ECOLOGY AND EVOLUTION OF AVIAN AND SWINE INFLUENZA VIRUSES IN GUATEMALA

by

LUCIA ORTIZ BATSCHE

(Under the Direction of Daniel R. Perez)

ABSTRACT

Wild aquatic birds are the major reservoir for influenza A viruses (FLUAV) and play a key role in viral maintenance and dissemination. Likewise, swine have a vital role in the adaptation, and reassortment of FLUAVs that may result in zoonotic strains of pandemic concern. Despite increased surveillance efforts, the diversity and ecology of FLUAVs circulating in wild birds and swine in Latin America remain understudied. For instance, the 2009 H1N1 influenza pandemic strain emerged in Mexico, but its circulation remained undetected in pigs before its detection in humans.

The studies presented here aimed to better define the ecology and evolution of FLUAVs in wild aquatic birds and swine, using samples from an understudied region of the world: Guatemala, in Central America. I used targeted sequencing on the FLUAV genome from wild birds and swine swabs, and bioinformatics tools to analyze full genome sequencing data generated by Next-Generation Sequencing technologies to characterize the potential genetic traits of these viruses to cross the interspecies barrier. The first part of this thesis comprises the results of analyzing six migratory seasons (2013-2019) of avian FLUAV surveillance efforts in Guatemala, a stopover site for

multiple overlapping flyways. Overall, I found great FLUAV diversity, in terms of subtype combinations and a high frequency of detection of viruses that are rarely found in breeding grounds in North America. My results provide additional support to previous observations that Guatemala can serve as a major geographic bottleneck with a significant role in the persistence and evolution of FLUAV in the region. The second part of this thesis describes the results of two years of FLUAV active surveillance in a commercial swine farm in Guatemala without a history of vaccination. We found a unique group of H1N1 pdm09 sequences that suggests independent evolution from similar viruses circulating in Central America. These viruses may represent the establishment of a novel genetic lineage with the potential to reassort with other co-circulating viruses, and whose zoonotic risk remains to be determined. My thesis has contributed to a better understanding of FLUAV ecology in Central America.

INDEX WORDS: Avian influenza, Swine influenza, Surveillance, Next-generation sequencing, Central America, great FLUAV diversity, Viral Evolution, Phylogenetics

ECOLOGY AND EVOLUTION OF AVIAN AND SWINE INFLUENZA VIRUSES IN GUATEMALA

by

LUCIA ORTIZ BATSCHE

B.S., Universidad del Valle de Guatemala, Guatemala, 2007

M.Sc., Staffordshire University, UK, 2013

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2022

ECOLOGY AND EVOLUTION OF AVIAN AND SWINE INFLUENZA VIRUSES IN GUATEMALA

by

LUCIA ORTIZ BATSCHE

Major Professor:	Daniel R. Perez
Committee:	Daniela Rajao
	Claudio Afonso
	Justin Bahl

Electronic Version Approved:

Ron Walcott
Vice Provost for Graduate Education and Dean of the Graduate School
The University of Georgia
August 2022

DEDICATION

This dissertation is dedicated to my family, especially to my sister who always will live in my heart.

ACKNOWLEDGEMENTS

I would like to thank my advisor, Dr. Daniel R. Perez for the opportunity to join his laboratory and pursue a scientific career. Thanks to him and his wife, Andrea, for all your advice and for having my back during difficult and joyful times. I would also like to thank all members of my dissertation committee Dr. Daniela Rajao, Dr. Claudio Afonso, and Dr. Justin Bahl for their valuable time, recommendations, and support.

I would like to express my gratitude to all my colleagues from Universidad del Valle de Guatemala, Maria Luisa, Ana Silvia, David, Dione, Jorge, and Silvia for their expertise, hard work, and friendship during this project, without them this work would not be possible. Thanks to Celia Cordón-Rosales for allowing me to be part of Centro de Estudios en Salud, which was a pivotal part of my personal and professional growth. I would also like to thank all my friends and colleagues in Guatemala, Juan Carlos, Mabel, Willy, Renata, Calu, Maru, and my former boss Dr Norma Padilla for their friendship, valuable support, and for encouraging me to pursue an academic career outside our home country. Thanks to all the former and present members of the Perez-Rajao Lab, Stivalis, Lucas, Zhimin, Silvia, Jojo, John, Brittany, Joaquin, Ginger, Sheroz, Claire, Flavio, Matias, Brianna, Pilar, and Ruby for all the constant support, advice and friendship.

Special thanks to my family, my father Sergio, my mother Maria Eugenia, my siblings Esteban and Anahí, my sister-in-law Lupita, my husband Gabriel, and my daughter Anahí Cristina for their immense love and for being the motivation in my life.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION

Influenza A virus (FLUAV) infects humans and a wide range of birds and mammals (1). The virus genome is negative single-stranded RNA, composed of 8 different segments: 6 internal gene segments (PB2, PB1, PA, NP, M, and NS) and 2 surface gene segments (HA and NA). FLUAVs are divided into subtypes based on the antigenic properties of the surface glycoproteins. Up to date, 18 HA (H1–H18) and 11 NA (N1–N11) subtypes have been described (2-4). Wild aquatic birds are the major reservoir for FLUAV diversity, maintenance, and dissemination (1). FLUAVs have been found in aquatic birds in many HA and NA subtype combinations (1), except for H17N10 and H18N11, only found in fruit bats (3, 4). FLUAVs have been isolated from approximately 100 species of wild birds from 12 of the 50 orders, including the aquatic orders *Anseriformes* and *Charadriiformes* (5). H5 and H7 subtypes are of particular significance because of their potential to cause outbreaks associated with high mortality in economically and agriculturally relevant poultry species (6). Eurasian origin highly pathogenic avian influenza viruses (HPAIV) of the H5 subtype are a major threat to the poultry industry worldwide, as highlighted by the destruction of many millions of birds in multiple countries (7), most recently related to the 2022 epizootics in Europe and the U.S. (United States). Human cases of HPAIV infection have also been reported. Although human-to-human transmission is limited, the fatality rate among humans may be over 50% (7).

The surface proteins of FLUAVs are under enormous selective pressure, and thus have high mutation rates that allow these viruses to escape the host immune system. These mutations allow FLUAVs to eventually cross species barriers, which may result in outbreaks of varying intensity in immunologically naïve populations, for instance, in poultry. Past influenza pandemics have resulted from zoonotic FLUAV strains that cross from animals to humans, although the mechanisms to cross such barriers remain unclear [5]. The most notorious example of animal FLUAV spillover to humans occurred during the H1N1 or "Spanish influenza" in 1918-1919, which many of its gene segments had ancestors of avian origin that adapted to humans (8). The 1957 pandemic was caused by a H2N2 subtype virus that contained novel HA, NA and PB1 gene segments of an avian-origin virus from the previous seasonal lineage (9). The pandemic strain from 1968 was a reassortant of the H2N2 virus that acquired novel avian-origin H3 and PB1 gene segments, followed by a re-emergence of the human 1957 H1N1 virus in 1977. Lastly, the most recent 2009 pandemic virus originated from multiple reassortment events from viruses that had been circulating in pigs for many years [8]. During this pandemic, a triple reasorted virus originated from an avian/human/swine reassortant swine virus reassorted with a Eurasian avian-like swine virus, resulting in the H1N1 pdm09 strain (10). The triple reassortant virus rise during the late 1990s, containing gene constellation of human (HA, NA, PB1), swine (NS, NP, and M) and avian (PB2 and PA) origin and became endemic among swine populations in North America (11). Multiple viruses, including H1N1, H1N2 and H3N2 have resulted from the co-circulation of H1N1 and H3N2 viruses in the US and Canada, leading to the establishment of new H1 lineages

(termed H1α, H1β, Hδ, Hγ) and similarly with H3 viruses (classified in four genetic clusters, I – IV) (11-14).

Swine may have a critical role as they are recognized as intermediary hosts for the adaptation and reassortment of FLUAVs that may result in zoonotic or pandemic strains [14]. Swine FLUAV impacts swine production due to losses caused by the disease and can result in zoonotic infections, posing a threat to public health. In swine farms, close contact with infected animals susceptible to influenza may also increase the risk for FLUAV transmission. Studies have demonstrated that swine farm workers develop antibodies against influenza virus strains that typically circulate in pigs [15, 16]. Bi-directional interspecies transmission events of FLUAVs between humans and pigs play a significant role in the generation of novel reassortant strains that may spread among humans and/or swine populations. Recent studies suggest that introduction of human viruses into pigs may be a major driver in the evolution of FLUAV lineages in Latin America [17]. Hence, the emergence of the 2009 FLUAV pandemics in Mexico highlights the importance to strengthen animal influenza surveillance in understudied regions (15).

Research objective and aims

One of the main purposes of surveillance programs is to understand the diversity of circulating viruses in a specific population, which can help with the early recognition of threats to human and animal health (16). With new sequencing technologies, it is relatively easy and possible to obtain whole virus genome sequence data for detailed

identification of molecular signatures of circulating FLUAV. Given the great diversity of FLUAV subtypes, their accurate identification is crucial to providing clues on their potential origin, phylogenetic relationships with other circulating viruses, and the establishment of epidemiological links during outbreaks. Decades of intensive surveillance of FLUAVs in wild birds have shown the importance of wild bird movements in spreading genetically diverse viruses between locations and facilitating the emergence of new strains, including those with the capacity to infect poultry, pigs, and humans, among other animal species. Despite increased surveillance efforts in the past decade or so, the diversity and ecology of viruses circulating in wild birds in Latin America remain understudied. Central America is located at an intermediate point between bird migratory routes and the majority of the more than one hundred aquatic bird species found in this area are migratory birds, including some of the main FLUAV reservoirs also found in North America [12]. Previous studies show that Central America´s unique location may provide a site where hundreds of different migratory bird populations converge yearly and thus, increasing the likelihood of interspecies virus transmission [13]. Published viral sequences from Guatemala together with newly generated sequencing data generated during this study, represent the most abundant genetic data set available from Central America. This dataset provides clues that better define the ecology of FLUAV in Central America and how this geographical area may play a role in the evolution of avian FLUAVs in natural reservoirs. Chapter 4 focuses on better understanding the contribution to virus diversity and the exchange of gene segments of FLUAVs from different continents, as well as defining in more detail the genetic structure of the FLUAV population in wild birds in Guatemala. My thesis work

provides a more comprehensive picture of the spread of FLUAV in an important region of Central America. My studies reveal unique phylogenetic patterns and unusual subtypes in Guatemala compared to other parts of the world. The data is useful in the context of pandemic preparedness, as well as its significance for animal health. In Chapter 5, I analyzed in more detail data related to the rarely found H14 subtype. My work led to the realization that Guatemala has since reported on the largest collection of H14 strains in the world. Interestingly, the H14 viruses are reassortants between Eurasian and North American lineage viruses and the question remains whether these viruses will remain endemic in the wild bird population in Guatemala.

Recent studies suggest that the introduction of human FLUAVs into swine is a major driver in the evolution of FLUAV lineages that are exclusive to Latin America [17]. In Central America, Guatemala is the country with the largest swine industry. Circulation of FLUAV of human origin has been documented previously in the swine population in Guatemala [18]; however, the contribution of humans to the molecular epidemiology of FLUAV in the swine population has not been investigated in detail. Chapter 6 focuses on the characterization of swine FLUAV from a commercial farm. The obtained viruses represent a unique group of sequences, suggesting independent evolution from other FLUAVs circulating in Central America.

The long-term goal of this research was to better define the ecology and evolution of FLUAVs in two animal populations of major importance in the evolution and interspecies transmission of FLUAVs: wild birds and swine. We used bioinformatics tools to analyze full genome sequencing data generated by next-generation sequencing

technologies (NGS) to characterize the potential genetic traits of these viruses to cross the interspecies barrier. The main research objectives are described below.

Specific aim 1: *To characterize the origin and genetic composition by full genome sequencing of FLUAVs circulating in wild birds in Guatemala, collected from 2013 to 2019.*

i. Identify circulating genotypes and gene constellations of sequenced viruses

ii. Establish phylogenetic relationships of sequenced viruses

iii. Perform time scale phylogenetic analyses of Guatemalan H14 including all available sequences of the H14 subtype from North America and Europe

iv. Identify potential events of inter-lineage reassortment (including the South American lineage)

Specific aim 2: *To establish the subtype diversity and analyze the year-to-year variation of FLUAVs circulating in wild aquatic birds during the winter migration in Guatemala.*

i. Expand analysis of temporal variations in FLUAV prevalence to determine temporal trends

Specific aim 3: *To characterize the origin and genetic composition by full genome sequencing of FLUAVs circulating in swine on a commercial farm in Guatemala (in absence of vaccination).*

ii. Better define the seasonal patterns of FLUAV infection in swine farms

iii. Establish phylogenetic relationships of sequenced viruses

iv. Identify amino acid variants within sequenced viruses

v. Identify molecular markers associated with anti-viral resistance

CHAPTER 2

LITERATURE REVIEW

Classification

FLUAV is one of 7 genera in the family *Orthomyxoviridae*: *Alphainfluenzavirus* (also referred to as type A influenza virus), *Betainfluenzavirus* (type B), *Gammainfluenzavirus* (type C), *Deltainfluenzavirus* (type D), *Isavirus, Thogotovirus* and *Quaranjavirus* (17). These viruses are characterized by a genome consisting of six to eight single-stranded RNA segments. FLUAVs are classified into subtypes by the antigenic properties of the surface glycoproteins hemagglutinin (HA) and neuraminidase (NA). Up to date, 18 HA and 11 NA subtypes have been described. With exception of H17N10 and H18N11, found in bats (3, 4), all 16 HA and 9 NA subtypes have been found in wild aquatic wild birds, which are considered primary hosts for FLUAVs (18).

Virion structure and organization

Virions are pleomorphic particles with a mixture of spherical (~100 nm in diameter) and filamentous (up to 300 nm in length) forms, protecting a segmented, negative, single stranded RNA genome of ~13.6kb. The virions consist of a host-derived lipid bilayer envelope with two main glycoproteins, HA and NA, protruding from the viral surface, in a ratio of four to one (19). And additional surface protein is also present, the ion channel protein (M2), in a ratio of one M2 channel per 101-102 HA molecules (20). Underneath the viral bilayer, the M1 protein forms a matrix holding the nuclear export protein (NEP)

and the viral ribonucleoproteins (vRNPs). Each vRNA is looped and wrapped around multiple copies the NP molecules (nucleoprotein) in a rod-like structure (21). The end of each vRNP is bounded by the viral polymerase proteins, composed of two "polymerase basic" and one "polymerase acidic" subunits (PB1, PB2, and PA) located at the end of the nucleocapsids (18). The vRNA have 15 nucleotides at the 3′ end and 16 nucleotides at the 5′ end that are complementary between the 5' and 3' ends. These conserved vRNA ends form a "corkscrew" RNA structure that interacts with the viral polymerase (22), which is required for transcription, replication, and packaging of the genome (23, 24).

The virus genome is composed of 8 segments numbered in order of decreasing length (Table 1.1). Segments 4, 5, and 6 encode a single protein, HA, NP, and NA, respectively. Segments 1, 2, 3, 7, and 8 encode at least two proteins produced from different open reading frames (ORF) by multiple mechanisms, described below.

The HA is encoded in segment 4 which varies in length with 1728-1778 nucleotides, depending on the subtype. HA is an integral membrane glycoprotein responsible for binding of virions to host cell sialic acid (SA) receptors, and mediating fusion between the viral and host cell membranes (further details of these receptors are presented in the following sections). HA is a homotrimer, composed of three identical subunits associated by noncovalent interactions. Each monomer is produced as a single polypeptide from a single ORF, and consist of two subunits HA1 and HA2, attached by a disulfide bond. These two subunits are cleaved by host-produced trypsin-like proteases from its precursor (HA0), a step required for viral infectivity. The HA molecule is composed of a globular head (HA1 subunits) that contains the receptor-binding site (RBS), as well as antigenic sites, and a stalk (HA2 subunits) (18).

Segment 6 also varies in length with 1453-1470 nucleotides, and codifies the NA protein, the second major surface antigen of FLUAVs. NA is an integral membrane glycoprotein responsible for cleavage the surface sialic-acid receptors to release new viral particles after budding from the cell (25), allowing viral particles to escape from the cell.

Segments 1 and 2 are longest segments, consisting of 2,341 nucleotides each. Segment 1 encodes the PB2 protein, but additionally recent studies showed that a novel protein is also encoded: PB2-S1, corresponding to nucleotides 1513 to 1894 of the PB2 mRNA. The function of this accessory protein is not yet well understood, but it might be implicated in a role for PB2 within the mitochondrial matrix (26, 27). Segment 2 encodes for the PB1 protein, but some strains may encode an 87 amino acid accessory protein (PB1-F2) in a +1 alternate reading frame (28). Segment 3 consists of 2,233 nucleotides and codifies for both PA and the fusion protein PA-X through ribosomal frameshifting (29). Segment 5 consists of 1565 nucleotides and codifies for the NP (nucleoprotein) that has a primary structural role, but also interacts with cellular polypeptides, including actin, components of the nuclear import and export apparatus and a nuclear RNA helicase (30). Segment 7 codifies for the M1 matrix protein, along with the M2 ion channel via mRNA splicing (31). Also, through mRNA splicing, segment 8 codifies the NS1 and NEP proteins (29). At the 3′ and 5′ ends, all segments are flanked by segment-specific untranslated regions (UTR) that contain unique “packing signals” and play a role in transcription, polyadenylation, and segment replication (32, 33).

Table 1.1 Summary of FLUAV segments and protein functions[1]

Segment	Protein	Size (nucleotides)	Size (nucleotides)	Function
1	PB2	2341	759	Cap-binding protein
2	PB1	2341	757	Replicase/Transcriptase
	PB1-F2		87	Pro-apoptotic activity
3	PA	2233	716	Endonuclease
	PA-X			Virulence Factor/Host cell protein shut off
4	HA	1728-1778	539-550	Receptor binding Membrane fusion
5	NP	1565	498	RNA binding protein
6	NA	1453-1470	454-468	Receptor cleavage
7	M1	1027	252	Capsid protein
	M2		97	Ion-channel, Unidirectional proton pump
8	NS1	890	230	Interferon antagonist
	NEP		121	viral RNP export from the host cell nucleus

[1]Modified from Bouvier and Palese, 2007 (19).

Replication cycle

Virus attachment: FLUAV binds to terminal sialic acid or SA moieties in glycoprotein or glycolipid receptors on the host cell surface. The most common SA is N-acetylneuraminic acid (Neu5Ac), which is nine-carbon acidic monosaccharides (34). The carbon-2 of the terminal SA can bind to the carbon-3 or carbon-6 of galactose, forming α-2,3- or α2,6-linkages. The HA spikes of FLUAVs can recognize either α-2,3 SA or α-2,6 SA conformations. In mammals, α-2,6 SA are predominant in tracheal epithelial cells, but also α-2,3 SA are present in less abundance the in human respiratory epithelium. In contrast, α-2,3 SA are predominant in the gut epithelium and respiratory track from avian species (35). Distribution of α-2,3 and α-2,6 SA in different host is an important determinant for FLUAV host range, as described in the next section. The HA protein is

present on the virus' surface as trimers of identical subunits. The HA is responsible for the viral entry and fusion to the host cell. The HA recognizes terminal SA on glycoproteins and glycolipids on the cell surface that serve as anchors from which the virus is internalized via endocytosis.

Virus entry: Following attachment of the influenza virus HA protein to sialic acid, the virus is endocytosed. Acidification by the fusion of the endocytic vesicle and a lysosome, leads to structural changes on the HA trimers that expose the fusion peptides via a spring-loaded mechanism resulting in the fusion of the viral and endosomal membranes. Endosome acidification activates also the M2 ion channel protein pump hydrogen ions from the endosome into the virus particle, disrupting protein-protein interactions, and thus, causing the matrix proteins to detach from the vRNPs (36). During virus replication, the HA protein is cleaved by host proteases forming the HA1 and HA2 subunits. The HA2 portion is thought to mediate the fusion of virus envelope with cell membranes. The pores produced during HA fusion allow the vRNPs to flow freely to the cell cytoplasm, completing the uncoating process (36, 37). The vRNPs are taken up by karyopherin proteins via protein-protein interactions (proteins in the P-complex and NP contain nuclear localization signals, NLS) and transported to the nucleus, the site of FLUAV transcription and replication.

Transcription and replication of the viral genome: The vRNPs serve as templates for two +sense RNAs: viral mRNAs and complementary. FLUAV mRNAs are capped and polyadenylated, whereas cRNAs represent full-length copies of the vRNAs. The FLUAV RNA-dependent RNA polymerase (RdRp, composed of the PB2, PB1, and PA subunits) primes viral mRNA synthesis via a cap-snatching mechanism utilizing capped

5’ ends (~12 nts) derived from host mRNAs. Polyadenylation of the viral mRNAs occurs due to a stuttering mechanism, whereby the RdRp moves back and forth over this stretch of U residues, leading to the formation of a poly(A) tail (38). Additional co- and post-transcription modifications occur (for instance, frameshift to generate PA-X mRNA or splicing in mRNAs from segments 7 and 8 to M2 and NEP mRNAs, respectively). The viral mRNAs are transported to the cytoplasm for translation into the corresponding viral protein products. Some of the early viral protein products include those with the ability to interferes with host mRNA translation (NS1, PA-X) and/or antagonize the host’s innate immune mechanisms (NS1). The HA, NA and M2 mRNAs are translated by ribosomes associated to the endoplasmic reticulum (ER) and their products transported by the secretory pathway. All other mRNAs are translated by ribosomes in the cytoplasm. Newly synthesized PB2, PB1, PA, and NP are imported into the nucleus for unprimed/uncapped/non-polyadenylated cRNA synthesis. Later in the infection cycle, newly synthesized M1 and NEP migrate to the nucleus, bind to progeny vRNPs, and together migrate back to the cytoplasm for virion assembly.

Assembly and release. The HA, NA and M2 envelope proteins are transported to the plasma membrane via the Golgi apparatus. Signals at the 5’ and 3’ ends of the each vRNA spanning non-coding and coding regions dictate packaging into virions (39). The vRNP/M1/NEP complexes are transported cytoplasm where additional copies of M1 provide the milieu for virion assembly through network of protein-protein interactions that involve M1, vRNPs and the cytoplasmic tails of HA and NA. Assembly is completed at this location by budding from the plasma membrane. The release of discrete virus particles from viral aggregates from the host cell is mediated by the NA protein via

cleavage of sialic acid residues from HA trimers and other glycostructures (40). Once these particles reach the lumen of extracellular compartments, the HA protein undergoes proteolytic cleavage by host proteases from the HA0 precursor into the covalently linked subunits HA1 and HA2, allowing the virus to start a new replication cycle (41).

Determinants of host range

Although the host range of FLUAVs is considered a polygenic trait, the specificity and affinity of HA for sialic acid (SA) receptors binding through the receptor binding site (RBS) are key determinants. RBS The head of the HA contains the sialic acid receptor-binding site as described previously, which is surrounded by the five predicted antigenic sites (Sa, Sb, Ca1, Ca2, and Cb in the H1 subtype) (19). The RBS forms a shallow pocket at the distal tip of the HA1 head creating three important secondary elements (130-loop, 190-helix, and 220-loop) with four highly conserved residues (Y98, W153, H183 and Y195) that contribute to receptor binding via van der Waals interactions. Avian α2,3 SA receptors interact with the 220-loop of HA, whereas human α2,6-linked SA receptors interact with the 190-helix (42).

The HA protein recognizes terminal sialic acid (SA) moieties present on glycoproteins and glycolipids on the surface of host cell membranes. In broad terms, the HA of FLUAVs recognize SAs in two conformations depending on its linkage (described in the previous sections) to the adjacent galactose: α-2,3 SA and α-2,6 SA (43). Although perhaps an oversimplification on the intricacies that modulate SA recognition by FLAUV, it is consistent with the relative expression of terminal SAs and tissue tropism of these viruses in various hosts. In wild aquatic birds, FLUAV replicates predominantly

in epithelial cells of the intestinal tract which are rich in α-2,3 SA. In contrast, in humans FLUAVs replicate in epithelial cells of the respiratory tract which mostly express α-2,6 SA receptors. Thus, FLUAVs from wild birds preferentially recognize α-2,3 SA, whereas those from humans prefer α-2,6 SA receptors. It is commonly accepted that a transition from α-2,3 SA to α-2,6 SA recognition is needed for FLUAVs to spread among humans. Intermediary and/or accidental hosts with alternative expression patterns of SA receptors are thought to provide an environment for FLUAVs to mutate and thus change receptor preference. Humans themselves have low levels of α-2,3 SA expression in the lower respiratory tract that could provide such environment. However, pigs have been long regarded as a potential mixing vessel for the emergence of novel FLUAV strains with α-2,6 SA due to their susceptibility to FLUAVs of avian and human origin consistent with the relatively high expression of both α-2,3 SA and α-2,6 SA in their respiratory tract (44-46). In ferrets, α2-6 SAs are also prevalent in the respiratory tract (47), by contrast to non-human primates, mice and dogs show higher expression of α2-3 SA (48-50).

Mutations on the HA have a direct impact on the FLUAV host range. For instance, the HA of the H1 subtype of avian origin usually presents the E190/G225, E190/D225, or D190/G225 combinations in the RBS, whereas H1 strains of human origin have transitioned to D190/D225 or D190/E225 combinations (51, 52). These amino acid transitions have recently in strains that went to almost strict α-2,3 SA binding to almost strict α-2,6 binding. Likewise, mutations in the HA of H3 and H9 viruses are critical for receptor specificity, including those present in position 226 and 228. Avian viruses usually present mutations Q226/G228 that confers dual-binding or α2,3-SA preference, in contrast to Q226L/G228S that confers α2,6-SA preference (50). The NA plays the

opposite function of the HA. While the HA attaches to the SA receptors, the NA removes the SA residues to release the viral progeny. Therefore, a coordination of the HA-NA specificity seems to be an important factor for viral adaptation in other hosts (53).

Besides the HA and NA, mutations in other gene segments shown to be a critical determinant of the host range. Mutations in PB2 may impact the replication in different host cells and at different temperatures. A change in the position of E627K in PB2 has resulted in increased virus replication and transmission in mammals (54). Additionally, M gene segment appears to play an important role for enhancing viral transmission as observed in the 2009 pandemic virus, where the M gene segment was a critical factor in the spread of the virus in humans (55).

FLUAV evolution

FLUAVs have high evolutionary rates, estimated at ∼10−3 nucleotide substitutions per site per year (56). However, each gene segment evolves differently depending on specific selective pressures and evolutionary constraints. FLUAVs display remarkable diversity due to two major mechanisms that drive influenza evolution: antigenic drift and antigenic shift. Antigenic drift arises when mutations are accumulated over time particularly in the HA, due to copying errors during viral replication caused by the lack of viral RNA polymerase proofreading activity. Such mutations alter the viral antigenic features and thus, the host immune system cannot longer neutralize it. Antigenic drift can also be observed on the NA with a slower mutation rate. Additionally, mutations can also occur on the internal genes having an impact particularly in the viral replication. Antigenic drift is a slow virus evolution event. For example, H3N2 FLUAV

subtypes that circulate seasonally in humans evolve from their precursor rising new antigenic strains every 3-5 years, whereas seasonal H1N1 viruses every 3-8 years (57). Emerged virus variants can arise as a specific clade, as a consequence of sporadic outbreaks (58-60) that eventually may replace the original strain through time (60).

Another major evolutionary mechanism that promotes natural FLUAV diversity is the exchange of gene segment of reassortment. Due to the segmented nature of the FLUAV genome, gene segments are exchanged in the new viral progeny during mixed infections, and thus new variants can arise with alternative HA and NA combinations. Reassortment involves rapid changes in the FLUAV genome. Antigenic shift occurs when reassortment rises new viruses with novel HA subtypes that are introduced and spread into a naïve population. These viruses are characterized by major phenotypic jumps from their ancestors arising as diverging lineages that proliferate and persist over time (61). Reassortment events between animal-origin and human-origin FLUAVs have played a major role in the past FLUAV pandemics. During the “Spanish flu” in 1918, an H1N1 that contain gene segments of avian-origin virus was transmitted to the naïve human population (8). The 1957 pandemic was caused by a H2N2 subtype virus that contained novel HA, NA and PB1 gene segments of an avian-origin virus from the previous seasonal lineage (9). The pandemic strain from 1968 was a reassortant of the H2N2 virus that acquired novel avian-origin H3 and PB1 gene segments, followed by a re-emergence of the human 1957 H1N1 virus in 1977. Lastly, the most recent 2009 pandemic virus originated from multiple reassortment events from viruses that had been circulating in pigs for many years [8], as detailed in the following sections.

Geographic separation of host species in nature is a major driver of FLUAV evolution. Certain subtypes of FLUAV predominate in wild ducks, determined by the geographical separation and migratory patterns of the bird species, resulting in the independent evolution of lineages. Phylogenetic studies show a clear separation between the North American and Eurasian lineages (62), but other independent lineages from Australia, Antarctica and more recently from South America have been also documented (63-65).

FLUAV in avian species

In birds, FLUAVs have been isolated from about 100 species of wild birds from 12 of the 50 orders, mainly from the orders *Anseriformes* (ducks, geese, and swans) and C*haradriiformes* (shorebirds, auks, gulls, among others) (5) reaching FLUAV prevalence levels >20% in the autumn migration season (66). These avian species are widely distributed, and many species can perform long migrations across different locations, allowing the movement and spread of FLUAV from distant regions around the world (67, 68). Aquatic wild birds show the greatest diversity of FLUAV, therefore they are recognized as the main natural virus reservoirs (1). They play an important role for the virus perpetuation in nature, since from the 18 HA (H1–H18) and 11 NA (N1–N11) described subtypes, H1-H16 and N1-N9 have been isolated from wild birds (2-4). In theory, all 144 HA-NA subtype combinations are possible, but some subtype combinations have been never reported in nature.

Replication of FLUAVs in aquatic wild birds and shorebirds occurs primarily in the intestinal tract, and sometimes in the respiratory tract (69-72). The main route of

transmission is fecal-oral transmission since FLUAV is secreted in high concentration in the feces, but the movement of animals also plays an important role in its dispersal in nature (71-73). Transmission by fecal material in the water is a common way of FLUAV dispersal to other waterfowl and other domestic avian species. Rates of FLUAV in wild birds in the western hemisphere is higher in August to September, where juvenile birds are infected as they congregate prior to migration, likely due to immunologically naïve population of juvenile birds (68). During migration, they continue to shed virus, but the positivity FLUAV rates decrease, possible to increase of natural population immunity (66). During migration, most of the bird species make multiple stops before arriving to their wintering grounds. This behavior may allow the congregation of birds from different species and flyways, and thus increasing the chance of FLUAV transmission and reassortment, originating new viruses than can be spread as birds continue its migration (74).

FLUAV have been sporadically isolated from shorebirds (C*haradriiformes*) differing in the HA and NA subtypes found in ducks. H4, H9, H11, H13 and H16 are predominant subtypes found in shorebirds and gulls (75, 76), in contrast to H3 and H6 found in ducks (77), suggesting a different pool of viruses among these two orders.

Even though several FLUAV subtypes have been isolated from wild birds, only H5 and H7 have become economically and agriculturally important due to their potential to become highly pathogenic when infecting poultry (6); however viral infections caused by any virus subtype may be aggravated by secondary infections or environmental conditions. FLUAV can cause asymptomatic infection in avian species, but chickens and

turkeys are the most susceptible species to present mild to fatal diseases varying considerably with the strain of the virus.

In poultry, FLUAV can be classified as low pathogenicity (LPAI) and high pathogenicity (HPAI) avian influenza viruses based on the severity of the disease. HPAI viruses originate from LPAI avirulent precursors, but HPAI are associated to the insertion of a polybasic amino acid sequence, typically a Lysine or Arginine at the cleavage site of H5 and H7 viruses (78-80). This insertion allows different ubiquitously-expressed proteases in the intracellular and extracellular environments to cleave HA protein into the HA1 and HA2 subunits, expanding the range of tissue that the virus can replicate, and being able to cause fatal infection in poultry (81). Signs of HPAI in poultry can include sudden death, as well as lack of energy, purple discoloration or swelling, diarrhea, nasal discharge, coughing, sneezing, and reduced egg production.

HPAI viruses may cause outbreaks with high mortality rates ranging from 75% to 100% (82, 83), and therefore causing important economic loses in the poultry industry. In areas where there is high contact between domestic influenza hosts with wild birds may be a risk factor for virus introduction in domestic hosts as LPAI present in migratory birds may become virulent through the passage in poultry (6).

HPAI viruses are not only a threat the poultry industry, but they are also representing an increasing threat to humans. Since 2004, the HPAI H5N1 virus has spread to many countries worldwide and has been responsible for the destruction of many millions of birds. To date, 18 countries worldwide have reported outbreaks of avian H5N1 in poultry (84) and from 2003 to January 2020, 861 confirmed human cases of infection with avian

influenza H5N1 virus from 17 countries with a 53% case-fatality rate were reported to WHO (85).

FLUAV in Blue-winged Teals

Although mallards (*Anas platyrhynchos*) and northern shovelers (*Anas clypeata*) are two of the duck species with the highest prevalence of FLUAV in North America (86-89), Blue-winged teals (BWTE) represent an important FLUAV host species, particularly in the wintering grounds. The BWTE may be a potential key host for virus dispersal given due to its long-distance migration (68, 74).

The Blue-winged Teal (BWTE, *Anas discors*), belong to the Order *Anseriformes* (family *Anatidae*, along with other species of ducks, geese, and swans). It is one of the most abundant species of dabbling ducks that breed at high densities in the northcentral states of the US and provinces of central Canada. It has a widespread distribution during winter from several southern states of US (including California, Texas and Louisiana) throughout the northern Neotropics (Mexico, Central America, the Caribbean, and northern South America) (90, 91). It is one of the first waterfowl species to migrate southward in fall, and one of the last to migrate northward in spring. Males begin fall migration earlier than females and immatures (92). The long-distance behavior (68, 93, 94), and the diversity of viruses found in BWTE, suggest that this bird species could be a candidate for targeted FLUAV surveillance in the Neotropics.

BWTE has been recognized as an important FLUAV reservoir (95), harboring a high diversity of virus subtypes at the wintering grounds, particularly in areas in the Mississippi and Central migration flyways (77, 89, 96-98). Studies at the wintering

grounds in North America and Europe showed that FLUAV prevalence in BWTE and waterfowl species tend to be lower in the wintering grounds, compared to those observed at the breeding grounds. Prevalence of FLUAV in BWTE is driven by host density and population immunity (77). Early migration suggests that BWTE migrates before periods of high FLUAV prevalence (89). Susceptible birds resulting from early migration could get infected later in the season during migration (89). A decreasing trend of the FLUAV prevalence toward the end of each migratory season has been observed in wintering grounds, including Guatemala, California and Texas (96, 97, 99, 100), likely to the accumulation of natural population immunity (66), as mentioned in the previous section.

Studies performed at breeding and wintering grounds show a distinctive seasonal pattern in the relative abundance of FLUAV subtypes. According to publicly available data from the Influenza Research Database (IRD), up to June 2022, 1,156 HA sequences have been published from BWTE from 1974 to 2020. These sequences are reported mainly from North America (US=840, Canada=183 and Mexico=3) and Central America (Guatemala=126), Caribbean (Barbados=1) and other regions (Hong Kong=1, Singapore=1 and Egypt=1). For the breeding grounds (northern states of the US and Canada) 636 (55%) records have been published, and 516 (45%) for the wintering grounds (California, Florida, Texas, Louisiana, Mexico, Guatemala, and Barbados). In the breeding grounds, the most common subtypes were H3 and H4, as well as N6 and N8 during the summer/fall season and H7 during the spring season, with around 6% of mixed infections, consistent with previous observations (77). Information of wintering grounds is limited and largely bias to regions from North America, where H3, H4, H7, N6 and N8 are the most common subtypes, also consistent with previous reports (77). Guatemala is

the only country in Central America consistently reporting sequences from BWTE, describing a different scenario of circulating subtypes that includes H14, H3, H5, N3 and N2 (99, 101). According to IRD data, in the wintering grounds, the percentage of mixed infections reported in BWTE is around 12%.

FLUAV in wild birds in Guatemala

In Guatemala, BWTE are present in a relative high abundance, particularly on the Pacific coast of Guatemala, with reports over counts >10,000 individuals, whereas other FLUAV host species (Northern Pintails, American Wigeons, Northern Shovelers, Green-winged Teals and Cinnamon Teals) are less frequently reported (counts < 250) (102). Central America´s unique location provides a site where hundreds of different migratory bird populations converge yearly, increasing the likelihood of interspecies virus transmission, and thus have the potential to contribute to virus diversity and the exchange of gene lineages from different continents (103). The majority of aquatic bird species found in Guatemala are migratory birds, including some of the main FLUAV reservoirs also found in North America (102).

In Guatemala, prevalence data of FLUAV is not available for poultry or other animal species, although the presence of H5N2 LPAI was first confirmed in poultry in 2000 (104). Since 2006, cloacal and tracheal swabs from aquatic birds have been collected to detect FLUAV and identify the subtypes circulating at different sites on the Pacific and Atlantic coasts of Guatemala (105). Multiple FLUAV subtypes and high genetic diversity in migratory and resident birds in Guatemala were detected, including subtypes that are

rarely detected globally like H14 (105). FLUAV prevalence detected by RRT-PCR ranged from 5% to 38%; with a decreasing trend in monthly prevalence within a migration season (99). Previous phylogenetic analysis indicates that these viruses are highly similar to those from North America (101, 105, 106). However, a depth phylogenetic analysis of FLUAV that circulates in Guatemala during 2010 to 2013 revealed a Eurasian virus PA segment introduced into the Americas in the early 2000s disseminated to Guatemala from ~2007 to 2010, most likely via the pacific flyway (103). Based on phylogenetic analyses of the many different FLUAV subtypes obtained in the country, Guatemala (and perhaps other areas of Central America) may behave as a bottleneck where FLUAVs from the Atlantic/Central flyways can mix with those from the Pacific flyway (103). Although some viruses detected in Guatemala were closely related to viruses detected in South America, viruses from the South American lineage have not yet been detected in Guatemala (103).

FLUAV in swine

Swine were considered as "mixing vessels" for influenza transmission due to its susceptibility with both human-origin and avian-origin FLUAV, and their potential for generating novel progeny viruses capable of replicating and spreading among humans (107, 108). The main explanation of the "mixing vessel" theory is that although FLUAV replication is specie- restricted, both avian and human viruses replicate well in swine due to the presence and distribution of α-2,3 and α-2,6 SA virus receptors in the respiratory mucosa of swine, as described before. These characteristics provide a favorable environment for reassortment of viruses with different receptor specificities during co-

infections. However, in recent years, new evidence came to light where the role of swine in the emergence of strains with pandemic potential is not that critical. The distribution of α -2,3 SA and α-2,6 SA in swine is like those of the human respiratory tract (46, 47, 109), suggesting that both species (swine and humans) are equally prone to avian-origin virus infections. Despite the questionable role of swine as true "mixing vessels", they are an important host for FLUAV as demonstrated in the H1N1 pdm2009, where the pandemic precursor virus circulated in the swine population for more than 10 years before it was detected in humans(110).

