Nanomicroarray and Multiplex Next-Generation Sequencing for Simultaneous Identification and Characterization of Influenza Viruses

Jiangqin Zhao, Viswanath Ragupathy, Jikun Liu, Xue Wang, Sai Vikram Vemula, Haja Sittana El Mubarak, Zhiping Ye, Marie L. Landry, Indira Hewlett

Conventional methods for detection and discrimination of influenza viruses are time consuming and labor intensive. We developed a diagnostic platform for simultaneous identification and characterization of influenza viruses that uses a combination of nanomicroarray for screening and multiplex next-generation sequencing (NGS) assays for laboratory confirmation. The nanomicroarray was developed to target hemagglutinin, neuraminidase, and matrix genes to identify influenza A and B viruses. PCR amplicons synthesized by using an adapted universal primer for all 8 gene segments of 9 influenza A subtypes were detected in the nanomicroarray and confirmed by the NGS assays. This platform can simultaneously detect and differentiate multiple influenza A subtypes in a single sample. Use of these methods as part of a new diagnostic algorithm for detection and confirmation of influenza infections may provide ongoing public health benefits by assisting with future epidemiologic studies and improving preparedness for potential influenza pandemics.

Influenza A virus consists of 8 negative, single-stranded RNA segments encoding 11 proteins: polymerase basic 1 and 2 (PB1 and PB2); polymerase acidic (PA); hemagglutinin (HA); nucleoprotein (NP); neuraminidase (NA); matrix (M1/2); and nonstructural (NS1/2). Influenza A viruses are classified into 18 HA subtypes (H1–H18) and 11 NA subtypes (N1–N11), determined on the basis of the antigenic differences in the surface glycoproteins HA and NA (*1–4*). All known HA subtypes of influenza A virus are found in aquatic birds, and some, including H1, H2, H3, H5, H7, and H9, have been reported to infect humans (*1*,*5–7*). Direct transmission of avian influenza A virus subtypes H5N1, H7N2, H7N3, H7N7, H9N2, and H10N7 from domestic poultry to humans has been reported (*8–13*).

Author affiliations: Food and Drug Administration, Silver Spring, Maryland, USA (J. Zhao, V. Ragupathy, J. Liu, X. Wang, S.V. Vemula, H.S. El Mubarak, Z. Ye, I. Hewlett); Yale University School of Medicine, New Haven, Connecticut, USA (M.L. Landry)

DOI: http://dx.doi.org/10.3201/eid2103.141169

In early 2009, a novel swine-origin virus, designated influenza A(H1N1)pdm09 (pH1N1), emerged in Mexico and spread rapidly around the world, causing a global influenza pandemic (*14*,*15*). This virus was generated by multiple reassortment events over 10 years (*16*,*17*) and continued to circulate in humans after the initial pandemic period, replacing the previously circulating seasonal H1N1 viruses. Influenza A(H3N2) variant virus (H3N2v) isolated from humans in the United States in 2011 was also generated through reassortment originating from swine, avian, and human viruses, including the M gene from pH1N1 virus (*18*,*19*). More recently, a novel avian-origin influenza A(H7N9) virus capable of poultry-to-human transmission was identified in China (*7;* http://www.who.int/influenza/human_animal_interface/influenza_h7n9/140225_H7N9RA_for_web_20140306FM.pdf). Diagnosis of infection with this virus is difficult because infection does not kill infected poultry, but the virus may post a substantial risk for a human pandemic because of a lack of immunity in the general population (*7*). As these viruses demonstrate, reassortment of pH1N1 virus with other circulating seasonal strains can produce virulent variants that can be transmitted to and among humans and that could emerge as a future pandemic strain (*15*,*20*,*21*). Therefore, it is critical to determine whether transmitted viruses have pandemic potential in humans during the influenza season.

Multiple influenza strains are usually prevalent during an influenza season. Increasing global travel results in rapid spread of novel influenza viruses from one geographic region to another (*13*,*22*). Current approaches for screening and characterizing novel influenza viruses require many steps and multiple assays. A single test has not been available for simultaneous identification of newly emerging strains from known or unknown subtypes of influenza viruses and the characterization of unique virulence factors or putative antiviral resistance markers.

We previously described a method for detection of avian influenza A(H5N1) and swine-origin pH1N1 viruses that used a nanotechnology-based, PCR-free, whole-genome

microarray assay (nanomicroarray) (*23*,*24*). In this article, we describe a new diagnostic platform for identification and characterization of subtypes of influenza A virus that uses nanomicroarray for screening and multiplex next-generation sequencing (NGS) for laboratory confirmation. We demonstrate that this platform enables accurate and simultaneous identification of multiple subtypes in a single sample. We used this platform to evaluate clinical nasopharyngeal swab specimens from patients with influenza-like illness that had tested positive for influenza virus to determine influenza virus subtype.

Materials and Methods

Oligonucleotide Design and Nanomicroarray Assay

The sequences for multiple capture and intermediate oligonucleotides per target gene were designed and prepared as described previously (*23*,*24*). The oligonucleotide sequences and details of the nanomicroarray assays are listed and described in the online Technical Appendix (http://wwwnc.cdc.gov/EID/article/21/3/14-1169-Techapp1.pdf).

Viruses and Clinical Samples

Information about influenza viruses used in this study is provided in the online Technical Appendix. Nasopharyngeal swab specimens from patients with symptoms of influenza-like illness were submitted to the Clinical Virology Laboratory at Yale–New Haven Hospital, New Haven, Connecticut, USA, during December 27–December 31, 2012. Samples were tested by using direct fluorescent antigen (DFA) test with SimulFluor reagents (Millipore, Billerica, MA, USA) and, in some cases, by real-time reverse transcription PCR (rRT-PCR), as requested by the patients' physicians. PCR was performed by using the Centers for Disease Control and Prevention rRT-PCR protocol for influenza as previously described (*25*). Samples for which DFA, rRT-PCR, or both gave results positive for influenza A were selected, de-identified, and sent to the Laboratory of Molecular Virology at the Food and Drug Administration in Silver Spring, Maryland, USA, for further testing (Table 1).

Viral RNA Extraction and rRT-PCR

A previously reported universal primer designed to amplify all 8 gene segments (*26*,*27*) was modified by adding 13-bp flanking sequence (5′-ACGACGGGCGACA-3′) at the 5′ end of each primer to enhance the annealing temperature and achieve high fidelity and yield in PCR amplification. Additional details of RNA extraction and rRT-PCR conditions are described in the online Technical Appendix.

NGS Assay

The concentration of PCR amplicons of all 8 gene segments of influenza A virus was measured by using the Qubit dsDNA BR Assay System (Covaris, Woburn, MA, USA); 1 ng of DNA product was processed for NGS sample preparation by using a Nextera XT DNA Sample Preparation Kit (Illumina, San Diego, CA, USA), according to the

Table 1. Detection of influenza A viruses in nasopharyngeal swab samples collected from naturally infected patients, Connecticut, USA, 2012–13 influenza season*

	Patient age,	Sample collection		Detection methods			
Patient ID	y/sex	Date, 2012	Location	DFA	rRT-PCR, C_t	Universal PCR	NGS
FLU001	47/F	Dec 30	Hamden, CT	+	ND	+	H3N2
FLU002	80/M	Dec 30	Milford, CT	+	17.5	+	H3N2
FLU004	35/F	Dec 30	Meriden, CT	+	ND	+	H3N2
FLU006	25/M	Dec 29	New Haven, CT	+	ND	+	H3N2
FLU007	23/F	Dec 30	New Haven, CT	+	ND	+	H3N2
FLU008	31/F	Dec 30	Trumbull, CT	+	28.1	+	H3N2
FLU009	68/M	Dec 30	Hamden, CT	+	ND	+	H3N2
FLU012	35/F	Dec 30	Rutledge, MO	+	ND	+	H3N2
FLU013	92/M	Dec 29	Woodbridge, CT	+	16.6	+	H3N2
FLU014	84/F	Dec 30	Chester, CT	+	15.5	+	H3N2
FLU017	66/F	Dec 29	Clinton, CT	+	24.3	+	H3N2
FLU018	17/M	Dec 31	New Haven, CT	+	19.9	+	H3N2
FLU021	63/F	Dec 30	New Haven, CT	+	ND	+	H3N2
FLU023	55/F	Dec 28	North Haven, CT	+	21.8	+	H3N2
FLU025	47/M	Dec 30	West Haven, CT	+	ND	+	H3N2
FLU026	32/F	Dec 28	West Haven, CT	+	ND	+	H3N2
FLU027	26/F	Dec 30	Bridgeport, CT	+	ND	+	H3N2
FLU028	89/F	Dec 29	Woodbridge, CT	I	21.0	+	H3N2
FLU033	82/F	Dec 29	Guilford, CT	+	19.8	+	H3N2
FLU034	37/F	Dec 27	New Haven, CT	+	ND	+	H3N2
FLU036	18/M	Dec 29	New Haven, CT	I	17.8	+	H3N2
FLU037	8/F	Dec 29	New Haven, CT	+	ND	+	H3N2
FLU038	21/M	Dec 27	West Haven, CT	+	ND	+	H3N2
FLU040	23/M	Dec 27	West Haven, CT	+	ND	+	H3N2

*C_t, cycle threshold value; DFA, direct fluorescent antigen test; I, inadequate cells for DFA; ID, identification; ND, not done; NGS, next-generation sequencing; rRT-PCR, real-time RT-PCR.

manufacturer's instructions. Briefly, the Nextera XT transposome fragmented PCR amplicons into a size of ≈500–700 bp and added adaptor sequences to the ends, enabling a 12-cycle PCR amplification to append additional unique dual index (i7 and i5) sequences at the end of each fragmented DNA for cluster formation. Mega-amplicons from influenza virus were internally marked with these dual-barcoded primers, which enabled multiplexing and simultaneous detection of different subtypes in the same run. After purification of PCR fragments and library normalization, sample pooling was performed by mixing equal volumes of each normalized DNA library, and the barcoded multiplexed library sequencing was performed on an Illumina MiSeq (Illumina). After automated cluster generation, sequencing was processed and genomic sequence reads obtained.

