Identification of A/H5N1 Influenza Viruses Using a Single Gene Diagnostic Microarray

ARTICLE *in* ANALYTICAL CHEMISTRY · FEBRUARY 2007

Impact Factor: 5.64 · DOI: 10.1021/ac061920o · Source: PubMed

10 AUTHORS, INCLUDING:

Erica D Dawson
Indevr, Inc.
19 PUBLICATIONS **468** CITATIONS
SEE PROFILE

Michael B Townsend
Centers for Disease Control and Prevention
12 PUBLICATIONS **346** CITATIONS
SEE PROFILE

Robert Kuchta
University of Colorado Boulder
91 PUBLICATIONS **2,671** CITATIONS
SEE PROFILE

Small Correspondence

Identification of A/H5N1 Influenza Viruses Using a Single Gene Diagnostic Microarray

Erica D. Dawson, Chad L. Moore, Daniela M. Dankbar, Martin Mehlmann, Michael B. Townsend, James A. Smagala, Catherine B. Smith, Nancy J. Cox, Robert D. Kuchta, and Kathy L. Rowlen

Anal. Chem., **2007**, 79 (1), 378-384 • DOI: 10.1021/ac061920o

Anal. Chem. **2007,** *79,* 378−384

Correspondence

Identification of A/H5N1 Influenza Viruses Using a Single Gene Diagnostic Microarray

Erica D. Dawson,[†] Chad L. Moore,[†] Daniela M. Dankbar,[†] Martin Mehlmann,[†] Michael B. Townsend,[†] James A. Smagala,[†] Catherine B. Smith,[‡] Nancy J. Cox,[‡] Robert D. Kuchta,[†] and Kathy L. Rowlen*[,†,§]

Department of Chemistry and Biochemistry, UCB 215, University of Colorado at Boulder, Boulder, Colorado 80309, Influenza Division, The Centers for Disease Control and Prevention, 1600 Clifton Road, Atlanta, Georgia 30333, and InDevR, LLC, 2100 Central Avenue, Suite 106, Boulder, Colorado 80301

In previous work, a simple diagnostic DNA microarray that targeted only the matrix gene segment of influenza A (MChip) was developed and evaluated with patient samples. In this work, the analytical utility of the MChip for detection and subtyping of an emerging virus was evaluated with a diverse set of A/H5N1 influenza viruses. A total of 43 different highly pathogenic A/H5N1 viral isolates that were collected from diverse geographic locations, including Vietnam, Nigeria, Indonesia, and Kazakhstan, representing human, feline, and a variety of avian infections spanning the time period 2003−2006 were used in this study. A probabilistic artificial neural network was developed for automated microarray image interpretation through pattern recognition. The microarray assay and subsequent subtype assignment by the artificial neural network resulted in correct identification of 24 "unknown" A/H5N1 positive samples with no false positives. Analysis of a data set composed of A/H5N1, A/H3N2, and A/H1N1 positive samples and negative controls resulted in a clinical sensitivity of 97% and a clinical specificity of 100%.

Based on our collective experience each year, it is well known that influenza viruses impact the human population. Type A influenza viruses, which tend to have the highest impact in terms of mortality and morbidity, are further categorized by "subtype" according to the antigenic characteristics of two surface glycoproteins, hemagglutinin (HA) and neuraminidase (NA). The term A/H5N1 refers to a type A influenza virus with an H5 HA protein and a N1 NA protein. Of course, the nature of the protein is related to the gene that codes for that protein, and subtyping may also be achieved by sequencing the appropriate genes.

Highly pathogenic avian A/H5N1 influenza was first isolated from a human in Hong Kong in 1997.[1] As of July 4, 2006, 229 human infections of this avian virus have been reported resulting in 131 deaths,[2] with the majority of cases clearly arising from direct avian-to-human transmission. Although no widespread human-to-human transmission of the H5N1 virus has been reported to date, there is evidence of limited human-to-human transmission. [3] The rapid emergence of avian A/H5N1 has sparked a great deal of public and scientific concern and the initiation of worldwide pandemic preparedness measures.[4,5]

It is known that the A/H5N1 virus has undergone significant reassortment with other avian viruses since its occurrence in humans in 1997, with a number of different reassortants being identified as early as 2001.[6,7] Antigenic analyses of A/H5N1 viruses from 2002 to 2003 have also shown significant antigenic drift compared to viruses from 1997 to 2001.[8] In addition, the evolution of the HA gene of A/H5N1 since 2004 has resulted in several different lineages, and according to phylogenetic analyses, there is evidence that the M gene has coevolved with the HA gene.[9] As this particular virus continues to mutate, it is important to have a diagnostic tool capable of detecting as many of the A/H5N1 variants as possible for reliable global strain surveillance efforts. In addition, it has been recently highlighted that new, simple, rapid diagnostic tools capable of identifying both subtypes currently

* To whom correspondence should be addressed. E-mail: rowlen@colorado.edu.
[†] University of Colorado at Boulder.
[‡] The Centers for Disease Control and Prevention.
[§] InDevR, LLC.

(1) Subbarao, K.; Klimov, A.; Katz, J.; Regnery, H.; Lim, W.; Hall, H.; Perdue, M.; Swayne, D.; Bender, C.; Huang, J.; Hemphill, M.; Rowe, T.; Shaw, M.; Xu, X.; Fukuda, K.; Cox, N. *Science* **1998**, *279*, 393−396.

