Multiyear Persistence of 2 Pandemic A/H1N1 Influenza Virus Lineages in West Africa

Martha I. Nelson,[1] Richard Njouom,[2] Cecile Viboud,[1] Mbayame N. D. Niang,[3] Hervé Kadjo,[4] William Ampofo,[5] Adedeji Adebayo,[6] Zekiba Tarnagda,[7] Mark A. Miller,[1] Edward C. Holmes,[1,8] and Ousmane M. Diop[3]

[1]Fogarty International Center, National Institutes of Health, Bethesda, Maryland; [2]National Influenza Center, Centre Pasteur du Cameroun, Yaoundé, Cameroon; [3]Institut Pasteur de Dakar, Senegal; [4]National Influenza Center, Institut Pasteur de Côte d'Ivoire, Abidjan Côte d'Ivoire; [5]Noguchi Memorial Institute for Medical Research, University of Ghana, Legon, Accra Ghana; [6]CDC-Nigeria, Federal Ministry of Health, Abuja, Nigeria; [7]Institut de Recherche en Sciences de la Santé, Bobo-Dioulasso, Burkina Faso; and [8]Marie Bashir Institute for Infectious Diseases and Biosecurity, School of Biological Sciences and Sydney Medical School, University of Sydney, New South Wales, Australia

Our understanding of the global ecology of influenza viruses is impeded by historically low levels of viral surveillance in Africa. Increased genetic sequencing of African A/H1N1 pandemic influenza viruses during 2009–2013 revealed multiyear persistence of 2 viral lineages within West Africa, raising questions about the roles of reduced air traffic and the asynchrony of seasonal influenza epidemics among West African countries in the evolution of independent lineages. The potential for novel influenza virus lineages to evolve within Africa warrants intensified influenza surveillance in Africa and other understudied areas.

Keywords. human influenza A virus; pandemic; phylogenetic analysis; Africa.

Despite strong seasonal bottlenecks, influenza A viruses (IAVs) persist in humans through continual global migration, which repeatedly reseeds viral diversity on local scales [1, 2]. It has been proposed that Southeast Asian countries are the most important global sources of antigenically novel IAVs, supplying Europe, North America, Africa, Latin America, and Oceania with novel variants on a continual basis [3]. Subsequent studies added complexity to this model, including roles for North America and other regions in the genesis of novel diversity [4] that are more consistent with a shifting meta-population model [5]. However, the lack of viral sequence data from a number of global regions, including Latin America, South Asia, and Africa, remains a major barrier to understanding the complex global ecology and evolution of IAVs.

The emergence of novel pandemic A/H1N1 (pH1N1) viruses in early 2009 resulted in a global expansion in IAV sequencing. Currently, >8000 full-length sequences of the main antigenic protein, the hemagglutinin (HA), are available through the Global Initiative on Sharing All Influenza Data (GISAID; www.gisaid.org). Notably, influenza surveillance increased in a number of African countries during 2009–2013 in response to the pandemic and as part of a larger trend of increased recognition of the seasonal influenza virus burden in Africa (see reviews [6, 7]). To elucidate the evolution of influenza viruses in Africa, we conducted a large-scale phylogenetic analysis of global pH1N1 influenza virus diversity during 2009–2013, including 299 pH1N1 HA sequences collected in 18 African countries. Our analysis identified 2 well-supported clades of pH1N1 viruses that each persisted for >1.5 years in West Africa, highlighting the need to further understand the ecology and evolution of IAVs in this understudied and relatively geographically isolated region.

METHODS

Full-length HA sequences of 8712 pandemic A/H1N1 influenza viruses collected globally during 2009–2013 were downloaded from GISAID (www.gisaid.org). These data include 299 HA sequences from pandemic A/H1N1 influenza viruses collected in 18 African countries representing 4 African subregions: 3 nations in North Africa (Algeria, Morocco, and Egypt), 7 nations in West Africa (Burkina Faso, Cameroon, Côte d'Ivoire, Ghana, Niger, Nigeria, and Senegal), 7 nations in East Africa (Djibouti, Ethiopia, Kenya, Madagascar, Tanzania, Uganda, and Zambia), and 1 Southern African nation (South Africa) (Supplementary Table 1; GenBank accession numbers KJ026367–KJ026452). Because Cameroon shares a long border with Nigeria and no other Central African countries had viral sequence data available that met the criteria for this study, for simplicity Cameroon was considered to be within West Africa in this analysis.

