ECOLOGY AND GEOGRAPHY OF AVIAN VIRUSES USING NICHE MODELS AND WILD BIRD SURVEILLANCE

BY

Richard A. J. Williams

Submitted to the graduate degree program in Ecology and Evolutionary Biology and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

_____
Chairperson A. Townsend Peterson

_____
Steve Benedict

_____
Rafe Brown

_____
Nick Komar

_____
Stuart MacDonald

_____
Marilyn Smith

_____
Val Smith

Date Defended: 7 December, 2010

The Dissertation Committee for Richard A. J. Williams
certifies that this is the approved version of the following dissertation:

ECOLOGY AND GEOGRAPHY OF AVIAN VIRUSES USING NICHE MODELS

AND WILD BIRD SURVEILLANCE

_____

Chairperson A. Townsend Peterson

Date approved: 7 December, 2010

ABSTRACT

The emergence of highly pathogenic avian influenza strain H5N1 (hereafter "H5N1"), and other bird-associated viruses, have raised serious concerns about impacts on human, livestock, and wildlife populations. Ecological niche modelling (ENM) techniques were used to test the hypothesis that spatial distributions of H5N1 cases are related to coarse-scale environmental features in West Africa and in the Middle East and north-eastern Africa. Areas of drought-sensitive vegetation phenology were identified as key to H5N1 transmission, notwithstanding a small minority of models which indicated more variable transmission environments. ENMs were further used to estimate the environmental distribution of H5N1 relative to host group (poultry, wild birds, etc) in Europe. Results revealed no distinct ecological niche requirements among H5N1 host groups, suggesting that transmission cycles are broadly interwoven. Finally, avian virus surveillance was carried out in Ghana and Peru to assess patterns of host association and to test the assumption that avian influenza (AI) prevalence is low or nil in land birds. 600 Peruvian land birds of 177 species were tested for AI using rRT-PCR, revealing an infection prevalence of 1.3%. 564 Ghanaian land birds of 146 species were tested for AI (and Alphaviruses, and Flaviviruses) using PCR techniques. Samples were negative for Alphaviruses and AI, but amplified one sequence of a Yaoundé-like Flavivirus. Results of AI surveillance highlight the spatial variation of AI prevalences. Nonetheless, the prevalence in Peru demonstrates that surveillance programs for monitoring spread and identification of AI viruses should not focus solely on water birds.

To Mary Williams, who died before I completed my dissertation. You went to great pains to encourage my inquisitive mind and to feed my passion for natural history. You are the rock on which this work is built.

ACKNOWLEDGEMENTS

First I thank my doctoral supervisor, Town Peterson, for accepting me in his lab, providing direction, ideas, funding, not to mention the freedom to carry out most of my work with my family in Spain. Town's broad-ranging research interests, extensive collaborations, and productivity are inspiring and have been instrumental in my development as a scientist, and indeed, the creation of this dissertation. My heartfelt thanks extend to other committee members, Steve Benedict, Rafe Brown, Nick Komar, Stuart MacDonald, Rob Moyle, Marilyn Smith, and Val Smith, particularly to Marilyn, Nick, Stuart, and Val for extending lab facilities and resources to me, as well as spending their time teaching me basic techniques and sharing their ideas with me.

This work was highly collaborative. My sincere thanks to all co-authors of the publications and manuscripts that form my dissertation, for helping me to develop my intellectual outlook, for financial and practical assistance, and for comments on manuscripts, particularly to Town, Joel Montgomery, Ana Vázquez, Marie-Paz Sánchez-Seco, Antonio Tenorio, and William Ampofo. My thanks to Xiangming Xiao and Folorunso Fasina, respectively, for remote sensing and Nigerian H5N1 locality data; to Ana Vázquez, Ivy Asante, Kofi Bonney, Bruno Ghersi, Victor Gonzaga, Shirley Odoom, Naiki Puplampu and Karen Segovia Hinostroza for lab assistance. Ana and Kofi taught me a great deal in the lab, and it was always a pleasure to work with them. Joel and William offered me warm hospitality during my visits to Peru and Ghana respectively, many thanks.

No chapter of my dissertation would have been possible without advice and support from a number of friends and colleagues. My thanks to Monica Papeş, Ryan Lash, Sean Maher, and Yoshinori Nakazawa for taking me step by step through the challenges presented by ecological niche modelling, GIS and use of satellite imagery, and Chandrashekhar Biradar for help with remotely sensed data; to Narayani Barve for developing her partial AUC program and sharing it with me; to Matt Davis for advice and practical assistance with phylogenetic analyses, software solutions and DVDs; to expert ornithologists Mike Andersen, Elisa Bonaccorso, Mark Robbins, Max Thompson, Roger Boyd, Adolfo Navarro, Luis Sanchez-Gonzalez for field sample collection, without whose efforts two chapters could not have been written. I especially thank Adolfo, for giving me a virus, Mike and Mark for answering a plethora of specimen-related questions, and Elisa, Mark and Max for being excellent field companions in Ghana. Thanks too, to Brett Benz, Árpi Nyári and Andres Lira for helping in countless ways. Laura Benitez was instrumental in introducing me to collaborators at the Instituto de Salud Carlos III (ISCII), and the wonderful world of warts.

My doctoral research was funded by diverse institutions: the Centers for Disease Control and Prevention (CDC)/National Institutes of Health; CDC grant FIS PI07/1308 of the Red de Investigacion de Centros de Enfermedaes Tropicales RD06/0021; ISCIII grant GR09/0040(MPY-1440/09); the National Biological Information Infrastructure of the U.S. Geological Survey; an agreement signed between ISCIII and the Spanish Ministry of Health and Social Policy for the surveillance of imported viral hemorrhagic fevers; the National Institutes of Health - Fogarty International Center; the University of Kansas (KU) Center for Latin American Studies Tinker Summer Field Research Grant;

KU Natural History Museum (NHM) Panorama Grant; the US DoD AFHSC-GEIS grant number GF0121_07_LI; Wildlife Conservation Society's Global Avian Influenza Network for Surveillance. I received additional funding from KU, from the Department of Ecology and Evolutionary Biology (EEB), the Kansas Field Station Small Grants Program, and the NHM – though this funding was not specifically used to obtain results for my dissertation, it provided me with training which was vital to this project.

I owe a great deal to a number of institutions without which I would not have been able to complete my doctoral research. Foremost, to KU, and especially EEB, and the NHM, for providing an outstanding academic platform – an excellent library, employment, classes, training, the oxygen of a thriving intellectual environment. Lab testing was carried out at four institutions. I gratefully acknowledge: the Noguchi Memorial Institute for Medical Research, University of Ghana; the Naval Medical Research Center Detachment (NMRCD), Peru; the School of Veterinary Medicine, Universidad Nacional Mayor de San Marcos, Peru; the National Center of Microbiology, Instituto de Salud Carlos III (ISCIII), Spain, in both the Laboratorio de Gripe y Virus Respiratorios and Arbovirus e Infecciones Víricas Importadas. I am grateful for assistance to a multitude of people in these labs in addition to those already named as collaborators. I would like to give my particular thanks to Cecilia Gonzalez of NMRCD, to Armando Gonzalez for extending facilities and hospitality at San Marcos, and to many workers at ISCIII – Noelia Reyes for technical assistance with AI testing, to Laura Herrero, and especially Ana Negredo for holding my hand through viral cloning, to Lourdes Hernández, Dolores de la Cruz and Francisca Molero for their advice, assistance and friendliness, and to Dr Francisco Pozo and Dr Inmaculada Casas for their collaboration with influenza testing. Many thanks to the Government of Ghana Wildlife Division,

John Mason, Alex and others of the Nature Conservation Research Centre, Ghana, and the staff of the Centro de Ornitologia y Biodiversidad (CORBIDI), Lima, Peru, for extensive assistance with logistics associated with field sampling of birds. Ghanaian field assistants provided many laughs, eye-opening experiences, and colourful memories, in addition to helping us get the work done – thank you.

I believe I have thanked almost everybody who had a direct impact on this work, and I apologize if I have omitted anybody who should have been mentioned. A number of people, not previously mentioned, have helped this work indirectly – through their friendship, discussions, as class mates, mentors and in many other ways: in no particular order, Juan Manuel Guayasamin, Bob Timm, Francine Abe, Kathryn Mickle, Shannon Devaney, Jeff Cole, Ed Wiley, Jorge Soberón, Ryan O'Leary, Mark Mort, Lori Schlenker, Andy Bentley, Dick and Kathy Robins, Barb Clauson, Victor Gonzalez, Ron Heinrich, Chris Hauffler, Cal Cink, Matthew Buechner, Ginger Young. Thanks too to Craig Martin for calming words. I am indebted to pre-KU mentors for helping to fledge me – Fred Naggs, Roger Bamber and other friends at the Natural History Museum, London, and to Diana Lieberman.

My parents, Mary and John Williams, provided my fundamental education, and so much more. It was probably tough for you sometimes. Thanks for sticking with it, and I hope I manage to do half as good a job with my own little bundles of joy. Juliet, you probably deserve a sliver of credit/blame as well for those formative years.

Finally, and most importantly, my thanks to Merche. You let me go on this crazy American trip, supported me throughout, despite reservations, and gave me invaluable

advice on numerous problems. You are always there for me, and I will always be for you.

TABLE OF CONTENTS

INTRODUCTION

My dissertation is divided into five chapters exploring the spatial and taxonomic distribution of pathogens in avian hosts. One important focus is the emerging zoonosis, highly pathogenic avian influenza strain H5N1 (hereafter "H5N1"). Additionally I carried out wild bird surveillance, testing more generally for avian influenza (AI) strains, as well as for the arbovirus genera, *Alphavirus* and *Flavivirus*.

There is a long history of human infection with influenza viruses. Influenza is a major cause of human morbidity and mortality, typically between 250,000 and 500,000 deaths, and between 3 and 5 million serious cases per year [(1)]. This is punctuated by influenza pandemics, such as the "Spanish 'flu" of 1918, with an estimated death toll of 20-100 million [(2-3)]. H5N1 is an emerging strain of influenza that is notable for its pathogenicity to man [(4)], decimation of poultry flocks [(5)] and, to a lesser extent, wild bird populations [(6)], and for its rapid hemispheric spread [(7)]. Human H5N1 infections can be fatal, indeed, the current mortality rate is 59% of confirmed cases [(4)], though to date, fears that the H5N1 emergence will cause a human influenza pandemic have proved unfounded. Nevertheless, focus on H5N1 is essential both to improve understanding of this notable pathogen, and as a model of dynamics in this virus group.

Genus *Flavivirus* contains a number of well-known human pathogens, such as Dengue, Japanese Encephalitis, Tick-borne Encephalitis, West Nile, and Yellow Fever viruses. Several flaviviruses, particularly those of the Japanese Encephalitis serogroup, have been particularly associated with avian hosts [(8)]. West Nile virus (WN) became notorious, following emergence in North America, as causal agent of significant human and avian mortality. Notably, WN rapidly invaded the Western Hemisphere in less than

a decade [9]. Genus *Alphavirus*, though less well-known than influenza and Flavivirus also contains notable human pathogens, such as Chikungunya, Eastern Equine Encephalitis and Sindbis viruses. A number of Alphaviruses are primarily associated with avian hosts [10].

My first chapter [11] tests the hypothesis that spatial distributions of veterinary and human H5N1 cases are related to coarse-scale environmental features in West Africa. Ecological niche models (ENMs) were used to predict the potential distribution of H5N1 in West Africa. Models were subjected to predictive challenges of the ability of the model to anticipate the spatial distribution of known cases. The second chapter [12] further explored the hypothesis that spatial distributions of H5N1 cases are predictably associated to coarse-scale environmental features, now in the Middle East and north-eastern Africa. Once again, ENMs were used to relate case occurrences to environmental parameters. Models were challenged by a variety of spatial stratification schemes testing the ability of models to predict case distributions in broadly unsampled areas.

Since the relative roles of different avian host groups to H5N1 transmission remain contentious [13-14], chapter three compares the ecological niche requirements of paired host groups (e.g. wild birds and poultry), using tools recently developed to analyse ecological niches and geographic distributions of species. If environmental signals of different host groups are significantly different, the groups are likely involved in distinct transmission cycles; in contrast, models for which similarity cannot be rejected imply no unique ecological niches, and potential linkage of transmission cycles.

Chapters four and five present results from field survey in Ghana and Peru for avian influenza (AI) in non-symptomatic wild birds, testing the assumption that AI prevalence is low or nil in land birds. Survey work in Ghana also tested for the presence of Alphaviruses and Flaviviruses with a view to characterize the diversity and host ecology of these viruses in a previously un-sampled region. Behavioural and ecological characteristics of results from the Peru survey were analysed to measure associations between numbers of positive versus negative influenza cases in individuals and species displaying particular characteristics (e.g. forest birds versus open country birds). In addition I explored the phylogenetic relationships between the Peru matrix sequence and 123 influenza matrix gene sequences obtained from GenBank.

INTRODUCTION REFERENCES

1. WHO. Influenza (Seasonal). 2009 [cited 2010 11th November]; Available from: http://www.who.int/mediacentre/factsheets/fs211/en/

2. Beveridge WIB. The chronicle of influenza epidemics. Hist Philos Life Sci. 1991;13:223-35.

3. Taubenberger JK. The origin and virulence of the 1918 "Spanish" influenza virus. Proc Am Philos Soc. 2006;150:86-112.

4. WHO. *Cumulative number of confirmed human cases of avian influenza A(H5N1) reported to WHO*. 2010 [cited 2010 24 November]; Available from: http://www.who.int/csr/disease/avian_influenza/country/cases_table_2010_11_19/en/index.html

5. Whitworth D, Newman SH, Mundkur T, Harris P, editors. Wild birds and avian influenza: an introduction to applied field research and disease sampling techniques. FAO Animal Production and Health Manual, No. 5 ed. Rome; 2007.

6. Chen H, Smith GJD, Zhang SY, Qin K, Wang J, Li KS, et al. H5N1 virus outbreak in migratory waterfowl. Nature. 2005;436:191-2.

7. Gauthier-Clerc M, Lebarbenchon C, Thomas F. Recent expansion of highly pathogenic avian influenza H5N1: a critical review. Ibis. 2007;149:202-14.

8. Gould EA, de Lamballerie X, Zanotto PMD, Holmes EC. Origins, evolution, and vector/host coadaptations within the genus flavivirus. Advances in Viral Research. 2003;59:277-314.

9. Komar N, Clark GG. West Nile virus activity in Latin America and the Caribbean. Pan American Journal of Public Health. 2006;19:112-7.

10. Powers AM, Brault AC, Shirako Y, Strauss EG, Kang WL, Strauss JH, et al. Evolutionary relationships and systematics of the alphaviruses. J Virol. 2001;75:10118-31.

11. Williams RAJ, Fasina FO, Peterson AT. Predictable ecology and geography of avian influenza (H5N1) transmission in Nigeria and West Africa. Trans R Soc Trop Med Hyg. 2008;102:471-9.

12. Williams RAJ, Peterson AT. Ecology and geography of avian influenza (HPAI H5N1) transmission in the Middle East and northeastern Africa. Int J Health Geogr. 2009;8:e47.

13. Weber TP, Stilianakis NI. Ecologic immunology of avian influenza (H5N1) in migratory birds. Emerg Infect Dis. 2007;13:1139-43.

14. Keawcharoen J, van Riel D, van Amerongen G, Bestebroer T, Beyer WE, van Lavieren R, et al. Wild ducks as long-distance vectors of highly pathogenic avian influenza virus (H5NI). Emerg Infect Dis. 2008;14:600-7.

CHAPTER 1: PREDICTABLE ECOLOGY AND GEOGRAPHY OF AVIAN INFLUENZA (H5N1) TRANSMISSION IN NIGERIA AND WEST AFRICA

Published as Williams, R.A.J., Folorunso O. Fasina, and A. Townsend Peterson, Predictable ecology and geography of avian influenza (H5N1) transmission in Nigeria and West Africa, Trans R Soc Trop Med Hyg (2008), doi:10.1016/j.trstmh.2008.01.016
Received 24 July 2007; received in revised form 24 January 2008; accepted 24 January 2008

ABSTRACT

The emerging virus strain termed highly pathogenic H5N1 avian influenza (hearafter "H5N1") has spread widely in the past decade and is now the focus of considerable concern in several sectors. We tested the hypothesis that spatial distributions of veterinary and human H5N1 cases are related to coarse-scale environmental features in West Africa. We used ecological niche models to associate Nigerian H5N1 occurrences with 1 km resolution digital data layers summarizing parameters of surface reflectance and landform. Predictive challenges included anticipating the spatial distribution of (i) random subsamples and (ii) spatially and temporally stratified subsamples of Nigerian occurrence data, and (iii) more limited occurrence data from across West Africa. In almost all tests, we found that H5N1 cases were occurring under predictable environmental conditions, suggesting that elements of the transmission cycle have some form of ecological determination, here measured as differences in land-surface reflectance and plant phenology through the year. Considerable additional work is needed to establish how these differences affect H5N1 transmission.

INTRODUCTION

Influenza A viruses are responsible for considerable human 3 morbidity and mortality, with annual estimates of up to 4 500 000 deaths globally [1]. Influenza mortality was even higher during the three major pandemics of the twentieth century: the Spanish influenza epidemic (1917—1919) killed an estimated 20 million to 100 million people [2-4]. Much speculation is focusing on a future influenza pandemic, and highly pathogenic H5N1 (H5N1) is receiving attention as the prime candidate [5-6]. At the time of writing, 337 human H5N1 cases had been documented from 59 countries [7], of which 207 (~60%) were fatal [8].

The main reservoirs of influenza A viruses (such as H5N1) are generally considered to be aquatic wild birds, principally Anseriformes (ducks, geese, swans) and Charadriiformes (gulls, terns, shorebirds) [9-11], although this assumption may not be universally applicable [12-14]. H5N1, however, is best known in domestic birds, particularly chickens and ducks, with numerous cases documented across Asia, Europe and Africa; 140 million domestic birds have been culled owing to fears of epidemic occurrence of H5N1 [15].

H5N1 was first isolated in 1996 from a farm goose in Guangdong Province, People's Republic of China [16]. By August, 2004, H5N1 had been detected in 10 East Asian countries, and by December 2005 had been detected widely in northern Asia, the Middle East and much of Europe. It was then confirmed in Africa (Nigeria), and by April 2006 had been detected in five African countries (Burkina Faso, Cameroon, Egypt, Nigeria, Niger); the list presently also includes Sudan, Ivory Coast, Ghana, Togo and Benin.

H5N1 was confirmed in Nigeria in February 2006, at a commercial poultry farm (chickens, geese, ostriches) in Kaduna Province — by the end of the initial outbreak 42 000 poultry cases had been reported [17]. Since that time, H5N1 has been detected in >150 instances (F.O.Fasina, unpublished data) in 20 Nigerian states (particularly in the northern and southwestern parts of the country), including one confirmed human death [18]. Three distinct H5N1 genotypes have been detected in the country [19], likely the result of introduction of two strains and one subsequent reassortment [20]. H5N1 was probably not present in Nigeria much before 2006, as serosurveys of samples collected in 1999—2004 were all negative [21].

Factors associated with risk of H5N1 transmission in local landscapes are poorly known — the only previous study [22] showed a clear association with domestic duck populations in rice-paddy agro-ecosystems, but such factors are unknown elsewhere in the distribution of the virus; indeed, the question of whether landscape scale risk factors exist must remain open to testing. Here, we use novel tools [ecological niche models, (ENMs)] to provide a landscape-scale perspective on the question of H5N1 risk assessment: we associate Nigerian H5N1 cases with annual variation in 'greenness' [Normalized Difference Vegetation Index (NDVI) values derived from imagery from the Advanced Very High Resolution Radiometer (AVHRR) satellite], and develop and test predictive spatial models of H5N1 occurrence. As such, this paper represents a first exploration of the ecological 'niche' of H5N1 cases in Nigeria and more broadly across West Africa. Although we are well aware of the unlikely nature of predictable environmental correlates, given the multiple, diverse factors associated with H5N1 transmission (migratory birds, poultry trade,

etc.), our results indicate that H5N1 occurs under consistent, predictable environmental circumstances in West Africa.

Our development of H5N1 ENMs is based on the concept of ecological niches defined as the set of conditions under which a species is able to maintain populations without immigration [23-24]. ENMs have seen considerable exploration recently as a means of estimating the dimensions of ecological niches of species based on incomplete sampling across distributions (see review in Peterson [25]). Widespread evolutionary conservatism in niche characteristics has been demonstrated, allowing accurate predictions of transmission of infectious diseases [26-27], invasive species' potential distributions [28] and projections of species' responses to climate change [29], etc. Overall, ENMs offer a powerful approach to understanding coarse-scale environmental correlates of presence and absence of species or biological phenomena across complex landscapes.

METHODS

Input data

The principal suite of occurrence information for this study was H5N1 case-occurrence data for January—April 2006 from the National Veterinary Research Institute, Nigeria, which consisted of 132 H5N1 detections (including isolations from two wild birds, the remainder from poultry; Figure 1.1); an additional 12 occurrences were drawn from the same source for November 2006—January 2007 for model testing. Textual descriptions of occurrence localities were converted to geographic coordinates accurate to the nearest 0.001° using the Alexandria Digital Library Gazetteer (http://middleware.alexandria.ucsb.edu/client/gaz/adl/index.jsp), GEOnet Names Server

(http://gnswww.nga.mil/geonames/GNS/index.jsp), and other sources [30]. In all, we were able to convert 72 of these locations into unique geographic coordinates; the attrition from 132 to 72 localities is largely because 38 duplicate occurrences from the same localities were discarded, and in 22 cases we were unable to confirm the coordinates of the locality. Finally, to characterize the broader distribution of H5N1 in West Africa, we searched the archives of the International Society for Infectious Disease (ProMED Avian Influenza archive) for West Africa (14 occurrences; Figure 1.1): Burkina Faso (four points), Ivory Coast (three), Ghana (two), Niger (two) and Cameroon (one), excluding two duplicated localities four localities (three from Niger, one from Ivory Coast) for which we were unable to locate coordinates of the reported site. Throughout, although the geographic coordinates assigned may not always fix the exposure point precisely, they represent a best guess as to its position, and are likely to represent the coarse-scale ecological conditions under which H5N1 transmission occurs. We believe that georeferencing imprecision is of a magnitude smaller than the resolution of our environmental grids, so the modelling process is not compromised.

Environmental data sets included twenty-four 1 km^2 resolution monthly composite remotely sensed data layers (April 1992—March 1993 and February 1995—January 1996), in each case presenting values of the NDVI (native spatial resolution 1 km). NDVI is derived from reflectance in the visible and near-infrared domains and as such is sensitive to photosynthetic activity and is closely correlated with photosynthetic mass [31] — the time series of NDVI values used here thus profile differences in land cover and plant phenology across landscapes. We also included four data sets summarizing aspects of topography — elevation, slope, aspect and compound topographic index (which summarizes tendency to

pool water) — from the U.S. Geological Survey's Hydro-1K data set (http://eros.usgs.gov/products/elevation/gtopo30/hydro/africa.html; native resolution 1 km). Climate data were not included in these analyses for lack of sufficiently high-resolution data sets across the region of interest. We purposefully included NDVI series from both an El Niño year (1992/1993) and a normal year (1995/1996) to take into account any effects that these global climate phenomena might have on West African landscapes.

The GARP algorithm

We used the Genetic Algorithm for Rule-set Prediction (GARP) [32] for ENM development. GARP uses an evolutionary-computing genetic algorithm to search for non-random associations between environmental variables and known occurrences of species, as contrasted 156 with environmental characteristics across the overall study area. GARP has been applied widely to questions of disease transmission [33-34], and its predictive ability has been tested under diverse circumstances [35-37]. Although GARP was highlighted for its relatively poor performance in recent comparative studies [36, 38], other recent studies have indicated considerably better performance [39-40] and some artifactual causation of the previous negative results [39, 41]. As such, we employed GARP in these analyses, although we highlight algorithm choice as an important issue remaining in risk analysis applications.

In general, we develop tests based on subsets of available occurrence information set aside prior to model development. Of data provided to GARP, the program divides occurrence data randomly into three subsets: training data (for rule development), intrinsic testing data (for evaluation of rules) and extrinsic testing data (for evaluation of model quality, see

below). Spatial predictions of presence versus absence can include two types of error: false negatives (areas of actual presence predicted absent) and false positives (areas of actual absence predicted present) [42] — rule performance in each of these dimensions is evaluated via the intrinsic testing data set. Change in predictive accuracy from one iteration to the next are used to evaluate whether particular rules should be incorporated into the model or not, and the algorithm runs either 1000 iterations or until convergence [32]. The final rule-set is then used to query the environmental data sets to identify areas fitting the rule set predictions to produce a hypothesis of the potential geographic distribution of the species [43].

Since GARP includes several random-walk elements, each replicate model developed produces distinct results, representing alternative solutions to the optimization challenge. Following best-practices approaches [35], we developed 100 replicates of each model. We filtered these replicates based on their error characteristics, retaining the 20 with lowest false negative rates (= percentage of testing occurrence points falling in areas not predicted to be suitable), and then retained the 10 (of the 20) closest to the median of proportional area predicted present, an index of false-positive error rates [35]. A consensus of these 'best subset' models was then developed by summing values for each pixel in the map to produce final predictions of potential distributions with 11 thresholds (integers from 0 to 10).

The customary approaches to spatial model validation (e.g. receiver operating characteristic, kappa statistics) are not applicable to situations in which presence-only data are the only information available [42, 44]. As such, we validated models using simple

calculations of binomial probabilities that coincidence of predictions and independent test data are no better than random, with the probability of *k* successes in *n* trials depending on *p*, the probability of success in any one trial — we estimated *p* as the proportion of the testing area predicted present, and *k* as the number of the *n* testing points that were successfully predicted [35]. Binomial probabilities were calculated for each of the 10 thresholds representing predictions of presence (1 = broad, 10 = narrow), in each case testing whether predictivity is better than that expected by chance. In one case, we explored the effects of spatial uncertainty regarding the localization of outbreak sites by calculating success in predicting areas of presence within 100 m of known occurrence sites, adjusting *p* appropriately to reflect the broader area of potential presence.

Modelling approach

This study focuses on the question of whether H5N1 occurrences in West Africa follow a consistent and predictable environmental regime. As such, we developed a series of tests of model predictivity, in each case with the models developed and the predictions tested being based on independent suites of occurrence data. Model tests were based on subsets of the 2006 Nigerian occurrence data described above, as well as on the 12 additional Nigerian occurrences from November 2006 to January 2007 (Figure 1.1); we also tested our Nigerian models with occurrence data from across West Africa. The basic design of testing was as follows:

Predictivity across training landscape

We divided the 72 Nigerian occurrences from 2006 into two equal groups at random, and used one group for model development and the other for model testing (hereafter 'RND'

tests). We also tested the ability of 2006-based ENMs to predict the spatial distribution of cases from November 2006 to January 2007 (hereafter 'YEAR' tests). This scheme assesses the ability of the modelling approach to anticipate the spatial distribution of H5N1 cases were sampling density to be increased, but across a region in which samples are already available.

Predictivity across space (medium scale)

We stratified the 72 Nigerian occurrences from 2006 spatially into quadrants above and below the median longitude and median latitude of the occurrence data. From this spatial stratification, we developed three pairs of quadrants: west versus east of the median longitude (hereafter 'EW' tests), north versus south of the median latitude (here after 'NS' tests), and on-diagonal (upper left-hand and lower right-hand quadrants) versus off-diagonal (lower left-hand and upper right-hand quadrants; hereafter 'DIAG' tests). In each case, we developed both reciprocal predictions, testing the ability of ENMs to anticipate the spatial distribution of HP-H5N1 cases in regions for which no sampling is available.

Predictivity across space (broader scale)

We projected ENMs trained using all 2006 Nigerian occurrence points onto the rest of West Africa, and tested their spatial predictions via their coincidence with the 14 cases for which geographic coordinates were available in other West African nations (hereafter 'WA' tests). These tests evaluated the ability of the ENMs to predict into even broader unsampled areas.

RESULTS

Almost all tests conducted in this study indicated that independent sets of test points coincided with ENM predictions significantly better than random expectations (Table 1.1). In other words, models based on known H5N1 occurrences can anticipate spatial distributions of independent samples of H5N1 based on their environmental attributes.

Predictivity across training landscape

The two reciprocal tests based on random subsamples of 2006 Nigerian occurrence data (RND tests) both indicated significant predictive power of the ENMs, with all thresholds showing predictivity better than random expectations (Table 1.1). For example, the broadest predictions (threshold >1 of 10 models predict present) predicted 65—72% of Nigeria as present, and predicted the spatial position of >86% of independent test occurrence points correctly ($P < 0.001$); the narrowest predictions (threshold of 10 of 10 models predict present) predicted 17% of Nigeria as present, and predicted the spatial position of 35—37.5% of independent test points correctly ($P < 0.001$).

Tests using 12 H5N1 occurrences subsequent to the main sample (YEAR tests) were somewhat successful. Regarding the raw model results, five of the 12 occurrences were predicted by all of the replicate models, and 10 of 12 occurrences were predicted by at least one of the replicate models; two of the 10 thresholds predicted spatial distributions better than randomly ($P < 0.05$; Table 1.1). However, we noted that several points lay close to predicted areas, so, in view of uncertainties of georeferencing, we explored the effects of introducing a degree of uncertainty into our tests: we traced a 100 m buffer around H5N1 occurrence sites, and noted a marked improvement in model performance. Here, six of 12

occurrences were predicted by all replicate models, and 11 of 12 by at least one model, and six of the 10 thresholds predicted H5N1 occurrences better than random expectations ($P < 0.05$).

Predictivity across space (medium scale)

We also assessed the ability of ENMs to anticipate spatial distributions of H5N1 occurrences in regions for which no input data were available — these three pairs of tests based on spatially stratified subsets of Nigerian H5N1 occurrence data also indicated, for the most part, significant predictive power of the ENM (Table 1.1, Figure 1.2). For both DIAG tests, all thresholds in both reciprocal tests were statistically significantly more predictive than random expectations ($P < 0.05$). For the NS tests, results were similar, except that two of 10 thresholds in the south predicts-north tests were not significant ($P > 0.05$). Finally, in the EW tests, nine of 10 thresholds were statistically significant in the west-predicts-east tests, but none was significant in the converse test — this ENM dramatically underpredicted H5N1 cases in the southern part of the country (Table 1.1).

More generally, H5N1 potential presence is predicted more broadly in northern Nigeria than in the southern part of the country; and some areas (e.g. along the south-eastern border) are predicted to be largely H5N1 free. The presence of the virus is predicted mostly in the savannah and woodland habitats of the north, whereas absence is predicted in montane areas, coastal mangroves, the freshwater swamps of the Niger Delta, and rainforest areas in the south. Areas of highest predicted H5N1 risk were not the 'greenest' areas, but rather were relatively dry and highly variable seasonally, as can be appreciated from NDVI profiles through the course of the year (Figure 1.3).

Predictivity across space (broader scale)

Given the initial signal of predictive power regarding H5N1 within Nigeria, we projected a Nigeria 2006-based ENM across West Africa to develop a broader-scale test of ENM predictive ability (Table 1.1, Figure 1.4). As within Nigeria, H5N1 potential for occurrence is predicted in the savannah and woodland belt and across the southern portion of the Sahel. The virus is not predicted to occur in the Sahara, montane zones, coastal mangroves or rainforest. In particular, we note that the Upper Guinean rainforest block (Guinea, Ivory Coast, Liberia, Sierra Leone, western Ghana) and the Lower Guinea rainforest block (Cameroon, southern Nigeria) are not predicted to be suitable for H5N1, but the savannah areas of the Dahomey Gap (Benin, eastern Ghana, Togo) are strongly predicted as suitable. The coincidence of the projection of the Nigerian ENM rule sets across West Africa coincided with the 14 independent test points better than random expectations at eight of 10 thresholds (all $P < 0.01$), except for the two most restrictive thresholds (both $P > 0.06$).

DISCUSSION

ENM applications to transmission are still preliminary, and are certainly relatively new to the field [25]. Given this novelty, two features of ENMs should be emphasized at this point to facilitate interpretation of model results. Firstly, ENMs are frequently coarse-resolution, distribution-wide views of biological phenomena that outline broad potential for disease occurrence; particular landscape features, management regimes (e.g. biosafety measures), and chance events (e.g. introduction of the pathogen) may prevent this potential from being realized, but the model results indeed indicate the coarser-scale potential for such occurrences. Secondly, given the potential nature of predictions of presence (compared to

ENM predictions of absence), false-positive error is much more serious in model evaluation than false-negative error. These features of ENM must be taken into account in any consideration of model predictions and their utility.