(2) World Health Organization. *Cumulative number of confirmed human cases of avian influenza A/(H5N1) reported to WHO.* Accessed July 16 2006. Avaliable from: http://www.who.int/csr/disease/avian_influenza/country/cases_table_2006_07_04/en/index.html.

(3) Centers for Disease Control and Prevention. Accessed November 6, 2006. Available from: http://www.cdc.gov/flu/avian/gen-info/avian-flu-humans.htm.

(4) Department of Health and Human Services. *HHS Pandemic Influenza Plan.* Accessed August 23 2006. Avaliable from: http://www.hhs.gov/pandemicflu/plan/pdf/HHSPandemicInfluenzaPlan.pdf.

(5) Stohr, K. *Vaccine* **2003**, *21*, 1744−1748.

(6) Li, K. S.; Guan, Y.; Wang, J.; Smith, G. J. D.; Xu, K. M.; Duan, L.; Rahardjo, A. P.; Puthavathana, P.; Buranathai, C.; Nguyen, T. D.; Estoepangestie, A. T. S.; Chaisingh, A.; Auewarakul, P.; Long, H. T.; Hanh, N. T. H.; Webby, R. J.; Poon, L. L. M.; Chen, H.; Shortridge, K. F.; Yuen, K. Y.; Webster, R. G.; Peiris, J. S. M. *Nature* **2004**, *430*, 209−213.

(7) Guan, Y.; Peiris, J. S. M.; Lipatov, A. S.; Ellis, T. M.; Dyrting, K. C.; Krauss, S.; Zhang, L. J.; Webster, R. G.; Shortridge, K. F. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 8950−8955.

(8) Guan, Y.; Poon, L. L. M.; Cheung, C. Y.; Ellis, T. M.; Lim, W.; Lipatov, A. S.; Chan, K. H.; Sturm-Ramirez, K. M.; Cheung, C. L.; Leung, Y. H. C.; Yuen, K. Y.; Webster, R. G.; Peiris, J. S. M. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 8156−8161.

10.1021/ac061920o CCC: $37.00
Published on Web 11/11/2006

circulating in humans (H1N1, H3N2) and the H5N1 influenza virus are needed for broad surveillance.[10] New diagnostic tools will be extremely important for informing health officials involved in vaccine development, for instituting quarantine measures in affected areas, and in the early administration of appropriate antiviral drugs.

A variety of methods have been used to identify and subtype influenza A viruses. The conventional method of laboratory diagnosis of influenza A is isolation of the virus in embryonated chicken eggs or in cell culture[11] followed by HA and/or NA subtyping by serological methods (e.g., hemagglutinin inhibition (HI) assay). Although viral isolation is a very sensitive method, it is tedious and time-consuming, with 1−2 weeks often needed to achieve subtype information. Methods based on reverse transcription-polymerase chain reaction (RT-PCR) and real-time RT-PCR have emerged as important diagnostic tools for detecting and subtyping influenza A.[12−20] PCR-based methods either detect a broad range of influenza A subtypes using a single primer set targeting a short portion of the relatively well-conserved matrix (M) gene segment[12−16] or detect a single influenza subtype by targeting a portion of the HA gene that is well-conserved for a particular subtype.[13,15] It is not possible to target all influenza subtypes using a single primer pair based on the HA gene, as there can be as little as 20% sequence homology between the different HA subtypes.[11] Several commercially available RT-PCR detection kits are available. For example, the artus Influenza/H5 LC RT-PCR Kit (Qiagen Diagnostics, Hamburg, Germany) can screen for all influenza A and B viruses and can detect A/H5 in a separate assay. Similarly, the TaqMan Influenza A/H5 Virus Detection Kit (version 1.0, Applied Biosystems, Foster City, CA) uses a single primer pair based on the M gene segment to detect influenza A and contains a separate assay for A/H5 detection using HA-specific primers. Although a duplex RT-PCR reagent kit (LightMix from TIB MolBiol, Berlin, Germany) is available, it only detects H5N1 since it utilizes specific primers based on the HA and NA genes. A single primer and probe set for PCR, which amplifies a portion of the HA gene of A/H5 viruses of Asian lineage, has been FDA approved.[21] While all of these existing singleplex RT-PCR-based methods are useful for rapid screening, they can only provide positive identification of a single subtype of influenza A.

In previous work, we demonstrated the ability to fully subtype influenza A viruses utilizing RT-PCR with a single primer pair for amplification of the M gene segment in conjunction with a low-density microarray, termed MChip.[22] The microarray assay utilized "capture" oligonucleotides immobilized on a functionalized glass substrate. Fragmented, amplified RNA was hybridized to the capture sequences, and the hybridization event was detected via a second fluorophore-functionalized "label" oligonucleotide in a sandwich assay. The microarray capture/label pairs were designed for either broad reactivity with a wide range of influenza A subtypes or more specific reactivity with a particular subtype.[22] Using only 15 capture sequences that targeted the relatively well-conserved matrix gene segment, this simple microarray exhibited distinct patterns in the relative fluorescence intensities for a variety of influenza A virus subtypes. In order to automate image interpretation and eliminate user subjectivity, an artificial neural network was used for pattern recognition. When evaluated in a blind study of patient samples, the MChip assay exhibited a clinical sensitivity of 95% and clinical specificity of 92%.