Multiple human and avian-origin viruses introduction events into swine population have been recorded worldwide (111), resulting in independent FLUAV evolution from a diverse pool of co-circulating viruses (112). These viruses are characterized by major phenotypic jumps from their ancestors arising as diverging lineages that proliferate and persist over time (61). In North America, in the 90´s, a H1N1 pdm2009 precursor virus denominated "classical swine H1N1 (csH1N1)" was epidemiological stable until late 1990s, when reassorted with a human seasonal H3N2 containing avian like segments, arising the "H3N2 triple reassortment virus or TRIG". This novel H3N2 contained segments from swine-origin (NP, M, and NS), human-origin (NP, M, and NS) and avian-origin (PB2 and PA) viruses (113), became established in the swine population and further evolve into four H3 clusters (cluster I to cluster IV) (112). The H3N2 TRIG virus resorted again with the classical swine H1N1, generating novel H1N1 and H1N2 viruses, which continued evolving to H1-α, H1-β, H1-γ clades, and further to H1-γ1 and H1-γ2 clades (114). Another reassortment event occurred between the H3N2 TRIG virus and H1 and N2 human seasonal viruses, generating two additional clades (H1-δ1 and H1-δ2 viral

lineages) (115). Finally in 2009, the H1N1pdm09 emerged from the H1 triple reassortment subtype (circulating in North America) and a H1N1 avian-like virus (circulating in Europe) (110) and became a divergent lineage from its ancestors.

In 1979 in Europe, an avian-origin H1N1 (termed European “avian-like” H1N1 or H1avN1) caused outbreaks in Germany and Belgium, and rapidly spread and became predominant in Europe (116, 117). During the 1980's in Europe, a human-origin H3N2 virus descendent from the 1968 pandemic virus was introduced and after its reassortment with the H1avN1 became established widespread throughout Europe and remains endemic to date (116, 118). Another reassortment event occurred in the mid-1990s, when H3N2 viruses reassorted with a human-seasonal H1N1 virus, generating a H1N2 (or “H1huN2”) virus lineage (119). These three major lineages co-circulated in Europe until the H1N1 pdm09 strain became established, and the H1N1 2009 pdm internal cassette became the dominant backbone in the region (120).

In Asia, human-like viruses of the H3N2 subtype have been documented in pigs since the 1970s (121). Due to pig importations from other regions, the European H1avN1 and H3N2, and North American TRIG lineage viruses were introduced. After, 2010, the H1N1 pdm09 strain became established and reassorted with endemic strains leading to a diverse pool of viruses (122-124). Less information is available for South America due to the limited surveillance in animals(125). However, genetically and antigenically different human derived H1N1, H1N2, and H3N2 lineages, as well as the H1N1 pdm09 have been reported in Chile, Argentina and Brazil (125-129).

Multiple introductions of human-origin viruses into swine, as well as documentation of reverse zoonotic transmission in different geographical regions are recognized to be

important in the generation of novel reassortants strains with the potential to spread not only in humans but also in swine populations (130-132), and thus, underscoring the importance of understanding how FLUAV evolves in these hosts.

Swine and FLUAV interspecies transmission

Zoonotic infections with swine-origin FLUAV are reported sporadically (131). These infections are usually associated with exposure to infected swine and do not usually result in the person-to-person spread. However, recent studies revealed that as a consequence of the increased FLUAV surveillance efforts in swine after the 2009 pandemic, FLUAV reverse zoonosis (transmission from humans to swine) is more frequent than swine to human zoonosis (133). The introduction of human viruses into swine affects the ecology of FLUAVs circulating in swine globally and the genetic diversity of swine FLUAV has been increased (133, 134). Moreover, studies suggest that the introduction of human viruses into swine affects the ecology of FLUAVs circulating in swine globally (134), maybe be a major driver in the evolution of FLUAV lineages that are exclusive to Latin America (111).

Some environments with high animal density, provide an environment for close contact between animals and humans. Such environment allows the virus to acquire mutations that help overcome natural barriers in the new host, and thus, facilitate the initial spillover between species and transmission. Interspecies transmission events of FLUAVs between humans and pigs play a significant role in the generation of novel reassortant strains that transmit among humans and/or swine populations.

In swine commercial farms, contact with animals susceptible to influenza may also increase the risk for influenza transmission since studies have demonstrated that individuals who work with swine (farmers, meat processing workers, and veterinarians) develop antibodies against influenza virus strains that typically circulate in pigs (135, 136), and are at increased risk of swine FLUAV infection (137). Most of the world's pork is produced in modern swine production systems with tens of thousands of susceptible animals concentrated in a small area, with intense contact between animals and swine workers. Thus, swine production farms provide a unique system to study how FLUAVs may jump and adapt at the human-animal interface. Interspecies transmission of FLUAVs between humans and swine plays a significant role in the generation of novel viral strains that may become established in the new host population. In addition, prolonged direct animal contact may result in bi-directional transmission of different influenza strains, circulating in both human and swine. Increased occupational exposure to swine, as well as a high degree of connection between migrating waterfowl, domestic, and swine highlight the important role of swine as a key species for virus adaptative evolution of FLUAV interspecies transmission (138). This occupational exposure to swine influenza viruses, may represent a potential source of introduction of new reassortant viruses, increasing the viral gene pool circulating in both populations. As an example, the introduction of the H1N1 pdm09 virus into humans in 2009 (110) has led to the replacement of the previously circulating seasonal H1N1 virus and changed the epidemiology of FLUAV circulating in humans.

Swine FLUAV in Guatemala

In Central America, Guatemala is the country with the largest swine industry. Since 2009, increased efforts have been made to establish the prevalence of FLUAVs in the swine population in the country, identifying the circulation of the classical swine H1N1, swine H3N2, as well as viruses closely related to the human seasonal H3N2 and human H1N1 pdm09(139). The reported serological prevalence in 2010 for the classical swine H1N1 and swine H3N2 FLUAV subtypes were 20% and 7% in commercial swine farms and 41% and 17% for backyard swine productions, respectively (139). Additionally, further studies performed by our group showed that viruses isolated from Guatemalan swine are closely related to human seasonal strains (H3N2) and the H1N1 pdm2009, which indicates transmission from humans to swine. Positivity rates in swine for FLUAV by RRT-PCR ranged from 12% to 16% in 2010 and 2011 respectively (140). Interestingly, positive farms tended to cluster in areas where FLUAVs have been detected from wild birds, bats, or humans (3, 105, 141), underscoring the potential role of swine in the interspecies transmission of FLUAV.

FLUAV in other hosts in Guatemala

Studies conducted on the Pacific coast of Guatemala have found in fruit bats (*Sturnira lilium*) hosting a novel subtype (H17N10) of FLUAV (3). Distinguished HA-like (HL) and NA-like (NL) proteins, as well as phylogenetic analysis show that this novel subtype, as well as H18N11 found in bats from Peru (4), are likely an ancient reservoir for a diverse group of influenza viruses, and thus a potentially important source of genetic variation (142). Seroprevalence studies indicate widespread circulation of

these two and perhaps additional bat FLUAV subtypes in several bat species (143). Interestingly, the bat H17N10 was discovered in nearby regions where FLUAV has been detected in other hosts, including humans, swine and wild birds (103, 105, 140).

Next-generation sequencing technologies in FLUAV surveillance

Next-generation sequencing (NGS) technologies allow massive parallel sequencing of both DNA and RNA in a relative short period of time. The main characteristic is that hundreds or thousands of targets can be analyzed at the same time, avoiding the need of multiple assays. In the last decade, NGS techniques have been incorporated to FLUAV genomic surveillance. Traditional surveillance is based on collection, analysis, and interpretation of data. Usually, the information provided consists in epidemiological data related to a certain pathogen or disease (144). Genomic surveillance is the generation of large amount of sequencing data shared among institutions (by open access sequence repositories), in order to establish the transmission routes of outbreaks and track mutations/variants of concern, as they arise at national and international levels (145). The most notorious example of the use of genomic surveillance data for FLUAV is *Flunet*, a global platform for monitoring FLUAV movement and interpreting the epidemiological data, critical for identifying FLUAV vaccines yearly (146).

New NGS technology have been used to improve the genomic characterization of traditional surveillance. Traditionally, after a real time PCR to screen FLAUV positive samples obtained during surveillance, all FLUAV-positive samples are subject to virus isolation and obtained isolated is later sequenced. However, during studies performed in

Guatemala using this approach less than 30% of IAV-positive samples could be isolated and later sequenced (99, 101, 140). New NGS-based approaches allow sequence FLAUV-positive samples directly from swabs, avoiding the need of virus isolation and increasing the number of sequenced samples (101). By combining the two approaches (first sequencing FLUAV-positive swabs and isolate and sequence only those that yielded incomplete genomes) the number of samples with complete genomes around two times greater, and additionally results are obtained in a shorter period of time (101).

CHAPTER 3

MATERIAL AND METHODS

Tracheal and cloacal swabs from wild birds

Tracheal and cloacal swabs were collected from hunter-harvested ducks during the winter migration season 2013-2019 on the Pacific coast of Guatemala. Samples were preserved in 1 mL of virus transport medium (VTM, Medium 199 with Hanks balanced salt solution, 2mM l-glutamine, 0.5% bovine serum albumin, 0.35g/L sodium bicarbonate, 2×106 IU/L penicillin, 200 mg/L streptomycin, 2×106 IU/L polymyxin B, 250 mg/L gentamycin, 0.5×106 IU/L nystatin, 60 mg/L ofloxacin, and 0.2g/L sulfamethoxazole) as previously described (105) and transported to the laboratory on ice and stored at −70°C until used. Permits for sampling bird species at sampling sites were obtained from the Center for Conservation Studies (CECON) and the National Council of Protected Areas (CONAP).

Nasal swabs from swine

Nasal swabs from fattening swine and sows with respiratory signs were collected from May 2016 to April 2018 on a swine farm located at Palin, Escuintla, Guatemala. No vaccination was performed prior and during the study. Pigs were identified and sampled twice per week by farmworkers when animals presented with signs of respiratory disease, including coughing, sneezing nasal and/or ocular discharges. A minimum of 32 nasal swabs were collected monthly to detect at least 1 positive sample based on a prevalence of 10% or higher, with a 95% confidence and in a group size n ≥

600 (147). Swine were tattooed with a unique identifier to avoid resampling. Rectal temperature, sex, respiratory signs, and pen´s density of each animal was recorded at the time of sampling. Nasal swabs were collected and preserved in 3 mL of VTM with antibiotics and antimycotics as described above (140). Animal sampling was conducted under approved animal use protocols from the Ministry of Agriculture, Livestock, and Food of Guatemala (MAGA), and the protocols were reviewed and approved by the Institutional Animal Use and Care Committee of the University del Valle de Guatemala.

RNA extraction

Viral RNA from tracheal swabs was extracted from 200 µL of supernatant using the QIAamp viral RNA kit (QIAGEN, Valencia, CA) according to the manufacturer's instructions and eluted in 60 µL of elution buffer. Viral RNA from cloacal swabs was extracted from 250 µL of supernatant with Trizol LS reagent (Invitrogen, Carlsbad, CA) using an organic extraction protocol described previously (148). Extracted RNA was resuspended in 100 µL of DEPC treated water. Viral RNA from nasal swabs was extracted with the MagMAX-96 AI/ND Viral RNA Isolation Kit (Ambion, Austin, TX) according to the manufacturer's instructions and eluted in a final volume of 50 µL. All extracted RNA was stored at −70°C until use.

FLUAV detection by RRT-PCR

Swabs were screened in duplicates for FLUAV by real-time reverse-transcriptase polymerase chain reaction (RRT-PCR) targeting the matrix gene, as described previously (140, 149, 150). Briefly, QuantiTect Probe RT-PCR (reverse transcription polymerase

chain reaction) Kit (QIAGEN, Hilden, Germany) was used to perform RRT-PCR reactions in the ABI 7300 Real-Time PCR System (Applied Biosystems, Foster City, CA). Each reaction contained 12.5 µL of kit-supplied 2X RT-PCR Master mix, 10 pmol of each primer, 0.3 µM probe, 0.25 µL of kit-supplied enzyme mix, 6.5U RNase inhibitor and 8 µL of RNA template. Thermal cycling conditions were as follows: one cycle of reverse transcription at 50ºC for 30 min and 94ºC for 15 min, followed by 45 cycles of denaturation at 94ºC for 1s and combined annealing and extension at 60ºC for 27s. Fluorescence signals were obtained at the end of each cycle after the annealing/extension step. Primer and probe sequences are described in Table 3.1. Swabs showing Ct values <40 were considered positives.

Multi-Segment amplification (MS-RTPCR)

Multi-Segment amplification of FLUAV genes (MS-RTPCR) was performed in all FLUAV-positive swabs regardless of whether they were of swine or avian origin. MS-RTPCR was performed as described previously (24) with minor modifications. Briefly, 2.5 uL of extracted RNA were used as a template in a 25 uL MS-RTPCR reaction (Superscript III high-fidelity RT-PCR kit, ThermoFisher, Waltham, MA); using Opti1-F1 (0.06 uM), Opti1-F2 (0.14 uM) and Opti1-R1 (0.2 uM) primers (101). The cycling conditions were as: 55°C for 2 minutes, 42°C for 1 hour, 5 cycles (94°C for 30 seconds, 44°C for 30 seconds, 68°C for 3 minutes), followed by 31 cycles (94°C for 30 seconds, 57°C for 30 seconds, 68°C for 3 minutes). Final extension at 68°C for 10 minutes. MS-RTPCR final product was analyzed in 1% agarose gel to corroborate whole genome amplification. Swabs from avian samples that tested negative for MS-RTPCR or

produced incomplete genomes during NGS were later isolated (see next section) and amplified again by MS-RT-PCR. (15, 24, 101).

Virus isolation from wild bird samples

Virus isolation was attempted in triplicate from samples that tested positive by RRT-PCR FLUAV. From each positive bird sample, a pool of the tracheal/cloacal swab sample (200 µL) was of inoculated into the allantoic cavity of 9-day-old Specific-Pathogen-Free (SPF) embryonated chicken eggs following the protocol described in the World Health Organization's Manual on Animal Influenza Diagnosis and Surveillance (151). Collected allantoic fluids were tested by hemagglutination assay to assess for the presence of FLUAV. Allantoic fluids that tested negative were diluted 1:2 with sterile phosphate-buffered saline (pH 7.4) with 1X Antibiotic and Antimycotic solution and inoculated in a new batch of three SPF eggs. The process was repeated three times as needed. Samples were considered negative for the presence of viable virus if no FLUAV growth was detected after three serial passages. FLUAV-positive allantoic fluid was aliquoted and stored at −70°C until later use.

Sequencing and genome assembly

MS-RTPCR products were sequenced using the Illumina platform as described previously (101) with minor modifications. Briefly, amplicons from MS-RTPCR reactions were cleaned by 0.45X of Agencourt AMPure XP Magnetic Beads (Beckman Coulter, Brea, CA) according to the manufacturer's protocol and eluted in 30uL of HyClone molecular biology water (Genesee Scientific, San Diego, CA). Amplicons were

quantified using Qubit buffer kit (ThermoFisher, Waltham, MA) in Qubit 3.0 fluorometer (ThermoFisher, Waltham, MA) and normalized to 0.2 ng/ul. Adaptors were added by tagmentation using the Nextera XT DNA library preparation kit (Illumina, San Diego Ca). The reaction was set as 60% of the suggested final volume. Samples were purified using 0.7X of Agencourt AMPure XP Magnetic Beads and analyzed on a Bioanalyzer using High Sensitivity DNA kit (Agilent, Santa Clara, CA) to determine the distribution of fragment size. Libraries were pooled and normalized to 1-4nM. After denaturation, the final loading concentration of the pooled libraries was 14pM. Libraries were sequenced using MiSeq Reagent Kit V2 300 cycles (Illumina, San Diego, CA). Genome assembly was performed using a customized pipeline developed at the Icahn School of Medicine at Mount Sinai Genome assembly was performed using a pipeline previously described (15). Briefly, adapters were removed by Cutadapt (152) and reads were mapped against reference sequences in the NCBI's Influenza Virus Resource at GenBank, and de novo assembly was performed using Trinity (153). Final consensus sequences were generated using the Burrows-Wheelereler alignment tool (BWA) (154) to map reads to contigs.

Minor variants detection

Variant calling was performed as described previously (155) using LoFreq 2.1.3.1 (156) following the Genome Analysis Toolkit best practices (157). The cutoff for minor variant analysis was arbitrarily set at 1,000× based on a similar analysis found in the literature (158, 159) with a minimum depth of coverage of 400 and a central base quality score of Q30 or higher.

Statistical analysis

For wild birds, the prevalence of FLUAV was calculated as the total number of positive samples detected by RRT-PCR by the total number of collected samples of the same species. All analyses were performed using GraphPad Prism v.9.1.0. Differences in prevalence among seasons were analyzed with $\chi 2$ assuming a two-sided alternative hypothesis was assumed with p-values <0.05 considered significant. To assess viral subtypes, the Simpson's diversity index and their respective confidence intervals within each season were calculated based on the number of different subtype combinations as previously described (160, 161). Seasons with the highest viral diversity based on the Simpson diversity index were compared using the Sørensen-Dice coefficient (162, 163).

For swine, the percentages of positive pigs detected by RRT-PCR were calculated by week, as the total number of positive detected by RRT-PCR by the total number of collected samples per week. To analyze the risk of IAV infection, odds ratios (ORs) were estimated using Stata 15.1 and GraphPad Prism v.9.1.0

Phylogenetic analysis

Independent phylogenetic analyses were performed for the surface genes (HA and NA) and internal genes (PB2, PB1, PA, NS, NP, and MP) of the sequenced viruses. The Additional FLUAV genome sequences from human- and swine-origin H1N1 pdm09-like viruses (from 2009 to 2019) from the Americas and avian-origin were downloaded from the Global Initiative on Sharing All Influenza Data (GISAID, http://platform.gisaid.org) and the Influenza Research Database (IRD, http://www.fludb.org). Sequences were aligned with MUSCLE 5.1 (164) and manually trimmed to keep the open reading frame

(ORF) of the gene of interest. Background sequences were subsampled to remove identical sequences and those sequences with less than 80% of the total length of the ORF using SeqKit 0.16.1 bioinformatic tool. Representative sequences were randomly selected by region and host. Best hit blasts were included in the final tree. The best-fit model of nucleotide substitution was determined for each gene using the Bayesian information criterion (BIC) obtained using jModelTest 2.1.10 (165). For the internal gene segments, phylogenetic trees were constructed using maximum-likelihood (ML) inference methods, using RAxML 8.2.12 (166) under the general time reversible GTR+G+I nucleotide substitution model with 1,000 bootstrap replicates. Trees were run at least twice to confirm topology. Surface gene phylogenies were inferred by time-scaled phylogenetic analysis using the Bayesian Markov chain Monte Carlo (MCMC) approach as implemented in BEAST 1.10.4. The final model of each gene was chosen based on the marginal likelihood estimators obtained in BEAST from several tested models of each dataset. Three independent analyses of 50 million generations were performed to ensure convergence, sampling every 1,000 states. The burn-in percentage of each dataset was identified using Tracer v1.7.1 and after its removal, results were combined using LogCombiner v1.10.4. The MCC trees were annotated using TreeAnnotator v1.10.4. Outlier sequences were identified using TemEst V1.5.3. When necessary, outlier sequences were removed, and the datasets were run again as described above. Nucleotide alignment, background sequences subsampling, BIC, and phylogenetic analyses (RaxML and Beast) were performed using the computational resources Sapelo2 available at The University of Georgia. The resulting trees were visualized in Figtree 1.4.4 and aesthetically modified using Inkscape v0.48.1 (https://inkscape.org). Gene lineages and

genotype assignments were performed based on the phylogenetic trees as described previously. Briefly, FLUAV genes with ≥96% nucleotide identity was classified in the same phylogenetic clade, whereas with ≥90% nucleotide identity within the same linage. Gene constellations of genes were defined as unique genotypes at the level of clades (167).

Amino acid sequence analysis

Nucleotide sequences were translated using SeqKit 0.16.1 bioinformatic tool as described in the previous section and visualized on Geneious Prime 2021.2.2. The resulting amino acid sequences were compared to prototypic reference sequences to calculate the frequency of phenotypic markers for mammalian transmission/pathogenicity and identify specific motifs in the sequence.

A complete list of reagents/supplies used for this study can be found at the Appendix.

Table 3.1 List of primer/probes used for this research work

Primer Name	Purification	Sequence (5´- 3´)	Purpose
M+25	STD	AGATGAGTCTTCTAACCGAGGTCG	RRT-PCR[1]
M-124	STD	TGCAAAAACATCTTCAAGTCTCTG	RRT-PCR[1]
M+64	HPLC	56-FAM/TCA GGC CCC CTC AAA GCC GA/36-TAMSp	RRT-PCR[1]
Opti-F1	HPLC	GTTACGCGCCAGTAGAAACAAGG	MS-RTPCR[2]
Opti-F2	HPLC	GTTACGCGCCAGCGAAAGCAGG	MS-RTPCR[2]
Opti-R2	HPLC	GTTACGCGCCAGTAGAAACAAGG	MS-RTPCR[2]

[1]Real-time Reverse Transcription–polymerase chain reaction; [2] Multi-Segment Reverse Transcription Polymerase Chain Reaction

CHAPTER 4

PERPETUATION OF INFLUENZA A VIRUS IN BLUE-WINGED TEALS IN GUATEMALA

Lucia Ortiz; Ginger Geiger, Lucas Ferreri, David Moran, Dione Méndez, Carlos Serrano, Ana Silvia Gonzalez-Reiche, Mayra Motta, Francisco Escobar, Daniela Rajao, Celia Cordon-Rosales, and Daniel R. Perez.

To be submitted to *Journal of Virology.*

Abstract

Influenza A viruses (FLUAVs) are highly diverse and such diversity is perpetuated in wild aquatic bird species, considered the natural hosts. Most of what is known about the ecology and transmission of FLUAV is due to long-standing surveillance efforts in North America and Eurasia. Our understanding of FLUAV ecology in natural hosts in other parts of the world, including Latin America, is still limited. Prior studies showed that Central America provides a site for the congregation and interaction of aquatic birds from different flyways, serving as hotspot for FLUAV. Gaps remain in the understanding of genetic information flow among FLUAV populations and whether there is seasonality of FLUAV transmission at the wintering grounds in the region. We conducted active FLUAV surveillance over 6 consecutive migrating seasons from 2013 to 2019. This effort resulted in the collection of 3,109 paired cloacal and tracheal samples from hunter-harvested aquatic birds. The overall percentage of RRT-PCR FLUAV-positive samples was 10.0% (95% CI: 10.00% – 11.11%). Most of the samples were collected from *Anas discors* (blue-winged teal) that showed a decreasing trend for RRT-PCR FLUAV-positivity at the end of each migration season. FLUAVs representing 12 HA and 8 NA subtypes in 26 different combinations were identified. The typically elusive H14 and other uncommon subtypes such as H12 and H8, typically uncommon for other parts of the world, were among the most common subtypes found in Guatemala. None of the subtype combinations were consistently found in all seasons; but subtype combinations H3N2, H8N4, and H11N3 were identified in at least four seasons.

Phylogenetic analyses revealed repeated introductions of Eurasia lineage viruses. Low pathogenicity avian influenza H7 subtype viruses identified in Guatemala were

related to sequences identified in H7 highly pathogenic avian influenza outbreaks in the US and Mexico in 2017. By comparing the phylogenetic topologies of each gene segment from 2007-2019, 127 unique genotypes were identified, and 59 gene constellations were found out of 205 full genome sequences. The more diverse genes among Guatemalan FLUAVs were PB2, PA and NP gene segments based on assignments to diversity of clades. In contrast, M1 and NS1 gene segments presented the lowest number of clades. However, we could trace the circulation of two M1 and NS1 gene segments from 2008-2009 within the Pacific flyway, suggesting their persistence for at least a decade. Combined with our previous studies, we provide the largest data set of FLUAV genomes from wild birds in Central America to date.

Introduction

Influenza A viruses (FLUAVs) belong to the family *Orthomyxoviridae*. FLUAV genomes are composed of 8 segments of negative single-stranded RNA corresponding to 6 internal gene segments (PB2, PB1, PA, NP, M, and NS) and 2 surface gene segments (HA and NA). FLUAVs are diverse, divided into subtypes, and widely distributed in wild avian species (1). Subtype classification is based on the antigenic properties of the surface glycoproteins. Up to date, 18 HA (H1–H18) and 11 NA (N1–N11) subtypes have been described, from which H1-H16 and N1-N9 have been isolated from wild birds (2-4). At least 90 species of wild birds from 12 of the 50 orders, from aquatic orders *Anseriformes* and *Charadriiformes* have been identified as harboring FLUAV (5). These species play a key role in FLUAV introduction, and maintenance (1), and occasionally can be a source of FLUAVs that cross over to poultry and mammals, causing outbreaks associated from mild to severe disease (83, 168, 169). In poultry, FLUAVs are classified as low pathogenicity avian influenza viruses (LPAIV) or high pathogenicity avian influenza viruses (HPAIV) (6). HPAIVs outbreaks are typically associated with high mortality in poultry. Typical HPAIV outbreaks are dealt with massive depopulation campaigns that threaten food security and incur in significant economic losses. For instance, the ongoing 2022 H5N1 epizootic of Eurasian origin H5N1 in the US has been detected in at least 35 states with over 40M affected birds states and whose total cost it is still to be determined (170). In China, the H7N9 LPAIV/HPAIV outbreak in 2013 led to an estimated loss of $ 1.24 billion into the poultry industry in ten provinces (171). From a public health perspective, both LPAIV and HPAIV of Asian origin have shown a

propensity for enhanced zoonotic profiles resulting in significant human infections and mortality (172).

Intensive efforts in FLUAV surveillance in wild birds, especially in North America and Eurasia, have shown the importance of wild bird movements in spreading genetically diverse viruses between locations. Wild bird species have been implicated in the emergence of novel strains, most notably those with the capacity to infect poultry, swine, and humans (173). In other parts of the world, including Latin America, surveillance for FLUAV remains underperformed, and it is implemented mostly in response to disease outbreaks (106, 174, 175). Avian FLUAVs are usually genetically separated by geographical barriers, resulting in the independent evolution of lineages. The North American and Eurasian lineages are widely described (62), but other independent lineages from Australia and Antarctica have been documented (63, 64). Recently, a distinct FLUAV genetic lineage was described in South America, providing evidence of the divergent evolution of these viruses in the Southern Hemisphere. Overall, these observations implicate, geographic barriers and migratory routes of a plethora of avian species as major drivers in the evolution and diversity of FLUAVs (65).

Central America is located at an intersection point between Northern and Southern migratory bird routes, as well as Atlantic and Pacific flyways in the Americas. The majority of the aquatic bird species found in this area are migratory birds, including some of the main FLUAV reservoirs found in North America [northern shovelers (*Anas clypeata*), northern pintails (*Anas acuta*), blue-winged teals (*Anas discors*), and American wigeons (*Anas americana)*] (102). Previous studies show that Central America´s unique location provides a site where hundreds of different migratory bird populations converge

yearly, increasing the likelihood of interspecies transmission and gene segment exchanged among FLUAVs, and thus contribute to virus diversity (103).

No FLUAV outbreaks in poultry have been reported in Central America. However, circulation of LPAIV in this area was confirmed with the isolation of an H5N2 Mexican lineage from poultry in 2000 (104). Phylogenetic analysis of the HPAIV H7N3 strain isolated from the outbreak in Mexico in 2012 strongly suggests its relatedness to LPAIV H7 strains isolated from wild birds circulating at that time in North and Central American regions. Interestingly, the Mexican HPAIV H7 strains cluster within the same group of Guatemalan LPAIV H7 viruses identified during 2010-2013 (103, 176, 177). No evidence of the introduction of FLUAVs of the South American lineage has been observed in Guatemala. In contrast, introductions of FLUAV gene segments from Eurasian lineage are detected in Guatemala (103). The sizable number of circulating subtypes (including H5 and H7) and the wide diversity of gene segment constellations is consistent with the convergence of multiple bird migration routes at wintering grounds in Guatemala (103).

Although previous surveillance in studies in Guatemala performed during 2008-2013 have generated viral sequence data, efforts have been directed to address the phylogenetic relations of circulating strains. There are still gaps in the knowledge of the genetic structure of avian FLUAV populations and the seasonality of influenza transmission at the wintering grounds in the region. The results published to date on this topic are limited. This information however is useful to determine the current level of FLUAV gene flow and patterns of transmission in a region that could contribute help to prevent the spread of the disease. This study aims to investigate long-term patterns in IAV

prevalence, and subtype variation, as well as define in more detail the genetic structure of the FLUAV population from wild birds in Guatemala. Previously published viral sequences from 2007 to 2013 from Guatemala, together with newly generated sequencing data obtained during this study, represent the largest data set of FLUAV genomes compiled to date for this region. It may provide a broader picture of the spread of FLAUV which could be used to early detect FLUAV strains with pandemic concern to animal and human health.

Materials and Methods

Methods used for this Chapter are described in detail in Chapter 3.

Results

FLUAV prevalence in Guatemala decreases at the end of each migration season

From November 2013 to March 2019, 3,109 paired cloacal and tracheal samples were collected from at least twenty species different species of aquatic birds, which were distributed in seven locations from three departments (Escuintla, Santa Rosa, and Jutiapa) located on the Pacific coast of Guatemala near the border with El Salvador (Table 4.1 and Figures 4.1 and 4.2A). Only 2 sites (La Gomera, Escuintla, and El Pumpo, Santa Rosa) out of 7 collection sites were sampled consecutively every season. The number of aquatic bird samples per season ranged from 206 to 999. More than 95% of collected samples (2,967/3,109) came from blue-winged teals (n=2,967), ranging from 201 to 973 per season. The overall percentage of RRT-PCR FLUAV-positive samples was 10.0% (95% CI: 10.00% – 11.11%); corresponding to 311 FLUAV-positive samples, with Ct values

ranging from 21.6 to 38.3. The detection of positive samples occurred mostly among members of the Anatidae family (Figure 4.2 A) with the highest number of positive samples in 2014-2015 ($p<0.0001$, $\chi2=104.0$, df=5) with 27.36% (95% CI: 21.67% – 33.91%) (Figure 4.2 B). FLUAV-positive samples were identified in green-winged teals [*Anas crecca*, 40.0% (95% CI: 11.76% – 76.93%), n=5], blue-winged teals [*Anas discors*, 10.35%, (95% CI: 9.30% – 11.49%), n=2,967], black-bellied whistling ducks [*Dendrocygna autumnalis*, 7.14%, (95% CI: 1.27% – 31.47%), n=14)], and northern shovelers [*Anas clypeata*, 1.52%, (95% CI: 0.27% – 8.10%), n=66). Only data from samples collected from blue-winged teals were used to estimate prevalence and seasonal variation since it represented the largest number. The low number of FLUAV positive samples from green-winged teals (n=2), black-bellied whistling ducks (n=1), and northern shovelers (n=1), led to their exclusion from analyses of FLUAV prevalence. Based on blue-winged teal FLUAV positive data, monthly prevalence (as determined by the total number of positive detected by RRT-PCR by the total number of collected samples of the same species) was statistically different in 5 out of 6 seasons: 2013-14 ($p<0.0001$, $\chi2=28.57$, df=4), 2014-15 ($p<0.0001$ using Fisher's exact test), 2016-17 ($p<0.0001$, $\chi2=6.443$, df=2), 2017-18 ($p<0.0001$, $\chi2=88.15$, df=4) and 2018-19 ($p<0.0001$, $\chi2=52.53$, df=3). The exception was observed for the 2015-16 season (p not significant) (Figure 4.3).

Significant FLUAV subtype diversity in blue-winged teal samples from Guatemala.

Whole virus genome sequences were obtained from 123 out of 311 FLUAV-positive samples (38.2%) either from the swab material (n=59) or virus isolate (n=63), following

the approach described by Ferreri et al (101). The HA and NA gene segments were successfully sequenced from an additional 13 swab samples that produced incomplete internal gene segment sequences. And from 17 additional swab samples at least one gene segment was sequenced; overall, 153 samples produced at least one complete gene segment. Of these sequences, only one sequence set corresponds to a virus isolate from a northern shoveler (H3N2) in the 2014-2015 season, whereas the rest of the set corresponds to samples obtained from blue-winged teal distributed among the 6 sampling seasons. FLUAV sequences from blue-winged teal samples included 12 HA and 8 NA subtypes in 26 different combinations (Figure 4.4). The H14 (n=25) subtype was detected with highest frequency, followed by H12 (n=20), H3 (n=18), H5 (n=14) and H11 (n=11) (Figure 4.4). The highest frequency of H14 was identified in the 2013-14 and, 2014-15, consistent with previous observations as previously described (99, 103, 105). The 2013-14 and 2014-15 seasons also produced also the uncommon H8, H10, in all 6 all seasons, H8 and H11 were found in all seasons except 2015-16, whereas H2 (2013-14) was identified only in a single season (Figure 4.4). H5 subtype viruses (n=14) were identified in seasons 2014-15 (n=8), 2015-16, (n=1), 2017-18 (n=5), and 2018-19; as well as seven H7 subtype strains (n=12) were detected in the 2013-14 (n=2), 2016-17 (n=3) and 2018-19 (n=7). The HA of H5 and H7 viruses had motifs consistent with LPAIV (H5 motif PQRETR*GLF and H7 motif PENPKTR*GLF, respectively). Among NA subtypes, N3 (n=75) was the most common, followed by N4 (n=30) and N2 (n=19) (Figure 4.4).

The most frequent FLUAV subtype combinations included H14N3 (n=14), H5N3 (n=10), H8N4 (n=10), and H7N3 (n=8), as well as uncommon subtype combinations in blue-winged teals (recovered ≤10 times in North America) (95) included H12N3 (n=5),

H12N4 (n=6), H12N7 (n=5) and mixed infections (n=20) described on Table 4.2. Finally, 17 blue-winged teal FLUAV positive samples (14%) could not be fully subtyped due to lack of amplification of HA and/or NA gene segments. No subtype combination was found in all 6 sampling seasons; but subtypes H3N2 (2013-14, 2014-15, 2015-16, 2017-18), H8N4 (2013-14, 2014-15, 2016-17, 2018-19), and H11N3 (2013-14, 2016-17, 2017-18, 208-19) were identified in at least four seasons. Subtypes H1N1 (2018-19), H2N2 (2013-14), H2N9 (2013-14), H5N2 (2013-14), H6N5 (2014-15), H7N7 (2013-14), H12N7 (2013-14), and H14N5 (2013-14) were found sporadically in a single season.

The co-circulation of various subtypes combinations over time was readily observed (Figure 4.5). The 2013-14 season presented the highest number of distinct subtype combinations (n=18) followed by the 2018-19 (n=11) and 2014-15 (n=10) seasons. The estimated Simpson diversity index (1-D) for each season is shown in Table 4.3, suggesting high diversity within each season. The Sørensen-Dice coefficient (from 0 to1, none to perfect overlap, respectively) was used to compare differences in the composition of virus populations between 2013-14 and 2018-19, and between 2013-14 and 2014-16 (seasons with the highest number of subtype combinations). The resulting values indicate high diversity of subtypes among the compared seasons (Table 4.4). Overall, high FLUAV diversity was found in blue-winged teal samples during all seasons as indicated by these indexes.

Evidence of multiple introductions of uncommon subtypes in Guatemala and reassortment with Eurasian origin FLAUVs.

To establish the phylogenetic relationships among the sequenced viruses, we used the predicted major ORF for each segment to construct phylogenetic trees by the maximum likelihood (ML) method. Among the HA gene, H5 and H7 viruses clustered with contemporary North American isolates from ducks. However, the HA of two of the Guatemalan H7 viruses (H7N3 and H7 mixed) clustered and shared 97% similarity with A/chicken/Tennessee/17-008279-4/2017(H7N3) isolated in the US during an HPAI outbreak in 2017 (178) (Figure 4.6). Molecular clock analyses of uncommon subtypes showed repeated introductions from the North American lineage viruses (Figure 4.7). The H8 viruses appeared to have been introduced in Guatemala once between 2014.3-2018.3 with closest relatives from the Pacific and Mississippi flyways. In contrast, the H10 viruses had 2 introductions in 2013.2-2015.9 and 2015.4-2016.3 with closest relatives in all 4 major North American flyways. Likewise, H12 viruses were introduced twice, in 2012.0-2015.0, and in 2012.7-2013.1 sharing phylogenies phylogenetic characteristics with viruses from the Atlantic and Mississippi flyways. These observations suggests that that Guatemala can serve as a re-distribution center location for FLUAVs from multiple flyways.

Regarding the NA subtype (Figure 4.8), all N2 gene segments clustered on an ancient divergent clade within the main North American lineage. N3 and N4 gene segments clustered in the main North American lineage interspersed with other North American and Guatemalan viruses previously reported (103).

The internal gene segments (PB2, PB1, PA, NP, M1, NS1) clustered by season together with other viruses isolated in Guatemala during 2010-2013 (Figures 4.9 and 4.10). These gene segments tended to cluster with those from the North American lineage, in agreement with previous observations (101, 103, 105). However, the PA, NP, NS1 gene segments from A/ BWTE/GT/CIP049H152-13/2017(H3N2), A/BWTE/GT/CIP049H116-22/2013(H5N3) and A/BWTE/GT/CIP049H151-22/2013(HxNx) clustered with strains of Eurasian lineage. Evidence of repeated introductions of Eurasian lineage gene segments was also apparent in FLAUV sequences from previous seasons (2010-213) (103). The PB2 gene segment of three viruses from Guatemala [A/BWTE/GT/CIP049H123-54/2014(H6N2), A/BWTE/GT/CIP049H125-108/2013(H14N4) and A/BWTE/GT/CIP049H134-61/2017(H12N4)] , and the HPAIV A/chicken/Jalisco/PAVX17170/2017(H7N3) isolated in Mexico in 2017 (179), clustered within the same clade within the North American lineage and shared >94% similarity. Likewise, the NP gene segment of one virus of these viruses [A/BWTE/GT/CIP049H125-108/2013(H14N4)] and the HPAIV A/chicken/Jalisco/716/2017(H7N3) isolated in Mexico in 2017 (179), clustered within the same clade with >94% similarity. In addition, for the NP gene segments, one samples [A/blue-winged teal/Guatemala/CIP049H151-22/2013(HxNx)] clustered within a separate minor lineage along other North American isolates and other Guatemalan viruses identified during 2012-2013. These basal lineages diverged more than a century ago from its their ancestor and have evolved as a separate lineage since then (180). Six samples [A/BWTE/GT/CIP049H147-10/2017(H5N3), A/BWTE/GT/CIP049H189-41/2019(mixed), A/ BWTE/GT/CIP049H116-84/2013(H6N2),

A/BWTE/GT/CIP049H123-21/2014(mixed), A/ BWTE/GT/CIP049H123-63/2014(H6N2), and BWTE/GT/CIP049H125-23/2015(H14N4)] carry an allele B NS1 gene segments, whereas the rest carry an allele A, both within the North American lineage.

