

Phylogeography of the Spring and Fall Waves of the H1N1/09 Pandemic Influenza Virus in the United States^{∇§}

Martha I. Nelson,^{1*} Yi Tan,¹ Elodie Ghedin,^{2,3} David E. Wentworth,^{4,5} Kirsten St. George,⁴ Laurel Edelman,⁶ Eric T. Beck,⁷ Jiang Fan,⁷ Tommy Tsan-Yuk Lam,⁸ Swati Kumar,⁷ David J. Spiro,³ Lone Simonsen,^{1,9} Cecile Viboud,¹ Edward C. Holmes,^{1,8} Kelly J. Henrickson,^{7¶} and James M. Musser^{10¶}

Fogarty International Center, National Institutes of Health, Bethesda, Maryland¹; Center for Vaccine Research, Department of Computational and Systems Biology, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania²; J. Craig Venter Institute, Rockville, Maryland³; Wadsworth Center, New York State Department of Health, Albany, New York⁴; Department of Biomedical Sciences, School of Public Health, State University of New York, Albany, New York⁵; Surveillance Data Inc., Plymouth Meeting, Pennsylvania⁶; Medical College of Wisconsin, Children's Hospital of Wisconsin, Milwaukee, Wisconsin⁷; Department of Biology, the Pennsylvania State University, University Park, Pennsylvania⁸; Department of Global Health, School of Public Health and Health Services, George Washington University, Washington, DC⁹; and Center for Molecular and Translational Human Infectious Diseases Research, the Methodist Hospital Research Institute, and Department of Pathology and Laboratory Medicine, the Methodist Hospital, Houston, Texas¹⁰

Received 20 August 2010/Accepted 27 October 2010

Spatial variation in the epidemiological patterns of successive waves of pandemic influenza virus in humans has been documented throughout the 20th century but never understood at a molecular level. However, the unprecedented intensity of sampling and whole-genome sequencing of the H1N1/09 pandemic virus now makes such an approach possible. To determine whether the spring and fall waves of the H1N1/09 influenza pandemic were associated with different epidemiological patterns, we undertook a large-scale phylogeographic analysis of viruses sampled from three localities in the United States. Analysis of genomic and epidemiological data reveals distinct spatial heterogeneities associated with the first pandemic wave, March to July 2009, in Houston, TX, Milwaukee, WI, and New York State. In Houston, no specific H1N1/09 viral lineage dominated during the spring of 2009, a period when little epidemiological activity was observed in Texas. In contrast, major pandemic outbreaks occurred at this time in Milwaukee and New York State, each dominated by a different viral lineage and resulting from strong founder effects. During the second pandemic wave, beginning in August 2009, all three U.S. localities were dominated by a single viral lineage, that which had been dominant in New York during wave 1. Hence, during this second phase of the pandemic, extensive viral migration and mixing diffused the spatially defined population structure that had characterized wave 1, amplifying the one viral lineage that had dominated early on in one of the world's largest international travel centers.

Following its initial description in March 2009, the pandemic H1N1/09 virus rapidly spread globally and was detected in at least 208 countries by December 2009. In the United States, an estimated 7,500 to 44,100 deaths were attributable to the H1N1/09 virus between May and December 2009 (21), so that the >18,000 global laboratory-confirmed human deaths that have been directly linked to the virus (2) are clearly a huge underestimate of the true number (21, 25). While the disease burden associated with the H1N1/09 pandemic is relatively low compared to that of past pandemics, most notably the devastating “Spanish flu” of 1918, the H1N1/09 virus has disproportionately impacted younger age groups, resulting in substantial life years lost (25).

The H1N1/09 virus is a novel reassortant, containing the PB2, PB1, PA, HA, NP, and NS segments from North American triple-reassortant swine viruses and NA and M segments derived from the Eurasian swine lineage (7). Phylogenetic analysis suggests that these segments have been circulating undetected in swine for a decade or more but have emerged in humans only recently, perhaps only several months prior to their initial detection in humans (22). Although phylogenetic resolution of the H1N1/09 virus is difficult because of the short duration of evolution in humans, seven discrete viral clades were identified globally during the first pandemic wave in the spring of 2009 (16).

The epidemiology of the three influenza pandemics of the 20th century (1918 [H1N1], 1957 [H2N2], and 1968 [H3N2]) has been studied intensively, particularly the variation in mortality and transmissibility among countries and cities and between successive waves (1, 6, 18). However, the viral gene sequence data acquired from these past pandemics has been insufficient to precisely reveal epidemiological patterns and dynamics, especially at the level of individual cities. Advances in the speed of genome sequencing technologies and increas-

* Corresponding author. Mailing address: Fogarty International Center, National Institutes of Health, 16 Center Drive, Building 16, Room 202, Bethesda, MD 20892. Phone: (301) 402-5203. Fax: (301) 496-8496. E-mail: nelsonma@mail.nih.gov.

§ Supplemental material for this article may be found at <http://jvi.asm.org/>.

¶ These authors contributed equally to this work.

∇ Published ahead of print on 10 November 2010.

ing public data access mean that the H1N1 influenza A pandemic of 2009 represents the first time that the epidemiological dynamics of multiple pandemic waves can be characterized at the molecular level and for individual localities.

