







# Influenza A Virus Field Surveillance at a Swine-Human Interface

 Benjamin L. Rambo-Martin,<sup>a</sup>  Matthew W. Keller,<sup>b</sup> Malania M. Wilson,<sup>c</sup> Jacqueline M. Nolting,<sup>d</sup>  Tavis K. Anderson,<sup>e</sup> Amy L. Vincent,<sup>e</sup> Ujwal R. Bagal,<sup>a</sup> Yunho Jang,<sup>c</sup> Elizabeth B. Neuhaus,<sup>c</sup> C. Todd Davis,<sup>c</sup>  Andrew S. Bowman,<sup>d</sup> David E. Wentworth,<sup>c</sup> John R. Barnes<sup>c</sup>

<sup>a</sup>Battelle Memorial Institute, Atlanta, Georgia, USA

<sup>b</sup>Oak Ridge Institute of Science and Education (ORISE), Oak Ridge, Tennessee, USA

<sup>c</sup>Influenza Division, National Center for Immunization and Respiratory Diseases (NCIRD), Centers for Disease Control and Prevention (CDC), Atlanta, Georgia, USA

<sup>d</sup>Department of Veterinary Preventive Medicine, The Ohio State University, Columbus, Ohio, USA

<sup>e</sup>National Animal Disease Center, Agricultural Research Service (ARS), U.S. Department of Agriculture (USDA), Ames, Iowa, USA

Benjamin L. Rambo-Martin and Matthew W. Keller contributed equally to this work. Author order was determined by ongoing collaborative agreements and in increasing order of seniority.

**ABSTRACT** While working overnight at a swine exhibition, we identified an influenza A virus (IAV) outbreak in swine, Nanopore sequenced 13 IAV genomes from samples we collected, and predicted in real time that these viruses posed a novel risk to humans due to genetic mismatches between the viruses and current pre-pandemic candidate vaccine viruses (CVVs). We developed and used a portable IAV sequencing and analysis platform called *Mia* (Mobile Influenza Analysis) to complete and characterize full-length consensus genomes approximately 18 h after unpacking the mobile lab. Exhibition swine are a known source for zoonotic transmission of IAV to humans and pose a potential pandemic risk. Genomic analyses of IAV in swine are critical to understanding this risk, the types of viruses circulating in swine, and whether current vaccines developed for use in humans would be predicted to provide immune protection. Nanopore sequencing technology has enabled genome sequencing in the field at the source of viral outbreaks or at the bedside or pen-side of infected humans and animals. The acquired data, however, have not yet demonstrated real-time, actionable public health responses. The *Mia* system rapidly identified three genetically distinct swine IAV lineages from three subtypes, A(H1N1), A(H3N2), and A(H1N2). Analysis of the hemagglutinin (HA) sequences of the A(H1N2) viruses identified >30 amino acid differences between the HA1 of these viruses and the most closely related CVV. As an exercise in pandemic preparedness, all sequences were emailed to CDC collaborators who initiated the development of a synthetically derived CVV.


**IMPORTANCE** Swine are influenza virus reservoirs that have caused outbreaks and pandemics. Genomic characterization of these viruses enables pandemic risk assessment and vaccine comparisons, though this typically occurs after a novel swine virus jumps into humans. The greatest risk occurs where large groups of swine and humans congregate. At a large swine exhibition, we used Nanopore sequencing and on-site analytics to interpret 13 swine influenza virus genomes and identified an influenza virus cluster that was genetically highly varied to currently available vaccines. As part of the National Strategy for Pandemic Preparedness exercises, the sequences were emailed to colleagues at the CDC who initiated the development of a synthetically derived vaccine designed to match the viruses at the exhibition. Subsequently, this virus caused 14 infections in humans and was the dominant U.S. variant virus in 2018.

**Citation** Rambo-Martin BL, Keller MW, Wilson MM, Nolting JM, Anderson TK, Vincent AL, Bagal UR, Jang Y, Neuhaus EB, Davis CT, Bowman AS, Wentworth DE, Barnes JR. 2020. Influenza A virus field surveillance at a swine-human interface. *mSphere* 5:e00822-19. <https://doi.org/10.1128/mSphere.00822-19>.

**Editor** Anice C. Lowen, Emory University School of Medicine

This is a work of the U.S. Government and is not subject to copyright protection in the United States. Foreign copyrights may apply.

Address correspondence to John R. Barnes, [fzq9@cdc.gov](mailto:fzq9@cdc.gov).

 New technology is allowing @CDCflu scientists to gather detailed genetic sequencing information from #flu viruses in the field. This new technology can rapidly speed up response time allowing scientists to better understand and act in an outbreak situation

**Received** 6 November 2019

**Accepted** 7 January 2020

**Published** 5 February 2020

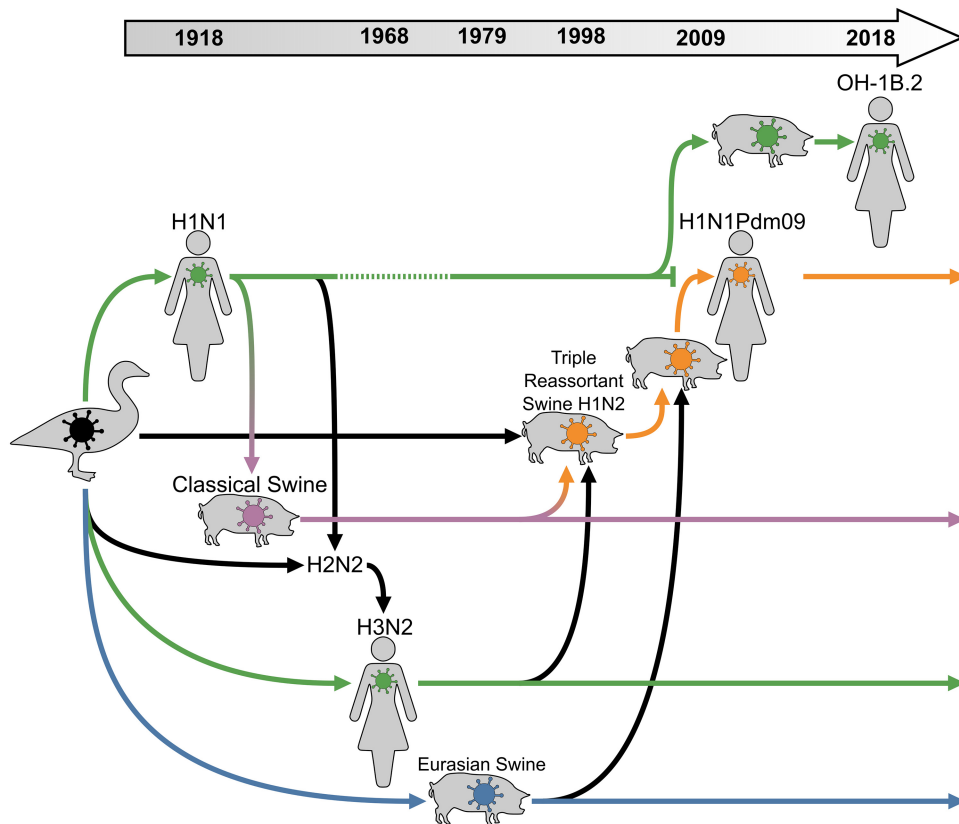
**KEYWORDS** influenza, mobile sequencing, Nanopore sequencing, pandemic preparedness, swine influenza

Mobile sequencing technology, specifically the MinION sequencing platform from Oxford Nanopore Technologies (Oxford, UK), has successfully generated rapid genomic surveillance data on-site during outbreaks of Ebola (1), Zika (2), and Lassa (3) viruses. These studies relied on the transfer of raw read data to centralized institutions harboring high-performance computational resources to perform extensive phylogenetic analyses. These analyses resulted in a deep understanding of the viral evolution during the outbreak, as well as an estimation of transmission chains. Their reliance on vast databases and computationally intensive algorithms limited their ability to influence real-time prevention strategies and potentially stop further transmission events. Furthermore, the lack of vaccines to these viruses meant that the real-time-derived sequence data could not be leveraged to identify a suitable vaccine. On the other hand, influenza viruses have multiple enzootic reservoirs, cause annual epidemics in humans, and rapidly evolve, which necessitates constant genomic surveillance and iterative vaccine development. Given this need, we have an opportunity to mitigate the risk of an influenza outbreak at its source in real time by analyzing the viral genomes present and, subsequently, developing the best-matched candidate vaccine virus (CVV).