Mixed infections and internal gene segment constellations suggest multiple reassortment events.

At least 20 mixed viral infections were identified. H3N2/N4 and H10N4/N7 were observed twice during the study, while the rest were observed only once. Three viruses presented a HA (H3, H7) in combination with two NA (N2, N3). Likewise, nine viruses had one HA (H3, H4, H5, H8, H10, H12, H14) in combination with two NA (N3, N4, N5, N7, N8). Additionally, eight viruses had two of both HA (H3, H4, H5, H6, H8, H10, H11, H12) and NA (N3, N4, N5) (Table 4.4).

Whole virus genome sequences (n=123) were used to determine the internal gene constellations of FLUAVs identified in Guatemala. Gene constellations were defined as unique gene combination (genotype) at the level of clades. Gene segments were divided into sub-lineages based on the topology of the phylogenetic trees and identity <90%, and further divided into clades based on >96% identity. These gene segments were also compared with other Guatemalan viruses previously published (for a total of 205 additional gene segment sequences). These analyses showed 59 unique genotypes and 127 gene constellations (Figure 4.11). The PB2, PA and NP segments were associated with 2 sub-lineages within the North American lineage, while PB1, M1 and NS1 with only one sub-lineage. Likewise, the PA, NP and NS1 were also associated with the

Eurasian lineage. NS1 presented both forms of the NS1 alleles (A and B), both within the North American lineage.

Overall, the Polymerase gene segments (PB2, PB1 and PA), as well as the NP gene segments were more diverse in terms of clade classification, being classified in 16, 17 and 14 respectively (including all lineages). M1 and NS1 presented the lowest number of clades with 5 and 7 respectively (including all lineages).

The main North American sub-lineage was divided ion clades (PB2=15, PB1=11, PA=6, NP=11, M1=6 and NS1=5). Minor and basal North American lineages, as well as the Eurasian lineage and NS1 allele B presented only one clade of each segment, except for the minor PA North American sub-lineage that were associated with 10 clades.

In general, gene constellations varied from season to season. However, two were detected over three consecutive seasons (one to 2012-13 to 2016-17 and 2014-15 to 2018-19). PB2, PB1, PA and NP gene segments of these two gene constellations were identified by BLAST (>99% identical) in North America circulating in different flyways during the time of collection of the Guatemalan isolates. However, the M1 and NS1 gene segments were traced to 2009 and 2008 respectively in viruses from diving ducks in the Pacific flyaway (Alaska and California), suggesting their persistence for at least a decade.

Molecular analysis shows markers for antivirus resistance and PB1-F2 N66S marker associated with increased virulence in mammals.

No mammalian-associated virulence markers in PB2 (E627K and D701N) (181-184), PA (S409N) (185), or NS1 protein (T92E) (186) were found for any of the viruses. One sample carried mutations correlated with for antivirus resistance (V27A and S31N on

M2) along with the PB1-F2 N66S marker, also associated to with increased virulence in mammals (187-189). Additionally, 15 samples presented two different lengths of the PB1-F2 protein (87 and 90 amino acids), both already described previously (189).

Discussion

Aquatic wild birds, particularly ducks are the major reservoir for FLUAV diversity and play a key role in their maintenance and dissemination (1). Surveillance in these species serves as an early warning for the detection of emergent influenza strains with potential implications for animal and/or human health. This study summarizes the results of FLUAV surveillance efforts performed during six migratory bird seasons (2013-2019) in Guatemala, a stopover site with multiple overlapping flyways, during six migratory bird seasons (2013-2019). Guatemala is one of the most biodiverse regions of the world and has vast regions of tropical rainforest. The Pacific coast supports the highest abundance of aquatic birds in Guatemala, especially migratory bird populations. It is one of the hot spots for migrating waterfowl, with more than 4,400 ducks counted in one season (190). Due to the interactions between aquatic birds, farming practices and human settlements, this region is candidate may potentiate gene segment exchange among FLAUV strains derived from different animal species. Initiated in 2007, the surveillance efforts in Guatemala have resulted in full genome characterization of 205 FLUAVs, 88 of which are newly to this reported here. Our efforts were biased towards blue-winged teal samples since those were the most common samples obtained from hunter-harvested birds. Nevertheless, such sampling provided an opportunity to realize the diverse FLUAV population in a rather small region near the Southern Pacific coast of Guatemala. We

used a workflow in which FLUAV-positive samples were sequenced directly from the swab samples, and only those that failed or yielded incomplete genomes were subjected to virus isolation and re-sequencing. The approach allowed us to improve the number of samples sequenced and reduce potential selection bias introduced by virus isolation prior to sequencing. Overall, we were able to obtain full-length genomes in ~40% of FLUAV-positive samples, which is about twice as high compared to previous studies using the conventional approach of virus isolation (99, 103, 105). A decreasing trend of FLUAV-prevalence in blue-winged teal was found in 5 out 6 migratory seasons consistent with previous reports (77, 95, 99, 100) likely due to development of natural immunity in the host population towards the end of each season (66). The exception was in the in 2015-16 season perhaps due to the limited number of samples collected. Similarly, previous studies performed over two consecutive seasons (2010-2013) found significant FLUAV diversity, in terms of subtype combinations and frequency of detection of viruses that are rarely found at the breeding grounds in North America. These previous studies also revealed unique phylogenetic characteristics in the Guatemalan viruses providing evidence that Guatemala (and perhaps other areas of Central America) can serve as a major geographic bottleneck that results in several bird species comingling in a small area, allowing FLUAVs from multiple North American migratory flyways to converge and exchange gene segments (103). Consistent with this notion, during 2013-2019 sampling period, we identified substantial genetic and subtype diversity of FLUAVs circulating in wild birds on the Pacific Coast in Guatemala. We frequently found ≥10 co-circulating subtypes during the same season, in a relatively small area, almost all from blue-winged teals, suggesting that this avian species could play a key role in maintaining

natural FLUAV diversity in Guatemala. Sequencing directly from swabs showed significant mixed infections (~20% of subtyped isolates), further supporting evidence that the wintering grounds in Guatemala provides an ideal environment for FLUAVs to reassort. Overall, the most common HA subtype found was H14 (discussed in chapter 5), followed by H12 and H8, which have been found in low frequencies in North America (95, 191). By contrast, H3, H4 and H6, commonly found in North America, showed with less frequency in blue winged teal samples of in Guatemala. High frequency of uncommon subtypes has been described in other regions (191), probably maintained during non-migration periods in other FLUAV-susceptible bird species. During migration, FLUAV spills overs to migratory birds in wintering grounds. Evidence of multiple HA subtype introductions in many internal gene segment constellations supports this observation. Additionally, identification of H5 and H7 subtypes was common in Guatemala. These virus subtypes pose a risk of spillover to poultry and other domestic animals, that can later lead to HPAI outbreaks. Our data showed LPAIV H7 strains from Guatemala were closely related to the HPAIV H7N3 strains isolated from poultry in the US in 2017.

Maintaining appropriate FLUAV surveillance in natural hosts can potentially help identify early warning signs for the presence of high consequence pathogens or their precursors. Our data, however, cannot be used to establish whether blue-winged teals serve as hosts for persistence of H5 and H7 FLUAVs, but close monitoring of this species is warranted. Interestingly, the N3 subtype was the most single dominant NA subtype in Guatemala in combination with a myriad of HA subtypes. It remains to be determined,

whether in Guatemala all these observations are a true representation of subtype prevalence in all bird species or just blue-winged teals.

The FLUAV sequences obtained during this study are positioned in multiple clades interspersed with North American origin FLUAVs. However, evidence of reassortment with Eurasian origin strains was also detected. In contrast, we found no evidence of reassortment with viruses from the South American lineage. Our data does not allow us to determine whether these intercontinental lineage FLUAV reassortants emerged in Guatemala or if they were introduced as such from other sources. Nevertheless, Guatemala is likely promoting further virus reassortment as observed by the 127 internal gene constellations detected within and/or across multiple seasons.

Two recurrent internal gene constellations were found at least in five seasons. The M1 and NS gene segments of these constellation were circulating in wild birds in other regions for more than a decade. Maintenance of M and NS gene segments for prolonged time was also observed in FLUAV from ducks of the Atlantic flyway of North America (192), most likely due to the lowest nucleotide substitution rates of these genes (193). Perpetuation of certain FLUAV genes in natural reservoirs in settings regions such as Alaska is explained in part from the environmental conditions. In these settings, the viruses might be preserved in frozen lake water while ducks are absent and transmitted during the next migratory season (194). However, such factors are different in Guatemala, where temperatures can reach the threshold throughout most of the year to would prevent perpetuation of FLUAV in the environment (195). Rather, these subtypes might be maintained by alternative FLUAV susceptible hosts. For instance, within same areas in other regions H8 has been detected frequently in poultry in other regions but it is

almost absent in wild birds (196) within the same areas. Nevertheless, even though our study may describe perhaps only a fraction of the potential FLUAV diversity found in Guatemala since due to sampling was biased towards blue-winged teals. However, such fraction provided a significant set example of FLUAV diversity comparable compared to other parts of the world considered host spots of FLUAV activity. Our study contributes to a better understanding of the ecology and potential reassortment dynamics of FLUAVs in a natural host in an understudied region in Central America.

Figure 4.1. Location of collection sites on the Pacific Coast in Guatemala.

Geographic coordinates are described in detail in Table 4.1.

Figure 4.2. Distribution of collected and FLUAV-positive samples in different species of aquatic birds during 2013-2019 on the Pacific Coast in Guatemala. A) Distribution by species sampled during this study. Members of the Anatidae family are shown in the brackets. The majority of samples were collected from *Anas discors* (blue-winged teals), which is present in relatively high abundance in Guatemala. B) Distribution by the season of collection. The percentage of RRT-PCR FLUAV-positive samples and their 95% CI is shown for each season.

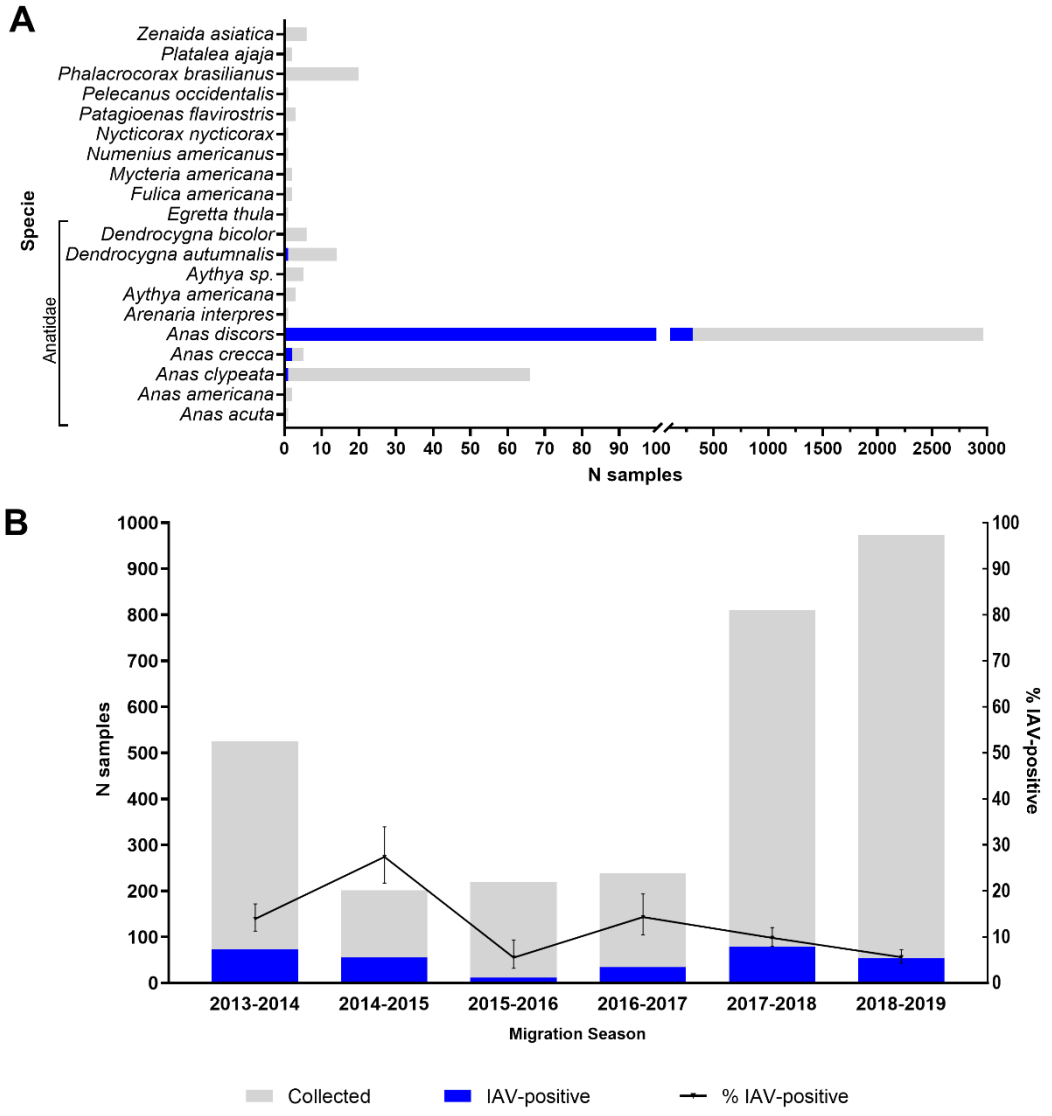

Figure 4.3. Temporal distribution of FLUAV prevalence in blue-winged teals collected during 2013-2019 on the Pacific Coast in Guatemala. Prevalence of FLUAV is shown by month of collection. Monthly prevalence was statistically different in five out six seasons: 2013-14, 2014-15, 2016-17, 2017-18 and 2018-19 with a decreasing trend towards the end of the season.

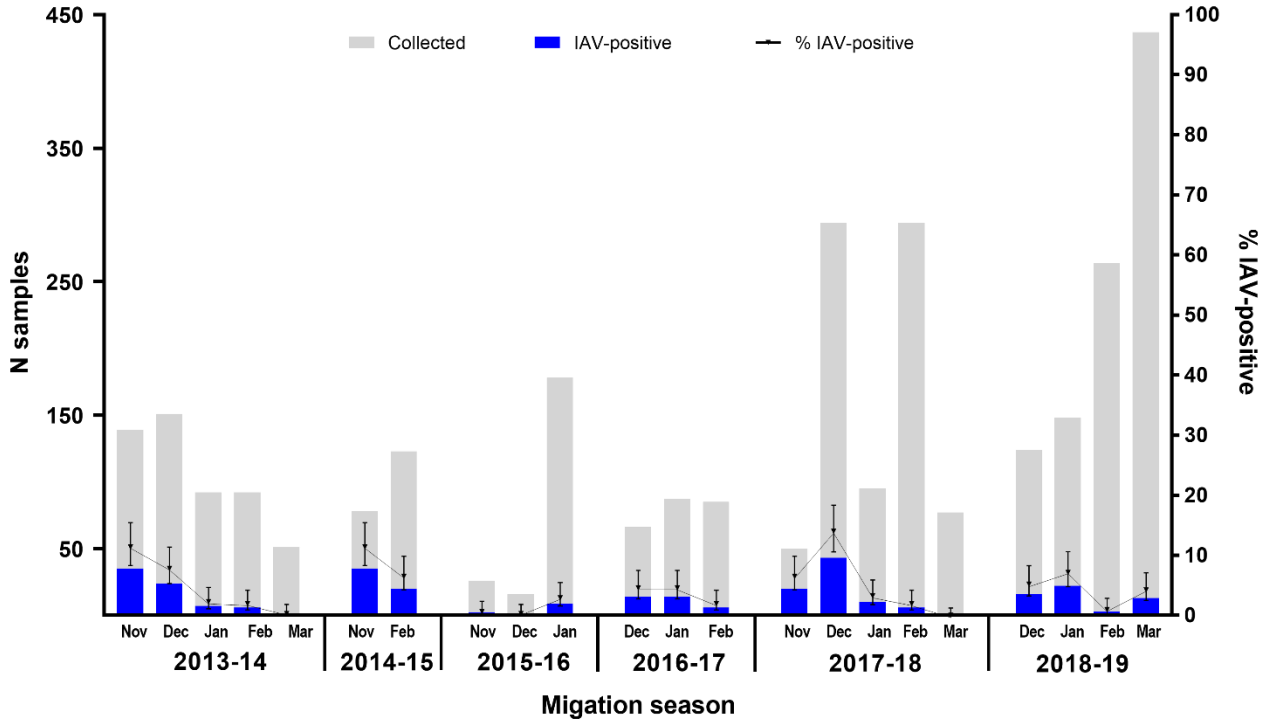

Figure 4.4. Virus subtype combinations identified in FLUAV-positive samples in different species of aquatic birds, Guatemala 2013-2019. Number of subtype combinations identified are presented at top. The gradient background color represents the number of HA or NA identified during the study, going from the smallest number, in dark flue, to the largest, in yellow. No detection of specific subtypes is represented a blank cell. Number of HA and NA per season of collection are presented at bottom. The most abundant HA and NA are shown in red (H14 and N3 respectively), follow by blue (H12 and N4) and yellow (H8 and N5). Less abundant HA and NA are presented in different colors.

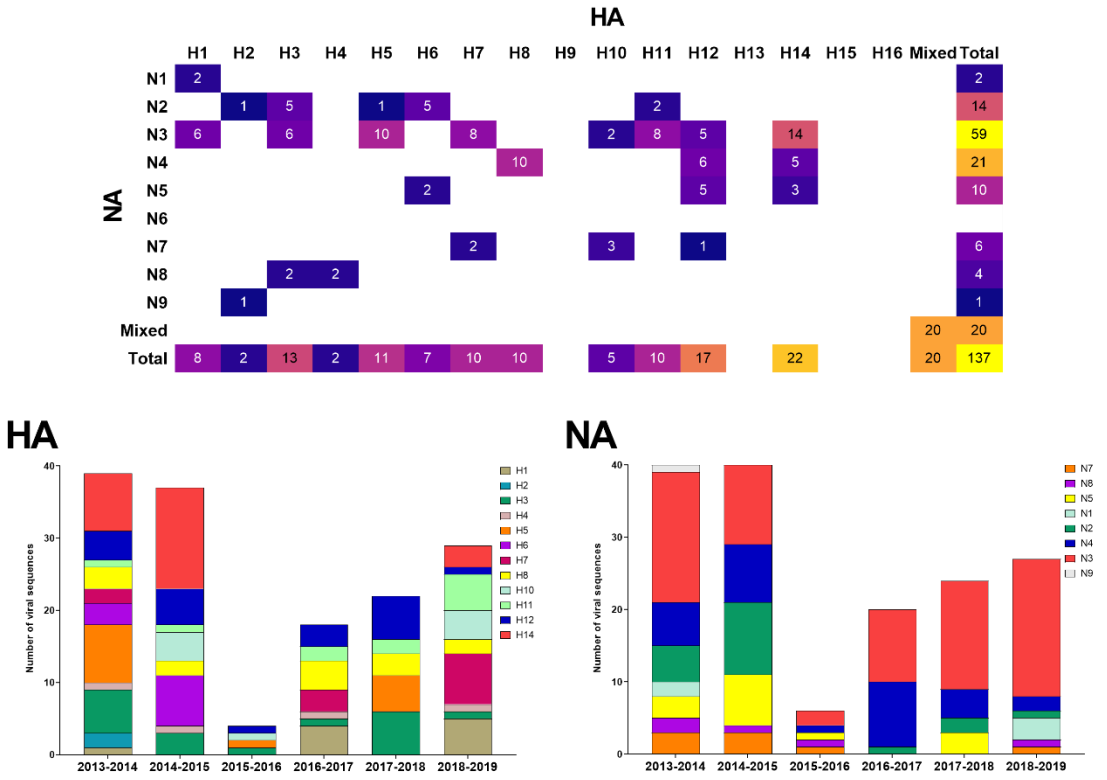

Figure 4.5. Variation of subtype diversity over time in FLUAV-positive samples in different species of aquatic birds, Guatemala 2013-2019. Subtypes diversity is presented by month of collection. The most abundant subtypes are shown in green (mixed), follow by red (H14N3), yellow (H5N3) and blue (H8N4). Less abundant HA and NA are presented in different colors.

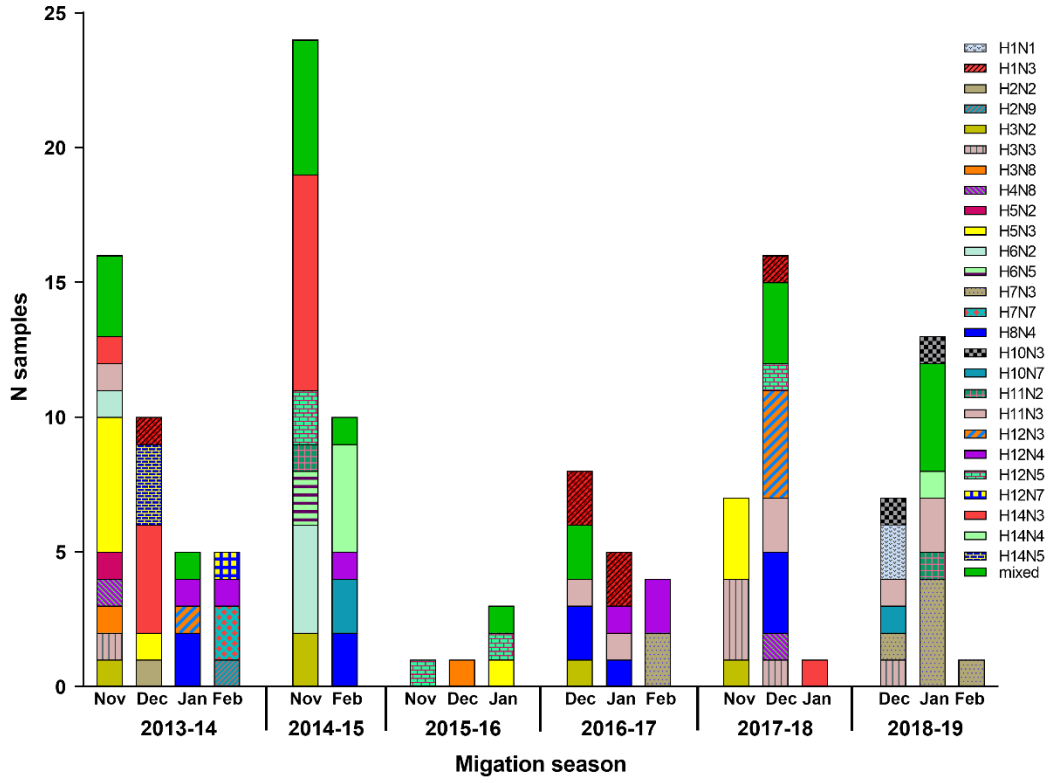

Figure 4.6. Phylogenies inferred using maximum likelihood methods for H5 and H7 from FLUAV-positive samples in different species of aquatic birds collected during 2013-2019 on the Pacific Coast in Guatemala. The tree is midpoint rooted. Scale bar on the bottom-left indicates number of nucleotide substitutions per site. Geographic location is indicated by color. For the H7 gene segment, Guatemalan isolates collected during 2016-2017 cluster together with A/Tennessee/17-008279-4/2017, a virus identified in Tennessee during an HPAI outbreak in 2017.

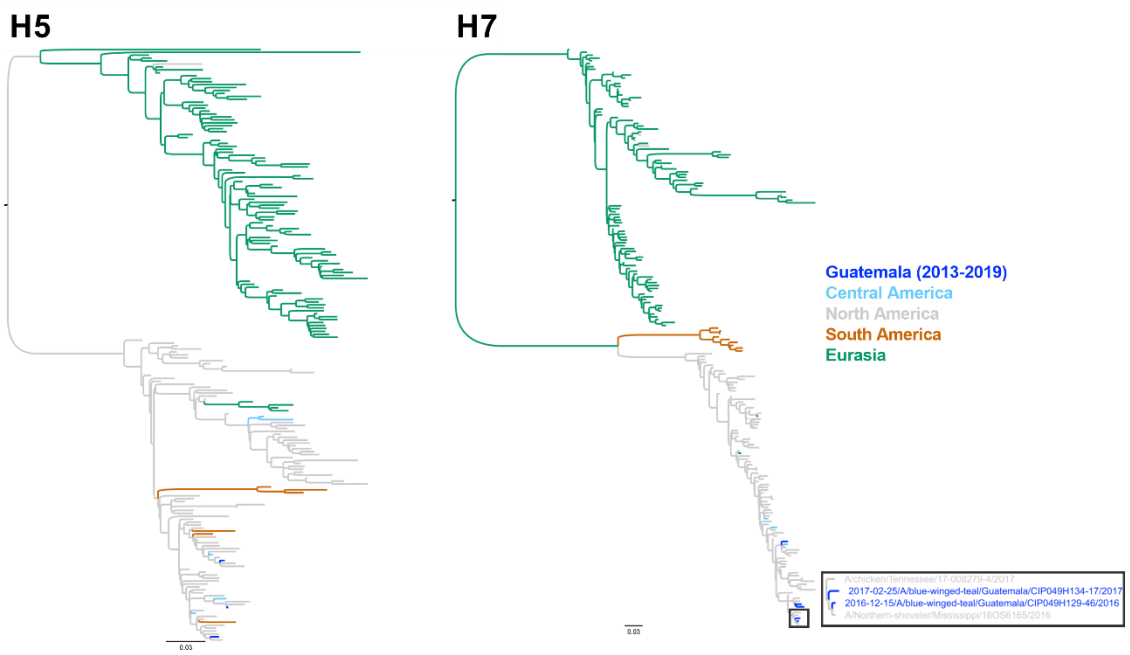

Figure 4.7. Time-scaled MCC trees inferred for uncommon Hemaglutinin (HA) gene found in this study (H8, H10 and H12) from FLUAV-positive samples in different species of aquatic birds, Guatemala 2013-2019. H8 viruses from 1968 to 2019 were included, H10 from 1977 to 2019 and H12 from 1977 to 2019. The H8 viruses were introduced once from the North American linege. In contrast, the H10 viruses had 2 introductions (shown with the arrows) from the same lineage. Geographic location is indicated by color.

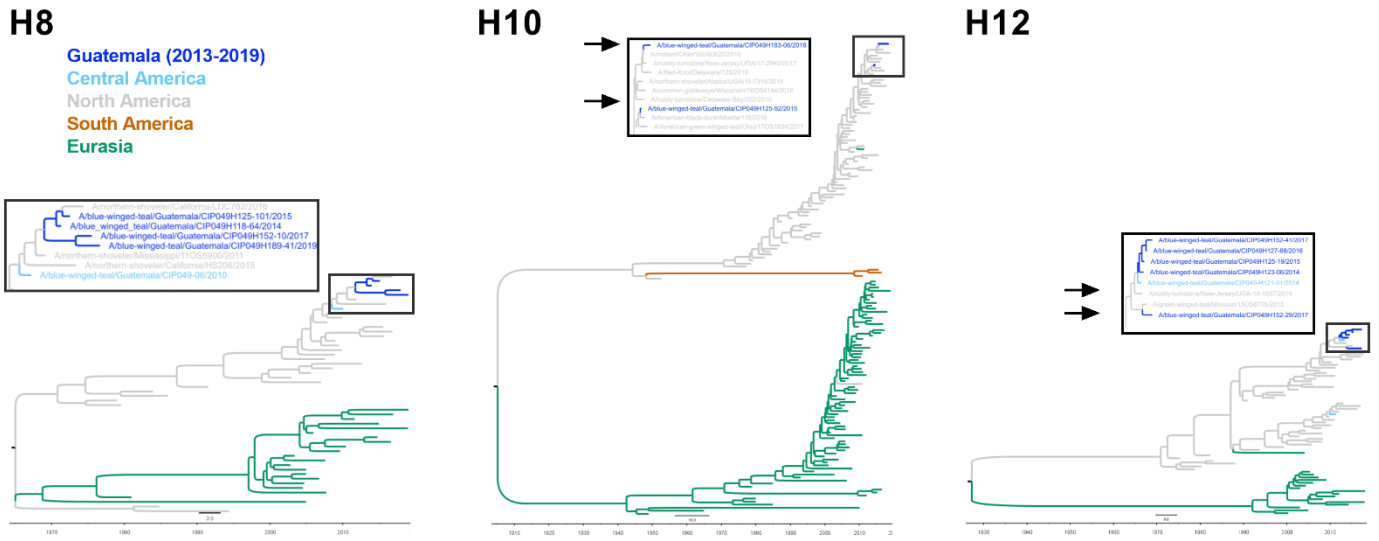

Figure 4.8. Phylogenies inferred using maximum likelihood methods for the most common Neuraminidase (NA) gene segment (N2, N3 and N4) from FLUAV-positive samples in different species of aquatic birds collected during 2013-2019 on the Pacific Coast in Guatemala. The tree is midpoint rooted. Scale bar on the bottom-left indicates number of nucleotide substitutions per site. Geographic location is indicated by color. An ancient divergent North American clade is shown in the blue box.

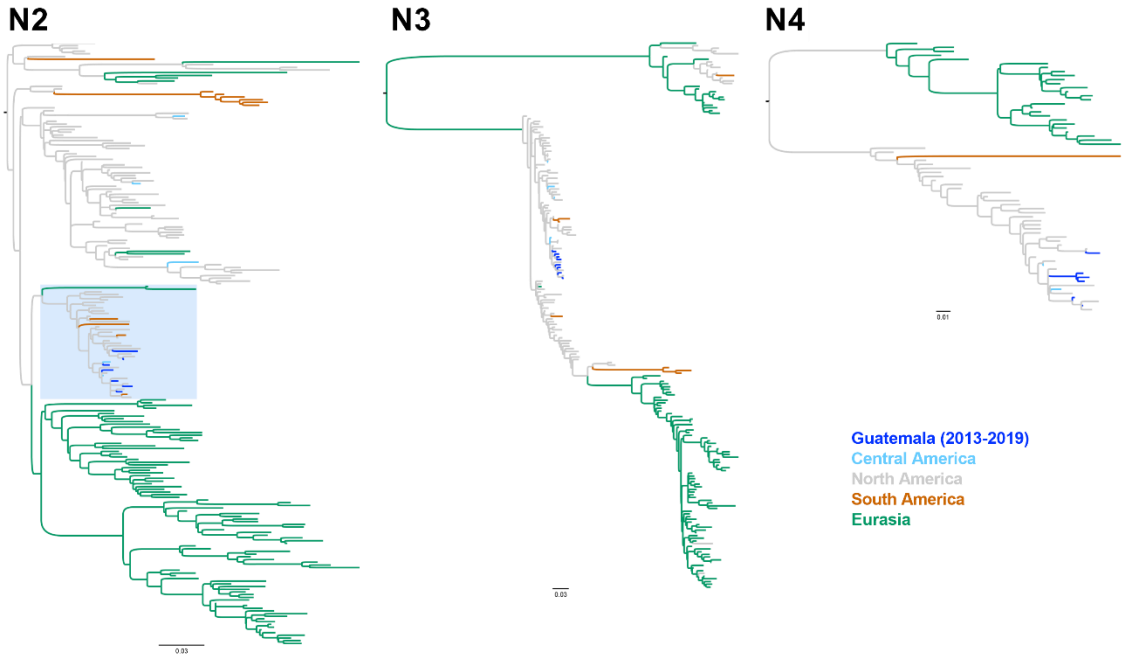

Figure 4.9. Phylogenies inferred using maximum likelihood methods for the Polymerase genes (PB2, PB1 and PA) from FLUAV-positive samples in different species of aquatic birds collected during 2013-2019 on the Pacific Coast in Guatemala. The tree is midpoint rooted. Scale bar on the bottom-left indicates number of nucleotide substitutions per site. Geographic location is indicated by color. The arrow in PA shows Guatemalan isolates collected during this study, as well as other viruses from Guatemala collected in 2012-2013 that cluster within the Eurasian lineage.

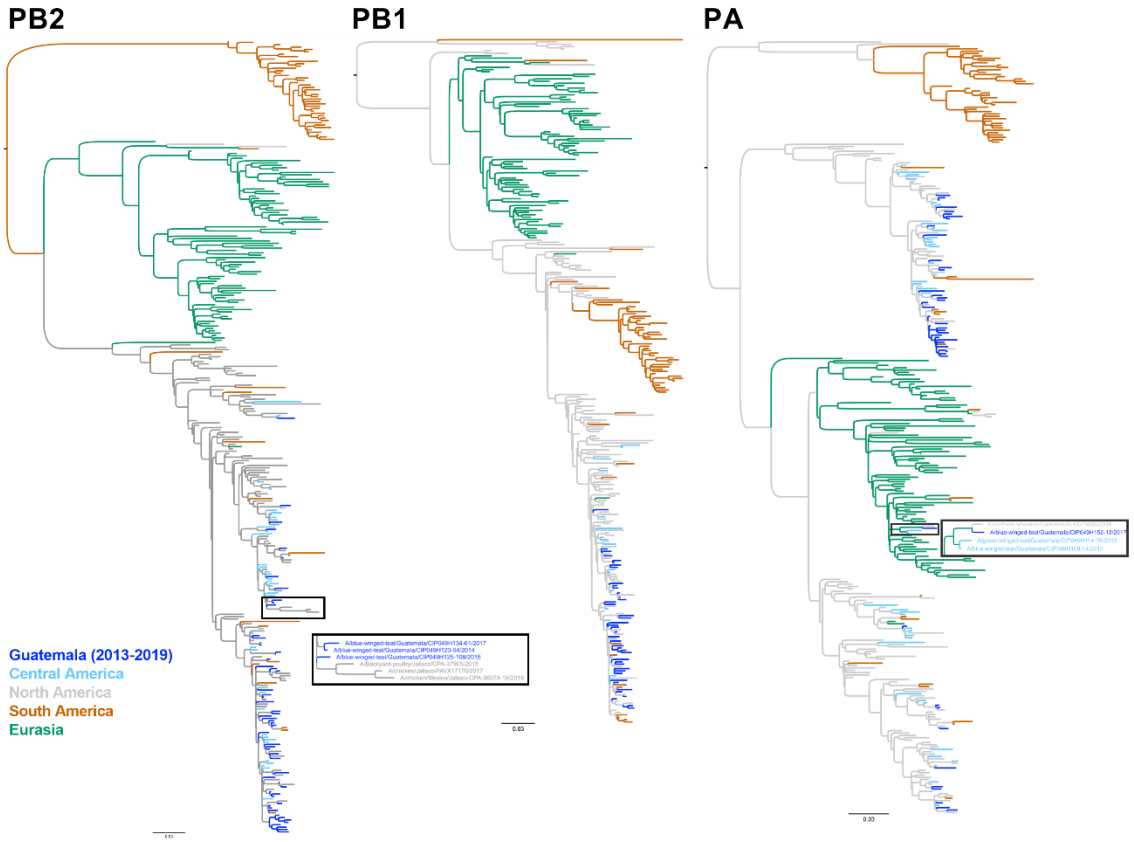

Figure 4.10. Phylogenies inferred using maximum likelihood methods for the internal gene segments (NP, M1 and NS1) from FLUAV-positive samples in different species of aquatic birds collected during 2013-2019 on the Pacific Coast in Guatemala. The tree is midpoint rooted. Scale bar on the bottom-left indicates number of nucleotide substitutions per site. Geographic location is indicated by color. The arrow in NP and NS1 shows Guatemalan isolates collected during this study, as well as other viruses from Guatemala collected in 2013 that clusters within the Eurasian lineage. For the NP gene segment, Guatemalan isolates collected in 2010-2011 and 2015, cluster together with A/chicken/Jalisco/716/2017, a virus identified in Mexico during an HPAI outbreak in 2017. For NS1, alleles A and B are shown.

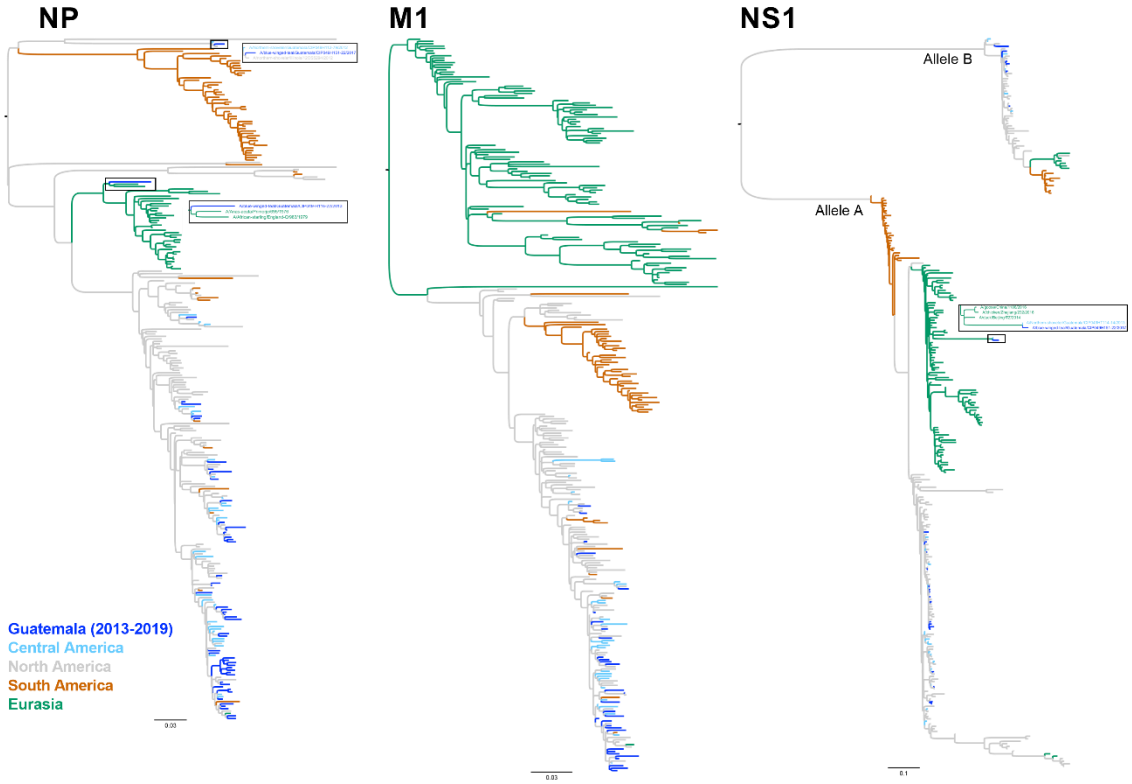

Figure 4.11. Internal genes constellations of FLUAV-positive samples in different species of aquatic birds collected in Guatemala from 2008-2019. A total of 205 viral sequences from Guatemala collected from 2008-2019 were included in the analysis, representing, 59 unique genotypes and 127 gene constellations. Blue, grey and green represent the main, minor and basal North American sub-lineages respectively, and red the Eurasian lineage described on Figures 4.6-4.7. Different clades of the North American lineages are represented by gradient color depending on the number of clades of each segment. The most frequent gene constellation was detected over a period of 6 years, viruses carrying this genotype are shown with an asterisk (*) symbol. Season and number of viruses used for the analysis are shown. A complete description of Guatemalan viruses can be found at Table 4.4.