Received 30 October 2013; accepted 15 January 2014; electronically published 19 January 2014.

Presented in part: Options for the Control of Influenza VIII conference, September 5, 2013, Cape Town, South Africa.

Correspondence: Martha I. Nelson, PhD, Division of International Epidemiology and Population Studies, Fogarty International Center, National Institutes of Health, 16 Center Dr, Room 202, Bethesda, MD 20892 (nelsonma@mail.nih.gov).

The Journal of Infectious Diseases 2014;210:121–5

DOI: 10.1093/infdis/jiu047

The entire data set of 8712 pH1N1 HA (H1) sequences was aligned using MUSCLE alignment software (version 3.3.81) [8], with manual correction. A maximum likelihood (ML) tree of these data was inferred by using RAxML software (version 7.2.6) [9] and a general time-reversible (GTR) model of nucleotide substitution with a gamma-distributed (Γ) rate variation among sites (data available on request). Given the extremely large size of the data set, 100 bootstrap replicates were generated using RAxML software, with the same GTR + Γ model of nucleotide substitution, to assess the robustness of individual nodes. The resulting phylogeny was visualized and midpoint rooted using FigTree software (version 1.4.0; available at: http://tree.bio.ed.ac.uk/software/figtree/), and amino acid substitutions were identified using a customized R script. Similarly, an ML phylogeny also was inferred for 7644 full-length (1407 nt) pH1N1 NA (N1) sequences that were obtained and analyzed using identical methods (Supplementary Figure 1). For the purposes of visualization (Figure 1), an ML tree also was inferred from a subset of the HA sequence data (1643 HA sequences) that included all 299 African viruses, 9 viruses identified within West African clades I and II (described below), and 300 HA sequences randomly sampled from each year during 2009–2013 (all sequences from 2013 were used because <300 sequences were available). The full tree used for Figure 1, including tip labels, is available in Supplementary Figure 2.

RESULTS

The majority of pH1N1 influenza viruses collected in Africa during 2009–2013 (256 of 299; 86%) are phylogenetically interspersed with viruses collected from other continents, as expected given the ability of influenza viruses to rapidly disseminate between continents. However, 43 African pH1N1 viruses (14%) are positioned within 2 phylogenetically distinct clades that comprised viruses collected predominantly in West Africa (labeled as West African clades I and II in Figure 1). Both West African clades I and II are defined by high bootstrap values (≥80%) and separated by long branch lengths on the HA tree, indicating that the lineages have evolved independently. A monophyletic cluster of 21 Kenyan viruses also was identified that contains viruses collected from March 2010 (A/Nairobi/20/2010) through November 2011 (A/Kenya/190/2011), indicative of 20 months of persistence (Figure 1). However, this cluster was not supported by high bootstrap values and therefore was not considered in our analysis. Both West African clades I and II are relatively small—15 and 37 isolates, respectively. However, the small size of these clades is more likely to be explained by the sparse sampling of influenza viruses in West Africa than by their low prevalence (Table 1).

West African clade I comprises14 isolates collected from 3 West African countries (Côte d'Ivoire, Ghana, and Senegal)