The ecological niche model predictions that we have developed for Nigeria and West Africa are exploratory, designed to test the basic hypothesis that environmental correlates exist. Although ENMs have been applied broadly to biodiversity questions [45-48], their application to disease systems remains preliminary [25]. Although several initial tests have been published [26, 28, 33, 49], the failings and biases of the technique in a disease transmission context are still being discovered and understood.

We recognize several limitations in our analyses. First, imprecision inherent in georeferencing infection sites sets a base level of error, and guarantees some predictive failures. Given that poultry is frequently traded and moved to markets, H5N1 infections may frequently appear at sites not coincident with transmission sites — a number of Nigerian HP-H5N1 cases were detected in poultry markets, to which infected birds were presumably transported over unknown distances from actual transmission sites. These factors — movement, transportation, trade and biosecurity measures on poultry farms — may impact the epidemiology of the disease, but we focus explicitly on ecological and geographical factors with the aim of developing a model of the ecological niche of the virus.

Another important challenge for these analyses is that of distinguishing true spatial and ecological biases in case distributions (i.e. the ecological niche!) from the spatial and

ecological biases in distributions of the major known H5N1 host in Nigeria (chickens). The total Nigerian chicken population is > 140 million birds, including 'backyard chickens', raised without biosecurity measures (~60), commercially farmed chickens under high biosecurity (~25%), and semi-commercial chickens, raised with some biosecurity measures (~15%)[50]. Most commercial birds (65%) are raised in the south-western part of the country, near Lagos [50].

Free mingling of backyard poultry and wild birds has been identified as a risk factor for H5N1 transmission [51-53]. In Nigeria, however, at least at the coarse scales examined herein, backyard chicken distributions and our reconstructed risk areas are virtually inverse: backyard chicken populations are highest and H5N1 predictions lowest in south-eastern Nigeria, and backyard chicken populations are lowest and H5N1 predictions highest in northern Nigeria (Figure 1.5). Moreover, H5N1 outbreak localities do not necessarily coincide with areas of high backyard chicken population — for example; the state with the highest backyard chicken populations (Imo, southwest Nigeria) has had no cases of H5N1, despite having roughly ten-fold higher density of backyard chickens as Plateau, the state with the highest number of H5N1 outbreak sites. Similarly, we observed little coincidence between H5N1 outbreaks and areas of high density of commercially farmed birds in the southeast [50]. This result coincides with experience in Thailand, where chickens are the most frequent victims of poultry H5N1 outbreaks, but outbreaks do not correspond to the distribution of backyard chickens [54].

Recent studies in Southeast Asia [22, 54] identified predictable foci of H5N1 activity based on free-range duck farming and rice-paddy cultivation. Although that association has clearly

and easily interpretable foundations, our results suggest that predictable ecology may be more pervasive in H5N1 geography than might have been expected. Several elements in the H5N1 transmission cycle could be responsible for this predictivity: ecological biases associated with initial arrival of virus propagules in a poultry population via migratory birds [55-56], transmission among Nigerian poultry flocks [54, 57], or even with transportation routes within Nigeria that might be responsible for communicating infections — most likely, the truth lies in a combination of such factors [58]. The precise basis for this predictivity has yet to be identified, but the existence of an environmental signal in H5N1 transmission may offer valuable clues as to its nature.

Interpretation of the nature of the environmental signal associated with high H5N1 transmission is complex. The NDVI data used in this study are correlated with photosynthetic mass [31], and our time series of NDVI images thus summarize patterns of vegetation phenology across landscapes. In the crudest sense, our NDVI profiles identify areas of drought-sensitive vegetation phenology as particularly key in HP-H5N1 transmission (Figure 1.3), but the details are still under study and exploration [34].

Perhaps most importantly, projecting the Nigerian ENMs across the entire region yielded a view of West African H5N1 distributions that was highly predictive of what independent test data could be assembled. Such validated model predictions offer the possibility of public health applications, providing information that may be used to prioritize surveillance and remediation activities. Similarly, such predictions may be helpful to policy makers planning expansions to and investment in the Nigerian chicken industry, particularly as regards biosecurity measures. The spatial limits of the predictivity we have documented

remain an open question — our initial demonstration of predictable H5N1 geography across West Africa awaits further testing and comparison with H5N1 occurrence information from other regions.

ACKNOWLEDGEMENTS

Thanks to Monica Papeş, Ryan Lash, and Yoshinori Nakazawa for their expert assistance with GIS and satellite imagery, and to the National Biological Information Infrastructure of the U.S. Geological Survey for financial support.

CHAPTER 1 REFERENCES

1. WHO. Influenza: Fact sheet 211. 2003 [cited 14th June, 2007]; Available from: http://www.who.int/mediacentre/factsheets/fs211/en/print.html

2. Beveridge WIB. The chronicle of influenza epidemics. Hist Philos Life Sci. 1991;13:223-35.

3. Johnson N, Mueller J. Updating the accounts: global mortality of the 1918-1920 "Spanish" influenza pandemic. Bull Hist Med. 2002;76:105-15.

4. Taubenberger JK. The origin and virulence of the 1918 "Spanish" influenza virus. Proc Am Philos Soc. 2006;150:86-112.

5. Guan Y, Poon LLM, Cheung CY, Ellis TM, Lim W, Lipatov AS, et al. H5N1 influenza: a protean pandemic threat. Proc Nat Acad Sci USA. 2004;101:8156-61.

6. Webby RJ, Webster RG. Are we ready for pandemic influenza? Science. 2003;302:1519-22.

7. WHO. H5N1 avian influenza: Timeline of major events. 2007 5 December 2007 [cited 7th May, 2007]; Available from: http://www.who.int/csr/disease/avian_influenza/Timeline_07_may_30.pdf

8. WHO. Cumulative number of confirmed human cases of avian influenza A/(H5N1) reported to WHO. 2007 [cited 2007 11 December]; Available from: http://www.who.int/csr/disease/avian influenza/Timeline 5 Dec 07.pdf

9. Suarez DL. Evolution of avian influenza viruses. Vet Microbiol. 2000;74:15-27.

10. Swayne DE, Suarez DL. Highly pathogenic avian influenza. Rev Sci Tech Off Int Epiz. 2000;19:463-82.

11. Webster RG, Bean WJ, Gorman OT, Chambers TM, Kawaoka Y. Evolution and ecology of influenza-A viruses. Microbiol Rev. 1992;56:152-79.

12. Tanimura N, Tsukamoto K, Okamatsu M, Mase M, Imada T, Nakamura K, et al. Pathology of fatal highly pathogenic H5N1 avian influenza virus infection in large-billed crows (*Corvus macrorhynchos*) during the 2004 outbreak in Japan. Vet Pathol. 2006;43:500-9.

13. Kou Z, Lei FM, Yu J, Fan ZJ, Yin ZH, Jia CX, et al. New genotype of avian influenza H5N1 viruses isolated from tree sparrows in China. J Virol. 2005;79:15460-6.

14. OIE. Rapport de mission: Mission to Russia to assess the avian influenza situation in wildlife and the national measures being taken to minimize the risk of international spread. 2005 [cited; Available from: http://www.oie.int/downld/Missions/2005/ReportRussia2005Final2.pdf

15. Webster RG, Peiris M, Chen HL, Guan Y. H5N1 outbreaks and enzootic influenza. Emerg Infect Dis. 2006;12:3-8.

16. Xu XY, Subbarao K, Cox NJ, Guo YJ. Genetic characterization of the pathogenic influenza A/Goose/Guangdong/1/96 (H5N1) virus: Similarity of its hemagglutinin gene to those of H5N1 viruses from the 1997 outbreaks in Hong Kong. Virology. 1999;261:15-9.

17. ProMED. Avian influenza - worldwide: Nigeria, OIE: ProMED archive number 20060208.0409, www.promedmail.org. ProMED-mail. 2006.

18. ProMED. Avian influenza, human (29): Nigeria, Confirmed: ProMED archive number 20070203.0439, www.promedmail.org. ProMED-mail. 2007.

19. Ducatez MF, Olinger CM, Owoade AA, De Landtsheer S, Ammerlaan W, Niesters HG, et al. Avian flu: multiple introductions of H5N1 in Nigeria. Nature. 2006;442:37.

20. Salzberg SL. Genome analysis linking recent European and African influenza (H5N1) viruses. Emerg Infect Dis. 2007;13:713-8.

21. Owoade AA, Ducatez MF, Muller CP. Seroprevalence of avian influenza virus, infectious bronchitis virus, reovirus, avian pneumovirus, infectious laryngotracheitis virus, and avian leukosis virus in Nigerian poultry. Avian Dis. 2006;50:222-7.

22. Gilbert M, Xiao XM, Chaitaweesub P, Kalpravidh W, Premashthira S, Boles S, et al. Avian influenza, domestic ducks and rice agriculture in Thailand. Agric Ecosyst Environ. 2007;119:409-15.

23. Grinnell J. The Niche Relationships of the Californian Thrasher. 1917;34:427-33.

24. Grinnell J. Geography and evolution. Ecology. 1924;5:225-9.

25. Peterson AT. Ecological niche modeling and spatial patterns of disease transmission. Emerg Infect Dis. 2007;12:1822-6.

26. Peterson AT, Sanchez-Cordero V, Ben Beard C, Ramsey JM. Ecologic niche modeling and potential reservoirs for Chagas disease, Mexico. Emerg Infect Dis. 2002;8:662-7.

27. Peterson AT, Bauer JT, Mills JN. Ecologic and geographic distribution of filovirus disease. Emerg Infect Dis. 2004;10:40-7.

28. Peterson AT, Shaw J. Lutzomyia vectors for cutaneous leishmaniasis in Southern Brazil: ecological niche models, predicted geographic distributions, and climate change effects. Int J Parasit. 2003;33:919-31.

29. Peterson AT, Ortega-Huerta MA, Bartley J, Sanchez-Cordero V, Soberon J, Buddemeier RH, et al. Future projections for Mexican faunas under global climate change scenarios. Nature. 2002;416:626—9.

30. Rand McNally. New Millenium World Atlas Deluxe. Skokie, Illinois, USA: Rand McNally & Co.; 1998.

31. Tucker CJ. Red and photographic infrared linear combinations for monitoring vegetation. Remote Sens Environ. 1979;8:127-50.

32. Stockwell DRB, Peters DP. The GARP modelling system: problems and solutions to automated spatial prediction. Int J Geogr Inf Sci. 1999;13:143-58.

33. Costa J, Peterson AT, Beard CB. Ecologic niche modeling and differentiation of populations of *Triatoma brasiliensis* Neiva, 1911, the most important Chagas' disease vector in northeastern Brazil (Hemiptera, Reduviidae, Triatominae). Am J Trop Med Hyg. 2002;67:516-20.

34. Peterson AT, Benz BW, Papeş M. Highly pathogenic H5N1 avian influenza: entry pathways into North America via bird migration. PLoS One. 2007;2:e261.

35. Anderson RP, Lew D, Peterson AT. Evaluating predictive models of species' distributions: criteria for selecting optimal models. Ecol Model. 2003;162:211-32.

36. Elith J, Graham CH, Anderson RP, Dudik M, Ferrier S, Guisan A, et al. Novel methods improve prediction of species' distributions from occurrence data. Ecography. 2006;29:129-51.

37. Peterson AT, Vieglais DA, Andreasen JK. Migratory birds modeled as critical transport agents for West Nile Virus in North America. Vector-Borne Zoonotic Dis. 2003;3:27-37.

38. Stockman AK, Beamer DA, Bond DA. An evaluation of a GARP model as an approach to predicting the spatial distribution of non-vagile invertebrate species. Divers Distrib. 2006;12:81—9.

39. McNyset K, Blackburn J. Does GARP really fail miserably? A response to Stockman et al. (2006). Divers Distrib. 2006;12:782-6.

40. Tsoar A, Allouche O, Steinitz O, Rotem D, Kadmon R. A comparative evaluation of presence-only methods for modelling species distribution. Divers Distrib. 2007;13:397-405.

41. Peterson AT, Papes M, Eaton M. Transferability and model evaluation in ecological niche modeling: a comparison of GARP and Maxent. Ecography. 2007;30:550-60.

42. Fielding AH, Bell JF. A review of methods for the assessment of prediction errors in conservation presence/absence models. Environ Conserv. 1997;24:38-49.

43. Soberón J, Peterson AT. Interpretation of models of fundamental ecological niches and species' distributional areas. Biodivers Inform. 2005;2:1-10.

44. Manel S, Dias JM, Buckton ST, Ormerod SJ. Alternative methods for predicting species distribution: an illustration with Himalayan river birds. J Appl Ecol. 1999;36:734-47.

45. Araujo MB, Pearson RG, Thuiller W, Erhard M. Validation of species-climate impact models under climate change. Global Change Biol. 2005;11:1504-13.

46. Guisan A, Zimmermann NE. Predictive habitat distribution models in ecology. Ecol Model. 2000;135:147-86.

47. Peterson AT. Uses and requirements of ecological niche models and related distributional models. Biodivers Inform. 2006;3:59-72.

48. Thuiller W. BIOMOD - optimizing predictions of species distributions and projecting potential future shifts under global change. Global Change Biol. 2003;9:1353-62.

49. Levine RS, Peterson AT, Benedict MQ. Geographic and ecologic distributions of the *Anopheles gambiae* complex predicted using a genetic algorithm. Am J Trop Med Hyg. 2004;70:105-9.

50. Adene DF, Oguntade AE. The structure and importance of the commercial and village based poultry industry in Nigeria. Rome, Italy: FAO; 2006.

51. Chotpitayasunondh T, Ungchusak K, Hanshaoworakul W, Chunsuthiwat S, Sawanpanyalert P, Kijphati R, et al. Human disease from influenza A (H5N1), Thailand, 2004. Emerg Infect Dis. 2005;11:201-9.

52. de Benedictis P, Joannis TM, Lombin LH, Shittu I, Beato MS, Rebonato V, et al. Field and laboratory findings of the first incursion of the Asian H5N1 highly pathogenic avian influenza virus in Africa. Avian Pathol. 2007;36:115-7.

53. WHO. Influenza pandemic risk assessment and preparedness in Africa. 2005 [cited 8th July, 2007]; Available from: http://www.afro.who.int/csr/epr/avian_flu/afr_avian_flu_31_10_05.pdf

54. Gilbert M, Chaitaweesub P, Parakarnawongsa T, Premashthira S, Tiensin T, Kalpravidh W, et al. Free-grazing ducks and highly pathogenic avian influenza, Thailand. Emerg Infect Dis. 2006;12:227-34.

55. Chen HL, Li YB, Li ZJ, Shi JZ, Shinya K, Deng GH, et al. Properties and dissemination of H5N1 viruses isolated during an influenza outbreak in migratory waterfowl in western China. J Virol. 2006;80:5976-83.

56. Olsen B, Munster VJ, Wallensten A, Waldenstrom J, Osterhaus A, Fouchier RAM. Global patterns of influenza A virus in wild birds. Science. 2006;312:384-8.

57. Yasue M, Feare CJ, Bennun L, Fiedler W. The epidemiology of H5N1 avian influenza in wild birds: why we need better ecological data. BioScience. 2006;56:923-9.

58. Kilpatrick AM, Chmura AA, Gibbons DW, Fleischer RC, Marra PP, Daszak P. Predicting the global spread of H5N1 avian influenza. Proc Natl Acad Sci USA. 2006;103:19368-73.

CHAPTER 1 FIGURES AND TABLES

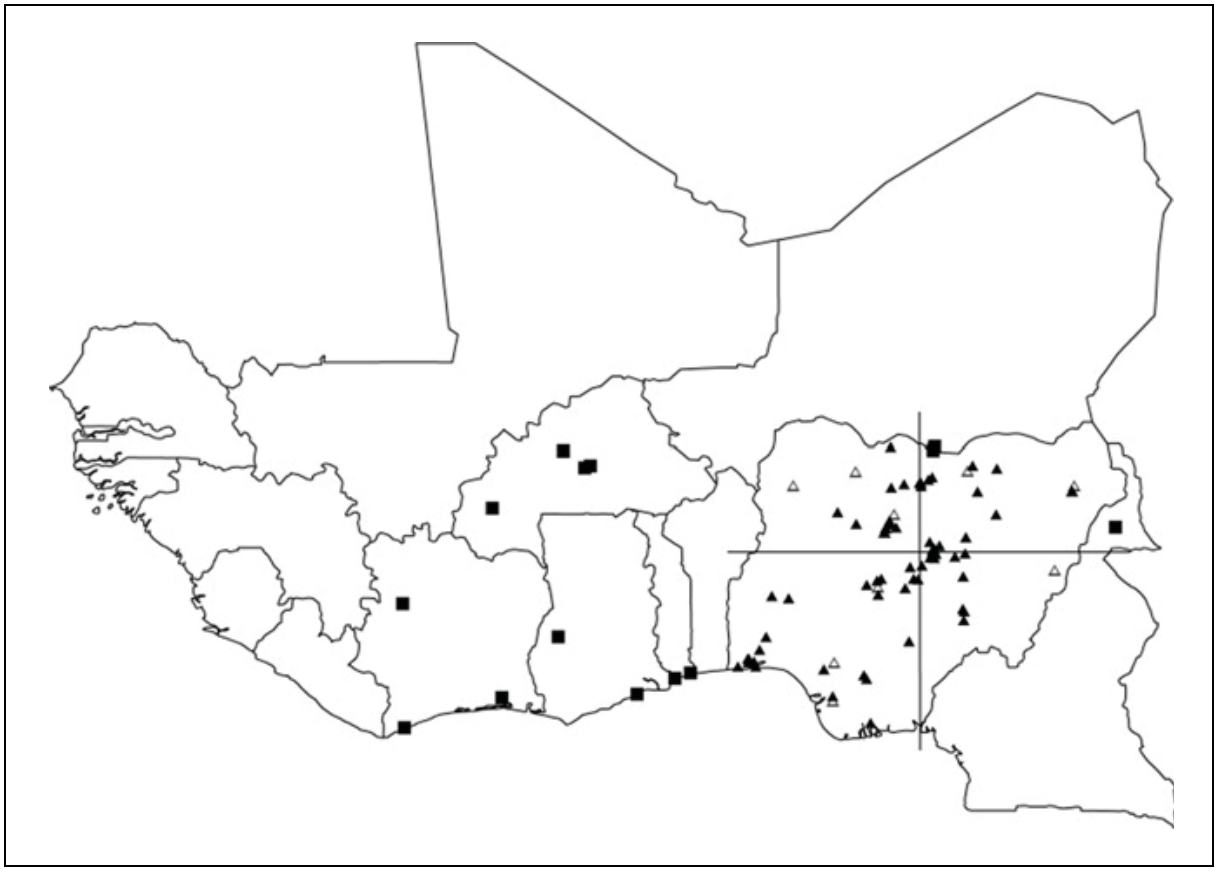

Figure 1.1 Occurrence data for highly pathogenic H5N1 in West Africa used in this study. Filled triangles = 2006 cases in Nigeria; open triangles = Nigerian ‘‘YEAR’’ cases (November 2006—January 2007); squares = occurrences from elsewhere in West Africa. The two dashed lines overlain on Nigeria indicate the median latitude and longitude used for spatial subsets of Nigerian occurrence data (EW, NS and DIAG tests).

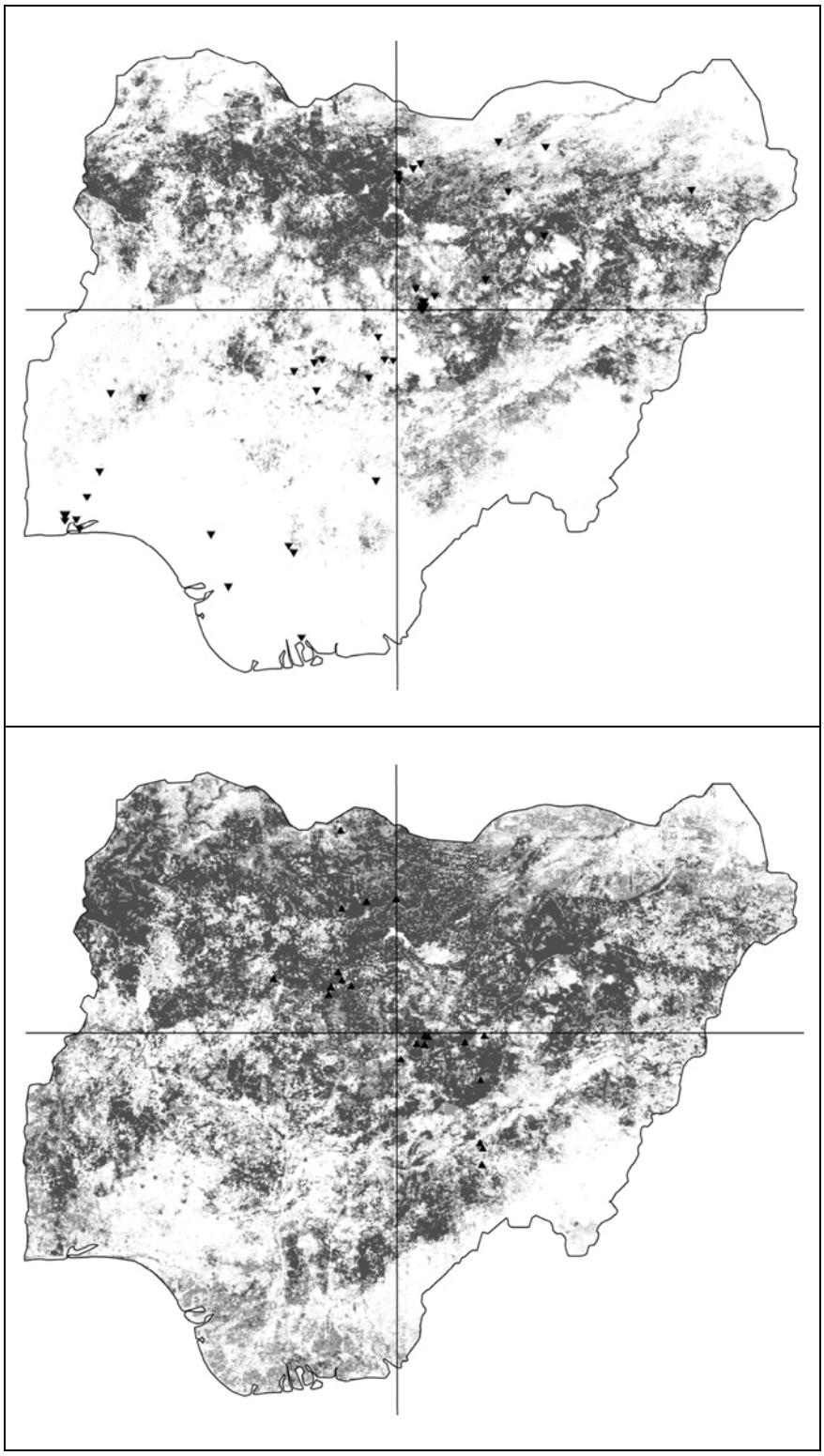

Figure 1.2 Example of spatially stratified tests of ENM predictions of highly pathogenic H5N1 distributions within Nigeria. Here, occurrences in on-diagonal quadrants were used to predict distributions of cases in off-diagonal quadrants, and vice versa. Model predictions are shown as ramps of model agreement in predictions: white = 10 of 10 models predict absence; light grey = 1—5 of 10 models predict potential presence; dark grey = 6—9 of 10 models predict potential presence; and darkest grey = all 10 models agree in predicting potential presence. Only independent test points are plotted on each map.

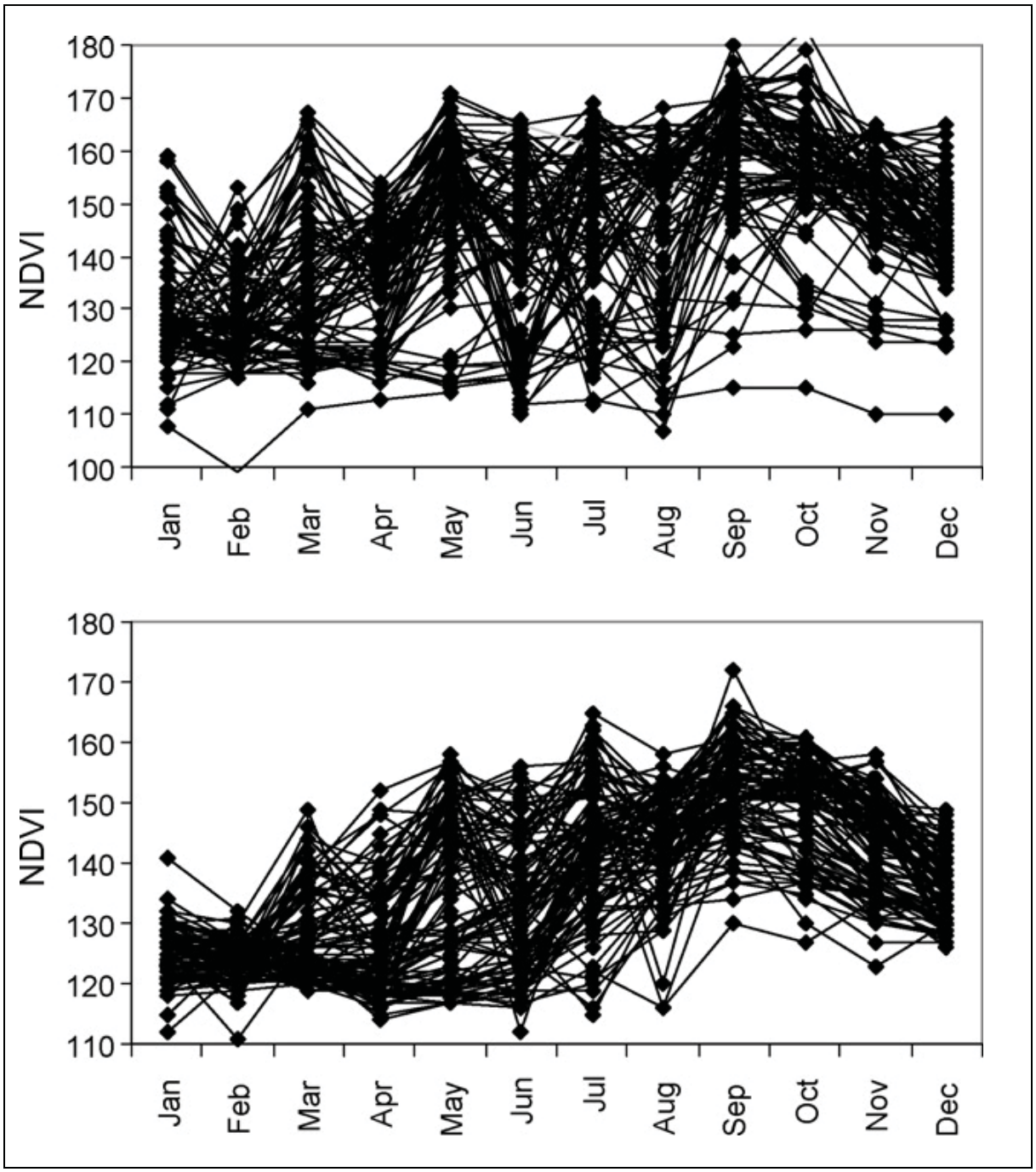

Figure 1.3 Summary of Normalized Difference Vegetation Index 'greenness' profiles of Nigeria through the year for (top) 85 randomly selected points of predicted absence, and (bottom) 96 randomly selected points of predicted presence. Note considerable reduction of variance and accentuated seasonality in the bottom graph.

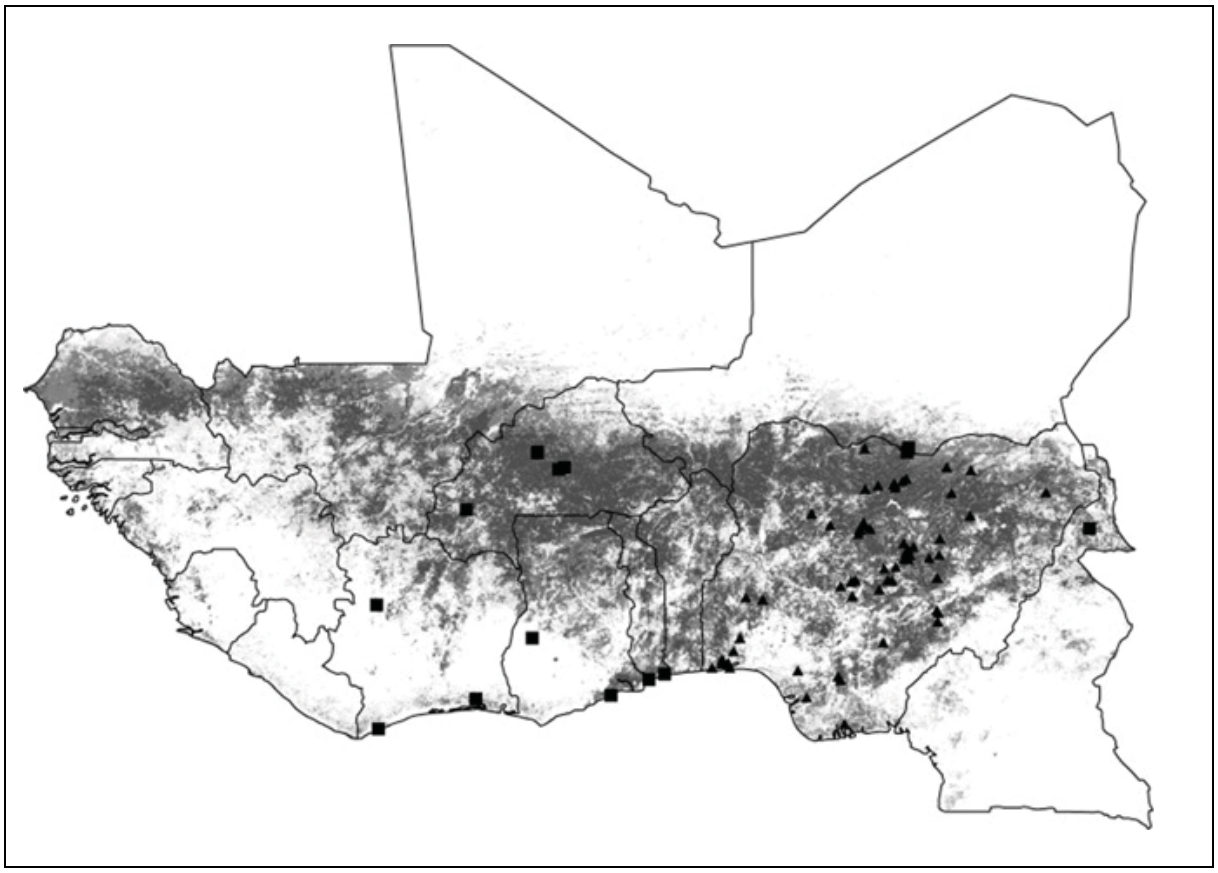

Figure 1.4 Regional projection across all of West Africa of highly pathogenic H5N1 ecological niche model results based on 2006 Nigeria occurrences (black triangles). Model predictions are shown as ramps of model agreement in predictions: white = 10 of 10 models predict absence; light grey = 1—5 of 10 models predict potential presence; dark grey = 6—9 of 10 models predict potential presence; and darkest grey = all 10 models agree in predicting potential presence. Solid triangles indicate independent test occurrence data from Nigeria; solid squares indicate independent test data from other West African countries.