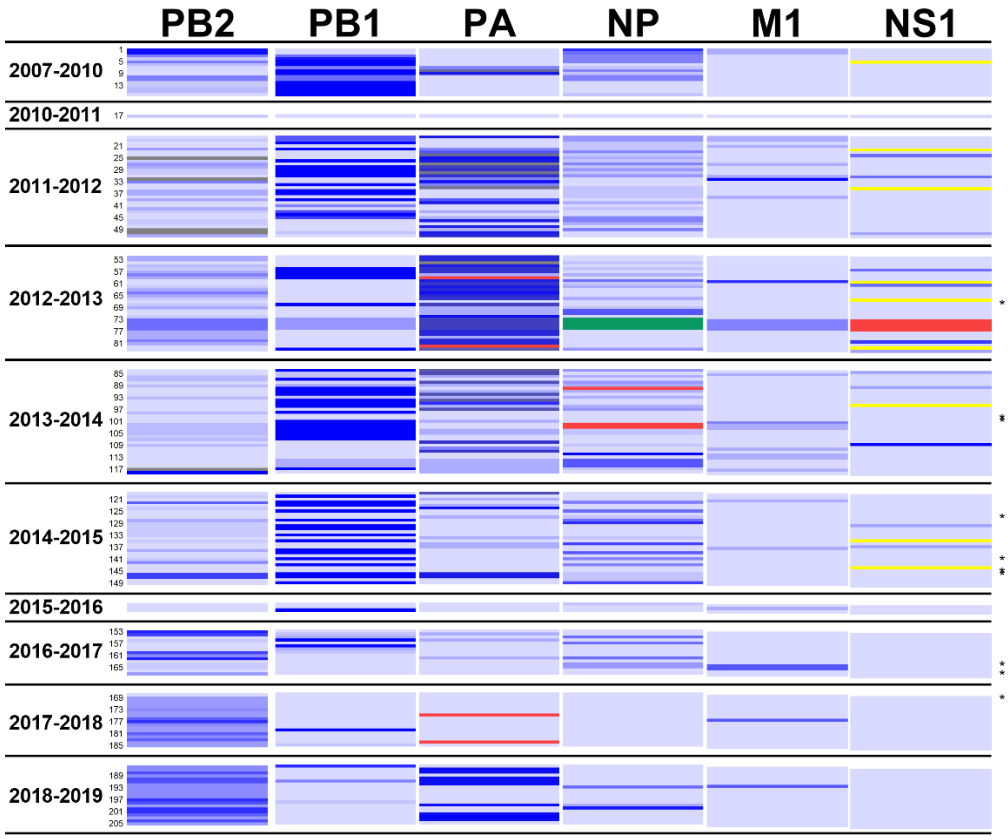

Table 4.1 Distribution of percentage of FLUAV RRT-PCR positive samples by location site[1].

Department	Location	Latitude	Longitude	Sampled seasons	% FLUAV RRT-PCR positive samples	
Escuintla	La Gomera	13.91756 13.9298 13.94427778	-91.0734 -91.11578 -91.11280556	All	9%	(153/1753)
	Sipacate	13.929868	-91.115819	2017-2018 2018-2019	8%	(4/49)
				Sub-total	*9%*	*(157/1802)*
Santa Rosa	La Candelaria	13.915722	-90.548917	2014-2015	16%	(20/127)
	Los Cerritos	13.95947	-90.2993	2015-2016	5%	(1/19)
	El Pumpo	13.89935 13.89772222	-90.4894 -90.49255556	All	9%	(91/986)
	El Zunso	13.92615	-90.61445	2017-2018	15%	(8/52)
				Sub-total	*9.4%*	*(100/1057)*
Jutiapa	Pasaco, Finca la Danta	13.88572222	-90.19591667	2013-2014	28%	(34/123)
				Sub-total	*28%*	*(34/123)*
				TOTAL	***10.0%***	***(311/3,109)***

[1]Location sites are shown in Figure 4.1

Table 4.2 FLUAV mixed infections identified in blue-winged teals during 2013-2019.

Subtype combination	Number of observations	Season	Source f sequencing material
H3/H6N2	1	2014-2015	swab
H7/H8N3	1	2016-2017	swab
H7/H10N3	1	2018-2019	viral isolate
H4N3/N4	1	2016-2017	swab
H5N3/N5	1	2017-2018	swab
H3N2/N4	2*	2013-2014, 2017-2018	viral isolate
H8N3/N4	1	2013-2014	swab
H4N2/N8	1	2014-2015	viral isolate
H12N3/N5	1	2014-2015	swab
H14N3/N5	1	2014-2015	swab
H14N2/N3	1	2014-2015	swab
H10N4/N7	2*	2014-2015, 2015-2016	viral isolate
H3/H6N1/N3	1	2013-2014	viral isolate
H3/H6N1/N3	1	2013-2014	viral isolate
H5/H12N3/N5	1	2017-2018	swab
H1/H8N3/N4	1	2018-2019	swab
H11/H14/N3	1	2018-2019	viral isolate
H1/H8N1/N3	1	2018-2019	swab

* Observed once in a single season

Table 4.3 Number of subtype combinations identified by season and diversity indexes.

Season	Number of subtype combinations	Simpson diversity index (95% CI)	Sørensen-Dice coefficient (CC)[1]
2013-14	18	0.888 (0.841-0.936)	---
2014-15	10	0.847 (0.779-0.915)	0.179
2015-16	3	0.667 (0.667-0.667)	---
2016-17	6	0.890 (0.748-0.870)	---
2017-18	7	0.827 (0.775-0.879)	---
2018-19	12	0.890 (0.840-0.940)	0.167

[1] compared with season 2013-14.

Table 4.4 Full-genome Guatemalan virus sequences from 2007-2013 available on public databases and newly virus sequences from 2013-2019 used for Chapter 4.

Virus number	Season	Strain
1	2007-10	A/blue-winged teal/GT/CIP049-01/2008(H7N9)
2	2007-10	A/blue-winged teal/GT/CIP049-02/2008(H7N9)
3	2007-10	A/blue-winged teal/GT/CIP049-10/2009(H11N2)
4	2007-10	A/blue-winged teal/GT/CIP049-11/2009(H11N2)
5	2007-10	A/blue-winged teal/GT/CIP049-03/2009(H11N2)
6	2007-10	A/blue-winged teal/GT/CIP049-04/2010(H8N4)
7	2007-10	A/blue-winged teal/GT/CIP049-05/2010(H3N8)
8	2007-10	A/blue-winged teal/GT/CIP049-06/2010(H8N4)
9	2007-10	A/blue-winged teal/GT/CIP049-12/2010(H5N4)
10	2007-10	A/blue-winged teal/GT/CIP049-13/2010(H5N4)
11	2007-10	A/blue-winged teal/GT/CIP049-14/2010(H8N4)
12	2007-10	A/blue-winged teal/GT/CIP049-07/2010(H8N4)
13	2007-10	A/blue-winged teal/GT/CIP049-08/2010(H5N3)
14	2007-10	A/blue-winged teal/GT/CIP049-15/2010(H8N4)
15	2007-10	A/blue-winged teal/GT/CIP049-09/2010(H5N3)
16	2010-11	A/blue-winged teal/GT/CIP049H098-32/2010(H5N3)
17	2011-12	A/blue-winged teal/GT/CIP049H101-24/2011(H3N8)
18	2011-12	A/blue-winged teal/GT/CIP049H101-29/2011(H4N2)
19	2011-12	A/blue-winged teal/GT/CIP049H101-36/2011(H4N8)
20	2011-12	A/blue-winged teal/GT/CIP049H102-05/2011(H4N2)
21	2011-12	A/blue-winged teal/GT/CIP049H102-08/2011(H1N3)
22	2011-12	A/blue-winged teal/GT/CIP049H102-09/2011(H3N2)
23	2011-12	A/blue-winged teal/GT/CIP049H102-10/2011(H3N2)
24	2011-12	A/blue-winged teal/GT/CIP049H102-11/2011((mixed)
25	2011-12	A/blue-winged teal/GT/CIP049H102-18/2011(H1N3)
26	2011-12	A/blue-winged teal/GT/CIP049H102-19/2011((mixed)
27	2011-12	A/blue-winged teal/GT/CIP049H102-23/2011((mixed)
28	2011-12	A/blue-winged teal/GT/CIP049H102-25/2011(H3N2)
29	2011-12	A/blue-winged teal/GT/CIP049H102-29/2011(H3N2)
30	2011-12	A/blue-winged teal/GT/CIP049H102-32/2011(H3N8)
31	2011-12	A/blue-winged teal/GT/CIP049H102-46/2011(H1N3)
32	2011-12	A/northern shoveler/GT/CIP049H103-05/2011(H1N3)
33	2011-12	A/blue-winged teal/GT/CIP049H103-13/2011(H3N2)
34	2011-12	A/blue-winged teal/GT/CIP049H103-20/2011(H3N2)
35	2011-12	A/blue-winged teal/GT/CIP049H103-27/2011((mixed)

Virus number	Season	Strain
36	2011-12	A/blue-winged teal/GT/CIP049H103-30/2011(H4N3)
37	2011-12	A/blue-winged teal/GT/CIP049H103-37/2011(H4N3)
38	2011-12	A/blue-winged teal/GT/CIP049H104-38/2011(H1N3)
39	2011-12	A/northern shoveler/GT/CIP049H104-69/2011(H11N3)
40	2011-12	A/blue-winged teal/GT/CIP049H104-93/2011(H3N3)
41	2011-12	A/blue-winged teal/GT/CIP049H104-99/2011(H12N5)
42	2011-12	A/blue-winged teal/GT/CIP049H105-01/2011(H1N3)
43	2011-12	A/blue-winged teal/GT/CIP049H105-04/2011(H11N3)
44	2011-12	A/blue-winged teal/GT/CIP049H105-05/2011(H5N3)
45	2011-12	A/blue-winged teal/GT/CIP049H105-08/2011(H3N3)
46	2011-12	A/blue-winged teal/GT/CIP049H105-15/2011(H14N3)
47	2011-12	A/blue-winged teal/GT/CIP049H105-31/2011(H14N3)
48	2011-12	A/blue-winged teal/GT/CIP049H105-32/2011(H1N3)
49	2011-12	A/blue-winged teal/GT/CIP049H106-62/2011(H14N6)
50	2011-12	A/northern shovelerGT/CIP049H107-88/2012(H7N3)
51	2012-13	A/blue-winged teal/GT/CIP049H108-02/2012(H14N3)
52	2012-13	A/blue-winged teal/GT/CIP049H108-04/2012(H14N3)
53	2012-13	A/blue-winged teal/GT/CIP049H108-11/2012(H14N3)
54	2012-13	A/blue-winged teal/GT/CIP049H108-39/2012(H4N2)
55	2012-13	A/blue-winged teal/GT/CIP049H108-45/2012(H3N2)
56	2012-13	A/blue-winged teal/GT/CIP049H108-53/2012(H4N2)
57	2012-13	A/blue-winged teal/GT/CIP049H108-67/2012(H3N2)
58	2012-13	A/blue-winged teal/GT/CIP049H109-06/2012(H4N2)
59	2012-13	A/blue-winged teal/GT/CIP049H109-14/2012(H3N3)
60	2012-13	A/blue-winged teal/GT/CIP049H109-15/2012(H3N2)
61	2012-13	A/blue-winged teal/GT/CIP049H109-28/2012(H3N2)
62	2012-13	A/blue-winged teal/GT/CIP049H109-38/2012(H3N8)
63	2012-13	A/blue-winged teal/GT/CIP049H109-49/2012(H14N3)
64	2012-13	A/blue-winged teal/GT/CIP049H109-67/2012(H14N3)
65	2012-13	A/blue-winged teal/GT/CIP049H109-76/2012(H14N3)
66	2012-13	A/blue-winged teal/GT/CIP049H109-86/2012(H4N2)
67	2012-13	A/blue-winged teal/GT/CIP049H110-04/2012((mixed)
68	2012-13	A/blue-winged teal/GT/CIP049H110-27/2012(H4N6)
69	2012-13	A/blue-winged teal/GT/CIP049H110-31/2012(H14N3)
70	2012-13	A/blue-winged teal/GT/CIP049H110-36/2012(H11N9)
71	2012-13	A/blue-winged teal/GT/CIP049H110-37/2012(H11N3)
72	2012-13	A/blue-winged teal/GT/CIP049H111-98/2012(H2N3)
73	2012-13	A/American wigeon/GT/CIP049H112-03/2012(H4N2)
74	2012-13	A/blue-winged teal/GT/CIP049H112-31/2012(H4N2)

Virus number	Season	Strain
75	2012-13	A/blue-winged teal/GT/CIP049H112-60/2012(H3N2)
76	2012-13	A/northern shoveler/GT/CIP049HT112-79/2012(H4N2)
77	2012-13	A/blue-winged teal/GT/CIP049H113-07/2013(H14N3)
78	2012-13	A/blue-winged teal/GT/CIP049H113-08/2013(H14N4)
79	2012-13	A/blue-winged teal/GT/CIP049H113-53/2013(H5N3)
80	2012-13	A/blue-winged teal/GT/CIP049H113-74/2013(H14N3)
81	2012-13	A/blue-winged teal/GT/CIP049H113-76/2013(H14N4)
82	2012-13	A/blue-winged teal/GT/CIP049H114-76/2013((mixed)
83	2013-14	A/blue-winged teal/GT/CIP049-I_H115-29/2013(H3N3)
84	2013-14	A/blue-winged teal/GT/CIP049-H116-05/2013((mixed)
85	2013-14	A/blue-winged teal/GT/CIP049-H116-07/2013(H5N3)
86	2013-14	A/blue-winged teal/GT/CIP049-H116-10/2013(H5N3)
87	2013-14	A/blue-winged teal/GT/CIP049-H116-120/2013(H14N3)
88	2013-14	A/blue-winged teal/GT/CIP049-I_H116-16/2013(H3N8)
89	2013-14	A/blue-winged teal/GT/CIP049-I_H116-17/2013(H3N2)
90	2013-14	A/blue-winged teal/GT/CIP049-H116-22/2013(H5N3)
91	2013-14	A/blue-winged teal/GT/CIP049-I_H116-48/2013((mixed)
92	2013-14	A/blue-winged teal/GT/CIP049-H116-50/2013(H5N3)
93	2013-14	A/blue-winged teal/GT/CIP049-H116-51/2013(H5N3)
94	2013-14	A/blue-winged teal/GT/CIP049-I_H116-76/2013((mixed)
95	2013-14	A/blue-winged teal/GT/CIP049-I_H116-84/2013(H6N2)
96	2013-14	A/blue-winged teal/GT/CIP049-H116-96/2013(H5N2)
97	2013-14	A/blue-winged teal/GT/CIP049-I_H116-97/2013(H4N8)
98	2013-14	A/blue-winged teal/GT/CIP049-I_H117-123/2013(H14N3)
99	2013-14	A/blue-winged teal/GT/CIP049-H117-125/2013(H1N3)
100	2013-14	A/blue-winged teal/GT/CIP049-I_H117-13/2013(H14N3)
101	2013-14	A/blue-winged teal/GT/CIP049-H117-130/2013(H2N2)
102	2013-14	A/blue-winged teal/GT/CIP049-I_H117-143/2013(H14N3)
103	2013-14	A/blue-winged teal/GT/CIP049-H117-34/2013(H14N3)
104	2013-14	A/blue-winged teal/GT/CIP049-H117-36/2013(H14N5)
105	2013-14	A/blue-winged teal/GT/CIP049-H117-38/2013(H14N5)
106	2013-14	A/blue-winged teal/GT/CIP049-H117-42/2013(H5N3)
107	2013-14	A/blue-winged teal/GT/CIP049-H117-99/2013(H14N5)
108	2013-14	A/blue-winged teal/GT/CIP049-H118-23/2014((mixed)
109	2013-14	A/blue-winged teal/GT/CIP049-I_H118-64/2014(H8N4)
110	2013-14	A/blue-winged teal/GT/CIP049-H118-81/2014(H8N4)
111	2013-14	A/blue-winged teal/GT/CIP049-H119-01/2014(H12N3)
112	2013-14	A/blue-winged teal/GT/CIP049-H119-16/2014(H12N4)
113	2013-14	A/blue-winged teal/GT/CIP049-I_H120-33/2014(H12N4)

Virus number	Season	Strain
114	2013-14	A/blue-winged teal/GT/CIP049-H121-01/2014(H12N7)
115	2013-14	A/blue-winged teal/GT/CIP049-H121-09/2014(H7N7)
116	2013-14	A/blue-winged teal/GT/CIP049-H121-14/2014(H7N7)
117	2013-14	A/blue-winged teal/GT/CIP049-H121-36/2014(H2N9)
118	2014-15	A/blue-winged teal/GT/CIP049H123-06/2014(H12N5)
119	2014-15	A/blue-winged teal/GT/CIP049H123-12/2014(H6N2)
120	2014-15	A/blue-winged teal/GT/CIP049H123-13/2014(H14N3)
121	2014-15	A/blue-winged teal/GT/CIP049H123-16/2014(H14N3)
122	2014-15	A/blue-winged teal/GT/CIP049H123-21/2014/(mixed)
123	2014-15	A/blue-winged teal/GT/CIP049H123-32/2014(H6N5)
124	2014-15	A/blue-winged teal/GT/CIP049H123-33/2014(H14N3)
125	2014-15	A/blue-winged teal/GT/CIP049H123-35/2014/(mixed)
126	2014-15	A/blue-winged teal/GT/CIP049H123-38/2014/(mixed)
127	2014-15	A/blue-winged teal/GT/CIP049H123-42/2014(H6N2)
128	2014-15	A/northern shoveler/GT/CIP049H123-52/2014(H3N2)
129	2014-15	A/blue-winged teal/GT/CIP049H123-54/2014(H6N2)
130	2014-15	A/blue-winged teal/GT/CIP049H123-57/2014(H14N3)
131	2014-15	A/blue-winged teal/GT/CIP049H123-58/2014(H14N3)
132	2014-15	A/blue-winged teal/GT/CIP049H123-59/2014(H14N3)
133	2014-15	A/blue-winged teal/GT/CIP049H123-61/2014(H14N3)
134	2014-15	A/blue-winged teal/GT/CIP049H123-63/2014(H6N2)
135	2014-15	A/blue-winged teal/GT/CIP049H123-65/2014/(mixed)
136	2014-15	A/blue-winged teal/GT/CIP049H123-74/2014(H3N2)
137	2014-15	A/blue-winged teal/GT/CIP049H123-75/2014(H14N3)
138	2014-15	A/blue-winged teal/GT/CIP049H123-76/2014/(mixed)
139	2014-15	A/blue-winged teal/GT/CIP049H124-27/2015(H12N5)
140	2014-15	A/blue-winged teal/GT/CIP049H125-101/2015(H8N4)
141	2014-15	A/blue-winged teal/GT/CIP049H125-108/2015(H14N4)
142	2014-15	A/blue-winged teal/GT/CIP049H125-14/2015(H14N4)
143	2014-15	A/blue-winged teal/GT/CIP049H125-23/2015(H14N4)
144	2014-15	A/blue-winged teal/GT/CIP049H125-32/2015(H8N4)
145	2014-15	A/blue-winged teal/GT/CIP049H125-59/2015(H14N4)
146	2014-15	A/blue-winged teal/GT/CIP049H125-67/2015(H10N7)
147	2014-15	A/blue-winged teal/GT/CIP049H125-82/2015(H10N7)
148	2014-15	A/blue-winged teal/GT/CIP049H125-92/2015/(mixed)
149	2015-16	A/blue-winged teal/GT/CIP049H126-15/2015(H3N8)
150	2015-16	A/blue-winged teal/GT/CIP049H127-52/2016/(mixed)
151	2015-16	A/blue-winged teal/GT/CIP049H127-88/2016(H12N5)
152	2016-17	A/blue-winged teal/GT/CIP049H129-24/2016(H3N2)

Virus number	Season	Strain
153	2016-17	A/blue-winged teal/GT/CIP049H129-27/2016(H1N3)
154	2016-17	A/blue-winged teal/GT/CIP049H129-39/2016(H11N3)
155	2016-17	A/blue-winged teal/GT/CIP049H129-40/2016(H8N4)
156	2016-17	A/blue-winged teal/GT/CIP049H129-46/2016/(mixed)
157	2016-17	A/blue-winged teal/GT/CIP049H129-50/2016/(mixed)
158	2016-17	A/blue-winged teal/GT/CIP049H129-67/2016(H8N4)
159	2016-17	A/blue-winged teal/GT/CIP049H130-06/2017(H1N3)
160	2016-17	A/blue-winged teal/GT/CIP049H130-31/2017(H11N3)
161	2016-17	A/blue-winged teal/GT/CIP049H130-52/2017(H1N3)
162	2016-17	A/blue-winged teal/GT/CIP049H130-65/2017(H8N4)
163	2016-17	A/blue-winged teal/GT/CIP049H134-17/2017(H7N3)
164	2016-17	A/blue-winged teal/GT/CIP049H134-36/2017(H12N4)
165	2016-17	A/blue-winged teal/GT/CIP049H134-61/2017(H12N4)
166	2016-17	A/blue-winged teal/GT/CIP049H134-78/2017(H7N3)
167	2017-18	A/blue-winged teal/GT/CIP049H147-14/2017(H3N3)
168	2017-18	A/blue-winged teal/GT/CIP049H147-30/2017(H3N2)
169	2017-18	A/blue-winged teal/GT/CIP049H147-31/2017(H5N3)
170	2017-18	A/blue-winged teal/GT/CIP049H147-32/2017(H3N2)
171	2017-18	A/blue-winged teal/GT/CIP049H147-41/2017(H5N3)
172	2017-18	A/blue-winged teal/GT/CIP049H147-46/2017(H3N2)
173	2017-18	A/blue-winged teal/GT/CIP049H148-12/2017/(mixed)
174	2017-18	A/blue-winged teal/GT/CIP049H152-10/2017(H8N4)
175	2017-18	A/blue-winged teal/GT/CIP049H152-13/2017(H3N2)
176	2017-18	A/blue-winged teal/GT/CIP049H152-20/2017(H12N3)
177	2017-18	A/blue-winged teal/GT/CIP049H152-33/2017(H8N4)
178	2017-18	A/blue-winged teal/GT/CIP049H152-37/2017(H8N4)
179	2017-18	A/blue-winged teal/GT/CIP049H152-41/2017/(mixed)
180	2017-18	A/blue-winged teal/GT/CIP049H152-46/2017/(mixed)
181	2017-18	A/blue-winged teal/GT/CIP049H152-66/2017(H11N3)
182	2017-18	A/blue-winged teal/GT/CIP049H152-69/2017(H11N3)
183	2017-18	A/blue-winged teal/GT/CIP049H152-71/2017(H12N5)
184	2017-18	A/blue-winged teal/GT/CIP049H152-80/2017(H4N8)
185	2018-19	A/blue-winged teal/GT/CIP049H158-26/2018(H14N3)
186	2018-19	A/blue-winged teal/GT/CIP049H183-04/2018(H7N3)
187	2018-19	A/blue-winged teal/GT/CIP049H183-06/2018(H10N7)
188	2018-19	A/blue-winged teal/GT/CIP049H183-15/2018(H1N1)
189	2018-19	A/blue-winged teal/GT/CIP049H185-18/2018(H5N3)
190	2018-19	A/blue-winged teal/GT/CIP049H186-15/2018(H10N3)
191	2018-19	A/blue-winged teal/GT/CIP049H186-37/2018(H11N3)

Virus number	Season	Strain
192	2018-19	A/blue-winged teal/GT/CIP049H186-51/2018(H1N1)
193	2018-19	A/blue-winged teal/GT/CIP049H187-16/2019(H11N2)
194	2018-19	A/blue-winged teal/GT/CIP049H187-26/2019(H10N3)
195	2018-19	A/blue-winged teal/GT/CIP049H188-09/2019(H3N7)
196	2018-19	A/blue-winged teal/GT/CIP049H188-12/2019(H3N7)
197	2018-19	A/blue-winged teal/GT/CIP049H189-08/2019/(mixed)
198	2018-19	A/blue-winged teal/GT/CIP049H189-11/2019(H7N3)
199	2018-19	A/blue-winged teal/GT/CIP049H189-15/2019/(mixed)
200	2018-19	A/blue-winged teal/GT/CIP049H189-19/2019(H14N4)
201	2018-19	A/blue-winged teal/GT/CIP049H189-31/2019(H11N3)
202	2018-19	A/blue-winged teal/GT/CIP049H189-32/2019(H11N3)
203	2018-19	A/blue-winged teal/GT/CIP049H189-40/2019(H7N3)
204	2018-19	A/blue-winged teal/GT/CIP049H189-41/2019/(mixed)
205	2018-19	A/blue-winged teal/GT/CIP049H190-41/2019(H7N3)

Acknowledgements

This study was supported by a contract (HHSN272201400008C) from the National Institute of Allergy and Infectious Diseases (NIAID) Centers for Influenza Research and Surveillance (CEIRS).

CHAPTER 5

UNUSUAL INCIDENCE OF H14 SUBTYPE INFLUENZA A VIRUS IN BLUE-WINGED TEALS IN GUATEAMALA

Lucia Ortiz; Ginger Geiger, Lucas Ferreri, David Moran, Dione Méndez, Carlos Serrano, Ana Silvia Gonzalez-Reiche, Mayra Motta, Francisco Escobar, Daniela Rajao, Celia Cordon-Rosales, and Daniel R. Perez.

To be submitted to *Journal of Virology*.

Abstract

Influenza A viruses (FLUAV) infects a wide range of birds and mammals. Wild aquatic birds are considered the natural hosts for in which 16 HA and 9 NA subtypes have been detected. However, with some subtypes, such as H14, have been rarely detected, such as H14. Since its first detection in Guatemala in 2011, it H14 became the predominant subtype over three consecutive migration seasons leading to the largest collection of H14 FLUAVs from a single region. Here, genomic characteristics and gene constellations of 16 new H14 FLUAVs full-length genome sequences are analyzed and compared to previously published H14 sequences from Guatemala and elsewhere.

Pairwise percentage of identity suggests high degree of similarity among Guatemalan and those of North American origin. Phylogenetic analyses revealed at least 3 independent introductions of H14 viruses from the North America lineage and its persistence for ~4 years in the country. Overall, the 36 H14 Guatemalan FLUAVs contained 18 unique internal gene and 21 gene constellations in blue-winged teals in Guatemala. Interestingly, the HA of the Guatemalan H14s showed three distinct motifs at the cleavage site, two of these containing Arginine instead of Lysine in the first and fourth positions, respectively.

Introduction

Influenza A viruses (FLUAV) infect a wide range of bird species and mammals, including humans [1]. The virus genome contains 8 segments of negative single-stranded RNA encoding 6 internal (PB2, PB1, PA, NP, M, and NS) and 2 surface (HA and NA) gene segments. Subtype classification is based on the antigenic properties of the surface glycoproteins. Up to date, 18 HA (H1–H18) and 11 NA (N1–N11) subtypes have been described (2-4). Wild aquatic birds are considered the natural hosts for 16 HA (H1-16) and 9 NA (N1-9) subtypes, playing an important role in the perpetuation of FLUAVs in nature. H3, H4, and H6 subtypes are usually detected at high frequencies in the breeding grounds in North America, while others such as H8, H12, H14, and H15 are rarely found (95, 191).

The H14 subtype of the Eurasian lineage was initially detected from four virus isolates obtained from mallard ducks and a herring gull in the former Soviet Union (Russia and Kazakhstan) in 1982 (143). More recently, four H14 strains were published in the GISAID and IRD databases (accessed April 11, 2022) from a garganey duck in Ukraine in 2006, a goose in Pakistan in 2014, and two in 2019 in Russia from a sandpiper and a common teal. In North America, H14 FLUAVs have been detected sporadically from 2010-2018 from different species of dabbling ducks such as blue-winged-teal, northern-shoveler, long-tailed-duck, mallards and scoter within the Pacific, Central, and Mississippi flyways (197-200). Phylogenetic analyses indicate that the HA segment of H14 viruses in the US was introduced from an Eurasian ancestor around 2010 (200). Consistent with these observations, serological studies suggest H14 subtype FLUAVs were absent in North American ducks before the initial detection in Wisconsin in 2010

(66). Overall, the number of recovered H14 FLUAVs around the world, except Guatemala, account for <25, of which 16 have complete genomes. In Guatemala, H14 subtype FLUAVs were first detected in 2011 from hunter-harvested blue-winged teals (BTWE) and quickly became the most prominent subtype detected in areas of the Southern Pacific coast of Guatemala. Since then, Guatemala has reported the highest number of H14 subtype FLUAVs from a single location (99, 101, 103, 105), all from BWTE samples. From 2014 to 2019, we identified 16 new H14 FLUAVs from Guatemala. Full-genome sequencing of these viruses was used to perform phylogenetic analyses of the entire set of H14 virus sequences available in public databases. These analyses indicate that Guatemalan H14 viruses were independently introduced in Guatemala from North American viruses at least 3 times.

Materials and Methods

Methods used for this Chapter are described in detail in Chapter 3.

Results

H14 FLUAVs from Guatemala are paired to with various NA subtypes and are like similar to those from North America.

During five seasons of avian FLUAV surveillance in Guatemala (seasons 2014-15, 2015-16, 2017-18, and 2018-19), 3,109 paired tracheal/cloacal swab samples were obtained from wild aquatic birds on the Pacific Coast of Guatemala (see Chapter 4). Following FLUAV's matrix (M) gene screening by RRT-PCR and sequencing by NGS, 16 H14 full-length genomes were obtained, all from blue-winged teal samples. Subtype

combinations included H14N3 (n=23), H14N4 (n=7), H14N5 (n=3), H14N6 (n=1), and two mixed infections (H14N3/N5 and H14N2/N3). Together with the H14 sequences already published, this new set of sequences brings the total of Guatemalan H14 strains to accounts for 40 full H14 HA gene segment sequences, corresponding to 36 H14 full-length virus genomes sequences. H14 FLUAVs are positioned within two major clades based on geographic regions: Eurasian and North America (103, 198). To determine the similarity among the Guatemalan H14 FLUAVs, pairwise percentage nucleotide identities were established using either concatenated or single gene ORF sequences. Genetic comparisons across the concatenated internal gene sequences (PB2, PB1, PA, HA, NP, M1, and NS1) from the 36 full genome sequences revealed 95.5% pairwise identity (Figure 5.1). At the individual gene level, the HA segment presented the highest % of pairwise identity (98.1%), followed by M1 (98.0%), PB1 (96.9%), NS1 (96.6%), NP (95.1%), PB2 (93.6%) and PA (92.9%). Identical HA of H14 sequences were detected within the Guatemalan blue-winged teal samples (24 out 36), most of them identified from samples collected the same day. Pairwise comparisons among NA gene segments were excluded from these comparisons since multiple NA subtypes were detected.

The 36 full-genome H14 virus sequences were also compared to the 12 H14 full-length virus genomes from North America (n=10, from Ohio, Texas, Wisconsin, California, Mississippi, and Missouri) and Eurasia (n=2, from Russia and Pakistan) available on IRD and NCBI GeneBank database (accessed April 11, 2022). The H14 Guatemalan and North American FLUAVs are similar, showing 93.4%. pairwise identity based on concatenated internal gene segments. In contrast, similar analysis reveals

pairwise identity of 88.7% compared to H14 Eurasian sequences (Figure 5.1). Additional analyses using single gene comparisons showed pairwise % identity between Guatemalan and North American sequences ranging from 97.2% to 85.8% (PB2:92.9%, PB1:95.8%, PA:89.9%, H14: 95.8%, NP: 94.3%, MP:97.2%, NS:85.8%). Guatemalan H14 viruses showed less similarity to Eurasian viruses. However, the NS1 appears to be more similar to Eurasian viruses, possibly because of the presence of both the A and B alleles in the set of sequences (PB2:84.2%, PB1:88.8%, PA:89.0%, H14: 89.5%, NP: 89.9%, MP:92.8%, NS:93.7%).

Multiple introductions of H14 subtype FLUAVs from the North American lineage with diverse gene constellations were detected in blue-winged teals in Guatemala.

Phylogenetic analyses showed the expected segregation of H14 FLUAVs into the Eurasian and North American clades (103, 198). The Guatemalan H14 FLUAVs clustered together with those from North America and consistent with date of collection. This analysis revealed at least three independent introductions of North American H14 FLUAVs viruses into Guatemala (Figure 5.2), two of them previously described (103). The closest ancestors of the HA of the H14 Guatemalan viruses were identified in the Mississippi flyway as previously observed (198, 201). After its introduction in 2011, the H14 viruses persisted in Guatemala until ~2014.9 to 2015.1 (95% highest posterior density [HPD]). The most recent H14 virus identified in Guatemala (sample collected in 2019) clustered in a separate clade with other North American viruses, indicating a recent independent introduction of the H14 subtype in the country. Similarly, all NAs from the

H14 Guatemalan strains grouped within contemporary North American viruses of the same subtype.

Blast searches of internal gene segments revealed high nucleotide identity (>95%) with contemporary sequences of viruses circulating within the Pacific and Mississippi flyways, consistent with previous observations (101, 103, 105). All ORF of internal genes of the H14 Guatemalan viruses grouped within the North American lineage, except for the NP gene of two H14N3 viruses collected in 2013 that grouped within the Eurasian lineage.

Thirty-five viruses carried the allele A; while one H14N4 virus (collected in 2015) carried the allele B. All H14 Guatemalan viruses grouped with North American samples within the same allele. Phylogenetic inference by the maximum likelihood method revealed internal genes grouped in subclades within a particular lineage. For the PB2, PB1, PA, and NP sequences, the H14 viruses were grouped into 6 to 7 subclades. Overall, the 36 H14 Guatemalan FLUAVs contained 18 unique internal gene and 21 gene constellations. Even for samples collected on the same day and with identical H14 HA sequences, they contained distinct gene segment constellations, highlighting the high virus diversity found in these viruses in Guatemala.

The HA of H14 Guatemalan FLUAVs contain three different predicted motifs at the HA cleavage site.

It is well accepted that most FLUAVs rely on host's trypsin like proteases for cleavage of the HA into to subunits, HA1 and HA2 to render infectious virus. Among the H14 Guatemalan HA sequences three cleavage site motifs were identified. Sequences

from samples collected in seasons 2010-11 and 2012-13, and one sample collected in season 2018-19 showed the motif PDKQTK'GLF, typical of H14 HA sequences previously identified in North American bird samples (197, 198). H14 viruses contained a Lysine (K) at instead or Arginine (R) in the cleavage site between HA1 and HA2 unlike others viruses (143). Samples collected from late 2013 to mid-2016 (seasons 2013-14, 2014-15, and 2015-16) contained an Arginine at the 1st position of the cleavage motif (PDKQTR'G) and one sample collected during the same period contained an additional Arginine at the first and fourth position (PDRQTR'G).

No mammalian-associated virulence markers were found in the remaining of the predicted ORFs, PB2 (E627K and D701N) (181-184), PA (S409N) (185), or NS1 protein (T92E) (186) nor any antiviral resistance markers (V27A and S31N on M2 (189) and H274Y on NA). Interestingly, 15/36 samples presented an Asparagine (N) in position 66 of PB2-F2 protein, associated with increased virulence in mammals (187, 188). Samples with N66 in PB1-F2 were 87aa long, in contrast to samples with S66 that were 90aa long. PB1-F2 proteins with either one of those lengths have been previously described in FLUAVs from ducks (189). All Guatemalan H14 FLUAVs encoded a full-length PA-X protein of 252 amino acids in length.

Discussion

During our more than a decade long FLUAV surveillance in wild birds in Guatemala, an unusual large number of H14 subtype viruses were detected (99, 101, 103, 105, 106). Prior to its initial isolation in North America in 2010, there was no evidence of significant circulation of this subtype in the Western hemisphere (66). The H14 subtype

FLUAVs appeared dominant in Guatemala since the first detection in 2011 (99, 101, 103), and have been detected ever since as far as in 2019. In this study, we compared all public full-length H14 FLUAV sequences from North America (n=10), Eurasia (n=2), and Guatemala (n=36), to provide clues about their origin in Guatemala.

Due to the limited information about this subtype, its ecology and how it is maintained in this region is unknown. One hypothesis to the high level of H14 detected in Guatemala may be due to epizootic events originated from a single introduction, followed by local clonal expansion of the newly introduced subtype that later are maintained in local populations (202). Studies of H3 subtype FLAUV epizootic events in mallards suggest that migrant birds play an important role as vectors of novel strains that are introduced and amplified by local bird populations (202). It is feasible that H14 FLUAVs are maintained in susceptible hosts during non-migration seasons, and later transmitted during migration season by a routinely sampled host, such as BWTE. H14 viruses from Guatemala were exclusively detected from hunter harvested BWTE (see Chapter 4). Therefore, our sampling is biased towards game bird species that are in large numbers. BWTE is one of the most abundant species of dabbling ducks in North America and has been found in breeding grounds from North American to wintering grounds in Mexico, Central America and parts of South America (203). Due to the wide distribution BWTE across the American continent, it is intriguing as to why the H14 subtype has become predominant only Central America. We believe that yet to be defined local conditions in Guatemala may have provided the adequate environment for this subtype to thrive. Such conditions explain in part the detection of H14 in Guatemala from 2011 to 2015, where all H14 HAs shared the same common ancestor. However, the distinct NA subtype

combinations and internal gene segment constellations of Guatemalan H14 viruses within and between migratory seasons indicate active reassortment and independent introductions from unknown sources. Phylogenetic analysis revealed that Guatemalan H14 viruses have gene segments closely related to those detected during an HPAI outbreak in Mexico in 2017 (179), highlighting the importance to closely monitor avian species since they can be precursors of HPAI viruses (204).