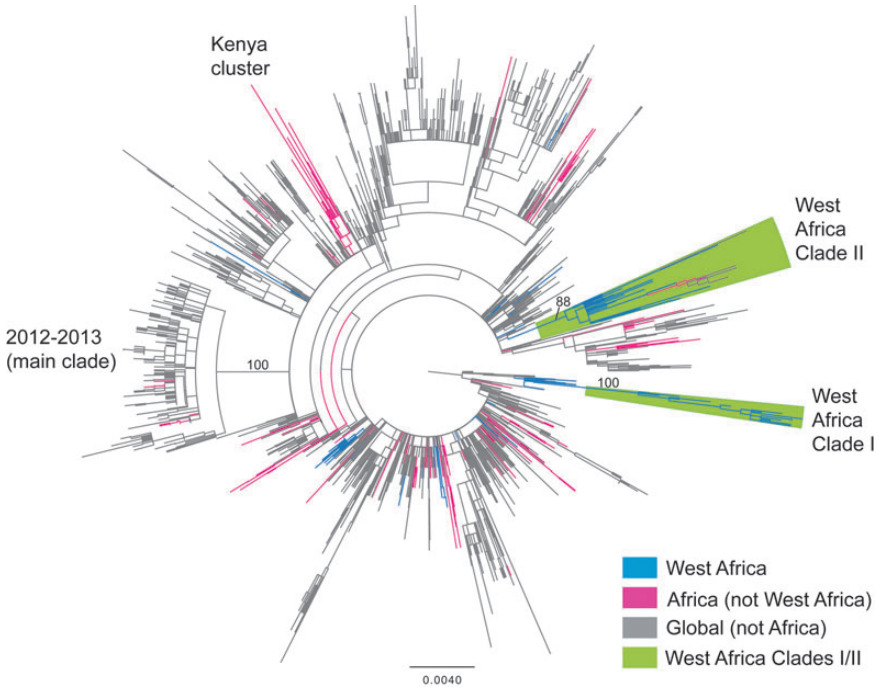

Figure 1. Phylogenetic relationships of H1 sequences from 1643 pandemic A/H1N1 influenza viruses collected globally during 2009–2013, estimated using a maximum likelihood method. West African clades I and II, the main global clade of 2012–2013 viruses, and the cluster of Kenyan viruses are labeled. Bootstrap values >70% are provided for West African clades I and II and the main 2012–2013 clade. The full phylogeny, including tip labels, is provided in Supplementary Figure 2.

Table 1. pH1N1 Influenza Viruses From Africa Used in This Study (n = 299)

Country	Viruses, No.			Viruses in Clade I or II, %
	Clade I	Clade II	Total[a]	
North Africa				
Algeria	0	0	7	0[b]
Egypt	0	0	21	0
Morocco	0	0	19	0
West Africa				
Burkina Faso	0	3	3	100[b]
Cameroon	0	13	20	65
Côte d'Ivoire	3	3	10	60
Ghana	2	3	55	9[b]
Niger	0	0	4	0[b]
Nigeria	0	4	10	40
Senegal	9	0	13	69
East Africa				
Djibouti	0	0	2	0[b]
Ethiopia	0	3	14	21
Kenya	0	0	66	0
Madagascar	0	0	14	0
Tanzania	0	0	22	0
Uganda	0	0	8	0[b]
Zambia	0	0	1	0[b]
Southern Africa				
South Africa	0	0	10	0
Total	**14**	**29**	**299**	**14**

[a] Total across the entire phylogenetic tree (ie, all viruses from a country).

[b] Percentages for countries with <10 influenza virus hemagglutinin sequences should be interpreted with caution.

and a singleton isolate from France (A/Lyon/CHU/43.28/2010). Bootstrap support for West African clade I is very high (100%), indicating that clade I circulated among West African countries for ≥1.8 years: from at least March–April 2011 in Senegal (A/Dakar/09/2011, A/Dakar/11/2011, A/Dakar/14/2011) to January 2013 in Côte d'Ivoire (A/Cote d'Ivoire/1575/2013). Two amino acid substitutions in the H1 sequences were observed among all 15 viruses in clade I: S145T and R276K (Supplementary Figure 3). Otherwise, the S145T and R276K substitutions were observed globally at extremely low levels, at <1% among the other 8790 pandemic H1 sequences. All 14 West African clade I isolates also clustered together on the NA global phylogeny, and 11 of 14 isolates in this cluster also were highly bootstrap supported (85%) (Supplementary Figure 1).

West African clade II is larger (37 isolates) and more diverse than clade I, containing 26 isolates from 5 West African countries (Burkina Faso, Cameroon, Côte d'Ivoire, Ghana, and Nigeria), 3 from Ethiopia in East Africa, and 8 from non-African countries in Europe (France, Italy, Norway, and Sweden) and the United States (Minnesota). Within West Africa, clade II has persisted for ≥1.9 years, from at least November 2010 in Nigeria and Cameroon (A/Nigeria/3310/2010, A/Cameroon/CAM13876HIN/2010, and A/Cameroon/10v-4655/2010) to October 2012 in Nigeria (A/Nigeria/7781/2012) (Table 1). In contrast, the East African isolates were collected during 2012, and all non-African isolates were collected from November 2012 to January 2013. Critically, 10 of the 11 non–West African isolates (except for a single outlier from France: A/Paris/1878/2012) cluster together within a highly supported subclade (bootstrap, 100%), consistent with 1 migration event of this lineage out of West Africa. Although the majority of West African isolates from clade II also cluster together on the NA phylogeny (Supplementary Figure 1), the non–West African isolates from Ethiopia, Europe, and the United States are positioned together in a different part of the NA tree, indicative of reassortment. Four amino acid substitutions in the H1 sequences were observed among 37/37 clade II viruses: A15T, N490D, T491K, and V537A (Supplementary Figure 3). The A15T, N490D, and T491K substitutions were observed globally at extremely low levels: <1% among the other 8790 pandemic H1 sequences.