In order to test the hypothesis that the M gene segment is a useful diagnostic target even for rapidly evolving influenza viruses, the MChip was evaluated in this study with a diverse set of A/H5N1 influenza viruses in combination with a new (probabilistic) artificial neural network.

EXPERIMENTAL SECTION

Samples. H5N1 samples were originally received by the Centers for Disease Control (CDC, Atlanta, GA) and subsequently propagated in the allantoic cavities of 11-day-old embryonated chicken eggs or in Madin-Darby canine kidney cells. All handling of H5N1 samples was performed at the CDC. All research with H5N1 viruses or reassortment viruses was conducted under biosafety level 3 containment, including enhancements required by the U.S. Department of Agriculture and the Select Agent Program. The samples were initially tested for hemagglutination, and the HA subtypes of all hemagglutination-positive samples were identified by HI assay. Viral RNA was extracted from viral isolates

(9) Aubin, J.-T.; Azebi, S.; Balish, A.; Banks, J.; Bhat, N.; Bright, R. A.; Brown, I.; Buchy, P.; Burguiere, A.-M.; Chen, H.-I.; Cheng, P.; Cox, N. J.; Crosier, A.; Curns, A.; Cuvelier, F.; Deng, G.; Desheva, J.; Desvaux, S.; Diep, N. H.; Donis, R. O.; Douglas, A.; Dowell, S. F.; Dung, N. T.; Edwards, L.; Fukuda, K.; Garten, R.; Govorkova, E.; Gregory, V.; Hampson, A.; Hanh, N. T. H.; Harper, S.; Hay, A.; Hoffmann, E.; Hulse, D.; Imai, M.; Itamura, S.; Jadhao, S.; Jeannin, P.; Kang, C.; Katz, J.; Kim, J.-H.; Klimov, A.; Kwon, Y.-k.; Lee, C.-W.; Lien, P. S.; Li, Y.; Lim, W.; Lin, Y. P.; Lindstom, S.; Loftin, L.; Mabry, J.; Mai, L. Q.; Maines, T.; Manuguerra, J.-C.; Mase, M.; Matsuoka, Y.; McCarron, M.; Medina, M.-J.; Nguyen, D.; Ninomiya, A.; Obuchi, M.; Odagiri, T.; Peiris, M.; Perdue, M. L.; Reynes, J.-M.; Robertson, J.; Rousseaux, C.; Saito, T.; Sangkitporn, S.; Shaw, M.; Simmerman, J. M.; Slomka, M.; Smith, C.; Sorn, S.; Spackman, E.; Stoehr, K.; Suarez, D. L.; Sung, H. W.; Swayne, D. E.; Tardy-Panit, M.; Tashiro, M.; Thawatsupha, P.; Tumpey, T.; Uyeki, T.; Tu, P. V.; van de Werf, S.; Vong, S.; Webby, R.; Webster, R.; Wood, J.; Xu, X.; Yi, G.; Zhang, W. *Emerg. Infect. Dis.* **2005**, *11*, 1515−1521.

(10) Lu, P. S. *Science* **2006**, *312*, 337.

(11) Knipe, D. M.; Howley, P. M. *Field's Virology*, 4th ed.; Lippincott, Williams and Wilkins: Philadelphia, 2001.

(12) Fouchier, R. A. M.; Bestebroer, T. M.; Herfst, S.; Van Der Kemp, L.; Rimmelzwaan, G. F.; Osterhaus, A. D. M. E. *J. Clin. Microbiol.* **2000**, *38*, 4096−4101.

(13) Spackman, E.; Senne, D. A.; Myers, T. J.; Bulaga, L. L.; Garber, L. P.; Perdue, M. L.; Lohman, K.; Daum, L. T.; Suarez, D. L. *J. Clin. Microbiol.* **2002**, *40*, 3256−3260.

(14) Whiley, D. M.; Sloots, T. P. *Diagn. Microbiol. Infect. Dis.* **2005**, *53*, 335−337.

(15) Ng, L. F. P.; Barr, I.; Nguyen, T.; Noor, S. M.; Tan, R. S.-P.; Agathe, L. V.; Gupta, S.; Khalil, H.; To, T. L.; Hassan, S. S.; Ren, E.-C. *BMC Infect. Dis.* **2006**, *6*, 40.

(16) DiTrani, L.; Bedini, B.; Donatelli, I.; Campitelli, L.; Chiappini, B.; DeMarco, M. A.; Delogu, M.; Buonavoglia, C.; Vaccari, G. *BMC Infect. Dis.* **2006**, *6*, 87.

(17) Poddar, S. K. *J. Virol. Methods* **2002**, *99*, 63−70.

(18) Wei, H.-L.; Bai, G.-R.; Mweene, A. S.; Zhou, Y.-C.; Cong, Y.-L.; Pu, J.; Wang, S.; Kida, H.; Liu, J.-H. *Virus Genes* **2006**, *32*, 261−267.

(19) Payungporn, S.; Phakdeewirot, P.; Chutinimitkul, S.; Theamboonlers, A.; Keawcharoen, J.; Oraveerakul, K.; Amonsin, A.; Poovorawan, Y. *Viral Immunol.* **2004**, *17*, 588−593.

(20) Payungporn, S.; Chutinimitkul, S.; Chaisingh, A.; Damrongwantanapokin, S.; Buranathai, C.; Amonsin, A.; Theamboonlers, A.; Poovorawan, Y. *J. Virol. Methods* **2006**, *131*, 143−147.