Influenza A viruses (IAVs) circulating in swine have the potential to infect humans and are capable of causing pandemics (4), as evidenced by the 2009 H1N1 virus pandemic (Fig. 1). Since that pandemic, there have been 465 human cases of swine-origin influenza viruses (termed variant influenza viruses) in the United States resulting from exposure to swine, most commonly during attendance at agricultural fairs, and sporadic cases have been reported periodically in other countries (5). Exhibitors travel with their swine across geographical regions in the United States to attend national, state, and local exhibitions. At these exhibitions, swine of various ages and from many geographical regions intermingle. Throughout the exhibition, swine are walked around the barns, shuttled into staging pens, and exhibited in a corral simultaneously with a dozen or more swine. The swine are typically shepherded by children, as many exhibition rules require that individuals presenting a swine within the exhibition ring be under the age of 20. The exhibition itself can last up to a week, during which time swine are living in immediate proximity to others and have direct contact with humans. Children in particular may be immunologically naive to swine influenza viruses that are genetically and antigenically distinct from recent seasonal influenza viruses or vaccines to which they have been exposed (<https://www.cdc.gov/flu/swineflu/variant-cases-us.htm>).

Due to the sporadic nature of these infections, the immunologically naive population affected, and the persistent circulation of evolutionarily diverse IAVs in the swine host, the Centers for Disease Control and Prevention and other World Health Organization (WHO) Collaborating Centers for Influenza and Essential Regulatory Laboratories have developed pre-pandemic CVVs that target specific swine influenza virus subtypes and HA gene lineages that have caused variant virus infections (6). To date, 11 CVVs representing antigenically diverse groups of swine influenza viruses have been proposed or already developed for pre-pandemic preparedness purposes (<https://www.who.int/influenza/vaccines/virus/en/>).

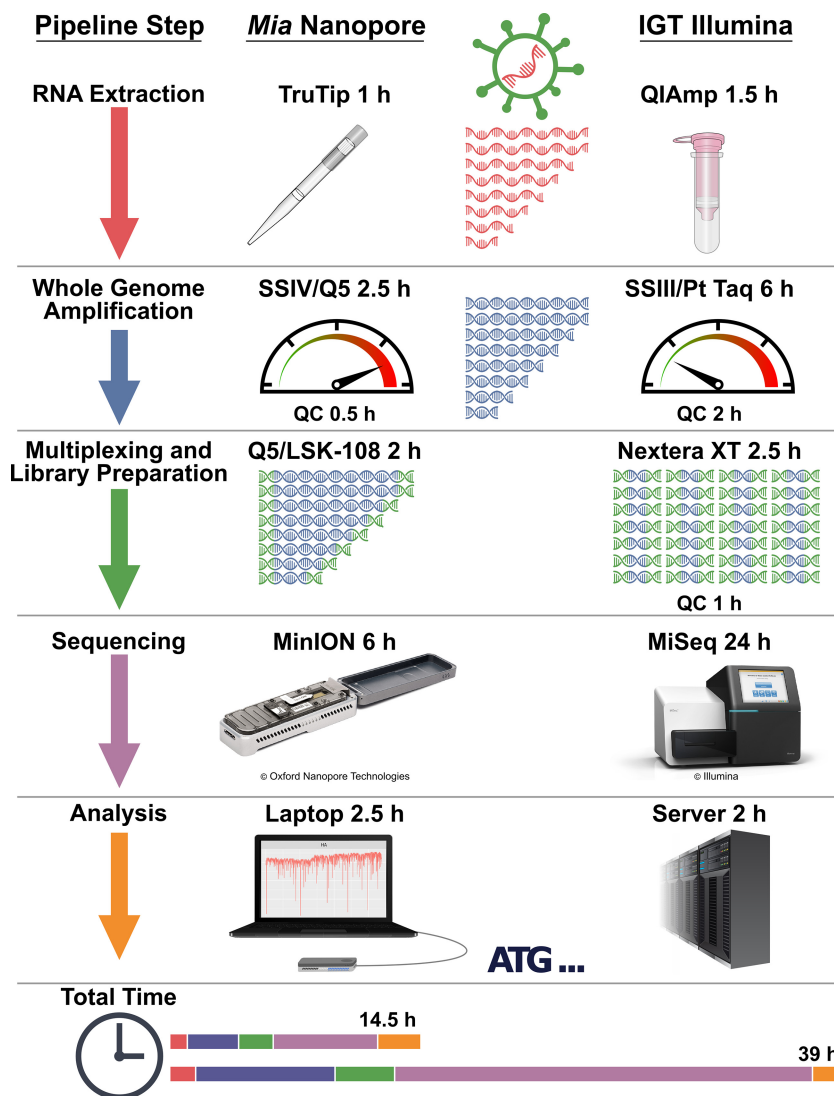
Collecting samples in the field, shipping them to centralized laboratories, sequencing their genomes, and analyzing them often takes several weeks or months. Thus, the current surveillance process is less able to contribute real-time data during the emergence of pandemic threats (7). The increased risk at this unique interface warrants concerted proactive surveillance and simultaneously serves as a proving ground for new real-time in-field genomic approaches to help prevent zoonotic infections and potential outbreaks. Here, we describe the development and deployment of a rapid and portable IAV sequencing pipeline we call *Mia* (Mobile Influenza Analysis [miə, MEE-uh]) and the genomic results obtained from this surveillance.



**FIG 1** Genetic relationship of human and swine virus lineages. Green, orange, purple, and blue arrows follow the evolution of hemagglutinin, while black errors follow other segments. Human seasonal lineages are green, classical swine-H1N1pdm09 is orange, the classical swine lineage is purple, and the Eurasian swine lineage is blue. In 2009, influenza A/H1N1Pdm09 virus displaced the previous human seasonal A/H1N1 virus from humans which continues circulating in swine. This virus reassorted, picking up the N2 segment from the triple-reassortant viruses into the A/H1N2 Other-Human (OH)-1B.2 lineage that caused human infections in the summer of 2018. The history of this HA can be traced back to the avian virus that caused the 1918 pandemic, through extinction and reemergence, and through decades of human seasonal activity. This figure was created using data from reference 31.