DISCUSSION

A central question in influenza virus epidemiology and evolution is whether viral lineages can persist at low levels of circulation on local and regional scales, or whether new viruses must be reseeded continually from a globally sustained gene pool. In the past 7 years, large-scale phylogenetic analyses of global influenza virus sequence data have provided important insights into viral migration and persistence [1–5]. The great intensification of IAV sequencing that occurred during the 2009 H1N1 pandemic provided the best opportunity to identify viral persistence on local scales, including low-frequency variants. Heightened influenza surveillance during New York State's first pandemic wave revealed that A/H3N2 seasonal influenza viruses persisted into the early summer, although this persistent lineage did not give rise to the state's subsequent fall epidemic virus [10]. In the United Kingdom, Baillie et al [11] detected 2 United Kingdom–specific pH1N1 lineages that persisted between the first and second pandemic waves of 2009, although the close timing of the 2 waves meant that the lineages persisted for <6 months in the United Kingdom.

Our analysis of viral evolution in West Africa revealed the sustained persistence of 2 pH1N1 clades over a nearly 2-year period, although increased sampling is required to confirm that isolates from other localities are not interspersed within these clades. The intensity of global sampling of pandemic A/H1N1 influenza viruses during this time (>8000 full-length HA sequences were included in this analysis) reduces the likelihood that the persistence observed here is an artifact of unsampled diversity. In fact, the relative sparseness of sampling in West Africa compared to other regions (115 of 8712 global sequences,

1.3%) means that our data are strongly biased *against* the detection of clades comprised solely of West African viruses. That air traffic volume within West Africa is lower than in other global regions further supports the plausibility of influenza virus lineage persistence for several years without widespread dissemination to other continents. The delayed appearance of the pandemic H1N1 virus in West Africa 6 months after the emergence of the virus in North America in March 2009 further supports the possibility that influenza virus evolution in West Africa is less strongly linked to other continents owing to reduced international air traffic [12].

The lack of synchrony among the variable seasonal patterns of geographically linked African countries may also facilitate the persistence of influenza viruses within West Africa. In contrast to temperate areas where influenza virus epidemics are strongly synchronized during winter and undergo strong bottlenecks during summertime, the virus may persist through continual migration among seasonally variable and asynchronous West African localities, similar to what is thought to occur within the Southeast Asian network [3]. In Senegal, which has been conducting influenza surveillance for more than a decade, peaks in influenza activity coincide with the rainy season during July through September [13]. Nigeria, Côte d'Ivoire, and Cameroon exhibit more variable patterns of influenza virus seasonality, although longer time series of data are required [14, 15]. The phylogenies indicate that both West African clades, clades I and II, evolved within multiple West African countries, with no evidence of long-term persistence within a single country, which is consistent with a role for viral migration between West African countries in the persistence of the virus. High levels of bootstrap support for a subclade within clade II that contains isolates from Ethiopia, Europe, and the United States indicates that clade II viruses may have disseminated from West Africa to East Africa and from East Africa to Europe and the United States around 2012. Additionally, the detection of a singleton virus from France within both clade I (A/Lyon/CHU/43.28/2010) and clade II (A/Paris/1878/2012) could relate to the frequency of air traffic between francophone West Africa and France. The lack of detection of either West African clades I or II in North or Southern Africa may also reflect low human mobility between these regions and West Africa. Further understanding of influenza virus migration within Africa and between Africa and other continents will require additional sequence data.