Bioinformatics Analysis

Sequencing reads of ≈300 bp were dynamically trimmed and sequence data were verified by FastQC software (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) before de novo assembly. The genome-contiguous assembly was constructed from MiSeq reads by using a de novo module in CLC genomics workbench software version 6.0.2 (CLC bio, Cambridge, MA, USA); minimum contiguous length was set at 800 for assembling consensus sequences (*28*). A comprehensive-read database was generated for the whole genome of the influenza virus tested. Sequences were further filtered so that the local database contained only 1 unique contig for each gene segment, and multiple contigs were generated for each sample. These representative sequences comprise the set of unique sequences from the dataset. A FASTA file with all unique contiguous sequences of each mega-amplicon was used to perform an all-by-all Identify Similae Sequences search in the Influenza Research Database (IRD, http://www.fludb.org), the Global Initiative on Sharing All Influenza Data database (http://platform.gisaid.org), and the National Center for Biotechnology Information database (http://www.ncbi.nlm.nih.gov). The top-scoring BLAST (http://blast.ncbi.nlm.nih.gov/) match was selected to identify the specific genome. Assembled sequences were aligned in ClustalW (http://www.clustal.org), and phylogenetic analysis was performed in MEGA by using the neighbor-joining method (*29*). All amplicons were accurately categorized into a typical subtype.

Results

Verification of Capture and Intermediate Oligonucleotides

A nanomicroarray for each target gene was designed, printed in-house, and tested separately by using the PCR products as templates to verify the ability of individual capture and intermediate oligonucleotides to detect a specific target gene. Ineffective capture oligonucleotides were replaced and retested. We amplified PCR products of HA, NA, and M genes for H7N2, H7N3, and H9N2 viruses separately or simultaneously in a single reaction using the corresponding 3 sets of specific primers. To identify the correct HA and NA gene segments for multiplex influenza subtyping, we fabricated a new nanomicroarray by pooling autologous 4 to 5 capture oligonucleotides for a specific gene and then printing them on the array substrate in triplicates. Each nanomicroarray subarray contains multiple gene spots for multiplex assays. The PCR products were hybridized on the array, and the specific signal profiles were correctly observed in the areas printed with corresponding gene-specific capture oligonucleotides (online Technical Appendix Figure 1). No interference or cross-hybridization was observed when multiple targets and intermediate oligonucleotides were included in the assay. The specific signal pattern showed the assay's ability to accurately discriminate influenza subtypes.

Amplification of Whole-Genome Segments

To further confirm subtypes detected in the nanomicroarray screening assay for final laboratory diagnosis, we redesigned universal primers to amplify whole-genome segments and separately tested 22 influenza A strains covering 10 subtypes and 3 influenza B viruses. All 8 segments of influenza A viruses were simultaneously amplified in a single reaction, resulting in multiple PCR products ranging in size from 500 to 2,500 bp (mega-amplicons). We also tested 3 influenza B viruses, B/Brisbane/60/2008 (Victoria lineage), B/Pennsylvania/7/2007 (Yamagata lineage), and B/Victoria/304/2006 (Victoria lineage), and 2 influenza A viruses, A/Panama/2007/1999 (H1N1) and A/ruddy turnstone/NJ/65/1985 (H7N3), and found several faint, nonspecific small bands no larger than 1 kb (data not shown).

Evaluation of Nanomicroarray Assay by using PCR Mega-amplicons

To include M gene capture oligonucleotides for influenza B viruses, we developed a new nanomicroarray (Figure 1). As shown in Figure 2, amplified PCR products of the matrix gene from influenza A and B viruses were specifically detected in the correct spot areas without cross-hybridization. The mega-amplicons of influenza A viruses were correctly identified and found to have a unique fingerprint for each influenza A virus tested. Each spot pattern represented a typical influenza virus subtype corresponding to the gene-specific capture oligonucleotides. We conclude that the current nanomicroarray assay can simultaneously discriminate influenza A from B viruses and specifically identify influenza A subtypes H1, H2, H3, H5, H7, H9, N1, N2, and N3.

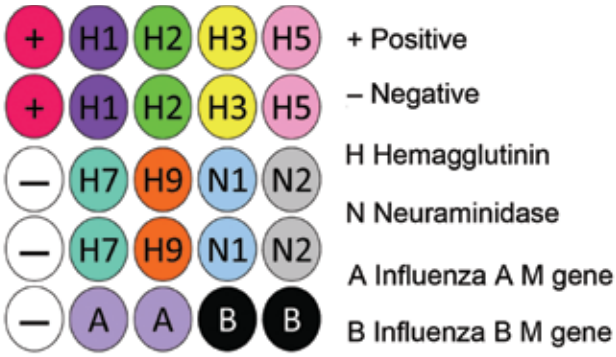


Figure 1. Nanomicroarray layout design for testing of samples for influenza A and B viruses. The microarray internal positive control capture is listed in online Technical Appendix Table 1 (http://wwwnc.cdc.gov/EID/article/21/3/14-1169-Techapp1.pdf). The negative control is the printing buffer. M, matrix protein.

NGS Confirmation of Influenza A Subtypes

A total of 17 mega-amplicons representing 10 subtypes of influenza A and 1 of influenza B were tested in the NGS assay. Multiple contiguous sequences were created automatically for each mega-amplicon by using a de novo assembly program in CLC, and 4–8 contigs supported by high coverage rate of sequence reads were generated (online Technical Appendix Table 2). The mega-amplicons of the A/Vietnam/1204/2004 (H5N1) strain yielded 8 contiguous sequences supported by 90,962 reads. Further BLAST search of 8 contiguous sequences in the IRD resulted in 8 mast BLAST reports. All contiguous sequences were found to correspond to 6 proteins (PB2, PB1, PA, NP, M, and NS) of A/Puerto Rico/8/1934 (H1N1) virus and 2 proteins (HA/CIP045/CY077101 and NA/HM006761) of A/Vietnam/1203/2004 (H5N1) with 99%–100% sequence identity. These results showed that all amplicons were correctly identified as the H5N1 laboratory strain.

The mega-amplicons from A/turkey/Virginia/4529/2002 (H7N2) and A/Minnesota/10/2012 (H3N2) strains resulted in 8 contigs, all correctly identified as the correct subtype. We found strong concordance in contiguous sequences and PCR fragments for each mega-amplicon. A de novo assembly program generated at least 7 contigs from faint band mega-amplicons for influenza B virus (B/Brisbane/60/2008); 2 showed good coverage (7,077 and 10,168), but the BLAST search indicated that none matched the gene sequence from this strain. Further investigation using freshly extracted RNA may be required.

Simultaneous NGS Discrimination of Multiple Subtypes in a Single Sample

We tested 4 influenza viruses obtained from the Centers for Disease Control and Prevention, A/Puerto Rico/8/1934 (H1N1), A/Vietnam/1203/2004 (H5N1), A/Minnesota/10/2012 (H3N2), and A/Anhui/1/2013 (H7N9), to determine the presence of H1, H3, H5, H7, N1, N2, and N9 subtypes. RNA was extracted from individual or mixed viral strains. The universal rRT-PCR was performed to amplify whole-genome segments in which the PCR mega-amplicons represented a similar pattern to individual or mixed viral samples (online Technical Appendix Figure 2). After the NGS assay followed by de novo assembly, 8 contigs were generated for the H5N1 and H3N2 viruses, and 15 contigs were generated for the mixed sample, supported by a high coverage rate of reads (Table 2). These 15 contigs from mixed samples exactly matched the gene segments of the

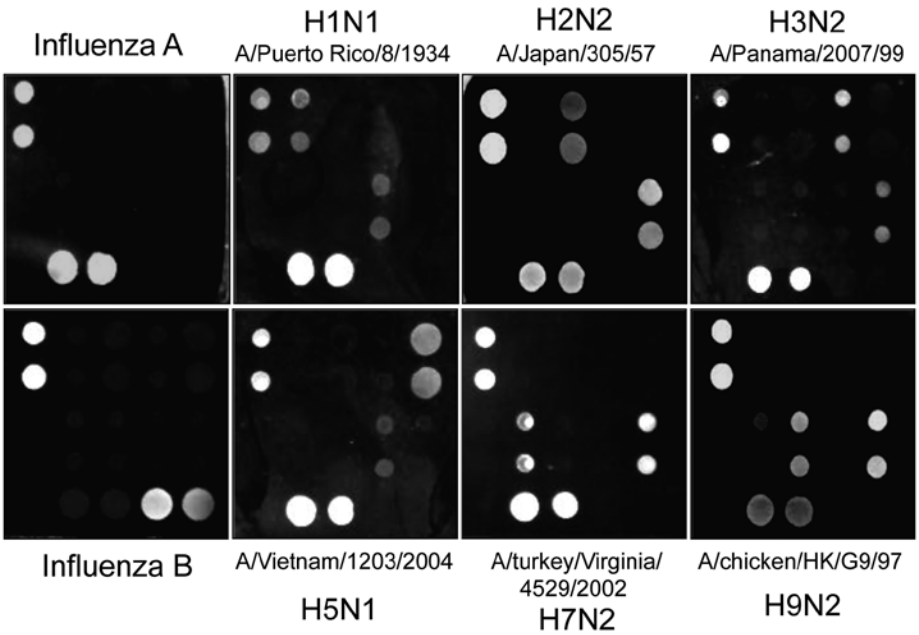


Figure 2. Portion of the microarray images for DNA oligonucleotides of influenza viruses after hybridization with PCR products. Lighter shades represent greater silver intensities for each gene. Typical nanomicroarray silver staining images represent the hits for specific types or subtypes indicated. The positive controls of influenza A and B (left panels) use PCR products amplified by pair-specific primes for matrix gene.