## RESULTS

**Development of the *Mia* platform.** We created a mobile influenza virus genomics platform to estimate the risk to humans posed by viruses within a local IAV outbreak by performing in-field extraction and amplification of influenza A viruses, complete genome sequencing, automated genome assembly, BLAST comparisons, phylogenetic analysis, and variant detection to candidate vaccine viruses. We used our high-throughput influenza virus sequencing pipeline as a template for developing *Mia* (Fig. 2). Our goal was a rapid workflow in a compact platform that is easily transported by two people on a commercial airliner. For RNA extraction, we elected to use Akonni Biosystems' TruTip rapid RNA kit due to its small footprint and demonstrated effectiveness on influenza A viruses (8). We chose the mini8 by miniPCR due to its small size, ability to run 8 samples at a time, low cost, low power needs, and overall simplicity. Due to the high fidelity of our rapid amplification strategy, we were able to simplify amplicon quality control from fragment analysis to fluorometric quantification, which is already a necessary step for amplicon pooling. For sequencing, we chose the Nanopore sequencing platform MiniION for its extremely small size, rapid sequencing speed, and live access to data as it is sequencing. Direct RNA sequencing of IAVs has recently been demonstrated, and while this technique is extremely fast (<2 h from RNA to sequencing), it is not yet suitable for clinical applications due to a lack of sensitivity and multiplexing (9). The final inventory is available in Table S1 posted at <https://figshare.com/s/b4cc885050283a40dfcd>.



**FIG 2** Timeline comparison of portable versus centralized surveillance pipelines. The *Mia* pipeline is modeled after the Influenza Genomics Team’s (IGT) high-throughput sequencing pipeline with special considerations for speed and portability. It is expected to obtain results from 24 samples in a maximum of 14 h and 30 min. The IGT pipeline is designed for throughput and accuracy and could obtain results from 96 samples in 39 h if operated continuously.

**Sample processing.** Our surveillance target was a large agricultural event featuring exhibition swine. The swine began arriving on Sunday (day 0) and received an initial veterinary screening upon entry. We began scouting swine on day 3 for influenza-like illness (ILI) and noted the locations in the exhibition barn of animals with clinical signs of respiratory disease. During the day, the barns were very crowded with swine, presenters and their families, farm hands, visitors, and event staff. To avoid interfering with the event’s proceedings, swabbing and sequencing were conducted during the event’s off-hours. On day 4, we swabbed ILI-identified swine and their pen neighbors ( $n = 94$ ) for influenza A virus (IAV) with the Flu Detect swine kits. We detected seven IAV-positive samples but suspected that additional samples collected from pen neighbors might be positive by sequence analysis. At 11 p.m. on day 4, we erected the *Mia* platform inside the exhibition barn and began the workflow on 24 samples collected from the rapid-test-positive swine and their immediate neighbors. Overnight, we extracted RNA, amplified the full genome, barcoded amplicons, and quantified the barcoded amplicons. Sample concentrations ranged from 12.7 to 41.9 ng/ $\mu$ l (Table S2

**TABLE 1** Sequencing results<sup>a</sup>

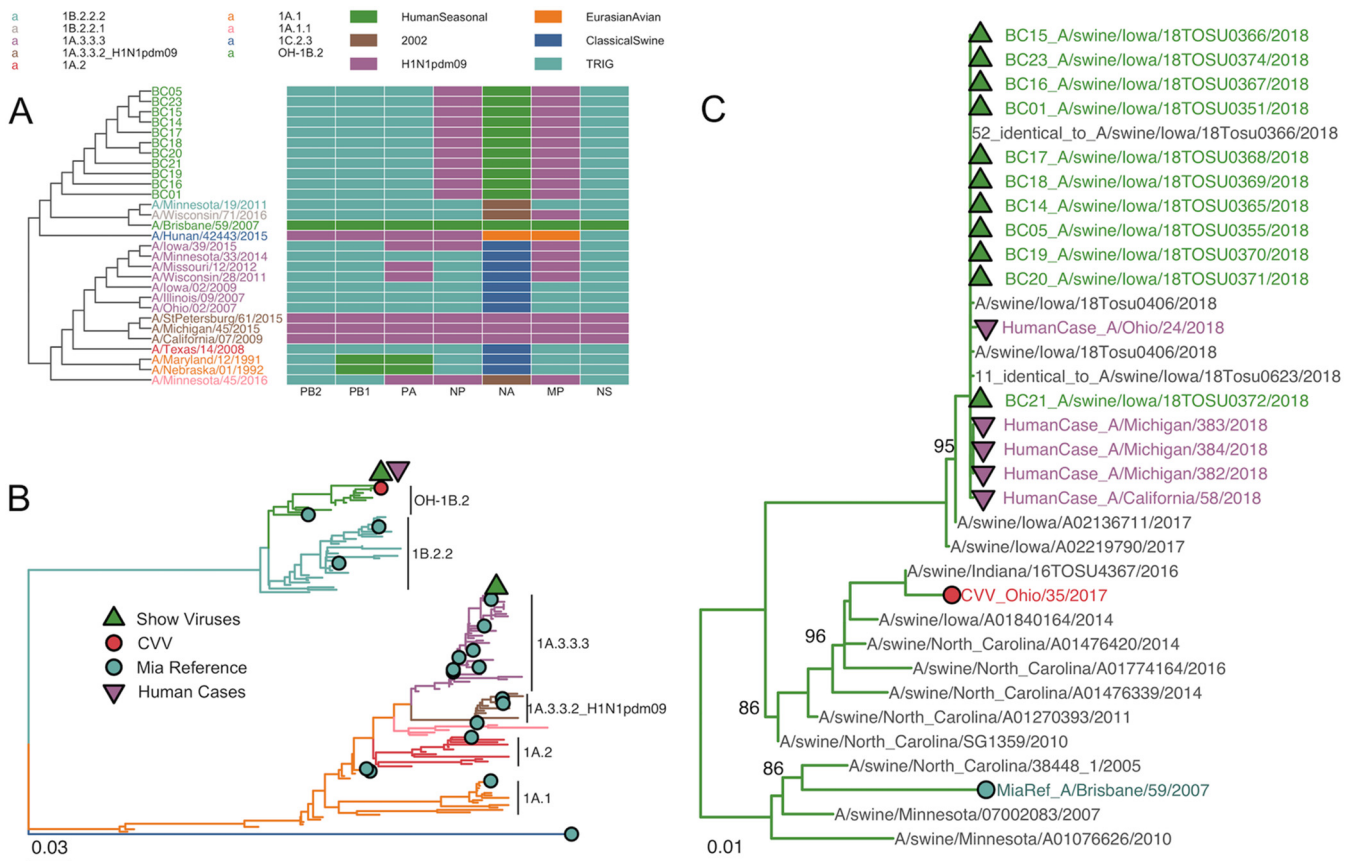
Barcode	IAV subtype	C <sub>T</sub>	Avg HA coverage <sup>a</sup>		Avg identity
			<i>Mia</i> Nanopore	IGT Illumina	
01	H1N2	29.1	+	+++	0.995
02	—	33.6	—	+	—
03	—	35.2	—	+	—
04	—	—	—	+	—
05	H1N2	27.3	+	+++	0.989
06	H1N1	28.0	+	+++	0.993
07	—	—	—	+	—
08	—	31.5	—	+	—
09	—	—	—	+	—
10	—	33.7	—	+	—
11	—	34.8	—	+	—
12	—	36.9	—	+	—
13	—	—	—	+	—
14	H1N2	23.5	+++	+++	0.997
15	H1N2	26.1	+++	+++	0.997
16	H1N2	26.0	+++	+++	0.997
17	H1N2	24.4	+++	+++	0.997
18	H1N2	23.1	+++	+++	0.997
19	H1N2	34.4	+++	+++	0.997
20	H1N2	24.1	+++	+++	0.997
21	H1N2	26.4	+++	+++	0.996
22	H3N2	30.9	+	+++	0.960
23	H1N2	28.6	+	+++	0.997
24	—	—	—	+	—

<sup>a</sup>Sequencing results. HA coverage is indicated for each technique by its threshold (20× coverage for Nanopore and 100× coverage for Illumina): — is for no coverage, + is for coverage below the threshold, and +++ is for coverage above the thresholds. Complete coverage data can be found in Table S3. Average identity is displayed for all segments.