(21) United States Food and Drug Administration. Accessed July 16 2006. Avaliable from: http://www.fda.gov/cdrh/reviews/K060159.pdf.

(22) Dawson, E. D.; Moore, C. L.; Smagala, J. A.; Dankbar, D. M.; Mehlmann, M.; Townsend, M. B.; Smith, C. B.; Cox, N. J.; Kuchta, R. D.; Rowlen, K. L. *Anal. Chem.* In press.

Table 1. A/H5N1 Virus Training Set

	specimen ID
1	A/chicken/Nigeria/42/2006
2	A/chicken/Nigeria/42/2006
3	A/chicken/Nigeria/42/2006
4	A/chicken/Nigeria/42/2006
5	A/chicken/Vietnam/NCVD15/2003
6	A/chicken/Vietnam/NCVD15/2003
7	A/chicken/Vietnam/NCVD30/2003
8	A/chicken/Vietnam/NCVD30/2003
9	A/chicken/Vietnam/NCVD-CDC22/2005
10	A/chicken/Vietnam/NCVD-CDC22/2005
11	A/chicken/Vietnam/NCVD-CDC33/2005
12	A/chicken/Vietnam/NCVD-CDC33/2005
13	A/chicken/Vietnam/NCVD-CDC33/2005
14	A/chicken/Vietnam/NCVD-CDC33/2005
15	A/chicken/Vietnam/NCVD-CDC36/2005
16	A/chicken/Vietnam/NCVD-CDC36/2005
17	A/chicken/Vietnam/NCVD-CDC37/2005
18	A/chicken/Vietnam/NCVD-CDC37/2005
19	A/chicken/Vietnam/NCVD-CDC42/2005
20	A/chicken/Vietnam/NCVD-CDC42/2005
21	A/chicken/Vietnam/NCVD-CDC51/2005
22	A/chicken/Vietnam/NCVD-CDC51/2005
23	A/chicken/Vietnam/NCVD-CDC52/2005
24	A/chicken/Vietnam/NCVD-CDC52/2005
25	A/duck/Vietnam/NCVD25/2003
26	A/duck/Vietnam/NCVD25/2003
27	A/duck/Vietnam/NCVD-CDC17/2005
28	A/duck/Vietnam/NCVD-CDC17/2005
29	A/duck/Vietnam/NCVD-CDC41/2005
30	A/duck/Vietnam/NCVD-CDC41/2005
31	A/duck/Vietnam/NCVD-CDC50/2005
32	A/duck/Vietnam/NCVD-CDC50/2005
33	A/goose/Kazakhstan/464/2005
34	A/goose/Kazakhstan/464/2005
35	A/goose/Kazakhstan/464/2005
36	A/goose/Kazakhstan/464/2005
37	A/goose/Kazakhstan/464/2005
38	A/goose/Kazakhstan/464/2005
39	A/Indonesia/CDC357/2006
40	A/Indonesia/CDC357/2006
41	A/Indonesia/CDC370T/2006
42	A/Indonesia/CDC370T/2006
43	A/Prachi Nburi/6231/2004
44	A/Prachi Nburi/6231/2004
45	A/Vietnam/JP20−2/2005
46	A/Vietnam/JP20−2/2005
47	A/Vietnam/JP36−2/2005
48	A/chicken/Vietnam/01/2004
49	A/swine/Vietnam/6774/2004
50	A/Vietnam/HN30408/05
51	A/Vietnam/JP4207/05
52	A/muscovy duck/Vietnam/04/2004
53	A/muscovy duck/Vietnam/17/2004
54	A/chicken/Vietnam/07/2004

using the Qiagen RNAeasy kit according to the manufacturer's instructions (Qiagen, Inc., Valencia, CA).

RNA Amplification. Reverse transcription was performed on viral RNA as described by Mehlmann et al.[23] PCR was performed on the first strand cDNA using the method and primers of Zou,[24] and products were visualized on a 1% agarose gel stained with ethidium bromide to evaluate amplification. The 5′ PCR primer included a T7 promoter site that allowed runoff transcription with

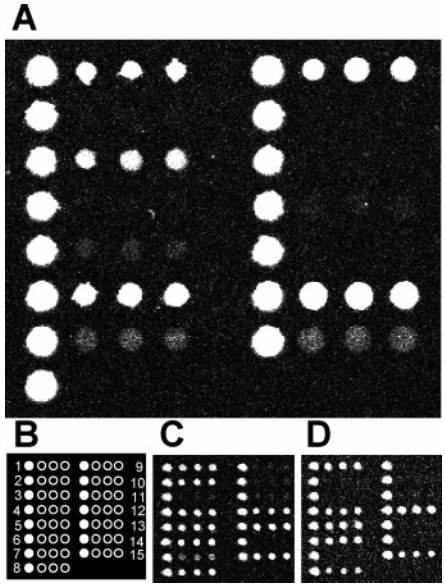

Figure 1. (A) Typical microarray fluorescence image for an A/H5N1 virus, (B) microarray layout with positive control sequences (closed circles) and capture sequences labeled numerically (open circles, spotted in triplicate), (C) typical A/H3N2 image, and (D) typical A/H1N1 image. Lighter shades represent greater fluorescence intensity.

T7 RNA polymerase (Invitrogen Corp., Carlsbad, CA). Transcribed RNA was stored at −20 °C for further use (within 24 h). Long-term storage was at −80 °C.