When the H14 subtype was initially sequenced, one feature that stood out was the presence of K instead of R at the -1 position in the HA cleavage site. Thus, previously described H14 HA showed the PDKQTK´G motif which is unusual among FLUAV subtypes. Our continuous surveillance and sequencing approach directly from swabs revealed two additional cleavage site motifs in the H14 HAs: PDKQTR´G, and PDRQTR´G, which place them closer to the more typical cleavage sites observed in other HA subtypes. It is commonly accepted that the HA cleavage is predicted to be a substrate for trypsin-like proteases found in the lumen of the intestinal tract of natural FLUAV hosts (197, 198). Cleavage of HA is necessary for virus infection that modulates virulence. At this stage it is unclear how these different cleavage site motifs in the HA affect replication, virulence and/or transmission of H14 viruses in natural and/or in potential accidental hosts (poultry, mammals) (205). It is also unclear whether our ability to detect different motifs (101) is related to sequencing directly from swabs as opposed to sequencing from virus isolates obtained through passage in chicken eggs (an artificial substrate with potential bottleneck effects). Overall, continued detection of H14 FLUAVs in Guatemala appears dictated by potential multiple drivers in the avian reservoir that warrant further studies.

Figure 5.1 Pairwise percentage of identity of H14 full-length genomes isolated from Guatemala (GT, n=36), North America (NA, n=10) and Eurasia (EU, n=2) during 1982-2019. Numbers 1-36 correspond to isolates from Guatemala, 37-45 to North American isolates, and 47-48 Eurasian isolates. Geographic location is indicated by a dotted square. The gradient background color represents the increase in pairwise nucleotide distances (smallest distance in yellow to the largest in dark red). A complete description of H14 viruses can be found in Table 5.1).

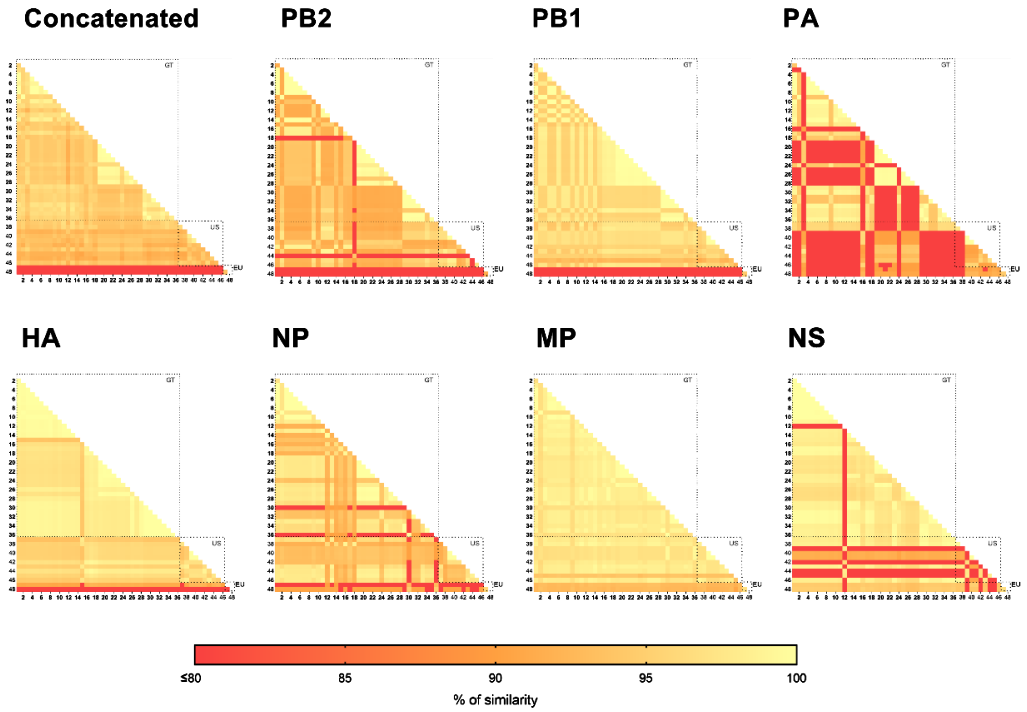

Figure 5.2. Time-scaled MCC tree inferred for H14 Guatemalan isolates collected during 2011-2019 (top) and phylogenies inferred using maximum likelihood methods for N3, N4 and N5 (bottom). The NA trees are midpoint rooted. Scale bar on the bottom-left indicates number of nucleotide substitutions per site. Geographic location and year of collection is indicated by color. HA cleavage motifs is shown for each isolate.

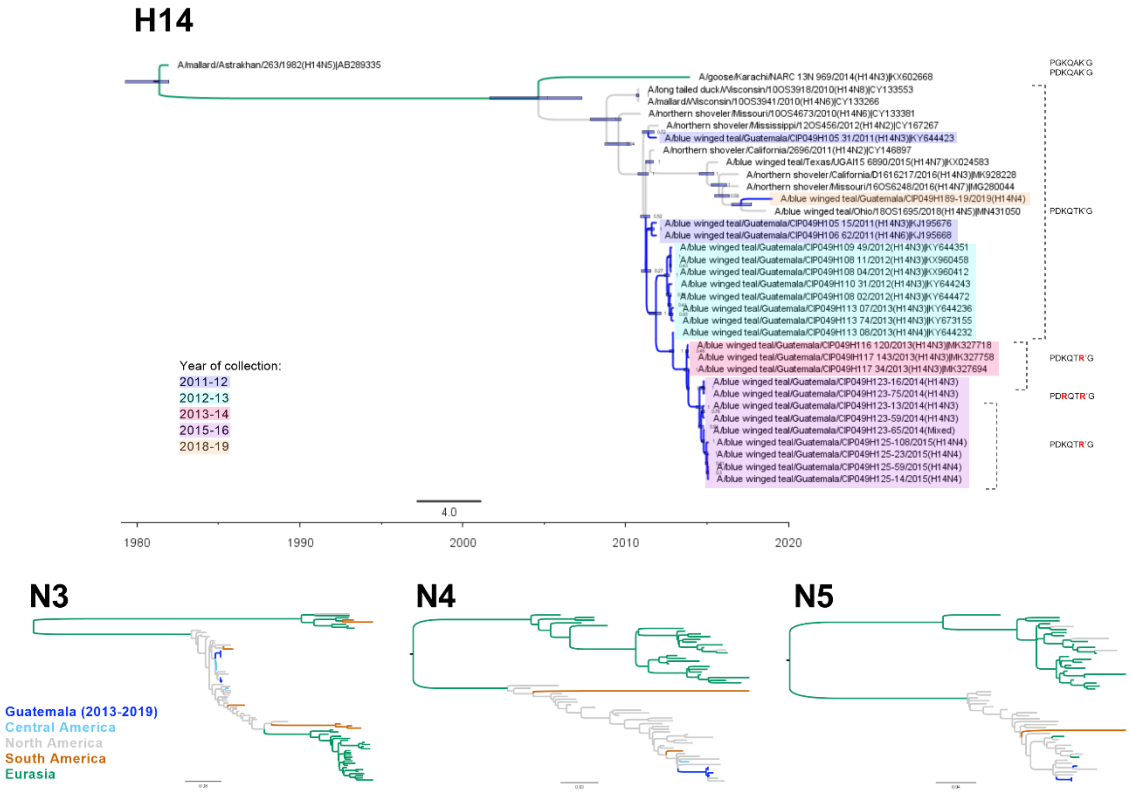

Figure 5.3. Genomic constellations of H14 Guatemalan isolates collected during 2011-2019. A total of 36 viral sequences from Guatemala collected from 2008-2019 were included in the analysis. Geographic location is indicated by color. Different clades with >95% similarity are shown with numbers.

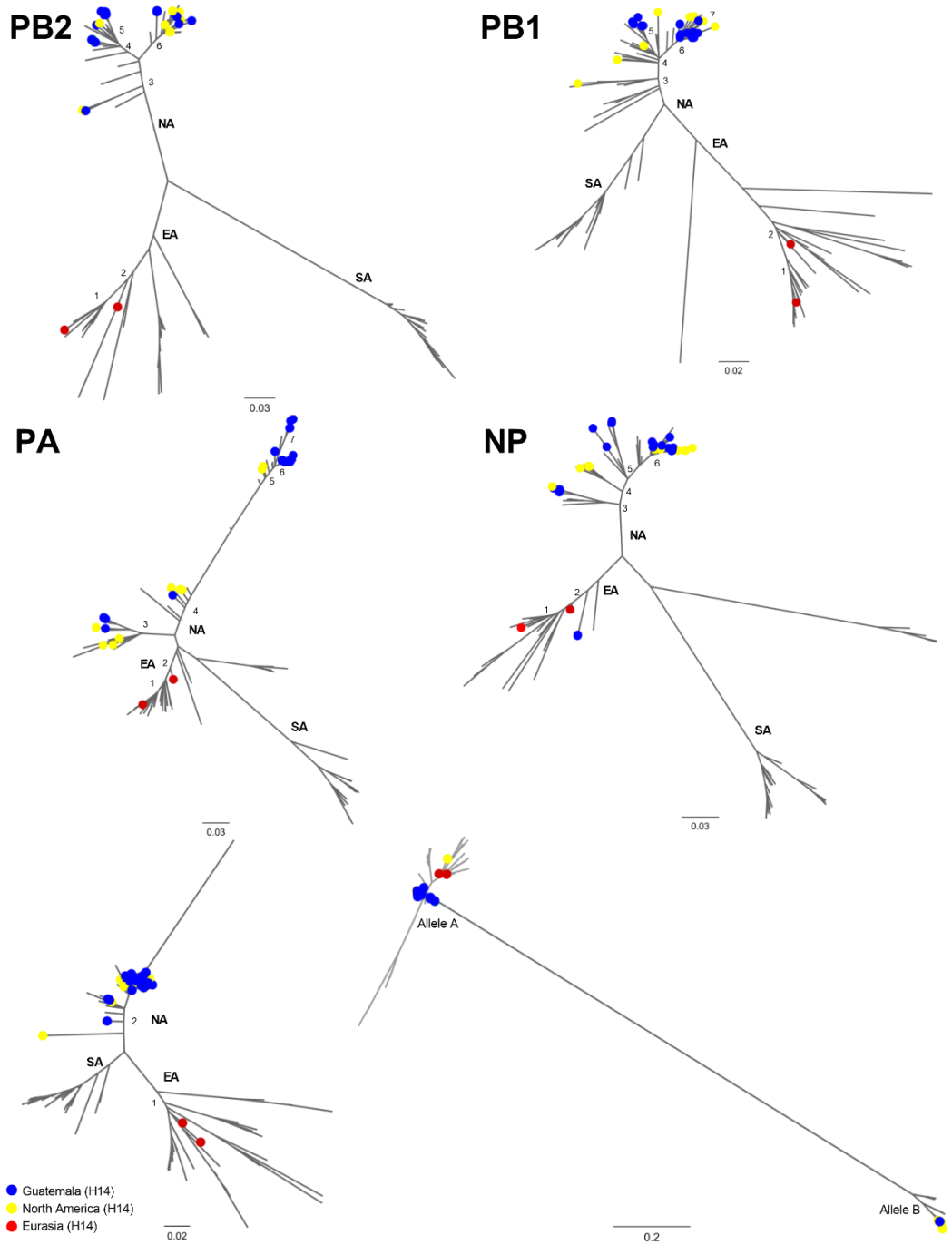

Table 5.1 Full-genome H14 strains available on public databases and newly H14 virus sequences from 2013-2019 used for Chapter 5.

Virus number	Strain	Region
1	A/blue-winged teal/Guatemala/CIP049H123-13/2014(H14N3)	Guatemala
2	A/blue-winged teal/Guatemala/CIP049H123-16/2014(H14N3)	Guatemala
3	A/blue-winged teal/Guatemala/CIP049H123-33/2014(H14N3)	Guatemala
4	A/blue-winged teal/Guatemala/CIP049H123-57/2014(H14N3)	Guatemala
5	A/blue-winged teal/Guatemala/CIP049H123-58/2014(H14N3)	Guatemala
6	A/blue-winged teal/Guatemala/CIP049H123-59/2014(H14N3)	Guatemala
7	A/blue-winged teal/Guatemala/CIP049H123-61/2014(H14N3)	Guatemala
8	A/blue-winged teal/Guatemala/CIP049H123-65/2014(Mixed)	Guatemala
9	A/blue-winged teal/Guatemala/CIP049H123-75/2014(H14N3)	Guatemala
10	A/blue-winged teal/Guatemala/CIP049H123-76/2014(Mixed)	Guatemala
11	A/blue-winged teal/Guatemala/CIP049H125-14/2015(H14N4)	Guatemala
12	A/blue-winged teal/Guatemala/CIP049H125-23/2015(H14N4)	Guatemala
13	A/blue-winged teal/Guatemala/CIP049H125-59/2015(H14N4)	Guatemala
14	A/blue-winged teal/Guatemala/CIP049H125-108/2015(H14N4)	Guatemala
15	A/blue-winged teal/Guatemala/CIP049H189-19/2019(H14N4)	Guatemala
16	A/blue-winged teal/Guatemala/CIP049H105-15/2011(H14N3)	Guatemala
17	A/blue-winged teal/Guatemala/CIP049H105-31/2011(H14N3)	Guatemala
18	A/blue-winged teal/Guatemala/CIP049H106-62/2011(H14N6)	Guatemala
19	A/blue-winged teal/Guatemala/CIP049H108-02/2012(H14N3)	Guatemala
20	A/blue-winged teal/Guatemala/CIP049H108-04/2012(H14N3)	Guatemala
21	A/blue-winged teal/Guatemala/CIP049H108-11/2012(H14N3)	Guatemala
22	A/blue-winged teal/Guatemala/CIP049H109-49/2012(H14N3)	Guatemala
23	A/blue-winged teal/Guatemala/CIP049H109-76/2012(H14N3)	Guatemala
24	A/blue-winged teal/Guatemala/CIP049H110-31/2012(H14N3)	Guatemala
25	A/blue-winged teal/Guatemala/CIP049H113-07/2013(H14N3)	Guatemala
26	A/blue-winged teal/Guatemala/CIP049H113-08/2013(H14N4)	Guatemala
27	A/blue-winged teal/Guatemala/CIP049H113-74/2013(H14N3)	Guatemala
28	A/blue-winged teal/Guatemala/CIP049H113-76/2013(H14N4)	Guatemala
29	A/blue-winged teal/Guatemala/CIP049-H116-120/2013(H14N3)	Guatemala
30	A/blue-winged teal/Guatemala/CIP049-H117-34/2013(H14N3)	Guatemala
31	A/blue-winged teal/Guatemala/CIP049-H117-36/2013(H14N5)	Guatemala
32	A/blue-winged teal/Guatemala/CIP049-H117-38/2013(H14N5)	Guatemala
33	A/blue-winged teal/Guatemala/CIP049-H117-99/2013(H14N5)	Guatemala

Virus number	Strain	Region
34	A/blue-winged teal/Guatemala/CIP049-I-H117-13/2013(H14N3)	Guatemala
35	A/blue-winged teal/Guatemala/CIP049-I-H117-123/2013(H14N3)	Guatemala
36	A/blue-winged teal/Guatemala/CIP049-I-H117-143/2013(H14N3)	Guatemala
37	A/blue-winged teal/Ohio/18OS1695/2018(H14N5)	USA
38	A/blue-winged teal/Texas/UGAI15-6890/2015(H14N7)	USA
39	A/long-tailed-duck/Wisconsin/10OS3918/2010(H14N8)	USA
40	A/long-tailed-duck/Wisconsin/10OS4225/2010(H14N6)	USA
41	A/mallard/Wisconsin/10OS3941/2010|1|11/18/2010(H14N6)	USA
42	A/northern-shoveler/California/2696/2011(H14N2)	USA
43	A/northern-shoveler/California/D1616217/2016(H14N3)	USA
44	A/northern-shoveler/Mississippi/12OS456/2012(H14N2)	USA
45	A/northern-shoveler/Missouri/10OS4673/2010(H14N6)	USA
46	A/northern-shoveler/Missouri/16OS6248/2016(H14N7)	USA
47	A/goose/Karachi/NARC-13N-969/2014(H14N3)	Pakistan
48	A/mallard/Astrakhan/263/1982(H14N5)	Russia

Acknowledgements

This study was supported by a contract (HHSN272201400008C) from the National Institute of Allergy and Infectious Diseases (NIAID) Centers for Influenza Research and Surveillance (CEIRS).

CHAPTER 6

INDEPENDENT EVOLOTION OF INFLUENZA A H1N1 IN PIGS FROM GUATEMALA

Lucia Ortiz; Ginger Geiger, Lucas Ferreri, David Moran, Dione Méndez, Ana Silvia Gonzalez-Reiche, Daniela Rajao, Celia Cordon-Rosales, and Daniel R. Perez

Submitted to *Journal of Virology*.

Abstract

Commercial swine farms provide unique systems for interspecies transmission of influenza A viruses (FLUAVs) at the animal-human interface. Bidirectional transmission of FLUAVs between pigs and humans plays a significant role in the generation of novel strains that become established in the new host population. Active FLUAV surveillance was conducted for two years on a commercial pig farm in Southern Guatemala with no history of FLUAV vaccination. Nasal swabs (n=2,094) from fattening pigs (6 to 24 weeks old) with respiratory signs were collected from May 2016 to February 2018. Swabs were screened for FLUAV by RRT-PCR and full virus genomes of FLUAV-positive swabs were sequenced by next-generation sequencing (NGS). FLUAV prevalence was 12.0% (95% CI: 10.6% – 13.4%) with two distinct periods of high infection. All samples were identified as FLUAVs of the H1N1 subtype within the H1 swine clade 1A.3.3.2 and whose ancestors are the human origin 2009 H1N1 influenza pandemic virus (H1N1 pdm09). Compared to the prototypic reference segment sequence, 10 amino acid signatures were observed on relevant antigenic sites on the hemagglutinin. We also found that Guatemalan swine-origin FLUAVs show independent evolution from other H1N1 pdm09 FLUAVs circulating in Central America. The zoonotic risk of these viruses remains unknown, but strongly calls for continued FLUAV surveillance in pigs in Guatemala.

Introduction

Influenza A viruses (FLUAV) infect a wide range of avian and mammalian hosts, including humans. The virus genome is composed of 8 segments of negative single-stranded RNA corresponding to 6 internal (PB2, PB1, PA, NP, M, and NS) and 2 surface (HA and NA) gene segments. Zoonotic FLUAV infections are relatively common and it is accepted that influenza pandemics have resulted from zoonotic FLUAV strains (206). FLUAV infections in swine have a significant economic impact on swine production due to losses caused by the disease. Swine farms present high animal density and provide an environment for close contact between animals and humans. Such environments facilitate the initial spillover between species and transmission (207). Interspecies transmission events of FLUAVs between humans and pigs play a significant role in the generation of novel reassortant strains that transmit among humans and/or swine populations (208). Recent studies suggest that the introduction of human-origin FLUAVs into pigs is a major driver in the independent evolution of FLUAV lineages depending on geographic location (111). The emergence of the 2009 H1N1 influenza pandemic virus (H1N1 pdm09) in Latin America with potential undetected circulation in pigs for several years undetected (15), highlights the need to improve influenza surveillance in these understudied regions to timely detect strains of zoonotic and/or pandemic concern.

In Central America, Guatemala is the country with the largest swine industry. Most of the swine production in Guatemala occurs in farrow-to-finish systems. According to the last census in 2021 by the Ministry of Agriculture, Livestock, and Food of Guatemala, the swine population was estimated to be 294,479 pigs, of which 32,518 were breeding stock (209). Circulation of FLUAV of human origin has been documented

previously in swine populations in Guatemala (140); however, only limited genetic data is available. In this study we conducted active surveillance to better understand the epidemiology of FLUAV in swine populations in commercial pig farms in Guatemala (in the absence of vaccination). Our results show that since its introduction in 2009, pdm09-like viruses disseminated and became the predominant subtype in Guatemalan pigs at the time of sampling (2016-2018). Phylogenetic analyses revealed that these viruses cluster with FLUAVs that circulated in 2009 and share nucleotide identities of >97%, representing a unique group of FLUAVs whose zoonotic risk remains to be determined.

Materials and Methods

Methods used for this Chapter are described in detail in Chapter 3.

Results

Sex, age, and rectal temperature were associated with FLUAV positivity in pigs

From May 2016 to April 2018, we performed active FLUAV surveillance in commercial farrow-to-finish farm located in Palin, Escuintla, ~60 km away from the Southern Pacific coast of Guatemala (Figure 6.1A). The farm production is divided in boars, sows, replacements, piglets, and growing areas, located in different facilities on the farm. Artificial insemination, gestation, and farrowing are performed routinely in the farm. Boars, sow, and replacements are in the same facility shared with quarantine, insemination, gestation, replacements, and maternity areas. Piglets are weaned at four weeks of age. After weaning, pigs are moved to a different area where they spend 6 weeks (from week four to week ten). Afterwards pigs are moved to finish stage, where

they spend approximately twelve more weeks (from week ten to week twenty-two) to get ready for market. The production cycle of each pig takes approximately 22 weeks. For the past 10 years, the farm reported a history of FLUAV exposure.

The survey included 31,093 pigs associated to 3,113 cases of respiratory disease. Nasal swabs tested positive for FLUAV by RRT- PCR in 251 out 2,094 samples; with an estimated prevalence of 12.0% (95% CI: 10.6% – 13.4%). Two periods of high FLUAV infection were detected, one from May 2016 to February 2017 and another from January 2018 to April 2018 with prevalence of 18.7% (95% CI: 16.4% – 21.2%) and 17.8% (95% CI: 14.2% – 22.1%), respectively (Figure 6.1B). Between these two periods, only one FLUAV-positive sample was detected in September 2017. The mean age of the pigs in the farm was 10.5 ± 6.6 weeks. FLUAV was detected in pigs as young as 5 weeks, and as old as 21 weeks. The mean age of FLUAV-positive pigs was 8.0 ± 2.4 weeks. Factors such as age [weaning pigs OR95%=4.1 (1.3, 20.7)], sex [male OR95%=1.4 (1.0, 1.8)] and rectal temperature [fever OR95%=2.7 (2.0, 3.6)] were associated with FLUAV positivity (Table 1). Borderline association was found for animal density, 201-300 animals per pen [OR95%=2.0 (1.0, 5.0)] and FLUAV detection. All animals presented coughing at the time of sampling.

Guatemalan swine-origin FLUAVs show independent evolution from other H1N1 pdm09 FLUAVs circulating in Central America

From the 251 FLUAV positive samples, 157 (62.5%) amplified at least one gene segment by multi-segment RT-PCR (MS-RT-PCR). These samples were subsequently sequenced by NGS, producing 57 complete genomes (53 from the 2016-17 period and 4

from the 2018 period). The total number of full-length sequences by segment and the number of unique Open Reading Frames (ORF) sequences are shown in (Table 2). To build the phylogenetic trees, only one representative from identical sequences was used resulting in 273 unique nucleotide sequences (225 from the 2016-17 period and 48 from the 2018 period).

Phylogenetic analyses of all gene segments show 2 clusters, one from samples collected from May 2016 to February 2017 and the other from samples collected in 2018. Within each cluster, all gene segments showed >99% sequence identity. Both clusters share a common ancestor derived from the H1N1 pdm09 FLUAV lineage. The HA segments of all samples belonged to the FLUAV H1 swine clade 1A.3.3.2 derived from the H1N1 pdm09 FLUAV lineage using the "swine H1 clade classification tool" of the Influenza Research Database (IRD) (210).

The HA and NA phylogenies of the Guatemalan FLUAVs show clear separation from other contemporary human H1N1 pdm09 FLUAVs and all swine FLUAVs detected in Central America whose sequences are available in the IRD and GISAID databases (Figure 6.2). The HA and NA gene segments of Central American human origin FLUAV isolates from 2009 are the most closely related ancestors of the HA and NA segments of the swine FLUAVs from Central America even though the samples in this study were collected during 2016-2018. It is estimated that these viruses were introduced early during the A(H1N1)pdm09 pandemic (~2010.8 [2009.4-2012.7 95% HPD] for HA and 2010.4 [2009.1-2012.3 95% HPD] for NA).

Phylogenetic analyses of the polymerase genes (PB2, PB1 and PA) showed similar clustering within H1N1 pdm09 FLUAV identified in humans in 2009 (Figure

6.3). NP and M1 genes from Guatemalan swine IAVs clustered with FLUAVs circulating in the Americas in swine and humans in 2009 respectively. Interestingly, the NS gene segment was shown in two clusters, those from the 2016-17 period associated with NS sequences of swine FLUAVs identified in Japan in 2017 and those from the 2018 period associated with swine FLUAVs circulating in Thailand between 2013-2017, consistent with possibly two independent introductions (Figure 6.4).

Amino acid signatures on the HA and NA surface proteins of Guatemalan swine origin H1N1 FLUAVs

Amino acid differences in HA observed in swine H1N1 pdm09 FLUAVs from Guatemala were compared to the reference strain A/California/04/2009 (H1N1) (Ca04, Figure 6.5 and Figure 6.8). The HA ORF encoded the same cleavage site sequence indistinguishable from the Ca04 strain (PSIQSR'GLF). In the rest of the HA ORF, up to 39 amino acid differences were observed compared to the Ca04 reference sequence (Figure 6.8), 25 of those on the HA1 globular head with 10 falling within antigenic sites Cb (S71F), Ca2 (P137S, H138Y, A141T), Ca1 (G170E, R205K, E235G) and Sb (S185N, S190T, A195S) (H1 numbering, mature protein). Minor variant frequency analysis revealed other mutations on the HA1 globular head in different animals (A141E, L161P, T184A, T184I, D187G, Q188R and E235K) with allele frequency ranging from 2.3% to 36.3% in at least 8 viruses from the 2016-2017 period.

Most amino acid differences in the HA1 region occurred at high frequency (>0.95) (Figure 6.6). Five of these amino acid positions were in common among the HA ORFs from the two high FLUAV infection periods (P83S, T197A, R205K in antigenic

site Ca1, I321V and D346E). The HA R45Q and the HA R45K were fixed in viruses from the 2016-2017 and 2018 periods, respectively (Figure 6.8). The HA L32V was the most common signature in viruses from both infection periods (>100 HA sequences), except for one virus each from 2016 and 2018 (L32M and L32I, respectively). S71F (Cb), P137S (Ca2) were fixed in viruses from the 2016-2017 and A141T and S190T in the 2018 period.

Amino acid signatures elsewhere in predicted ORFs of swine FLUAVs from Guatemala.

In the NA ORF, 36 amino acids differences were found compared to the N1 NA from the Ca04 reference strain (Figure 6.8), but most were found fixed in NA sequences from the 2018 period (18 amino acid signatures in 35 sequences). In contrast, only 5 amino acid signatures were identified in >90% out 101 sequences in the NA sequences from the 2016-2017 period. Only the Y100H signature was in common among sequences between the two periods, whereas the NA K331R and NA K331N were fixed in strains from the 2016-2017 and 2018 periods, respectively. None of these differences were present in either in the catalytic site (118, 151, 152, 224, 276, 292, 371, and 406, N2 numbering) or the framework residues (119, 156, 178, 179, 198, 222, 227, 274, 277, 294, and 425, N2 numbering) (211), and no amino acid changes in known drug binding sites were found.

In the internal gene segments, eight substitutions that appeared in both periods were found in PB1, PA, NP and M2 (Figure 6.7 and Figures 6.8-6.11). Compared to the Ca04 reference, the PB1 fixed mutations I179M, K353R, and N455D are located on predicted RNA-dependent RNA polymerase motifs. PA presented fixed mutations in the

NLS region (P224S) and the PB1 binding domain (Y650F), while NP showed a mutation overlapping the PB2 binding and RNA binding domains (D53E). M2 showed two fixed mutations found in both periods, located at amino acid positions 55 and 61 (F55I and R61K). All swine origin FLUAV Guatemalan viruses presented a truncated form of the PB1-F2 protein of 11 aa, due to premature stop codons at positions 12 and 58, in common with the Ca04 reference strain (212). All swine origin FLUAV Guatemalan viruses encode the PA-X functional protein (232 aa), as well as the ORFs for PA-N155 (562 aa) and PA-N182 (535 aa). No alternative fixed mutations were found in PA-X in either period compared to Ca04. In contrast, PA-N155 and PA-N182 ORFs show two mutations, P70S/Y498F and P43S/Y469F, respectively (Figure 6.11).

No mammalian-associated virulence markers in PB2 (E627K, D701N), PB1-F2 (N66S), NS (S42P, D92E and V149A), and M2 (V27A) were found for any of the analyzed sequences. All swine origin FLUAV Guatemalan viruses showed the PA S409N signature predictive of increased virulence and the M2 S31N marker of amantadine resistance like other Eurasian swine lineage M segments (185, 213, 214).

Discussion

Close contact between humans and pigs in swine production systems may result in bidirectional transmission of different FLUAVs between the two species. Systematic surveillance of FLUAV in swine production farms provides a unique opportunity to study how these viruses may jump and adapt at the human-animal interface and identify novel strains that may become established in the new host population. Based on an extensive FLUAV weekly surveillance in pigs from 2016 to 2018, we detected an overall FLUAV

prevalence of 12% in animals with respiratory disease, similar to previous studies in pigs in Guatemala (140). It must be noted that sick animals were identified and sampled based on observation of coughing (regardless of other signs), which may have resulted in underrepresentation of sampling and consequent underestimation of the FLUAV prevalence found in this study. Younger pigs showed the highest positivity rate of FLUAV and were associated with higher rates of infection, consistent with previous surveillance studies (140, 215). Factors such as sex (males) and fever were also associated with FLUAV positivity.

During the span of 2 years, we detected two high infection periods but not a seasonality pattern. Only the H1N1 subtype from the H1N1 pdm09 lineage was identified, despite other FLUAV subtypes being reported in humans and swine in Guatemala and Central America at the time of the surveillance period (140, 216). National records from Guatemala indicate minimal to zero circulation of H1N1 pdm09 FLUAVs in humans at the time of pig sampling for this study. This observation suggests a prior introduction of H1N1 pdm09 FLUAVs into pigs that remained endemic in the swine population in Guatemala. Interestingly, a sharp reduction in FLUAV-positive samples was observed for approximately 12 months, during most of 2017. This drop in detection could be explained by an increase in prevention and biosafety practices following the first year of the study (annual seasonal influenza vaccination of workers; quarantine of sick pigs; use of hand disinfectant and mask respirators by workers; and enforcement of the shower-in/shower-out procedures), as reported by farm owners. However, the reason for the subsequent increase in detections seen in 2018 remains

unknown, although it is consistent with the notion of introduction of a virus strain through replacement animals brought into the farm.

We obtained sequences from >75% of the FLUAV-positive samples and nearly 25% of complete genomes by NGS directly from the original swabs. This methodology allowed us to improve the number of characterized samples and reduce potential selection bias introduced by virus isolation prior to sequencing (101). Interestingly, all samples presented high number of defective RNAs, as it was observed in agarose gels and later by the valley-like shape in the graphs of the polymerase genes (PB2, PB1, and PA) of the NGS coverage read maps (Figure 6.12). These particles are truncated forms of FLUAV generated by most viruses during virus replication that retain the terminal sequences necessary for virus packaging (217) and are mostly present in the polymerase segments (218, 219). The function is not fully understood; but it is hypothesized that they could play a role in maintaining low levels of replication of infectious virus (220).

Phylogenetic analyses showed that the swine origin FLUAV Guatemalan viruses are clearly segregated from other H1N1 pdm09 viruses in the Americas, suggesting independent evolution of these viruses after introduction and subsequent circulation in pigs, consistent with reports in other regions (208). At least two independent introductions were observed as noted with the separated clusters of Guatemalan samples. Interestingly, the viruses seemed to have been introduced in 2010 during the H1N1 pdm09 pandemic and persisted as separate clades for ~6-8 years before detected by our surveillance. Long persistence sustained prevalence of human-origin FLUAVs in environments with significant interaction between two FLUAV hosts might lead to the generation of strains of pandemic concern.

Little human seasonal influenza sequence data is available from Guatemala and Central America and much less swine FLUAV sequence data exists. This gap in part explains the results of the phylogenetic relationships with the most closely related ancestors. Independent virus evolution events have been documented previously in pigs, particularly after introduction of human FLUAVs, showing long phylogenetic branches between the swine strains and their putative human FLUAV ancestor, as shown previously from other Latin American countries (111, 128). Here, we describe the circulation and evolution of H1N1 pdm09 lineage viruses in Guatemala that may represent the establishment of a novel genetic lineage with the potential to reassort with cocirculating viruses. The mutations found in relevant HA1 antigenic sites may lead to differences in antigenic relationships with other H1N1 pdm09 viruses of human- or swine-origin; however, the effect of these mutations and zoonotic risk remains to be determined. These observations highlight the need for increased and sustained influenza surveillance within understudied regions.

Figure 6.1. Location of study site (A) and number of collected, FLUAV-positive, and sequenced samples per week (B) during a two-years FLUAV surveillance in a commercial farm in Guatemala. Sampled site represents the largest swine-producing region in the country with history of FLUAV circulation. Red dotted line indicates the minimum number of samples required per week to achieve the desired sample size. Two high infection periods were detected by RRT-PCR one from May 2016 to February 2017, and another one from January 2018 to April 2018.

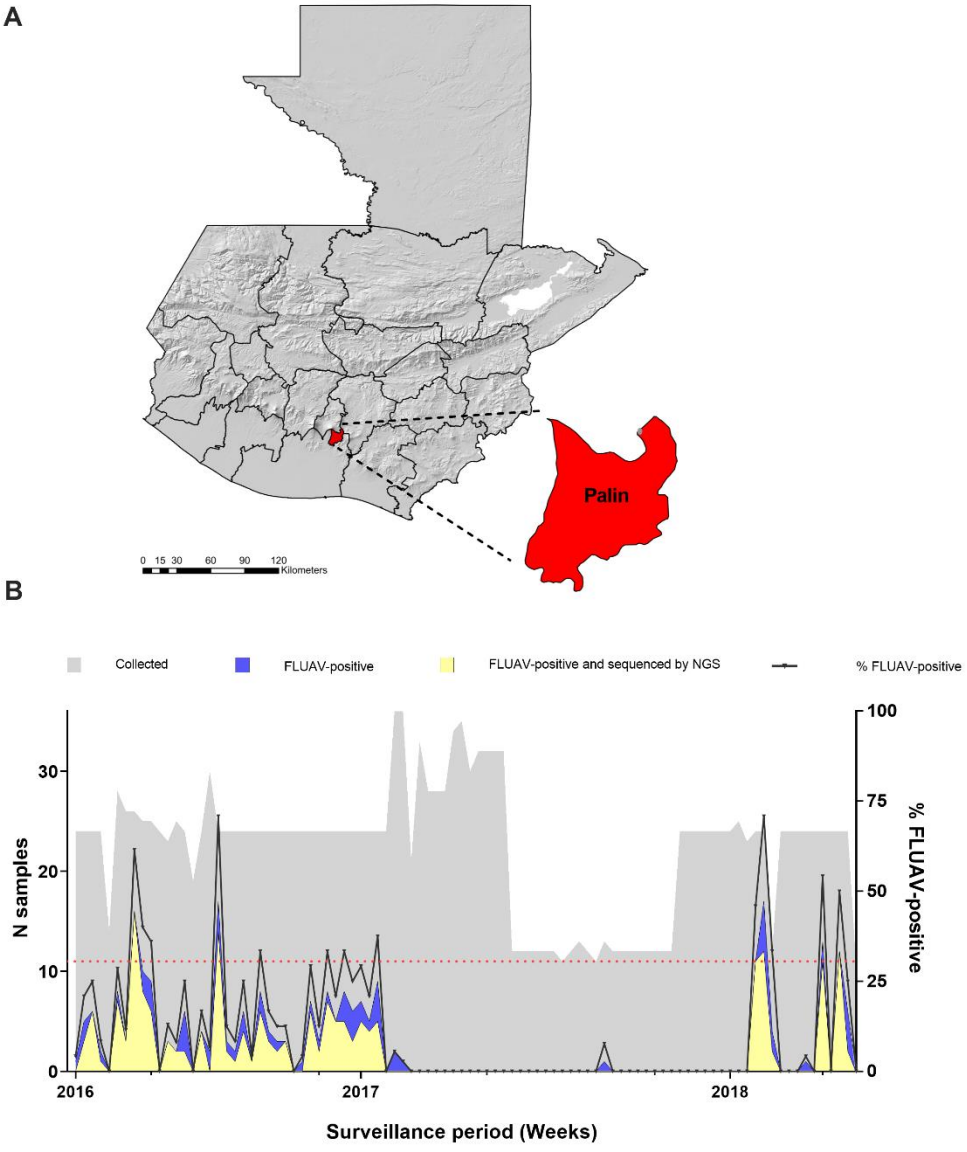

Figure 6.2. Surface genes (H1 and N1) time scaled MCC trees inferred for H1N1 pdm09 from swine and human coding sequences (2008-2019) for 273 (H1) and 321 (N1) sequences. Reference strain Ca04 is highlighted with a star symbol. Identical sequences of Guatemalan viruses were removed. The swine viruses from Guatemala are marked with a blue circle (May 2016 to February 2017 cluster) and a green rectangle (2018 cluster).

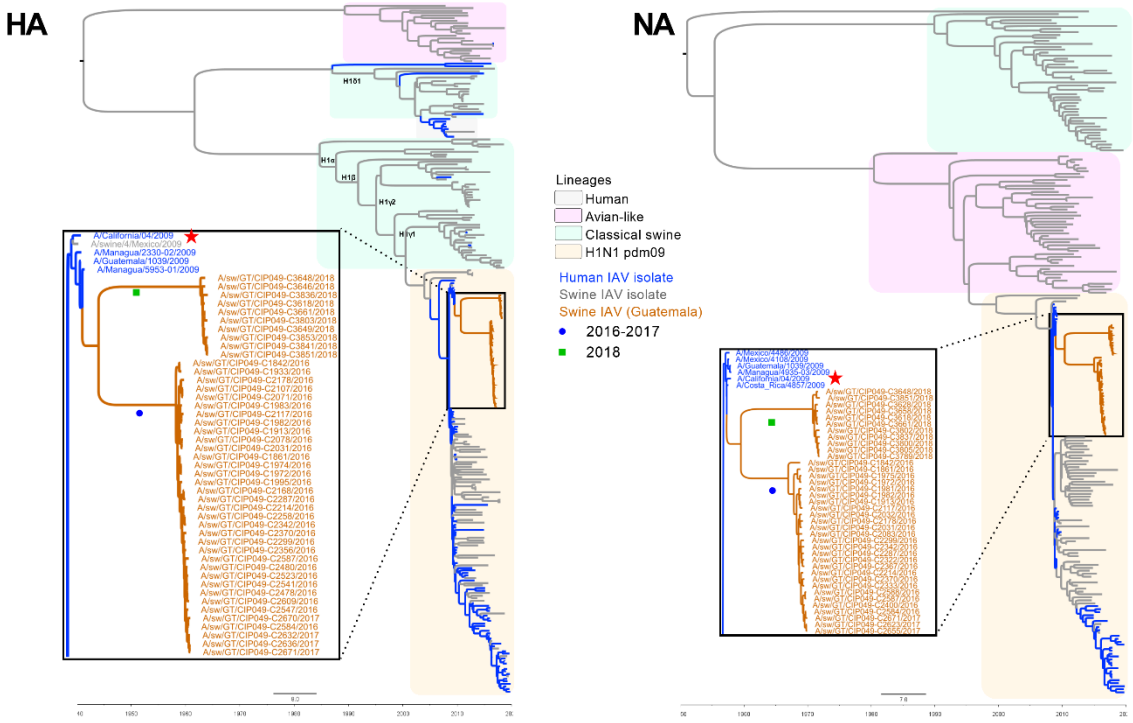

Figure 6.3 Internal genes (PB2, PB1, PA) phylogenetic inference for H1N1 pdm09 from swine and human coding sequences (2008-2019) for 390 (PB2), 378 (PB1) and 399 (PA) sequences. Maximum Likelihood phylogenetic inference using the best-fit model. Reference strain Ca04 is highlighted with a star symbol. Identical sequences of Guatemalan viruses were removed. The swine viruses from Guatemala are marked with a blue circle (May 2016 to February 2017 cluster) and a green rectangle (2018 cluster).