[<https://figshare.com/s/b4cc885050283a40dfcd>]). We normalized and pooled the bar-coded samples to 1 μg total. The final library amounted to 279 ng. The library was ready for sequencing at 6 a.m., and at this time, people were arriving at the barn to begin exhibition activities. To avoid disrupting the event, once the sequencing was initiated, we transported the sequencer attached to the laptop powered by the laptop's battery to a hotel room.

**Base calling and genome assembly.** Sequencing produced 99,988 reads that were base called and demultiplexed with Albacore v2.2.7. A total of 67,105 reads passed a basic quality control (QC) threshold Q value of ≥7, and 22,552 of these reads were assigned to one of 24 barcodes (Fig. S1 [<https://figshare.com/s/b4cc885050283a40dfcd>]). The distribution of read lengths followed an expected pattern, with peaks corresponding to IAV segment lengths (Fig. S1 at the URL mentioned above). We assembled demultiplexed reads with IRMA (10) and obtained full IAV genomes for 13 of the 24 samples, including all samples that had threshold cycle (C<sub>T</sub>) values of <30 (C<sub>T</sub> 30 ≈ 100 50% egg infective doses per ml [EID<sub>50</sub>/ml]; data not shown) (Table 1; see also Table S3 and Fig. S2 to S21 at the URL mentioned above).

**Real-time phylogenetic analysis.** *Mia* produced automated per-segment nucleotide alignments from all sequenced samples and a reference panel of variant IAVs that were transmitted from swine to humans between 2003 and 2016. The reference panel of viruses has broad diversity, including swine-origin viruses of classical swine lineages (1A.1 [H1-α], 1A.2 [H1-β], and 1A.3.3.3 [H1-γ]), human seasonal lineages (1B.2.1 [H1-δ2] and 1B.2.2 [H1-δ1]), Eurasian avian-like lineages (1C.1 and 1C.2), an avian-origin North American lineage, and human seasonal lineages (H3, H1 pdm09, N1, and N2) (11–14). Each segment alignment was automatically built into separate phylogenetic trees alongside a heat map displaying the top BLAST result per segment against the annotated references (Fig. S22 to S29 [<https://figshare.com/s/b4cc885050283a40dfcd>]). Viruses collected during our investigation included 1 H3N2 (barcode 22), 1 H1N1 (barcode 06), and 11 H1N2 (barcodes 01, 05, 14 to 21, and 23) viruses. These 11 H1N2



**FIG 3** Phylogenetic relationships of HA-H1 sequences from exhibition-sampled swine, broad swine, swine variant, and late summer 2018 human cases of swine variant IAVs. (A) Re-creation of the phylogenetic cladogram and genome constellation heat map displayed in *Mia*'s R Shiny application using the swine influenza virus H1 HA global nomenclature system. This automatically generated figure showed us that our sampled viruses had HA sequences most similar to the Other-Human-1B.2 (OH-1B.2) A/Brisbane/59/2007 *Mia* reference sequence and that they all shared a similar genome constellation. (B) Maximum likelihood phylogeny of swine IAV HA-H1 sequences incorporating the swine variant viruses *Mia* used as references, including the nearest candidate vaccine virus (Ohio/35/2017) to our outbreak of 1B.2.1 IAVs. (C) Detail of the maximum likelihood phylogeny of H1 1B.2.1 and OH-1B.2 sequences including our field-processed IAVs (BCxx, green), 65 additional field samples that were laboratory processed (blue), the nearest CVV (red), *Mia* reference (light blue), and five human cases of swine variant (purple) IAVs. Bootstrap values are annotated on the internal nodes.

viruses contained identical genome constellations. Both the HA and NA gene segments were closely related to 2017/2018 H1N2 influenza viruses circulating in the U.S. swine population and have been sporadically detected in previous H1N2v zoonotic infections. The HA gene segments belonged to the H1- $\delta$ 2 or clade 1B.2.1, using the swine influenza virus H1 HA global nomenclature system. This lineage is common to swine in the United States and is derived from human seasonal H1N1 virus that was transmitted from humans to swine in the early 2000s (15). Each of the H1N2 viruses possessed an NA gene derived from the 1998 human seasonal H3N2 lineage now found in U.S. swine influenza viruses. This NA is a distinct gene lineage from the NA gene found in the closest CVV, A/Ohio/35/2017 (2002 human seasonal NA lineage). The PB2, PB1, PA, and NS gene segments were of triple-reassortant internal gene (TRIG) cassette origin commonly found in swine influenza viruses in the United States. The NP and M gene segments belonged to the 2009 H1N1 pandemic virus lineage (Fig. 3A).

**Real-time comparison of HA proteins to candidate vaccine viruses.** *Mia* also provided amino acid difference tables of the mature HA protein sequence versus a set of WHO candidate vaccine viruses (CVVs). A/Ohio/9/2015 and A/Ohio/13/2017 were the nearest CVVs to the H1N1 and H3N2 viruses and had 2 and 0 antigenic changes, respectively. These two viruses also had relatively low sequence coverage of the HA segments (Fig. S4 and S19 [<https://figshare.com/s/b4cc885050283a40dfcd>]). A/Ohio/35/2017 was the most similar CVV in our database to the 11 H1N2 viruses and had 32



**FIG 4** Mature H1 amino acid sequence alignment of H1 1B.2.1 sequences from *Mia* Nanopore sequencing versus A/Ohio/35/2017, the nearest CVV. Red stars indicate antigenic sites. This figure was created postfield for publication purposes; however, the same information was viewable in an alignment viewer in the field as well as in *Mia*'s dashboard table.

amino acid differences compared to their consensus sequence (Fig. 4) (see also Fig. S30 and Table S4 at the URL mentioned above). Seven differences were identified in predicted H1 antigenic sites, including 4 differences in antigenic site A or B (16). More generally, the H1N2 virus HA genes were only ~93% identical to the nucleotide sequence of A/Ohio/35/2017 and, as discussed above, had NA genes belonging to different ancestral lineages. Based on the genetic differences and markers of antigenic drift detected by the analysis, the HA and NA consensus sequences, including untranslated regions (UTRs), were electronically transmitted to the CDC's Vaccine Preparedness Team in the Influenza Division to initiate the development of a synthetic candidate vaccine virus. We were able to start this intervention from the field roughly 18 h after the initial setup of *Mia* (Table S5 [<https://figshare.com/s/b4cc885050283a40dfcd>]).

**Postfield analyses.** To assess the MinION-derived sequence accuracy, we transported RNA and DNA amplicons created on-site to the CDC in Atlanta, GA, for confirmatory sequencing via the Illumina MiSeq platform using the standard Influenza Genomic Team (IGT) surveillance pipeline (10). On the MiSeq platform, 13 samples had completed genomes with average coverage of  $\geq 100\times$  for each segment (IGT's standard QC threshold). Of those 13 samples, 12 had a MinION-derived mean coverage of  $\geq 5\times$  across segments, with 8 segments having  $\geq 20\times$  coverage across HA (Table S3 [<https://figshare.com/s/b4cc885050283a40dfcd>]), the primary target of protective immune responses and the most important protein for vaccine development. There was high concordance ( $R^2 = 0.87$ ; Fig. S31 at the URL mentioned above) of the mean segment coverage produced by MinION and MiSeq for field-derived amplicons. Consensus identities averaged 99.3% across the 13 genomes (Table S3 and Fig. S32 at the URL mentioned above).