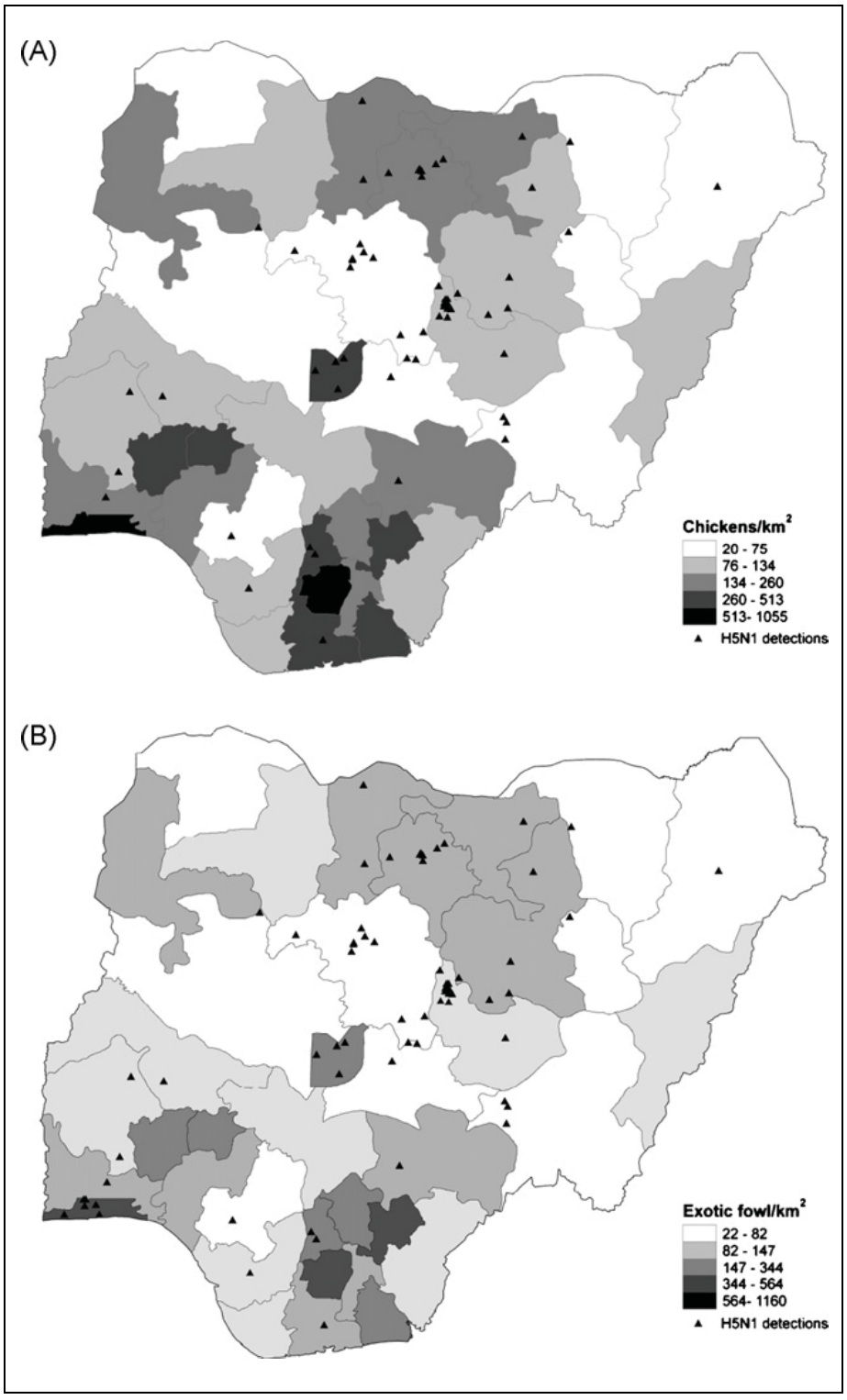
(A)
Chickens/km²
20 - 75
76 - 134
134 - 260
260 - 513
513- 1055
H5N1 detections
(B)
Exotic fowl/km²
22 - 82
82 - 147
147 - 344
344 - 564
564- 1160
H5N1 detections

Figure 1.5 Summary of population density (individuals/km^2) in Nigeria for (A) backyard chickens and (B) exotic fowl (ducks, guinea fowls, ostriches, and pigeons). Solid triangles show the distribution of highly pathogenic H5N1 cases. *Source*: Federal Ministry of Agriculture and Rural Development, Federal Department of Livestock and Pest Control Services in Adene and Oguntade [(50)].

Table 1.1 Summary of model predictions and tests in this study, illustrated by information for the threshold ≥ 5 of 10 best subset models predicting potential for presence

	Sample size (train/test)	Prop. area[a]	*k* (= no. of successes)	Cumulative binomial probability	No. of thresholds significant[b]
RND					
RND1 predicts RND2	37/34	0.466	24	0.008	10
RND2 predicts RND1	34/37	0.487	24	0.003	10
NS					
N predicts S	35/36	0.045	10	<0.001	10
S predicts N	36/35	0.385	23	<0.001	8
EW					
W predicts E	36/35	0.502	23	0.034	9
E predicts W	35/36	0.455	9	0.987	0
DIAG					
On predicts Off	24/47	0.592	21	<0.001	10
Off predicts On	47/24	0.197	17	0.003	10
YEAR					
2006 predicts 2007	72/12	0.610	9	0.094	2
WA					
Nigeria predicts West Africa	72/14	0.272	9	<0.001	8

[a] Prop. area indicates the proportion of the test region predicted present at that threshold.
[b] Number of thresholds (out of 10) for which model predictions were significantly better than random expectations.

CHAPTER 2: ECOLOGY AND GEOGRAPHY OF AVIAN INFLUENZA (HPAI H5N1) TRANSMISSION IN THE MIDDLE EAST AND NORTH-EASTERN AFRICA

Published as Williams and Peterson*, 2009, International Journal of Health Geographics* 2009, **8**:47 http://www.ij-healthgeographics.com/content/8/1/47

ABSTRACT

Background: The emerging highly pathogenic avian influenza strain H5N1 ("H5N1") has spread broadly in the past decade, and is now the focus of considerable concern. We tested the hypothesis that spatial distributions of H5N1 cases are related consistently and predictably to coarse-scale environmental features in the Middle East and north-eastern Africa.

We used ecological niche models to relate virus occurrences to 8 km resolution digital data layers summarizing parameters of monthly surface reflectance and landform. Predictive challenges included a variety of spatial stratification schemes in which models were challenged to predict case distributions in broadly unsampled areas.

Results: In almost all tests, H5N1 cases were indeed occurring under predictable sets of environmental conditions, generally predicted absent from areas with low NDVI values and minimal seasonal variation, and present in areas with a broad range of and appreciable seasonal variation in NDVI values. Although we documented significant predictive ability of our models, even between our study region and West Africa, case occurrences in the Arabian Peninsula appear to follow a distinct environmental regime.

Conclusion: Overall, we documented a variable environmental "fingerprint" for areas suitable for H5N1 transmission.

BACKGROUND

Highly pathogenic avian influenza of the strain H5N1 (hereafter "H5N1") has received considerable attention as an emerging virus with human pandemic potential [1,2] since it was first shown to be the cause of human morbidity and mortality in Hong Kong in 1997 [3]. To date, however, its most serious impacts have been on domestic poultry: millions of domestic birds have been killed by H5N1 infection, and >230 million domestic birds have been culled to contain the spread of the virus [4]. In contrast to the dramatic publicity, relatively few human cases are confirmed: at the time of writing, 385 human H5N1 cases had been documented, of which 243 (63.1%) were fatal [5], from 60 countries [6]. Human cases however, may eventually prove to be significantly underreported, reducing case-fatality rates.

Until spring 2005, H5N1 was restricted to East and Southeast Asia [6]. Between May and June 2005, however, >6000 birds of 8 wild waterbird species were found dead at Qinghai Lake, in central China: H5N1 was detected in 15 birds of 6 wild species [7], some migratory, fuelling fears of broader spread [8]. This event apparently marked a turning point in the spread of the virus: by early 2006, it had been detected widely across South Asia, Western Europe, and parts of Africa [6]. However, whether this rapid spread resulted from accelerated dispersal or from improved surveillance detecting existing infections remains debatable [9].

The first Middle Eastern detection of H5N1 was in Turkey in October 2005, in a flock of "backyard" turkeys (see Table 2.1). Further detections followed in 7 Balkan countries (Bosnia-Herzegovina, Bulgaria, Croatia, Greece, Romania, Serbia and Montenegro, and Slovenia; November 2005 – March 2006), more broadly in the Middle East (Egypt, Iraq, Iran, Israel, Jordan, Kuwait, Palestinian Territories; November 2005 – March 2006), and the Caucasus (Azerbaijan and Georgia; January – February 2006) by March 2006. The virus was detected in Sudan and Djibouti in April 2006 and in Saudi Arabia in March 2007 [6]. Countries in the region yet to record cases include the richest (Bahrain, Qatar, United Arab Emirates) and the poorest (Eritrea, Somalia, Yemen).

The concept of ecological niche describes the distinct ecological requirements that determine occurrences of organisms and other biological phenomena (including disease transmission, such as H5N1), and niches are customarily defined at relatively coarse spatial scales to avoid complexities of biotic interactions. The variables used to define it are described in Methods. Here, we use ecological niche modeling to provide a landscape-scale perspective on the ecological context of H5N1 occurrences across the Middle East and northeastern Africa (Figure 2.1), following protocols developed in an earlier study in West Africa [10]. In the previous study, we associated H5N1 case occurrences with month-to-month variation in 'greenness,' in the form of Normalized Difference Vegetation Index (NDVI) values derived from the Advanced Very High Resolution Radiometer (AVHRR) satellite, in an evolutionary-computing environment. We thus produced ecological niche models of H5N1 occurrence that had significant predictive ability, suggesting that H5N1occurs under consistent and predictable environmental circumstances in West Africa. In this study, we demonstrate consistent, predictable environmental conditions associated

with H5N1 occurrences across the Middle East and northeastern Africa, albeit not without notable exceptions.

RESULTS

Most of the 9 tests conducted as part of this study indicated that independent test points coincided with ENM predictions significantly better than random expectations (see Table 2.2), although not without exceptions. In other words, in general, models based on known H5N1 occurrences were able to anticipate spatial distributions of independent samples of H5N1 based on their environmental attributes. The details of these test results follow.

Predictivity across study region

The model based on all OIE points showed significant predictive ability when tested with the ProMed human case-occurrence data (see Figure 2.2; Table 2.2). Potential for H5N1 occurrence was predicted along the major rivers of the region (Euphrates, Nile, Tigris), across most of the Caucasus, southern Sudan, and in Ethiopia, Greece, northern and western Iran, southern Somalia, and Turkey. The virus was not predicted to have high probability of occurrence in the Sahara, nor more generally in arid areas. Egypt was largely predicted unsuitable, except for the fertile, densely populated Nile Valley. This model's predictions were significantly better than random expectations at all 10 thresholds; for example, at the 5 models out of 10 threshold, this model predicted 82.4% of the independent testing points in just 41.2% of the region ($P < 0.001$).

Single testing regions

These analyses tested the ability of models based on known occurrences across three subregions to predict patterns of occurrence in the fourth subregion. These tests indicated, for the most part, significant predictive power of the models (see Figure 2.3; Additional file 2.2). All thresholds of prediction were significant for prediction of occurrences in Levant-Iran by the remaining three regions, 8 of 10 thresholds were significant for predictions in northeastern Africa, and 7 of 10 were significant for predictions in Balkans-Caucasus. The model predicting distributions in the Arabian Peninsula performed more weakly than the other models, with only 4 of 10 thresholds significant and considerable deviation from coincidence when inspected visually (Figure 2.3).

Single predictor regions

Predictions of independent points across landscapes based on single training regions were less successful (see Figure 2.3; Table 2.2). Indeed, only 2 of 4 models showed any predictive ability. Predictions from northeastern Africa to the rest of the region were significant at 8 of 10 thresholds, and projections from Levant-Iran to the rest of the region were significant at 5 of 10 thresholds. Projections based on models trained in the Arabian Peninsula and Balkans-Caucasus showed no significant ability when challenged to predict occurrences in the remaining regions. Once again, visually, the Arabian Peninsula models performed particularly poorly (Figure 2.3).

Partial ROC analyses

The partial ROC analyses (see Table 2.2) were largely consistent with the cumulative binomial probability results (see Table 2. 2). According to these tests, all single-testing-

region predictions were successful (i.e., $P \leq 0.001$) while 2 of 4 single predictor regions (Levant-Iran, northeastern Africa) were significantly better than random ($P \leq 0.005$). The partial ROC evaluation of the overall prediction of the ProMed data was similarly significant ($P < 0.01$).

The NDVI data used in this study summarize photosynthetic mass of vegetation, and how this quantity changes through the year. Models based on case occurrences from across the region were compared in detail in terms of environmental conditions reconstructed as suitable versus unsuitable (Figure 2.4), approximating a visualization of the ecological niche estimated by each model. In the all region model, H5N1 was predicted absent from areas with low NDVI values and low seasonality, but present in areas with a broad range of NDVI values (from low to high) that showed marked seasonal variation. In contrast, the Arabian Peninsula model predicted presence in low NDVI areas with minimal seasonality, and absence from areas showing a broad range of NDVI values (from low to high) and seasonal variation. As such, the model with the least predictive ability (i.e., the Arabian Peninsula model) was the inverse of the one that had good predictive ability (i.e., the all-region model). It is interesting to compare these results to those from our previous West African models [10]. There, virus presence was predicted mostly in savannah and woodland habitats, whereas absence was predicted in montane areas, coastal mangroves, the freshwater swamps of the Niger Delta, and from rainforest areas: areas of highest predicted H5N1 risk were highly variable seasonally, just as with our all-region model.

The spatial limits of the predictivity we have documented remain an open question [10]. The initial demonstration of predictable H5N1 geography across West Africa is now supported

by replication of the modeling protocol across the Middle East. Projection of the Middle East model to West Africa, and testing with independent points from that region [10,11] (N = 101;) demonstrated significant predictivity at all thresholds with both the binomial test, and the partial ROC approach. This new prediction (Figure 2.5) is broadly quite similar to the first West African prediction [10], although differences are evident. In particular, the Middle East model predicts H5N1presence in forest and mountains, whereas the West African model did not. The two models are based upon different sets of environmental layers, so some level of difference is not surprising.

DISCUSSION

Our results are generally consistent with earlier predictions of the ecological niche of H5N1 in West Africa [10]. Most Middle Eastern and northeastern African models predicted suitable areas for H5N1 transmission in human-habitable areas, such as the Nile Valley, the Levant, the Fertile Crescent, and the savannas of southern Sudan. The major difference between the two sets of models is that most Middle Eastern and northeastern African models predicted suitability in montane areas (Caucasus, Ethiopian Highlands, northern and western Iran, and Turkey), whereas the West African models focused prediction of suitable areas in lowlands. These models agree most clearly in implicating areas with greatest seasonal variation as representing high H5N1 risk.

The major exception to the conclusion of predictivity of H5N1 in the Middle East and northeastern Africa were predictions involving the Arabian Peninsula, which were not generally statistically significantly better than random expectations. Indeed, in several areas, Arabian models were *inverse* to the rest of our predictions, predicting absence in

areas of presence and vice versa. That is to say, models based on Arabian Peninsula points predicted H5N1 presence in deserts, but not in mountains, the Levant, the Fertile Crescent, or in the Sudanese savannah, and only at low levels of model agreement in the Nile Valley (see Figure 2.6; Table 2.2).

It is interesting that Arabian models should produce predictions so inconsistent with those from the rest of the study area (see table 2.2): for example, the distribution of Arabian Peninsula occurrences could not be predicted with any confidence by models trained in the remainder of the region, and conversely, Arabian Peninsula points were unable to predict occurrences across the Balkans, Caucasus, Levant, Iran, or northeastern Africa unsuccessfully. Three major H5N1 outbreaks occurred in the Arabian Peninsula: in Kuwait, Ar-Riyad (city), and southern Ar-Riyad (province), none of which is predicted strongly by models trained elsewhere (Figures 2.2 and 2.3). Given the rather extreme arid conditions in the region, the Arabian Peninsula seems a harsh environment for both poultry and poultry diseases. We suspect that Arabian H5N1 occurs chiefly or only in human-subsidized habitats that would permit poultry to be raised: indeed, 26 of 30 reported Saudi Arabian cases were detected in commercial farms containing thousands to hundreds of thousands of poultry [(11)]. Perhaps, Arabian occurrence points reflect something other than the "ecological niche" of H5N1 in the subregion; for example, they may reflect principally the conditions under which poultry can be raised, albeit with considerable subsidy of water and shade, irrespective of disease distributions. We should add, though we suspect that such is not the case, the total lack of predictivity in the Arabian Peninsula raises the more troubling possibility that the correspondence between NDVI and disease occurrence in the rest of the region is coincidental. It is possible that H5N1 distribution is not driven by factors

correlated with NDVI seasonality, but by something that cannot be detected in the remotely sensed landscape.

Gilbert et al. [12] mapped the geographic distribution of suitable conditions for H5N1 across Southeast Asia, finding close associations between free-grazing domestic ducks in rice paddies and H5N1 cases. This result suggests that transmission risk could be mapped successfully in Southeast Asia, where duck production and rice cultivation are both extensive and intertwined, and that duck production may be an important driver of H5N1 persistence. The authors stated that large numbers of Anatidae concentrate in the Nile Delta, and that the Hadejia Jama'are river system of Nigeria is also an important area of duck production. FAO reports a combined domestic duck and goose population of 18.3 million for Egypt in 2004 [13], presumably concentrated in the Nile Delta and Valley (along with virtually the entire human population and all productive agricultural land), joined in winter by large flocks (several hundreds of thousands [14]) of wild aquatic birds. Figures are unavailable for domestic Anatidae in Nigeria, although numbers of undifferentiated "exotic poultry" (ducks, geese, turkeys, guinea fowl, ostriches, etc.) in the 5 states bordering Hadejia Jama'are were around 7.5 million birds in 2003 [15]. Egypt and Nigeria both produce substantial rice crops (on 613 000 and 2 725 000 ha of land, respectively) [16].

Although total area under rice cultivation and total Anatidae populations are far higher in East Asia than in Egypt, the ratio of domestic Anatidae to area of rice production is considerably higher than in Thailand and Vietnam (see Table 2.3), and about the same as that found in China. If grazing of domestic Anatidae in rice paddies does play an important role in driving H5N1 persistence and if duck-raising in the Middle East parallels that in

East Asia, we might, expect persistence in China, Egypt, and Iran, all countries with higher duck-to-rice production area ratios than Thailand (Table 2.3). On the other hand, cases of H5N1 have been numerous and widespread in Turkey, despite low numbers of Anatidae and little rice cultivation, suggesting that duck grazing in rice paddies is not the only factor in H5N1 transmission and persistence. Free mingling of backyard poultry and wild birds has been identified as a risk factor for H5N1 transmission [17, 18]. In Egypt, most domestic Anatidae are considered to be backyard (64% of ducks and "all" geese), whereas the majority of chickens (63%) are produced in commercial operations, apparently typifying the poultry industry of North Africa and the Middle East [19].

Our models and predictions cannot shed new light on the comparative roles of poultry and wild birds in H5N1transmission. One of the most important challenges for our analyses is distinguishing true ecological biases in case distributions (i.e., the ecological niche!) from the spatial and ecological biases in distributions of H5N1 hosts. In some regions (Nile Delta, Fertile Crescent, Levant, Turkey, western Iran), our predictions showed marked coincidence with poultry distributions (Figure 2.7). However, our models failed to predict the high poultry concentrations in western Saudi Arabia and the Arabian Gulf states as forming part of the potential distribution of H5N1, despite detections in Kuwait; as noted previously, our ability to predict H5N1 distribution patterns in the Arabian Peninsula was poor in all comparisons.

CONCLUSION

H5N1 detection data used for the development of these models are dominated by transmission among flocks of several poultry species. Given that detection data are so variable in terms of species composition (i.e., taxa, and number of taxa affected), husbandry method (high biosecurity, backyard, etc), origin (home-hatched, purchased, native-hatched, imported legally or illegally), and domestication, it is hard to define mechanisms driving transmission. We do not, however, find that our models are simply reproducing the spatial distributions of poultry flocks. Several ecologically-biased elements in the H5N1 transmission cycle could explain the predictivity we detected: introduction of H5N1 by migratory birds [20,21], transmission among poultry flocks [22,23], areas important for importation of poultry or hobby birds (legal or illegal) [24], or even transportation routes (e.g., roads, rivers). Inconsistencies in predictions based on H5N1 occurrences from different subregions suggest that certain of these factors may have greater importance in some subregions than in others. In the Middle East, at least, we observe coincidence between human populations and H5N1 cases, although, of course, this observation may simply point to the fact that influenza surveillance is more intensive in populated areas.

METHODS

Ecological Niche Models

The ecological niche models (ENMs) developed in this study are based on the idea that organisms and other biological phenomena (including disease transmission) have distinct ecological requirements that determine their occurrences in time and space [25]. In general, disease applications of ENM balance between focusing on individual species in the transmission system and using the integration of the whole system as a "black box"

determining transmission to some species or biological phenomenon of interest [26,27]. In this contribution, given the as-yet poorly characterized avian reservoir of H5N1, we focus on all cases of H5N1, effectively treating the transmission system as a black box. We thus attempt to model the transmission of a single pathogen based on its appearance in a multi-species system (i.e., the subset of animals in which H5N1 has been detected), in this case, dominated by distributions of domestic birds. In this sense, we deviate somewhat from the classical ENM approaches, which are based on single species occurrence-environment correspondence. ENMs have been developed via diverse methodological approaches [28-31]; however, the method most frequently applied to questions of disease transmission has been the Genetic Algorithm for Rule-set Prediction (GARP), an evolutionary-computing approach [32,33].

Input data

This study was based on H5N1 animal case-occurrence data for 2005–2008 from the Middle East and north-eastern Africa. Data were drawn from the World Organisation for Animal Health (OIE) [11], consisting of 610 unique locations, including isolations from wild birds, zoo birds, commercial poultry, and backyard poultry (Figure 2.1). This survey of occurrences includes birds assumed to be raised under strict biosecurity control, as well as birds raised with none; it similarly includes birds raised in strictly monospecific farms, multispecies assemblies mingling freely with wild birds (and other fauna), and even pets in a children's kindergarten. The database is composed of detections in at least 18 species of birds, although reporting standards are not consistent, so all too frequently information about hosts is either vague or absent. Most detections occurred in anthropogenic habitats. Our study area included 25 countries and one territory, ranging from Greece to the

northwest, Somalia to the southwest, Georgia to the north, and Iran to the east. We assembled a complementary set of 17 unique and non-overlapping human cases occurrences from the archives of the International Society for Infectious Disease (ProMed Avian Influenza archive) [34] from the region (Figure 2.2) with which to test model predictions. All textual descriptions of occurrence localities were converted to geographic coordinates accurate to the nearest 0.01° using the GeoNet Names Server http://gnswww.nga.mil/geonames/GNS/index.jsp, Alexandria Digital Library Gazetteer http://middleware.alexandria.ucsb.edu/client/gaz/adl/index.jsp, and other sources [35].

We based ENM development on the 610 OIE localities for which geographic coordinates were provided with a precision of at least 0.01°; duplicate localities (i.e., multiple occurrences in the same 8 × 8 km grid square) were discarded. Geographic coordinates in the OIE data set were drawn from global positioning system recordings for the point of detection of H5N1 cases [11]. They thus specify the spatial position of H5N1 occurrences, and probably represent the coarse-scale ecological conditions under which H5N1 transmission occurs. Given that the spatial pattern of H5N1 outbreaks has been on rather fine spatial scales, our previous experience with niche modeling and H5N1 outbreaks indicates that spatial resolutions on the order of 1–10 km are necessary, making use of climate-based data layers impractical. Environmental data sets included 12 monthly composite remotely-sensed data layers for Nov 1999 – Oct 2000, each summarizing maximum Normalized Difference Vegetation Index (NDVI; native spatial resolution 8 × 8 km) values [36]; although not exactly coincident with occurrence data temporally, these data provided an exemplar year of landscape variation in greenness. As NDVI is derived from reflectance in the visible and near-infrared domains, and as such is sensitive to

photosynthetic activity and is closely correlated with photosynthetic mass [36], the NDVI time series used here summarizes aspects of land cover and vegetation phenology across the region. A year 2001 MODIS-based vegetation continuous fields dataset summarizing percent tree cover was also used (native spatial resolution 500 m) [37]. Finally, we also included 3 data sets summarizing aspects of topography: slope, aspect, and compound topographic index (which summarizes tendency to pool water), from the U.S. Geological Survey's Hydro-1K data set (native resolution 1km) [38]. We deliberately excluded data on elevation from the study to avoid confusion caused by indirect variables. Climate data were not included in these analyses for lack of sufficiently high-resolution data sets across the region.

The GARP algorithm

The Genetic Algorithm for Rule-set Prediction (GARP) has been applied widely to questions of disease transmission [26, 39], and its predictive ability has been tested under diverse circumstances [30,40,41]. Although GARP has seen criticism in some comparative studies [30], more recent studies have indicated considerably better performance [42,43] and some artifactual causation of previous results [44]. As such, we used GARP for ENM development.

In general, we developed tests based on spatially stratified subsets of available occurrence information set aside prior to model development. Of occurrence data actually input into GARP, the program divides occurrence data randomly into three subsets: training data (25%; for rule development), intrinsic testing data (25%; for evaluation of rules) and extrinsic testing data (50%; for evaluation of model quality, see below). Spatial predictions

of presence versus absence can include two types of error: false negatives (areas of actual presence predicted absent) and false positives (areas of actual absence predicted present) [45] – rule performance in each of these dimensions is evaluated via the intrinsic testing data set. Changes in predictive accuracy from one iteration to the next are used to evaluate whether particular rules should be incorporated into the model or not, and the algorithm runs either 1000 iterations or until convergence [33]. The final rule-set is then used to query the environmental data sets across the study region to identify areas fitting the rule set predictions to produce a hypothesis of the potential geographic distribution of the species [25]. Since GARP processing includes several random-walk components, each replicate model produces distinct results, representing alternative solutions to the optimization challenge. Following best-practices approaches [40], we developed 100 replicates of each model. We filtered these replicates based on their error characteristics, retaining the 20 with lowest false negative rates (= percentage of independent testing points falling in areas not predicted to be suitable), and then retained the 10 (of the 20) closest to the median of proportional area predicted present, an index of false-positive error rates [40]. A consensus of these 'best subset' models was then developed by summing values for each pixel in the map to produce final predictions of potential distributions with 11 thresholds (integers from 0 to 10).

Modeling and testing approach

This study focuses on the question of whether H5N1transmission in the Middle East and northeastern Africa occurs under a consistent and predictable set of environmental conditions. As such, we developed a series of tests of model predictivity; in each case, models were developed and predictions tested using spatially independent suites of

occurrence data. Model tests were based on 4 spatial subsets of the Middle Eastern and northeastern African occurrence data (Figure 2.1): Arabian Peninsula (Bahrain, Kuwait, Oman, Qatar, Saudi Arabia, United Arab Emirates, Yemen; $N = 31$), Balkans-Caucasus (Armenia, Azerbaijan, Cyprus, Georgia, Greece, Turkey; $N = 175$), Levant-Iran (Iran, Iraq, Israel, Lebanon, Palestinian Territories, Syria; $N = 18$), and northeastern Africa (Djibouti, Egypt, Eritrea, Ethiopia, Somalia, Sudan; $N = 386$).

The basic design of testing included three schemes for subdividing available occurrence data, as follows:

1. *Single testing regions*: We combined each possible set of 3 subregional occurrence datasets to develop ENMs that were tested with the fourth subregion. Total 4 tests.

2. *Single predictor regions*: Occurrence data for each subregion were used to develop predictive models that were projected to the rest of the region for testing (e.g., Arabian Peninsula data points used to build predictions for the combination of Levant-Iran, northeastern Africa, and Balkans-Caucasus). Total 4 tests.

3. *Predictivity across study region*: We developed ENM predictions based on all OIE veterinary cases in the region, and tested its prediction based on coincidence of predictions with the 17 independent ProMed human cases. One test.

The customary approaches to spatial model validation (e.g. simple receiver operating characteristic, kappa statistics) are not applicable to situations in which presence only data

are the only information available [45,46]. As such, we validated models using two approaches. First, we calculated binomial probabilities that observed coincidence of predictions and independent test data is no better than random, with the probability of k successes in N trials depending on p, the probability of success in any one trial; we estimated p as the proportion of the testing area predicted present, and k as the number of the N testing points successfully predicted by the model prediction [40]. Binomial probabilities were calculated for each of the 10 thresholds representing predictions of presence (1 = broad, 10 = narrow), in each case testing whether predictivity was better than expected by chance.

Second, we followed Phillips et al. [47] in modifying receiver operating characteristic curves (ROCs) so as not to depend on absence data. We calculated the area under the curve (AUC) of the ROC, a statistical technique that has become a dominant tool in evaluating the accuracy of models predicting distributions of species [47]. However, when comparing two ROCs, AUC systematically undervalues models that do not provide predictions across the entire spectrum of proportional areas in the study area (such as GARP, the modeling approach used here) [48]. In addition current ROC approaches inappropriately weight the two error components (omission and commission) equally. Accordingly, we use a modification of ROC that remedies these problems: partial-area ROC approaches that evaluate only over the spectrum of the prediction and that allow for differential weighting of the two error components [48].

We carried out partial ROC analyses [48] for each model, all based solely on independent testing points not used to train the models in areas from distinct regions(s) to which models

were projected. AUCs were limited to the proportional areas over which models actually made predictions, and only omission errors of <5% were considered (i.e., E = 5% [48]). We calculated partial AUCs using a program based on the trapezoid method [49] kindly developed by N. Barve (in prep.), and present our ROC results as the ratio of the model AUC to the null expectation ("AUC ratio") [48]. Bootstrapping manipulations to permit evaluation of statistical significance of AUCs (as compared with null expectations) were achieved by resampling 50% of the test points with replacement 1000 times from the overall pool of testing data; one-tailed significance of differences in AUC (i.e. elevation above the line of null expectation) was assessed by counting the number of bootstrap replicates with AUC ratios <1.

ACKNOWLEDGEMENTS

Thanks to Monica Papeş, and Yoshinori Nakazawa for expert assistance with GIS and satellite imagery. Thanks to Wildlife Conservation Society's Global Avian Influenza Network for Surveillance and the U.S. Geological Survey's National Biological Information Infrastructure for funding this work.

CHAPTER 2 REFERENCES

1. Guan Y, Poon LLM, Cheung CY, Ellis TM, Lim W, Lipatov AS, et al. H5N1 influenza: a protean pandemic threat. Proc Nat Acad Sci USA 2004; 101:8156-8161.

2. Webby RJ, Webster RG: Are we ready for pandemic influenza? Science 2003; 302:1519-1522.

3. de Jong JC, Claas ECJ, Osterhaus A, Webster RG, Lim WL: A pandemic warning? Nature 1997; 389:554.

4. Whitworth D, Newman SH, Mundkur T, Harris P, editors. Wild birds and avian influenza: an introduction to applied field research and disease sampling techniques. FAO Animal Production and Health Manual, No. 5 ed. Rome; 2007.

5. WHO. *Cumulative number of confirmed human cases of avian influenza A(H5N1) reported to WHO*. 2009 30 December 2009 [cited 2009 16 September]; Available from: http://www.who.int/csr/disease/avian_influenza/country/cases_table_2009_12_30/en/index.html

6. WHO. H5N1 avian influenza: Timeline of major events. 2009 [cited 2009 February]; Available from: http://www.who.int/csr/disease/avian_influenza/Timeline_09_02_23.pdf

7. Chen HX, Shen HG, Li XL, Zhou JY, Hou YQ, Guo JQ, et al. Seroprevalence and identification of influenza A virus infection from migratory wild waterfowl in China (2004-2005). J Vet Med B. 2006;53:166-70.

8. Liu J, Xiao H, Lei F, Zhu Q, Qin K, Zhang XW, et al. Highly pathogenic H5N1 influenza virus infection in migratory birds. Science. 2005;309:1206.

9. Peterson AT, Benz BW, Papeş M. Highly pathogenic H5N1 avian influenza: entry pathways into North America via bird migration. PLoS One. 2007;2:e261.

10. Williams RAJ, Fasina FO, Peterson AT. Predictable ecology and geography of avian influenza (H5N1) transmission in Nigeria and West Africa. Trans R Soc Trop Med Hyg. 2008;102:471-9.

11. OIE. Update on highly pathogenic avian influenza in animals (Type H5 and H7). 2009 [cited 2009 22 September 2009]; Available from: http://www.oie.int/downld/avian%20influenza/A_AI-Asia.htm

12. Gilbert M, Xiao X, Pfeiffer DU, Epprecht M, Boles S, Czarnecki C, et al. Mapping H5N1 highly pathogenic avian influenza risk in Southeast Asia. Proc Natl Acad Sci USA. 2008;105:4769-74.

13. FAO - AGAL Animal Production and Health Division. GLiPHA, Global Livestock Production and Health Atlas. 2008 [cited 10th July]; Available from: http://www.fao.org/ag/aga/glipha/index.jsp

14. Hughes RH, Hughes JS. A directory of African wetlands. Gland, Switzerland and Cambridge, UK IUCN; 1992.

15. Adene DF, Oguntade AE. The structure and importance of the commercial and village based poultry industry in Nigeria. Rome, Italy: FAO; 2006.

16. International Rice Research Institute. Rice around the World. 2008 [cited 2008 10th July]; Available from: http://www.irri.org/science/ricestat/

17. Chotpitayasunondh T, Ungchusak K, Hanshaoworakul W, Chunsuthiwat S, Sawanpanyalert P, Kijphati R, et al. Human disease from influenza A (H5N1), Thailand, 2004. Emerg Infect Dis. 2005;11:201-9.