Microarray Experiments. Amplified RNA was fragmented as previously described.[25] Microarrays were prepared in-house as described by Dawson et al.[22] Hybridization of fragmented RNA was performed as described previously.[22] Briefly, a solution of 5 μL of fragmented RNA was added to 15 μL of the hybridization solution, and this 20 μL was applied to the microarray under a glass coverslip for 2 h and subsequently washed to remove any material bound nonspecifically. A positive control immobilized capture sequence and corresponding label sequence were included and used as an internal reference for hybridization efficiency. If the positive control indicated a failure in hybridization (due, for example, to poor substrate quality), the microarray was discarded. All hybridized slides were scanned using an Axon Gene Pix Professional 4200A scanner, and the collected fluorescence images were subsequently analyzed using VersArray Analyzer software, version 4.5 (Bio-Rad Laboratories, Hercules, CA). Mean raw intensity values were extracted from all replicate spots of each capture sequence in a single image, and the highest intensity average signal was normalized to 100. Note that patient samples can have widely variable viral loads, ranging from 10^2 to 10^9 viral particles/mL.[11] Although the absolute intensities on the microarray may differ from sample to sample as a result of varying viral load, efficiency of PCR amplification, or perhaps less efficient hybridization, the pattern of *relative* signal intensities was found to be independent of the starting concentration for a wide range of samples (data not shown but available upon request).

Probabilistic Neural Network Analysis. A probabilistic neural network (PNN) was utilized to perform pattern recognition

(23) Mehlmann, M.; Townsend, M. B.; Stears, R. L.; Kuchta, R. D.; Rowlen, K. L. *Anal. Biochem.* **2005**, *347*, 316−323.
(24) Zou, S. *J. Clin. Microbiol.* **1997**, *35*, 2623−2627.
(25) Townsend, M. B.; Dawson, E. D.; Mehlmann, M.; Smagala, J. A.; Dankbar, D. M.; Moore, C. L.; Smith, C. B.; Cox, N. J.; Kuchta, R. D.; Rowlen, K. L. *J. Clin. Microbiol.* **2006**, *44*, 2863−2871.

Table 2. PNN Probability Distribution for "Unknown" A/H5N1 Viruses and Negative Controls[a]

	probability distribution						
	H1N1	H3N2	H5N1	Neg	specimen ID	known subtype	correct ID?
1	0.01	0.00	*1.00*	0.00	A/Iraq/207-NAMRU3/2006	H5N1	
2	0.00	0.00	*1.00*	0.00	A/chicken/Vietnam/NCVD-CDC3/2005	H5N1	√
3	0.00	0.00	*1.00*	0.00	A/environment/Vietnam/NCVD-CDC54/2005	H5N1	√
4	0.00	0.00	*1.00*	0.00	A/duck/Kulon Progo/BBVET/IX/2004	H5N1	√
5	0.00	0.00	*1.00*	0.00	A/Anhui/2/2005	H5N1	√
6	0.00	0.00	*1.00*	0.00	A/Indonesia/CDC390/2006	H5N1	√
7	0.00	0.00	*0.99*	0.01	A/environment/Vietnam/NCVD-CDC60/2005	H5N1	√
8	0.00	0.00	*1.00*	0.00	A/muscovy duck/Vietnam/NCVD28/200	H5N1	√
9	0.00	0.00	*1.00*	0.00	A/environment/Vietnam/NCVD-CDC53/2005	H5N1	√
10	0.00	0.00	*1.00*	0.00	A/Anhui/1/2005	H5N1	√
11	0.00	0.00	*1.00*	0.00	A/chicken/Vietnam/NCVD-CDC37/2005	H5N1	√
12	0.00	0.00	*1.00*	0.00	A/chicken/Vietnam/NCVD-CDC23/2005	H5N1	√
13	0.00	0.00	*1.00*	0.00	A/duck/Kazakhstan/467/2005	H5N1	√
14	0.00	0.00	*1.00*	0.00	A/Indonesia/CDC326T/2006	H5N1	√
15	0.00	0.00	*1.00*	0.00	A/chicken/Vietnam/NCVD-CDC49/2005	H5N1	√
16	0.01	0.00	*0.99*	0.00	A/feline/Indonesia/CDC1/2006	H5N1	√
17	0.01	0.16	*0.83*	0.00	A/Indonesia/CDC624/2006	H5N1	√
18	0.01	0.00	*0.99*	0.00	A/Indonesia/CDC623/2006	H5N1	√
19	0.10	0.00	*0.90*	0.00	A/Indonesia/CDC594/2006	H5N1	√
20	0.08	0.01	*0.91*	0.00	A/Indonesia/CDC595/2006	H5N1	√
21	0.04	0.00	*0.96*	0.00	A/Indonesia/CDC597/2006	H5N1	√
22	0.02	0.00	*0.98*	0.00	no strain designation	H5N1	√
23	0.05	0.00	*0.95*	0.00	A/Indonesia/CDC596/2006	H5N1	√
24	0.05	0.00	*0.96*	0.00	A/Indonesia/CDC599/2006	H5N1	√
25	0.00	0.00	0.00	*1.00*	negative	neg	√
26	0.02	0.00	0.00	*0.91*	negative	neg	√
27	0.00	0.00	0.00	*1.00*	negative	neg	√
28	0.00	0.00	0.00	*1.00*	negative	neg	√
29	0.00	0.00	0.00	*1.00*	SARS	neg	√
30	0.00	0.20	0.00	*0.80*	RSV	neg	√
31	0.00	0.00	0.00	*1.00*	hPIV3	neg	√