As our real-time *Mia* database contained human cases of swine IAV, we performed additional postfield phylogenetic analyses using the outbreak data, the *Mia* database,

and additional swine IAV sequences that represent the genetic diversity and evolutionary relationships of contemporary swine IAV (11, 12, 17). Compared to 126 swine H1 sequences, our *Mia*-identified swine H1N2 viruses were in a monophyletic clade with contemporary 1B.2.1 (H1- $\delta$ 2) HA genes from the human seasonal lineage (Fig. 3B and C). These viruses predominantly pair with N2 sequences that have been circulating in swine since at least 1998 as part of a triple-reassortant internal gene (TRIG) group of viruses (12, 13) (Fig. 1) (see also Fig. S36 [<https://figshare.com/s/b4cc885050283a40dfcd>]). Across all segments, MiSeq-derived data mirror the tree topology created in real time in the field from MinION-derived sequences (Fig. 3) (see also Fig. S22 to S29 and S33 to S42 in the URL mentioned above).

To compare the viral diversity we found in our MinION-derived viral sequences to what fully existed at the exhibition, we performed our standard IAV surveillance procedure in which we sampled 425 additional swine regardless of ILI. From these samples, we detected 136 (32%) IAV-positive specimens by real-time reverse transcription-PCR (rRT-PCR) and successfully sequenced 65 full genomes on a MiSeq platform. We found that the diversity of viruses collected and sequenced by more traditional surveillance methods was also detected in similar proportions by our *Mia* field sequencing (Fig. 3) (see also Fig. S33 to S42 [<https://figshare.com/s/b4cc885050283a40dfcd>]).

## DISCUSSION

We successfully developed a mobile next-generation sequencing (NGS) suite called *Mia* and deployed it to a high-priority swine-human interface that resulted in our identification of an influenza A virus outbreak. We created *Mia* by selecting and optimizing laboratory equipment that fits into a cooler and two standard-sized suitcases (Fig. S43 [<https://figshare.com/s/b4cc885050283a40dfcd>]), which can be set up and operated in the field by two people (Fig. S44 at the URL mentioned above) to produce a high-quality multiplexed NGS amplicon library in 7 h. Our automated analysis pipeline performs base calling, quality control, genome assembly, phylogenetic analysis, BLAST searches, and amino acid comparisons to current CVVs without an Internet connection. We are able to monitor all of these results in real time on the laptop performing the analyses. *Mia* will be continuously improved and adapted for use in harsher field environments (see Text S1 at the URL mentioned above).

In the field, the data generated indicated that while multiple lineages were cocirculating, the numerically dominant subtype, H1N2, was genetically different from the most similar WHO pre-pandemic CVV, including variation in known antigenic sites. Moreover, we determined that these viruses posed a risk of causing disease in a young population, as similar H1 viruses disappeared from seasonal circulation in humans in 2010 and were replaced by the 2009 H1N1 pandemic virus. Therefore, the sequence data were used to initiate the synthesis of an optimal CVV approximately 18 h after unpacking *Mia*. This proactive strategy of identifying the predominant swine influenza viruses in exhibition pigs prior to the start of annual agricultural fairs in the United States proved to be very useful, as zoonotic infections caused by the predominant A(H1N2) virus found in exhibition pigs were detected in humans exposed at fairs in July and August 2018 ([https://gis.cdc.gov/grasp/fluview/Novel\\_Influenza.html](https://gis.cdc.gov/grasp/fluview/Novel_Influenza.html)). Had this virus caused a severe outbreak or pandemic, our proactive surveillance efforts and vaccine derivation would have provided an approximate 8-week time advantage for vaccine manufacturing.

By returning samples from the field, we were able to confirm the sequencing results with our Illumina sequencing pipeline (10). Without considering a coverage threshold for the Nanopore data, we sequenced 13 full genomes in the field. These same 13 samples were also successfully sequenced in the lab on an Illumina MiSeq platform. By comparing MiSeq- and MinION-produced sequences, we saw that the maximum MinION-derived sequence accuracy is achieved at 10 $\times$  to 20 $\times$  mean coverage (Fig. S32 [<https://figshare.com/s/b4cc885050283a40dfcd>]). If we apply a 20 $\times$  coverage threshold to the Nanopore data, only 4 genomes were successful in the field but with 8 HA



segment sequences passing. Importantly, the HA sequence is the most critical for making vaccine strain determinations. Also, we have higher confidence in consensus sequences from low-coverage viruses that are near identical to other viruses with higher coverage. As we are likely sampling the same transmission chain in real time, we can use the global sequence alignment to manually estimate the true sequence, particularly in fixing the frameshift mutations that are MinION's most common error. In future experiments, we can ensure that coverage thresholds are met by monitoring coverage in real time and more deliberately deciding when to stop sequencing, rather than simply sequencing for a defined 6 h as we did here. Illumina sequencing also allowed us to confirm the accuracy of our in-field-generated consensus sequences (99.3% overall). This level of accuracy confirms the validity of the suite of analyses that we performed in the field with the Nanopore data.

Influenza A viruses (IAVs) are a perpetual threat to global health security, both as human-endemic seasonal viruses and the more insidious pandemic viruses. IAV pandemics result from zoonotic transmission of IAVs to humans followed by onward human-to-human spread in populations lacking sufficient immunity. The swine-human and avian-human interfaces are of keen interest for future in-field surveillance efforts due to the pandemic risk of avian and swine influenza viruses that circulate in these settings (18, 19). *Mia* will be a useful tool to enhance the current centralized surveillance framework and can be deployed into resource-poor settings and operated by two technicians. This technology can serve to bolster IAV surveillance by monitoring important transmission interfaces for emerging viruses that have pandemic potential. Moreover, it might not always be possible to return samples from the field. In such cases, portable Nanopore sequencing and data transmission can take advantage of a distributed sequencing network while maintaining the advantages of a centralized database. This has the potential to greatly expand the reach of the current global influenza virus surveillance network.

## MATERIALS AND METHODS

**Logistics.** Our inventory was finalized after three full practice runs and consisted of three main pieces of luggage to transport *Mia* (Table S1 and Fig. S43 [<https://figshare.com/s/b4cc885050283a40dfcd>]). We used a Pelican Air 1615 case (internal dimensions, 75.16 cm length by 39.37 cm width by 23.82 cm depth; Torrance, CA, USA) to transport plastic consumables, ambient temperature liquids, personal protective equipment (PPE), racks, pipettes, and power strips. This is a larger suitcase that was checked during air travel. We used a smaller Pelican 1510 case (internal dimensions, 50.3 cm length by 27.94 cm width by 19.3 cm depth) to transport the electronic equipment, including the MinION systems, thermocyclers, vortex, microcentrifuge, and Qubit fluorometer. While this case is extremely rugged and would likely fare well against the checked baggage handling, we were concerned about damage during luggage inspection and opted to carry this onto the plane. We also carried on the laptop inside a backpack. Our cold chain was maintained during transportation in an Ozark Trail 26-quart certified bear-resistant cooler (internal dimensions, 40.64 cm length by 22.86 cm width by 27.3 cm depth; Potosi, MO, USA). To simultaneously transport materials at  $-20^{\circ}\text{C}$  and  $4^{\circ}\text{C}$ , the cooler was partitioned with a 7.6-cm piece of Styrofoam cut to a snug fit. All enzymes and reagents requiring storage at  $-20^{\circ}\text{C}$  were stored in a cold block and surrounded with frozen cold packs in the bottom of the cooler. Primers and flow cells were stored at  $4^{\circ}\text{C}$  on top of the Styrofoam and were surrounded by refrigerated cold packs. We maintained the cold chain on-site by storing reagents in a hotel refrigerator/freezer before repacking the cooler to transport the reagents to the exhibition barn. The SuperScript IV reverse transcriptase was shipped to the barn on dry ice at the insistence of the manufacturer. Since ethanol at a concentration over 70% cannot be transported on a commercial airliner, it was shipped via ground transportation. For a workbench, we used a small folding table found at the event and purchased plastic tablecloths, Lysol, and extra garbage bags locally. To avoid disrupting the event, we set up *Mia* inside a horse stall near the pigs and worked overnight (Fig. S44 [<https://figshare.com/s/b4cc885050283a40dfcd>]).