18. de Benedictis P, Joannis TM, Lombin LH, Shittu I, Beato MS, Rebonato V, et al. Field and laboratory findings of the first incursion of the Asian H5N1 highly pathogenic avian influenza virus in Africa. Avian Pathol. 2007;36:115-7.

19. Taha FA. The poultry sector in middle-income countries and its feed requirements: the case of Egypt. Washington, DC; 2003.

20. Chen HL, Li YB, Li ZJ, Shi JZ, Shinya K, Deng GH, et al. Properties and dissemination of H5N1 viruses isolated during an influenza outbreak in migratory waterfowl in western China. J Virol. 2006;80:5976-83.

21. Olsen B, Munster VJ, Wallensten A, Waldenstrom J, Osterhaus A, Fouchier RAM. Global patterns of influenza A virus in wild birds. Science. 2006;312:384-8.

22. Gilbert M, Xiao X, Domenech J, Lubroth J, Martin V, Slingenbergh J. Anatidae Migration in the Western Palearctic and Spread of Highly Pathogenic Avian Influenza H5N1 Virus. Emerg Infect Dis. 2006;12:1650-6.

23. Yasue M, Feare CJ, Bennun L, Fiedler W. The epidemiology of H5N1 avian influenza in wild birds: why we need better ecological data. BioScience. 2006;56:923-9.

24. Kilpatrick AM, Chmura AA, Gibbons DW, Fleischer RC, Marra PP, Daszak P. Predicting the global spread of H5N1 avian influenza. Proc Natl Acad Sci USA. 2006;103:19368-73.

25. Soberón J, Peterson AT. Interpretation of models of fundamental ecological niches and species' distributional areas. Biodivers Inform. 2005;2:1-10.

26. Peterson AT. Ecological niche modeling and spatial patterns of disease transmission. Emerg Infect Dis. 2007;12:1822-6.

27. McCormack J, Peterson A, Bonaccorso E, Smith T. Speciation in the highlands of Mexico: genetic and phenotypic divergence in the Mexican jay (*Aphelocoma ultramarina*). Mol Ecol. 2008;17:2505-21.

28. Austin MP, Nicholls AO, Margules CR. Measurement of the realized qualitative niche: environmental niches of five *Eucalyptus* species. Ecol Monogr. 1990;60:161-77.

29. Carpenter G, Gillison AN, Winter J. DOMAIN: a flexible modeling procedure for mapping potential distributions of animals and plants. Biodivers Conserv. 1993;2:667-80.

30. Elith J, Graham CH, Anderson RP, Dudik M, Ferrier S, Guisan A, et al. Novel methods improve prediction of species' distributions from occurrence data. Ecography. 2006;29:129-51.

31. Pearson RG, Dawson TP, Berry PM, Harrison PA. SPECIES: a spatial evaluation of climate impact on the envelope of species. Ecol Model. 2002;154:289-300.

32. Stockwell DRB, Noble IR. Induction of sets of rules from animal distribution data: a robust and informative method of analysis. Math Comput Simul. 1992;33:385-90.

33. Stockwell DRB, Peters DP. The GARP modelling system: problems and solutions to automated spatial prediction. Int J Geogr Inf Sci. 1999;13:143-58.

34. ProMED-mail. Avian Influenza Archive. 2008 [cited 2008 15th July]; Available from: http://www.promedmail.org

35. Rand McNally. New Millenium World Atlas Deluxe. Skokie, Illinois, USA: Rand McNally & Co.; 1998.

36. Tucker CJ. Red and photographic infrared linear combinations for monitoring vegetation. Remote Sens Environ. 1979;8:127-50.

37. Hansen M, DeFries R, Townshend JR, Carroll M, Dimiceli C, Sohlberg R. MOD44B: Vegetation Continuous Fields Collection 3, Version 3.0.0. University of Maryland, College Park; 2003.

38. USGS. HYDRO1K Elevation Derivative Database. 2001 [cited 2009 10th July]; Available from: http://edc.usgs.gov/products/elevation/gtopo30/hydro/index.html

39. Costa J, Peterson AT, Beard CB. Ecologic niche modeling and differentiation of populations of *Triatoma brasiliensis* Neiva, 1911, the most important Chagas' disease

vector in north-eastern Brazil (Hemiptera, Reduviidae, Triatominae). Am J Trop Med Hyg. 2002;67:516-20.

40. Anderson RP, Lew D, Peterson AT. Evaluating predictive models of species' distributions: criteria for selecting optimal models. Ecol Model. 2003;162:211-32.

41. Peterson A, Papeş M, Kluza D. Predicting the potential invasive distributions of four alien plant species in North America. Weed Sci. 2003;51:863-8.

42. McNyset K, Blackburn J. Does GARP really fail miserably? A response to Stockman et al. (2006). Divers Distrib. 2006;12:782-6.

43. Tsoar A, Allouche O, Steinitz O, Rotem D, Kadmon R. A comparative evaluation of presence-only methods for modelling species distribution. Divers Distrib. 2007;13:397-405.

44. Peterson AT, Papeş M, Eaton M. Transferability and model evaluation in ecological niche modeling: a comparison of GARP and Maxent. Ecography. 2007;30:550-60.

45. Fielding AH, Bell JF. A review of methods for the assessment of prediction errors in conservation presence/absence models. Environ Conserv. 1997;24:38-49.

46. Manel S, Dias JM, Buckton ST, Ormerod SJ. Alternative methods for predicting species distribution: an illustration with Himalayan river birds. J Appl Ecol. 1999;36:734-47.

47. Phillips SJ, Anderson RP, Schapire RE. Maximum entropy modeling of species geographic distributions. Ecol Model. 2006;190:231-59.

48. Peterson A, Papeş M, Soberon J. Rethinking receiver operating characteristic analysis applications in ecological niche modeling. Ecol Model. 2008;213:63-72.

49. Burden RL, Faires JD. Numerical Analysis. 8th ed. Belmont, California: Thomson Books; 2005.

CHAPTER 2 FIGURES AND TABLES

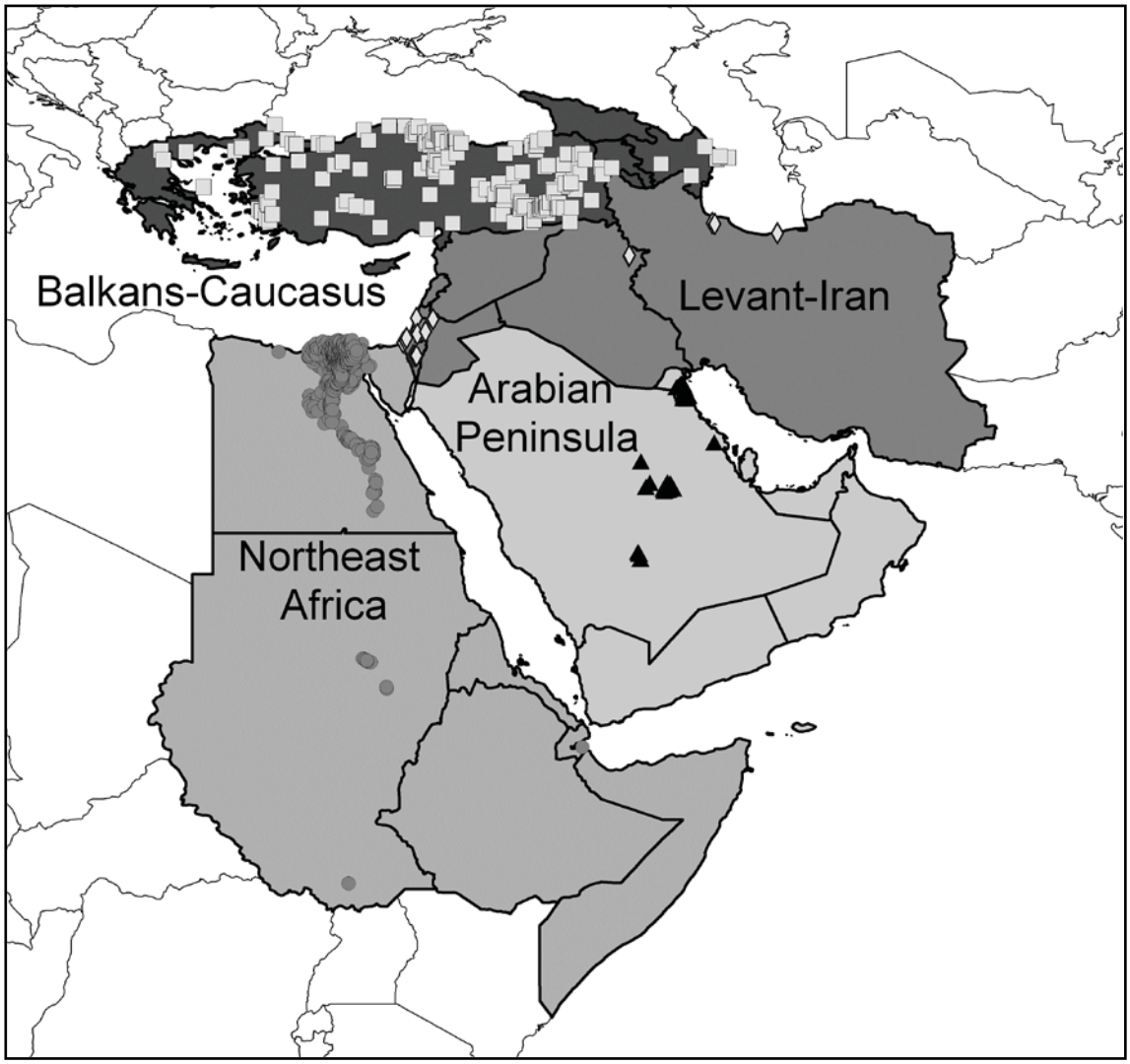

Figure 2.1 Occurrence data for H5N1 in the Middle East and northeastern Africa, and regional divisions used in this study.

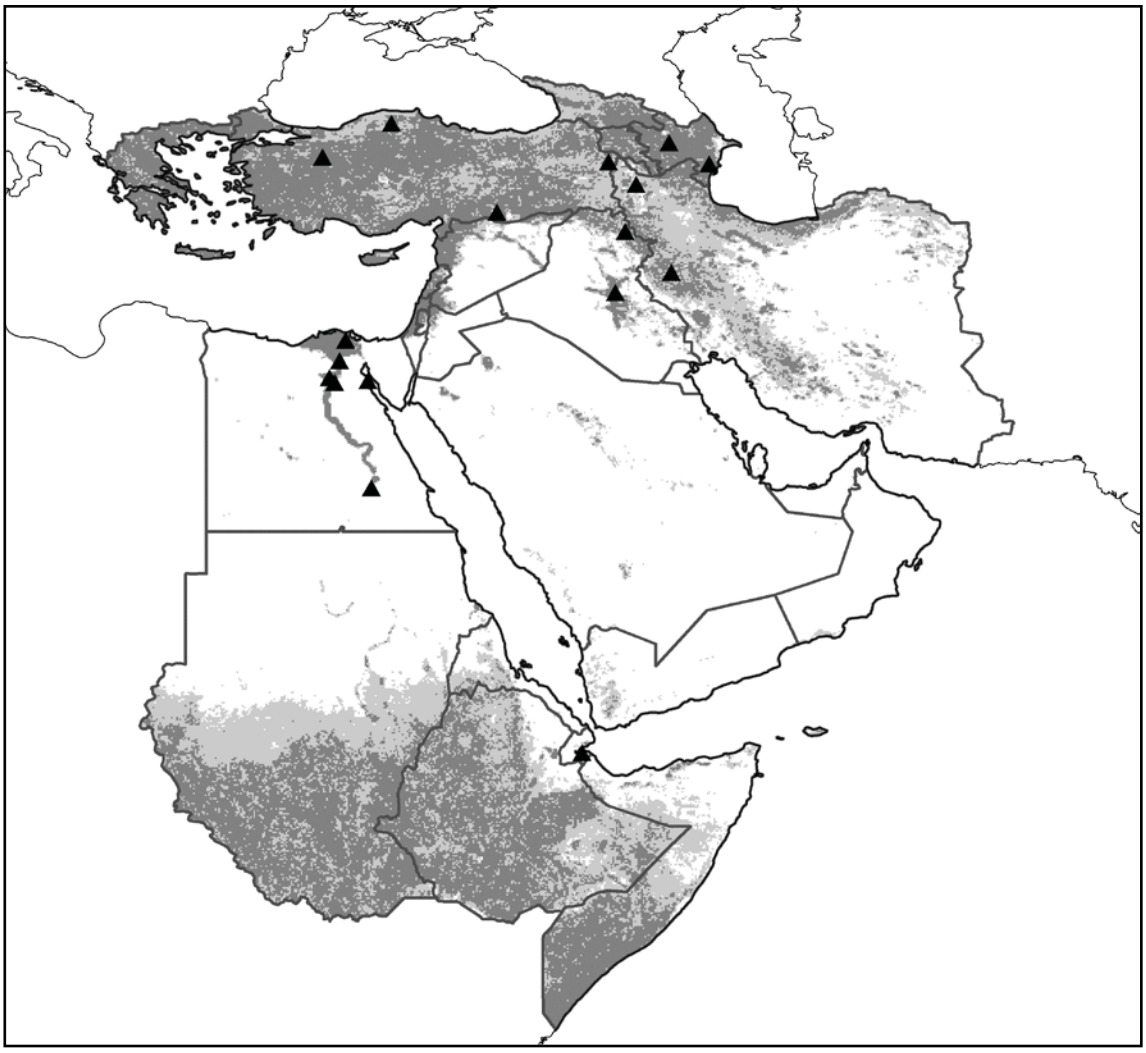

Figure 2.2 Regional projection across the Middle East and northeastern Africa of H5N1 ecological niche model results based on all OIE case occurrence points. Model predictions are shown as ramps of model agreement in predictions: light grey = 5–9 models predict potential presence, dark grey = all models agree in predicting potential presence. Black triangles indicate independent test data (N = 17) from the region drawn from the ProMed archive of human case reports.

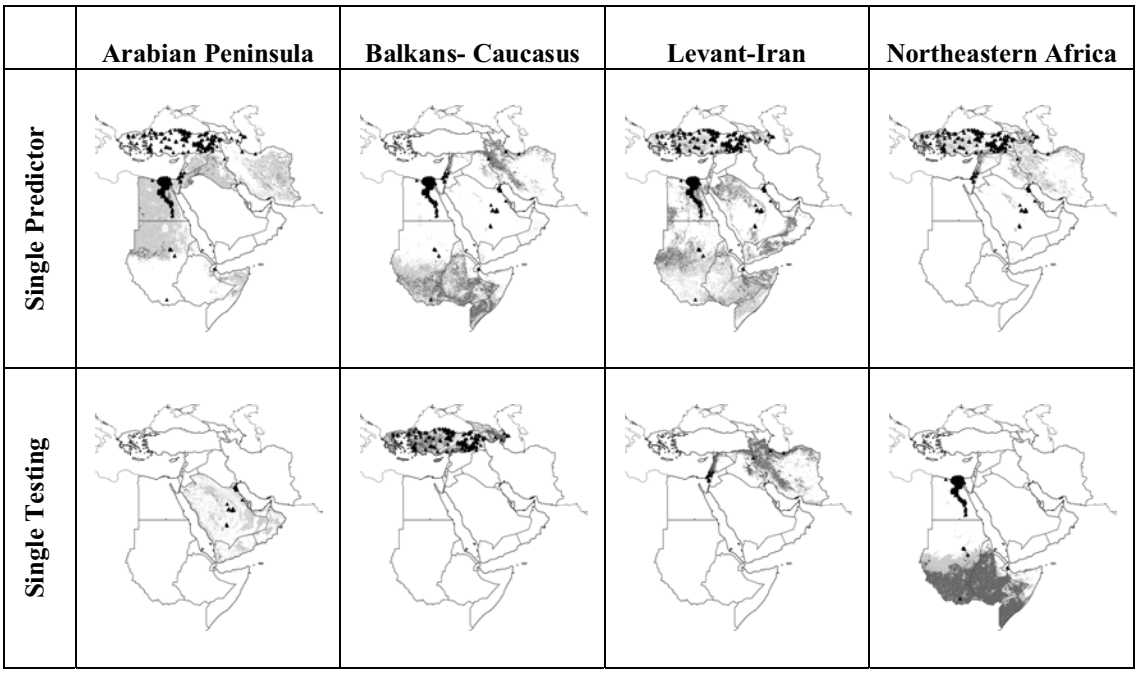

Figure 2.3 Spatially stratified tests of ENM predictions of H5N1 distributions in the Middle East and northeastern Africa. Here, occurrences from each subregion predict distributions of cases in the rest of the region, and vice versa. Model predictions are shown as ramps of model agreement in predictions: light grey = 5–9 models predict potential presence, dark grey = all models agree in predicting potential presence. Only independent test points are plotted on maps. The dense cluster of testing points along the lower Nile River in north-eastern Africa as testing region analyses covers an area predicted to be suitable.

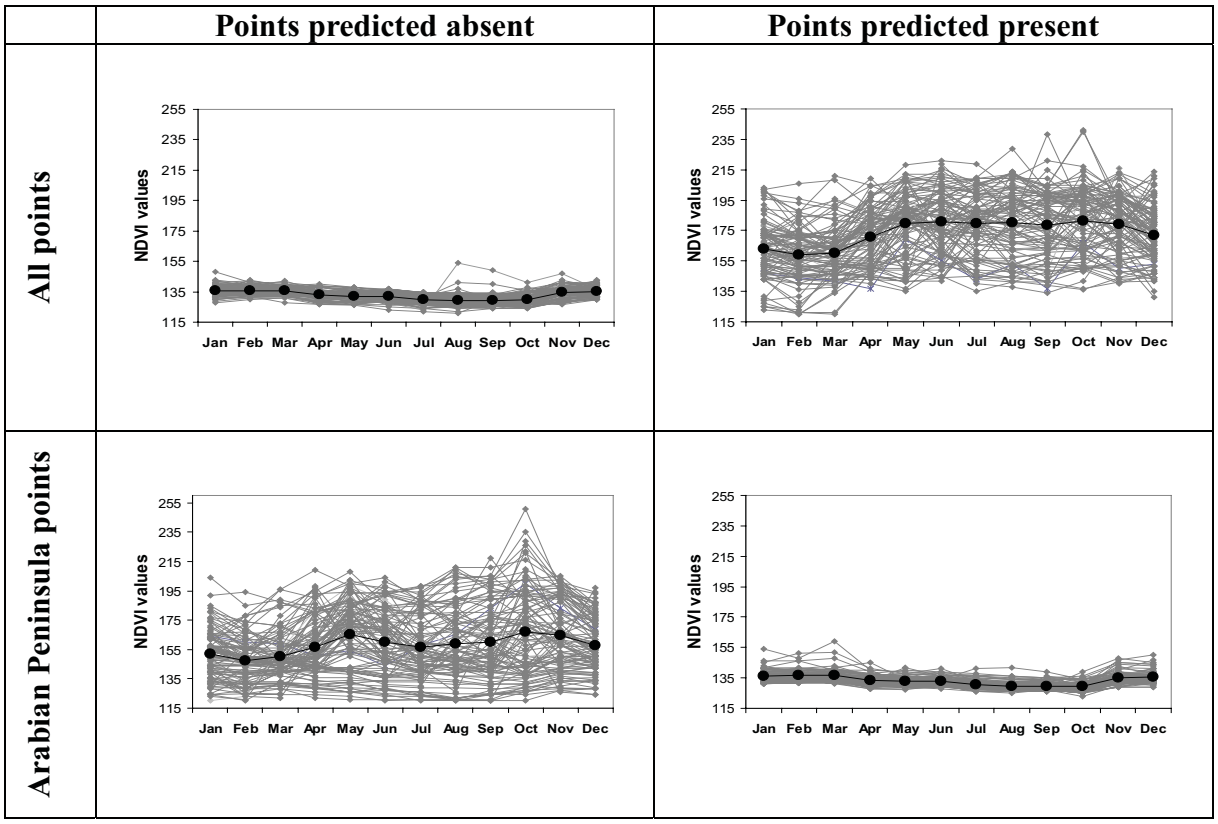

Figure 2.4 Summary of Normalized Difference Vegetation Index (NDVI) 'greenness' profiles of the Middle East and northeastern Africa through one year for models based on the entire region (top) and for models based only on the Arabian Peninsula. In each case, we show NDVI values for 100 randomly selected points of predicted absence versus 100 randomly selected points of predicted presence. Median values are shown in bold.

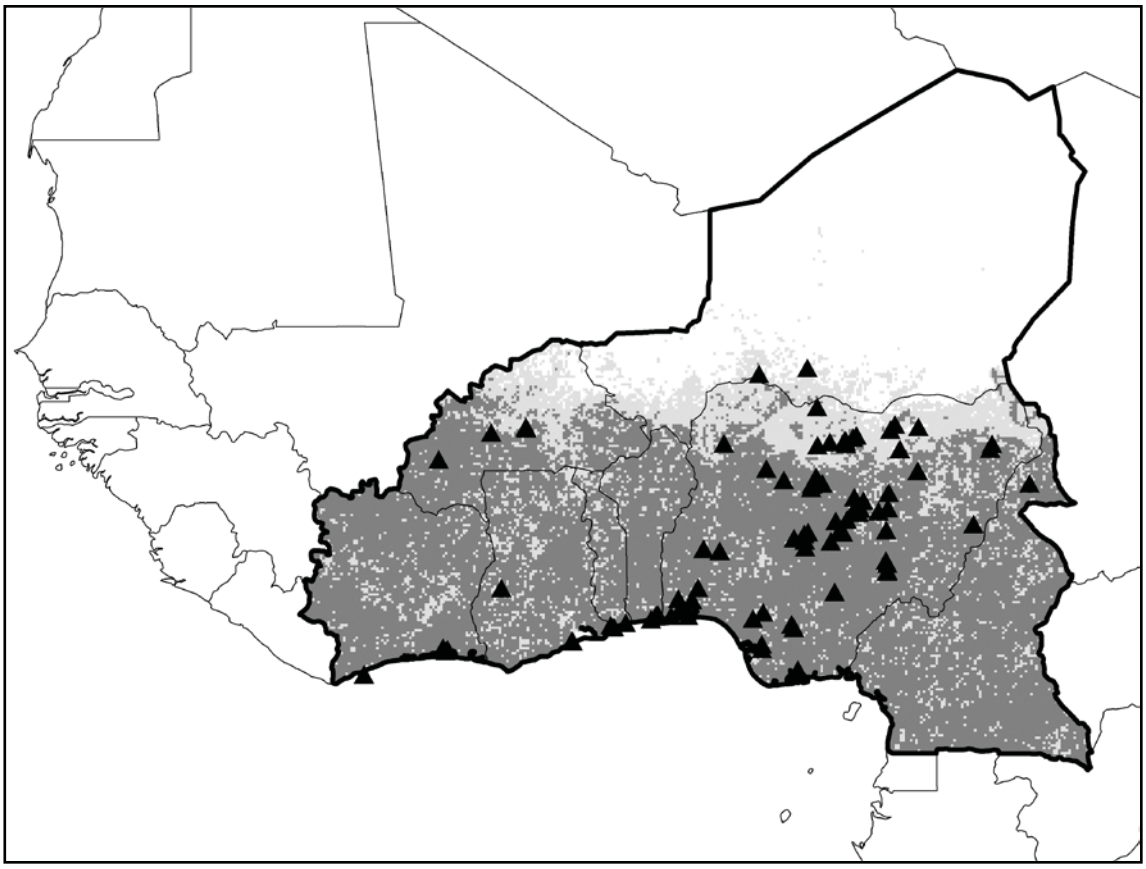

Figure 2.5 Regional projection across West Africa of H5N1 ecological niche model. Results based on OIE case occurrence points and environmental layers for the Middle East and northeastern Africa. Model predictions are shown as ramps of model agreement in predictions: light grey = 5–9 models predict potential presence, dark grey = all models agree in predicting potential presence. Black triangles indicate independent test data (*N* = 101) from the region [10,11]. Study area is delineated by bold border.

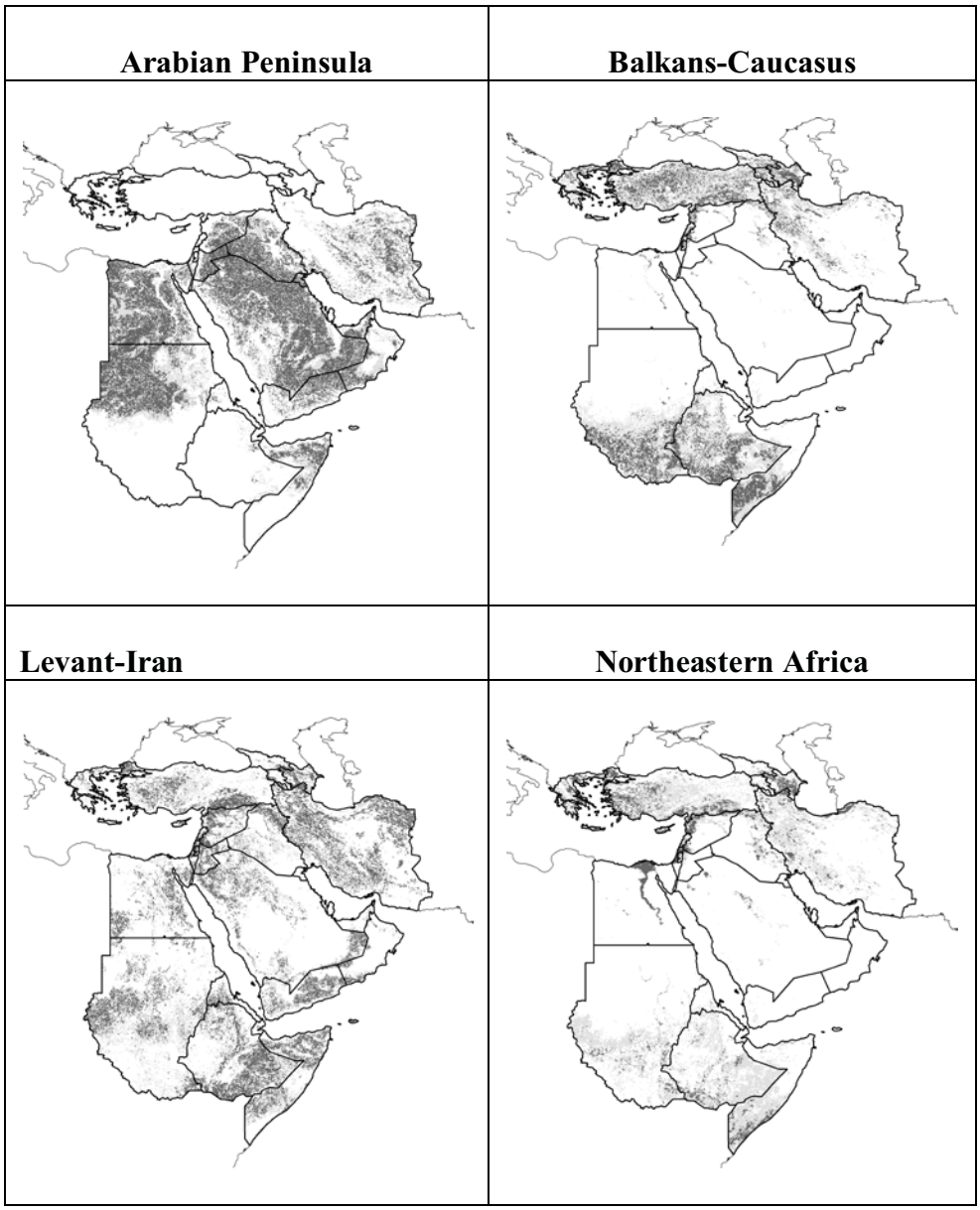

Figure 2.6 Projections of H5N1 occurrences from a single subregion across the whole region. Light grey = 5–9 models predict potential presence, dark grey = all models agree in predicting presence. Note the contrast between the Arabian Peninsula and the other three predictions.

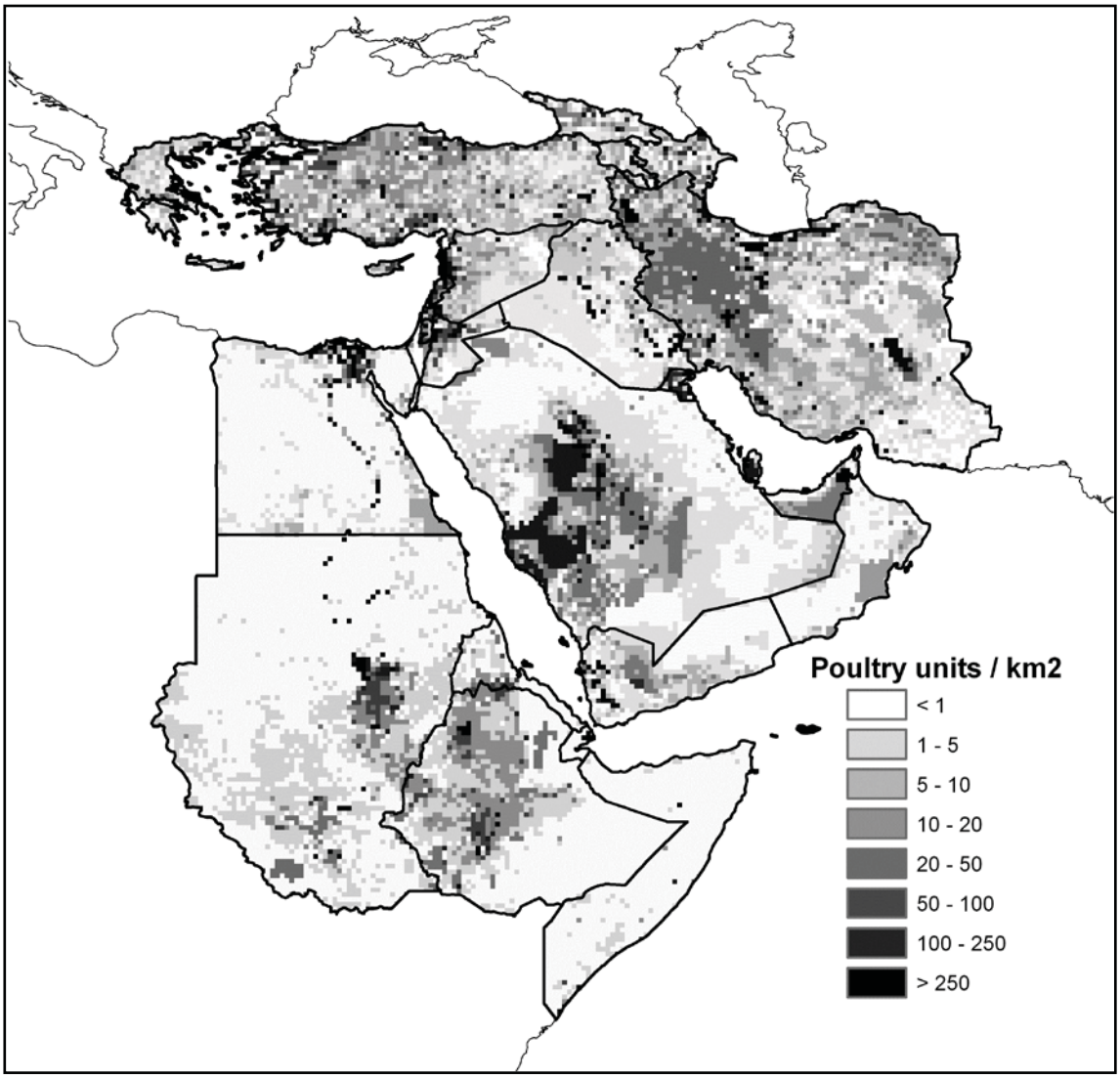

Figure 2.7 Density of poultry in the Middle East and north-eastern Africa (units per km^2) [13].

Table 2.1 Summary of H5N1 detections from countries across the Middle East and northeastern Africa reported by OIE[a] [(4)], ProMed[b] [(34)], and WHO[c] [(5)]. Note that numbers of wild bird cases seem to be unreliable: on one hand, these numbers are over reported in Egypt, where cases in birds captive in Giza Zoo are counted as "wild", and probably underreported from Azerbaijan, where "die-offs" yielded only 3 positive detections. Clearly, however, poultry cases far outnumber wild cases, and numbers of birds culled to prevent disease spread are higher still.