[a] Probabilities of >0.80 are in italic type, and checkmarks indicate correct identification with an 80% probability threshold. "Neg" indicates the sample is known to be influenza A negative. Samples 29−31 are influenza-like illnesses included as negative controls: SARS CoA (coronavirus that causes severe acute respiratory syndrome), RSV (respiratory syncytial virus), hPIV3 (human parainfluenza virus type 3).

on the relative fluorescence intensities from the microarray data and subsequent classification of virus subtype.[22] While the first MChip utilized an artificial neural network,[22] the output scores were not a direct measure of probability that the unknown belonged to a specified category. Here a PNN was developed in order provide a quantitative degree of certainty regarding the classification of an unknown.[26] Specifically, when queried with an unknown, the output of a PNN for a specified category is the probability that the unknown is a member of that category. The MChip PNN was trained to recognize the fluorescence relative intensity patterns associated with each subtype using normalized input data for a set of "known" samples (referred to as the training set). The output categories were (H3N2, H1N1, H5N1, and negative).

The Weka (Waikato Environment for Knowledge Analysis) software package, version 3.4.8 was used to create the PNN model,[27] applying the multilayer perceptron function with feed-forward back-propagation. Training was based on minimization of squared error between the desired and obtained outputs for validating examples.[28] The PNN was built around 17 inputs: 15 normalized fluorescence intensities, the maximum mean intensity (MAX), and the sample category (i.e., H1N1, H3N2, H5N1, or negative). The MAX input was added to differentiate negative samples from influenza A positive samples, since the MAX value for negative samples was generally very low. The training set consisted of processed microarray data (including replicate experiments) from influenza A positive samples: (57 A/H3N2, 23 A/H1N1, 54 A/H5N1) and 29 negative samples, which included influenza B positive samples along with other respiratory pathogens (e.g., RSV). A list of the A/H5N1 viral isolates and replicates used in training is given in Table 1. During training, the PNN utilized 25% of the total training set for validation and error minimization. Once the PNN is adequately trained, it can be used to query any unknown. In this case, the trained PNN was used to identify a set unknowns composed of 24 distinct A/H5N1 samples. Then, in order to address how well it would perform with a broader range of subtypes, it was used to identify a set of unknowns composed of the two human-adapted influenza A viruses (H3N2, H1N1) as well as the H5N1 viruses. It is worthwhile noting that the number of human A/H5N1 viral isolates used in this study represents a remarkable 8% of the total number of human infections reported worldwide as of July 4, 2006.

(26) Berrar, D. P.; Downes, C. S.; Dubitzky, W. In *Pacific Symposium on Biocomputing: Lihue, Hawaii*; Altman, R. B., et al., Eds.; World Scientific: River Edge, NJ, 2003; Vol. 8, pp 5−16.

(27) Whitten, I. H.; Eibe, F. *Data Mining: Practical Machine Learning Tools and Techniques*, 2nd ed.; Morgan Kaufmann: San Francisco, 2005.

(28) Vasilyev, A.; Kapishnikov, A. *Proceedings of International Conference on Modeling and Simulation of Business Systems (MOSBIUS 2003)*, Vilnius, Lithuania, May 13−14, 2003; pp 79−81.

Table 3. PNN Probability Distribution for "Unknown" Influenza A H3N2, H1N1 and H5N1 Viruses and Negative Controls