**Diagnostic screening and sampling.** To confirm the presence of IAV and to identify the location of outbreak clusters, we screened 94 swine displaying ILI using the Flu Detect influenza A virus swine rapid tests (Bridgewater, NJ, USA) via nasal swabs. We performed the screening early in the morning to locate IAV-positive swine and returned to the same location for sampling and sequencing later that evening.

We selected 24 swine for respiratory sample collection and sequencing, which included the seven positives from the screen and their immediate neighbors. We sampled these swine via sterile gauze nasal wipes and submersion in 5 ml brain heart infusion (BHI) medium (20, 21).

**RNA extraction.** Immediately after sampling, we extracted RNA from the BHI samples using the Akonni TruTip rapid RNA kit (Frederick, MD, USA), according to the manufacturer's instructions, with minor modifications (8). We extracted the samples in a 96-well deep block using a manual 12-channel 1,000- $\mu\text{l}$  pipette. Before lysis, we diluted the 70- $\mu\text{l}$  samples with 180  $\mu\text{l}$  of water and spiked them with

0.5  $\mu\text{g}$  of Qiagen carrier RNA (Venlo, the Netherlands). We eluted the samples in 70  $\mu\text{l}$  of water that had been warmed to 75°C.

**Amplification, barcoding, library preparation, and sequencing.** We amplified the full influenza A virus genome using a fast multisegment reverse transcription PCR (fMRT-PCR). For cDNA synthesis, we used a 5  $\mu\text{M}$  forward primer mix that is MBTuni-12 and MBTuni-12.4 combined in a 1:4 ratio to increase the affinity for the polymerase genes. The primer sequences are available in Table S07 (<https://figshare.com/s/b4cc885050283a40dfcd>). For primer annealing, we combined 10  $\mu\text{l}$  of RNA, 1  $\mu\text{l}$  forward primer mix, 1  $\mu\text{l}$  of 10 mM dinucleoside triphosphates (dNTPs), and 6  $\mu\text{l}$  of nuclease-free water. We heated the mixture to 65°C for 5 min and then cooled the mixture on an ice block. For cDNA synthesis, we added 4  $\mu\text{l}$  of 5 $\times$  SuperScript IV (SSIV) buffer, 1  $\mu\text{l}$  100 mM dithiothreitol (DTT), 1  $\mu\text{l}$  RNaseOUT RNase inhibitor (40 U/ $\mu\text{l}$ ), and 1  $\mu\text{l}$  of SuperScript IV reverse transcriptase to the annealed RNA. We incubated the reaction mixture at 42°C for 10 min, 53°C for 5 min, and 80°C for 10 min.

For full-genome amplification, we used a 10  $\mu\text{M}$  mixture of MBTuni-12, MBTuni-12.4, and MBTuni-13 primers in a 2:3:5 ratio. We combined 5  $\mu\text{l}$  cDNA, 5  $\mu\text{l}$  primer mix, 25  $\mu\text{l}$  2 $\times$  Q5 polymerase mix (New England Biosciences), and 15  $\mu\text{l}$  nuclease-free water. We amplified the full genome by cycling the mixture as follows: 98°C for 30 s; 5 cycles of 98°C for 10 s, 45°C for 15 s, and 72°C for 1.5 min; 30 cycles of 98°C for 10 s, 64°C for 15 s, and 72°C for 1.5 min; and a final extension of 72°C for 5 min. We cleaned the amplicons with 0.5 $\times$  AMPure beads (Beckman Coulter) and washed them with 80% ethanol before resuspending them in 25  $\mu\text{l}$  nuclease-free water.

We used three miniPCR mini8 thermocyclers (Cambridge, MA, USA) under laptop control for elution buffer heating, fMRT-PCR, barcoding, and library preparation. Following barcoding, we quantified the products using a Qubit DNA broad-range/high-sensitivity (BR/HS) assay (Thermo Fisher Scientific, Waltham, MA, USA). The quantification data allowed us to normalize and pool the barcodes and served as our only quality control check during the molecular workflow. Once pooled, we prepared the amplicons for Nanopore sequencing using the SQK-LSK 108 library kit (Oxford Nanopore Technologies, Oxford, UK). This protocol required cooling to 20°C; however, the temperature in the barn was roughly 30°C, and the fan-cooled mini8 thermocyclers could not get below ambient temperature. To solve this issue, we placed the thermocycler on a bed of ice, which effectively lowered the ambient temperature and allowed the mini8 to cool to 20°C. We sequenced the pooled samples on a MinION Mk 1B Nanopore sequencer using the flow cell FAH58363 equipped with R9.5.1 chemistry.

**Real-time analyses.** For real-time analyses, we used a high-performance Dell Precision 7720 laptop (Round Rock, TX, USA) with 64 Gb of ram, four-core Intel Xeon 3.00 GHz central processing unit (CPU) E3-1505M v6 (2 threads/core), and two 1-Tb solid state hard drives, partitioned separately into a Windows 10 OS and Ubuntu 16.04 OS. *Mia*'s custom analytical application runs on the Ubuntu partition and initiates automatically via cron on MiniON read file (fast5) origination during active sequencing. We base called and demultiplexed fast5 read files with Albacore v2.2.7 (Oxford Nanopore Technologies, Oxford, UK) and assembled subsequent fastq files into consensus sequences with IRMA v0.6.7 (10). We used a custom MinION configuration module for IRMA that changes the default "FLU" module parameters as follows: dropping the median read Q score filter from 30 to 7, raising the minimum read length from 125 to 150 bases, raising the frequency threshold for insertion and deletion refinement from 0.25 to 0.75 and 0.6 to 0.75, respectively, and lowering the Smith-Waterman mismatch penalty from 5 to 3 and the gap open penalty from 10 to 6. We checked consensus sequences against a local database containing clade-annotated IAV sequences using BLASTn v2.7.1+ (22) and used top BLAST matches to interpret a sample's IAV genome constellation. We aligned consensus sequences against reference sequences with MUSCLE v3.8.31 (23). We built phylogenetic trees with FastTree double precision version 2.1.8 (24) with a generalized time-reversible model, branch lengths optimized under a gamma distribution, and local support values produced from bootstrap sampling 10,000 times and visualized the trees in R v3.4.4 with GTree (25). We determined amino acid differences to each candidate vaccine virus sequence in our database with a custom Python script that translates DNA consensus sequences and MUSCLE aligns the mature HA1 protein sequences. Analytical and processing status results were fed into a SQLite v3.11.0 database (Hwaci, Inc., Charlotte, NC) and visualized in a Web browser with an interactive R Shiny (26) application.