Although the intensity of influenza surveillance in Africa still lags behind that of other continents, these findings suggest that substantial viral diversity circulates within Africa, including viral lineages that are unique to the region but capable of disseminating to other continents. Small sample sizes at country levels necessitate cautious inference of clade prevalence, and it remains unknown whether clades I and II comprised substantial proportions of the H1N1 diversity in West African countries. The possibility of minor variants evolving locally within West Africa undermines the assumption that a vaccine matched to globally dominant lineages will necessarily protect against these local lineages, although more data is clearly required. Further knowledge of the viral lineages that circulate within Africa, including antigenic characterization, is required to understand the full diversity and global ecology of influenza viruses in humans and to inform vaccination strategies within Africa.

Supplementary Data

Supplementary materials are available at *The Journal of Infectious Diseases* online (http://jid.oxfordjournals.org/). Supplementary materials consist of data provided by the author that are published to benefit the reader. The posted materials are not copyedited. The contents of all supplementary data are the sole responsibility of the authors. Questions or messages regarding errors should be addressed to the author.

Notes

Acknowledgments. We thank Dr Workenesh Ayele, PhD at Ethiopian Health and Nutrition Research Institute for providing A/H1N1 HA sequences from Ethiopia, East Africa, for comparison with sequences from West Africa.

Financial support. This work was supported by the Multinational Influenza Seasonal Mortality Study, an ongoing international collaborative effort to understand influenza epidemiology and evolution, led by the Fogarty International Center, National Institutes of Health, with funding from the Office of Global Affairs at the Department of Health and Human Services (C. V. and M. I. N.). E. C. H. is supported by an National Health and Medical Research Council (NHMRC) Australia fellowship.

Potential conflicts of interest. All authors: No reported conflicts.

All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

References

1. Nelson MI, Simonsen L, Viboud C, Miller MA, Holmes EC. Phylogenetic analysis reveals the global migration of seasonal influenza A viruses. PLoS Pathog **2007**; 3:1220–8.
2. Rambaut A, Pybus OG, Nelson MI, Viboud C, Taubenberger JK, Holmes EC. The genomic and epidemiological dynamics of human influenza A virus. Nature **2008**; 453:615–9.
3. Russell CA, Jones TC, Barr IG, et al. The global circulation of seasonal influenza A (H3N2) viruses. Science **2008**; 320:340–6.
4. Bedford T, Cobey S, Beerli P, Pascual M. Global migration dynamics underlie evolution and persistence of human influenza A (H3N2). PLoS Pathog **2010**; 6:e1000918.
5. Bahl J, Nelson MI, Chan KH, et al. Temporally structured metapopulation dynamics and persistence of influenza A H3N2 virus in humans. Proc Natl Acad Sci U S A **2011**; 108:19359–64.
6. Katz MA, Schoub BD, Heraud JM, et al. Influenza in Africa: Uncovering the Epidemiology of a Long-Overlooked Disease. J Infect Dis **2012**; 206 (Suppl 1):S1–S4.
7. Yazdanbakhsh M, Kremsner PG. Influenza in Africa. PLoS Med **2009**; 6:e1000182.
8. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res **2004**; 32:1792–7.
9. Stamatakis A. RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics **2006**; 22:2688–90.

10. Ghedin E, Wentworth DE, Halpin RA, et al. Unseasonal transmission of H3N2 influenza A virus during the swine-origin H1N1 pandemic. J Virol **2010**; 84:5715–8.
11. Baillie GJ, Galiano M, Agapow PM, et al. Evolutionary dynamics of local pandemic H1N1/2009 influenza virus lineages revealed by whole-genome analysis. J Virol **2012**; 86:11–8.
12. Nzussouo NT, Michalove J, Diop OM, et al. Delayed 2009 pandemic influenza A virus subtype H1N1 circulation in West Africa, May 2009-April 2010. J Infect Dis **2012**; 206(Suppl 1):S101–7.
13. Niang MN, Dosseh A, Ndiaye K, et al. Sentinel surveillance for influenza in Senegal, 1996–2009. J Infect Dis **2012**; 206(Suppl 1):S129–35.
14. Dalhatu IT, Medina-Marino A, Olsen SJ, et al. Influenza viruses in Nigeria, 2009–2010: results from the first 17 months of a national influenza sentinel surveillance system. J Infect Dis **2012**; 206(Suppl 1):S121–8.
15. Heraud JM, Njouom R, Rousset D, et al. Spatiotemporal circulation of influenza viruses in 5 African countries during 2008–2009: a collaborative study of the Institut Pasteur International Network. J Infect Dis **2012**; 206(Suppl 1):S5–13.