H5N1 and H3N2 strains, similar to the results obtained with the individual sample, as expected. HA and NA genomic sequences of the H7N9 subtype virus were identified as strain A/Anhui/1/2013 from the Global Initiative on Sharing All Influenza Data database. Because PB2 (2,341 bp), PB1 (2,341 bp), and PA (2,233 bp) genes have very similar sizes, direct separation of each gene from co-infected samples is not possible by conventional sequencing methods. These findings demonstrated that the NGS assay can simultaneously identify and confirm the presence of ≥1 influenza subtypes in a single sample.

Evaluation of Nasopharyngeal Swab Samples by using NGS Assays

We performed universal RT-PCR and NGS assays on 24 nasopharyngeal swab samples obtained from patients who had received a diagnosis of influenza. These samples were initially tested by DFA, rRT-PCR, or both and found to be positive for influenza A virus. After decoding, all samples were found to be positive by using the universal RT-PCR detection method, indicating presence of influenza A infection (Table 1). When mega-amplicons representing the 24 patient samples were tested in the NGS, a total of 32.8 million reads were obtained, and multiple contigs were generated for each sample (online Technical Appendix Table 3). A BLAST search of each contig in the IRD database identified the genome corresponding to the influenza A(H3N2) subtype. The coverage of influenza A(H3N2) genomes in the NGS assay was 96.7% (31.7/32.8 million) of raw reads and 76.6% (183/239) of total contigs. A total of 95.3% (183/192 contigs) of the influenza A(H3N2) genome was amplified and sequenced; the average depth of coverage for each contig was 3,259. Of these genomes, 71% (136/192) of segments yielded full-length sequences; HA genes were 96% (23/24); NP, 96% (23/24); NA, 88% (21/24); M, 88%, (21/24); and NS, 79% (19/24). The average breadth of coverage was 100% for HA, NA, NP, M, and NS genes and 93% for PB2, PB1, and PA genes.

Phylogenetic analysis of each of the 8 segments separately for all isolates showed that all genes clustered together in the H3N2 radiation with a high bootstrap value (data not shown). None of the M genes closely clustered with the M genes from the pH1N1 or H3N2v viruses (Figure 3), which suggests that these viruses are not H3N2v (*18*,*19*). The genotype of 24 influenza A(H3N2) viruses was determined as [A,D,B,3A,A,2A,B,1A] by using FluGenotyping (http://www.flugenome.org), which indicates that the same lineage virus is circulating in this region. The HA genes from most samples shared very high identity with A/Boston/DOA2–206/2013(H3N2) and the NA genes with A/Boston/DOA2-141/2013(H3N2) strain. After completing these analyses, 181 gene sequences were deposited into GenBank (accession nos. KJ741883–KJ742063).

Discussion

We report the development of a novel diagnostic platform for simultaneous detection, typing, and whole-genome characterization of influenza viruses that uses a combination of nanomicroarray and high-throughput NGS approaches. First, we designed capture and intermediate oligonucleotides for H1, H2, H3, H5, H7, H9, N1, N2, and N3 of influenza A virus and M genes of influenza B virus and evaluated these oligonucleotides in a nanomicroarray assay. Second, we modified previously reported universal primers (*26*,*27*) and used them to amplify the whole genome of influenza A viruses for validation of the nanomicroarray assay. Finally, we confirmed results by using the NGS assay. This protocol enables random accessing of a variety of target genes for simultaneous identification and final sequence-based confirmation of influenza virus infection.

Designing multiple capture and intermediate oligonucleotides with sequences covering the entire genome ensures specific capture of multiple target genes on the nanomicroarray and subsequent detection with a universal nanoparticle probe regardless of mutation, deletion, and

Table 2. Summary of results from NGS data analysis for influenza A(H3N2) and A(H5N1) viruses obtained from the Centers for Disease Control and Prevention*

			Gene segment (length, bp)							
Strain	NGS total contigs/reads	Findings	PB2 (2,341)	PB1 (2,341)	PA (2,233)	HA (1,778)	NP (1,565)	NA (1,413)	M (1,027)	NS (890)
A/Vietnam/1203/2004(H5N1)	8/125,438	Length, bp	2,032	1,801	1,841	1,747	1,120	1,259	1,075	894
		Read count	7,604	2,019	18,762	16,925	6,490	24,192	20,127	920
A/Minnesota/10/2012(H3N2)	8/162,155	Length, bp	2,251	2,192	2,081	1,754	1,556	1,639	1,171	876
		Read count	30,002	16,392	19,253	15,194	21,818	8,119	25,879	18,519
A/Vietnam/1203/2004(H5N1) and A/Minnesota/10/2012(H3N2)	15/150,756	Contigs, H5	2	1	13	14	4	6	7	
		Length, bp	2,013	2,115	1,815	1,746	1,476	1,193	936	
		Read count	7,003	6,857	4,359	3,735	854	10,895	15,127	
		Contigs, H3	15	5	10	11	3	12	8	9
		Length, bp	2,530	2,393	1,959	1,465	1,565	1,633	1,103	925
		Read count	15,521	6,753	4,113	7,743	13,808	11,147	5,411	6,110

*de novo assembly module was used in CLC Genomics Workbench software version 6.0.2 (CLC bio, Cambridge, MA, USA) for result handing. Minimum contiguous length was set for 800 to assemble the consensus sequences. NGS, next-generation sequencing; PB, polybasic; PA, polymerase; HA, hemagglutinin; NP, nucleoprotein; NA, neuraminidase; M, matrix; NS, nonstructural.

Figure 3. Phylogenetic analysis of the matrix (M) gene sequences obtained from nasopharyngeal swab samples from patients who had received a diagnosis of influenza in Connecticut, USA, during the 2012–13 influenza season (see Table 1). Analysis was performed by using the neighbor-joining module in MEGA (*29*) with the Kimura 2-parameter method. The reference subtypes were fetched from the Influenza Research Database (http://www.fludb.org) and used to construct the tree. Bootstrap values >70% are shown. The M genes identified in this study are indicated by black circles; reference M genes are indicated by black squares for influenza A(H3N2)v and black triangles for pandemic influenza A(H1N1) 2009 (pH1N1) virus. Scale bar indicates 2% genetic distance.

influenza reassortment. Furthermore, this design is adaptable for other applications and enables direct detection and subtyping of an unknown sample without previous knowledge of types and subtypes. In the current format, >50 degenerate capture oligonucleotides cover 12 influenza viral target genes, enabling direct detection of any combination of ≈20 subtypes in a single sample, identification of influenza A subtypes in a single assay, and differentiation of influenza A from B viruses. An optimal nanomicroarray assay, which is a reformatted portable device modified for use in point-of-care settings, should include target genes from most influenza A and B viruses as well as for other respiratory viral pathogens. The assay should be easily performed by an untrained technician for sample testing in the field without enzymatic reactions, and results should be in a form that can easily be visualized by the naked eye. In comparison to other conventional detection methods for targeting each gene of influenza A and B viruses separately, the nanomicroarray assay is a one-test-fits-all approach for diagnosis of influenza virus infections that can provide results in <1.5 hours, making this method relatively cost- and time-effective. More important, the nanomicroarray assay can detect emerging and reassortant viruses, and those samples can be sent to centralized laboratories that perform the NGS assay for final sequence confirmation.

Gene segments in most influenza viruses isolated from humans can be adapted from animals, as shown by genetic changes in influenza A(H7N9) isolates from poultry and humans (*7*,*30*). These studies indicated that more changes were acquired during the human infection process. Determining the nature and frequency of co-infection associated with influenza A virus will be critical if an unknown sample contains a novel strain or >1 HA or NA gene subtype. NGS is a powerful tool facilitating diagnosis on a large scale, including high-throughput and simultaneous identification of ≥96 samples barcoded by using dual index primers and detection of >9,216 genes in a single sequencing run. By using a universal primer adapted to fabricate the mega-amplicons, we showed that the NGS assay is capable of accurately subtyping any influenza A virus and detecting multiple known and unknown influenza genes in a single assay. Bioinformatic skills and mathematics tools, combined with epidemiologic studies, are useful in facilitating prediction of potential subtypes according to the genetic matrix composition of influenza genomic segments for a new, emerging, and reassorted strain whereby the subtypes can be confirmed (*31*–*34*). Influenza A viruses representing 11 subtypes were accurately detected in this study, and the 2 mixed influenza viruses were discriminated by using this sequencing-based diagnostic platform.

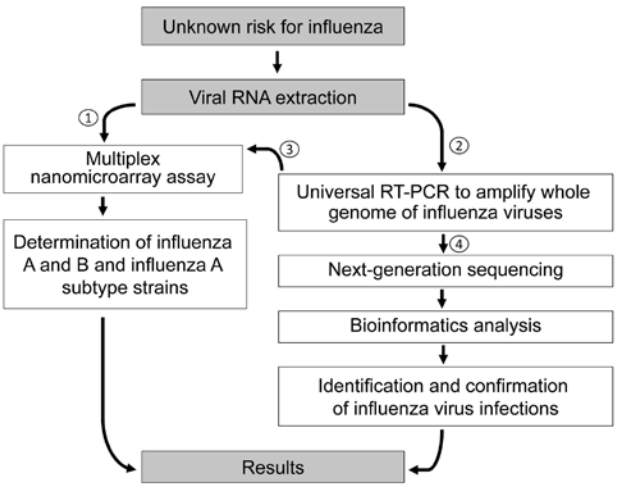


Figure 4. Diagnostic algorithm for identification of an unknown risk for influenza by using nanomicroarray and next-generation sequencing (NGS) assays. To determine the virus type for a suspected influenza virus infection, viral RNA is extracted from a patient sample and initially analyzed in nanomicroarray assay for screening and determining the influenza A and B viruses (*1*). Once a novel, emerging, or co-infected influenza A and B virus is found, universal reverse transcription PCR (RT-PCR) is performed to generate whole-genome mega-amplicons (*2*), which can then be retested on the nanomicroarray assay to confirm the initial finding (*3*) or sent to the central laboratory performing the NGS assay and data analysis for final sequence confirmation (*4*).