Despite the wealth of individual gene and complete genome sequence data available for H1N1/09, it is unknown how the molecular epidemiology of this virus varies spatially across the United States and between the first spring pandemic wave and the second fall wave. To determine whether the spring and fall waves were associated with different spatial patterns, we conducted a phylogenetic analysis of whole-genome H1N1/09 sequences collected from 1 April 2009 to 3 March 2010 from three intensively sampled localities in the United States: Houston, TX (Southwestern United States), Milwaukee, WI (Midwestern United States), and New York State (Northeastern United States), against a background of H1N1/09 sequences collected from other areas in North America, as well as globally. We also linked the observed evolutionary patterns with epidemiological data collected from the same localities over this time period.

MATERIALS AND METHODS

Epidemiological data. SDI Health, LLC (SDI) maintains a sample of hospital encounters, including emergency department, short-stay (<23 h), and inpatient events, derived from approximately 550 hospitals, which covers about 20% of such events in the United States. Data analyzed included discharge diagnoses coded for influenza according to the International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM; <http://www.cdc.gov/nchs/icd/icd9cm.htm>).

Sample collection, viral isolation, and RNA extraction. Virus specimens from Houston were collected from patient specimens processed at four Methodist hospitals at geographically dispersed sites in the metropolitan area. Patient specimens from Milwaukee were collected at Children's Hospital of Wisconsin (CHW), Dynacare Laboratories (DL), and their associated clinics and were tested for the presence of pandemic H1N1/09 virus using reverse transcription (RT)-PCR assays described previously (4, 10, 16). Specimens in New York State were collected from patients in primary care clinics and hospitals throughout the state and submitted to the Wadsworth Center for H1N1/09 testing using the real-time RT-PCR assay previously described (17). Extraction and amplification of the entire viral RNA genome from samples from Milwaukee and New York State were performed at the Wadsworth Center, New York State Department of Health, in Albany, NY, using methods described previously (16, 26).

Sequencing of viral RNA genome. Whole-genome sequencing was performed at the J. Craig Venter Institute (JCVI) in Rockville, MD, using methods described previously (16). Oligonucleotide primers were designed using a computational PCR primer design pipeline developed at JCVI (13). Sequencing, genome assembly, and closure reactions were performed as described previously (8). All sequences were submitted to GenBank and are publicly available at the National Center for Biotechnology Information (NCBI) Influenza Virus Resource (<http://www.ncbi.nlm.nih.gov/genomes/FLU/FLU.html>) (3).

Phylogenetic analysis. The sequence data used in this study came from 1,312 pandemic H1N1/09 sequences collected in North America (Canada, Mexico, and the United States) between 1 April 2009 and 3 March 2010, which were downloaded from the National Center for Biotechnology Information (NCBI) Influenza Virus Resource (<http://www.ncbi.nlm.nih.gov/genomes/FLU/FLU.html>) available on GenBank (3). This data set includes 226 H1N1/09 sequences collected from Houston, 319 sequences from Milwaukee, and 343 sequences from New York State. Only viruses for which whole-genome sequences and the exact dates of collection were available were used.

Nucleotide alignments were manually constructed for the coding regions of the eight genome segments of each sequence using the Se-AI program (20). To maximize phylogenetic resolution, and given no evidence of genomic reassortment, the eight segments were concatenated. To infer the evolutionary relationships for the complete H1N1/09 data set analyzed here, we employed the maximum likelihood (ML) method available in the PhyML program (9). Because of the great similarity of the sequences, such that multiple substitutions at single nucleotide sites can effectively be ignored, this analysis utilized the simple HKY85 model of nucleotide substitution and employed SPR branch-swapping.

Support for major clusters on the tree was provided by the approximate likelihood ratio test (aLRT) available in PhyML. Two additional phylogenetic trees were inferred: one including sequences that were collected only during wave 1, defined here as 1 April 2009 to 31 July 2009 ($n = 674$ sequences), and a second including sequences that were collected during wave 2, defined here as 1 August 2009 to 3 March 2010 ($n = 638$ sequences). The FigTree program (<http://tree.bio.ed.ac.uk/software/figtree/>) was used to annotate each phylogeny by coloring tip names by geographical location. H1N1/09 clades were identified based on strong support for individual nodes and long branch lengths, adopting the nomenclature and clade delineations defined previously (16). Some clades that were identified previously were not as clearly defined here, particularly clades 2 and 3. However, due to their small size, these clades are not important to our study.

To provide a global context for our North American study, an ML tree was inferred for a subset of the North American data plus 282 whole-genome sequences that were collected globally (Europe, Asia, Australia, and South and Central America) and available on GenBank (total of 1,356 sequences) (3).

Spatial analysis. To determine the extent and pattern of geographic structure in the wave 1 and wave 2 H1N1/09 populations, we assigned each sequence in the data set a single-digit character state (1, 2, 3, or 4) reflecting the locality of origin (Houston, Milwaukee, New York State, or other North American locality, respectively). Because of the very large data set available for analysis, computational constraints meant that we were forced to base our analysis on a single representation of epidemiological history. Accordingly, the minimum number of changes in character state needed to produce the observed distribution of character states on the ML tree for these data (described above) was then estimated by using the parsimony method in PAUP* (23). To determine the number expected under the null hypothesis of random mixing among North American localities, the character states of all sequences were randomized 10,000 times and the analysis was repeated. This analysis was conducted separately for the wave 1 and wave 2 data sets.