Country	Date of first detection [a]	Species of first detection [b]	No. of human cases [c]	No. of wild cases [a]	No. of poultry or captive bird cases [a]	No. of poultry culled [a]
Turkey	6 Oct 2005	*Meleagris gallopavo*	12	16	9178	289 812
Kuwait	11 Nov 2005	*Phoenicopterus ruber*	0	0	131	466 996
Iraq	9 Jan 2006	*Homo sapiens*	3	0	652	3478
Azerbaijan	29 Jan 2006	*Cygnus* sp.	8	3	1	296 000
Iran	2 Feb 2006	*Cygnus* sp.	0	153	14	475
Greece	9 Feb 2006	*Cygnus* sp.	0	17	0	0
Egypt	17 Feb 2006	Poultry	50	19	1 075 920	8 840 215
Sudan	20 Feb 2006	Poultry	0	0	87 370	107 327
Georgia	23 Feb 2006	*Cygnus* sp.	0	10	0	0
Israel	16 Mar 2006	*Gallus gallus*, *Meleagris gallopavo*	0	0	15 212	256 414
Palestinian Terr.	23 Mar 2006	Poultry	0	0	5000	40 800
Jordan	23 Mar 2006	*Gallus gallus*, *Meleagris gallopavo*	0	0	21	18 000
Djibouti	5 Apr 2006	Local poultry	1	0	4	18
Saudi Arabia	12 Mar 2007	*Struthio camelus*	0	0	12 606	5 310 290
TOTALS			74	218	1 206 109	15 629 825

Table 2.2 Summary of model predictions, binomial tests and partial ROC tests in this study, illustrated by information for the threshold >5 of 10 best subsets models predicting potential for presence. "Prop. area" indicates the proportion of the test region predicted present at that threshold. Also provided is the number of thresholds (out of 10) for which model predictions were significantly better than random expectations. Values under Max, Min, and Mean characterize distributions of AUC ratios (maximum, minimum, and mean) across 1000 bootstrap replicates, and the number of bootstrap replicates falling at or below unity. Subregion names are denoted by AP (Arabian Peninsula), BC (Balkans-Caucasus), LI (Levant-Iran), NA (northeastern Africa).

Region	Sample size (train/test)	Prop. area	# of successes	Cumulative binomial probability	# of thresholds significant	Max	Min	Mean	# of replicates ≤1
Single-testing									
AP / LI / NA → BC	435/176	0.836	158/175	0.004	7	1.611	1.140	1.457	0
AP / BC / NA → LI	592/18	0.412	13/18	0.002	10	1.708	1.071	1.217	0
AP / BC / LI →	224/386	0.524	277/386	0.000	8	1.081	1.012	1.035	0
BC / LI / NA → AP	579/31	0.355	12/31	0.282	4	1.271	1.021	1.065	0
Single predictor									
AP → BC / LI /	31/579	0.351	69/579	0.999	0	1.002	0.991	0.995	997
BC → AP / LI / NA	186/435	0.260	101/435	0.902	0	0.981	0.976	0.977	1000
LI → AP / BC / NA	18/592	0.428	404/592	0.000	5	1.656	0.951	1.272	2
NA → AP / BC / LI	386/224	0.230	124/224	0.000	10	1.120	0.996	1.043	1
Across study									
OIE veterinary cases → ProMed human cases	610/17	0.412	14/17	0.000	10	1.657	0.992	1.209	9

Table 2.3 Comparison of populations of domestic Anatidae and area under rice cultivation in 5 H5N1 affected countries. Data drawn from [a]Food and Agriculture Organization – Global Livestock Production and Health Atlas [(13)] and [b]International Rice Research Institute [(16)].

	Total poultry, 2004 [a]	**Domestic Anatidae, 2004[a]**	**Area of rice cultivation, 2006 ha[b]**	**Anatidae / rice ha**
China	4 735 229 952	875 230 000	29 380 000	29.79
Egypt	112 150 000	18 300 000	613 000	29.85
Iran	284 600 000	2 600 000	620 000	4.19
Thailand	187 270 000	17 270 000	10 073 000	1.71
Vietnam	252 000 000	75 000 000	7 324 000	10.24

CHAPTER 3: CONTINENT-WIDE ASSOCIATION OF H5N1 OUTBREAKS IN WILD AND DOMESTIC BIRDS IN EUROPE

ABSTRACT

Highly pathogenic avian influenza genotype H5N1 (hereafter "H5N1") was first detected in Europe in 2005, and has since been documented continentwide in wild birds and poultry, but the relative roles of each host group in transmission remain contentious. Using recently developed tools for analysis of ecological niche requirements and geographic distributions of species, we compare ecological niche requirements between paired host groups (poultry versus wild birds, Anseriformes versus Falconiformes, swans versus non-swan Anseriformes). If environmental signals of different host groups are significantly different, the groups are likely involved in distinct transmission cycles; in contrast, models for which similarity cannot be rejected imply no unique ecological niches, and potential linkage of transmission cycles. In 24 similarity tests, we found significant similarity (11/24) or no significant differences (1/24). Although 2 of 24 analyses found significant differences, neither was unequivocal, so we conclude an overall signal of niche similarity among groups. We thus could not document distinct ecological niches for H5N1 occurrences in different host groups, and conclude that transmission cycles are broadly interwoven.

INTRODUCTION

The highly-pathogenic avian influenza strain H5N1 spread from Southeast Asia and East Asia into Europe and Africa in 2005 [1]. H5N1 has now been detected in 62 countries, including 26 in Europe [2]. European H5N1 outbreaks appeared to differ from those in other regions in that numbers of distinct outbreaks in wild birds and poultry were roughly equal: 379 reports for domestic birds against 294 in wild birds for 2006. This apparent leap in numbers of wild birds infected may be explained in part by more intensive surveillance than elsewhere, as nearly 121,000 wild birds were sampled in the EU in 2006 [3]. Alternatively, these contrasts may point to viral evolution after 2005 [4], and consequent increased pathogenicity to wild bird species—for this reason, ecological or environmental differences among H5N1 strains transmitted in different host groups are of considerable interest.

Questions regarding the relative importance and linkage of transmission among wild birds and poultry remain controversial [5-6]; although, both likely play roles [7]. Avian influenza strains are known to be maintained in wild aquatic bird reservoirs [8]: *Anas platyrhynchos* has been identified as a possible H5N1 reservoir [9]. Other factors, implicated in the transmission cycle include illegal wild animal-smuggling [10], infected poultry feed [11], and undocumented poultry trade.

Previous studies have attempted to understand regional-scale H5N1 ecology and geography, implicating various factors as correlates of H5N1 occurrence: e.g. duck

abundance and rice cropping intensity in Southeast Asia [1], and high seasonal variation in NDVI (Normalized Difference Vegetation Index) values in studies in Africa and the Middle East [12-13]. Factors important to H5N1 transmission probably vary among regions: for instance, Germany has suffered annual outbreaks since 2006, despite negligible rice cultivation and few free-grazing domestic ducks.

Ecological niche models (ENMs) and associated techniques offer unique opportunities to study ecological associations of biological phenomena across broad regions. ENMs reconstruct coarse-resolution environmental and ecological requirements that determine geographic distributions [14], and have been used to explore diverse topics in distributional ecology, including disease distributions and transmission risks [13]. Here, we use ENM-based niche-comparison approaches [15] to develop detailed comparisons of ecological niches among different potential host groups for H5N1 to produce a first quantitative test of environmental connectivity of H5N1 transmission between wild and domestic birds.

MATERIAL AND METHODS

Input Data

Occurrence data were drawn from OIE [2], which compiled 494 poultry (including chickens, ducks, geese, quail, turkeys kept at all biosecurity levels) and 605 wild bird (of at least 21 species) laboratory-confirmed H5N1 detections across our study region (0°-58°E, 34°-60°N) during 2005-2008. We discarded points duplicated spatially at 0.01° resolution. To

minimize problems caused by the non-independence of cases, we also rarefied points within spatial clusters to densities similar to those of surrounding areas, leaving a total of 89 poultry and 90 wild bird occurrence points for analysis (Table 3.1). Separately, we explored subsets of the wild bird data, to permit comparisons of wild Anseriformes (N = 102) versus wild Falconiformes (N = 21), and wild swans (*Cygnus* spp., N = 63) versus wild non-swan Anseriformes (N = 39). Because these latter datasets did not appear to manifest clumped distributions, we included them in their entirety.

Niche models were developed using two distinct and comparable suites of environmental data layers. The first was based on multi-temporal remotely-sensed vegetation indices derived from the MODIS sensor onboard the Terra satellite: Normalized Difference Vegetation Index (NDVI) [16], Enhanced Vegetation Index (EVI) [17], and Land Surface Water Index (LSWI) [18]. NDVI, used to estimate vegetation growth and biomass production, is calculated as the following ratio of near infrared (NIR) and red spectral bands:

$$NDVI = \frac{\rho_{nir} - \rho_{red}}{\rho_{nir} + \rho_{red}}$$

(Equation 3.1)

EVI provides a similar vegetation measure, but corrects additionally for error due to aerosol reflectance and canopy background signals:

$$EVI = 2.5 \times \frac{\rho_{nir} - \rho_{red}}{\rho_{nir} + (6 \times \rho_{red} - 7.5 \times \rho_{Blue}) + 1}$$

(Equation 3.2)

Finally, LSWI is calculated as the following ratio of near infrared and shortwave infrared (SWIR) bands:

$$LSWI = \frac{\rho_{nir} - \rho_{swir}}{\rho_{nir} + \rho_{swir}}$$

(Equation 3.3)

In each case, we used maximum, mean, minimum, and range for the year 2005, supplemented with information on topographic features, including aspect, compound topographic index, and slope from the Hydro-1K digital elevation model data set [19], and elevation from GTOPO30 (all resampled to a spatial resolution of 0.01°) [20]. The second data set consisted of 7 "bioclimatic" variables from the 10' WorldClim data set [21]: annual mean temperature, mean diurnal range, maximum temperature of warmest month, minimum temperature of coldest month, annual precipitation, precipitation of wettest month and precipitation of driest month (all resampled to a spatial resolution of 0.0083°).

Ecological Niche Models

Ecological niche models (ENMs) have been developed via diverse methodological approaches, which have different strengths and weaknesses. For these analyses, we chose the algorithm Maxent, an approach that is known to perform particularly well in interpolation challenges [22]. Maxent is a method for characterizing probability distributions

from incomplete information that has been applied as a method for estimating ecological niches and inferring species' distributions from presence data [23]. Maxent has seen considerable success in model-comparison studies [22-23]. Although Maxent encounters challenges in estimating niches and geographic distributions across broad, unsampled regions [24], it is excellent for relatively densely-sampled landscapes and interpolation challenges, as is the challenge in this study. A further benefit of using the Maxent algorithm is that it generates area under the curve (AUC) values as a measure of the predictive power of the model. Finally, Maxent is integral to the software available for background similarity testing that is central to this study [15]. Maxent fits a probability distribution for occurrence of the biological phenomenon in question to the set of pixels across the study region, assuming that the best model will maximize the entropy of the probability distribution, subject to the constraint that pixel values where the biological phenomenon has been detected should reflect presence at higher probability values. We used default parameters for compatability with tools for testing niche similarity (see below), choosing logistic output format. Outputs were imported into ArcView 3.3 as floating-point grids, and then thresholded to the lowest predicted value associated with any known detection locality [25].

Quantifying Niche Similarity

We follow Warren et al. [15] and Soberón [26] in considering the overlap between maps of habitat suitability in environmental space as a measure of ecological niche similarity, using the *D* and *I* similarity metrics to quantify the similarity between two probability distributions. These statistics assume probability distributions defined over geographic

space, in which $p_{X,i}$ (or $p_{Y,i}$) denotes the probability assigned by the ENM for species X (or Y) to cell *i*. The *D* metric, Schoener's statistic for niche overlap [27], is calculated as:

$$D(p_X, p_{XY}) = 1 - \frac{1}{2} \sum |p_{X,i} - p_{Y,1}|.$$

(Equation 3.4)

The *I* metric was modified from the Hellinger distance [28] to be comparable to more conventional measures of niche overlap, and is calculated as:

$$I\ (p_X,\ p_Y) = 1 - \frac{1}{2}\ H(p_X,\ p_Y).$$

(Equation 3.5)

Values for each metric range from 0 (no overlap) to 1 (identical). Following Warren et al. [15], we used randomization tests that ask whether two species' modeled niches are more similar or more different than random expectations. Importantly, this allows specification of an area of analysis ("the background"), which we equate with M, the area accessible to a species over relevant time periods [14].

Hence, using ENMTools [15], we compared *D* and *I* values for paired ENMs for each of the paired groups of H5N1 hosts (poultry versus wild birds, Anseriformes versus Falconiformes, non-swans versus swans). We assumed a 300 km buffer around known case occurrences as a hypothesis of M. Analyses are conducted bi-directionally, so we compared the niche model for one host group to the background of the other, and vice versa. Numbers of points sampled from the background were set at observed sample sizes. In each test, 100

replicate randomizations were conducted to estimate probabilities for each test to estimate probabilities to the nearest 1%.

Given the unknown nature of the associations (positive or negative) between influenza occurrences between these groups, we used a two-tailed null hypothesis: that niche overlap observed among host groups is not real difference, but rather is explicable by differences in the background landscapes for each group. The hypothesis is rejected if observed similarity between models falls outside the 95% confidence limits of the null distribution, with greater than expected niche difference defined as <2.5%, and greater than expected similarity as >97.5%, of the distribution.

RESULTS

Niche models for H5N1 in poultry versus wild birds and swans versus other Anseriformes produced similar maps; covering much of our study area, save for high-elevation regions. Models of occurrences in Anseriformes versus Falconiformes were less similar, with the latter omitting much of southern Europe. Generally, models based on climatic data were more restricted in area predicted suitable than models based on remotely-sensed data (Fig. 3.1). The predictive power of all models, as measured by AUC values, was high (0.840 – 0.970: see Table 3.1), although this approach has limitations [29]. Jackknife evaluations of variable contributions in each model are available on request.

Background similarity tests assess whether two sets of occurrences were drawn from the same environmental niche, taking into account the availability of conditions across the region inhabited. No background similarity tests based on remotely-sensed layers indicated ecological niche difference: 7 of 12 pairs under comparison were more similar than random expectations, and in 5of 12 pairs the null hypothesis could not be rejected (Table 3.2). That is to say, in remotely sensed environmental dimensions, comparisons of ecological niches of poultry and wild birds, Anseriformes and Falconiformes, and *Cygnus* and non-*Cygnus* Anseriformes were unable to reject the null hypothesis of niche similarity. The similarity of niches between paired comparisons is shown by subtracting the remotely sensed ENM developed using H5N1 points detected in poultry from the remotely sensed wild bird model (Fig. 3.2A). Both background similarity tests assessing the H5N1 poultry versus wild bird niches were more similar than random expectations (Fig. 3.2B, 3.2C).

Background tests of niche similarity based on climate layers yielded similar, though more mixed, results (Table 3.2). The null hypothesis of similarity could not be rejected in comparisons of Falconiformes and Anseriformes, and swans and non-swan Anseriformes. However, although environments associated with wild bird and poultry H5N1 occurrences were indistinguishable from random similarity in two comparisons, they differed significantly in climatic dimensions in the remaining two comparisons.

To summarize, in all but one randomization test (that based on poultry data and climate layers), all pairs of model comparisons were found to be significantly similar (11/24), or

not significant from random expectations (11/24). Although some climate-based comparisons of case occurrences in poultry and wild birds detected significant differences, they were contradicted by comparisons based on remotely-sensed data. Hence, overall, the signal among our 24 comparisons was one of similarity among environments in which H5N1 was detected among host groups.

DISCUSSION

Throughout our analyses, no clear signal emerged to suggest distinct ecological niches for H5N1 case occurrences in different host groups. Among wild bird groups, the picture was completely consistent: H5N1 cases in Falconiformes and Anseriformes, and cases in swans and non-swan Anseriformes, occurred under a single set of environmental conditions. This result points to functional linkage of transmission cycles, and suggests that wild bird H5N1 case occurrences represent a single, coherent biological phenomenon. The picture was less clear for H5N1 in poultry and wild birds. Here, remotely-sensed assessments detected greater-than-expected niche similarity, but climate-based tests were more complex: comparisons of poultry H5N1 cases against the background of wild bird cases were significantly non-similar, but the converse test was indistinguishable from random. Hence, most evidence (6/8 tests) point towards environmental similarity of wild bird and poultry H5N1 cases.

In sum, we find no consistent signal of distinct ecological niches for H5N1 occurrences among the host bird groups tested herein. Kilpatrick et al [7] established likely pathways for

H5N1 introductions into 52 countries by mapping phylogenetic data for H5N1 isolates to wild bird movements and trade in poultry and wild birds. They determined that 26 introductions (including 20/23 European introductions) were probable wild bird introductions, 11 were probable poultry introductions, and the remainder could not be assigned. Genomic analysis of H5N1 surface proteins (hemagglutinin and neuraminidase) found no association between genotype and host [30], suggesting that the strain is not transmitted selectively to specific host groups. Our findings, though they cannot establish the transmission event introducing an isolate into a specific country, support this idea further: host-specific transmission pathways do not exist, and that H5N1 circulates freely with respect to host group. That is, within Europe, we found no consistent evidence indicating distinct transmission cycles in different avian hosts.

ACKNOWLEDGEMENTS

Xiangming Xiao, and A. Townsend Peterson are co-authors on the manuscript associated with this chapter. Thanks to Chandrashekhar Biradar for help with remotely sensed data, and to Monica Papeş, Yoshinori Nakazawa, and Sean Maher for expert advice on analyses. Thanks to the Wildlife Conservation Society, the National Biological Information Infrastructure of the U.S. Geological Survey, and the Centers for Disease Control and Prevention/National Institutes of Health for funding this work.

CHAPTER 3 REFERENCES

1. Gilbert M, Xiao X, Pfeiffer DU, Epprecht M, Boles S, Czarnecki C, et al. Mapping H5N1 highly pathogenic avian influenza risk in Southeast Asia. Proc Natl Acad Sci USA. 2008;105:4769-74.

2. OIE. Update on highly pathogenic avian influenza in animals (Type H5 and H7). 2009 [cited 2009 22 September 2009]; Available from: http://www.oie.int/downld/avian%20influenza/A_AI-Asia.htm

3. Guan Y, Poon LLM, Cheung CY, Ellis TM, Lim W, Lipatov AS, et al. H5N1 influenza: a protean pandemic threat. Proc Nat Acad Sci USA. 2004;101:8156-61.

4. Whitworth D, Newman S, Mundkur T, Harris P, editors. Wild birds and avian influenza: an introduction to applied field research and disease sampling techniques. Rome: FAO; 2007.

5. Weber TP, Stilianakis NI. Ecologic immunology of avian influenza (H5N1) in migratory birds. Emerg Infect Dis. 2007;13:1139-43.

6. Keawcharoen J, van Riel D, van Amerongen G, Bestebroer T, Beyer WE, van Lavieren R, et al. Wild ducks as long-distance vectors of highly pathogenic avian influenza virus (H5NI). Emerg Infect Dis. 2008;14:600-7.

7. Kilpatrick AM, Chmura AA, Gibbons DW, Fleischer RC, Marra PP, Daszak P. Predicting the global spread of H5N1 avian influenza. Proc Natl Acad Sci USA. 2006;103:19368-73.

8. Swayne DE, Suarez DL. Highly pathogenic avian influenza. Rev Sci Tech Off Int Epiz. 2000;19:463-82.

9. Brown JD, Stallknecht DE, Beck JR, Suarez DL, Swayne DE. Susceptibility of North American ducks and gulls to H5N1 highly pathogenic avian influenza viruses. Emerg Infect Dis. 2006;12:1663-70.

10. Van Borm S, Thomas I, Hanquet G, Lambrecht N, Boschmans M, Dupont G, et al. Highly pathogenic H5N1 influenza virus in smuggled Thai eagles, Belgium. Emerg Infect Dis. 2005;11:702-5.

11. Harder TC, Teuffert J, Starick E, Gethmann J, Grund C, Fereidouni S, et al. Highly pathogenic avian influenza virus (H5N1) in frozen duck carcasses, Germany, 2007. Emerg Infect Dis. 2009;15:272-9.

12. Williams RAJ, Peterson AT. Ecology and geography of avian influenza (HPAI H5N1) transmission in the Middle East and northeastern Africa. Int J Health Geogr. 2009;8:e47.

13. Williams RAJ, Fasina FO, Peterson AT. Predictable ecology and geography of avian influenza (H5N1) transmission in Nigeria and West Africa. Trans R Soc Trop Med Hyg. 2008;102:471-9.

14. Soberón J, Peterson AT. Interpretation of models of fundamental ecological niches and species' distributional areas. Biodivers Inform. 2005;2:1-10.

15. Warren DL, Glor RE, Turelli M. Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. Evolution. 2008;62:2868-83.

16. Tucker CJ. Red and photographic infrared linear combinations for monitoring vegetation. Remote Sens Environ. 1979;8:127-50.

17. Huete A, Didan K, Miura T, Rodriguez EP, Gao X, Ferreira LG. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. Remote Sens Environ. 2002;83:195-213.

18. Xiao XM, Boles S, Liu JY, Zhuang DF, Liu ML. Characterization of forest types in northeastern China, using multi-temporal SPOT-4 VEGETATION sensor data. Remote Sens Environ. 2002;82:335-48.

19. USGS. HYDRO1k Elevation Derivative Database. 2001 [cited 2009 10th July]; Available from: http://edc.usgs.gov/products/elevation/gtopo30/hydro/index.html

20. USGS. GTOPO30. Global 30 arc second elevation data. 2006 [cited 2009 27th July]; Available from: http://edc.usgs.gov/products/elevation/gtopo30/gtopo30.html

21. Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology. 2005;25:1965-78.

22. Elith J, Graham CH, Anderson RP, Dudik M, Ferrier S, Guisan A, et al. Novel methods improve prediction of species' distributions from occurrence data. Ecography. 2006;29:129-51.

23. Phillips SJ, Anderson RP, Schapire RE. Maximum entropy modeling of species geographic distributions. Ecol Model. 2006;190:231-59.

24. Peterson AT, Papes M, Eaton M. Transferability and model evaluation in ecological niche modeling: a comparison of GARP and Maxent. Ecography. 2007;30:550-60.

25. Pearson RG, Raxworthy CJ, Nakamura M, Peterson AT. Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. J Biogeogr. 2007;34:102-17.

26. Soberón J. Grinnellian and Eltonian niches and geographic distributions of species. Ecol Lett. 2007;10:1115-23.

27. Schoener TW. Anolis lizards of Bimini - resource partitioning in a complex fauna. Ecology. 1968;49:704-26.

28. Van der Vaart AW. Asymptotic statistics. Cambridge, UK: Cambridge University Press; 1998.

29. Lobo JM, Jimenez-Valverde A, Real R. AUC: a misleading measure of the performance of predictive distribution models. Glob Ecol Biogeogr. 2008;17:145-51.

30. Janies D, Hill AW, Guralnick R, Habib F, Waltari E, Wheeler WC. Genomic analysis and geographic visualization of the spread of avian influenza (H5N1). Syst Biol. 2007;56:321-9.

CHAPTER 3 FIGURES AND TABLES

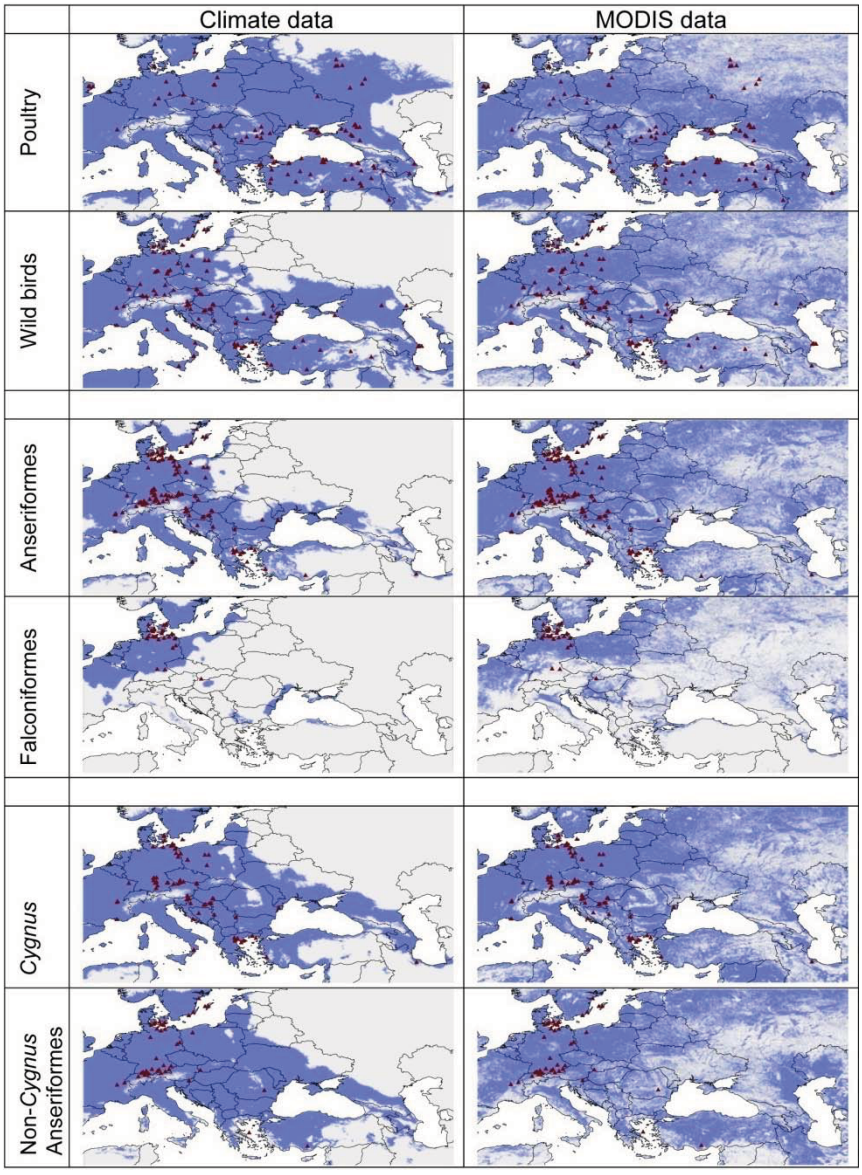

Fig. 3.1. Ecological niche models for H5N1 detections in distinct avian groups across Europe. Dark grey indicates potential H5N1 presence (based on a least training presence threshold); light grey indicates absence. Triangles show cases used for model training. Each dataset was modeled using climatic and remotely-sensed environmental datasets.

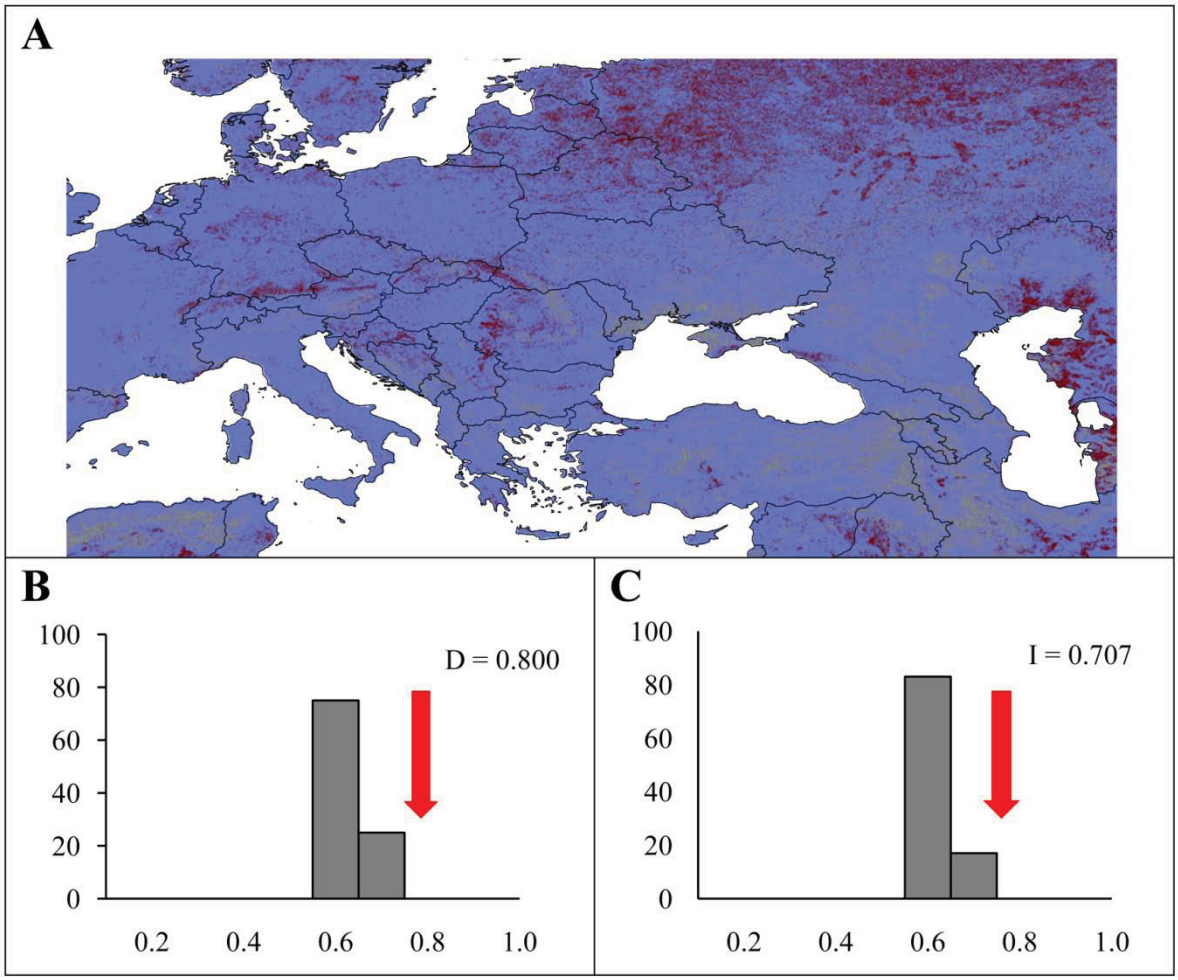

Fig. 3.2. (A) Difference between suitability predicted for wild birds and poultry, based on remotely-sensed environmental data. Grey indicates agreed prediction between models, white indicates areas predicted by poultry but not wild bird, and black indicates the reverse. (B) Histograms illustrating measured overlap (arrow) and distribution of background similarity among random replicate models for *D*, and (C) *I* metrics, for the wild bird remotely-sensed model.

Table 3.1. Numbers of H5N1 cases used for generating, and AUC values generated by, each ENM.

		AUC	
Test group	**Sample size**	**Climate**	**Remotely- sensed**
Poultry	89	0.841	0.898
Wild birds	90	0.907	0.926
Anseriformes	102	0.943	0.938
Falconiformes	21	0.977	0.960
Cygnus	63	0.907	0.875
non-*Cygnus* Anseriformes	39	0.936	0.908

Table 3.2. Background similarity tests using Schoener's D (*D*) and the Hellinger distance (*I*) for climatic and remotely-sensed data. The null hypothesis of niche similarity is rejected if test statistic values are below 0.025 (indicated with asterisks); reciprocal tests are presented (A-B, B-A).

		Climate			Remote sensing		
		Overlap	Background		**Overlap**	Background	
Test comparison		**value**	**A-B**	**B-A**	**Value**	**A-B**	**B-A**
[A]Poultry *versus*	D	0.676	0.00*	0.90	0.800	1.00	1.00
[B]Wild birds	I	0.649	0.00*	0.31	0.707	1.00	0.99
[A]Anseriformes *versus*	D	0.476	0.21	1.00	0.357	0.19	0.94
[B]Falconiformes	I	0.661	0.94	1.00	0.538	0.19	1.00
[A]*Cygnus versus*	D	0.785	0.95	0.78	0.754	0.73	1.00
[B]non-*Cygnus* Anseriformes	I	0.842	1.00	1.00	0.681	0.74	1.00

CHAPTER 4: YAOUNDÉ-LIKE VIRUS IN RESIDENT WILD BIRD, GHANA

TO THE EDITOR:

Yaoundé virus (YAOV) is a mosquito-borne *Flavivirus* within the *Japanese encephalitis virus* (JEV) serocomplex which includes human pathogens such as JEV and *West Nile virus* (WNV) [(1)]. YAOV has been isolated from one bird (*Bycanistes sharpii*), 2 mammal species (*Praomys* sp. and *Cavia porcellus*), and 10 mosquito species from four sub-Saharan African nations (Cameroon, Central African Republic, Congo, and Senegal) [(2)]. The virus was first isolated from a *Culex nebulosus* pool collected from degraded semi-deciduous forest near Nkolbisson, Cameroon, in 1968 [(3)]. Although YAOV has not been detected in humans [(2)], many other JEV group taxa have been.