Sequence analysis of 24 clinical samples revealed that 23 (92%) segments contained an amino acid substitution at position E627K in PB2 gene. Mutation of glutamic acid (E) at PB2 residue 627 to lysine (K) favors adaptation to the mammalian host; such mutations have been found in human isolates of highly pathogenic avian influenza viruses of the H7N7 and H7N9 subtypes (*7*,*35*,*36*). These mutations might confer high virulence to the virus by enhancing replication efficiency, increasing polymerase activity and disease severity of avian influenza viruses in mammals (*37*).

Of the 24 M genes of the samples we tested, 21 (88%) had a single S31N mutation in the transmembrane region of the M2 protein, which has been found to confer resistance to amantadine (*7*,*38*). The emergence of E627K(PB2) and S31N(M2) mutations in tested samples suggests that human host infection in the Connecticut region in the 2012–2013 seasons might be poultry-to-human transmission associated with disease severity (*39*,*40*). This observation highlights an increased risk to public health and the need to continually monitor isolates obtained from mammal reservoirs for genetic variation. This information may help guide clinical treatment and assessment of epidemiology during the epidemic season.

The assay we evaluated is a minimally manipulated procedure that greatly reduces the number of amplifications and omits fragment separation and purification. It is therefore suitable for identification of any strains of influenza virus. An ongoing study using this assay has simultaneously detected and confirmed influenza A(H3N2), pH1N1, and influenza B viruses in >100 nasopharyngeal swab samples (J. Zhao et al., unpub. data).

This detection platform provides a new, accurate, and rapid method to refine the differential diagnosis of influenza by selecting a single test or a small set of tests to determine the strain or strains present in a single clinical sample. We propose a new diagnostic algorithm based on this combined platform for identification and characterization of infection risks of unknown influenza strains (Figure 4). For testing a suspected influenza virus infection, this detection platform takes 2–3 days to perform NGS assay and data analysis. However, it provides whole-genome characterization and a final report in matrix type by which a potential pandemic prevalence strain can be predicted, with data including the genetic variant, amino acid signatures for virulence factors, and drug-resistance– and host-adaptation–associated mutations.

Future studies need to be conducted to reformat the current microarray to a point-of-care setting and to expand testing of clinical samples to other geographic regions and additional influenza virus types/subtypes. The NGS assay involves sample preparation and generates massive sequence data for the final report for test interpretation, which requires a higher level of performance for clinical assay

validation. Development of an automated assembly and analysis pipeline can make the bioinformatics analysis of transferring raw reads to the specific genomic identification more efficient. This molecular diagnostic platform has the potential for monitoring newly emerging or re-emerging viral reassortants derived from different precursors and could be included as a part of pandemic influenza surveillance strategies for efficient prevention and timely implementation of treatment to protect and improve public health.

Acknowledgments

We are thankful to FDA Center for Biologics Evaluation and Research core facility staff for help with some NGS assays and oligosynthesis.

This work was funded through the FDA Center for Biologics Evaluation and Research intramural and Medical Countermeasures Initiative funds.

The findings and conclusions in this article have not been formally disseminated by the Food and Drug Administration and should not be construed to represent any Agency determination or policy.

Dr Zhao is a virologist and reviewer at the Food and Drug Administration, Silver Spring, Maryland. His research interests include new technologies and tools for diagnosis of infectious viral pathogens.

References

1. Webster RG, Bean WJ, Gorman OT, Chambers TM, Kawaoka Y. Evolution and ecology of influenza A viruses. Microbiol Rev. 1992;56:152–79.
2. Röhm C, Zhou N, Suss J, Mackenzie J, Webster RG. Characterization of a novel influenza hemagglutinin, H15: criteria for determination of influenza A subtypes. Virology. 1996;217:508–16. http://dx.doi.org/10.1006/viro.1996.0145
3. Fouchier RA, Munster V, Wallensten A, Bestebroer TM, Herfst S, Smith D, et al. Characterization of a novel influenza A virus hemagglutinin subtype (H16) obtained from black-headed gulls. J Virol. 2005;79:2814–22. http://dx.doi.org/10.1128/JVI.79.5.2814-2822.2005
4. Tong S, Zhu X, Li Y, et al. New world bats harbor diverse influenza A viruses. PLoS Pathog 2013; 9(10): e1003657.
5. Horimoto T, Kawaoka Y. Pandemic threat posed by avian influenza A viruses. Clin Microbiol Rev. 2001;14:129–49. http://dx.doi.org/10.1128/CMR.14.1.129-149.2001
6. Hilleman MR. Realities and enigmas of human viral influenza: pathogenesis, epidemiology and control. Vaccine. 2002;20:3068–87. http://dx.doi.org/10.1016/S0264-410X(02)00254-2
7. Gao R, Cao B, Hu Y, Feng Z, Wang D, Hu W, et al. Human infection with a novel avian-origin influenza A (H7N9) virus. N Engl J Med. 2013;368:1888–97. http://dx.doi.org/10.1056/NEJMoa1304459
8. Arzey GG, Kirkland PD, Arzey KE, Frost M, Maywood P, Conaty S, et al. Influenza virus A (H10N7) in chickens and poultry abattoir workers, Australia. Emerg Infect Dis. 2012;18:814–6. http://dx.doi.org/10.3201/eid1805.111852
9. Cheng VC, Chan JF, Wen X, Wu WL, Que TL, Chen H, et al. Infection of immunocompromised patients by avian H9N2 influenza A virus. J Infect. 2011;62:394–9. http://dx.doi.org/10.1016/j.jinf.2011.02.007
10. Koopmans M, Wilbrink B, Conyn M, Natrop G, van der Nat H, Vennema H, et al. Transmission of H7N7 avian influenza A virus to human beings during a large outbreak in commercial poultry farms in the Netherlands. Lancet. 2004;363:587–93. http://dx.doi.org/10.1016/S0140-6736(04)15589-X
11. Ostrowsky B, Huang A, Terry W, Anton D, Brunagel B, Traynor L, et al. Low pathogenic avian influenza A (H7N2) virus infection in immunocompromised adult, New York, USA, 2003. Emerg Infect Dis. 2012;18:1128–31. http://dx.doi.org/10.3201/eid1807.111913
12. Tweed SA, Skowronski DM, David ST, Larder A, Petric M, Lees W, et al. Human illness from avian influenza H7N3, British Columbia. Emerg Infect Dis. 2004;10:2196–9. http://dx.doi.org/10.3201/eid1012.040961
13. Pabbaraju K, Tellier R, Wong S, Li Y, Bastien N, Tang JW, et al. Full-genome analysis of avian influenza A(H5N1) virus from a human, North America, 2013. Emerg Infect Dis. 2014;20:887–91. http://dx.doi.org/10.3201/eid2005.140164
14. Gallaher WR. Towards a sane and rational approach to management of influenza H1N1 2009. Virol J. 2009;6:51. http://dx.doi.org/10.1186/1743-422X-6-51
15. Vijaykrishna D, Poon LL, Zhu HC, Ma SK, Li OT, Cheung CL, et al. Reassortment of pandemic H1N1/2009 influenza A virus in swine. Science. 2010;328:1529. http://dx.doi.org/10.1126/science.1189132
16. Smith GJ, Vijaykrishna D, Bahl J, Lycett SJ, Worobey M, Pybus OG, et al. Origins and evolutionary genomics of the 2009 swine-origin H1N1 influenza A epidemic. Nature. 2009;459:1122–5. http://dx.doi.org/10.1038/nature08182
17. Garten RJ, Davis CT, Russell CA, Shu B, Lindstrom S, Balish A, et al. Antigenic and genetic characteristics of swine-origin 2009 A(H1N1) influenza viruses circulating in humans. Science. 2009;325:197–201. http://dx.doi.org/10.1126/science.1176225
18. Lindstrom S, Garten R, Balish A, Shu B, Emery S, Berman L, et al. Human infections with novel reassortant influenza A(H3N2) v viruses, United States, 2011. Emerg Infect Dis. 2012;18:834–7. http://dx.doi.org/10.3201/eid1805.111922
19. Nelson MI, Vincent AL, Kitikoon P, Holmes EC, Gramer MR. Evolution of novel reassortant A/H3N2 influenza viruses in North American swine and humans, 2009–2011. J Virol. 2012;86:8872–8. http://dx.doi.org/10.1128/JVI.00259-12
20. Kitikoon P, Vincent AL, Gauger PC, Schlink SN, Bayles DO, Gramer MR, et al. Pathogenicity and transmission in pigs of the novel A(H3N2)v influenza virus isolated from humans and characterization of swine H3N2 viruses isolated in 2010–2011. J Virol. 2012;86:6804–14. http://dx.doi.org/10.1128/ JVI.00197-12
21. Webby RJ, Webster RG. Emergence of influenza A viruses. Philos Trans R Soc Lond B Biol Sci. 2001;356:1817–28. http://dx.doi.org/10.1098/rstb.2001.0997
22. Chang SY, Lin PH, Tsai JC, Hung CC, Chang SC. The first case of H7N9 influenza in Taiwan. Lancet. 2013;381:1621. http://dx.doi.org/10.1016/S0140-6736(13)60943-5
23. Zhao J, Wang X, Ragupathy V, Zhang P, Tang W, Ye Z, et al. Rapid detection and differentiation of swine-origin influenza a virus (H1N1/2009) from other seasonal influenza A viruses. Viruses. 2012;4:3012–9. http://dx.doi.org/10.3390/v4113012
24. Zhao J, Tang S, Storhoff J, Marla S, Bao YP, Wang X, et al. Multiplexed, rapid detection of H5N1 using a PCR-free nanoparticle-based genomic microarray assay. BMC Biotechnol. 2010;10:74. http://dx.doi.org/10.1186/1472-6750-10-74
25. Landry ML, Ferguson D. Cytospin-enhanced immunofluorescence and impact of sample quality on detection of novel swine origin (H1N1) influenza virus. J Clin Microbiol. 2010;48:957–9. http://dx.doi.org/10.1128/JCM.01678-09
26. Zhou B, Donnelly ME, Scholes DT, St George K, Hatta M, Kawaoka Y, et al. Single-reaction genomic amplification