Nucleotide diversity. The average pairwise nucleotide diversity among sequences was estimated for six H1N1/09 virus populations: Houston (wave 1), Houston (wave 2), New York State (wave 1), New York State (wave 2), Milwaukee (wave 1), and Milwaukee (wave 2). For this analysis we employed a maximum likelihood (ML) method available in the MEGA4 package (24), with the standard deviation estimated using 100 bootstrap replicates.

RESULTS

Epidemiological analysis. Among the three intensively sampled U.S. localities of Houston, TX, Milwaukee, WI, and New York State, we detect marked heterogeneities in the epidemiology of the pandemic H1N1/09 virus in the first wave during the spring of 2009. Based upon emergency department visits coded as influenza, both Milwaukee and New York State experienced severe pandemic outbreaks during April to June 2009 (Fig. 1), observations that have been described previously (11, 12, 19). In contrast, no increase in influenza virus activity was detectable during the spring of 2009 in Houston (Fig. 1), despite its relative geographic proximity to Mexico (<400 miles), where the H1N1/09 pandemic is thought to have originated. Subsequently, Houston experienced a major increase in H1N1/09 activity beginning in August 2009, marking the beginning of the second pandemic wave, and New York State and Milwaukee experienced second pandemic waves approximately 1 month later. Based on these data, we defined two temporally distinct waves of the pandemic as follows: wave 1 is defined as April to July 2009, while wave 2 is defined as August 2009 to March 2010.

Phylogenetic analysis. To understand the evolutionary basis for these geographical differences in pandemic activity, we conducted a phylogenetic analysis on a collection of whole-genome sequences from 1,312 pandemic H1N1/09 viruses collected in North America from 1 April 2009 to 3 March 2010. This data set comprises 424 "background" sequences collected

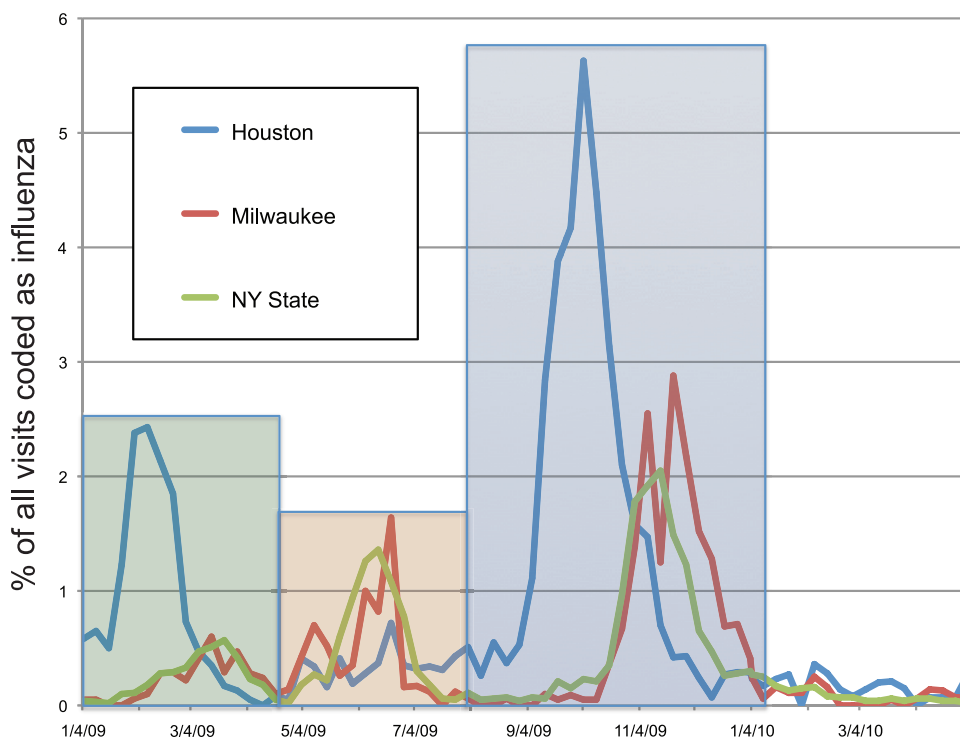


FIG. 1. Proportion of all emergency department visits that are coded as influenza in Houston, Milwaukee, and New York State, 2009 to 2010. Data collected by SDI, Plymouth Meeting, PA. Peaks in influenza virus activity associated with the 2008–2009 seasonal H1N1 epidemic (shaded green), the spring wave of pandemic H1N1/09 viruses (shaded red), and the fall wave of pandemic H1N1/09 viruses (shaded blue) are depicted.

from other parts of the United States, Canada, and Mexico and 888 sequences from three intensively sampled U.S. localities: 226 H1N1/09 sequences collected from Houston, 319 sequences from Milwaukee, and 343 sequences from New York State. These data include both the spring pandemic wave (wave 1) (51% of all sequences) and the subsequent fall wave (wave 2) (49%) (Table 1).