We sampled and tested 551 free-ranging birds and 12 backyard chickens for *Alphavirus*, *Flavivirus* and influenza A virus infections using PCR techniques. Sampling was conducted in Ghana, at Ankasa Conservation Area (N =305, 70 species), Western Region, and Gbele Resource Reserve (N =259, 73 species), Upper West Region, in October-November 2007. Birds were obtained by mist netting and selective harvesting with shotguns; all birds were apparently healthy at the time of collection. Buccal and cloacal swabs were collected for up to 30 individuals per species, and mixed tissue samples (brain, gonad, intestine, kidney and lung) for up to 10 individuals per species. Samples were frozen immediately in liquid nitrogen, and voucher specimens were prepared for all samples and deposited at the University of Kansas Biodiversity Institute. 95.5% of individuals sampled were resident,

2.9% were Palearctic migrants, and the remainder (1.6%) were species whose populations include internal African migrants.

Viral RNA was extracted from homogenized tissue pools using RNEasy minikit (QIAGEN, Valencia, CA, USA), and from swab pools prepared in VTM (containing 2.5% veal infusion broth, 0.5% BSA, 100 μg/ml gentamicin sulfate and 2 μg/ml Fungizone) using QIAamp Viral RNA Mini Kit (QIAGEN). Tissues were homogenized using TissueRuptor (QIAGEN). Molecular RNA detection techniques were used to detect genomic material from alphaviruses [4], flaviviruses [5,6], and influenza viruses [7,8]. A flavivirus was detected from one pool of tissues from seven birds using the nested RT-PCR technique [5], other tests resulted negative. The 100 nt amplification product was detected by electrophoresis, and purified using QIAquick PCR Purification Kit (QIAGEN). Sequencing reactions were performed with ABI Prism BigDye Terminator Cycle Sequencing v.3.1 Ready Reaction (Applied Biosystems, Foster City, CA, USA), and analyzed using an ABI PRISM model 3730 automated sequencer (Applied Biosystems). Assembly of consensus sequences and translation into amino acid sequences was performed with EditSeq (DNASTAR Inc., Madison, WI, USA). Comparisons with published sequences of known flaviviruses (excluding primer binding sites) were performed by searches with the FASTA program in the EMBL database (www.ebi.ac.uk/embl) to identify the detected agent and study the level of homology. Phylogenetic analyses were developed relative to 13 *Flavivirus* sequences, all aligned using ClustalX (http://www.clustal.org/). Phylogenetic trees were constructed by Bayesian analysis, using MrBayes v. 3.1.2 (http://mrbayes.csit.fsu.edu/), with 10,000,000 cycles for the Markov chain Monte Carlo algorithm, and specifying TBEV (GenBank Acc

No: NC_001672) as outgroup. Bayesian posterior probabilities were calculated from the consensus of 10,000 trees.

One tissue pool of seven resident wild birds (Table 4.1) captured among the two sites was positive for *Flavivirus* NS5 sequence. Attempts to determine viral host identity by testing samples from seven individuals from the *Flavivirus* positive pool individually, using the generic RT-nested PCR [4], were negative. Triturated tissue (and associated swab storage buffer) from individuals in the positive pool were inoculated into C6/36 (*Aedes albopictus* cells) and Vero (African green monkey cells) monolayer cells, no cytopathic effects were detected after three blind passages. Supernatant samples drawn from all passages of cell cultures were negative using the generic RT-nested PCR [5]. Further attempts to amplify a more informative sequence using nine generic RT-nested PCR protocols (derived from [1], and ISCIII designed primers) were unsuccessful.

The putative YAOV sequence (GenBank accession number HQ290163) showed 92% sequence identity with a published YAOV strain (EU074036), with 8 nt differences (Table 4.2), none producing amino acid replacements. The phylogenetic analyses (data not shown) confirmed the close relation of our sequence and YAOV (93% posterior probability). The virus we detected is thus YAOV, or one that is quite closely related.

To our knowledge no JEV group virus infection is known from Ghana; however, WNV presence has been inferred from a human serologic study that found a seroprevalence of

27.9% in adults [9]. Given the indirect nature of serological tests, we suggest that WNV and/or some WNV-like agent(s) would be a more objective inference. Although WNV is likely present in Ghana, we note that the epidemiology and serologic characteristics of YAOV remain unstudied in humans; thus, Ghanaian flaviviruses merit further investigation.

ACKNOWLEDGEMENTS

Ana Vázquez, Ivy Asante, Kofi Bonney, Shirley Odoom, Naiki Puplampu, William Ampofo, Marie-Paz Sánchez-Seco, Antonio Tenorio and Townsend Peterson are co-authors on the manuscript associated with this chapter. We thank field companions including E. Bonaccorso, M. Robbins, and M. Thompson; laboratory companions, including A. Negredo; and L. Benitez for introduction to ISCIII. We thank Noelia Reyes for technical assistance and Dr Francisco Pozo and Dr Inmaculada Casas for their collaboration in the Influenza and Respiratory Viruses Laboratory at the National Center of Microbiology (ISCIII). This study was supported by the US National Biological Information Infrastructure, the GAINS program of the Wildlife Conservation Society, and the Centers for Disease Control and Prevention, grants FIS PI07/1308 of the Red de Investigacion de Centros de Enfermedaes Tropicales RD06/0021, and the agreement signed between the Institute of Health Carlos III and the Spanish Ministry of Health and Social Policy for the surveillance of imported viral hemorrhagic fevers. Influenza work was supported by grant GR09/0040 (MPY-1440/09) ISCIII.

CHAPTER 4 REFERENCES

1. Kuno G, Chang GJ, Tsuchiya KR, Karabatsos N, Cropp CB. Phylogeny of the genus *Flavivirus*. J Virol. 1998;72:73-83.

2. Adam F, Digoutte JP. Virus d'Afrique (base de données). 2005 [cited 2010 July 1]; Available from: http://www.pasteur.fr/recherche/banques/CRORA

3. Digoutte JP. Rapport sur le fonctionnement technique de l'Institut Pasteur de Dakar. Année 1991. Dakar: Institut Pasteur de Dakar; 1992.

4. Sánchez-Seco MP, Rosario D, Quiroz E, Guzmán G, Tenorio A. A generic nested-RT-PCR followed by sequencing for detection and identification of members of the *Alphavirus* genus. J Virol Methods. 2001 Jun;95:153-61.

5. Sánchez-Seco MP, Rosario D, Domingo C, Hernández L, Valdés K, Guzmán MG, et al. Generic RT-nested-PCR for detection of flaviviruses using degenerated primers and internal control followed by sequencing for specific identification. J Virol Methods. 2005 Jun;126:101-9.

6. Pierre V, Drouet MT, Deubel V. Identification of mosquito-borne *Flavivirus* sequences using universal primers and reverse transcriptase-polymerase chain-reaction. Res Virol. 1994 Mar-Apr;145:93-104.

7. Spackman E, Senne DA, Myers TJ, Bulaga LL, Garber LP, Perdue ML, et al. Development of a real-time reverse transcriptase PCR assay for type A influenza virus and the avian H5 and H7 hemagglutinin subtypes. J Clin Microbiol. 2002 Sep;40:3256-60.

8. Centers for Disease Control and Prevention (CDC). CDC protocol of realtime RTPCR for influenza A(H1N1). 2009 October 6 [cited 2009 16 October]; Revision 2:

[Available from:

www.who.int/csr/resources/publications/swineflu/realtimeptpcr/en/index.html]

9. Wang WJ, Sarkodie F, Danso K, Addo-Yobo E, Owusu-Ofori S, Allain JP, et al. Seroprevalence of *West Nile virus* in Ghana. Viral Immuol. 2009 Feb;22:17-22.

CHAPTER 4 TABLES

Table 4.1: Seven individuals making up *Flavivirus* positive tissue pool 30. Swab samples were collected from four of the individuals in tissue pool 30. All 11 samples were tested individually for Flavivirus presence. All species are resident to Africa.

Order	Family	Species	Tissue	Swab	Site
Galliformes	Phasianidae	*Francolinus bicalcaratus*	T30	-	Gbele
Musophagiformes	Musophagidae	*Tauraco macrorhynchus*	T30	-	Ankasa
Passeriformes	Muscicapidae	*Alethe diademata*	T30	S19	Ankasa
Passeriformes	Pycnonotidae	*Andropadus virens*	T30	-	Ankasa
Passeriformes	Pycnonotidae	*Andropadus virens*	T30	S12	Ankasa
Passeriformes	Oriolidae	*Oriolus brachyrhynchus*	T30	S39	Ankasa
Passeriformes	Passeridae	*Petronia dentata*	T30	S17	Gbele

Table 4.2: Comparison at nucleotide level of sequenced fragment and *Yaoundé virus* (YAOV) available on GenBank (catalog numbers provided). Sequences from *Japanese Encephalitis virus* (JEV; 22 nt differences) and *West Nile virus* lineage 1 (WNV1; 21 nt differences) - taxa closely related to YAOV – are included to illustrate the close relationship between our sequence and YAOV. Dash indicates coincident nucleotide representation.

Strain	**0'**				**50'**
YAOV: EU074036	AAAACGTGAG	AAGAAGCCAG	GGGAATTCGG	AAAGGCAAAA	GGAAGCCGAG
Ave-Ghana: HQ290163	----------	-----A----	----------	---A------	--------G-
JEV: M18370	----A-A---	--------T-	-A--G--T--	---A--T---	------A-G-
WNVI: DQ211652	---GA-A---	--A--A--C-	-A--G-----	------C--G	------A---

	51'				**100'**
YAOV: EU074036	CCATCTGGTA	CATGTGGCTC	GGAGCCCGAT	TCTTGGAGTT	TGAAGCCCTT
Ave-Ghana: HQ290163	----T-----	----------	-----T----	-TC-------	---G------
JEV: M18370	----T----T	---------T	-----A--G-	ATC-A-----	------TT-G
WNVI: DQ211652	----T----T	----------	-----T--C-	-TC-------	C--G--T--G

CHAPTER 5: INFLUENZA A VIRUS INFECTIONS IN NON-MIGRANT LAND BIRDS IN ANDEAN PERU

ABSTRACT

As part of ongoing surveillance for influenza viruses (AI) in Peruvian birds, we sampled 600 land birds of 177 species, using real-time reverse transcription–PCR testing. The study addressed the assumption that AI prevalence is low or nil among land birds, a hypothesis that was not supported by the results—rather, we found Influenza A infections at relatively high prevalences in birds of the orders Apodiformes (hummingbirds) and Passeriformes (songbirds). Surveillance programs for monitoring spread and identification of AI viruses should thus not focus solely on water birds.

INTRODUCTION

The emergence and hemispheric spread of highly pathogenic avian influenza strain H5N1 (hereafter "H5N1") in the past 13 years led to renewed surveillance for avian influenza viruses (AI) and study of influenza ecology. H5N1 is pathogenic to humans [(1)] (though human pandemic fears have yet to be realised), has decimated some poultry flocks [(2)] and wild bird communities [(3)], and is notable for its rapid hemispheric range expansion [(4)]. Given rapid evolution in influenza strains, especially the potential for gene swapping, surveillance in poorly sampled groups and regions is a priority.

Wild birds, particularly of the aquatic bird orders, Anseriformes (ducks and geese) and Charadriiformes (shorebirds, gulls and terns), are widely considered to be AI reservoirs, particularly for low-pathogenic strains [(5)]. As a result, surveillance for AI and H5N1 has tended towards surveillance in water birds. Despite the apparent association with water birds, however, AI strains (including H5N1) have been detected more broadly across class Aves [(6-7)]. However, 86% of 213,115 birds sampled by USDA and other US agencies were Anseriformes or Charadriiformes [(8)]; sampling in the EU was similarly unbalanced (78.4% of 111,621 identified individuals were aquatic birds [(6)]). Nonetheless, known H5N1 hosts are diverse, and land birds, such as *Passer montanus*, may play an important transmission role [(9)]. While low pathogenic AI is associated with wild birds, highly pathogenic strains are best-known in poultry: most H5N1 cases have been detected in poultry [(2)]; only one other highly pathogenic AI strain has been detected, in wild *Sterna hirundo* in South Africa [(10)].

Information on AI host distribution and genotypes in wild birds in South America is sparse, as a result of limited surveillance [(11)]. To date, AI has been detected in wild birds in

Argentina [12-13], Bolivia [11], Brazil [14], Chile [15], and Peru [16], and from the Caribbean [17]. Surveys in North America have detected AI strains in birds that migrate between North and South America [18-19]. Highly pathogenic influenza has been detected once in South America: H7N3 in poultry at Los Lagos, Chile [15]. The highly pathogenic strain of H5N1 has not been detected in the Western Hemisphere.

We conducted broad–scale influenza surveillance among land birds, without regard to previous assumptions regarding taxonomic distribution in wild hosts, to assess the distribution of AI strains among land birds in Peru. We tested 600 individuals assignable to 177 species of 35 families in 14 orders for presence of influenza matrix gene using real-time Reverse Transcription-PCR (rRT-PCR). We further assessed whether influenza prevalence varies according to host characteristics, including taxonomy, feeding habits, habitat preferences, etc. As our sample included only one unambiguously migratory species, and as sampling was carried out in the boreal summer, this study represents a baseline of AI prevalence prior to the annual influx of boreal migrant species and any AI strains that they might introduce into the environment. This is the first broad-scale analysis of influenza distribution in South American land birds and one of few worldwide [7]; to our knowledge, we document the first detection of influenza in hummingbirds (family Trochilidae), and for that matter in all avian host species found positive herein.

METHODS

Field methods

In June 2008, we sampled birds from two mostly forested, Andean sites in Ayacucho Department, Peru (Figure 1), within 3 km of the towns of Ccano (12.785°S, 73.995°W) and

Tutumbaro (12.733°S, 73.956°W), at elevations of 2800-3100 m and 1800-2000 m, respectively. Sampling was conducted by mist netting and selective harvesting with shotguns; all individuals were apparently healthy when sampled.

Specimens were prepared following standard procedures as vouchers for the identification of each sample and to document host sex, age, and condition. Voucher specimens were deposited at the University of Kansas Natural History Museum and Centro de Ornitologia y Biodiversidad (CORBIDI), Lima, Peru. Ages of host individuals were determined by plumage characteristics, skull ossification, and dissection to determine presence or absence of a bursa of Fabricius. Buccal and cloacal swabs and tissue samples from intestine, and lung were collected. In total, samples were collected from 600 individuals of 177 species (Table 1).

RNA extraction and viral characterization

Buccal (N = 194 pools of 535 individuals) and cloacal (N = 192 pools of 538 individuals) swabs, and intestine (N = 191 pools of 569 individuals) and lung (N = 194 pools of 561 individuals) tissue samples were pooled by species (maximum 6 individuals/pool). Viral RNA was extracted directly from swabs using the QIAamp Viral RNA Mini Kit (Qiagen, Valencia, CA, USA). Intestine and lung tissue were homogenized using a high-speed mechanical homogenizer (Mixer Mill MM300; Qiagen) for 1 minute at 8000 rpm, and viral RNA was extracted from the supernatant using the RNEasy Mini Kit (Qiagen).

All samples were tested for influenza Matrix gene by rRT-PCR using the Centers for Disease Control and Prevention (CDC) protocol released on 6 October 2009 (revision 2) (www.who.int/csr/resources/publications/swineflu/realtimeptpcr/en/index.html). Invitrogen

SuperScript™III Platinum® One-Step Quantitative Kit (Invitrogen Inc, Carlsbad, CA, USA) was used for nucleic acid amplification, with a 20 µl reaction mixture containing: 5.5 µl of nuclease-free water, 0.5 µl of each primer (20 nmol), probe and kit-supplied RT-PCR super script enzyme, 12.5 µl of kit-supplied RT-PCR master mix, and 5 µl of extracted viral RNA. The primers/probe for reverse transcription of viral RNA genome were: GAC CRA TCC TGT CAC CTC TGA C (forward); AGG GCA TTY TGG ACA AAK CGT CTA (reverse); FAM-TGC AGT CCT CGC TCA CTG GGC ACG-BHQ-1 (probe). RRT-PCR was carried out in an ABI Prism 7500 machine (Applied Biosystems™, Foster City, CA, USA), with the following conditions for the RT step (50°C for 30 min and 95°C for 2 min) and for the PCR amplification (95°C for 15 sec and 55°C for 30 sec for 45 cycles), collecting fluorescence data during the 55°C incubation step. Results were read before 40 cycles were completed. Individual samples from positive pools were tested further to establish the identity of the positive individual.

All samples determined positive by rRT-PCR were processed for conventional PCR to amplify the haemagglutinin (HA), matrix (M), and neuraminidase (NA) genes. The genome segments of AI were amplified by RT-PCR, following Hoffmann et al. [20]. Briefly, 4 µl of RNA were transcribed into cDNA using AMV reverse transcriptase (Promega, Madison, WI) following the manufacturer´s protocol, and 500 ng of Uni12 primer in a 30 µl reaction mix. The RT reaction was performed at 42°C for 60 min. 1 µl of the RT-reaction was used for each PCR reaction. The cDNA was amplified using the Expand High-Fidelity PCR system (Roche Diagnostics, Mannheim, Germany) following manufacturer´s protocols. The final concentration of Mg^{2+}-ions was 1.5 mM; primer concentrations were 1.5 µM. The first cycle consisted of 4 min at 94°C, followed by 30 cycles with the following conditions:

94°C for 20 sec, 58°C for 30 sec, and 72°C for 7 min, and a final extension of 7 min at 72°C.

Following amplification, 10 µl of each reaction mixture was analysed by electrophoresis with a 2% agarose gel containing 0.5 µg/ml of ethidium bromide in TBE buffer gels. Products were visualised under UV light. A 1 Kb DNA ladder was included on each gel. Bands from this amplification were purified using a QIAquick PCR Purification Kit (Quiagen) and directly sequenced with the BigDye 3.1 terminator kit (Applied Biosystems™) on an ABI 3730 (Applied Biosystems™).

We attempted viral isolation to obtain reference material, and confirm the findings of rRT-PCR: positive pools, and all individual samples from positive pools (stored prior to RNA extraction) were centrifuged and filtered (Millex 0.45 µm) before inoculation of 0.2 ml into the allantoic cavity of 4 specific-pathogen-free 9-day-old chicken embryos. Eggs were incubated for 6 d; survival was checked daily. Allantoic fluid of each egg was tested for haemagglutinating agents by a direct haemagglutination assay [21]. First-passage negative pools were repassaged three times to confirm influenza absence.

Statistical testing

Behavioural and ecological characteristics (age, body mass, bursa presence, diet, foraging strata, habitat, mixed flocking tendency, migratory status, phylogeny, and sociability) of each species were gleaned from the appropriate literature (22-32). Odds ratios (ORs) and 95% confidence intervals (CIs) were calculated to measure associations between numbers of positive versus negative influenza cases in individuals and species displaying particular

characteristics (e.g. forest birds versus open country birds). Relationships were tested using Pearson's chi-square, or Fisher's exact test when sample size was <100. Associations were considered statistically significant if p value was <0.05.

The effects of body mass (the continuous variable) were tested by binary logistic regression. Each characteristic was tested overall and by site. For the purposes of statistical analysis, we made the conservative assumption that only one individual was positive in the two pools for which we were unable to determine the infected individual's identity. We could not establish the age for one positive individual. All statistical analyses were carried out using Minitab 12, (Minitab Inc., PA, USA) and SigmaPlot 11.0 (Systat Software Inc., IL, USA).

Phylogenetic analysis

We compared our sequence to published sequences using NCBI's Basic Local Alignment Search Tool (available at http://www.ncbi.nlm.nih.gov/blast/Blast.cgi) to identify closely related strains; sequence alignments were carried out with ClustalX [33] using default parameters. The phylogenetic analysis covered a 922 nt region and included123 influenza matrix gene sequences obtained from GenBank. The dataset represented all available HA and NA subtypes, selected to maximise geographic and taxonomic diversity; nevertheless, sequences were dominated by AI strains isolated from Anseriformes and Charadriiformes, from established sampling localities and collected within the last decade (Appendix Table 5.1). We used the program MrBayes v. 3.1.2 [34] for construction of phylogenetic trees from our RNA data, as this method is based on the likelihood function (the quantity that is proportional to the probability of observing the data conditional on a tree; see [34-35] for

reviews). Four independent analyses were run for 10^7 generations with four Markov chains under default heating values, sampling every 10^3 generations. Stationarity of MCMC analyses was determined by plotting−lnL against generation time, and the "burn -in" trees sampled prior to stationarity (first 10% of trees sampled were discarded). The sample H1N7UK92 was designated as the outgroup taxon in all analyses.

RESULTS

Eight of 196 pools (4.1%) of all sampling types were positive by rRT-PCR (Table 1; Appendix Table 5.2). The identity of six of the eight AI-positive individuals was established; the identity of the positive individual within two pools could not be established. Positive species were all from the orders Apodiformes (3 of 19 species; 15.8%; all Trochilidae) and Passeriformes (5 of 127 species; 3.9%), these orders constituted 10.7% and 76.3% of the species in the overall sample, respectively (Table 2). Among Passeriformes, two positive pools were detected in the New World warbler family Parulidae (two of six species; 33.3%). Four positive pools were detected from each site. Five of 194 buccal and one of 192 cloacal swab and three of 194 lung and zero of 191 intestine pools tested positive. Only one sample, of *Cinnycerthia peruana*, tested positive in more than one sample type (buccal swab and lung). All sample types (i.e. lung, intestine, etc.) were tested from all positive pools, with the exception of the hummingbird *Chlorostilbon mellisugus*, for which insufficient lung was available for testing. No positive species was associated with water; five species sampled that are associated with water tested negative for influenza among samples available.

Three individuals from rRT-PCR-positive pools tested positive for the AI matrix gene (using primer pair Bm-M-1 / Bm-M-1027R), and one 922 nt sequence (GenBank accession number HQ420901) was obtained, from the woodcreeper *Campylorhamphus pucherani* (Passeriformes; Dendrocolaptidae). Two individuals from rRT positive pools were positive for neuraminidase (using primer pair Bm-NA-1 / Bm-NA-1413R, which is designed to detect NA strains 1, 2, 4, 5 and 8). We found no evidence H5N1 presence, though full strain-level characterization proved impossible. No samples were positive by the haemagglutination assay, and none showed obvious cytopathic affects.

Statistical testing

We found a significant association (Appendix 5.3) of AI prevalence and host order, with Apodiformes (hummingbirds; 3 of 19 species and 75 individuals) more prone to AI infection than all other individuals (5 of 158 species and 525 individuals; OR 5.7; 95% CI, 1.4-24.1; $p = 0.012$; OR 4.3; 95% CI, 1.1-16.8; $p = 0.031$; for species and individuals, respectively). AI prevalences were significantly higher in Apodiformes species than in non-Apodiformes and Passeriformes – the later showing significantly lower prevalences than those in Apodiformes. Families Parulidae (New World warblers), and Trochilidae (hummingbirds) appeared to show higher-than-expected prevalences as compared with other families.

We found a significant association between elevated AI prevalence and nectarivory as opposed to other dietary habits (though see discussion), and use of forest edge habitats. We found no associations between overall influenza prevalence and adult status, body mass, foraging strata, mixed flocking behavior, sex, or sociability, or in other dietary, and habitat

variables tested. There was no variation in influenza prevalence associations by site (Appendix Table 5.4).

Phylogenetic analysis

BLAST searches revealed that our influenza strain shared >97% identity with the 100 most similar influenza sequences. Maximum identity was found to strains of H10N7; however, at least 22 influenza serotypes were found amongst the 100 most similar sequences. All 100 closely related strains were detected in at least three bird orders, and all were from North America. Our phylogenetic tree confirmed close affinities between our sequence and other North American sequences: generally, same-hemisphere sequences grouped together with only three of 124 sequences clearly grouping with isolates from outside their hemisphere of origin (Figure 2A).

Although phylogenetic analysis identified same continent-specific clades (e.g., Oceania), some genotypes appeared more widespread within their hemisphere than others (Figure 2B). Sequences did not appear to group according to host species type (land or water bird; Figure 2C) although the picture was less clear when analyzing phylogeny by host order (Figure 2D). Some clades were hosted by only Anseriformes or Charadriiformes (which, of course, dominate the sampling as well). Additionally, two of three Australian sequences from Procellariformes showed close identity (Figure 2D). Frequently sequences showed close relationships to sequences detected within the same decade (Figure 2E), though this effect was correlated with hemisphere, and to a lesser extent continent, of detection. Our sequence clustered with sequences from North American Anseriformes, all collected between 2000 and 2009. At this early stage in exploratory analyses, then, we can see some

limited evidence for phylogenetic structure among wild-bird AI isolates in terms of continent of origin, time of detection, and host type. Clear associations cannot be confirmed without a much-broader analysis.

DISCUSSION

This study is the first broad survey for influenza distributions in land birds in South America. We are not aware of previous detection of AI in any of the influenza positive species we identified. This work also describes the first AI detection in hummingbirds (Apodiformes: Trochilidae), to the best of our knowledge. It is noteworthy that prevalence is higher among hummingbird taxa than in other bird groups. All hummingbird species are nectarivorous, and prevalence among nectarivorous species was significantly higher than in non-nectarivores (OR 3.4; 95% CI, 0.9-13.7; $p = 0.045$). Nectarivory is a plausible transmission route for respiratory diseases. However, as other nectarivores are few in the communities sampled, the high prevalence found among hummingbirds may be explained by factors unrelated to feeding type. The use of edge habitats was also positively associated with AI prevalences, of note as it suggests a mechanism for transmission between forest and open habitats.

Statistical analysis of natural history characteristics suggests that influenza infections are broadly distributed with regard to ecological characteristics of host species. For instance, influenza infection prevalence was higher in Parulidae and Trochilidae than in other families. Parulidae are highly social, insectivorous species that generally travel in mixed flocks; in our sample, all forest birds. Behavioural traits such as sociability and mixed species flocking are plausible risks for AI transmission, though, overall, we found them not

to be significantly associated with prevalences. By contrast, Trochilidae are solitary, aggressively territorial, largely nectarivorous, and (in our sample) mixed in use of forest and more open habitats. As in a parallel study from our research group based on samples from southern China [7], we did not detect significant associations between influenza prevalences in open-country (1.74%) and forest (1.17%) species (cumulative binomial probability, $p>0.05$). We detected no variation of AI prevalence among sites.

Overall prevalence in Ayacucho (1.3%) was triple that documented recently in Peruvian water birds (0.38%) tested in the Lima region of Peru [16]. Higher detection rates may result from superior field collection protocols (swabs and tissue *versus* fecal samples), rather than reflecting higher prevalences in land birds than in water birds. Ayacucho prevalence is nonetheless lower than the 2.3% found in our parallel survey of southern Chinese birds, and somewhat lower than the prevalence in non-migratory birds in that study (1.8%) [7]. Our prevalence in Passeriformes (1.1%) is lower than the prevalence found in Anseriformes (3.0%), somewhat higher than that in Charadriiformes (0.8%) and an order of magnitude higher than that in Passeriformes (0.1%) in a recent European study [6], and considerably lower than prevalences found in a water bird dominated surveillance in North America (9.7% and 11% in 2007 and 2008, respectively) [8]. However, it is clear from our findings that resident land birds may play a non-negligible role in hosting AI, at least in East Asia and South America.

Our phylogenetic analysis revealed that six of seven South American sequences included within our analysis grouped with other Western Hemisphere sequences. Of these, four formed a well-supported South American clade (Figure 2B) echoing Alvarez *et al*. [12] that

some South American M gene isolates were most closely related to other South American isolates. Two sequences, our own and Cinnamon Teal Bolivia H7N3 [11], were most closely related to North American taxa, demonstrating that other South American M gene isolates were most closely related to North American wild bird lineages, as found by Spackman *et al*. [11, 36]. The remaining South American sequence, a "pandemic" H1N1 strain, detected on a Chilean domestic turkey farm [37], had closest affinities to African and Asian isolates of wild bird origin. Intriguingly, the African taxa (Tern South Africa H5N3, 1959) shared 100% M gene identity with the only known highly pathogenic wild bird strain apart from H5N1: Tern South Africa H5N3, 1961. "Pandemic" H1N1, first detected in La Gloria, Mexico, was widely assumed to be of Western Hemisphere origin [38]; our findings suggest origins may be more complicated, and merits further study. Sequences from land and water birds did not form distinct clades (Figure 3A), although our dataset was biased heavily towards isolates from water birds (108 of 124 sequences), echoing the ideas of Chen and Holmes [39] that geographic location and time, rather than host species, shape patterns of influenza diversity. In our analysis, the picture was less clear as regards host order (Figure 3B). Our sequence (detected in order Passeriformes) clustered with sequences from North American Anseriformes. However, the analysis was biased because 99 of 124 sequences available were isolated from hosts in the orders Anseriformes and Charadriiformes.

CONCLUSIONS

AI was detected in eight of 600 individuals of 177 bird species sampled in Ayacucho Department, Peru. All positive species were Apodiformes (prevalence 15.8%; significantly higher than other groups) and Passeriformes (3.9%). Two avian families, Parulidae (New

World warblers), and Trochilidae (hummingbirds), and two natural history characteristics (nectarivory, and use of forest edge habitat) showed significantly higher prevalences than expected by chance.

Phylogenetic analysis indicated that our sequence shares greatest identity with sequences originating in North American Anseriformes. Geography and time appear better predictors of relatedness of influenza M gene than host taxonomy [39]: hemisphere and (to a lesser extent) continent, and decade of detection were good predictors of relatedness, while host habit (i.e., land *versus water* birds) and host order were poor predictors, with the caveat that our analysis was dominated by sequences detected in two water bird orders. Our findings concur with Peterson et al. [7], that prevalence of influenza strains is not negligible in land birds, and that to prioritize surveillance on water birds is misguided, particularly given that absolute numbers of land birds are vastly higher than of water birds [40].

ACKNOWLEDGEMENTS

Karen Segovia Hinostroza, Bruno M. Ghersi, Victor Gonzaga, A. Townsend Peterson, and Joel, M. Montgomery are co-authors on the manuscript associated with this chapter. We thank ornithologists Mike Andersen, Roger Boyd, Adolfo Navarro, Luis Sanchez-Gonzalez, Mark Robbins; the staff of the Centro de Ornitologia y Biodiversidad (CORBIDI), Lima, Peru, for extensive assistance with logistics associated with field sampling of birds; M. Papeş for assistance with maps; A.E. Gonzalez for extending facilities and hospitality at the School of Veterinary Medicine, Universidad Nacional Mayor de San Marcos. The study was funded by the National Institutes of Health - Fogarty International Center and the US DoD AFHSC-GEIS grant number GF0121_07_LI.

CHAPTER 5 REFERENCES

1. Claas ECJ, de Jong JC, van Beek R, Rimmelzwaan GF, Osterhaus A. Human influenza virus A/HongKong/156/97 (H5N1) infection. Vaccine. 1998;16:977-8.

2. Whitworth D, Newman S, Mundkur T, Harris P, editors. Wild birds and avian influenza: an introduction to applied field research and disease sampling techniques. Rome: FAO; 2007.

3. Chen H, Smith GJD, Zhang SY, Qin K, Wang J, Li KS, et al. H5N1 virus outbreak in migratory waterfowl. Nature. 2005;436:191-2.

4. Gauthier-Clerc M, Lebarbenchon C, Thomas F. Recent expansion of highly pathogenic avian influenza H5N1: a critical review. Ibis. 2007;149:202-14.

5. Swayne DE, Suarez DL. Highly pathogenic avian influenza. Rev Sci Tech Off Int Epiz. 2000;19:463-82.

6. Hesterberg U, Harris K, Stroud D, Guberti V, Busani L, Pittman M, et al. Avian influenza surveillance in wild birds in the European Union in 2006. Influenza Other Respir Viruses. 2009;3:1-14.

7. Peterson AT, Bush SE, Spackman E, Swayne DE, Ip HS. Influenza A virus infections in land birds, People's Republic of China. Emerg Infect Dis. 2008;14:1644-6

8. DeLiberto TJ, Swafford SR, Nolte DL, Pedersen K, Lutman MW, Schmit BB, et al. Surveillance for highly pathogenic avian influenza in wild birds in the USA. Integr Zool. 2009;4:426-39.

9. Kou Z, Li YD, Yin ZH, Guo S, Wang ML, Gao XB, et al. The survey of H5N1 flu virus in wild birds in 14 provinces of China from 2004 to 2007. PLoS One. 2009;4: e6926.

10. Becker WB. The isolation and classification of Tern virus: influenza A-Tern South Africa-1961. J Hyg. 1966;64:309-20.

11. Spackman E, McCracken KG, Winker K, Swayne DE. An avian influenza virus from waterfowl in South America contains genes from North American avian and equine lineages. Avian Dis. 2007;51:273-4.