accelerates sequencing and vaccine production for classical and swine origin human influenza a viruses. J Virol. 2009;83:10309–13. http://dx.doi.org/10.1128/JVI.01109-09
27. Hoffmann E, Stech J, Guan Y, Webster RG, Perez DR. Universal primer set for the full-length amplification of all influenza A viruses. Arch Virol. 2001;146:2275–89. http://dx.doi.org/10.1007/s007050170002
28. Treangen TJ, Salzberg SL. Repetitive DNA and next-generation sequencing: computational challenges and solutions. Nat Rev Genet. 2012;13:36–46.
29. Tamura K, Dudley J, Nei M, Kumar S. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol Biol Evol. 2007;24:1596–9. http://dx.doi.org/10.1093/molbev/msm092
30. Chen Y, Liang W, Yang S, Wu N, Gao H, Sheng J, et al. Human infections with the emerging avian influenza A H7N9 virus from wet market poultry: clinical analysis and characterisation of viral genome. Lancet. 2013;381:1916–25. http://dx.doi.org/10.1016/S0140-6736(13)60903-4
31. Yurovsky A, Moret BM. FluReF, an automated flu virus reassortment finder based on phylogenetic trees. BMC Genomics. 2011;12(Suppl 2):S3. http://dx.doi.org/10.1186/1471-2164-12-S2-S3
32. Rabadan R, Levine AJ, Krasnitz M. Non-random reassortment in human influenza A viruses. Influenza Other Respir Viruses. 2008;2:9–22. http://dx.doi.org/10.1111/j.1750-2659.2007.00030.x
33. Lun AT, Wong JW, Downard KM. FluShuffle and FluResort: new algorithms to identify reassorted strains of the influenza virus by mass spectrometry. BMC Bioinformatics. 2012;13:208. http://dx.doi.org/10.1186/1471-2105-13-208
34. Jonges M, Meijer A, Fouchier RA, Koch G, Li J, Pan JC, et al. Guiding outbreak management by the use of influenza A(H7Nx) virus sequence analysis. Euro Surveill. 2013;18:20460.
35. Munster VJ, de Wit E, van Riel D, Beyer WE, Rimmelzwaan GF, Osterhaus AD, et al. The molecular basis of the pathogenicity of the Dutch highly pathogenic human influenza A H7N7 viruses. J Infect Dis. 2007;196:258–65. http://dx.doi.org/10.1086/518792
36. Mok CK, Lee HH, Lestra M, Nicholls JM, Chan MC, Sia SF, et al. Amino acid substitutions in polymerase basic protein 2 gene contribute to the pathogenicity of the novel A/H7N9 influenza virus in mammalian hosts. J Virol. 2014;88:3568–76. http://dx.doi.org/10.1128/JVI.02740-13
37. Hatta M, Gao P, Halfmann P, Kawaoka Y. Molecular basis for high virulence of Hong Kong H5N1 influenza A viruses. Science. 2001;293:1840–2. http://dx.doi.org/10.1126/science.1062882
38. Zaraket H, Saito R, Suzuki Y, Suzuki Y, Caperig-Dapat I, Dapat C, et al. Genomic events contributing to the high prevalence of amantadine-resistant influenza A/H3N2. Antivir Ther. 2010;15:307–19. http://dx.doi.org/10.3851/IMP1538
39. Liu Q, Lu L, Sun Z, Chen GW, Wen Y, Jiang S. Genomic signature and protein sequence analysis of a novel influenza A (H7N9) virus that causes an outbreak in humans in China. Microbes Infect. 2013;15:432–9. http://dx.doi.org/10.1016/j.micinf.2013.04.004
40. Chen Y, Liang W, Yang S, Wu N, Gao H, Sheng J, et al. Human infections with the emerging avian influenza A H7N9 virus from wet market poultry: clinical analysis and characterization of viral genome. Lancet. 2013;381:1916–25. http://dx.doi.org/10.1016/S0140-6736(13)60903-4

Address for correspondence: Jiangqin Zhao or Indira Hewlett, Laboratory of Molecular Virology, CBER/FDA, Bldg 52/72, Rm 4308, 10903 New Hampshire Ave, Silver Spring, MD 20993-0002, USA; email: jiangqin.zhao@fda.hhs.gov or indira.hewlett@fda.hhs.gov

Article DOI: http://dx.doi.org/10.3201/eid2103.141169

Nanomicroarray and Multiplex Next-Generation Sequencing for Simultaneous Identification and Characterization of Influenza Viruses

Technical Appendix

Design of capture and intermediate oligonucleotides

The sequences for capture and intermediate oligonucleotides were designed and prepared as described previously. In brief, by using nucleotide sequences available in the National Center for Biotechnology Information (NCBI) influenza resource (http://www.ncbi.nlm.nih.gov/genomes/FLU/) and the Influenza Primer Design Resource (IPDR) (http://www.ipdr.mcw.edu/fludb/search), we performed multiple sequence alignments of HA, NA and M genes with MEGA 5 and Vector NTI Advance 11 (Invitrogen, Foster City, CA). Over 120 known sequences of HA and NA genes of different subtypes were selected from ≈500 hit count analyses of each gene covering sequences of viruses from different geographic regions (i.e., Northern America, Europe, and Asia) and different time periods. The selected nucleotide sequences were further aligned to identify significant divergent or conserved regions for detection and differentiation of subtypes. Multiple degenerate capture and intermediate oligonucleotides (usually four or five, respectively) complementary to the conserved regions covering the entire genome of each subtype were designed. This critical design ensures that the target gene is captured on the microarray, hybridized with intermediate oligonucleotides and detected using the gold nanoparticle probe. The designed intermediate oligonucleotides modified with 5′-amino-C_6-modifier while a 25-mer poly (A) tail were added at the 3′ end of intermediate oligonucleotides during synthesis (Integrated DNA Technologies, Coralville, IA, USA). Capture oligonucleotides that did not bind to any known sequence of influenza A viruses were included as array internal positive controls. The oligonucleotides sequences are listed in Technical Appendix Table 1.

Viruses and clinical samples

Influenza virus isolates were propagated in 9–11 day-old embryonated chicken eggs at 33°C for 48h and then aliquoted and stored at −70°C until use in the FDA Center for Biologics Evaluation and Research. Virus infectivity was determined by plaque assay using MDCK cells or 50% egg infectious dose (EID_{50}). Some influenza reference strains were provided by Dr Stephen Lindstrom (Centers for Disease Control and Prevention, Atlanta, GA) and Dr Maryna Eichelberger (Food and Drug Administration, Silver Spring, MD) or purchased from ZeptoMetrix (ZeptoMetrix Corp., Buffalo, NY). Fifteen different strains, A/Puerto Rico/8/1934 (H1N1), A/California/04/2009 (pH1N1), A/Japan/305/1957 (H2N2), A/Panama/2007/1999 (H3N2), A/Brisbane/10/2007 (H3N2), A/Minnesota/10/2012 (H3N2), A/Indiana/08/2011 (H3N2v), A/Vietnam/1203/2004 (H5N1), A/turkey/Virginia/4529/2002 (H7N2), A/ruddy turnstone/NJ/65/1985 (H7N3), A/chicken/Hong Kong/G9/1997 (H9N2), A/Anhui/1/2013 (H7N9), B/Brisbane/60/2008 (Victoria lineage), B/Pennsylvania/7/2007 (Yamagata lineage), and B/Victoria/304/2006 (Victoria lineage), were selected for the nanomicroarray and NGS assays.