Three maximum-likelihood (ML) phylogenetic trees were inferred from these data: (i) wave 1 ($n = 674$ sequences), (ii) wave 2 ($n = 638$ sequences), and (iii) all data ($n = 1,312$ sequences) (Fig. 2 and 3 and Fig. S1 in the supplemental material, respectively). As a reference, a global ML tree also was inferred from a subset of these North American sequences and 282 whole-genome sequences from isolates collected in Asia, Europe, Australia, and South and Central America (see Fig. S2 in the supplemental material). Given (i) the unprecedented intensity of temporal sampling of this viral population,

(ii) how recently the H1N1/09 emerged in humans—estimated to be January or February 2009 (22)—and (iii) its rapid spread and phylogenetic diversification, there was predictably a lack of strong statistical support for large-scale phylogenetic clades. However, for the purposes of estimating the number of viral introductions into a specific locality and for describing overall changes in viral population structure that occurred between waves 1 and 2, we employ the clade nomenclature previously defined by Nelson et al. (16). Of the seven clades that previously were identified globally (16), six were detectable in our phylogenies of North American H1N1/09 sequences: clade 1, clade 2, clade 3, clade 5, clade 6, and clade 7 (Fig. 2 and 3). Clade 4, which has been detected only in Asia (16) and is exceptionally small (11 sequences; see Fig. S2 in the supplemental material), was not identified in this North American viral population.

During wave 1 in New York State, sequences were collected

TABLE 1. Summary of sample collection^a

Locality	Wave 1		Wave 2		Total no. of sequences
	No. of sequences	Collection dates (mo/day/yr)	No. of sequences	Collection dates (mo/day/yr)	
Houston	82	04/28/09-07/29/09	144	08/15/09-02/06/10	226
Milwaukee	183	04/28/09-06/30/09	136	09/07/09-12/18/09	319
New York State	193	04/24/09-7/30/09	150	08/06/09-03/01/10	343
North America	216	04/01/09-07/27/09	208	08/05/09-02/13/10	424
Total	674		638		1,312

^a Shown are the number of sequences collected from Houston, Milwaukee, New York State, and other localities in Canada, the United States, and Mexico that are used in Fig. 2 and 3 and in Fig. S1 in the supplemental material.

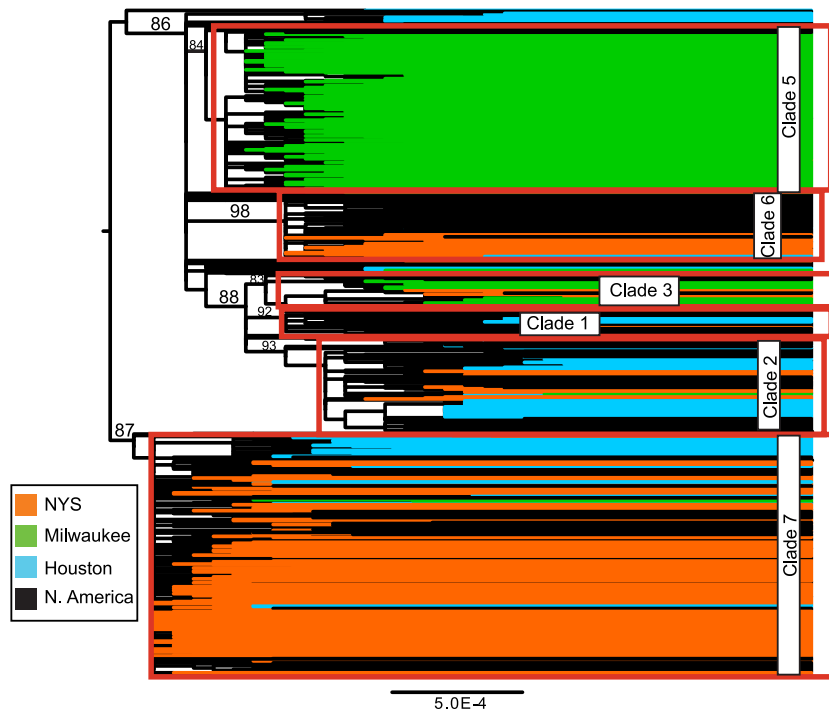


FIG. 2. Evolutionary relationships of the concatenated whole-genome sequences from 674 sequences of pandemic H1N1/09 viruses collected in North America during the first pandemic wave, from 1 April 2009 to 29 July 2009. The maximum likelihood (ML) tree is midpoint rooted for clarity, with all branch lengths drawn to scale. Numbers on branches indicate approximate likelihood ratio test values and provide an indication of statistical support for individual nodes. Sequences shaded in orange were collected in New York State; sequences shaded in green were collected in Milwaukee; sequences shaded in blue were collected in Houston; sequences shaded in black were collected in other North American localities.

from all six clades identified in North America (Fig. 2). Of these, clade 7 rapidly became dominant in New York during wave 1, representing 78% of all New York sequences (151/193), presumably reflecting a strong founder effect. In Milwaukee, the sequences collected during wave 1 were members of four clades (clade 2, clade 3, clade 5, and clade 7), indicating a minimum of four separate introductions. During Milwaukee's first pandemic wave, clade 5 viruses rapidly became dominant, representing 87% of all sequences collected (160/183), again indicative of a strong founder effect. At least four viral introductions occurred during wave 1 in Houston, based on the identification of clade 1, clade 2, clade 6, and clade 7 in this population. In addition, three small phylogenetically distinct clades that were not identified previously (16) contain sequences from Houston, which likely represent further introductions into this locality. Unlike the situations in New York State and Milwaukee, no single viral lineage appears to have been dominant during wave 1 in Houston, with what appears to be relatively low levels of onward transmission of each viral introduction.