12. Alvarez P, Mattiello R, Rivailler P, Pereda A, Davis CT, Boado L, et al. First isolation of an H1N1 avian influenza virus from wild terrestrial non-migratory birds in Argentina. Virology. 2010;396:76-84.

13. Pereda AJ, Uhart M, Perez AA, Zaccagnini ME, La Sala L, Decarre J, et al. Avian influenza virus isolated in wild waterfowl in Argentina: evidence of a potentially unique phylogenetic lineage in South America. Virology. 2008;378:363-70.

14. Senne DA. Avian influenza in North and South America, 2002-2005. Avian Dis. 2007;51:167-73.

15. Suarez DL, Senne DA, Banks J, Brown IH, Essen SC, Lee CW, et al. Recombination resulting in virulence shift in avian influenza outbreak, Chile. Emerg Infect Dis. 2004;10:693-9.

16. Ghersi BM, Blazes DL, Icochea E, Gonzalez RI, Kochel T, Tinoco Y, et al. Avian influenza in wild birds, central coast of Peru. Emerg Infect Dis. 2009;15:935-8.

17. Douglas KO, Lavoie MC, Kim LM, Afonso CL, Suarez DL. Isolation and genetic characterization of avian influenza viruses and a Newcastle Disease virus from wild birds in Barbados: 2003-2004. Avian Dis. 2007;51:781-7.

18. Dusek RJ, Bortner JB, DeLiberto TJ, Hoskins J, Franson JC, Bales BD, et al. Surveillance for high pathogenicity avian influenza virus in wild birds in the Pacific Flyway of the United States, 2006-2007. Avian Dis. 2009;53:222-30.

19. Hanson BA, Luttrell MP, Goekjian VH, Niles L, Swayne DE, Senne DA, et al. Is the occurrence of avian influenza virus in Charadriiformes species and location dependent? J Wildl Dis. 2008;44:351-61.

20. Hoffmann E, Stech J, Guan Y, Webster RG, Perez DR. Universal primer set for the full-length amplification of all influenza A viruses. Arch Virol. 2001;146:2275-89.

21. Swayne D, Senne D, Beard C. Avian influenza. In: Swayne D, Glisson J, Jackwood M, editors. A laboratory manual for the isolation and identification of avian pathogens. 4th ed. Kennett Square, PA: International Book Distributing Co.; 1998. p. 150–5.

22. Hoyo Jd, Elliott A, Sargatal J, Christie DA, BirdLife I. Handbook of the birds of the world. Barcelona: Lynx Edicions; 1992-2007.

23. Schulenberg TS, Douglas F. Stotz DF, Daniel F. Lane DF, John P., et al. Birds of Peru: Princeton University Press; 2007.

24. Ridgely RS, Tudor G. The birds of South America: volume 1: the Oscine Passerines. Austin, USA: University of Texas Press; 1989.

25. Ridgely RS, Tudor G. The birds of South America: volume 2: the Suboscine Passerines. Austin, USA: University of Texas Press; 1994.

26. Madge S, Hilary Burn H. Crows and jays. London, UK: Christopher Helm; 1999.

27. Jaramillo A, Burke P. New World blackbirds: the Icterids. London, UK: Christopher Helm; 1999.

28. Isler ML, Isler PR. Tanagers. 2nd ed. London, UK: Christopher Helm; 1999.

29. Curson J, Quinn D, Beadle D. New World warblers: an identification guide. London, UK: Christopher Helm; 1994.

30. Byers C, Olsson U, Curson J. Buntings and sparrows: a guide to the buntings and North American sparrows. London, UK: Christopher Helm; 1995.

31. Winkler H, Christie D. Woodpeckers: a guide to the woodpeckers, piculets and wrynecks of the world. London, UK: Christopher Helm; 1995.

32. Restall R, Rodner C, Lentino M. Birds of northern South America: an identification guide, volume 1: species accounts. London, UK: Chistopher Helm; 2006.

33. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, et al. Clustal W and Clustal X version 2.0. Bioinformatics. 2007;23:2947-8.

34. Huelsenbeck JP, Ronquist F. MRBAYES: bayesian inference of phylogenetic trees. Bioinformatics. 2001;17:754-5.

35. Holder M, Lewis PO. Phylogeny estimation: traditional and Bayesian approaches. Nature Rev Gen. 2003;4:275-84.

36. Spackman E, McCracken KG, Winker K, Swayne DE. H7N3 avian influenza virus found in a South American wild duck is related to the Chilean 2002 poultry outbreak, contains genes from equine and north American wild bird lineages, and is adapted to domestic turkeys. J Virol. 2006;80:7760-4.

37. Mathieu C, Moreno V, Retamal P, Gonzalez A, Rivera A, Fuller J, et al. Pandemic (H1N1) 2009 in breeding turkeys, Valparaiso, Chile. Emerg Infect Dis. 2010;16:709-11.

38. Fraser C, Donnelly CA, Cauchemez S, Hanage WP, Van Kerkhove MD, Hollingsworth TD, et al. Pandemic potential of a strain of influenza A (H1N1): early findings. Science. 2009;324:1557-61.

39. Chen R, Holmes EC. Frequent inter-species transmission and geographic subdivision in avian influenza viruses from wild birds. Virology. 2009;383:156-61.

40. BirdLife International. Datazone: species accounts 2009. [Downloaded from http://www.birdlife.org on 15/10/2009].

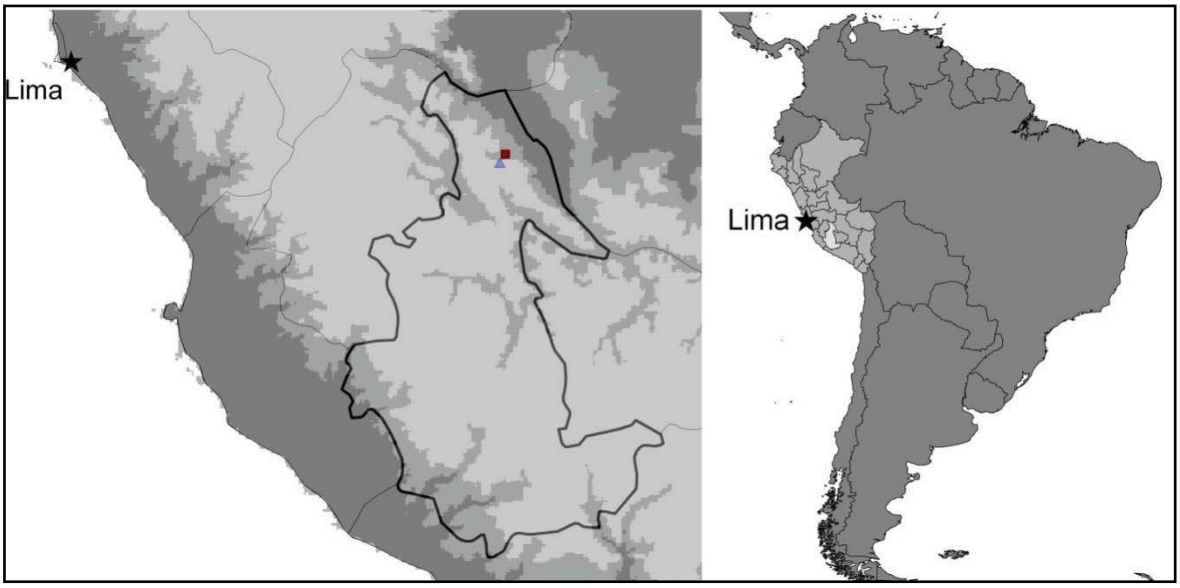

Figure 5.1: Ayacucho, Peru, showing the two Andean sites from which land birds were collected and tested for influenza A virus infections. Prevalences were 1.5% (n = 265) in Ccano (triangle): and 1.2% (n = 335) in Tutumbaro (square). Elevation is indicated by shading; sea level-2000m, (dark grey), 2000–3100m (medium grey), >3100m (light grey).

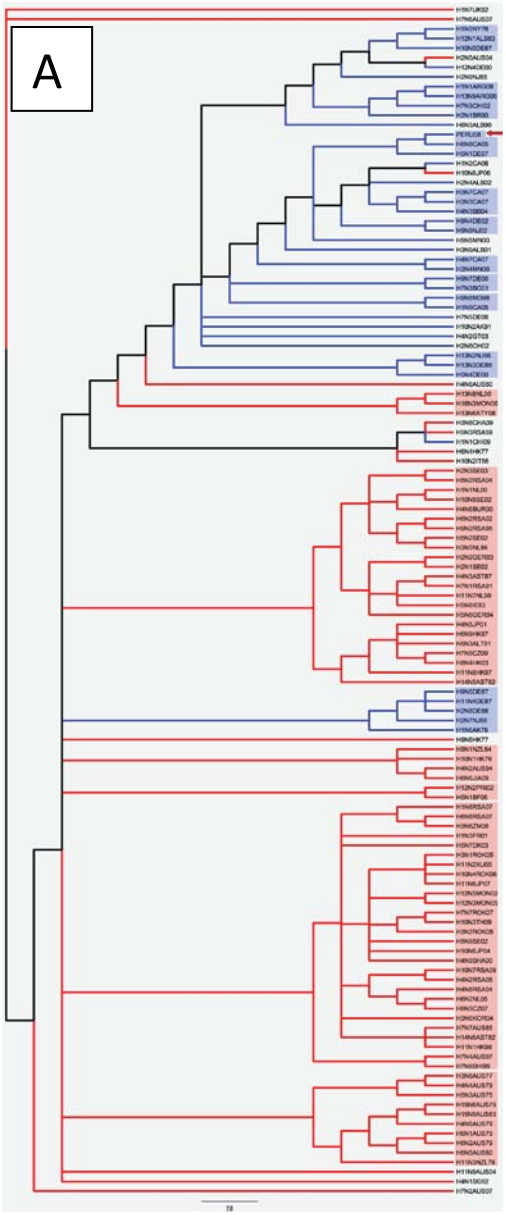

Figure 5.2A: Bayesian phylogenetic tree based on alignment of influenza matrix nucleotide sequences from representatives of all available avian influenza HA/NA subtypes. Our sample, labelled PERU08 (Genbank Accession Number HQ420901), is indicated by a red arrow. A shows hemisphere of detection (blue = western hemisphere; red = eastern hemisphere). Branches are colour-coded as stated above; where most closely related sequences group by hemisphere highlighted clades are colour-coded as stated above.

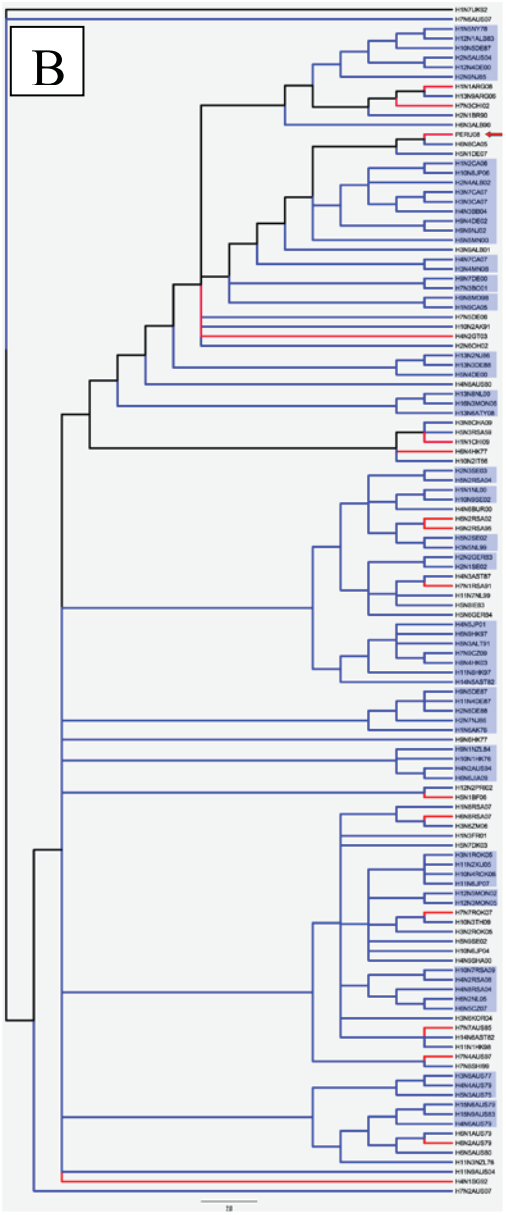

Figure 5.2B: Bayesian phylogenetic tree showing continent of detection - Africa (purple), Asia (red), Australia (grey), Europe (light blue), North America (blue) and South America (green). Branches are colour-coded as stated above; where most closely related sequences group by continent, highlighted clades are colour-coded as stated above.

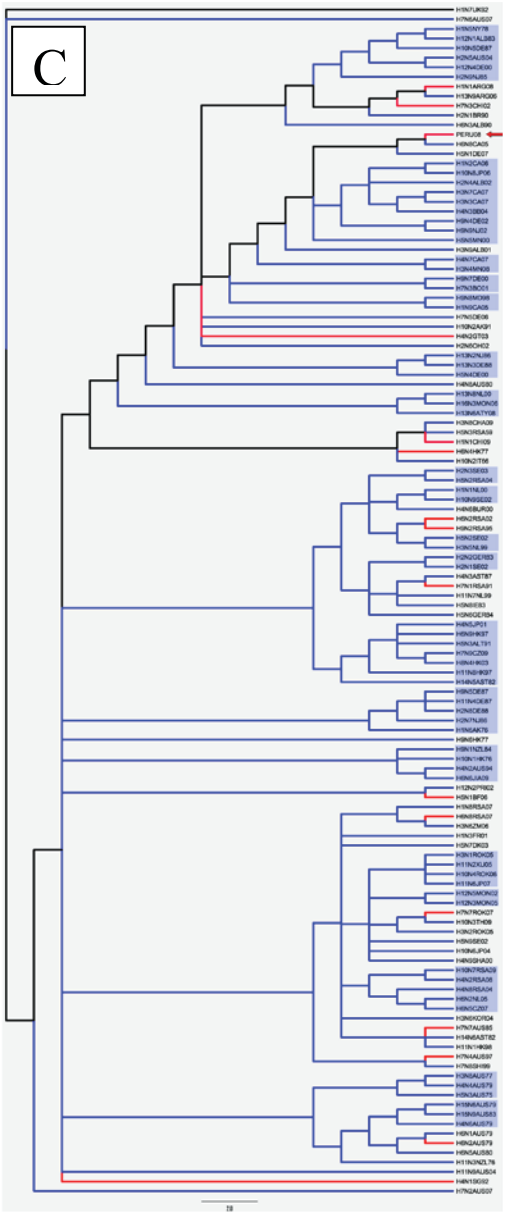

Figure 5.2C: Bayesian phylogenetic tree showing which sequences were detected in land birds (red) versus water birds (blue). Branches are colour-coded as stated above; where most closely related sequences group by habitat, highlighted clades are colour-coded as stated above.

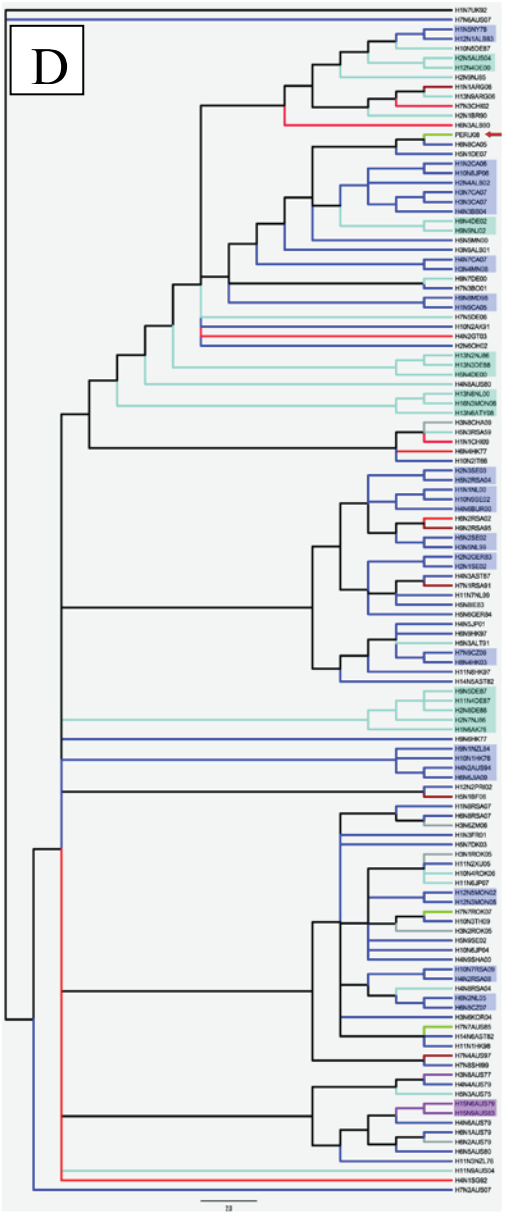

Figure 5.2D: Bayesian phylogenetic tree showing order of detection - Anseriformes (blue), Charadriiformes (light blue), Galliformes (red), other land birds (brown), other water birds (grey), Passeriformes (green), Procellariformes (purple) and Artiodactyla (black.). Branches are colour-coded as stated above; where most closely related sequences group by taxa (D), highlighted clades are colour-coded as stated above.

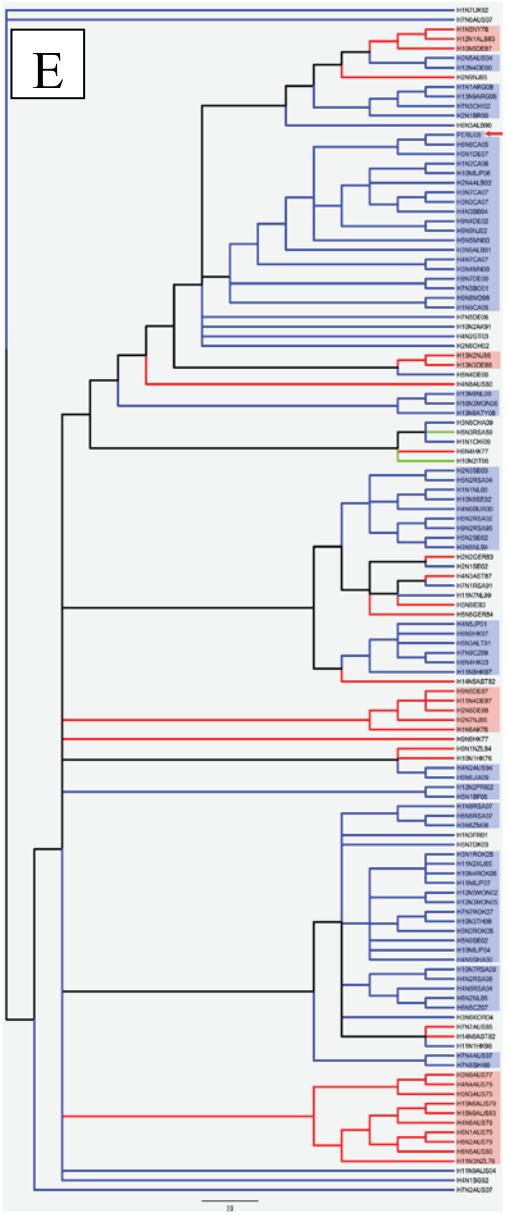

Figure 5.2E: Bayesian phylogenetic tree showing the time period of detection - 1950-1969 (green), 1970-1989 (red), 1990-2009 (blue). In each tree, branches are colour-coded as stated above; time, highlighted clades are colour-coded as stated above.

Table 5.1. Prevalence of influenza A virus in avian orders and families at 2 sites, Ayacucho, Peru.

Order	Family	Ccano	Tutumbaro	Total
Anseriformes	Anatidae		0/2	0/2
Apodiformes	Trochilidae	0/38	3/37	3/75
Caprimulgiformes	Caprimulgidae	0/1		0/1
Columbiformes	Columbidae	0/3	0/3	0/6
Coraciiformes	Momotidae		0/3	0/3
Cuculiformes	Cuculidae		0/2	0/2
Falconiformes	All	0/4	0/2	0/6
	Accipitridae	0/2	0/2	0/4
	Falconidae	0/2		0/2
Galliformes	Cracidae	0/2	0/4	0/6
Gruiformes	Eurypigidae		0/1	0/1
Passeriformes	All	4/204	1/254	5/458
	Cardinalidae	0/1	0/3	0/4
	Cinclidae	0/1	0/4	0/5
	Connopophagidae		0/3	0/3
	Corvidae	0/2	0/8	0/10
	Cotingidae	0/8	0/4	0/12
	Emberizidae	0/8	0/20	0/28
	Formicariidae	0/5	0/4	0/9
	Furnariidae	1/20	0/29	1/49
	Hirundidae		0/4	0/4
	Icteridae	0/10	0/2	0/12
	Parulidae	1/10	1/27	2/37
	Pipridae		0/1	0/1
	Rhinocryptidae	0/6	0/3	0/9
	Thamnophilidae		0/6	0/6
	Thraupidae	1/51	0/59	1/110
	Tityridae	0/2		0/2
	Troglodytidae	1/16	0/10	1/26
	Turdidae	0/8	0/17	0/25
	Tyrannidae	0/56	0/50	0/106
Piciformes	All	0/6	0/17	0/23
	Bucconidae		0/2	0/2
	Capitonidae		0/4	0/4
	Picidae	0/5	0/7	0/12
	Ramphastidae	0/1	0/4	0/5
Psittaciformes	Psittacidae	0/1	0/4	0/5
Strigiformes	Strigidae	0/1		0/1
Trogoniformes	Trogonidae	0/3	0/8	0/11
		4/263	4/337	8/600

Table 5.2. Prevalences of influenza A infections among families and orders in which infections were detected. Overall prevalences are shown at bottom of table

Positive orders	**Positive families**	**% positive Ccano**	**% positive Tutumbaro**	**% positive Total**
Apodiformes	Trochilidae	0.00	8.11	4.00
Passeriformes	All	1.99	0.39	1.09
	Furnariidae	5.00	0.00	0.20
	Parulidae	10.00	3.70	5.41
	Thraupidae	1.96	0.00	0.01
	Troglodytidae	6.25	0.00	0.04
All		1.52	1.19	1.33

Appendix Table 5.1: Influenza A virus sequences included in phylogenetic analysis, including labels used in the figure, GenBank accession number, continent of origin, and avian host order.

Label	GenBank Acc. No.	Continent	Avian host order
PERU08	HQ420901	South America	Passeriformes
H1N1CHI09	GQ866232	South America	Galliformes
H1N1ARG08	GQ223719	South America	Tinamiformes
H1N1NL00	CY060181	Europe	Anseriformes
H1N2CA06	FJ520170	North America	Anseriformes
H1N3FR01	AM157381	Europe	Anseriformes
H1N5NY78	CY014969	North America	Anseriformes
H1N6AK76	CY015164	North America	Charadriiformes
H1N7UK92	U85985	Europe	Artiodactyla
H1N8RSA07	GQ404706	Africa	Anseriformes
H1N9CA05	CY053798	North America	Anseriformes
H2N1BR90	CY005414	South America	Charadriiformes
H2N1SE02	CY060399	Europe	Anseriformes
H2N2GER83	CY005766	Europe	Anseriformes
H2N3SE03	CY060427	Europe	Anseriformes
H2N4ALB02	CY003985	North America	Anseriformes
H2N5AUS04	AB275864	Oceania	Charadriiformes
H2N6OH02	GU053445	North America	Anseriformes
H2N7NJ86	CY003888	North America	Charadriiformes
H2N8DE88	CY004555	North America	Charadriiformes
H2N9NJ85	CY003864	North America	Charadriiformes
H3N1ROK05	EU301244	Asia	Unknown aquatic bird
H3N2ROK05	EU301249	Asia	Unknown aquatic bird
H3N3CA07	CY032731	North America	Anseriformes
H3N4MN08	CY048900	North America	Anseriformes
H3N5NL99	CY060264	Europe	Anseriformes
H3N6KOR04	GU953255	Asia	Anseriformes
H3N6ZM06	AB470299	Africa	Pelecaniformes
H3N7CA07	CY034188	North America	Anseriformes
H3N8CHA09	GQ404596	Asia	Ciconiiformes
H3N8AUS77	CY028653	Oceania	Procellariiformes
H3N9ALB01	CY004700	North America	Anseriformes
H4N1SG92	EU014144	Asia	Galliformes
H4N2RSA08	GQ404714	Africa	Anseriformes
H4N2AUS94	CY045248	Oceania	Anseriformes
H4N2GT03	GU052344	North America	Galliformes
H4N3BB04	DQ236168	North America	Anseriformes
H4N3AST87	EU580551	Europe	Anseriformes
H4N4AUS79	CY005673	Oceania	Anseriformes
H4N5JP01	AB266089	Asia	Anseriformes
H4N6BUR00	EU580595	Asia	Anseriformes
H4N6AUS79	CY005680	Oceania	Anseriformes
H4N7CA07	CY032911	North America	Anseriformes
H4N8RSA04	EF041496	Africa	Anseriformes
H4N8AUS80	CY005699	Oceania	Charadriiformes
H4N9SHA00	EF597292	Asia	Anseriformes

Label	GenBank Acc. No.	Continent	Avian host order
H5N1DE07	CY043929	North America	Charadriiformes
H5N1BF06	AM503006	Africa	Accipitriformes
H5N2RSA04	FJ519980	Africa	Anseriformes
H5N2SE02	GU052565	Europe	Anseriformes
H5N3ALT91	GQ227555	Asia	Anseriformes
H5N3RSA59	GU052815	Africa	Charadriiformes
H5N3AUS75	EF566312	Oceania	Charadriiformes
H5N4DE00	DQ107462	North America	Charadriiformes
H5N5MN00	EU871914	North America	Anseriformes
H5N6GER84	CY005770	Europe	Anseriformes
H5N7DK03	DQ251445	Europe	Anseriformes
H5N8IE83	GU052854	Europe	Anseriformes
H5N9SE02	GU052869	Europe	Anseriformes
H6N1AUS79	CY005666	Oceania	Anseriformes
H6N2AUS79	CY028244	Oceania	Gruiformes
H6N2RSA02	DQ408525	Africa	Galliformes
H6N2NL05	CY041387	Europe	Anseriformes
H6N3ALB90	CY004243	North America	Anseriformes
H6N4HK77	DQ107436	Asia	Galliformes
H6N5AUS80	CY005692	Oceania	Anseriformes
H6N5CZ07	GQ404644	Europe	Anseriformes
H6N6JIA09	GU324777	Asia	Anseriformes
H6N8CA05	CY043809	North America	Anseriformes
H6N8RSA07	GQ404698	Africa	Struthioniformes
H6N9HK97	AF250485	Asia	Anseriformes
H7N1RSA91	GU052955	Africa	Struthioniformes
H7N2AUS07	CY033162	Oceania	Anseriformes
H7N3BO01	DQ525415	South America	Anseriformes
H7N3CHI02	AY303657	South America	Galliformes
H7N4AUS97	GU053103	Oceania	Struthioniformes
H7N5DE06	CY039328	North America	Charadriiformes
H7N6AUS07	CY061611	Oceania	Anseriformes
H7N7ROK07	FJ750855	Asia	Passeriformes
H7N7AUS85	CY024779	Oceania	Passeriformes
H7N8SHI99	AB280939	Asia	Anseriformes
H7N9CZ09	GQ404575	Europe	Anseriformes
H8N4HK03	EF597299	Europe	Anseriformes
H9N1NZL84	CY005747	Oceania	Anseriformes
H9N2RSA95	AF508684	Africa	Struthioniformes
H9N4DE02	GU051538	North America	Charadriiformes
H9N5DE87	CY004413	North America	Charadriiformes
H9N6HK77	CY005640	Asia	Anseriformes
H9N7DE00	DQ107507	North America	Charadriiformes
H9N8MD98	DQ021759	North America	Anseriformes
H9N9NJ02	GU051565	North America	Charadriiformes
H10N1HK76	GQ404604	Asia	Anseriformes
H10N2IT66	GQ404583	Europe	Anseriformes
H10N2AK91	CY015169	North America	Anseriformes
H10N3TH09	CY062599	Asia	Anseriformes
H10N4ROK06	EU239835	Asia	Charadriiformes
H10N5DE87	CY004428	North America	Charadriiformes
H10N6JP04	AB428692	Asia	Anseriformes
H10N7RSA09	GQ404729	Africa	Anseriformes

Label	GenBank Acc. No.	Continent	Avian host order
H10N8JP06	AB428709	Asia	Anseriformes
H10N9SE02	CY060361	Europe	Anseriformes
H11N1HK98	AF250486	Asia	Unknown aquatic bird
H11N2XU05	GQ219713	Asia	Anseriformes
H11N3NZL76	CY005740	Oceania	Anseriformes
H11N4DE87	AY664478	North America	Charadriiformes
H11N6JP07	AB428719	Asia	Anseriformes
H11N7NL99	GU052210	Europe	Anseriformes
H11N8HK97	AF250491	Asia	Anseriformes
H11N9AUS04	CY029882	Oceania	Charadriiformes
H12N1ALB83	CY005344	North America	Anseriformes
H12N2PRI02	DQ787809	Asia	Anseriformes
H12N3MON05	GQ907343	Asia	Anseriformes
H12N4DE00	CY005358	North America	Charadriiformes
H12N5MON02	AB428679	Asia	Anseriformes
H13N2NJ86	CY004451	North America	Charadriiformes
H13N3DE88	CY005382	North America	Charadriiformes
H13N6ATY08	GU953279	Asia	Charadriiformes
H13N8NL00	AY684906	Europe	Charadriiformes
H13N9ARG06	EU523138	South America	Charadriiformes
H14N5AST82	CY014605	Europe	Anseriformes
H14N6AST82	CY005394	Europe	Anseriformes
H15N6AUS79	GU052261	Oceania	Procellariiformes
H15N9AUS83	CY005718	Oceania	Procellariiformes
H16N3MON06	GQ907295	Asia	Charadriiformes

Appendix Table 5.2. Summary of individual birds tested, Ayacucho, Peru. Kansas University Natural History Museum (KUNHM) provided the permanent tissue number. M, migratory; O, open country; R, resident; F, forested; B, both. †Individual birds infected with avian influenza virus. All species, excepting *Turdus nigriceps*, were resident.