Reverse transcription PCR

Viral RNA was extracted directly from allantoic fluid or cell culture supernatants with QIAamp Viral RNA Mini Kit (QIAGEN, Valencia, CA). The purified RNA was quantified using a NanoDrop UV spectrometer (NanoDrop Technologies, Inc., Rockland, DE). Viral RNA was first transcribed into cDNA using a SuperScript III First-Strand Synthesis System for RT-PCR (Invitrogen) according to the manufacturer's instructions. The cDNA was then used as a template for RT-PCR. To evaluate capture oligonucleotides in the nanomicroarray assay, PCR primers sets for amplification of HA, NA, and M genes of H2N2, H7N2, H7N3 and H9N2 were designed and are listed in Technical Appendix Table 1. The PCR products (size: 1747 bp, 1335 bp, and 1015 bp) of three gene segments were amplified simultaneously for these viral subtypes. In addition, a universal primer set previously reported to amplify all eight gene segments was modified by adding a 13 bp in length of flanking sequence (5′-ACGACGGGCGACA-3′) at the 5′ end of each primer to enhance the annealing temperature and achieve high fidelity and yield in PCR amplification. Reverse transcription (RT) was performed with a uni12 primer SuperScript III First-Strand Synthesis System for RT-PCR. For amplification of all eight gene segments,

PCR was performed in a total volume of 30 µL containing 1 µL of cDNA, 15 µL of 2xPCR buffer (Extensor Hi-Fidelity ReddyMix PCR Master Mix, ABgen House, Surrey, UK), 2.5 pmol of forward and 2.5 pmol of reverse primers (unifluaf and unifluar). Reaction conditions included one cycle at 94°C, 5 min, 35 cycles at 94°C, 30 sec, 50°C, 40 sec, 68°C, 2.4 min, and one cycle at 68°C, 7 min. The PCR products were electrophoresed with 2.0% agarose gel slabs in Tris-acetate-EDTA buffer to observe multiple amplicons. PCR products were quantified using an Agilent 2100 Bioanalyzer (Agilent Technologies, Inc., Wilmington, DE), detected using the nanomicroarray assay and finally sequenced using the NGS assay. PCR was performed in Clinical Virology Laboratory at Yale New Haven Hospital using the CDC real-time reverse transcription PCR (RT-PCR) protocol for influenza (http://www.who.int/csr/resources/publications/swineflu/realtimeptpcr/en/index.html) on an ABI 7500 instrument (Applied Biosystems, Foster City, CA) (*25*).

Nanomicroarray printing and testing of samples

The printing solution contained 50 mM sodium phosphate at pH 8.5, 65% Pronto solution, 0.05% SDS, 0.01% glycerol, and 5 µM capture oligonucleotides. Four to five captures were prepared individually or mixed for identification of a typical gene segment. The nanomicroarray format was designed and the capture oligonucleotide array printed on CodeLink Activated slides in a double- or triple-spot format using an OMNIGrid Accent printer (Genomic Solutions Inc., Marlboroug, MI). Each slide contained 10 identical sub-arrays segregated by a hybridization gasket, thus allowing simultaneous testing of 10 samples per slide. Aqueous DNA-conjugated gold NP-probe and silver staining solutions were purchased from Nanosphere Inc (Northbrook, IL, USA). One µL of 1× to 10× diluted PCR products was used as template in nanomicroarray assay, RNA/DNA samples and 10 nM of the intermediate oligonucleotides were mixed in 100 µL of the hybridization buffer containing 5×SSC, 0.05% sorbitan mono-9-octa decenoate poly(oxy-1,1-ethanediyl), 0.05% Tween-20 and 40% formamide, applied to the nanomicroarray and incubated for 30 min at 40°C with orbital shaking at 500 rpm. PCR amplicons were denatured at 95°C for 5 min and cooled on ice for 2 min before being loaded on to the array. After three washes with wash buffer A (0.5 N $NaNO_3$, 0.01% SDS and 0.05% Tween-20) and one rinse with wash buffer B (0.5 N $NaNO_3$), the universal nanoparticle (NPs) probe was added to the slide and incubated for 30 min at 40°C. Slides were subsequently stained

with the Silver Enhancer A and B solutions for 5 min at room temperature. The light-scattering signal produced by silver-enhanced gold NPs was captured using a photosensor and converted to a TIFF image using a Verigene reader for analysis of the nanomicroarry assay data described previously. The resulting TIFF images were analyzed using GenePix Pro 7 software (Molecular Devices, San Francisco, CA).

Technical Appendix Table 1. Oligonucleotide sequences for capture (c), intermediate (i), and PCR

Oligo ID	Gene	Purpose	Sequences (5′ to 3′)
cFluBM1	Flu B M	capture	TGGGAAAGARTTTGACCTAGACTCTGCYTTGGAATGGATA
cFluBM2	Flu B M	capture	ATTCAATGCAAGTAAAACTAGGAACGCTCTGTGCTTTRTG
cFluBM3	Flu B M	capture	GGAAGGRATTGCAAAGGATGTAATGGAAGTGCTAAAGCAG
cFluBM4	Flu B M	capture	GAGACAATAAACAGAGAGGTATCAATTTTGAGACACAGTT
cH21	H2	capture	RAAYGTCACYGTGACTCAYGCCAAGGACATTCTTGAGAAR
cH22	H2	capture	YTGYAGCATTGCMGGATGGCTCCTTGGAAATCCAGAATGT
cH23	H2	capture	GGTAATYTAATTGCACCAGAGTATGGWTTCAAAATATCGA
cH24	H2	capture	TGTTTGGRGCAATAGCTGGTTTTATAGARGGNGGATGGCA
cH25	H2	capture	CRTATGATTATCCCAARTATGAAGARGAGTCYAARCTRAA
cH71	H7	capture	GACAAAATATGTCTTGGGCACCATGCTGTGGCAAATGGAA
cH72	H7	capture	GACAARATWTGYCTTGGGCAYCATGCYGTGSCAAAYGGRA
cH73	H7	capture	GTTCTTCTTTCTATGCAGAGATGAAGTGGTTGCTGTCGAA
cH74	H7	capture	GWTCYTCWTTCTATGCRGARATGAARTGGYTDYTGTCRAA
cH75	H7	capture	TTCACTTTCAATGGGGCATTCATAGCCCCTGACAGGGCAA
cH76	H7	capture	TTCASYTTCAATGGGGCATTCATAGCYCCWGAYMGDGYAA
cH77	H7	capture	TGGGAGGGTCTCATCAATGGATGGTATGGTTTCAGACATC
cH78	H7	capture	TGGGARGGTYTSATYRAYGGRTGGTATGGYTTCAGRCATC
cH79	H7	capture	GGTTTAGCTTCGGGGCATCATGTTTTCTTCTTCTAGCCAT
cH710	H7	capture	GGTTTAGCTTCGGGGCATCATGYTTYMTWCTTCTDGCCAT
cH91	H9	capture	ATGGGATGCTRTGTGCAACAAAYCTGGGACRTCCYCTYAT
cH92	H9	capture	CAAGTGTGRCAACAGAAGATATAAATAGRACCTTCAAACC
cH93	H9	capture	GGGAGGTTGGTCAGGRYTAGTYGCTGGTTGGTATGGGTTC
cH94	H9	capture	TGAYCAGTGCATGGAGACAATTCGGAACGGGACCTAYAAC
cN791	N2†	capture	AAATCAGAAGATAATAACAATTGGCTCCGTCTCTCTAACC
cN792	N2†	capture	AAATCAGAAGATAATARCAATTGGYTCYGTYTCYCTAACY
cN793	N2†	capture	TGGGAACCAGACAAGTTTGCATAGCATGGTCCAGCTCAAG
cN794	N2†	capture	TGGSAACCARACAAGTKTGYATAGCATGGTCCAGCTCAAG
cN795	N2†	capture	GAAGTGCTCAGCATATAGAGGAATGTTCCTGTTATCCCCG
cN796	N2†	capture	GRAGTGCTCAGCATRTRGAGGAATGYTCCTGTTAYCCCCG
cN797	N2†	capture	ATCAATAGGTGTTTTTATGTGGAGTTAATAAGAGGAAGGC
cN798	N2†	capture	ATYAAYAGGTGTTTTTATGTRGAGTTRATAAGRGGAAGRC
cN31	N3	capture	TACCGAATTGCAGTGACACTATAATAACATACAATAATAC
cN32	N3	capture	CAGACTCCATTAAATCATGGAGAAAGGACATATTGAGAAC
cN33	N3	capture	TGGATGAGAATCAACAACGAGACTATACTGGAAACAGGGT
cN34	N3	capture	TAGTTACTTTCTGTGGATTAGACAATGAACCTGGATCGGG
pCtrl		capture	ACTGTTTGTTATCTTGTTATCGTTATCTGA
iFluBM1	Flu B M	intermediate	TGACAGAAGATGGAGAAGGCAAAGCAGAACTAGCAGAAAA*
iFluBM2	Flu B M	intermediate	TGAAGCATTTGAAATAGCAGAAGGCCATGAAAGCTCAGCG*
iFluBM3	Flu B M	intermediate	TCTCAGCTATGAACACAGCAAAAACAATGAATGGAATGGG*
iFluBM4	Flu B M	intermediate	GCTCTCCATTTCRTGGCTTGGACAATAGGRCATTTGAATC*
iFluBM5	Flu B M	intermediate	GAYCACATARTAATTGAGGGGCTTTCTGCYGAAGAGATAA*
iH21	H2	intermediate	CCARATATGYATYGGRTACCATKCCAATAATTCCACAGAG*
iH22	H2	intermediate	YTAYCCAGGCAGYTTCAATGATTATGARGAATTGAAACAT*
iH23	H2	intermediate	CAACTGGWGGTTCWCGGGCCTGYGCRGTRTCTGGYAAYCC*
iH24	H2	intermediate	YTTTCACAAYRTYCACCCAYTGACAATWGGTGAGTGCCCC*
iH25	H2	intermediate	ACAGCAATGAYCARGGATCAGGRTATGCAGCAGACAAAGA*
iH26	H2	intermediate	GGAAGAYGGRTTTCTWGATGTRTGGACATAYAATGCYGAR*
iH71	H7	intermediate	GGACCTCCCCAATGTGATCAATTCCTGGAGTTTTCCTCTG*
iH72	H7	intermediate	GGWCCWCCMCARTGYGAYCAATTYCTRGARTTTKMSKWTG*
iH73	H7	intermediate	AGTTGATAACAGTAAGAAGCTCAAAATACCAGCAATCATT*
iH74	H7	intermediate	AGTTGATAACAGTAGGAAGCTCGAAATACCAGCAATCATT*
iH75	H7	intermediate	AGYTRATAACAGTWGGRAGYTCBAADTAYCARCARTCHTT*
iH76	H7	intermediate	TCTGGCTACAGGAATGAGAAATGTTCCAGAGAAACCAAAG*
iH77	H7	intermediate	TTTGGCTACAGGAATGAGAAACGTCCCAGAGAACCCCAAG*
iH78	H7	intermediate	DYTGGCWACWGGRATGARRAAYGTYCCHGARAHYCCMAAR*