A strikingly different pattern of clade structure was identified during wave 2. In particular, clade 7 was dominant in all three localities, as well as in North America as a whole (Fig. 3) and globally (see Fig. S2 in the supplemental material). Indeed, we observed that nearly 95% of sequences (605 of 638) from wave 2 belong to clade 7, with all other clades shrinking or disappearing entirely, as in the case of clade 1 and, most notably, clade 5 from Milwaukee, even at a global scale (Fig. S2). While the highly interspersed pattern of sequences from

Houston, Milwaukee, and New York State during wave 2 signifies that there were frequent introductions of clade 7 into these localities (Fig. 3), estimating the exact number of introductions into each locality is extremely difficult with the current data because of the lack of strongly supported phylogenetic clusters.

Spatial analysis. Our parsimony-based analysis of geographic structure provides a quantitative measure of the differences in population structure that we observed for the trees inferred for waves 1 and 2 (Fig. 2 and 3, respectively). This analysis reveals a significant geographically based population structure in both waves ($P < 0.0001$), which is expected given the highly localized scale of sampling and the nature of influenza virus transmission. However, the proportion of location changes required by parsimony to generate the observed phylogenetic pattern (Fig. 2 and 3), versus 10,000 randomized trees, is closer to 0 in wave 1 (range, 0.20 to 0.24) than in wave 2 (range, 0.40 to 0.45), indicating that the wave 1 population is more genetically structured by location. This is presumably because the H1N1/09 virus had undergone substantial geographical mixing by the fall of 2009, while wave 1 was still dominated by the initial founder effects.

Nucleotide diversity. Our analysis of pairwise nucleotide diversity within each of these three viral populations during waves 1 and 2 further supports the phylogenetic results described above. Most notably, the nucleotide diversity of the H1N1/09 populations in New York State and Milwaukee is approximately half that of Houston during wave 1, reflecting the strong founder effects observed in the former populations

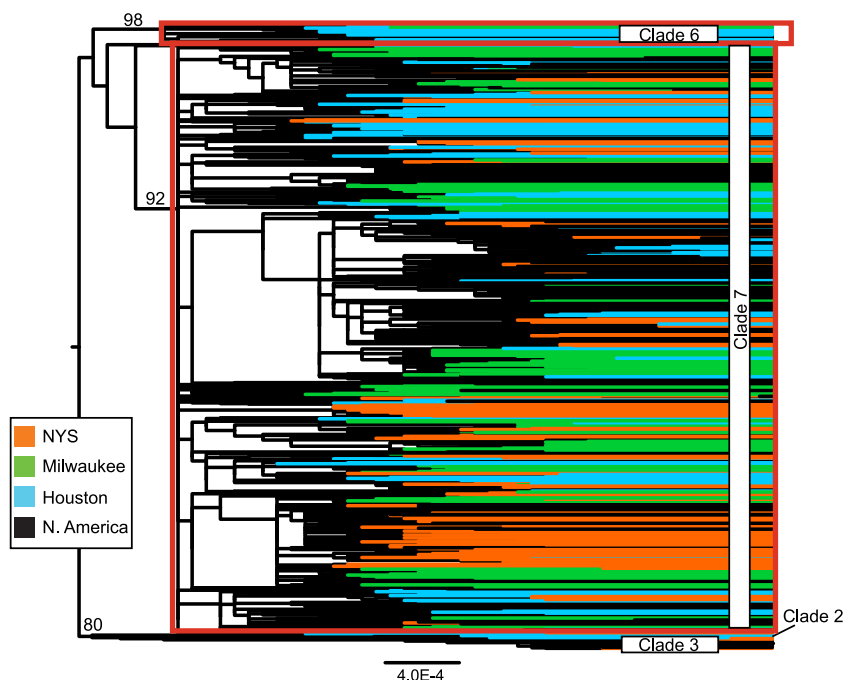


FIG. 3. Evolutionary relationships of the concatenated whole-genome sequences from 638 sequences of pandemic H1N1/09 viruses collected in North America during the second pandemic wave, from 5 August 2009 to 3 March 2010. All rooting, labels, and color coding are identical to those used in Fig. 2.

(clade 7 and clade 5, respectively) (Table 2). In contrast, all three localities exhibit similar levels of genetic diversity during wave 2, reflecting the greater degree of mixing at this time. Although fewer clades circulate during wave 2, with clade 7 strongly dominant, more genetic diversity is observed in each of these localities during the second wave due to the genetic diversification of clade 7 through time and the extensive geographical mixing of the viruses within clade 7.