Order	Family	Scientific name	Site	Tissue Number	No. Positive	Sample	Habitat type
Anseriformes	Anatidae	*Merganetta armata*	Tutumbaro	16745, 16754	0	2	B
Apodiformes	Trochilidae	*Aglaiocercus kingi*	Tutumbaro	16677, 16856, 16885	0	3	O
		Amazilia viridicauda	Tutumbaro	16712, 16752, 16799	0	3	O
		Boissonneaua matthewsii	Ccano	17008	0	1	F
		Chaetocercus mulsant	Tutumbaro	**16815**†, 16891	1	2	O
		Chalcostigma ruficeps	Ccano	17134, 17145, 17149, 17190	0	4	F
		Chlorostilbon mellisugus	Tutumbaro	**16861**†	1	1	O
		Coeligena coeligena	Tutumbaro	16616, ***16640***, 16678, ***16679***, 16689, 16710, 16720	1	7	O
		Coeligena torquata	Ccano	16983, 16996, 17162	0	3	F
			Tutumbaro	16790	0	1	F
		Coeligena violifer	Ccano	16962, 16977, 16984, 17002, 17050, 17063, 17064, 17078	0	8	O
		Colibri coruscans	Ccano	16924, 16988, 17100, 17101	0	4	O
			Tutumbaro	16707, 16877, 16911	0	3	O
		Colibri thalassinus	Ccano	16953, 16961, 17054, 17142, 17164	0	5	O
			Tutumbaro	16859, 16867, 16894	0	3	O
		Doryfera ludoviciae	Tutumbaro	16688, 16704, 16715, 16722, 16758	0	5	F
		Eriocnemis luciani	Ccano	17150, 17155, 17185	0	3	O
		Eutoxeres condamini	Tutumbaro	16626, 16645, 16676, 16692, 16736, 16737	0	6	O
		Heliangelus amethysticollis	Ccano	17019, 17069	0	2	F
		Lafresnaya lafresnayi	Ccano	16994, 17157	0	2	O
		Lesbia victoriae	Ccano	17125	0	1	O
		Metallura tyrianthina	Ccano	16999, 17035, 17055, 17105, 17168	0	5	F
		Phaethornis guy	Tutumbaro	16619, 16630, 16791	0	3	B
Caprimulgiformes	Caprimulgidae	*Caprimulgus longirostris*	Ccano	17184	0	1	B
Columbiformes	Columbidae	*Geotrygon frenata*	Ccano	16957	0	1	F
		Leptotila verreauxi	Tutumbaro	16921	0	1	O
		Patagioenas fasciata	Ccano	17075	0	1	F
			Tutumbaro	16850, 16888	0	2	F
		Zenaida auriculata	Ccano	17189	0	1	O
Coraciiformes	Momotidae	*Momotus aequatorialis*	Tutumbaro	16813, 16827, 16830	0	3	F
Cuculiformes	Cuculidae	*Piaya cayana*	Tutumbaro	16728, 16940	0	2	F

Order	Family	Scientific name	Site	Tissue Number	No. Positive	Sample	Habitat type
Falconiformes	Accipitridae	*Accipiter striatus*	Ccano	17183	0	1	B
		Buteo leucorrhous	Ccano	17031	0	1	F
		Buteo magnirostris	Tutumbaro	16870, 16898	0	2	F
	Falconidae	*Falco sparverius*	Ccano	17084, 17113	0	2	O
Galliformes	Cracidae	*Chamaepetes goudotii*	Tutumbaro	16733, 16739	0	2	F
		Ortalis guttata	Tutumbaro	16842, 16862	0	2	F
		Penelope montagnii	Ccano	17092, 17119	0	2	F
Gruiformes	Eurypigidae	*Eurypyga helias*	Tutumbaro	16818	0	1	F
Passeriformes	Cardinalidae	*Pheucticus aureoventris*	Tutumbaro	16857	0	1	O
		Saltator aurantiirostris	Ccano	17127	0	1	O
		Saltator maximus	Tutumbaro	16600, 16907	0	2	F
	Cinclidae	*Cinclus leucocephalus*	Ccano	17040	0	1	F
			Tutumbaro	16706, 16718, 16810, 16836	0	4	F
	Conopophagidae	*Conopophaga castaneiceps*	Tutumbaro	16623, 16666, 16884	0	3	F
	Corvidae	*Cyanocorax yncas*	Tutumbaro	16700, 16796, 16871	0	3	O
		Cyanolyca viridicyana	Ccano	16952, 17037	0	2	F
			Tutumbaro	16779, 16800, 16906, 16908, 16915	0	5	F
	Cotingidae	*Ampelion rubrocristatus*	Ccano	16916, 16959, 17077, 17096	0	4	O
		Pipreola arcuata	Ccano	17000, 17014, 17060	0	3	F
		Pipreola intermedia	Ccano	16925	0	1	F
			Tutumbaro	16648	0	1	F
		Pipreola pulchra	Tutumbaro	16882	0	1	F
		Rupicola peruviana	Tutumbaro	16757, 16771	0	2	F
	Emberizidae	*Atlapetes melanopsis*	Ccano	17175, 17194, 17196	0	3	O
		Atlapetes tricolor	Ccano	16970	0	1	O
			Tutumbaro	16629, 16659, 16667, 16669, 16686, 16713, 16846	0	7	O
		Buarremon brunneinucha	Tutumbaro	16603, 16611, 16622, 16697	0	4	F
		Buarremon torquatus	Ccano	16657	0	1	F
			Tutumbaro	16928, 16967, 17005, 17058, 17079	0	5	F
		Zonotrichia capensis	Ccano	16933, 16951, 16978	0	3	O
			Tutumbaro	16776, 16804, 16829, 16854	0	4	O
	Formicariidae	*Grallaria rufula*	Ccano	16918, 16958, 17034, 17161	0	4	F
		Grallaria squamigera	Ccano	17120	0	1	F
		Grallaria sp.	Tutumbaro	16698, 16717	0	2	F
		Grallaricula flavirostris	Tutumbaro	16705, 16748	0	2	F
	Furnariidae	*Anabacerthia striaticollis*	Tutumbaro	16631, 16902	0	2	F
		Asthenes ottonis	Ccano	17176	0	1	O
		Campylorhamphus pucherani	Ccano	**17071**†	1	1	F
		Cranioleuca marcapatae	Ccano	17106, 17135	0	2	F
		Dendrocincla tyrannina	Ccano	16956	0	1	F
			Tutumbaro	16753, 16782	0	2	F
		Lepidocolaptes lacrymiger	Tutumbaro	16767, 16864	0	2	F
		Lochmias nematura	Ccano	16971, 17124	0	2	F
			Tutumbaro	16708, 16922	0	2	F
		Margarornis squamiger	Ccano	17004, 17087, 17090, 17141	0	4	F
		Premnoplex brunnescens	Tutumbaro	16671, 16780	0	2	F
		Premnornis guttuligera	Tutumbaro	16617, 16896	0	2	F
		Pseudocolaptes boissonneautii	Ccano	16989, 17181	0	2	F
			Tutumbaro	16909	0	1	F
		Synallaxis azarae	Tutumbaro	16730, 16784, 16817, 16835, 16844, 16845	0	6	O
		Synallaxis unirufa	Ccano	16927, 16968, 17028, 17029, 17107, 17115	0	6	F
		Syndactyla rufosuperciliata	Tutumbaro	16643, 16649, 16658, 16725	0	4	F

Order	Family	Scientific name	Site	Tissue Number	No. Positive	Sample	Habitat type
		Thripadectes holostictus	Tutumbaro	16606, 16772, 16939	0	3	F
		Thripadectes scrutator	Ccano	17172	0	1	F
		Xenops rutilans	Tutumbaro	16873	0	1	F
		Xiphocolaptes promeropirhynchus	Tutumbaro	16807, 16900	0	2	O
	Hirundinidae	*Pygochelidon cyanoleuca*	Tutumbaro	17136, 17153, 17154, 17165	0	4	O
	Icteridae	*Amblycercus holosericeus*	Ccano	17018, 17025, 17083, 17130, 17143	0	5	F
		Cacicus chrysonotus	Ccano	16976, 16982, 16997, 17023, 17082	0	5	F
		Psarocolius atrovirens	Tutumbaro	16740, 16768	0	2	F
	Incertae sedis	*Chlorospingus ophthalmicus*	Tutumbaro	16769	0	1	F
		Chlorospingus parvirostris	Tutumbaro	16727, 16761, 16762, 16788, 16803, 16843	0	6	F
	Parulidae	*Basileuterus coronatus*	Tutumbaro	16599, 16613, 16618, **16620†**, 16627, 16633, 16651, 16653, 16719, 16826	1	10	F
		Basileuterus luteoviridis	Ccano	***16937, 16954, 16960, 16986, 17011, 17041***	1	6	F
		Basileuterus signatus	Tutumbaro	16610, 16614, 16687, 16923	0	4	F
		Basileuterus tristriatus	Tutumbaro	16615, 16637, 16732, 16760, 16765	0	5	F
		Myioborus melanocephalus	Ccano	17042, 17043, 17051, 17126	0	4	F
			Tutumbaro	16696, 16876, 16893	0	3	F
		Myioborus miniatus	Tutumbaro	16624, 16682, 16690, 16872, 16920	0	5	F
	Pipridae	*Xenopipo unicolor*	Tutumbaro	16866	0	1	F
	Rhinocryptidae	*Scytalopus atratus*	Tutumbaro	16601, 16742, 16808	0	3	O
		Scytalopus parvirostris	Ccano	16943, 16993, 17129, 17159, 17173, 17174	0	6	F
	Thamnophilidae	*Drymophila caudata*	Tutumbaro	16639, 16744	0	2	F
		Thamnophilus caerulescens	Tutumbaro	16683, 16743, 16823, 16892	0	4	B
	Thraupidae	*Anisognathus igniventris*	Ccano	17095, 17114, 17144, 17151	0	4	F
		Anisognathus lacrymosus	Ccano	**17030†**	1	1	B
		Anisognathus sumptuosus	Tutumbaro	16794	0	1	F
		Buthraupis montana	Ccano	16917, 16969, 17128, 17169, 17188	0	5	B
		Chlorornis riefferii	Ccano	17001, 17036, 17097, 17103, 17110	0	5	F
		Conirostrum albifrons	Ccano	17117	0	1	O
		Conirostrum cinereum	Ccano	16926, 17133	0	2	O
		Conirostrum sitticolor	Ccano	17068	0	1	F
		Creurgops verticalis	Tutumbaro	16665, 16851	0	2	F
		Dacnis cayana	Tutumbaro	16914	0	1	B
		Delothaupis castaneoventris	Ccano	16995, 17081	0	2	O
		Diglossa brunneiventris	Ccano	17013, 17086, 17152, 17191	0	4	B
		Diglossa caerulescens	Tutumbaro	16801	0	1	O
		Diglossa cyanea	Ccano	16966, 16981, 17016, 17024, 17080, 17178	0	6	B
			Tutumbaro	16832	0	1	B
		Dubusia taeniata	Ccano	17012, 17170, 17182	0	3	F
			Tutumbaro	16789, 16802, 16809, 16838	0	4	F

Order	Family	Scientific name	Site	Tissue Number	No. Positive	Sample	Habitat type
		Hemispingus atropileus	Ccano	16935, 16936, 16963, 16975, 17111	0	5	F
		Hemispingus frontalis	Tutumbaro	16644, 16654, 16662, 16714, 16729, 16899	0	6	F
		Hemispingus melanotis	Tutumbaro	16604, 16709, 16833	0	3	F
		Hemispingus superciliaris	Ccano	16948, 16965	0	2	F
		Hemispingus xanthophthalmus	Ccano	17059, 17108, 17180	0	3	F
		Pipraeidea melanonota	Tutumbaro	16919	0	1	O
		Ramphocelus carbo	Tutumbaro	16881	0	1	O
		Sporophila luctuosa	Ccano	17112,17158	0	2	O
			Tutumbaro	16912, 16929, 16930, 16931	0	4	O
		Tangara nigroviridis	Tutumbaro	16883, 16901	0	2	F
		Tangara parzudakii	Tutumbaro	16887	0	1	F
		Tangara vassorii	Tutumbaro	16750	0	1	F
		Tangara viridicollis	Tutumbaro	16774, 16786	0	2	F
		Tangara xanthocephala	Tutumbaro	16770, 16793, 16805, 16816, 16875, 16879	0	6	F
		Thlypopsis ruficeps	Ccano	17052, 17179	0	2	B
			Tutumbaro	16684, 16781, 16814, 16840	0	4	B
		Thraupis cyanocephala	Ccano	16932, 17047, 17138	0	3	O
			Tutumbaro	16681, 16755, 16775, 16812, 16841, 16847	0	6	O
		Thraupis episcopus	Tutumbaro	16905, 16910	0	2	O
		Tiaris obscurus	Tutumbaro	16664, 16672, 16837	0	3	O
	Tityridae	*Pachyramphus versicolor*	Ccano	17067, 17091	0	2	B
	Troglodytidae	*Cinnycerthia peruana*	Ccano	**16980**†, 16998, 17026, 17056, 17085, 17099, 17131, 17132	1	8	F
		Henicorhina leucophrys	Ccano	17003	0	1	F
			Tutumbaro	16612, 16621, 16650, 16656, 16726, 16763	0	6	F
		Thryothorus coraya	Tutumbaro	16831, 16858	0	2	F
		Troglodytes aedon	Ccano	17007, 17137, 17146	0	3	O
			Tutumbaro	16660, 16735	0	2	O
		Troglodytes solstitialis	Ccano	17098, 17139, 17156, 17186	0	4	F
	Turdidae	*Catharus fuscater*	Ccano	16938	0	1	F
			Tutumbaro	16723, 16863	0	2	O
		Entomodestes leucotis	Tutumbaro	16605, 16661, 16695, 16702	0	4	F
		Myadestes ralloides	Tutumbaro	16602, 16641, 16655, 16773	0	4	F
		Turdus fuscater	Ccano	17006, 17048, 17057, 17163	0	4	F
			Tutumbaro	16741, 16878	0	2	F
		Turdus nigriceps	Tutumbaro	16797, 16834	0	2	F
		Turdus serranus	Ccano	17027, 17070, 17195	0	3	F
			Tutumbaro	16721, 16759, 16853	0	3	F
	Tyrannidae	*Anairetes parulus*	Ccano	17009, 17072	0	2	F
		Conopias cinchoneti	Tutumbaro	16913	0	1	O
		Contopus fumigatus	Ccano	17065	0	1	B
			Tutumbaro	16874, 16895	0	2	B
		Elaenia albiceps	Ccano	16950, 17122, 17171, 17187	0	4	O
		Elaenia obscura	Tutumbaro	16609, 16663, 16890, 16897	0	4	O
		Elaenia pallatangae	Ccano	17053, 17073, 17116, 17121, 17140, 17193, 17197, 17198	0	8	O
			Tutumbaro	16824	0	1	O
		Hemitriccus granadensis	Ccano	17033	0	1	F

Order	Family	Scientific name	Site	Tissue Number	No. Positive	Sample	Habitat type
		Knipolegus poecilurus	Tutumbaro	16820, 16821, 16822	0	3	F
		Mecocerculus leucophrys	Ccano	16946, 17074	0	2	B
		Mecocerculus poecilocercus	Tutumbaro	16868	0	1	F
		Mecocerculus stictopterus	Ccano	17017, 17094, 17123, 17167	0	4	F
		Mionectes macconnelli	Tutumbaro	16880	0	1	F
		Mionectes olivaceus	Tutumbaro	16941	0	1	F
		Mionectes striaticollis	Ccano	16972, 17010, 17088, 17104	0	4	F
			Tutumbaro	16625, 16635, 16652, 16668, 16699, 16724, 16825	0	7	F
		Myiarchus cephalotes	Tutumbaro	16646	0	1	O
		Myiarchus tuberculifer	Ccano	16973, 16985, 17022, 17039	0	4	O
			Tutumbaro	16747, 16819	0	2	O
		Myiophobus flavicans	Tutumbaro	16634 , 16636, 16903	0	3	F
		Myiotheretes fuscorufus	Ccano	17089	0	1	F
		Ochthoeca cinnamomeiventris	Ccano	16934, 16979, 17021, 17062	0	4	F
			Tutumbaro	16685, 16711, 16731, 16751, 16764	0	5	F
		Ochthoeca frontalis	Ccano	16947, 16992, 17020, 17032, 17061, 17066	0	6	F
		Ochthoeca pulchella	Ccano	16964	0	1	F
			Tutumbaro	16746, 16792	0	2	F
		Ochthoeca rufipectoralis	Ccano	16945, 16990, 17102, 17118, 17160, 17177	0	6	B
		Phylloscartes ophthalmicus	Tutumbaro	16860, 16886	0	2	F
		Pseudotriccus pelzelni	Tutumbaro	16798	0	1	F
		Pseudotriccus ruficeps	Ccano	16944, 16955, 16974, 16991, 17044, 17045	0	6	F
		Pyrrhomyias cinnamomeus	Ccano	17038, 17093	0	2	O
			Tutumbaro	16628, 16716, 16777, 16778, 16783	0	5	O
		Sayornis nigricans	Tutumbaro	16766, 16869, 16889	0	3	O
		Tyrannus melancholichus	Tutumbaro	16608, 16828, 16848, 16849, 16904	0	5	O
Piciformes	Bucconidae	*Malacoptila fulvogularis*	Tutumbaro	16675, 16949	0	2	F
	Capitonidae	*Eubucco versicolor*	Tutumbaro	16638, 16670, 16785, 16806	0	4	F
	Picidae	*Campephilus haematogaster*	Tutumbaro	16703, 16734, 16855	0	3	F
		Piculus rivolii	Ccano	16942, 16987, 17049	0	3	F
		Picumnus dorbignyanus	Tutumbaro	16632, 16647, 16680, 16691	0	4	F
		Veniliornis nigriceps	Ccano	17147, 17148	0	2	F
	Ramphastidae	*Andigena hypoglauca*	Ccano	17015	0	1	F
		Aulacorhynchus coeruleicinctus	Tutumbaro	16607, 16642, 16674, 16694	0	4	F
Psittaciformes	Psittacidae	*Bolborhynchus orbygnesius*	Ccano	17109	0	1	F
		Hapalopsittaca melanotis	Tutumbaro	16795	0	1	O
		Pionus tumultuosus	Tutumbaro	16693, 16756, 16787	0	3	B
Strigiformes	Strigidae	*Glaucidium jardinii*	Ccano	17192	0	1	B
Trogoniformes	Trogonidae	*Pharomachrus auriceps*	Tutumbaro	16852, 16865	0	2	B
		Trogon personatus	Ccano	17046, 17166, 17706	0	3	F
			Tutumbaro	16673, 16701, 16738, 16749, 16811, 16839	0	6	F

Appendix Table 5.3: Associations between positive influenza cases and ecological characteristics, in individuals and species, using Pearsons chi-square test, or Fisher's exact test when sample size was <100 (indicated by *). In some cases an individual or species was classed in more than one category, for instance, in the category of habitat some species were classed both as using forest and open habitat. These results were expressed as a relative odds ratio with 95% confidence intervals. A p value <0.05 was considered to be statistically significant.

Ecological characteristic	*k* (positive individuals)	*N* (Total individuals)	Odds ratio	95% Confidence limits	P-value
Order – individuals	**8**	**600**			
Apodiformes	3	75			
v. all others	5	525	4.333	1.121 – 16.809	0.031
v. negative orders	0	67	Inf	0.707 – Inf	NS
v. Passeriformes	5	458	3.775	0.976 – 14.650	0.055
Other orders	0	67			
v. Apodiformes	3	75	0.000	0.000 - 1.414	0.098
v. all others	8	533	0.000	0.000 – 3.815	NS
v. Passeriformes	5	458	0.000	0.000 – 5.278	NS
Passeriformes	5	458			
v. Apodiformes	3	75	0.265	0.068 – 1.024	0.055
v. all others	3	142	0.511	0.133 – 1.960	NS
v. negative orders	0	67	Inf	0.189 – Inf	NS
Order - species	**8**	**177**			
Apodiformes	3	19			
v. non-Apodiformes	5	158	5.737	1.387 – 24.132	0.012
v. negative orders	0	30	Inf	1.342 – Inf	0.053*
v. Passeriformes	5	128	4.612	1.113 – 19.445	0.033
Negative orders	0	30			
v. Apodiformes	3	19	0.000	0.000 – 0.745	0.053*
v. Apodiformes-Passeriformes	8	147	0.000	0.000 – 2.297	NS
v. Passeriformes	5	128	0.000	0.000 – 3.271	NS
Passeriformes	5	128			
v. Apodiformes	3	19	0.217	0.051 – 0.899	0.033
v. non-Passeriformes	3	49	0.623	0.157 – 2.455	NS
v. negative orders	0	30	Inf	0.306 – Inf	NS

Ecological characteristic	*k* (positive individuals)	*N* (Total individuals)	Odds ratio	95% Confidence limits	P-value
Age	**7**	**600**			
Adult	6	383	3.438	0.538 – 21.848	NS
Immature	1	217	0.291	0.046 – 1.857	NS
Bursa	**7**	**600**			
Absent	7	534	Inf	0.225 – Inf	NS
Present	0	66	0.00	0.00 – 4.446	NS
Diet – individuals	**8**	**593**			
Carnivore	0	39	0.000	0.000 – 6.872	NS
Insectivore	8	539	Inf	0.208 – Inf	NS
(excluding Apodiformes)	5	464	0.458	0.119 – 1.755	NS
Frugivore	1	247	0.197	0.031 – 1.238	0.092
Nectarivore	3	90	3.434	0.891 – 13.275	0.076
Graminivore	0	129	0.000	0.000 – 1.711	NS
Diet – species	**8**	**175**			
Carnivore	0	17	0.000	0.000 – 4.450	NS
Insectivore	8	151	Inf	0.337 – Inf	NS
(excluding Apodiformes)	5	132	0.546	0.132 – 2.075	NS
Frugivore	1	75	0.180	0.022 – 1.492	NS
Nectarivore	3	24	4.171	1.027 – 17.166	0.045
Graminivore	0	39	0.000	0.000 – 1.619	NS
Family – individuals	**8**	**600**			
Trochilidae	3	75	4.333	1.121 – 16.809	0.031
Furnariidae	1	49	1.619	0.256 – 10.388	NS
Parulidae	2	37	5.305	1.186 – 24.025	0.026
Thraupidae	1	110	0.633	0.101 – 4.002	NS
Troglodytidae	1	26	3.240	0.505 – 21.287	NS
Tyrannidae	0	106	0.000	0.000 – 2.223	NS
Family – species	**8**	**177**			
Trochilidae	3	19	5.737	1.387 – 24.132	0.012
Furnariidae	1	18	1.277	0.196 – 8.632	NS
Parulidae	2	6	16.600	2.915 – 100.815	<0.000
Thraupidae	1	34	0.589	0.092 – 3.849	NS
Troglodytidae	1	5	5.893	0.803 – 46.748	0.091
Tyrannidae	0	28	0.000	0.000 – 2.501	NS
Foraging strata – individuals	**8**	**564**			
Ground	2	141	1.000	0.229 – 4.393	NS
Under storey	8	417	Inf	0.743 – Inf	0.091
Mid-canopy	6	336	2.055	0.469 – 8.969	NS
Canopy	2	226	0.494	0.113 – 2.164	NS

Ecological characteristic	*k* (positive individuals)	*N* (Total individuals)	Odds ratio	95% Confidence limits	P-value
Ground	2	40	1.044	0.232 – 4.759	NS
Under storey	8	116	Inf	0.922 – Inf	0.059
Mid-canopy	6	100	2.011	0.446 – 8.961	NS
Canopy	2	73	0.404	0.091 – 1.817	NS
Habitat type - individuals	**8**	**600**			
Open	4	230	1.619	0.439 – 5.969	NS
Forest	5	429	0.660	0.172 – 2.527	NS
Edge	8	391	Inf	1.129 – Inf	0.037
Mixed	1	54	1.453	0.230 – 9.298	NS
Habitat type - species	**8**	**177**			
Open	**4**	**66**	1.726	0.456 – 6.538	NS
Forest	**5**	**131**	0.569	0.143 – 2.244	NS
Edge	**8**	**109**	Inf	1.374 – Inf	0.022
Mixed	**1**	**18**	1.277	0.196 – 8.632	NS
Mixed flock - individuals	**8**	**595**			
No	**4**	**241**	1.477	0.401 – 5.443	NS
Yes	4	354	0.677	0.184 – 2.496	NS
Mixed flock - species	**8**	**175**			
No	4	76	1.319	0.349 – 4.990	NS
Yes	4	99	0.758	0.200 – 2.866	NS
Sex – individuals	**6**	**577**			
Female	3	228	1.538	0.352 – 6.722	NS
Male	3	349	0.650	0.149 – 2.843	NS
Sociability - individuals	**8**	**582**			
Solitary	4	349	0.664	0.180 – 2.447	NS
Social	4	233	1.507	0.409 – 5.554	NS
Sociability - species	**8**	**173**			
Solitary	4	104	0.650	0.172 – 2.461	NS
Social	4	69	1.538	0.406 – 5.827	NS

Appendix Table 5.4: Associations between positive influenza cases and ecological characteristics, at two Andean sites, compared using Pearsons chi-square test, or Fisher's exact test when sample size was <100 (indicated by *). Results were expressed as a relative odds ratio with 95% confidence intervals. A p value <0.05 was considered to be statistically significant.

Ecological characteristic	*k* (positive individuals)	*N* (Total individuals)	Odds ratio	95% Confidence limits	P-value
Diet – individuals					
Ccano	**4**	**262**			
Tutumbaro	**4**	**331**			
Insectivore					
Ccano	4	236	1.289	0.349 – 4.754	NS
Tutumbaro	4	303	0.776	0.210 – 2.862	NS
Frugivore					
Ccano	1	104	inf	0.358 - inf	NS
Tutumbaro	0	143	0.00	0.000 - 2.796	NS
Nectarivory					
Ccano	0	53	0.000	0.000 - 0.859	0.066*
Tutumbaro	3	37	inf	1.164 - inf	0.066*
Diet - species					
Ccano	**4**	**91**			
Tutumbaro	**4**	**115**			
Insectivore					
Ccano	4	77	1.342	0.355 – 5.073	NS
Tutumbaro	4	102	0.745	0.197 - 2.815	NS
Frugivore					
Ccano	1	77	inf	0.181 - inf	NS
Tutumbaro	0	54	0.000	0.000 – 5.515	NS
Nectarivory					
Ccano	0	13	0.000	0.000- 1.363	NS*
Tutumbaro	3	15	inf	0.733 - inf	NS*
Family - individuals					
Ccano	4	263			
Tutumbaro	4	337			
Trochilidae					
Ccano	0	38	0.000	0.000 – 1.206	NS*
Tutumbaro	3	37	inf	0.829 - inf	NS*
Furnariidae					
Ccano	1	20	inf	0.376 - inf	NS*
Tutumbaro	0	29	0.000	0.000 - 2.657	NS*
Parulidae					
Ccano	1	10	2.889	0.274 – 30.654	NS*
Tutumbaro	1	27	0.346	0.033 – 3.656	NS*

Ecological characteristic	*k* (positive individuals)	*N* (Total individuals)	Odds ratio	95% Confidence limits	P-value
Thraupidae					
Ccano	1	51	inf	0.300 - inf	NS
Tutumbaro	0	59	0.000	0.000 - 3.332	NS
Family - species					
Ccano	4	92			
Tutumbaro	4	117			
Trochilidae*					
Ccano	0	8	0.000	0.000 - 1.588	NS*
Tutumbaro	3	11	inf	0.630 - inf	NS*
Furnariidae*					
Ccano	1	10	Inf	0.308 – Inf	NS*
Tutumbaro	0	12	0.000	0.000 – 3.250	NS*
Parulidae*					
Ccano	1	3	1.000	0.054 – 18.522	NS*
Tutumbaro	1	3	1.000	0.054 – 18.522	NS*
Thraupidae					
Ccano	1	17	inf	0.335 - inf	NS*
Tutumbaro	0	22	0.000	0.000 to 2.988	NS*
Feeding strata - individuals					
Ccano	4	257			
Tutumbaro	4	307			
Ground					
Ccano	1	52	1.725	0.176 – 16.876	NS
Tutumbaro	1	89	0.580	0.059 – 5.667	NS
Under storey					
Ccano	4	190	1.199	0.324 – 4.437	NS
Tutumbaro	4	227	0.834	0.225 – 3.087	NS
Mid-canopy					
Ccano	3	157	3.468	0.490 – 24.401	NS
Tutumbaro	1	179	0.288	0.041 – 2.041	NS
Canopy					
Ccano	0	118	0.000	0.000 - 3.523	NS
Tutumbaro	1	108	inf	0.284 – inf	NS
Feeding strata – species					
Ccano	4	89			
Tutumbaro	4	107			
Ground					
Ccano	1	19	1.611	0.158 – 16.398	NS*
Tutumbaro	1	30	0.621	0.061 – 6.312	NS*
Under storey					
Ccano	4	65	1.148	0.301 – 4.380	NS
Tutumbaro	4	74	0.871	0.228 – 3.326	NS
Mid-canopy					
Ccano	3	53	1.890	0.361 – 9.820	NS
Tutumbaro	2	65	0.529	0.102 – 2.766	NS
Canopy					
Ccano	0	40	0.000	0.000 – 4.452	NS*
Tutumbaro	1	46	inf	0.220 - inf	NS*

Ecological characteristic	*k* (positive individuals)	*N* (Total individuals)	Odds ratio	95% Confidence limits	P-value
Habitat - individuals					
Ccano	4	263			
Tutumbaro	4	337			
Open					
Ccano	1	103	0.405	0.057 – 2.887	NS
Tutumbaro	3	127	2.468	0.346 – 17.434	NS
Forest					
Ccano	4	192	5.021	0.746 – 33.620	NS
Tutumbaro	1	237	0.199	0.030 – 1.341	NS
Edge					
Ccano	4	185	1.116	0.301 – 4.134	NS
Tutumbaro	4	206	0.896	0.242 – 3.319	NS
Mixed					
Ccano	1	32	inf	0.176 – inf	NS*
Tutumbaro	0	22	0.000	0.000 to 5.684	NS*
Habitat - species					
Ccano	4	92			
Tutumbaro	4	117			
Open*					
Ccano	1	35	0.412	0.057 – 3.055	NS*
Tutumbaro	3	45	2.429	0.327 – 17.577	NS*
Forest					
Ccano	4	69	5.046	0.773 – 34.209	NS
Tutumbaro	1	83	0.198	0.029 – 1.364	NS
Edge					
Ccano	4	64	1.067	0.279 – 4.081	NS
Tutumbaro	4	68	0.938	0.245 – 3.587	NS
Mixed					
Ccano	1	12	inf	0.188 - inf	NS*
Tutumbaro	0	9	0.000	0.000 - 5.530	NS*
Mixed flock - individual					
Ccano	4	263			
Tutumbaro	4	337			
No					
Ccano	1	105	0.426	0.060 – 3.035	NS
Tutumbaro	3	136	2.346	0.330 – 16.565	NS
Yes					
Ccano	3	158	3.774	0.534 – 26.549	NS
Tutumbaro	1	196	0.265	0.038 – 1.874	NS
Mixed flock - species					
Ccano	4	92			
Tutumbaro	4	117			
No					
Ccano	1	40	0.402	0.056 – 2.959	NS*
Tutumbaro	3	50	2.489	0.249 – 24.895	NS*
Yes					
Ccano	3	52	3.918	0.395 – 38.831	NS
Tutumbaro	1	65	0.255	0.026 – 2.529	NS

Ecological characteristic	*k* (positive individuals)	*N* (Total individuals)	Odds ratio	95% Confidence limits	P-value
Apodiformes					
Ccano	0	38	0.000	0.000 to 1.206	NS*
Tutumbaro	3	37	inf	0.829 - inf	NS*
Passeriformes					
Ccano	4	204	0.506	0.752 – 33.865	NS
Tutumbaro	1	254	0.198	0.030 – 1.329	NS
Sociability – individual					
Ccano	4	231			
Tutumbaro	4	331			
Solitary					
Ccano	1	154	0.427	0.044 – 4.146	NS
Tutumbaro	3	199	2.342	0.331 – 16.474	NS
Social					
Ccano	3	139	2.824	0.398 – 19.911	NS
Tutumbaro	1	129	0.354	0.050 – 2.516	NS
Sociability – species					
Ccano	4	90			
Tutumbaro	4	113			
Solitary					
Ccano	1	56	0.424	0.059 – 3.074	NS
Tutumbaro	3	73	2.357	0.325 – 16.820	NS
Social					
Ccano	3	47	4. 364	0.598 – 31.267	NS
Tutumbaro	1	65	0.229	0.032 – 1.671	NS

CONCLUSION

This dissertation presents five chapters that evaluate spatial and taxonomic distributions of AI, primarily of strain H5N1, as well as other avian viruses. Each chapter stands alone as a single manuscript, that is already published or that has been submitted for publication. Nevertheless, together they construct a broad view of occurrence and patterns of viruses that affect birds, and specifically address whether it is possible to predict the ecological niche of H5N1, how similar ecological niche requirements are in paired H5N1 host groups, and the assumption that AI prevalence is low or nil in land birds.

In Chapter 1 I demonstrate that H5N1 cases were found to occur under predictable environmental conditions, suggesting that elements of the transmission cycle have some form of ecological determination, here measured as differences in land-surface reflectance and plant phenology through the year. Generally models predicted in Chapter 2 concurred with Chapter 1; H5N1 cases were found to occur under predictable sets of environmental conditions: absent from areas with low NDVI values and minimal seasonal variation, and present in areas with a broad range of and appreciable seasonal variation in NDVI values. However, case occurrences in the Arabian Peninsula appeared to occur in a distinct environmental regime, suggesting that there is variable environmental "fingerprint" area suitable for H5N1 transmission.

In Chapter 3 I explore the similarity of ecological niche requirements in paired H5N1 host groups, and found significant niche similarity (13/24) or no significant differences (9/24) in almost all tests. Although 2 of 24 analyses found significant differences in niche, these

analyses were contradicted by others, suggesting an overall signal of niche similarity among groups. I thus could not document distinct ecological niches for H5N1 occurrences in different host groups, and conclude that transmission cycles are broadly interwoven.

Results of survey work in Ghana (Chapter 4) and Peru (Chapter 5) testing the assumption that AI prevalence is low or nil in land birds were equivocal, with prevalences of 0% and 1.3%, respectively. Apparently AI prevalences vary spatially. Nonetheless, the prevalence in Peru is not insignificant, and shows that surveillance programs for monitoring spread and identification of AI viruses should thus not focus solely on water birds. In Peru, AI infections were at relatively high prevalences in birds of the orders Apodiformes (hummingbirds; 4%) and Passeriformes (songbirds; 1.1%). Additionally testing in Ghana for other viruses (Alphaviruses and Flaviviruses) yielded one sequence of a Yaoundé-like *Flavivirus*, constituting only the second known avian host record of this virus.