	probability distribution					known	correct
	H1N1	H3N2	H5N1	Neg	specimen ID	subtype	ID?
1	*0.99*	0.00	0.01	0.00	A/Florida/03/2004	H1N1	√
2	*0.99*	0.01	0.00	0.00	n/a	H1N1	√
3	*0.98*	0.00	0.01	0.00	n/a	H1N1	√
4	*0.99*	0.00	0.00	0.00	n/a	H1N1	√
5	*0.99*	0.00	0.00	0.01	n/a	H1N1	√
6	0.02	*0.98*	0.00	0.00	n/a	H3N2	√
7	0.02	*0.98*	0.00	0.00	n/a	H3N2	√
8	0.01	*0.99*	0.00	0.00	n/a	H3N2	√
9	0.01	*0.99*	0.01	0.00	n/a	H3N2	√
10	0.01	*0.99*	0.00	0.00	n/a	H3N2	√
11	0.00	*1.00*	0.00	0.00	n/a	H3N2	√
12	0.00	*0.99*	0.00	0.01	n/a	H3N2	√
13	0.01	*0.99*	0.00	0.00	n/a	H3N2	√
14	0.01	*0.99*	0.00	0.00	n/a	H3N2	√
15	0.00	*0.99*	0.00	0.00	n/a	H3N2	√
16	0.00	0.27	0.73	0.00	A/Indonesia/CDC624/2006	H5N1	√
17	0.00	0.00	*1.00*	0.00	A/Iraq/207-NAMRU3/2006	H5N1	√
18	0.00	0.00	*1.00*	0.00	A/chicken/Vietnam/NCVD-CDC3/2005	H5N1	√
19	0.00	0.00	*1.00*	0.00	A/environment/Vietnam/NCVD-CDC54/2005	H5N1	√
20	0.00	0.00	*1.00*	0.00	A/duck/Kulon Progo/BBVET/IX/2004	H5N1	√
21	0.00	0.00	*1.00*	0.00	A/Anhui/2/2005	H5N1	√
22	0.00	0.00	*1.00*	0.00	A/Indonesia/CDC390/2006	H5N1	√
23	0.00	0.00	*1.00*	0.00	A/environment/Vietnam/NCVD-CDC60/2005	H5N1	√
24	0.00	0.00	*1.00*	0.00	A/muscovy duck/Vietnam/NCVD28/200	H5N1	√
25	0.00	0.00	*1.00*	0.00	A/environment/Vietnam/NCVD-CDC53/2005	H5N1	√
26	0.00	0.00	*1.00*	0.00	A/Anhui/1/2005	H5N1	√
27	0.00	0.00	*1.00*	0.00	A/chicken/Vietnam/NCVD-CDC37/2005	H5N1	√
28	0.00	0.00	*1.00*	0.00	A/chicken/Vietnam/NCVD-CDC23/2005	H5N1	√
29	0.00	0.00	*1.00*	0.00	A/duck/Kazakhstan/467/2005	H5N1	√
30	0.00	0.00	*1.00*	0.00	A/Indonesia/CDC326T/2006	H5N1	√
31	0.00	0.00	*1.00*	0.00	A/chicken/Vietnam/NCVD-CDC49/2005	H5N1	√
32	0.01	0.00	*0.99*	0.00	A/feline/Indonesia/CDC1/2006	H5N1	√
33	0.00	0.01	*0.99*	0.00	A/Indonesia/CDC623/2006	H5N1	√
34	0.03	0.00	*0.97*	0.00	A/Indonesia/CDC594/2006	H5N1	√
35	0.03	0.01	*0.97*	0.00	A/Indonesia/CDC595/2006	H5N1	√
36	0.01	0.00	*0.99*	0.00	A/Indonesia/CDC597/2006	H5N1	√
37	0.00	0.00	*1.00*	0.00	no strain designation	H5N1	√
38	0.01	0.00	*0.99*	0.00	A/Indonesia/CDC596/2006	H5N1	√
39	0.01	0.00	*0.99*	0.00	A/Indonesia/CDC599/2006	H5N1	√
40	0.00	0.53	0.00	0.47	RSV	neg	x
41	0.00	0.00	0.00	*1.00*	hPIV3	neg	√
42	0.00	0.00	0.00	*1.00*	hPIV3	neg	√
43	0.00	0.00	0.01	*0.99*	SARS	neg	√
44	0.00	0.00	0.00	*1.00*	negative	neg	√
45	0.01	0.00	0.00	*0.98*	negative	neg	√
46	0.00	0.00	0.01	*0.99*	negative	neg	√
47	0.00	0.00	0.00	*0.99*	negative	neg	√

[a] Probabilities of >0.95 are in italic type, and checkmarks indicate correct identification at this 95% probability threshold. "Neg" indicates the sample is known to be influenza A negative. Samples 40−43 are influenza-like illnesses included as negative controls: SARS CoA (coronavirus causing severe acute respiratory syndrome), RSV (respiratory syncytial virus), hPIV3 (human parainfluenza virus type 3). N/a, not available.

RESULTS AND DISCUSSION

It was previously established that distinct patterns of relative fluorescence intensity response on the MChip could be used to subtype influenza viruses.[22] In that work, the majority of samples examined were clinical patient samples infected with human influenza A subtype H3N2 or H1N1. In a blind study of 53 clinical samples, the MChip assay exhibited a remarkable 95% sensitivity and 92% specificity. Considering the potential pandemic nature of the avian influenza A/H5N1 virus, the work described here was aimed at extending the MChip analysis to include a wide variety of A/H5N1 viral isolates in an effort to assess its potential for rapid identification and subtyping of a rapidly evolving influenza virus.

A/H5N1 Microarray Results. Given the amount of variability exhibited by the A/H5N1 subtype in recent years, it was necessary to evaluate a diverse set of A/H5N1 influenza viruses. These viruses were collected from human, feline, and a variety of avian hosts during the time period spanning 2003−2006 in varied geographic locations including Vietnam, Nigeria, Indonesia, and Kazakhstan. Six of these samples were from the infection of an Indonesian family in which human-to-human transmission of the virus was suspected.[3] A total of 24 A/H5N1 viral isolates, 3 pathogens that cause influenza-like illness, including the SARS-causing coronavirus, and 4 negative controls, were used to test the MChip assay and associated trained probabilistic neural network. A typical MChip

microarray image for an A/H5N1 sample is shown in Figure 1A. For comparison, representative images of human A/H3N2 and A/H1N1 influenza positive samples are shown in Figure 1C and D, respectively. Figure 1 shows clear visual distinction in the relative fluorescence intensity pattern of "hits" between the different viral subtypes. It is important to note that the patterns shown in Figure 1 are representative images and that within each subtype a certain amount of variability exists in the resulting pattern depending on the particular strain being examined. For example, the H3N2 pattern shown in Figure 1C is similar but not identical to the pattern shown for an H3N2 virus in our previous work.[22] Because the training set includes a wide variety of strains analyzed for a particular subtype, even different strains of the same subtype can be accurately identified in the neural network analysis.

PNN Results. Microarray image data was processed, as described in the Experimental Section, to provide *relative* fluorescence intensity values for each triplicate set of spots to the trained PNN for identification. The probability distribution for each of the four output categories is summarized in Table 2. While the study was not conducted "blind" (i.e., the viral subtype was known to the users), these samples were "unknown" to the trained neural network (i.e., they were not used for training or validating the PNN).