DISCUSSION

An intensified sampling of the H1N1/09 virus in three defined U.S. localities reveals a strong geographically defined viral population structure, particularly within wave 1, followed by geographic dispersal in wave 2. Separate analyses of phylogenetic relationships, the strength of spatial structure, and nucleotide diversity within spatiotemporally defined groups consistently support this conclusion. Furthermore, the heterogeneities in virus population structure during wave 1 (i.e.,

multiple circulating clades) reflect the strong heterogeneities in the epidemiology of the pandemic in the three studied localities.

Our results therefore suggest that wave 1 and wave 2 represent two distinct phases in the molecular epidemiology of the H1N1/09 pandemic: a first wave dominated by strong founder effects in areas of high disease transmission, resulting in distinct spatial heterogeneities, and a second wave characterized by extensive global migration and increased spatial mixing, resulting in the global dominance of a single lineage. Although multiple introductions of the virus occur in both waves, founder effects appear to be particularly important in localities that experience a major first wave, resulting in point-source outbreaks of a single dominant lineage: clade 5 in Milwaukee and clade 7 in New York State. The dominance of clade 5 in Milwaukee is especially noteworthy, given the very low levels of this clade that are observed in other sampled North American localities and the global disappearance of clade 5 during wave 2.

Although Houston did not experience a major pandemic wave in the spring of 2009, the genetic diversity of wave 1 is higher in Houston than in Milwaukee or New York State, owing to (i) multiple viral introductions, (ii) the absence of a single dominant lineage, and (iii) perhaps the geographical proximity of the city to Mexico, where the pandemic is thought to have originated, based upon the basal phylogenetic position of Mexican sequences on the tree. Clearly these findings indicate that the absence of a major spring wave in Houston is not due to any diminished importation or initial presence of the H1N1/09 virus in the city but rather to the fact that no single viral introduction resulted in high onward transmission. Earlier termination of school terms may have reduced pandemic ac-

TABLE 2. Nucleotide diversity of H1N1/09 virus populations^a

Locality/wave	Nucleotide diversity (base substitutions per site)	SD
Houston (wave 1)	1.72×10^{-3}	2.2×10^{-4}
Milwaukee (wave 1)	8.9×10^{-4}	1.4×10^{-4}
New York State (wave 1)	8.9×10^{-4}	1.1×10^{-4}
Houston (wave 2)	3.04×10^{-3}	1.4×10^{-4}
Milwaukee (wave 2)	2.73×10^{-3}	1.3×10^{-4}
New York State (wave 2)	2.94×10^{-3}	1.1×10^{-4}

^a Shown are data for collections in Houston, Milwaukee, and New York State during wave 1 (March to July 2009) and wave 2 (August 2009 to March 2010) of the pandemic.

tivity during the spring in Southern U.S. regions (generally late May, compared with mid-June in Northern U.S. regions), a hypothesis that certainly warrants further study. School cycles have been shown to affect the timing of the onset of the fall 2009 pandemic wave in the United States (5), but the role of school term timing in the intensity of the spring wave in different regions remains unclear.

Finally, it is tempting to propose that the global dominance of clade 7 during the second pandemic wave may relate to its initial foothold in New York State during the first wave, facilitating its rapid global spread via New York City's high international interconnectivity. In contrast, we detect no evidence of continued transmission during the second pandemic wave of clade 5, following its dominance during wave 1 in Milwaukee. Given the lack of amino acid changes or any likely differences in fitness between clade 5 and clade 7 (16), the differential success of clade 7 during wave 2 most likely is attributable to stochastic differences (i.e., founder effects) in the early geographical patterning of these lineages during wave 1.

Although those H1N1/09 viruses characterized to date are very similar at both the genetic and antigenic levels (such that it is very difficult to detect genomic reassortment events involving viruses from different clades), the intensity of the sampling of the early dissemination of the H1N1/09 virus in the United States allows us to detect a strong geographical structure in human influenza at the scale of the United States for the first time. This is in marked contrast to previous large-scale phylogenetic analyses of seasonal influenza virus epidemics which have not detected strong spatial patterns at the scale of the United States (15) or within New York State (14), due to the cocirculation of multiple divergent lineages and likely differences in sampling. It remains unknown whether the spatial patterns we observe during the first pandemic wave are unique to these specific pandemic circumstances or could be detected during seasonal influenza epidemics with greatly increased sampling in discrete localities.

ACKNOWLEDGMENTS

This project has been funded in part with federal funds from the National Institute of Allergy and Infectious Diseases (NIAID), National Institutes of Health, Department of Health and Human Services, under contract number HHSN272200900007C and by NIAID grants UO1-AI070428, U01-387 AI077988, and U01-AI066584. The whole-genome sequencing of viral samples from Houston, Milwaukee, and New York State was funded by the Influenza Genome Sequencing Project, administered by NIAID/NIH. This research was conducted in the context of the MISMS study, an ongoing international collaborative effort to understand influenza epidemiological and evolutionary patterns, led by the Fogarty International Center, National Institutes of Health (www.origem.info/misms). Funding for the MISMS project comes in part from the Office of Global Health Affairs' International Influenza Unit in the Office of the Secretary of the Department of Health and Human Services.