*Mia*-processed reads in 4,000-read batches for the first 5 iterations, followed by 20,000-read batches for another 5 iterations and then 40,000-read batches until reads are no longer produced by the sequencer. This stepwise batching provides the user an understanding of the run's quality within minutes of sequencing, allowing the user to decide if it is worth continuing the run while maximizing computational resources. Results seen within the first hour of a sequencing track across a 6- or even 48-h run, with the extended run time increasing coverage and resolving a scant number of consensus bases.

**Postfield analyses.** We transported field-derived RNA and amplicons on dry ice back to the CDC in Atlanta, GA, and processed them with the Influenza Genomics Team's standard Illumina MiSeq influenza surveillance pipeline (San Diego, CA, USA). We processed and assembled reads with IRMA's Flu-sensitive module (10) and calculated the identity of MinION-derived consensus sequences based on alignment to the sample's MiSeq-derived consensus, which is considered the gold standard. Out of 425 swine sampled through our standard surveillance efforts at the exhibition, 136 swine tested positive for IAV by RT-PCR. Complete genome sequencing via our standard Illumina pipeline was attempted on 79 of those and produced 65 full genomes.

We performed phylogenetic analyses with the MiSeq-produced sequences and an annotated set of swine IAV sequences. The annotated data set was constructed by downloading all publicly available swine IAV data collected in North America from the Influenza Research Database (27) on 1 June 2018. Each gene was aligned with MAFFT v7.294b (28), and the best-known maximum likelihood phylogeny for

each alignment was inferred using IQ-TREE v1.6.2 (29), with the model of molecular evolution automatically selected during the tree search. Each gene was classified to an evolutionary lineage and/or phylogenetic clade (12, 14, 17). Subsequently, we selected 3 to 10 viruses from each named clade that captured the evolutionary relationships among clades and that represented the major HA, NA, and whole-genome constellations circulating in U.S. swine (12, 17). We then aligned the MiSeq and annotated sequences with MUSCLE v3.8.31 (23) and inferred maximum likelihood trees using IQ-TREE v1.6.8 (29), implementing a generalized time-reversible model of nucleotide substitution with gamma-distributed rate heterogeneity across sites while allowing for a proportion of invariable sites and 1,000 ultrafast bootstrap approximations (30) (iqtree -m GTR+G+I -bb 1000).

**Ethics approval.** This study was approved under IACUC animal use protocol number 2009A0134-R3.

**Data availability.** The raw MinION fast5 read data and MiSeq-derived consensus sequences are deposited at the NCBI under BioProject number [PRJNA528211](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA528211). Reference sequences reside in various public databases under the accession numbers listed in Table S6 (<https://figshare.com/s/b4cc88505283a40dfcd>). The pipeline code is available at [https://github.com/CDCgov/Mia\\_publication](https://github.com/CDCgov/Mia_publication).

## ACKNOWLEDGMENTS

The research reported in this publication was supported by the office of Advanced Molecular Detection (AMD grant CAN 939018 C) at the Centers for Disease Control and Prevention. The Council of State and Territorial Epidemiologists supported this project with funds provided by the Centers for Disease Control and Prevention under cooperative agreement no. 5U38OT000143-05. Nasal wipes were collected as part of influenza A virus surveillance supported by the Centers of Excellence for Influenza Research and Surveillance, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Department of Health and Human Services contract HHSN272201400006C.

We thank bluegrass legend Mindy Barringer from the CDC's Division of Communication Services for the Fig. 2 visual display.

The mention of trade names or commercial products in this article is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the USDA.

ONT paid for travel for Matthew W. Keller to present at the Nanopore community meeting in San Francisco, CA, in November 2018. ONT paid Matthew W. Keller \$1,000 to travel and present this research at ASM NGS in Tysons, VA, in September 2018.

## REFERENCES

- Quick J, Loman NJ, Duraffour S, Simpson JT, Severi E, Cowley L, Bore JA, Koundouno R, Dudas G, Mikhail A, Ouédraogo N, Afrough B, Bah A, Baum JH, Becker-Ziaja B, Boettcher J-P, Cabeza-Cabrero M, Camino-Sanchez A, Carter LL, Doerrbecker J, Enkirch T, Dorival IGG, Hetzelt N, Hinzmann J, Holm T, Kafetzopoulou LE, Koropogui M, Kosgey A, Kuisma E, Logue CH, Mazzarelli A, Meisel S, Mertens M, Michel J, Ngabo D, Nitzsche K, Pallash E, Patrono LV, Portmann J, Repits JG, Rickett NY, Sachse A, Singethan K, Vitoriano I, Yemanaberhan RL, Zekeng EG, Trina R, Bello A, Sall AA, Faye O, et al. 2016. Real-time, portable genome sequencing for Ebola surveillance. *Nature* 530:228–232. <https://doi.org/10.1038/nature16996>.
- Faria NR, Sabino EC, Nunes MRT, Alcantara LCJ, Loman NJ, Pybus OG. 2016. Mobile real-time surveillance of Zika virus in Brazil. *Genome Med* 8:97. <https://doi.org/10.1186/s13073-016-0356-2>.
- Kafetzopoulou LE, Pullan ST, Lemey P, Suchard MA, Ehichioya DU, Pahlmann M, Thielebein A, Hinzmann J, Oestereich L, Wozniak DM, Efthymiadis K, Schachten D, Koenig F, Matjeschek J, Lorenzen S, Lumley S, Ighodalo Y, Adomeh DI, Olorok T, Omomoh E, Omiunu R, Agbukor J, Ebo B, Aiyepada J, Ehbodaghe P, Osiemi B, Ehikhametalor S, Akhilomen P, Airende M, Esumeh R, Mueobonam E, Giwa R, Ekanem A, Igenegbale G, Odigie G, Okonofua G, Enigbe R, Oyakhilome J, Yerumoh EO, Odi A, Aire C, Okonofua M, Atafo R, Tobin E, Asogun D, Akpede N, Okokhere PO, Rafiu MO, Iraoyah KO, Iruolagbe CO, et al. 2019. Metagenomic sequencing at the epicenter of the Nigeria 2018 Lassa fever outbreak. *Science* 363:74–77. <https://doi.org/10.1126/science.aau9343>.
- Garten RJ, Davis CT, Russell CA, Shu B, Lindstrom S, Balish A, Sessions WM, Xu X, Skepner E, Deyde V, Okomo-Adhiambo M, Gubareva L, Barnes J, Smith CB, Emery SL, Hillman MJ, Rivaviller P, Smagala J, de Graaf M, Burke DF, Fouchier RAM, Pappas C, Alpuche-Aranda CM, López-Gatell H, Olivera H, López I, Myers CA, Faix D, Blair PJ, Yu C, Keene KM, Dotson PD, Boxrud D, Sambol AR, Abid SH, St. George K, Bannerman T, Moore AL, Stringer DJ, Blevins P, Demmler-Harrison GJ, Ginsberg M, Kriner P, Waterman S, Smole S, Guevara HF, Belongia EA, Clark PA, Beatrice ST, Donis R, et al. 2009. Antigenic and genetic characteristics of swine-origin 2009 A(H1N1) influenza viruses circulating in humans. *Science* 325:197–201. <https://doi.org/10.1126/science.1176225>.
- Bowman AS, Walia RR, Nolting JM, Vincent AL, Killian ML, Zentkovich MM, Lorbach JN, Lauterbach SE, Anderson TK, Davis CT, Zanders N, Jones J, Jang Y, Lynch B, Rodriguez MR, Blanton L, Lindstrom SE, Wentworth DE, Schiltz J, Averill JJ, Forshey T. 2017. Influenza A(H3N2) virus in swine at agricultural fairs and transmission to humans, Michigan and Ohio, USA, 2016. *Emerg Infect Dis* 23:1551–1555. <https://doi.org/10.3201/eid2309.170847>.
- World Health Organization. 2018. Antigenic and genetic characteristics of zoonotic influenza viruses and development of candidate vaccine viruses for pandemic preparedness. World Health Organization, Geneva, Switzerland. [https://www.who.int/influenza/vaccines/virus/201809\\_zoonotic\\_vaccinevirusupdate.pdf](https://www.who.int/influenza/vaccines/virus/201809_zoonotic_vaccinevirusupdate.pdf).
- Gardy JL, Loman NJ. 2018. Towards a genomics-informed, real-time, global pathogen surveillance system. *Nat Rev Genet* 19:9–20. <https://doi.org/10.1038/nrg.2017.88>.
- Chandler DP, Griesemer SB, Cooney CG, Holmberg R, Thakore N, Mokhiber B, Belgrader P, Knickerbocker C, Schied J, St. George K. 2012. Rapid, simple influenza RNA extraction from nasopharyngeal samples. *J Virol Methods* 183:8–13. <https://doi.org/10.1016/j.jviromet.2012.03.002>.
- Keller MW, Rambo-Martin BL, Wilson MM, Ridenour CA, Shepard SS, Stark TJ, Neuhaus EB, Dugan VG, Wentworth DE, Barnes JR. 2018. Direct RNA sequencing of the coding complete influenza A virus genome. *Sci Rep* 8:14408. <https://doi.org/10.1038/s41598-018-32615-8>.
- Shepard SS, Meno S, Bahl J, Wilson MM, Barnes J, Neuhaus E. 2016. Viral