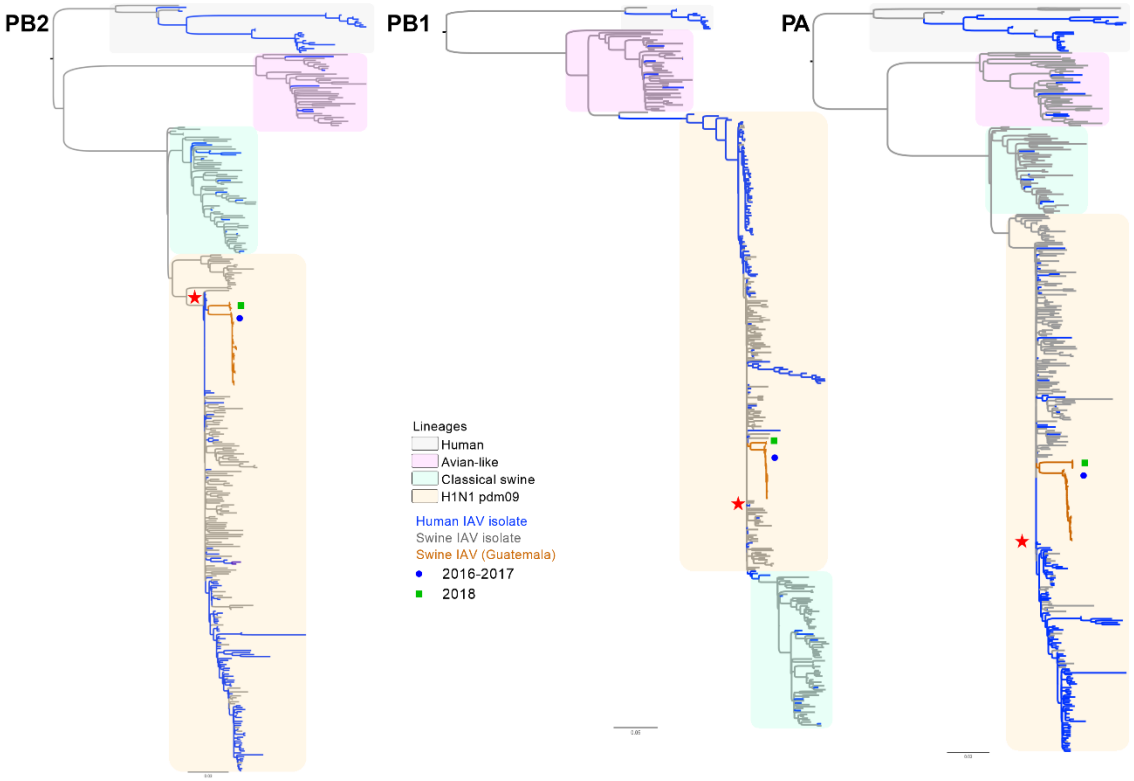

Figure 6.4 Internal genes (NP, M1, and NS1) phylogenetic inference for H1N1 pdm09 from swine and human coding sequences (2008-2019) for 360 (NP), 347 (M1), and 378 (NS1) sequences. Maximum Likelihood phylogenetic inference using the best-fit model. Reference strain Ca04 is highlighted with a star symbol. Identical sequences of Guatemalan viruses were removed. The swine viruses from Guatemala are marked with a blue circle (May 2016 to February 2017 cluster) and a green rectangle (2018 cluster).

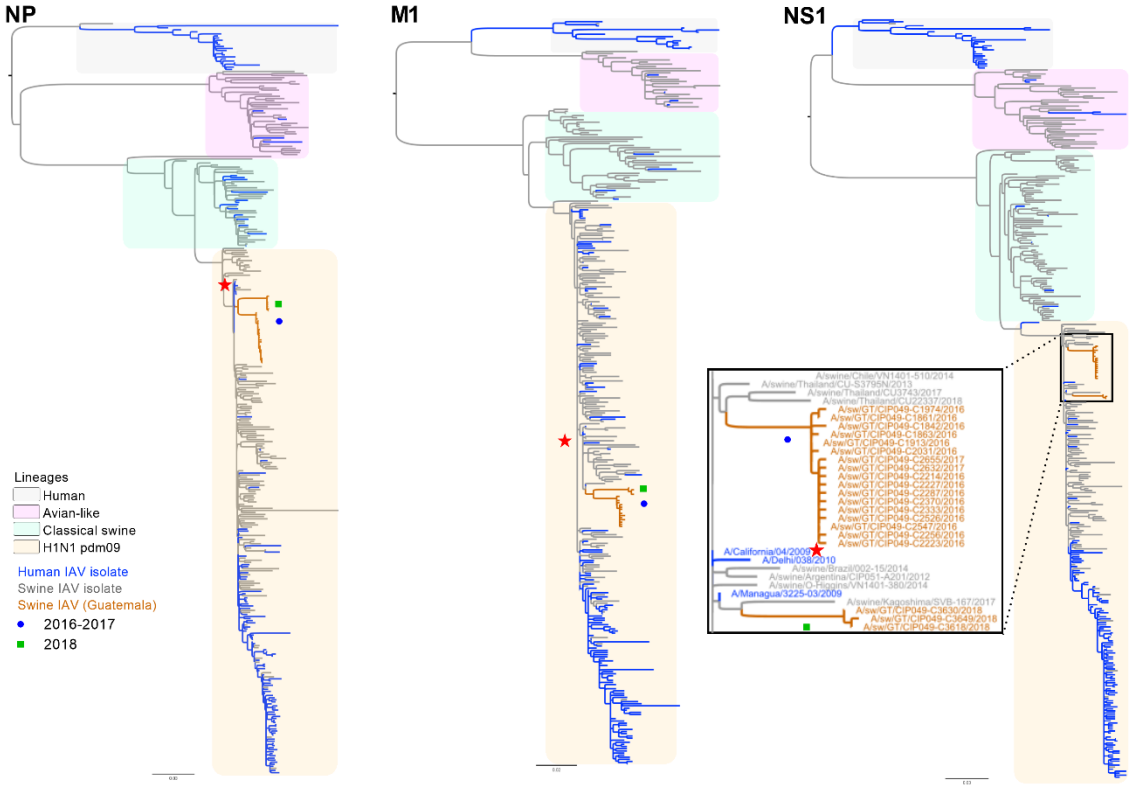

Figure 6.5 Distribution of amino acid mutations in the HA1 domain (H1 numbering without the signal peptide) of unique H1N1 pdm09 swine viruses from Guatemala aligned to Ca04 (GenBank accession no. GQ117044.1). Dots represent amino acids identical to the reference strain. The antigenic sites (Cb, Sa2, Ca1 and Sb) are shown in colored boxes.

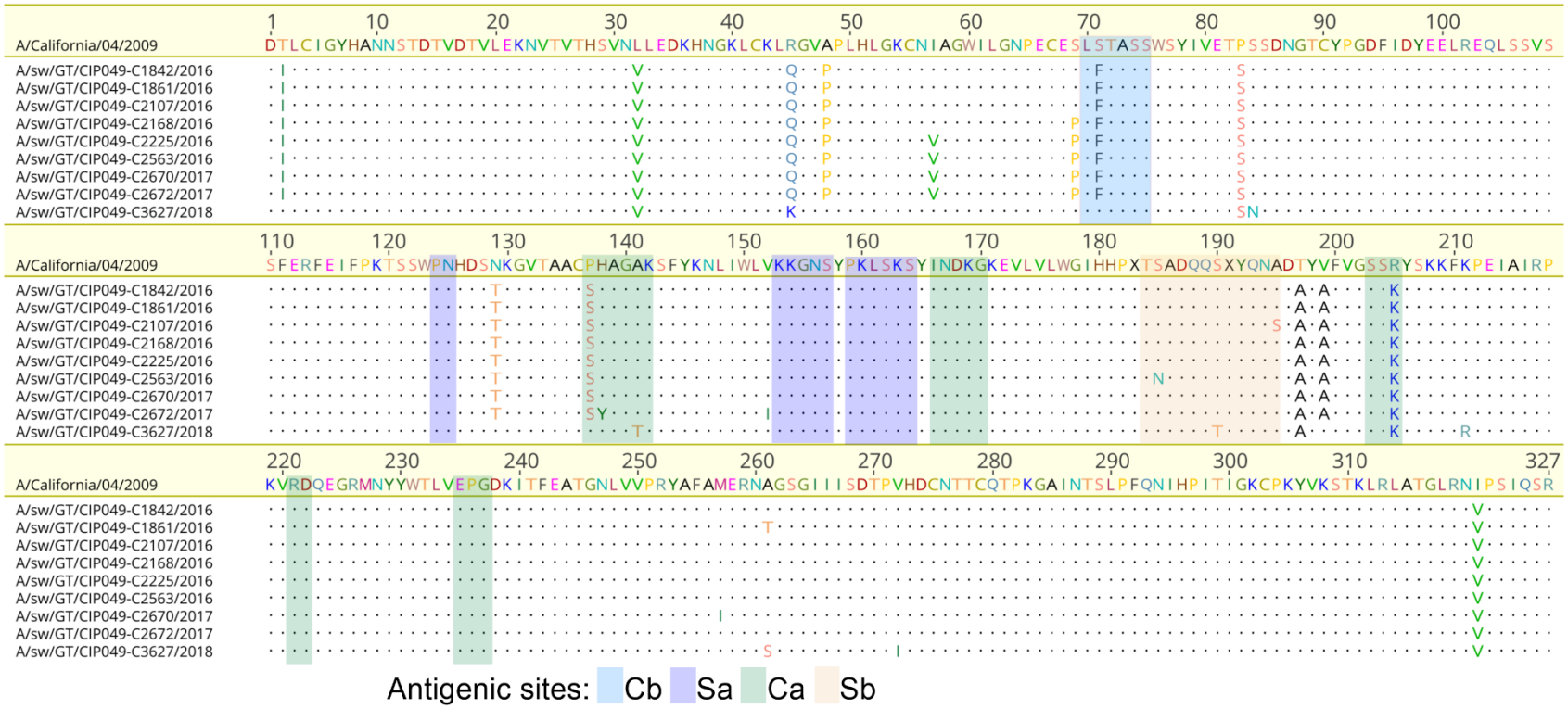

Figure 6.6. Distribution and frequency of amino acid point mutations in proteins of surface genes (HA and NA) from H1N1 pdm09 swine viruses from Guatemala in comparison to the reference strain Ca04 (GenBank accession no. GQ117044.1 and MN371610.1). Known protein domains or important sites of pdm09 haplotypes available in public databases are represented by colored blocks. Amino acid position is shown in the he x-axis.

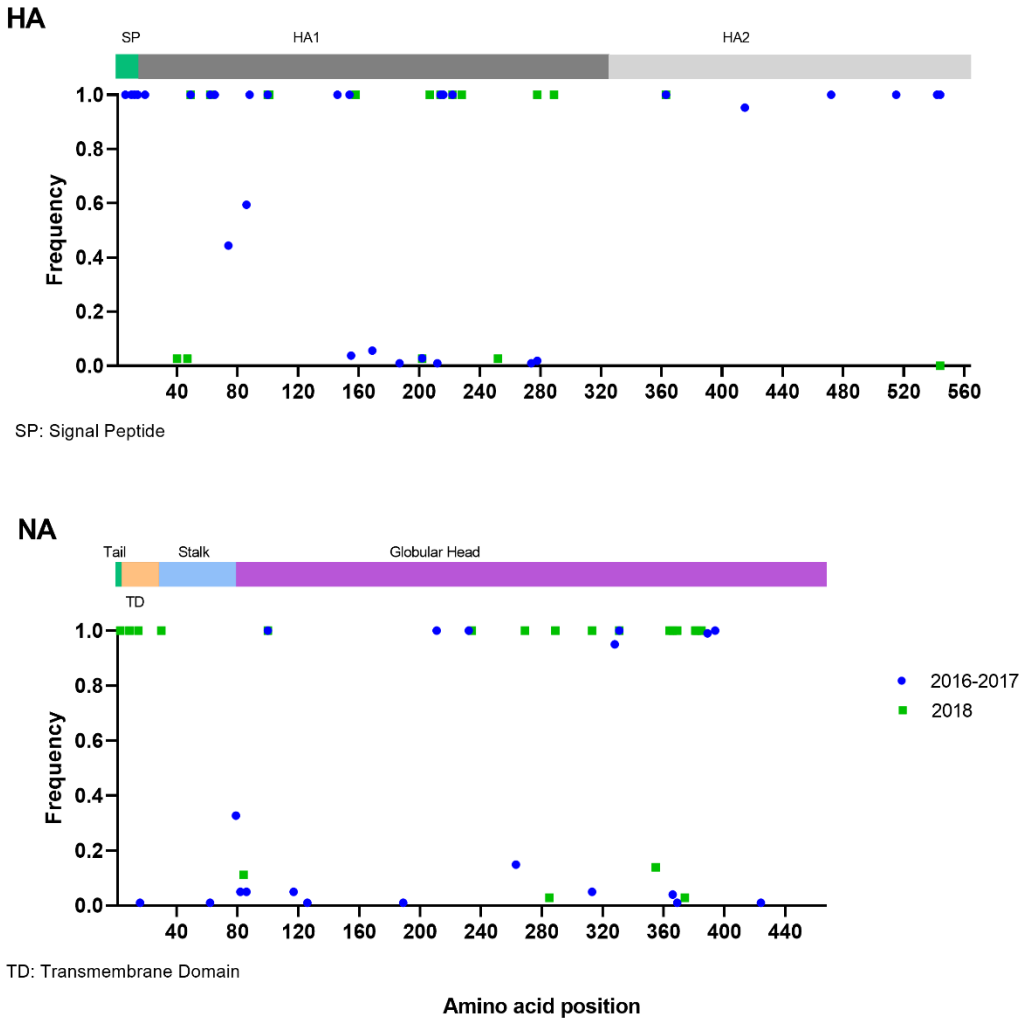

Figure 6.7. Distribution and frequency of amino acid point mutations in PB2, PB1, PA, NP, M1, M2, NS1 and NEP from swine H1N1 pdm09 FLUAVs from Guatemala in comparison to the reference strain Ca04 (GenBank accession no. MN371615.1, MN371613.1, MN371611.1, MN371617.1, FJ969513.1, and FJ969514.1). Known protein domains or important sites of pdm09 haplotypes available in public databases are represented by colored blocks

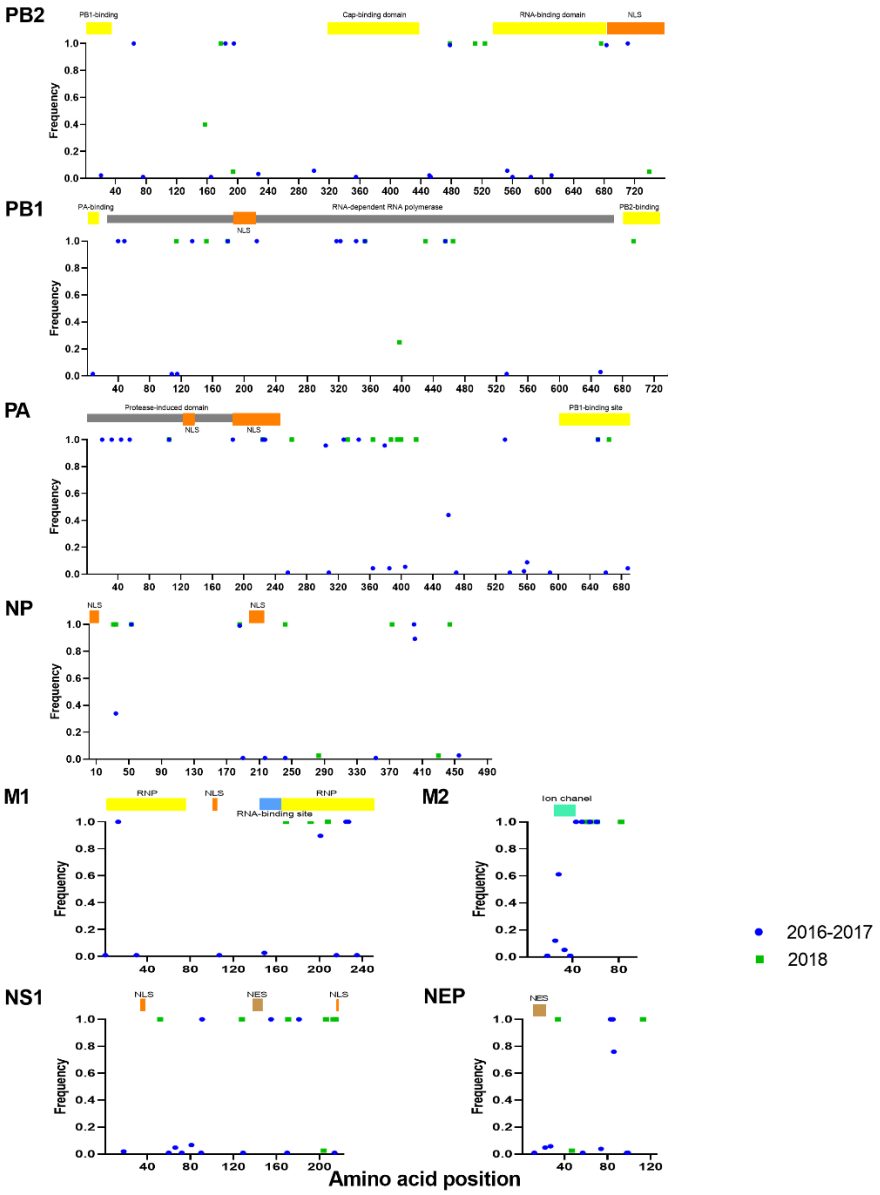

Figure 6.8. HA (H1 numbering) and NA (N2 numbering) unique amino acids of H1N1 pdm09 swine viruses from Guatemala aligned to Ca04 (GenBank accession no. GQ117044.1 and MN371610.1). Dots represent amino acids identical to Ca04. The antigenic sites for HA (Cb, Sa2, Ca1 and Sb) are shown in green boxes; whereas viruses collected during May 2016 to February 2017 are shown in orange and those collected in 2018 in grey.

HA

A/California/04/2009	number of identical amino acid sequences	V	Y	F	T	T	N	V	L	R	A	I	S	S	P	S	N	P	H	A	V	G	S	N	T	V	R	K	E	M	A	V	I	D	N	N	E	V	V
						2	23	30	32	45	48	57	69	Cb 71	83	84	129	Ca2 137	138	141	152	Ca1 170	Sb 185	190	195	197	199	Ca1 205	211	Ca1 235	257	261	272	321	346	398	455	498	525	527
A/Guatomala/1039/2009		.	.	.	.	.	.	.	I	K	.	.	.	F	S	.	.	.	.	.	.	.	.	.	.	A	.	K	.	.	.	.	.	V	.	.	.	.	.	.
A/sw/GT/CIP049-C1842/2016	5	A	N	L	I	I	.	.	V	Q	P	.	.	F	S	.	T	S	.	.	.	.	.	.	.	A	A	K	.	.	.	.	.	V	E	.	T	K	I	I
A/sw/GT/CIP049-C1861/2016	2	A	N	L	I	I	.	.	V	Q	P	.	.	F	S	.	T	S	.	.	.	.	.	.	.	A	A	K	.	.	.	T	.	V	E	S	T	K	I	I
A/sw/GT/CIP049-C1913/2016	35	A	N	L	I	I	.	.	V	Q	P	.	.	F	S	.	T	S	.	.	.	.	.	.	.	A	A	K	.	.	.	.	.	V	E	S	T	K	I	I
A/sw/GT/CIP049-C2107/2016	1	A	N	L	I	I	.	.	V	Q	P	.	.	F	S	.	T	S	.	.	.	.	.	S	.	A	A	K	.	.	.	.	.	V	E	S	T	K	I	I
A/sw/GT/CIP049-C2214/2016	16	A	N	L	I	I	.	.	V	Q	P	.	P	F	S	.	T	S	.	.	.	.	.	.	.	A	A	K	.	.	.	.	.	V	E	S	T	K	I	I
A/sw/GT/CIP049-C2356/2016	1	A	N	L	I	I	.	.	V	Q	P	V	P	F	S	.	T	S	.	.	.	E	.	.	.	A	A	K	.	.	.	.	.	V	E	S	T	K	I	I
A/sw/GT/CIP049-C2541/2016	3	A	N	L	I	I	.	.	V	Q	P	V	P	F	S	.	T	S	.	.	.	.	N	.	.	A	A	K	.	.	.	.	.	V	E	S	T	K	I	I
A/sw/GT/CIP049-C2547/2016	35	A	N	L	I	I	.	.	V	Q	P	V	P	F	S	.	T	S	.	.	.	.	.	.	.	A	A	K	.	.	.	.	.	V	E	S	T	K	I	I
A/sw/GT/CIP049-C2809/2016	1	A	N	L	I	I	.	.	M	Q	P	V	P	F	S	.	T	S	.	.	.	.	.	.	.	A	A	K	.	.	.	.	.	V	E	S	T	K	I	I
A/sw/GT/CIP049-C2832/2017	2	A	N	L	I	I	.	.	V	Q	P	V	P	F	S	.	T	S	.	.	I	.	.	.	.	A	A	K	.	.	.	.	.	V	E	S	T	K	I	I
A/sw/GT/CIP049-C2870/2017	1	A	N	L	I	I	.	.	V	Q	P	V	P	F	S	.	T	S	.	.	.	.	.	.	.	A	A	K	.	.	I	.	.	V	E	S	T	K	I	I
A/sw/GT/CIP049-C2871/2017	4	A	N	L	I	I	.	.	V	Q	P	V	P	F	S	.	T	S	Y	.	I	.	.	.	.	A	A	K	.	.	.	.	.	V	E	S	T	K	I	I
A/sw/GT/CIP049-C3618/2018	33	.	.	.	.	.	.	.	V	K	.	.	.	.	S	N	.	.	.	T	.	.	.	T	.	A	.	K	R	.	.	S	I	V	E	.	.	.	.	.
A/sw/GT/CIP049-C3646/2018	1	.	.	.	.	.	.	.	V	K	.	.	.	.	S	N	.	.	.	T	.	.	N	T	.	A	.	K	R	.	.	S	T	V	E	.	.	.	.	.
A/sw/GT/CIP049-C3649/2018	1	.	.	.	.	.	.	.	V	K	.	.	.	.	S	N	.	.	.	T	.	.	.	T	.	A	.	K	R	G	.	S	I	V	E	.	.	.	.	.
A/sw/GT/CIP049-C3681/2018	1	.	.	.	.	.	D	.	V	K	.	.	.	.	S	N	.	.	.	T	.	.	.	T	.	A	.	K	R	.	.	S	I	V	E	.	.	.	.	.
A/sw/GT/CIP049-C3841/2018	1	.	.	.	.	.	.	.	I	K	.	.	.	.	S	N	.	.	.	T	.	.	.	T	.	A	.	K	R	.	.	S	I	V	E	.	.	.	.	.
A/sw/GT/CIP049-C3853/2018	1	.	.	.	.	.	.	I	V	K	.	.	.	.	S	N	.	.	.	T	.	.	.	T	.	A	.	K	R	.	.	S	I	V	E	.	.	.	.	V

NA

A/California/04/2009	number of identical sequences	P	T	M	T	I	V	S	S	K	A	Y	I	P	N	I	A	V	I	M	S	T	Q	P	K	N	S	S	N	I	T	G	N	N	I	V	V
		3	9	15	16	30	62	79	82	84	86	100	117	126	189	211	232	234	263	269	285	289	313	328	331	355	364	366	369	374	381	382	385	386	389	394	424
A/Guatomala/1039/2009		.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
A/sw/GT/CIP049-C1842/2016	5	.	.	.	.	.	.	P	.	S	H	M	.	.	.	.	.	.	.	.	.	.	R	.	R	.	.	.	.	.	.	.	.	.	M	I	.
A/sw/GT/CIP049-C1861/2016	54	.	.	.	.	.	.	.	.	.	H	.	.	.	.	.	.	.	.	.	.	.	.	S	R	.	.	.	.	.	.	.	.	.	M	L	.
A/sw/GT/CIP049-C1972/2016	4	.	.	.	.	.	.	.	.	.	H	.	.	.	.	.	.	.	.	.	.	.	.	S	R	.	.	N	.	.	.	.	.	.	M	L	.
A/sw/GT/CIP049-C1981/2016	1	.	.	.	.	.	.	.	.	.	H	.	.	.	S	.	.	.	.	.	.	.	.	S	R	.	.	.	.	.	.	.	.	.	M	L	.
A/sw/GT/CIP049-C2117/2016	1	.	.	.	.	.	.	.	.	.	H	.	.	.	.	.	.	.	.	.	.	.	.	S	R	.	.	.	.	T	.	.	.	.	M	L	.
A/sw/GT/CIP049-C2299/2016	1	.	.	.	A	.	.	.	.	.	H	.	.	.	.	.	.	.	.	.	.	.	.	S	R	.	.	.	.	.	.	.	.	.	M	L	I
A/sw/GT/CIP049-C2342/2016	1	.	.	.	.	.	L	.	.	.	H	.	.	.	.	.	.	.	.	.	.	.	.	S	R	.	.	.	.	.	.	.	.	.	M	L	.
A/sw/GT/CIP049-C2367/2016	1	.	.	.	.	.	.	.	.	.	H	.	.	.	.	.	.	.	.	.	.	.	.	S	R	.	.	.	.	.	.	.	.	.	.	L	.
A/sw/GT/CIP049-C2380/2016	17	.	.	.	.	.	.	P	.	.	H	.	.	.	.	.	.	.	.	.	.	.	.	S	R	.	.	.	.	.	.	.	.	.	M	L	.
A/sw/GT/CIP049-C2584/2016	15	.	.	.	.	.	.	P	.	.	H	.	.	.	.	.	.	.	V	.	.	.	.	S	R	.	.	.	.	.	.	.	.	.	M	L	.
A/sw/GT/CIP049-C2587/2016	1	.	.	.	.	.	.	P	.	.	H	.	.	H	.	.	.	.	.	.	.	.	.	S	R	.	.	.	.	.	.	.	.	.	M	L	.
A/sw/GT/CIP049-C3615/2018	24	S	A	I	.	T	.	.	.	.	H	.	.	.	.	I	A	I	.	I	.	I	R	.	N	.	N	N	K	.	I	E	K	.	.	.	.
A/sw/GT/CIP049-C3628/2018	5	S	A	I	.	T	.	.	.	.	H	.	.	.	.	I	A	I	.	I	.	I	R	.	N	K	N	N	K	.	I	E	K	.	.	.	.
A/sw/GT/CIP049-C3648/2018	1	S	A	I	.	T	.	.	.	.	H	.	.	.	.	I	A	I	.	I	F	I	R	.	N	.	N	N	K	.	I	E	K	.	.	.	.
A/sw/GT/CIP049-C3805/2018	1	S	A	I	.	T	.	.	.	.	H	.	.	.	.	I	A	I	.	I	.	I	R	.	N	.	N	N	K	.	I	E	K	S	.	.	.
A/sw/GT/CIP049-C3837/2018	1	S	A	I	.	T	.	.	.	.	H	.	.	.	.	I	A	I	.	I	.	I	R	.	N	.	N	N	K	V	I	E	K	.	.	.	.
A/sw/GT/CIP049-C3851/2018	4	S	A	I	.	T	.	.	.	R	H	.	.	.	.	I	A	I	.	I	.	I	R	.	N	.	N	N	K	.	I	E	K	.	.	.	.

Figure 6.9 PB2, PB1, PA and NP unique amino acids of H1N1 pdm09 swine viruses from Guatemala aligned to Ca04 (GenBank accession no. MN371615.1, MN371613.1, MN371611.1 and MN371617.1). Dots represent amino acids identical to Ca04. Viruses collected during May 2016 to February 2017 are shown in orange and those collected in 2018 in grey

PB2

A/California/04/2009	number of identical sequences	21 T	64 M	76 T	157 K	165 E	178 T	184 A	194 Q	195 D	227 V	300 Q	355 R	451 I	453 S	478 I	511 V	524 T	553 I	560 V	584 V	611 D	676 T	683 T	711 N	739 R
A/sw/GT/CIP049-C1842/2016	5	.	I	.	.	.	.	T	.	N	.	K	.	.	.	.	.	.	V	.	.	.	.	A	S	.
A/sw/GT/CIP049-C1861/2016	68	.	I	.	.	.	.	T	.	N	.	.	.	.	.	.	.	.	.	.	.	.	.	A	S	.
A/sw/GT/CIP049-C1972/2016	1	.	I	.	.	.	.	T	.	N	.	.	.	.	.	.	.	.	.	.	.	.	.	.	S	.
A/sw/GT/CIP049-C1973/2016	1	.	I	.	.	.	.	T	.	N	.	.	G	.	.	.	.	.	.	.	.	.	.	A	S	.
A/sw/GT/CIP049-C2107/2016	1	.	I	I	.	.	.	T	.	N	.	.	.	.	.	.	.	.	.	I	.	.	.	A	S	.
A/sw/GT/CIP049-C2356/2016	1	.	I	.	.	.	.	T	.	N	.	.	.	.	.	.	.	.	.	.	I	.	.	A	S	.
A/sw/GT/CIP049-C2370/2016	1	.	I	.	.	.	.	T	.	N	.	.	.	.	.	I	.	.	.	.	.	.	.	A	S	.
A/sw/GT/CIP049-C2408/2016	1	.	I	.	.	.	.	T	.	N	.	.	.	.	P	.	.	.	.	.	.	.	.	A	S	.
A/sw/GT/CIP049-C2585/2016	2	.	I	.	.	.	.	T	.	N	.	.	.	T	.	.	.	.	.	.	.	.	.	A	S	.
A/sw/GT/CIP049-C2623/2017	1	.	I	.	.	K	.	T	.	N	.	.	.	.	.	.	.	.	.	.	.	.	.	A	S	.
A/sw/GT/CIP049-C2650/2017	1	.	I	.	.	.	.	T	.	N	.	.	.	.	.	.	.	.	.	.	.	E	.	A	S	.
A/sw/GT/CIP049-C2651/2017	1	N	I	.	.	.	.	T	.	N	.	.	.	.	.	.	.	.	.	.	.	E	.	A	S	.
A/sw/GT/CIP049-C2655/2017	1	N	I	.	.	.	.	T	.	N	.	.	.	.	.	.	.	.	.	.	.	.	.	A	S	.
A/sw/GT/CIP049-C2671/2017	3	.	I	.	.	.	.	T	.	N	I	.	.	.	.	.	.	.	.	.	.	.	.	A	S	.
A/sw/GT/CIP049-C3617/2018	12	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	I	S	.	.	.	.	I	.	.	.
A/sw/GT/CIP049-C3804/2018	6	.	.	.	N	.	A	.	.	.	.	.	.	.	.	.	I	S	.	.	.	.	I	.	.	.
A/sw/GT/CIP049-C3806/2018	1	.	.	.	N	.	A	.	K	.	.	.	.	.	.	.	I	S	.	.	.	.	I	.	.	.
A/sw/GT/CIP049-C3878/2018	1	.	.	.	N	.	A	.	.	.	.	.	.	.	.	.	I	S	.	.	.	.	I	.	.	Q

PB1

A/California/04/2009	number of identical sequences	8 L	40 M	48 Q	108 L	114 V	115 Q	134 N	152 S	179 I	216 G	317 M	322 I	342 I	353 K	397 I	430 K	455 N	465 R	533 N	652 A	694 N	740 F	743 I
A/sw/GT/CIP049-C1842/2016	60	.	I	K	.	.	.	.	.	M	S	V	V	V	R	.	.	D	.	.	.	.	.	.
A/sw/GT/CIP049-C1973/2016	1	.	I	K	.	.	K	.	.	M	S	V	A	V	R	.	.	D	.	.	.	.	.	.
A/sw/GT/CIP049-C1982/2016	1	.	I	K	I	.	.	.	.	M	S	V	V	V	R	.	.	D	.	.	.	.	.	.
A/sw/GT/CIP049-C2117/2016	1	.	I	K	.	.	.	.	.	M	S	V	V	V	R	.	.	D	.	.	.	.	.	T
A/sw/GT/CIP049-C2173/2016	2	.	I	K	.	.	.	.	.	M	S	V	V	V	R	.	.	D	.	.	S	.	.	.
A/sw/GT/CIP049-C2367/2016	1	.	I	K	.	.	.	.	.	M	S	V	V	V	R	.	.	D	.	S	.	.	.	.
A/sw/GT/CIP049-C2500/2016	1	P	I	K	.	.	.	.	.	M	S	V	V	V	R	.	.	D	.	.	.	.	.	.
A/sw/GT/CIP049-C2651/2017	1	.	I	K	.	.	.	.	.	M	S	V	V	V	R	.	.	D	.	.	.	.	S	.
A/sw/GT/CIP049-C3627/2018	3	.	.	.	.	I	.	N	T	M	.	.	.	.	R	.	R	D	K	.	.	H	.	.
A/sw/GT/CIP049-C3804/2018	1	.	.	.	.	I	.	N	T	M	.	.	.	.	R	M	R	D	K	.	.	H	.	.

PA

A/California/04/2009	number of identical sequences	20 A	32 T	44 V	55 D	105 F	186 S	224 P	227 E	256 K	261 L	304 L	308 I	327 E	332 P	346 Q	364 S	379 V	385 K	387 V	395 S	400 P	405 S	419 D	460 M	470 L	532 L	538 E	556 Q	560 P	589 L	650 Y	660 A	664 K	688 E
A/sw/GT/CIP049-C1842/2016	4	T	I	I	N	H	N	S	D	.	.	.	.	A	.	H	.	.	.	.	.	.	.	.	.	.	M	.	.	.	.	F	.	.	G
A/sw/GT/CIP049-C1861/2016	28	T	I	I	N	H	N	S	D	.	.	I	.	A	.	H	.	I	.	.	.	.	.	.	.	.	M	.	.	.	.	F	.	.	.
A/sw/GT/CIP049-C1968/2016	8	T	I	I	N	H	N	S	D	.	.	I	.	A	.	H	.	I	.	.	.	.	.	.	.	.	M	.	.	S	.	F	.	.	.
A/sw/GT/CIP049-C2004/2016	1	T	I	I	N	H	N	S	D	.	.	I	.	A	.	H	.	I	.	.	.	.	.	.	.	.	M	.	.	.	I	F	.	.	.
A/sw/GT/CIP049-C2083/2016	1	T	I	I	N	H	N	S	D	Q	.	I	.	A	.	H	.	I	.	.	.	.	.	.	.	.	M	.	.	.	.	F	.	.	.
A/sw/GT/CIP049-C2173/2016	2	T	I	I	N	H	N	S	D	.	.	I	.	A	.	H	.	I	.	.	.	.	.	.	.	.	M	.	K	.	.	F	.	.	.
A/sw/GT/CIP049-C2213/2016	5	T	I	I	N	H	N	S	D	.	.	I	.	A	.	H	.	I	.	.	.	.	G	.	.	.	M	.	.	.	.	F	.	.	.
A/sw/GT/CIP049-C2224/2016	27	T	I	I	N	H	N	S	D	.	.	I	.	A	.	H	.	I	.	.	.	.	.	.	I	.	M	.	.	.	.	F	.	.	.
A/sw/GT/CIP049-C2256/2016	1	T	I	I	N	H	N	S	D	.	.	I	.	A	.	H	.	I	.	.	.	.	.	.	.	P	M	.	.	.	.	F	.	.	.
A/sw/GT/CIP049-C2258/2016	1	T	I	I	N	H	N	S	D	.	.	I	.	A	.	H	.	I	.	.	.	.	.	.	.	.	M	.	.	.	.	F	G	.	.
A/sw/GT/CIP049-C2408/2016	1	T	I	I	N	H	N	S	D	.	.	I	.	A	.	H	.	I	.	.	.	.	.	.	I	.	M	G	.	.	.	F	.	.	.
A/sw/GT/CIP049-C2547/2016	1	T	I	I	N	H	N	S	D	.	.	I	V	A	.	H	.	I	.	.	.	.	.	.	I	.	M	.	.	.	.	F	.	.	.
A/sw/GT/CIP049-C2622/2017	6	T	L	I	N	H	N	S	D	.	.	I	.	A	.	H	.	I	.	.	.	.	.	.	I	.	M	.	.	.	.	F	.	.	.
A/sw/GT/CIP049-C2636/2017	4	T	I	I	N	H	N	S	D	.	.	I	.	A	.	H	N	I	R	.	.	.	.	.	I	.	M	.	.	.	.	F	.	.	.
A/sw/GT/CIP049-C3615/2018	31	.	.	.	.	L	.	S	.	.	S	.	.	.	S	.	G	.	.	I	G	S	.	N	.	.	.	.	.	.	.	F	.	R	.