We thank Sue Kehl, Nate Ledebouer, Ruoyan Chen, Jessica Trost, Teresa Patitucci, Lorraine Witt, Meredith VanDyke, Elizabeth Davis, and Kate Gaffney from the MRVP and Medical College of Wisconsin, Sara Griesemer at the Wadsworth Center, NYSDOH, Joshua Cherry from the National Center for Biotechnology Information (NCBI/NIH), and Patricia Cernoch and Randall J. Olsen of the Methodist Hospital System for their help in this study. We are also greatly indebted to all of those who submitted H1N1/09 virus sequences to GenBank's Influenza Virus Resource. The Madin-Darby canine kidney (MDCK) cells used for viral isolation were kindly provided by

Xiyang Xu at the Centers for Disease Control and Prevention (Atlanta, GA).

REFERENCES

- Andreasen, V., C. Viboud, and L. Simonsen. 2008. Epidemiological characterization of the 1918 influenza pandemic summer wave in Copenhagen: implications for pandemic control strategies. *J. Infect. Dis.* **197**:270–278.
- Anonymous. 2010. World Health Organization pandemic (H1N1) 2009—update 111. Accessed 3 August 2010. http://www.who.int/csr/don/2010_7_30/en/index.html.
- Bao, Y., P. Bolotov, D. Dernovoy, B. Kiryutin, L. Zaslavsky, T. Tatusova, J. Ostell, and D. Lipman. 2008. The influenza virus resource at the National Center for Biotechnology Information. *J. Virol.* **82**:596–601.
- Bose, M. E., E. T. Beck, N. Ledebouer, S. C. Kehl, L. A. Jurgen, T. Patitucci, L. Witt, E. LaGue, P. Darga, J. He, J. Fan, S. Kumar, and K. J. Henrickson. 2009. Rapid semiautomated subtyping of influenza virus species during the 2009 swine origin influenza A H1N1 virus epidemic in Milwaukee, Wisconsin. *J. Clin. Microbiol.* **47**:2779–2786.
- Chao, D. L., M. E. Halloran, and I. M. Longini, Jr. 2010. School opening dates predict pandemic influenza A(H1N1) outbreaks in the United States. *J. Infect. Dis.* **202**:877–880.
- Chowell, G., C. Viboud, L. Simonsen, M. A. Miller, and R. Acuna-Soto. 2010. Mortality patterns associated with the 1918 influenza pandemic in Mexico: evidence for a spring herald wave and lack of preexisting immunity in older populations. *J. Infect. Dis.* **202**:567–575.
- Garten, R. J., C. T. Davis, C. A. Russell, B. Shu, S. Lindstrom, A. Balish, W. M. Sessions, X. Xu, E. Skepner, V. Deyde, M. Okomo-Adhiambo, L. Gubareva, J. Barnes, C. B. Smith, S. L. Emery, M. J. Hillman, P. Rivaller, J. Smagala, M. de Graaf, D. F. Burke, R. A. Fouchier, C. Pappas, C. M. Alpuche-Aranda, H. Lopez-Gatell, H. Olivera, I. Lopez, C. A. Myers, D. Faix, P. J. Blair, C. Yu, K. M. Keene, P. D. Dotson, Jr., D. Boxrud, A. R. Sambol, S. H. Abid, K. St. George, T. Bannerman, A. L. Moore, D. J. Stringer, P. Blevins, G. J. Demmler-Harrison, M. Ginsberg, P. Kriner, S. Waterman, S. Smole, H. F. Guevara, E. A. Belongia, P. A. Clark, S. T. Beatrice, R. Donis, J. Katz, L. Finelli, C. B. Bridges, M. Shaw, D. B. Jernigan, T. M. Uyeki, D. J. Smith, A. I. Klimov, and N. J. Cox. 2009. Antigenic and genetic characteristics of swine-origin 2009 A(H1N1) influenza viruses circulating in humans. *Science* **325**:197–201.
- Ghedini, E., N. A. Sengamalay, M. Shumway, J. Zaborsky, T. Feldblyum, V. Subbu, D. J. Spiro, J. Sitz, H. Koo, P. Bolotov, D. Dernovoy, T. Tatusova, Y. Bao, K. St. George, J. Taylor, D. J. Lipman, C. M. Fraser, J. K. Taubenberger, and S. L. Salzberg. 2005. Large-scale sequencing of human influenza reveals the dynamic nature of viral genome evolution. *Nature* **437**:1162–1166.
- Guindon, S., and O. Gascuel. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* **52**:696–704.
- He, J., M. E. Bose, E. T. Beck, J. Fan, S. Tiwari, J. Metallo, L. A. Jurgens, S. C. Kehl, N. Ledebouer, S. Kumar, W. Weisburg, and K. J. Henrickson. 2009. Rapid multiplex reverse transcription-PCR typing of influenza A and B virus, and subtyping of influenza A virus into H1, 2, 3, 5, 7, 9, N1 (human), N1 (animal), N2, and N7, including typing of novel swine origin influenza A (H1N1) virus, during the 2009 outbreak in Milwaukee, Wisconsin. *J. Clin. Microbiol.* **47**:2772–2778.
- Kumar, S., M. J. Chusid, R. E. Willoughby, P. L. Havens, S. C. Kehl, N. A. Ledebouer, P. Simpson, M. Vandyske, E. Davis, K. Gaffney, S. H. Li, M. E. Bose, and K. J. Henrickson. 2010. Epidemiologic observations from passive and targeted surveillance during the first wave of the 2009 H1N1 influenza pandemic in Milwaukee, WI. *Viruses* **2**:782–795.
- Lessler, J., N. G. Reich, D. A. Cummings, and New York City Department of Health and Mental Hygiene Swine Influenza Investigation Team. 2009. Outbreak of 2009 pandemic influenza A (H1N1) at a New York City school. *N. Engl. J. Med.* **361**:2628–2636.
- Li, K., A. Brownley, T. B. Stockwell, K. Beeson, T. C. McIntosh, D. Busam, S. Ferreira, S. Murphy, and S. Levy. 2008. Novel computational methods for increasing PCR primer design effectiveness in directed sequencing. *BMC Bioinformatics* **9**:191.
- Nelson, M. I., L. Simonsen, C. Viboud, M. A. Miller, J. Taylor, K. St. George, S. B. Griesemer, E. Ghedin, N. A. Sengamalay, D. J. Spiro, I. Volkov, B. T. Grenfell, D. J. Lipman, J. K. Taubenberger, and E. C. Holmes. 2006. Stochastic processes are key determinants of short-term evolution in influenza A virus. *PLoS Pathog.* **2**:e125.
- Nelson, M. I., L. Edelman, D. J. Spiro, A. R. Boyne, J. Bera, R. Halpin, N. Sengamalay, E. Ghedin, M. A. Miller, L. Simonsen, C. Viboud, and E. C. Holmes. 2008. Molecular epidemiology of A/H3N2 and A/H1N1 influenza virus during a single epidemic season in the United States. *PLoS Pathog.* **4**:e1000133.
- Nelson, M. I., D. J. Spiro, D. Wentworth, E. T. Beck, J. Fan, E. Ghedin, R. Halpin, J. Bera, E. Hine, K. Proudfoot, T. B. Stockwell, X. Lin, S. Griesemer, S. Kumar, M. Bose, C. Viboud, E. C. Holmes, and K. J. Henrickson. 2009. The early diversification of influenza A/H1N1pdm. *PLoS Curr.* **1**:RRN1126.
- Novel Swine-Origin Influenza A (H1N1) Virus Investigation Team, F. S. Dawood, S. Jain, L. Finelli, M. W. Shaw, S. Lindstrom, R. J. Garten, L. V.