- deep sequencing needs an adaptive approach: IRMA, the iterative refinement meta-assembler. *BMC Genomics* 17:708. <https://doi.org/10.1186/s12864-016-3030-6>.
11. Rajao DS, Anderson TK, Kitikoon P, Stratton J, Lewis NS, Vincent AL. 2018. Antigenic and genetic evolution of contemporary swine H1 influenza viruses in the United States. *Virology* 518:45–54. <https://doi.org/10.1016/j.virol.2018.02.006>.
  12. Gao S, Anderson TK, Walia RR, Dorman KS, Janas-Martindale A, Vincent AL. 2017. The genomic evolution of H1 influenza A viruses from swine detected in the United States between 2009 and 2016. *J Gen Virol* 98:2001–2010. <https://doi.org/10.1099/jgv.0.000885>.
  13. Kitikoon P, Nelson MI, Killian ML, Anderson TK, Koster L, Culhane MR, Vincent AL. 2013. Genotype patterns of contemporary reassorted H3N2 virus in US swine. *J Gen Virol* 94:1236–1241. <https://doi.org/10.1099/vir.0.51839-0>.
  14. Anderson TK, Macken CA, Lewis NS, Scheuermann RH, Van Reeth K, Brown IH, Swenson SL, Simon G, Saito T, Berhane Y, Ciacci-Zanella J, Pereda A, Davis CT, Donis RO, Webby RJ, Vincent AL. 2016. A phylogeny-based global nomenclature system and automated annotation tool for H1 hemagglutinin genes from swine influenza A viruses. *mSphere* 1:e00275-16. <https://doi.org/10.1128/mSphere.00275-16>.
  15. Centers for Disease Control and Prevention (CDC). 2009. Update: novel influenza A (H1N1) virus infections-worldwide, May 6, 2009. *MMWR Morb Mortal Wkly Rep* 58:453–458.
  16. Igarashi M, Ito K, Yoshida R, Tomabechei D, Kida H, Takada A. 2010. Predicting the antigenic structure of the pandemic (H1N1) 2009 influenza virus hemagglutinin. *PLoS One* 5:e8553. <https://doi.org/10.1371/journal.pone.0008553>.
  17. Walia RR, Anderson TK, Vincent AL. 2019. Regional patterns of genetic diversity in swine influenza A viruses in the United States from 2010 to 2016. *Influenza Other Respir Viruses* 13:262–273. <https://doi.org/10.1111/irv.12559>.
  18. Peiris JSM, de Jong MD, Guan Y. 2007. Avian influenza virus (H5N1): a threat to human health. *Clin Microbiol Rev* 20:243–267. <https://doi.org/10.1128/CMR.00037-06>.
  19. Cox NJ, Trock SC, Burke SA. 2014. Pandemic preparedness and the influenza risk assessment tool (IRAT), p 119–136. *In Compans RW, Oldstone MBA (ed), Influenza pathogenesis and control - volume I*. Springer International Publishing, Cham, Switzerland.
  20. Edwards JL, Nelson SW, Workman JD, Slemmons RD, Szablewski CM, Nolting JM, Bowman AS. 2014. Utility of snout wipe samples for influenza A virus surveillance in exhibition swine populations. *Influenza Other Respir Viruses* 8:574–579. <https://doi.org/10.1111/irv.12270>.
  21. Nolting JM, Szablewski CM, Edwards JL, Nelson SW, Bowman AS. 2015. Nasal wipes for influenza A virus detection and isolation from swine. *J Vis Exp* 2015:53313. <https://doi.org/10.3791/53313>.
  22. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden T. 2009. BLAST+: architecture and applications. *BMC Bioinformatics* 10:421. <https://doi.org/10.1186/1471-2105-10-421>.
  23. Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32:1792–1797. <https://doi.org/10.1093/nar/gkh340>.
  24. Price MN, Dehal PS, Arkin AP. 2009. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. *Mol Biol Evol* 26:1641–1650. <https://doi.org/10.1093/molbev/msp077>.
  25. Yu G, Smith DK, Zhu H, Guan Y, Lam TT-Y. 2017. ggtree: an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods Ecol Evol* 8:28–36. <https://doi.org/10.1111/2041-210X.12628>.
  26. Chang W, Cheng J, Allaire J, Xie Y, McPherson J. 2018. shiny: Web application framework for R. R package version 1.1.0. <https://cran.r-project.org/web/packages/shiny/index.html>.
  27. Zhang Y, Aevermann BD, Anderson TK, Burke DF, Dauphin G, Gu Z, He S, Kumar S, Larsen CN, Lee AJ, Li X, Macken C, Mahaffey C, Pickett BE, Reardon B, Smith T, Stewart L, Suloway C, Sun G, Tong L, Vincent AL, Walters B, Zaremba S, Zhao H, Zhou L, Zmasek C, Klem EB, Scheuermann RH. 2017. Influenza Research Database: an integrated bioinformatics resource for influenza virus research. *Nucleic Acids Res* 45:D466–D474. <https://doi.org/10.1093/nar/gkw857>.
  28. Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* 30:772–780. <https://doi.org/10.1093/molbev/mst010>.
  29. Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* 32:268–274. <https://doi.org/10.1093/molbev/msu300>.
  30. Hoang DT, Vinh LS, Chernomor O, Minh BQ, von Haeseler A. 2018. UFBoot2: improving the ultrafast bootstrap approximation. *Mol Biol Evol* 35:518–522. <https://doi.org/10.1093/molbev/msx281>.
  31. Taubenberger JK, Kash JC. 2010. Influenza virus evolution, host adaptation, and pandemic formation. *Cell Host Microbe* 7:440–451. <https://doi.org/10.1016/j.chom.2010.05.009>.