NP

A/California/04/2009	number of identical sequences	31 R	34 G	53 D	186 V	190 A	217 V	242 V	283 L	294 E	353 I	373 T	400 K	401 A	430 S	444 V	455 D
A/sw/GT/CIP049-C1842/2016	11	.	.	E	I	.	.	.	.	.	.	.	R	.	.	.	.
A/sw/GT/CIP049-C1913/2016	55	.	.	E	I	.	.	.	.	.	.	.	R	T	.	.	.
A/sw/GT/CIP049-C2342/2016	1	.	.	E	I	.	I	I	.	.	.	.	R	T	.	.	.
A/sw/GT/CIP049-C2370/2016	1	.	.	E	.	V	.	.	.	.	.	.	R	T	.	.	.
A/sw/GT/CIP049-C2400/2016	1	.	S	E	I	.	.	.	.	.	V	.	R	T	.	.	.
A/sw/GT/CIP049-C2408/2016	31	.	S	E	I	.	.	.	.	.	.	.	R	T	.	.	.
A/sw/GT/CIP049-C2671/2017	1	.	S	E	I	.	.	.	.	X	.	.	R	T	.	.	E
A/sw/GT/CIP049-C2672/2017	3	.	S	E	I	.	.	.	.	.	.	.	R	T	.	.	E
A/sw/GT/CIP049-C3618/2018	35	K	S	E	I	.	.	I	.	.	.	A	.	.	.	I	.
A/sw/GT/CIP049-C3631/2018	1	K	S	E	I	.	.	I	.	.	.	A	.	.	N	I	.
A/sw/GT/CIP049-C3646/2018	1	K	S	E	I	.	.	I	I	.	.	A	.	.	.	I	.

Figure 6.10. M1, M2, NS1 and NEP unique amino acids of H1N1 pdm09 swine viruses from Guatemala aligned to Ca04 (GenBank accession no. FJ969513.1, FJ969514.1). Dots represent amino acids identical to Ca04. Viruses collected during May 2016 to February 2017 are shown in orange and those collected in 2018 in grey.

M1

A/California/04/2009	number of identical sequences	13	30	107	149	169	192	201	208	216	225	227	235
		S	S	I	A	T	M	E	Q	M	S	A	E
A/sw /GT/ClP049-C1842/2016	12	T	.	.	.	.	.	.	.	.	.	.	.
A/sw /GT/ClP049-C1929/2016	3	T	.	.	S	.	.	D	.	.	.	.	.
A/sw /GT/ClP049-C2178/2016	1	T	.	T	.	.	.	D	.	T	.	.	.
A/sw /GT/ClP049-C2256/2016	1	T	N	.	.	.	.	D	.	.	.	.	.
A/sw /GT/ClP049-C2547/2016	98	T	.	.	.	.	.	D	.	.	.	.	.
A/sw /GT/ClP049-C2623/2017	1	T	.	.	.	.	.	D	.	.	.	.	G
A/sw /GT/ClP049-C3618/2018	41	.	.	.	.	I	I	.	R	.	S	A	.

M2

A/California/04/2009	number of identical sequences	18	25	28	33	38	43	48	52	55	58	61	82	95
		R	P	I	I	L	T	F	Y	F	G	R	S	E
A/sw /GT/ClP049-C1842/2016	39	.	.	.	.	.	I	L	.	I	.	K	.	.
A/sw /GT/ClP049-C1861/2016	5	.	.	.	V	.	I	L	.	I	.	K	.	.
A/sw /GT/ClP049-C1975/2016	1	K	.	.	V	.	I	L	.	I	.	K	.	.
A/sw /GT/ClP049-C2224/2016	1	.	.	F	.	.	I	L	.	I	.	K	.	.
A/sw /GT/ClP049-C2299/2016	13	.	L	F	.	.	I	L	.	I	.	K	.	.
A/sw /GT/ClP049-C2479/2016	1	.	.	F	.	.	I	L	.	I	.	K	.	G
A/sw /GT/ClP049-C2534/2016	1	.	.	F	.	M	I	L	.	I	.	K	.	.
A/sw /GT/ClP049-C2547/2016	54	.	.	F	.	.	I	L	.	I	.	K	.	.
A/sw /GT/ClP049-C2673/2017	1	.	.	F	.	.	I	L	.	I	E	K	.	.
A/sw /GT/ClP049-C3618/2018	41	.	.	.	.	.	.	.	H	I	.	K	N	.

NS1

A/California/04/2009	number of identical sequences	18	52	60	66	72	81	90	91	128	129	155	170	171	181	204	206	213	214	215
		I	L	V	E	E	I	L	S	I	V	A	T	Y	L	R	C	S	L	P
A/sw /GT/ClP049-C1842/2016	5	.	.	.	D	.	.	.	F	I	.	.	.	.	I	.	.	.	.	.
A/sw /GT/ClP049-C1974/2016	1	.	.	A	.	.	.	.	F	I	.	.	.	.	I	.	.	.	.	.
A/sw /GT/ClP049-C2214/2016	7	.	.	.	.	.	V	.	F	I	.	.	.	.	I	.	.	.	.	.
A/sw /GT/ClP049-C2227/2016	2	V	.	.	.	.	.	.	F	I	.	.	.	.	I	.	.	.	.	.
A/sw /GT/ClP049-C2256/2016	1	.	.	.	.	.	.	.	F	I	.	.	.	.	I	.	.	.	P	.
A/sw /GT/ClP049-C2287/2016	1	.	.	.	.	.	.	F	F	I	.	.	.	.	I	.	.	.	.	.
A/sw /GT/ClP049-C2333/2016	1	.	.	.	.	D	.	.	F	I	.	.	.	.	I	.	.	.	.	.
A/sw /GT/ClP049-C2526/2016	1	.	.	.	.	.	.	.	F	I	.	.	A	.	I	.	.	.	.	.
A/sw /GT/ClP049-C2547/2016	84	.	.	.	.	.	.	.	F	I	.	.	.	.	I	.	.	.	.	.
A/sw /GT/ClP049-C2655/2017	1	.	.	.	.	.	.	.	F	I	I	.	.	.	I	.	.	.	.	.
A/sw /GT/ClP049-C3618/2018	38	.	I	.	.	.	.	.	.	.	.	A	.	H	.	.	S	P	.	S
A/sw /GT/ClP049-C3649/2018	1	.	I	.	.	.	.	.	.	.	.	A	.	H	.	K	S	P	.	S

NEP

A/California/04/2009	number of identical sequences	12	22	27	34	47	57	74	83	85	86	89	98	99	113
		I	G	D	R	E	Y	E	M	H	R	A	T	F	I
A/sw /GT/ClP049-C1842/2016	5	.	E	.	.	.	.	.	I	.	.	.	.	.	.
A/sw /GT/ClP049-C1861/2016	2	.	.	G	.	.	.	.	I	.	.	.	.	.	.
A/sw /GT/ClP049-C1913/2016	14	.	.	.	.	.	.	.	I	.	.	.	.	.	.
A/sw /GT/ClP049-C1972/2016	4	.	.	G	.	.	.	V	I	.	.	.	.	.	.
A/sw /GT/ClP049-C2026/2016	1	.	.	.	.	.	.	.	I	.	K	.	A	.	.
A/sw /GT/ClP049-C2256/2016	1	.	.	.	.	.	H	.	I	.	K	.	.	.	.
A/sw /GT/ClP049-C2526/2016	1	M	.	.	.	.	.	.	I	.	K	.	.	.	.
A/sw /GT/ClP049-C2547/2016	75	.	.	.	.	.	.	.	I	.	K	.	.	.	.
A/sw /GT/ClP049-C2650/2017	1	.	.	.	.	.	.	.	I	.	K	.	.	L	.
A/sw /GT/ClP049-C3618/2018	38	.	.	.	L	.	.	.	.	H	.	T	.	.	L
A/sw /GT/ClP049-C3649/2018	1	.	.	.	L	K	.	.	.	H	.	T	.	.	L

Figure 6.11. PA-X, PA-N155 and PA-N182 unique amino acids of H1N1 pdm09 swine viruses from Guatemala aligned to Ca04 (GenBank accession no. MN371611.1). Dots represent amino acids identical to Ca04. Viruses collected during May 2016 to February 2017 are shown in orange and those collected in 2018 in grey.

PA-X

A/California/04/2009	number of identical sequences	20	32	44	55	105	186	204	205	207	210	212	219	227	229
		A	T	V	D	F	S	N	L	L	L	A	S	K	L
A/sw/GT/CIP049-C1842/2016	83	T	I	I	N	H	N	S	.	S	.	.	F	I	S
A/sw/GT/CIP049-C2541/2016	2	T	I	I	N	H	N	S	.	S	.	.	.	I	S
A/sw/GT/CIP049-C2622/2017	6	T	L	I	N	H	N	S	.	S	.	.	F	I	S
A/sw/GT/CIP049-C3615/2018	31	.	.	.	.	L	.	.	S	.	P	E	.	.	.

PA-N155

A/California/04/2009	number of identical sequences	32	70	73	102	107	150	154	173	178	192	210	225	231	233	241	246	251	265	306	316	378	384	402	406	435	496	506	510	534
		S	P	E	K	L	L	I	E	P	Q	S	V	K	V	S	P	S	D	M	L	L	E	Q	P	L	Y	A	K	E
A/sw/GT/CIP049-C1842/2016	4	N	S	D	.	.	.	.	A	.	H	.	.	.	.	.	.	.	.	.	.	M	.	.	.	.	F	.	.	G
A/sw/GT/CIP049-C1861/2016	28	N	S	D	.	.	I	.	A	.	H	.	I	.	.	.	.	.	.	.	.	M	.	.	.	.	F	.	.	.
A/sw/GT/CIP049-C1968/2016	8	N	S	D	.	.	I	.	A	.	H	.	I	.	.	.	.	.	.	.	.	M	.	.	S	.	F	.	.	.
A/sw/GT/CIP049-C2004/2016	1	N	S	D	.	.	I	.	A	.	H	.	I	.	.	.	.	.	.	.	.	M	.	.	.	I	F	.	.	.
A/sw/GT/CIP049-C2083/2016	1	N	S	D	Q	.	I	.	A	.	H	.	I	.	.	.	.	.	.	.	.	M	.	.	.	.	F	.	.	.
A/sw/GT/CIP049-C2173/2016	2	N	S	D	.	.	I	.	A	.	H	.	I	.	.	.	.	.	.	.	.	M	.	K	.	.	F	.	.	.
A/sw/GT/CIP049-C2213/2016	5	N	S	D	.	.	I	.	A	.	H	.	I	.	.	.	.	G	.	.	.	M	.	.	.	.	F	.	.	.
A/sw/GT/CIP049-C2224/2016	34	N	S	D	.	.	I	.	A	.	H	.	I	.	.	.	.	.	.	I	.	M	.	.	.	.	F	.	.	.
A/sw/GT/CIP049-C2256/2016	1	N	S	D	.	.	I	.	A	.	H	.	I	.	.	.	.	.	.	.	P	M	.	.	.	.	F	.	.	.
A/sw/GT/CIP049-C2258/2016	1	N	S	D	.	.	I	.	A	.	H	.	I	.	.	.	.	.	.	.	.	M	.	.	.	.	F	G	.	.
A/sw/GT/CIP049-C2408/2016	1	N	S	D	.	.	I	.	A	.	H	.	I	.	.	.	.	.	.	I	.	M	G	.	.	.	F	.	.	.
A/sw/GT/CIP049-C2547/2016	1	N	S	D	.	.	I	V	A	.	H	.	I	.	.	.	.	.	.	I	.	M	.	.	.	.	F	.	.	.
A/sw/GT/CIP049-C2636/2017	4	N	S	D	.	.	I	.	A	.	H	N	I	R	.	.	.	.	.	I	.	M	.	.	.	.	F	.	.	.
A/sw/GT/CIP049-C3615/2018	31	.	S	.	.	S	.	.	.	S	.	G	.	.	I	G	S	.	N	.	.	.	.	.	.	.	F	.	R	.

PA-N182

A/California/04/2009	number of identical sequences	5	43	46	75	80	123	127	146	151	165	183	198	204	206	214	219	224	238	279	289	351	357	375	379	408	469	479	483	507
		S	P	E	K	L	L	I	E	P	Q	S	V	K	V	S	P	S	D	M	L	L	E	Q	P	L	Y	A	K	E
A/sw/GT/CIP049-C1842/2016	4	N	S	D	.	.	.	.	A	P	H	.	.	.	.	.	.	.	.	.	.	M	.	.	.	.	F	.	.	G
A/sw/GT/CIP049-C1861/2016	28	N	S	D	.	.	I	.	A	P	H	.	I	.	.	.	.	.	.	.	.	M	.	.	.	.	F	.	.	.
A/sw/GT/CIP049-C1968/2016	8	N	S	D	.	.	I	.	A	P	H	.	I	.	.	.	.	.	.	.	.	M	.	.	S	.	F	.	.	.
A/sw/GT/CIP049-C2004/2016	1	N	S	D	.	.	I	.	A	P	H	.	I	.	.	.	.	.	.	.	.	M	.	.	.	I	F	.	.	.
A/sw/GT/CIP049-C2083/2016	1	N	S	D	Q	.	I	.	A	P	H	.	I	.	.	.	.	.	.	.	.	M	.	.	.	.	F	.	.	.
A/sw/GT/CIP049-C2173/2016	2	N	S	D	.	.	I	.	A	P	H	.	I	.	.	.	.	.	.	.	.	M	.	K	.	.	F	.	.	.
A/sw/GT/CIP049-C2213/2016	5	N	S	D	.	.	I	.	A	P	H	.	I	.	.	.	.	G	.	.	.	M	.	.	.	.	F	.	.	.
A/sw/GT/CIP049-C2224/2016	34	N	S	D	.	.	I	.	A	P	H	.	I	.	.	.	.	.	.	I	.	M	.	.	.	.	F	.	.	.
A/sw/GT/CIP049-C2256/2016	1	N	S	D	.	.	I	.	A	P	H	.	I	.	.	.	.	.	.	.	P	M	.	.	.	.	F	.	.	.
A/sw/GT/CIP049-C2258/2016	1	N	S	D	.	.	I	.	A	P	H	.	I	.	.	.	.	.	.	.	.	M	.	.	.	.	F	G	.	.
A/sw/GT/CIP049-C2408/2016	1	N	S	D	.	.	I	.	A	P	H	.	I	.	.	.	.	.	.	I	.	M	G	.	.	.	F	.	.	.
A/sw/GT/CIP049-C2547/2016	1	N	S	D	.	.	I	V	A	P	H	.	I	.	.	.	.	.	.	I	.	M	.	.	.	.	F	.	.	.
A/sw/GT/CIP049-C2636/2017	4	N	S	D	.	.	I	.	A	P	H	N	I	R	.	.	.	.	.	I	.	M	.	.	.	.	F	.	.	.
A/sw/GT/CIP049-C3615/2018	31	.	S	.	.	S	.	.	.	.	.	G	.	.	I	G	S	.	N	.	.	.	.	.	.	.	F	.	R	.

Figure 6.12. Coverage plot of sequenced samples. Gray lines show the coverage distribution from each individual sample. The red line depicts the geometric mean.

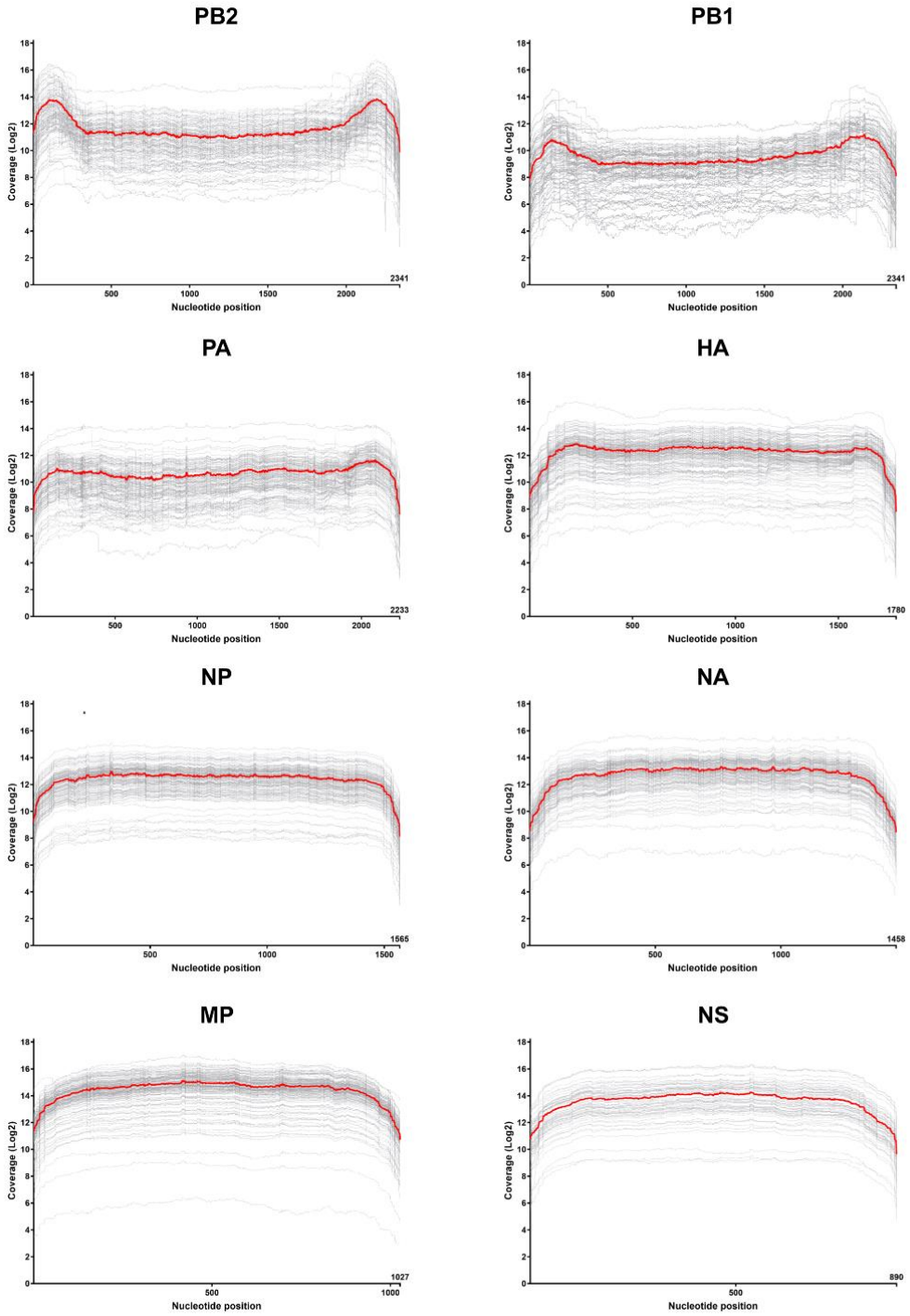

Table 6.1 Risk factors associated with FLUAV detection by RRT-PCR in sampled pigs in Guatemala during 2016-2018.

Characteristic		**FLUAV positive (%) (n=251)**		**FLUAV negative (%) n=(1,843)**		**OR (95% IC)**		**P-value**
Sex	Female	103	(41.0)	894	(48.5)		Referent	
	Male	148	(59.0)	949	(51.5)	**1.4**	**(1.0,1.8)**	**0.0262**
Age	Weaning (4-10wks)	232	(92.4)	1236	(67.1)	**4.1**	**(1.3,20.7)**	**0.0098**
	Juvenile (11-17wks)	11	(4.4)	296	(16.1)	0.8	(0.2,4.7)	0.7618
	Semi-adult (18-20wks)	5	(2.0)	245	(13.3)	0.4	(0.1,3.0)	0.2695
	Adult (>20 wks)	3	(1.2)	66	(3.5)		Referent	
Rectal temperature	<39.7b	162	(64.5)	1527	(82.9)		Referent	
	>39.8	89	(35.5)	316	(17.1)	**2.7**	**(2.0,3.6)**	**0.0000**
Density (animals/pen)	<100	8	(3.2)	100	(5.4)		Referent	
	101-200	24	(9.6)	315	(17.1)	1.0	(0.4,2.5)	0.9084
	201-300	110	(43.8)	672	(36.5)	2.0	(1.0,5.0)	0.0558
	301-400	107	(42.6)	703	(38.1)	1.9	(0.9,4.7)	0.0871
	>400	2	(0.8)	53	(2.9)	0.5	(0.0,2.5)	0.3428

[a:] Normal temperature ranges between 38.7–39.7. OR: odds ratio.

Table 6.1 Total number of sequences obtained from sampled pigs in Guatemala during 2016-2018.

Segment	Number of total full-length sequences	Number of ORF nucleotide sequences (a)		Number of unique amino acid sequences	
		2016-2017	2018	2016-2017	2018
PB2	106	86 (38)	20 (5)	14	4
PB1	71	68 (29)	4 (2)	8	2
PA	122	91(39)	31(5)	14 (PA)	1 (PA)
				3 (PA-X)	1 (PA-X)
				13 (PA-N155)	1 (PA-N155)
				13 (PA-N182)	1 (PA-N182)
HA	144	106 (35)	38 (10)	12	6
NP	140	103 (25)	37 (8)	8	3
NA	137	101 (27)	36 (11)	11	6
M	157	116 (15)	41 (4)	6 (M1)	1 (M1)
				9 (M2)	1 (M2)
NS	143	104 (17)	39 (3)	10 (NS1)	2 (NS1)
				9(NEP)	2 (NEP)
TOTAL	1,020	775 (225)	246 (48)		

[a] Used for phylogenetic analysis

Acknowledgements

This study was supported by a contract (HHSN272201400008C) from the National Institute of Allergy and Infectious Diseases (NIAID) Centers for Influenza Research and Surveillance (CEIRS).

CHAPTER 7

SUMMARY AND CONCLUSSIONS

Influenza A viruses (FLUAV) infect a wide range of birds and mammals, including humans. Aquatic wild birds, particularly members of the *Anseriformes* and *Charadriiformes* orders are the major reservoir of FLUAV. Members of these two families of aquatic birds play a key role in maintaining and increasing FLUAV diversity and spread. On the other hand, swine have a rather unique role in the emergence of FLUAVs of public health concern. The main goal of my dissertation was to better define the ecology and evolution of FLUAVs in two animal populations with major importance for pandemic preparedness, but from which limited information exists: wild birds and swine in Guatemala.

In Chapter 4, I analyzed the results of sustained avian FLUAV surveillance efforts performed in Guatemala during six migratory seasons (2013-2019). I aimed to investigate long-term patterns in FLUAV prevalence and subtype variation, as well as to define in more detail the genetic structure of FLUAV population from wild birds in Guatemala. My thesis work focused on a relatively small geographic location in the Southernmost part of the Pacific coast of Guatemala near the border with El Salvador. I relied on hunter-harvested aquatic bird samples, which were largely biased to samples from blue winged teals. Although, FLUAVs were isolated also from a handful of other avian species, my analyses of FLUAVs prevalence and seasonality contemplated only those derived from blue winged teal samples. Thus, it is possible that my analyses revealed

only a fraction of the true FLUAV diversity that exist in Guatemala, although it is consistent with other parts of the world. I was able to sequence the full genome of twice as many virus isolates compared to previous studies from Guatemala, using the workflow proposed by Ferreri et al 2019 (101) in which NGS was prioritized using swab samples directly. This approach allowed me to provide a comprehensive picture of FLUAVs circulating in Guatemala without the confounding effects of bottlenecks that result from growing viruses in artificial systems (e.g. tissue culture or eggs). I found a great diversity of FLUAVs, in terms of subtype combinations and high frequency of detection of viruses that are rarely found at the breeding grounds in North America. Co-circulation of several subtypes and a high number of mixed infections found during the same season in a relatively small area, almost all from blue-winged teals suggests that this avian species plays a key role in the natural evolution and diversification of FLUAVs in Guatemala. I found a significant number of LPAIV H5 and H7 subtype strains highly similar to those implicated in HPAI outbreaks in poultry in the region. Such finding highlights the importance of continuous monitoring of natural FLUAV avian species to detect strains more promptly with major animal and/or human health implications. I found frequent evidence of reassortment between North American and Eurasian lineage FLUAVs, indicating the potential interconnection between these two large FLUAV reservoirs in Guatemala. My data also identified FLUAVs with multiple gene constellations from different flyways, consistent with a bottleneck effect of landmass on migratory flyways.

One of the most important findings of my thesis work is the realization of the highest number of H14 subtype strains detected in a single region (Chapter 5). Through my work, Guatemala represents the largest collection of H14 viruses worldwide. Overall, the gene

segment constellation of the Guatemalan H14 viruses are like those found in North America, but there is evidence of repeated introductions from Eurasia, as denoted the H14 Guatemalan viruses with NP segments from that region. The phylogenetic analyses revealed repeated introductions of H14 isolates from North America. Despite high similarity in the HA segment among H14 viruses, they show many internal gene constellations adding to the notion of free segment exchanges among FLUAVs in Guatemala. The persistence of H14 FLUAVs in Guatemala is the result of complex ecological drivers that warrant further studies beyond the scope of my thesis work. The HA of the Guatemalan H14 strains showed three distinct cleavage site motifs, two containing Arginine instead of Lysine in the first and fourth positions, respectively. It is unclear whether these changes affect virus replication, transmission and/or pathogenesis in either natural or "accidental" (poultry, mammals) hosts. Future studies should contemplate addressing these differences in order to more fully characterize H14 subtype FLUAVs. Through my thesis work, I provide additional support to the notion that Guatemala plays an important role in the persistence and evolution of FLUAV in the Americas.

In Chapter 6, I described and analyze data resulting from two years of active FLUAV surveillance in a commercial swine farm without history of vaccination. Commercial swine farms provide unique systems for interspecies transmission of FLUAVs at the animal-human interface. Bidirectional transmission of FLUAVs between pigs and humans plays a significant role in the generation of novel viral strains that become established in the new host population. My results found a general FLUAV prevalence like reported in previous studies in Guatemala (140). Over the 2-year surveillance, two

high infection periods but no seasonality patterns were detected. I sequenced >75% of the FLUAV-positive samples and ~25% of complete genomes by NGS directly from the original swabs. This methodology allowed us to improve the number of characterized samples and reduce potential selection bias introduced by virus isolation prior to sequencing. All samples were identified as FLUAVs of the H1N1 subtype within the H1 swine clade 1A.3.3.2 and whose ancestor are the human origin 2009 H1N1 pmd09 FLUAV. The mutations found in relevant HA1 antigenic sites may lead to differences in antigenic relationships compared to similar strains of human- or swine-origin; however, the effect of these mutations remain unclear. My effort suggests that Guatemalan FLUAVs from swine represent a unique group of sequences, having independent evolution pathways from similar viruses circulating in Central America. These viruses may represent the establishment of a novel genetic lineage with the potential to reassort with other co-circulating viruses, and whose zoonotic/pandemic risk remains to be determined.

Considering the findings presented in this dissertation, future directions of the work should aim at:

1. Include other sampling methodologies to capture other species of birds and other FLUAV susceptible hosts. My thesis work was largely limited to hunter harvested duck samples and those were biased towards blue winged teal due to its relatively high abundance. Other sampling methodologies including live bird captures and environmental sampling to obtain samples from other species, including resident species would provide a further understanding of

the local dynamics of FLUAV in Guatemala. By incorporating other FLUAV susceptible hosts, it may provide answers to how virus diversity and especially uncommon subtypes (including H14) are maintained in the wintering grounds.

2. Expand the FLUAV surveillance to the Central American region to identify alternative inter-lineage reassortments (including those that might result from gene segment exchange with the South American lineage). This study provides data from a relatively small region and incorporating other regions from Central America will contribute to better characterizing the patterns of virus reassortment during one or more migration seasons.
3. Evaluate pathogenicity, viral replication, and transmissibility of H14 strains containing alternative HA cleavage motifs. Limited studies with H14 strains were performed in the past showing similarities in viral replication and shedding. However, it is not clear the impact of the HA cleavage motifs changes found in this study may have on the pathogenesis of the virus; therefore, more studies are needed to evaluate the effects of these mutations on these traits.
4. Evaluate effects of HA1 mutations on antigenic sites in Guatemalan swine viruses to determine the effect of these mutations and its zoonotic risk.
5. Expand FLUAV surveillance to the human-animal interface to characterize local virus populations and their antigenic properties to improve our understanding of the FLUAV ecology on settings with high interaction of FLUAV susceptible hosts. Long persistence of human-origin viruses in environments with high interaction, such as swine commercial farms might

lead to the generation of strains with pandemic concern; therefore, it is important to determine the risk these viruses represent to the public health.

REFERENCES

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

2. Gamblin SJ, Skehel JJ. 2010. Influenza hemagglutinin and neuraminidase membrane glycoproteins. J Biol Chem 285:28403-9.

3. Tong S, Li Y, Rivailler P, Conrardy C, Castillo DA, Chen LM, Recuenco S, Ellison JA, Davis CT, York IA, Turmelle AS, Moran D, Rogers S, Shi M, Tao Y, Weil MR, Tang K, Rowe LA, Sammons S, Xu X, Frace M, Lindblade KA, Cox NJ, Anderson LJ, Rupprecht CE, Donis RO. 2012. A distinct lineage of influenza A virus from bats. Proc Natl Acad Sci U S A 109:4269-74.

4. Tong S, Zhu X, Li Y, Shi M, Zhang J, Bourgeois M, Yang H, Chen X, Recuenco S, Gomez J, Chen LM, Johnson A, Tao Y, Dreyfus C, Yu W, McBride R, Carney PJ, Gilbert AT, Chang J, Guo Z, Davis CT, Paulson JC, Stevens J, Rupprecht CE, Holmes EC, Wilson IA, Donis RO. 2013. New world bats harbor diverse influenza A viruses. PLoS Pathog 9:e1003657.

5. Alexander DJ. 2000. A review of avian influenza in different bird species. Veterinary microbiology 74:3-13.

6. Kalthoff D, Globig A, Beer M. 2010. (Highly pathogenic) avian influenza as a zoonotic agent. Veterinary Microbiology 140:237-245.

7. WHO. 2020. Cumulative number of confirmed human cases for avian influenza A(H5N1) reported to WHO, 2003-2020, *on* World Health Organization. https://www.who.int/influenza/human_animal_interface/2020_01_20_tableH5N1.pdf?ua=1. Accessed June 13, 2021.

8. Taubenberger JK, Reid AH, Lourens RM, Wang R, Jin G, Fanning TG. 2005. Characterization of the 1918 influenza virus polymerase genes. Nature 437:889-93.

9. Kilbourne E. 2006. Influenza Pandemics of the 20th Century. Emerging Infectious Disease journal 12:9.

10. Neumann G, Noda T, Kawaoka Y. 2009. Emergence and pandemic potential of swine-origin H1N1 influenza virus. Nature 459:931-939.

11. Webby RJ, Swenson SL, Krauss SL, Gerrish PJ, Goyal SM, Webster RG. 2000. Evolution of swine H3N2 influenza viruses in the United States. J Virol 74:8243-51.

12. Karasin AI, Carman S, Olsen CW. 2006. Identification of human H1N2 and human-swine reassortant H1N2 and H1N1 influenza A viruses among pigs in Ontario, Canada (2003 to 2005). J Clin Microbiol 44:1123-6.

13. Lewis NS, Anderson TK, Kitikoon P, Skepner E, Burke DF, Vincent AL. 2014. Substitutions near the hemagglutinin receptor-binding site determine the antigenic evolution of influenza A H3N2 viruses in U.S. swine. J Virol 88:4752-63.

14. Vincent AL, Ma W, Lager KM, Gramer MR, Richt JA, Janke BH. 2009. Characterization of a newly emerged genetic cluster of H1N1 and H1N2 swine influenza virus in the United States. Virus Genes 39:176-85.

15. Mena I, Nelson MI, Quezada-Monroy F, Dutta J, Cortes-Fernández R, Lara-Puente JH, Castro-Peralta F, Cunha LF, Trovão NS, Lozano-Dubernard B, Rambaut A, van Bakel H, García-Sastre A. 2016. Origins of the 2009 H1N1 influenza pandemic in swine in Mexico. eLife 5:e16777.

16. Fouchier RA, Osterhaus AD, Brown IH. 2003. Animal influenza virus surveillance. Vaccine 21:1754-7.

17. McCauley J. 2014. International Committee on Taxonomy of Viruses. http://www.ictvonline.org. Accessed

18. Webster RG, Bean WJ, Gorman OT, Chambers TM, Kawaoka Y. 1992. Evolution and ecology of influenza A viruses. Microbiological Reviews 56:152 - 179.

19. Palese P, Shaw M. 2007. Fields virology. Orthomyxoviridae: The Viruses and Their Replication, 5th ed; Lippincott Williams & Wilkins, Wolters Kluwer Business: Philadelphia, PA, USA:1647-1689.

20. Zebedee SL, Lamb RA. 1988. Influenza A virus M2 protein: monoclonal antibody restriction of virus growth and detection of M2 in virions. Journal of virology 62:2762-2772.

21. Noda T, Kawaoka Y. 2010. Structure of influenza virus ribonucleoprotein complexes and their packaging into virions. Rev Med Virol 20:380-91.

22. Flick R, Hobom G. 1999. Interaction of influenza virus polymerase with viral RNA in the 'corkscrew' conformation. J Gen Virol 80 (Pt 10):2565-2572.

23. Fields BN. 2007. Fields' virology, vol 1. Lippincott Williams & Wilkins.

24. Zhou B, Donnelly ME, Scholes DT, St George K, Hatta M, Kawaoka Y, Wentworth DE. 2009. Single-reaction genomic amplification accelerates sequencing and vaccine production for classical and Swine origin human influenza a viruses. J Virol 83:10309-13.

25. Shaw ML, Palese P. 2013. Orthomyxoviridae, p 1151-1185. *In* Knipe DM, Howley PM (ed), Fields Virology, vol II. Lippincott Williams & Wilkins, Philadelphia, PA.

26. Yamayoshi S, Watanabe M, Goto H, Kawaoka Y. 2015. Identification of a Novel Viral Protein Expressed from the PB2 Segment of Influenza A Virus. Journal of virology 90:444-456.

27. Long JC, Fodor E. 2016. The PB2 Subunit of the Influenza A Virus RNA Polymerase Is Imported into the Mitochondrial Matrix. J Virol 90:8729-38.

28. Chen W, Calvo PA, Malide D, Gibbs J, Schubert U, Bacik I, Basta S, O'Neill R, Schickli J, Palese P, Henklein P, Bennink JR, Yewdell JW. 2001. A novel influenza A virus mitochondrial protein that induces cell death. Nature Medicine 7:1306-1312.

29. Briedis DJ, Lamb RA. 1982. Influenza B virus genome: sequences and structural organization of RNA segment 8 and the mRNAs coding for the NS1 and NS2 proteins. J Virol 42:186-93.

30. Portela A, Digard P. 2002. The influenza virus nucleoprotein: a multifunctional RNA-binding protein pivotal to virus replication. J Gen Virol 83:723-734.

31. Lamb RA, Lai CJ, Choppin PW. 1981. Sequences of mRNAs derived from genome RNA segment 7 of influenza virus: colinear and interrupted mRNAs code for overlapping proteins. Proceedings of the National Academy of Sciences of the United States of America 78:4170-4174.

32. Furuse Y, Oshitani H. 2011. Evolution of the influenza A virus untranslated regions. Infect Genet Evol 11:1150-4.

33. Bouvier NM, Palese P. 2008. The biology of influenza viruses. Vaccine 26 Suppl 4:D49-D53.

34. Varki A, Schauer R. 2009. Sialic Acids. *In* Varki A, Cummings RD, Esko JD, Freeze HH, Stanley P, Bertozzi CR, Hart GW, Etzler ME (ed), Essentials of Glycobiology, 2nd ed, Cold Spring Harbor (NY).

35. Couceiro JN, Paulson JC, Baum LG. 1993. Influenza virus strains selectively recognize sialyloligosaccharides on human respiratory epithelium; the role of the host cell in selection of hemagglutinin receptor specificity. Virus Res 29:155-65.

36. Martin K, Helenius A. 1991. Transport of incoming influenza virus nucleocapsids into the nucleus. Journal of virology 65:232-244.

37. Sieczkarski SB, Whittaker GR. 2005. Viral entry. Curr Top Microbiol Immunol 285:1-23.

38. Poon LL, Pritlove DC, Fodor E, Brownlee GG. 1999. Direct evidence that the poly(A) tail of influenza A virus mRNA is synthesized by reiterative copying of a U track in the virion RNA template. Journal of virology 73:3473-3476.

39. Fujii Y, Goto H, Watanabe T, Yoshida T, Kawaoka Y. 2003. Selective incorporation of influenza virus RNA segments into virions. Proceedings of the National Academy of Sciences of the United States of America 100:2002-2007.

40. Palese P, Tobita K, Ueda M, Compans RW. 1974. Characterization of temperature sensitive influenza virus mutants defective in neuraminidase. Virology 61:397-410.

41. Garten W, Braden C, Arendt A, Peitsch C, Baron J, Lu Y, Pawletko K, Hardes K, Steinmetzer T, Böttcher-Friebertshäuser E. 2015. Influenza virus activating host

proteases: Identification, localization and inhibitors as potential therapeutics. Eur J Cell Biol 94:375-83.

42. Shi Y, Wu Y, Zhang W, Qi J, Gao GF. 2014. Enabling the 'host jump': structural determinants of receptor-binding specificity in influenza A viruses. Nat Rev Microbiol 12:822-31.

43. Nicholls JM, Chan RW, Russell RJ, Air GM, Peiris JS. 2008. Evolving complexities of influenza virus and its receptors. Trends Microbiol 16:149-57.

44. Rogers GN, D'Souza BL. 1989. Receptor binding properties of human and animal H1 influenza virus isolates. Virology 173:317-22.

45. Ma W, Kahn RE, Richt JA. 2008. The pig as a mixing vessel for influenza viruses: Human and veterinary implications. J Mol Genet Med 3:158-66.

46. Trebbien R, Larsen LE, Viuff BM. 2011. Distribution of sialic acid receptors and influenza A virus of avian and swine origin in experimentally infected pigs. Virol J 8:434.

47. Nelli RK, Kuchipudi SV, White GA, Perez BB, Dunham SP, Chang KC. 2010. Comparative distribution of human and avian type sialic acid influenza receptors in the pig. BMC Vet Res 6:4.

48. Ning ZY, Luo MY, Qi WB, Yu B, Jiao PR, Liao M. 2009. Detection of expression of influenza virus receptors in tissues of BALB/c mice by histochemistry. Vet Res Commun 33:895-903.

49. Nelson MI, Worobey M. 2018. Origins of the 1918 Pandemic: Revisiting the Swine "Mixing Vessel" Hypothesis. Am J Epidemiol 187:2498-2502.

50. Gagneux P, Cheriyan M, Hurtado-Ziola N, van der Linden EC, Anderson D, McClure H, Varki A, Varki NM. 2003. Human-specific regulation of alpha 2-6-linked sialic acids. J Biol Chem 278:48245-50.

51. Xu R, McBride R, Nycholat CM, Paulson JC, Wilson IA. 2012. Structural characterization of the hemagglutinin receptor specificity from the 2009 H1N1 influenza pandemic. J Virol 86:982-90.

52. Zhang W, Shi Y, Qi J, Gao F, Li Q, Fan Z, Yan J, Gao GF. 2013. Molecular basis of the receptor binding specificity switch of the hemagglutinins from both the 1918 and 2009 pandemic influenza A viruses by a D225G substitution. J Virol 87:5949-58.

