note