As observed in Table 2, 21 of the 24 A/H5N1 samples were correctly assigned as A/H5N1 with ≥95% probability. Only one virus was assigned as A/H5N1 with less than 90% certainty (no. 17 = 0.83). Furthermore, six of the seven negative controls were correctly assigned with a probability of ≥98%. If 80% certainty is used as the cutoff value for positive identification, all viruses were correctly identified, resulting in both clinical sensitivity and specificity of 100%.[30] It is interesting to note that the most recent examples of A/H5N1 from the family cluster in Indonesia were all correctly identified despite controversial reports of significant genetic mutation.[29] Significantly, there were no false positive or false negative assignments. If the probability threshold is raised to 95%, 4 samples out of the 31 were incorrectly identified and the corresponding clinical sensitivity and specificity would be 88 and 100%, respectively. It should be noted that sample 30 was omitted from the 95% threshold calculation, since it does not constitute a false negative in terms of the clinical sensitivity.

PNN with H3N2, H1N1, and H5N1 Subtypes. In order to evaluate how well the PNN performed when queried with a set of unknowns composed of the two human-adapted influenza A virus subtypes (H3N2 and H1N1) as well as the H5N1 subtype, additional influenza A positive samples were pooled with the H5N1 set and subsequently tested. The results from this study, in which all four categories were represented, are summarized in Table 3. Only 2 of the 47 samples were not definitively identified (samples 16 and 40). All of the other samples were correctly identified with >95% certainty. The resulting clinical sensitivity was 97% with 100% specificity. Again, note that sample 40 was not included in this sensitivity calculation as it does not meet the definition of a false negative.

Other Subtypes. Our efforts to date have focused on the human-adapted influenza A virus subtypes (H3N2 and H1N1), as they are both currently circulating in humans, and on the H5N1 subtype of influenza A due to its pandemic potential. In addition, however, samples representing a variety of other subtypes have also been examined. Several example images are shown in Figure 2 representing a variety of avian viruses (H7N3, H7N2, H10N7, H9N2, H10N?) and an equine H3N8 virus, all of which gave different and unique relative fluorescence patterns. For example, sequence 9 is positive in Figure 2B, C, and F, but negative for Figure 2A, D, and E. These results indicate that the differences in the microarray relative fluorescence intensity patterns observed for the H3N2, H1N1, and H5N1 samples may be extended to other subtypes as well. It should be noted that human patient samples positive for influenza A subtypes other than H3N2 and H1N1 are extremely rare, the H5N1 subtype being of greatest concern recently with less than 300 cases worldwide since 2003. However, the neural network need not be trained to recognize all possible subtypes. It is envisioned that the MChip analysis would be used simply to address the following critical questions: (1) is a patient infected with influenza A virus, and (2) if so, is the virus a typical human-adapted subtype or is it atypical? In a real-world application, a human sample that tests positive for influenza A but is not identified as either of the currently circulating subtypes (H3N2 or H1N1) would be cause for immediate concern and would direct health officials to a more detailed investigation of the sample.

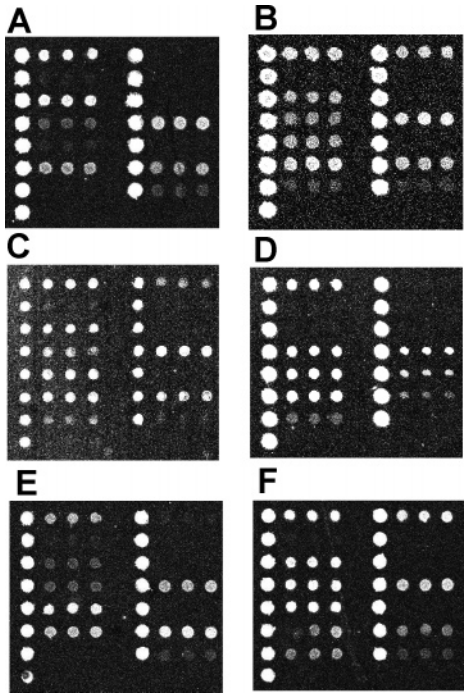

Figure 2. MChip results for other subtypes: (A) equine H3N8, (B) avian H10N?, (C) avian H7N3, (D) avian H7N2, (E) avian H10N7, and (F) avian H9N2. Microarray layout is given in Figure 1.

Summary. The MChip assay is a simple, inexpensive, single-gene microarray diagnostic that exhibited 97% sensitivity and 100% specificity for identification of influenza A H5N1, H3N2, and H1N1 viruses. The primary analytical advantage of this new method, which includes automated data interpretation, over current influenza diagnostics is its ability to simultaneously screen and provide full subtype information for a variety of influenza A viruses,

(29) Butler, D. *Nature* **2006**, *442*, 114−115.
(30) Loong, T.-W. *BMJ (Br. Med. J. Clin. Res. Ed.)* **2003**, *327*, 716−719.

including a rapidly emerging influenza virus such as avian A/H5N1.

ACKNOWLEDGMENT

The authors gratefully acknowledge funding from NIH/NIAID (U01AI056528). E.D.D. acknowledges partial support from NSF. The authors also acknowledge Dr. Triono with the Indonesian Ministry of Health for selected Indonesian A/H5N1 samples used to provide the viral isolates utilized in this study.

Received for review October 10, 2006. Accepted October 27, 2006.

AC061920O