Oligo ID	Gene	Purpose	Sequences (5′ to 3′)
iH79	H7	intermediate	TTTGCATAAAGAATGGAAACATGCGGTGCACTATTTGTAT*
iH710	H7	intermediate	THTGYRTRAAGAATGGAAACATGCRGTGCACTATTTGTAT*
iH91	H9	intermediate	ATCGTYGAAAGACCATCGGCYGTTAATGGAWTGTGTTACC*
iH92	H9	intermediate	GTGGTAACTGTGYAGTGCAATGTCARACWGAAARAGGTGG*
iH93	H9	intermediate	GACATATGGRCWTATAAYGCAGAATTGCTAGTRCTGCTTG*
iH94	H9	intermediate	CTTACAAAATCCTCACCATTTATTCGACTGTCGCCTCATC*
iN791	N2	intermediate	TGGTGGAGACATTTGGATAACAAGAGAGCCTTATGTGTCG*
iN792	N2	intermediate	WGGTGGAGAYATYTGGRTRACAAGAGARCCTTATGTRTCR*
iN793	N2	intermediate	TGCATCAATGGGTCTTGTACAGTAGTAATGACTGATGGAA*
iN794	N2	intermediate	TGYATCAATGGRWCYTGTACRGTAGTAATGACKGATGGAA*
iN795	N2	intermediate	CAGGAGTGAAAGGATGGGCCTTTGACAGTGGAGATGATGT*
iN796	N2	intermediate	CAGGAGTGAAAGGGTGGGCCTTTGACAGTGGRRATGAYRT*
iN797	N2	intermediate	GCACTTATGGAACAGGCTCATGGCCTGATGGGGCGAACAT*
iN798	N2	intermediate	GYACYTATGGAACAGGCTCATGGCCTGATGGGGCGAAYAT*
iN31	N3	intermediate	GAATCCAAATCAGAAGATAATAACAATCGGGGTAGTGAAT*
iN32	N3	intermediate	GACAGAACACCATATAGGTCTCTGATCCGATTCCCAATAG*
iN33	N3	intermediate	CTTGTGTTGTTACTGTCACAGATGGCCCTGCTGCTAATAG*
iN34	N3	intermediate	ACACTGGTGTCCAACAATGATTGGTCAGGCTATTCAGGTA*
uni12‡	Flu A	RT	AGCAAAAGCAGG
uni13‡	Flu A	PCR	AGTAGAAACAAGG
unifluaf	Flu A	PCR	ACGACGGGCGACAAGCAAAAGCAGG
unifluar	Flu A	PCR	ACGACGGGCGACAAGTAGAAACAAGG
FluAMf	Flu A M	PCR	TCTAACCGAGGTCGAAACG
FluAMr	Flu A M	PCR	TGACAAAATGACCATCGT
FluBMf	Flu B M	PCR	TCGCTGTTTGGAGAC
FluBMr	Flu B M	PCR	TTTATTTGCTGACATTGATTAC
H22f	H2	PCR	AGCAAAAGCAGGGGTTAT
H22r	H2	PCR	AGTAGAAACAAGGGTG
N22f	N2	PCR	TCAGGGAGCAAAAGCAGGAG
N22r	N2	PCR	AGTAGAAACAAGGAGT
H7273f	H7	PCR	CATTCATTGCTTGTGTGCT
H7273r	H7	PCR	TCTCAAACTATATACAAAT
N72r	N2	PCR	ATAGGCATGAAATTGAT
N7292f	N2	PCR	CAGGAGTGAAAATGAATC
N22r	N2	PCR	AGTAGAAACAAGGAGT
H7273f	H7	PCR	CATTCATTGCTTGTGTGCT
H7273r	H7	PCR	TCTCAAACTATATACAAAT
N72r	N2	PCR	ATAGGCATGAAATTGAT
N7292f	N2	PCR	CAGGAGTGAAAATGAATC
N92r	N2	PCR	ATAGGCATGAAGTTGAT
N73f	N3	PCR	GAATCCAAATCAGAAGATAATA
N73r	N3	PCR	TACTTGGGCATAAACCCAAT
H92f	H9	PCR	GTTCCTGTGACACATGCCAAAG
H92r	H9	PCR	AAGGCAGCAAACCCCATT
N22f	N2	PCR	TCAGGGAGCAAAAGCAGGAG
N22r	N2	PCR	AGTAGAAACAAGGAGT

*25-mer poly (A) tail added at 3′ end of each intermediate oligonucleotide. H: hemagglutinin; N: neuraminidase; M: matrix; f: foward; r: reverse.
†Oligonucleotide designed from H7N2 and H9N2 subtypes.
‡Sequences from previous publications (*1,2*).

Technical Appendix Table 2. Summary of results from NGS data analysis for reference strains

			Gene segment (length, bp)							
Strains	NGS Total contigs/reads	Finding	PB2 (2341)	PB1 (2341)	PA (2233)	HA (1778)	NP (1565)	NA (1413)	M (1027)	NS (890)
A/Puerto Rico/8/1934	8 / 122036	length (bp)	667	559	1552	1781	1210	1433	1028	1108
(H1N1)		read count	24	36	544	2056	6795	6549	196	51107
A/California/04/2009	7 / 93996	length (bp)	1846	2017	1827	1549	1555	754	1029	0
(pH1N1)		read count	3546	8335	3546	3996	9980	1246	18843	0
A/Japan/305/1957	8 / 60234	length (bp)	1658	1119	1341	1771	1206	1046	1114	882
(H2N2)		read count	7444	4253	3596	1621	4563	7302	2030	4253
A/Panama/2007/1999	4 / 106464	length (bp)	0	605	0	780	0	918	1026	0
(H3N2)		read count	0	423	0	130	0	178	8289	0
A/Vietnam/1203/2004	8 / 90962	length (bp)	2140	2205	2865	920	1580	1408	1042	697
(H5N1)		read count	8535	6103	14290	3170	14637	5537	18915	11715
A/turkey/Virginia/	8 / 63852	length (bp)	1780	1810	1671	1707	1570	1423	1031	886
4529/2002 (H7N2)		read count	632	334	2054	5978	8464	13057	9367	3440
A/RuddyTurnstone/	4 / 116382	length (bp)	0	509	0	0	1570	790	0	601
NJ/65/1985 (H7N3)		read count	0	6257	0	0	1267	4810	0	6257

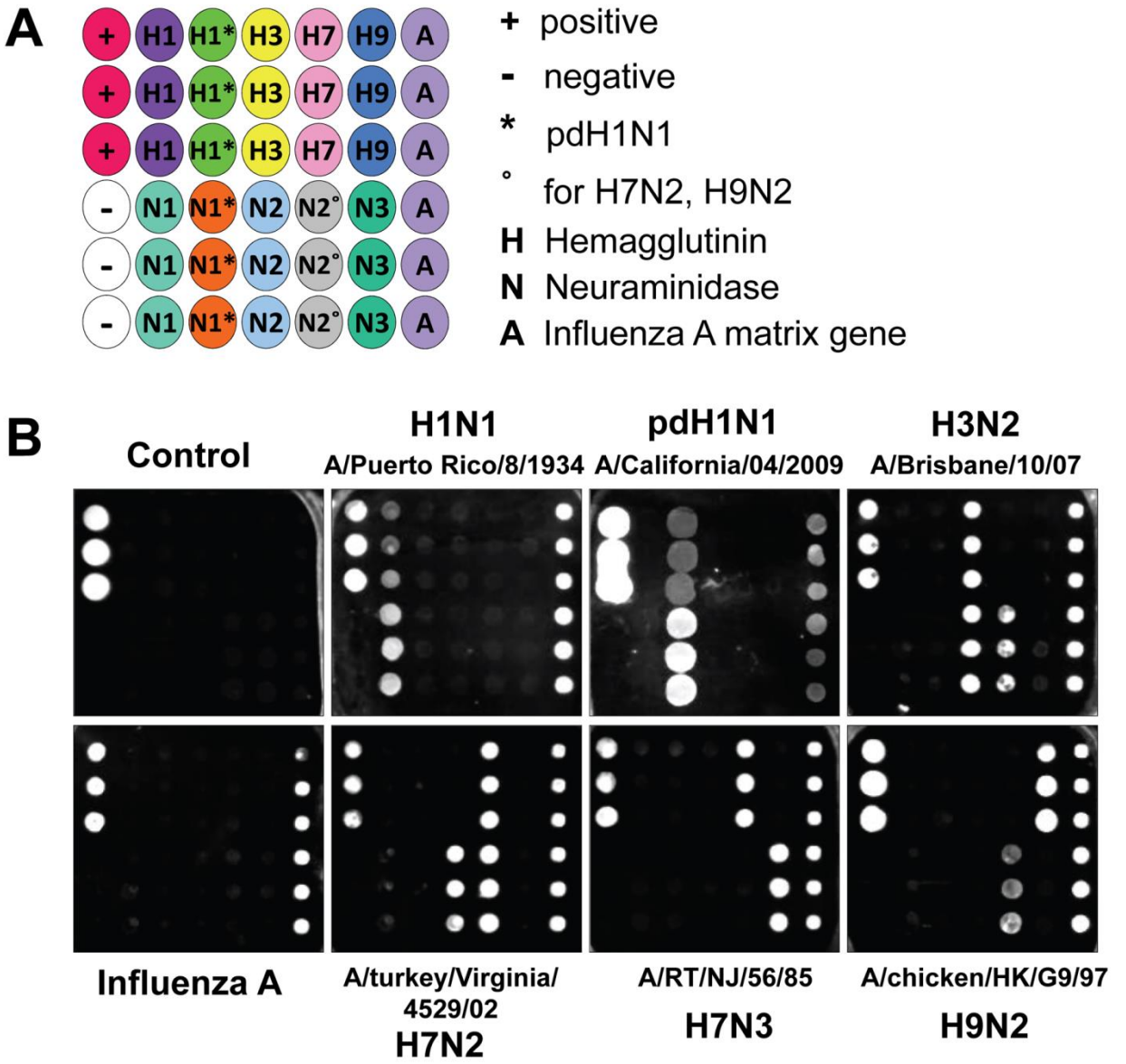
Strains	NGS Total contigs/reads	Finding	Gene segment (length, bp) PB2 (2341)	PB1 (2341)	PA (2233)	HA (1778)	NP (1565)	NA (1413)	M (1027)	NS (890)
A/chicken/Hong	7 / 65234	length (bp)	0	1756	1869	1729	1546	1458	939	690
Kong/G9/97 (H9N2)		read count	0	6549	4963	6746	8456	5632	8908	4594
A/Minnesota/10/2012	8 / 111564	length (bp)	1598	1925	2063	1099	1567	1470	1031	884
(H3N2)		read count	349	1816	4021	6847	19867	14500	21819	14762
A/Indiana/08/2011	8 / 66564	length (bp)	620	649	872	1762	1380	1054	1036	900
(H3N2v)		read count	33	32	2572	2557	1931	1366	21325	13026
B/Brisbane/60/2008	7 / 57860	length (bp)	694	944	713	635	520	626	794	890
		read count	7011	10168	674	237	36	30	45	0