- Gubareva, X. Xu, C. B. Bridges, and T. M. Uyeki. 2009. Emergence of a novel swine-origin influenza A (H1N1) virus in humans. *N. Engl. J. Med.* **360**:2605–2615.
18. Olson, D. R., L. Simonsen, P. J. Edelson, and S. S. Morse. 2005. Epidemiological evidence of an early wave of the 1918 influenza pandemic in New York City. *Proc. Natl. Acad. Sci. U. S. A.* **102**:11059–11063.
19. Presanis, A. M., D. De Angelis; New York City Swine Flu Investigation Team, A. Hagy, C. Reed, S. Riley, B. S. Cooper, L. Finelli, P. Biedrzycki, and M. Lipsitch. 2009. The severity of pandemic H1N1 influenza in the United States, from April to July 2009: a Bayesian analysis. *PLoS Med.* **6**:e1000207.
20. Rambaut, A. 2002. Sequence alignment editor. <http://tree.bio.ed.ac.uk/software/seal/>.
21. Reed, C., and J. M. Katz. 2010. Serological surveys for 2009 pandemic influenza A H1N1. *Lancet* **375**:1062–1063.
22. Smith, G. J., D. Vijaykrishna, J. Bahl, S. J. Lycett, M. Worobey, O. G. Pybus, S. K. Ma, C. L. Cheung, J. Raghwani, S. Bhatt, J. S. Peiris, Y. Guan, and A. Rambaut. 2009. Origins and evolutionary genomics of the 2009 swine-origin H1N1 influenza A epidemic. *Nature* **459**:1122–1125.
23. Swofford, D. L. 2003. PAUP*: phylogenetic analysis using parsimony (*and other methods) version 4.0. Sinauer Associates, Sunderland, MA.
24. Tamura, K., J. Dudley, M. Nei, and S. Kumar. 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* **24**:1596–1599.
25. Viboud, C., M. Miller, D. Olson, M. Osterholm, and L. Simonsen. 2010. Preliminary estimates of mortality and years of life lost associated with the 2009 A/H1N1 pandemic in the US and comparison with past influenza seasons. *PLoS Curr.* **2**:RRN1153.
26. Zhou, B., M. E. Donnelly, D. T. Scholes, K. St. George, M. Hatta, Y. Kawaoka, and D. E. Wentworth. 2009. Single-reaction genomic amplification accelerates sequencing and vaccine production for classical and swine origin human influenza A viruses. *J. Virol.* **83**:10309–10313.