De novo assembly module was used in CLC Genomics Workbench software (v6.0.2) package for result handing to set parameter for mapping reads back to contiguous, similarity fraction, 0.9; length fraction, 0.5; mismatch cost 2; insertion cost 3 and deletion cost 3. Minimum contiguous length sets on 800 to assemble the consensus sequences and coverage is over 1000 reads. The sample subtype was verified using Influenza Research Database (IRD) for Identify Similar Sequences (ISS) (vBLASTN 2.2.22).

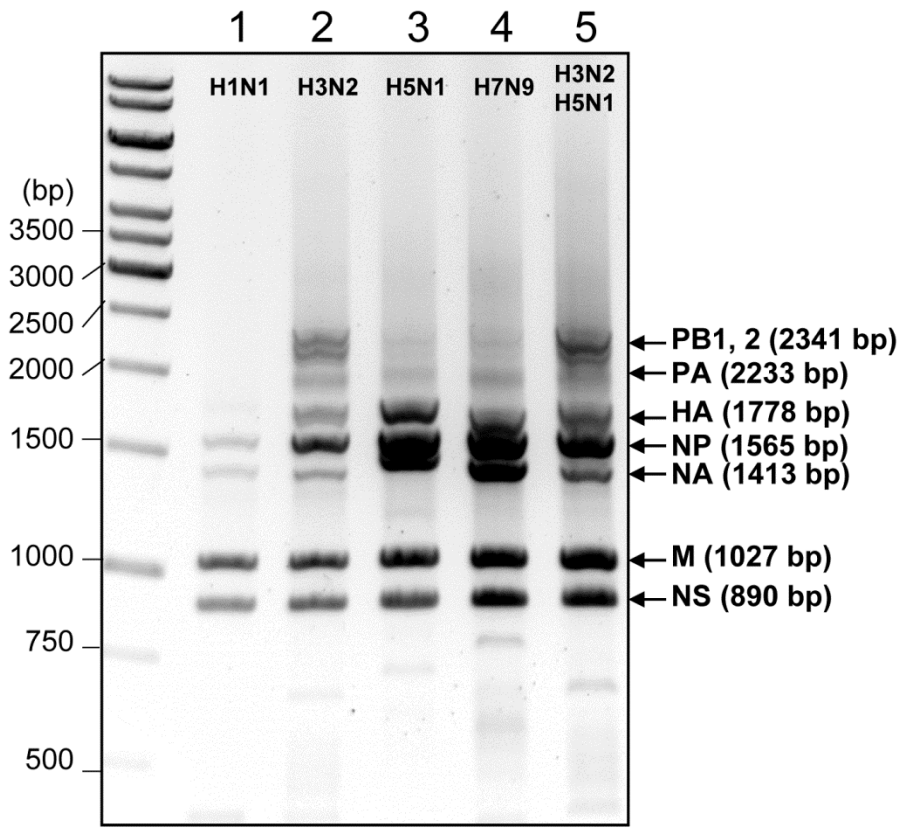
Technical Appendix Table 3. NGS data for detection and subtype of 24 clinical samples.

Patient ID	NGS Detection Total contigs	Ave. reads	Flu A contigs	Ave. reads	de novo assembling and bioinformatics analysis of segment (bp) PB2 (2341)	PB1 (2341	PA (2233)	HA (1778)	NP (1565)	NA (1413)	M (1027)	NS (890)	verified subtype
Flu001	8	216567	8	216567	2086	1922	2161	1771	1576	1479	1038	1086	H3N2
Flu002	8	233887	8	233887	2183	2204	2232	1771	1577	1613	1248	1060	H3N2
Flu004	7	208896	7	208896	2132	2017	1860	1895	1707	1817	1053	0	H3N2
Flu006	22	87902	8	203553	2302	1702	1870	1751	1769	1560	1208	1103	H3N2
Flu007	10	187419	8	232023	2123	1712	2060	1751	1576	1479	1251	891	H3N2
Flu008	20	50685	7	104105	2352	0	2371	1849	1620	1473	1037	945	H3N2
Flu009	8	85170	7	97168	1627	1754	2031	1764	1717	1693	1037	0	H3N2
Flu012	10	72106	7	72913	2105	2488	1791	1763	1575	1192	0	1068	H3N2
Flu013	8	194200	8	194200	1815	2071	2039	1762	1576	1297	1203	1121	H3N2
Flu014	9	243391	8	272331	2166	2447	1983	1828	1574	1484	1062	1042	H3N2
Flu017	10	95049	7	135026	2096	2572	2198	1764	1781	1680	0	977	H3N2
Flu018	7	76522	7	76522	1936	1957	1436	1705	1726	1901	0	891	H3N2
Flu021	8	119828	8	119828	2259	2351	2244	1952	1657	1594	1201	893	H3N2
Flu023	9	83396	8	93662	2196	2156	1929	1765	1574	1508	1325	1027	H3N2
Flu025	12	118080	8	170548	2349	2350	1990	1765	1572	1473	1036	899	H3N2
Flu026	12	155380	7	264292	1991	2475	2183	1909	1777	1557	1227	0	H3N2
Flu027	8	185997	8	185997	2352	2321	2071	1766	1579	1476	1650	893	H3N2
Flu028	8	163622	8	163622	2352	2348	2091	1765	1577	1475	1245	893	H3N2
Flu033	8	230527	8	230527	2244	2523	2390	1748	1631	1434	1243	892	H3N2
Flu034	12	132682	8	198210	2209	2011	2195	1761	1567	1467	1256	1115	H3N2
Flu036	8	229192	8	229192	2439	2611	1963	1881	1774	1677	1163	1077	H3N2
Flu037	8	120771	8	120771	2541	2153	2330	2094	1454	1476	1254	1231	H3N2
Flu038	10	103751	7	146822	2170	2091	2190	1959	1571	1473	1399	0	H3N2
Flu040	9	125152	7	160395	1976	1968	2062	1038	1776	1472	1035	0	H3N2

The sample subtype was verified using Influenza Research Database (IRD) for Identify Similar Sequences (ISS). All of the 181 genome sequences from 24 difference strains were submitted to the NCBI GenBank under accession nos. KJ741883–KJ742063 in April 2014.



Technical Appendix Figure 1. Microarray layout (A) and sample image (B) for different subtypes of influenza A viruses. Positive control capture (closed red circles), negative control which uses printing buffer as capture (closed white circles), degenerated HA, NA and matrix gene captures (filled as variable color of closed circles) are indicated. A portion of the microarray images for DNA oligonucleotide following hybridization with PCR products are shown and light shades represent greater silver intensities for each genes. Typical nanomicroarray silver staining images represent the hits for specific subtypes indicated.



Technical Appendix Figure 2. Identification of whole genomic PCR amplicons from one or two mixed influenza A viruses. Viral RNA was extracted and RT-PCR was performed using adapted universal primer set . PCR products were electrophoresed on a 2% agarose gel. Lane 1: A/Puerto Rico/8/1934 (H1N1). Lane 2: A/Minnesota/10/2012 (H3N2). Lane 3: A/Vietnam/1203/2004 (H5N1). Lane 4: A/Anhui/1/2013 (H7N9). Lane 5: A/Minnesota/10/2012 (H3N2) and A/Vietnam/1203/2004 (H5N1).

References

1. Hoffmann E, Stech J, Guan Y, Webster RG, Perez DR. Universal primer set for the full-length amplification of all influenza A viruses. Arch Virol. 2001;146:2275–89. PubMed http://dx.doi.org/10.1007/s007050170002

2. Zhou B, Donnelly ME, Scholes DT, St George K, Hatta M, Kawaoka Y, et al. Single-reaction genomic amplification accelerates sequencing and vaccine production for classical and swine origin human influenza a viruses. J Virol. 2009;83:10309–13. PubMed http://dx.doi.org/10.1128/JVI.01109-09