53. Chen Q, Huang S, Chen J, Zhang S, Chen Z. 2013. NA proteins of influenza A viruses H1N1/2009, H5N1, and H9N2 show differential effects on infection initiation, virus release, and cell-cell fusion. PLoS One 8:e54334.

54. Steel J, Lowen AC, Mubareka S, Palese P. 2009. Transmission of Influenza Virus in a Mammalian Host Is Increased by PB2 Amino Acids 627K or 627E/701N. PLOS Pathogens 5:e1000252.

55. Chou YY, Albrecht RA, Pica N, Lowen AC, Richt JA, García-Sastre A, Palese P, Hai R. 2011. The M segment of the 2009 new pandemic H1N1 influenza virus is critical for its high transmission efficiency in the guinea pig model. J Virol 85:11235-41.

56. Chen R, Holmes EC. 2006. Avian influenza virus exhibits rapid evolutionary dynamics. Mol Biol Evol 23:2336-41.

57. Petrova VN, Russell CA. 2018. The evolution of seasonal influenza viruses. Nature Reviews Microbiology 16:47-60.

58. Puenpa J, Vongpunsawad S, Österback R, Waris M, Eriksson E, Albert J, Midgley S, Fischer TK, Eis-Hübinger AM, Cabrerizo M, Gaunt E, Simmonds P, Poovorawan Y. 2016. Molecular epidemiology and the evolution of human coxsackievirus A6. J Gen Virol 97:3225-3231.

59. Chu P-Y, Tyan Y-C, Chen Y-S, Chen H-L, Lu P-L, Chen Y-H, Chen B-C, Huang T-S, Wang C-F, Su H-J, Shi Y-Y, Sanno-Duanda B, Lin K-H, Motomura K. 2015. Transmission and Demographic Dynamics of Coxsackievirus B1. PloS one 10:e0129272-e0129272.

60. Tee KK, Lam TT-Y, Chan YF, Bible JM, Kamarulzaman A, Tong CYW, Takebe Y, Pybus OG. 2010. Evolutionary Genetics of Human Enterovirus 71: Origin, Population Dynamics, Natural Selection, and Seasonal Periodicity of the VP1 Gene. Journal of Virology 84:3339-3350.

61. Poon AFY, Walker LW, Murray H, McCloskey RM, Harrigan PR, Liang RH. 2013. Mapping the shapes of phylogenetic trees from human and zoonotic RNA viruses. PloS one 8:e78122-e78122.

62. Obenauer JC, Denson J, Mehta PK, Su X, Mukatira S, Finkelstein DB, Xu X, Wang J, Ma J, Fan Y, Rakestraw KM, Webster RG, Hoffmann E, Krauss S, Zheng J, Zhang Z, Naeve CW. 2006. Large-scale sequence analysis of avian influenza isolates. Science 311:1576-80.

63. Hurt AC, Vijaykrishna D, Butler J, Baas C, Maurer-Stroh S, Silva-de-la-Fuente MC, Medina-Vogel G, Olsen B, Kelso A, Barr IG, González-Acuña D. 2014. Detection of evolutionarily distinct avian influenza a viruses in antarctica. mBio 5:e01098.

64. Bulach D, Halpin R, Spiro D, Pomeroy L, Janies D, Boyle DB. 2010. Molecular analysis of H7 avian influenza viruses from Australia and New Zealand: genetic diversity and relationships from 1976 to 2007. Journal of virology 84:9957-9966.

65. Rimondi A, Gonzalez-Reiche AS, Olivera VS, Decarre J, Castresana GJ, Romano M, Nelson MI, van Bakel H, Pereda AJ, Ferreri L, Geiger G, Perez DR. 2018. Evidence of a fixed internal gene constellation in influenza A viruses isolated from wild birds in Argentina (2006-2016). Emerg Microbes Infect 7:194.

66. Latorre-Margalef N, Grosbois V, Wahlgren J, Munster VJ, Tolf C, Fouchier RAM, Osterhaus ADME, Olsen B, Waldenström J. 2013. Heterosubtypic immunity to influenza A virus infections in mallards may explain existence of multiple virus subtypes. PLoS pathogens 9:e1003443-e1003443.

67. Stallknecht DE, Shane SM. 1988. Host range of avian influenza virus in free-living birds. Vet Res Commun 12:125-41.

68. Olsen B, Munster VJ, Wallensten A, Waldenstrom J, Osterhaus AD, Fouchier RA. 2006. Global patterns of influenza a virus in wild birds. Science 312:384-8.

69. Webster RG, Hinshaw VS, Bean WJ, Jr., Turner B, Shortridge KF. 1977. Influenza viruses from avian and porcine sources and their possible role in the origin of human pandemic strains. Dev Biol Stand 39:461-8.

70. Easterday BC, Trainer DO, Tumova B, Pereira HG. 1968. Evidence of infection with influenza viruses in migratory waterfowl. Nature 219:523-4.

71. Slemons RD, Easterday BC. 1977. Type-A influenza viruses in the feces of migratory waterfowl. J Am Vet Med Assoc 171:947-8.

72. Slemons RD, Easterday BC. 1978. Virus replication in the digestive tract of ducks exposed by aerosol to type-A influenza. Avian Dis 22:367-77.

73. Easterday BC, Tumova B. 1972. Avian influenza viruses: in avian species and the natural history of influenza. Adv Vet Sci Comp Med 16:201-22.

74. Fouchier RA, Munster VJ. 2009. Epidemiology of low pathogenic avian influenza viruses in wild birds. Rev Sci Tech 28:49-58.

75. Stallknecht DE, Luttrell MP, Poulson R, Goekjian V, Niles L, Dey A, Krauss S, Webster RG. 2012. Detection of avian influenza viruses from shorebirds: evaluation of surveillance and testing approaches. Journal of wildlife diseases 48:382-393.

76. Verhagen JH, Poen M, Stallknecht DE, van der Vliet S, Lexmond P, Sreevatsan S, Poulson RL, Fouchier RAM, Lebarbenchon C. 2020. Phylogeography and Antigenic Diversity of Low-Pathogenic Avian Influenza H13 and H16 Viruses. J Virol 94.

77. Ramey AM, Poulson RL, González-Reiche AS, Wilcox BR, Walther P, Link P, Carter DL, Newsome GM, Müller ML, Berghaus RD, Perez DR, Hall JS, Stallknecht DE. 2014. Evidence for Seasonal Patterns in the Relative Abundance of Avian Influenza Virus Subtypes in Blue-Winged Teal (Anas discors). Journal of Wildlife Diseases 50:916-922, 7.

78. Bogs J, Veits J, Gohrbandt S, Hundt J, Stech O, Breithaupt A, Teifke JP, Mettenleiter TC, Stech J. 2010. Highly pathogenic H5N1 influenza viruses carry virulence determinants beyond the polybasic hemagglutinin cleavage site. PLoS One 5:e11826.

79. Suarez DL. 2010. Avian influenza: our current understanding. Animal Health Research Reviews 11:19-33.

80. Lee DH, Torchetti MK, Killian ML, Berhane Y, Swayne DE. 2017. Highly Pathogenic Avian Influenza A(H7N9) Virus, Tennessee, USA, March 2017. Emerg Infect Dis 23:1860-3.

81. Swayne DE, Suarez DL. 2000. Highly pathogenic avian influenza. Rev Sci Tech 19:463-82.

82. Pantin-Jackwood MJ, Swayne DE. 2009. Pathogenesis and pathobiology of avian influenza virus infection in birds. Rev Sci Tech 28:113-36.

83. Swayne DE, Pantin-Jackwood M. 2006. Pathogenicity of avian influenza viruses in poultry. Dev Biol (Basel) 124:61-7.

84. OIE. 2013-01-30 2013. Outbreaks of highly pathogenic avian influenza (subtype H5N1) in poultry notified to the OIE from the end of 2003 to 30 January 2013, *on* Animal Health Information Department, World Organization for Animal Health. http://www.oie.int/fileadmin/Home/eng/Animal_Health_in_the_World/docs/pdf/graph_avian_influenza/graphs_HPAI_30_01_2013.pdf. Accessed 2013-02-04.

85. WHO. 2020. Avian Influenza Weekly Update Number 723, *on* World Health Organization. https://www.who.int/docs/default-source/wpro---documents/emergency/surveillance/avian-influenza/ai-20200110.pdf?sfvrsn=223ca73f_38. Accessed 2020-01-25.

86. Goekjian VH, Smith JT, Howell DL, Senne DA, Swayne DE, Stallknecht DE. 2011. Avian influenza viruses and avian paramyxoviruses in wintering and breeding waterfowl populations in North Carolina, USA. J Wildl Dis 47:240-5.

87. Hill NJ, Takekawa JY, Cardona CJ, Meixell BW, Ackerman JT, Runstadler JA, Boyce WM. 2012. Cross-seasonal patterns of avian influenza virus in breeding and wintering migratory birds: a flyway perspective. Vector Borne Zoonotic Dis 12:243-53.

88. Hinshaw VS, Webster RG, Turner B. 1980. The perpetuation of orthomyxoviruses and paramyxoviruses in Canadian waterfowl. Canadian Journal of Microbiology 26:622-629.

89. Stallknecht DE, Shane SM, Zwank PJ, Senne DA, Kearney MT. 1990. Avian Influenza Viruses from Migratory and Resident Ducks of Coastal Louisiana. Avian Diseases 34:398-405.

90. Carbonell M, Garvin J. 2002. Know your ducks: A waterfowl identification guide for the Caribbean, Central America and Northern South America. , p 71. Ducks Unlimited.

91. Rappole JH, Morton ES, Lovejoy TE, Ruos JL. 1995. Nearctic avian migrants in the Neotropics: maps, appendices, and bibliography., Second ed. Smithsonian Institution, Washington.

92. Gammonley JH, Frederickson LH. 1995. 13.1.8. Life History and Management of the Blue-winged Teal, p 1-7, Waterfowl Management Handbook 13. U.S. Department of the Interior, National Biological Service, Washington, D.C.

93. Botero JE, Rusch DH. 1988. Recoveries of North American waterfowl in the neotropics. Waterfowl in winter University of Minnesota Press, Minneapolis, MN:469-482.

94. Ramey AM, Walther P, Link P, Poulson RL, Wilcox BR, Newsome G, Spackman E, Brown JD, Stallknecht DE. 2014. Optimizing Surveillance for South American Origin Influenza A Viruses Along the United States Gulf Coast Through Genomic Characterization of Isolates from Blue-winged Teal (Anas discors). Transbound Emerg Dis doi:10.1111/tbed.12244.

95. Diskin ER, Friedman K, Krauss S, Nolting JM, Poulson RL, Slemons RD, Stallknecht DE, Webster RG, Bowman AS. 2020. Subtype Diversity of Influenza A Virus in North American Waterfowl: a Multidecade Study. J Virol 94.

96. Ferro PJ, Budke CM, Peterson MJ, Cox D, Roltsch E, Merendino T, Nelson M, Lupiani B. 2010. Multiyear surveillance for avian influenza virus in waterfowl from wintering grounds, Texas coast, USA. Emerg Infect Dis 16:1224-30.

97. Ferro PJ, Khan O, Peterson MJ, Batchuluun D, Reddy SM, Lupiani B. 2012. Avian influenza virus surveillance in hunter-harvested waterfowl, Texas coast, September 2009-January 2010. Avian Dis 56:1006-9.

98. Hanson BA, Swayne DE, Senne DA, Lobpries DS, Hurst J, Stallknecht DE. 2005. Avian influenza viruses and paramyxoviruses in wintering and resident ducks in Texas. J Wildl Dis 41:624-8.

99. Gonzalez-Reiche AS, Muller ML, Ortiz L, Cordon-Rosales C, Perez DR. 2016. Prevalence and Diversity of Low Pathogenicity Avian Influenza Viruses in Wild Birds in Guatemala, 2010-2013. Avian Dis 60:359-64.

100. Hill NJ, Takekawa JY, Cardona CJ, Meixell BW, Ackerman JT, Runstadler JA, Boyce WM. 2011. Cross-Seasonal Patterns of Avian Influenza Virus in Breeding and Wintering Migratory Birds: A Flyway Perspective. Vector-Borne and Zoonotic Diseases 12:243-253.

101. Ferreri LM, Ortiz L, Geiger G, Barriga GP, Poulson R, Gonzalez-Reiche AS, Crum JA, Stallknecht D, Moran D, Cordon-Rosales C, Rajao D, Perez DR. 2019. Improved detection of influenza A virus from blue-winged teals by sequencing directly from swab material. Ecology and evolution 9:6534-6546.

102. Eisermann K. 2006. Evaluation of waterbird populations and their conservation in Guatemala. Guatemala City, Guatemala.

103. Gonzalez-Reiche AS, Nelson MI, Angel M, Müller ML, Ortiz L, Dutta J, van Bakel H, Cordon-Rosales C, Perez DR. 2017. Evidence of Intercontinental Spread and Uncommon Variants of Low-Pathogenicity Avian Influenza Viruses in Ducks Overwintering in Guatemala. mSphere 2:e00362-16.

104. Senne DA. 2007. Avian influenza in North and South America, 2002-2005. Avian diseases 51:167-73.

105. Gonzalez-Reiche AS, Morales-Betoulle ME, Alvarez D, Betoulle JL, Muller ML, Sosa SM, Perez DR. 2012. Influenza a viruses from wild birds in Guatemala belong to the North American lineage. PLoS One 7:e32873.

106. Gonzalez-Reiche AS, Perez DR. 2012. Where Do Avian Influenza Viruses Meet in the Americas? Avian Diseases 56:1025-1033.

107. Ito T, Couceiro JN, Kelm S, Baum LG, Krauss S, Castrucci MR, Donatelli I, Kida H, Paulson JC, Webster RG, Kawaoka Y. 1998. Molecular basis for the

generation in pigs of influenza A viruses with pandemic potential. J Virol 72:7367-73.

108. Scholtissek C. 1990. Pigs as 'Mixing Vessels' for the Creation of New Pandemic Influenza A Viruses. Medical Principles and Practice 2:65-71.

109. Van Poucke SG, Nicholls JM, Nauwynck HJ, Van Reeth K. 2010. Replication of avian, human and swine influenza viruses in porcine respiratory explants and association with sialic acid distribution. Virol J 7:38.

110. Smith GJ, Vijaykrishna D, Bahl J, Lycett SJ, Worobey M, Pybus OG, Ma SK, Cheung CL, Raghwani J, Bhatt S, Peiris JS, Guan Y, Rambaut A. 2009. Origins and evolutionary genomics of the 2009 swine-origin H1N1 influenza A epidemic. Nature 459:1122-5.

111. Nelson MI, Viboud C, Vincent AL, Culhane MR, Detmer SE, Wentworth DE, Rambaut A, Suchard MA, Holmes EC, Lemey P. 2015. Global migration of influenza A viruses in swine. Nat Commun 6:6696.

112. Walia RR, Anderson TK, Vincent AL. 2019. Regional patterns of genetic diversity in swine influenza A viruses in the United States from 2010 to 2016. Influenza and Other Respiratory Viruses 13:262-273.

113. Zhou NN, Senne DA, Landgraf JS, Swenson SL, Erickson G, Rossow K, Liu L, Yoon K-j, Krauss S, Webster RG. 1999. Genetic Reassortment of Avian, Swine, and Human Influenza A Viruses in American Pigs. Journal of Virology 73:8851-8856.

114. Vincent AL, Ma W, Lager KM, Gramer MR, Richt JA, Janke BH. 2009. Characterization of a newly emerged genetic cluster of H1N1 and H1N2 swine influenza virus in the United States. Virus genes 39:176-185.

115. Lewis NS, Russell CA, Langat P, Anderson TK, Berger K, Bielejec F, Burke DF, Dudas G, Fonville JM, Fouchier RAM, Kellam P, Koel BF, Lemey P, Nguyen T, Nuansrichy B, Peiris JSM, Saito T, Simon G, Skepner E, Takemae N, consortium E, Webby RJ, Van Reeth K, Brookes SM, Larsen L, Watson SJ, Brown IH, Vincent AL. 2016. The global antigenic diversity of swine influenza A viruses. eLife 5:e12217.

116. Simon G, Larsen LE, Dürrwald R, Foni E, Harder T, Van Reeth K, Markowska-Daniel I, Reid SM, Dan A, Maldonado J, Huovilainen A, Billinis C, Davidson I, Agüero M, Vila T, Hervé S, Breum SØ, Chiapponi C, Urbaniak K, Kyriakis CS, consortium E, Brown IH, Loeffen W. 2014. European surveillance network for influenza in pigs: surveillance programs, diagnostic tools and Swine influenza virus subtypes identified in 14 European countries from 2010 to 2013. PloS one 9:e115815-e115815.

117. Hervé S, Garin E, Calavas D, Lecarpentier L, Ngwa-Mbot D, Poliak S, Wendling S, Rose N, Simon G. 2019. Virological and epidemiological patterns of swine influenza A virus infections in France: Cumulative data from the RESAVIP surveillance network, 2011–2018. Veterinary Microbiology 239:108477.

118. Castrucci MR, Campitelli L, Ruggieri A, Barigazzi G, Sidoli L, Daniels R, Oxford JS, Donatelli I. 1994. Antigenic and sequence analysis of H3 influenza virus haemagglutinins from pigs in Italy. J Gen Virol 75 (Pt 2):371-9.

119. Brown IH, Harris PA, McCauley JW, Alexander DJ. 1998. Multiple genetic reassortment of avian and human influenza A viruses in European pigs, resulting in the emergence of an H1N2 virus of novel genotype. The Journal of general virology 79 (Pt 12):2947-2955.

120. Watson SJ, Langat P, Reid SM, Lam TT, Cotten M, Kelly M, Van Reeth K, Qiu Y, Simon G, Bonin E, Foni E, Chiapponi C, Larsen L, Hjulsager C, Markowska-Daniel I, Urbaniak K, Dürrwald R, Schlegel M, Huovilainen A, Davidson I, Dán Á, Loeffen W, Edwards S, Bublot M, Vila T, Maldonado J, Valls L, Brown IH, Pybus OG, Kellam P. 2015. Molecular Epidemiology and Evolution of Influenza Viruses Circulating within European Swine between 2009 and 2013. J Virol 89:9920-31.

121. Shortridge KF, Webster RG. 1979. Geographical distribution of swine (Hsw1N1) and Hong Kong (H3N2) influenza virus variants in pigs in Southeast Asia. Intervirology 11:9-15.

122. Yu H, Hua RH, Zhang Q, Liu TQ, Liu HL, Li GX, Tong GZ. 2008. Genetic evolution of swine influenza A (H3N2) viruses in China from 1970 to 2006. J Clin Microbiol 46:1067-75.

123. Yu H, Zhang PC, Zhou YJ, Li GX, Pan J, Yan LP, Shi XX, Liu HL, Tong GZ. 2009. Isolation and genetic characterization of avian-like H1N1 and novel

ressortant H1N2 influenza viruses from pigs in China. Biochem Biophys Res Commun 386:278-83.

124. Guan Y, Shortridge KF, Krauss S, Li PH, Kawaoka Y, Webster RG. 1996. Emergence of avian H1N1 influenza viruses in pigs in China. J Virol 70:8041-6.

125. Nelson M, Schaefer R, Gava D, Cantão ME, Ciacci-Zanella JR. 2015. Influenza A Viruses of Human Origin in Swine, Brazil. Emerging Infectious Disease journal 21:1339.

126. Rajão DS, Costa ATR, Brasil BSAF, Del Puerto HL, Oliveira FG, Alves F, Braz GF, Reis JKP, Guedes RMC, Lobato ZIP, Leite RC. 2013. Genetic characterization of influenza virus circulating in Brazilian pigs during 2009 and 2010 reveals a high prevalence of the pandemic H1N1 subtype. Influenza and Other Respiratory Viruses 7:783-790.

127. Pereda A, Cappuccio J, Quiroga MA, Baumeister E, Insarralde L, Ibar M, Sanguinetti R, Cannilla ML, Franzese D, Cabrera OEE, Craig MI, Rimondi A, Machuca M, Debenedetti RT, Zenobi C, Barral L, Balzano R, Capalbo S, Risso A, Perfumo CJ. 2010. Pandemic (H1N1) 2009 Outbreak on Pig Farm, Argentina. Emerging Infectious Disease journal 16:304.

128. Cappuccio JA, Pena L, Dibárbora M, Rimondi A, Piñeyro P, Insarralde L, Quiroga MA, Machuca M, Craig MI, Olivera V, Chockalingam A, Perfumo CJ, Perez DR, Pereda A. 2011. Outbreak of swine influenza in Argentina reveals a non-contemporary human H3N2 virus highly transmissible among pigs. Journal of General Virology 92:2871-2878.

129. Nelson M, Culhane MR, Rovira A, Torremorell M, Guerrero P, Norambuena J. 2015. Novel Human-like Influenza A Viruses Circulate in Swine in Mexico and Chile. PLoS currents. 7(ecurrents.outbreaks.c8b3207c9bad98474eca3013fa933c. doi:10.1371/currents.outbreaks.c8b3207c9bad98474eca3013fa933ca6.

130. Nelson MI, Gramer MR, Vincent AL, Holmes EC. 2012. Global transmission of influenza viruses from humans to swine. The Journal of general virology 93:2195-2203.

131. Myers KP, Olsen CW, Gray GC. 2007. Cases of swine influenza in humans: a review of the literature. Clin Infect Dis 44:1084-8.

132. Morens DM, Taubenberger JK. 2014. A possible outbreak of swine influenza, 1892. Lancet Infect Dis 14:169-72.

133. Nelson MI, Vincent AL. 2015. Reverse zoonosis of influenza to swine: new perspectives on the human-animal interface. Trends in microbiology 23:142-153.

134. Nelson MI, Wentworth DE, Culhane MR, Vincent AL, Viboud C, LaPointe MP, Lin X, Holmes EC, Detmer SE. 2014. Introductions and evolution of human-origin seasonal influenza a viruses in multinational swine populations. J Virol 88:10110-9.

135. Gray GC, McCarthy T, Capuano AW, Setterquist SF, Olsen CW, Alavanja MC. 2007. Swine workers and swine influenza virus infections. Emerg Infect Dis 13:1871-8.

136. Krumbholz A, Lange J, Dürrwald R, Hoyer H, Bengsch S, Wutzler P, Zell R. 2010. Prevalence of antibodies to swine influenza viruses in humans with occupational exposure to pigs, Thuringia, Germany, 2008–2009. Journal of Medical Virology 82:1617-1625.

137. Myers KP, Olsen CW, Setterquist SF, Capuano AW, Donham KJ, Thacker EL, Merchant JA, Gray GC. 2006. Are swine workers in the United States at increased risk of infection with zoonotic influenza virus? Clin Infect Dis 42:14-20.

138. Ren H, Jin Y, Hu M, Zhou J, Song T, Huang Z, Li B, Li K, Zhou W, Dai H, Shi W, Yue J, Liang L. 2016. Ecological dynamics of influenza A viruses: cross-species transmission and global migration. Sci Rep 6:36839.

139. Serrano L, Veliz Y, Motta M. 2011. Evaluación y monitoreo nacional de influenza: utilizando la prueba de inhibición de la hemoaglutinación (HI) para detección de cerdos seropositivos. Editorial Académica Española, Guatemala.

140. Gonzalez-Reiche AS, Ramírez AL, Müller ML, Orellana D, Sosa SM, Ola P, Paniagua J, Ortíz L, Hernandez J, Cordón-Rosales C, Perez DR. 2017. Origin, distribution, and potential risk factors associated with influenza A virus in swine in two production systems in Guatemala. Influenza and other respiratory viruses 11:182-192.

141. Lindblade KA, Arvelo W, Gray J, Estevez A, Frenkel G, Reyes L, Moscoso F, Moir JC, Fry AM, Olsen SJ. 2010. A comparison of the epidemiology and clinical presentation of seasonal influenza A and 2009 pandemic influenza A (H1N1) in Guatemala. PLoS One 5:e15826.

142. Wu Y, Wu Y, Tefsen B, Shi Y, Gao GF. 2014. Bat-derived influenza-like viruses H17N10 and H18N11. Trends Microbiol 22:183-91.

143. Kawaoka Y, Yamnikova S, Chambers TM, Lvov DK, Webster RG. 1990. Molecular characterization of a new hemagglutinin, subtype H14, of influenza A virus. Virology 179:759-67.

144. Abat C, Chaudet H, Rolain JM, Colson P, Raoult D. 2016. Traditional and syndromic surveillance of infectious diseases and pathogens. Int J Infect Dis 48:22-8.

145. Hamilton WL, Fieldman T, Jahun A, Warne B, Illingworth CJR, Jackson C, Blane B, Moore E, Weekes MP, Peacock SJ, De Angelis D, Goodfellow I, Gouliouris T, Török ME, Cambridge C-g. 2021. Applying prospective genomic surveillance to support investigation of hospital-onset COVID-19. The Lancet Infectious diseases doi:10.1016/S1473-3099(21)00251-6:S1473-3099(21)00251-6.

146. WHO. 2013. Global Influenza Surveillance and Response System (GISRS), *on* World Health Organization. http://www.who.int/influenza/gisrs_laboratory/en/. Accessed February 12, 2013.

147. Sargento E. 2018. Epitools Epidemiological Calculators. Ausvet. http://epitools.ausvet.com.au. Accessed

148. Spackman E, Suarez DL. 2008. Avian influenza virus RNA extraction from tissue and swab material. Methods Mol Biol 436:13-8.

149. Slomka MJ, Densham AL, Coward VJ, Essen S, Brookes SM, Irvine RM, Spackman E, Ridgeon J, Gardner R, Hanna A, Suarez DL, Brown IH. 2010. Real time reverse transcription (RRT)-polymerase chain reaction (PCR) methods for detection of pandemic (H1N1) 2009 influenza virus and European swine influenza A virus infections in pigs. Influenza Other Respir Viruses 4:277-93.

150. Spackman E, Senne DA, Myers TJ, Bulaga LL, Garber LP, Perdue ML, Lohman K, Daum LT, Suarez DL. 2002. Development of a real-time reverse transcriptase PCR assay for type A influenza virus and the avian H5 and H7 hemagglutinin subtypes. Journal of clinical microbiology 40:3256-3260.

151. WHO. 2002. WHO manual on animal influenza diagnosis and surveillance. World Health Organization, Geneva.

152. Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. 2011 17:3 %J EMBnet.journal.

153. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol 29:644-52.

154. Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25:1754-60.

155. Ferreri LM, Geiger G, Seibert B, Obadan A, Rajao D, Lowen AC, Perez DR. 2022. Intra- and inter-host evolution of H9N2 influenza A virus in Japanese quail. Virus evolution 8:veac001-veac001.

156. Wilm A, Aw PP, Bertrand D, Yeo GH, Ong SH, Wong CH, Khor CC, Petric R, Hibberd ML, Nagarajan N. 2012. LoFreq: a sequence-quality aware, ultra-sensitive variant caller for uncovering cell-population heterogeneity from high-throughput sequencing datasets. Nucleic Acids Res 40:11189-201.

157. Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, del Angel G, Levy-Moonshine A, Jordan T, Shakir K, Roazen D, Thibault J, Banks E, Garimella KV, Altshuler D, Gabriel S, DePristo MA. 2013. From FastQ Data to High-Confidence Variant Calls: The Genome Analysis Toolkit Best Practices Pipeline. Current Protocols in Bioinformatics 43:11.10.1-11.10.33.

158. Grubaugh ND, Gangavarapu K, Quick J, Matteson NL, De Jesus JG, Main BJ, Tan AL, Paul LM, Brackney DE, Grewal S, Gurfield N, Van Rompay KKA, Isern S, Michael SF, Coffey LL, Loman NJ, Andersen KG. 2019. An amplicon-based sequencing framework for accurately measuring intrahost virus diversity using PrimalSeq and iVar. Genome Biol 20:8.

159. McCrone JT, Lauring AS. 2016. Measurements of Intrahost Viral Diversity Are Extremely Sensitive to Systematic Errors in Variant Calling. Journal of virology 90:6884-6895.

160. Simpson EH. 1949. Measurement of Diversity. Nature 163:688-688.

161. Grundmann H, Hori S, Tanner G. 2001. Determining Confidence Intervals When Measuring Genetic Diversity and the Discriminatory Abilities of Typing Methods for Microorganisms. Journal of Clinical Microbiology 39:4190-4192.

162. Dice LR. 1945. Measures of the Amount of Ecologic Association Between Species. Ecology 26:297-302.

163. Sorensen TA. 1948. A method of establishing groups of equal amplitude in plant sociology based on similarity of species content and its application to analyses of the vegetation on Danish commons. Biol Skar 5:1-34.

164. Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792-7.

165. Posada D. 2008. jModelTest: Phylogenetic Model Averaging. Molecular Biology and Evolution 25:1253-1256.

166. Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30:1312-1313.

167. Huang Y, Robertson GJ, Ojkic D, Whitney H, Lang AS. 2014. Diverse inter-continental and host lineage reassortant avian influenza A viruses in pelagic seabirds. Infection, Genetics and Evolution 22:103-111.

168. Baigent SJ, McCauley JW. 2003. Influenza type A in humans, mammals and birds: determinants of virus virulence, host-range and interspecies transmission. Bioessays 25:657-71.

169. Schrauwen EJ, Fouchier RA. 2014. Host adaptation and transmission of influenza A viruses in mammals. Emerging microbes & infections 3:e9-e9.

170. USDA. 2022. HPAI 2022 Confirmed Detections. https://www.aphis.usda.gov/aphis/ourfocus/animalhealth/animal-disease-information/avian/avian-influenza/hpai-2022/2022-hpai-commercial-backyard-flocks. Accessed 05/22/22.

171. Qi X, Jiang D, Wang H, Zhuang D, Ma J, Fu J, Qu J, Sun Y, Yu S, Meng Y, Huang Y, Xia L, Li Y, Wang Y, Wang G, Xu K, Zhang Q, Wan M, Su X, Fu G, Gao GF. 2014. Calculating the burden of disease of avian-origin H7N9 infections in China. BMJ Open 4:e004189.

172. Horimoto T, Kawaoka Y. 2001. Pandemic threat posed by avian influenza A viruses. Clinical microbiology reviews 14:129-149.

173. Vandegrift KJ, Sokolow SH, Daszak P, Kilpatrick AM. 2010. Ecology of avian influenza viruses in a changing world. Annals of the New York Academy of Sciences 1195:113-128.

174. Melville DS, Shortridge KF. 2006. Spread of H5N1 avian influenza virus: an ecological conundrum. Letters in applied microbiology 42:435-7.

175. Butler D, Ruttimann J. 2006. Avian flu and the New World. Nature 441:137-9.

176. Kapczynski DR, Pantin-Jackwood M, Guzman SG, Ricardez Y, Spackman E, Bertran K, Suarez DL, Swayne DE. 2013. Characterization of the 2012 highly pathogenic avian influenza H7N3 virus isolated from poultry in an outbreak in Mexico: pathobiology and vaccine protection. J Virol 87:9086-96.

177. Lopez-Martinez I, Balish A, Barrera-Badillo G, Jones J, Nuñez-García TE, Jang Y, Aparicio-Antonio R, Azziz-Baumgartner E, Belser JA, Ramirez-Gonzalez JE, Pedersen JC, Ortiz-Alcantara J, Gonzalez-Duran E, Shu B, Emery SL, Poh MK, Reyes-Teran G, Vazquez-Perez JA, Avila-Rios S, Uyeki T, Lindstrom S, Villanueva J, Tokars J, Ruiz-Matus C, Gonzalez-Roldan JF, Schmitt B, Klimov A, Cox N, Kuri-Morales P, Davis CT, Diaz-Quiñonez JA. 2013. Highly pathogenic avian influenza A(H7N3) virus in poultry workers, Mexico, 2012. Emerging infectious diseases 19:1531-1534.

178. Li L, DeLiberto TJ, Killian ML, Torchetti MK, Wan X-F. 2018. Evolutionary pathway for the 2017 emergence of a novel highly pathogenic avian influenza A(H7N9) virus among domestic poultry in Tennessee, United States. Virology 525:32-39.

179. Youk S, Lee D-H, Ferreira HL, Afonso CL, Absalon AE, Swayne DE, Suarez DL, Pantin-Jackwood MJ. 2019. Rapid evolution of Mexican H7N3 highly pathogenic avian influenza viruses in poultry. PLOS ONE 14:e0222457.

180. Worobey M, Han GZ, Rambaut A. 2014. A synchronized global sweep of the internal genes of modern avian influenza virus. Nature 508:254-7.

181. Hatta M, Gao P, Halfmann P, Kawaoka Y. 2001. Molecular basis for high virulence of Hong Kong H5N1 influenza A viruses. Science 293:1840-2.

182. Li Z, Chen H, Jiao P, Deng G, Tian G, Li Y, Hoffmann E, Webster RG, Matsuoka Y, Yu K. 2005. Molecular basis of replication of duck H5N1 influenza viruses in a mammalian mouse model. J Virol 79:12058-64.

183. Shinya K, Hamm S, Hatta M, Ito H, Ito T, Kawaoka Y. 2004. PB2 amino acid at position 627 affects replicative efficiency, but not cell tropism, of Hong Kong H5N1 influenza A viruses in mice. Virology 320:258-66.

184. Subbarao EK, London W, Murphy BR. 1993. A single amino acid in the PB2 gene of influenza A virus is a determinant of host range. J Virol 67:1761-4.

185. Yamayoshi S, Yamada S, Fukuyama S, Murakami S, Zhao D, Uraki R, Watanabe T, Tomita Y, Macken C, Neumann G, Kawaoka Y. 2014. Virulence-affecting amino acid changes in the PA protein of H7N9 influenza A viruses. J Virol 88:3127-34.

186. Ayllon J, García-Sastre A. 2015. The NS1 protein: a multitasking virulence factor. Curr Top Microbiol Immunol 386:73-107.

187. Košík I, Hollý J, Russ G. 2013. PB1-F2 expedition from the whole protein through the domain to aa residue function. Acta Virol 57:138-48.

188. Krumbholz A, Philipps A, Oehring H, Schwarzer K, Eitner A, Wutzler P, Zell R. 2011. Current knowledge on PB1-F2 of influenza A viruses. Med Microbiol Immunol 200:69-75.

189. Furuse Y, Suzuki A, Oshitani H. 2009. Large-Scale Sequence Analysis of M Gene of Influenza A Viruses from Different Species: Mechanisms for Emergence

and Spread of Amantadine Resistance. Antimicrobial Agents and Chemotherapy 53:4457-4463.

190. Defensores de la Naturaleza, Ducks Unlimited, USAC. 2007. Conteo de anátidas en Guatemala, informe final temporada 2006/2007.

191. Wille M, Latorre-Margalef N, Tolf C, Halpin R, Wentworth D, Fouchier RAM, Raghwani J, Pybus OG, Olsen B, Waldenström J. 2018. Where do all the subtypes go? Temporal dynamics of H8–H12 influenza A viruses in waterfowl. Virus Evolution 4:vey025.

192. Huang Y, Wille M, Dobbin A, Walzthöni NM, Robertson GJ, Ojkic D, Whitney H, Lang AS. 2014. Genetic Structure of Avian Influenza Viruses from Ducks of the Atlantic Flyway of North America. PLOS ONE 9:e86999.

193. Bahl J, Vijaykrishna D, Holmes EC, Smith GJ, Guan Y. 2009. Gene flow and competitive exclusion of avian influenza A virus in natural reservoir hosts. Virology 390:289-97.

194. Ito T, Okazaki K, Kawaoka Y, Takada A, Webster RG, Kida H. 1995. Perpetuation of influenza A viruses in Alaskan waterfowl reservoirs. Arch Virol 140:1163-72.

195. David ES, Virginia HG, Benjamin RW, Rebecca LP, Justin DB. 2010. Avian Influenza Virus in Aquatic Habitats: What Do We Need to Learn? Avian Diseases 54:461-465.

196. Bergervoet SA, Pritz-Verschuren SBE, Gonzales JL, Bossers A, Poen MJ, Dutta J, Khan Z, Kriti D, van Bakel H, Bouwstra R, Fouchier RAM, Beerens N. 2019. Circulation of low pathogenic avian influenza (LPAI) viruses in wild birds and poultry in the Netherlands, 2006-2016. Sci Rep 9:13681.

197. Boyce WM, Schobel S, Dugan VG, Halpin R, Lin X, Wentworth DE, Lindsay LL, Mertens E, Plancarte M. 2013. Complete Genome Sequence of a Reassortant H14N2 Avian Influenza Virus from California. Genome Announcements 1:e00543-13.

198. Ramey AM, Poulson RL, González-Reiche AS, Perez DR, Stallknecht DE, Brown JD. 2014. Genomic characterization of H14 subtype Influenza A viruses in

new world waterfowl and experimental infectivity in mallards (Anas platyrhynchos). PLoS One 9:e95620.

199. Bowman AS, Nolting JM, Massengill R, Baker J, Workman JD, Slemons RD. 2015. Influenza A Virus Surveillance in Waterfowl in Missouri, USA, 2005-2013. Avian diseases 59:303-308.

200. Nolting J, Fries AC, Slemons RD, Courtney C, Hines N, Pedersen J. 2012. Recovery of H14 influenza A virus isolates from sea ducks in the Western Hemisphere. PLoS currents 4:RRN1290-RRN1290.

201. Fries AC, Nolting JM, Danner A, Webster RG, Bowman AS, Krauss S, Slemons RD. 2013. Evidence for the Circulation and Inter-Hemispheric Movement of the H14 Subtype Influenza A Virus. PLOS ONE 8:e59216.

202. Verhagen JH, van Dijk JGB, Vuong O, Bestebroer T, Lexmond P, Klaassen M, Fouchier RAM. 2014. Migratory birds reinforce local circulation of avian influenza viruses. PloS one 9:e112366-e112366.

203. Rohwer FC, Johnson WP, Loos ER. 2002. Blue-winged Teal: Anas Discors. Birds of North America, Incorporated.

204. Krauss S, Stucker KM, Schobel SA, Danner A, Friedman K, Knowles JP, Kayali G, Niles LJ, Dey AD, Raven G, Pryor P, Lin X, Das SR, Stockwell TB, Wentworth DE, Webster RG. 2015. Long-term surveillance of H7 influenza viruses in American wild aquatic birds: are the H7N3 influenza viruses in wild birds the precursors of highly pathogenic strains in domestic poultry? Emerging microbes & infections 4:e35-e35.

205. Veits J, Weber S, Stech O, Breithaupt A, Gräber M, Gohrbandt S, Bogs J, Hundt J, Teifke JP, Mettenleiter TC, Stech J. 2012. Avian influenza virus hemagglutinins H2, H4, H8, and H14 support a highly pathogenic phenotype. Proceedings of the National Academy of Sciences 109:2579-2584.

206. Kuiken T, Fouchier R, Rimmelzwaan G, van den Brand J, van Riel D, Osterhaus A. 2011. Pigs, Poultry, and Pandemic Influenza: How Zoonotic Pathogens Threaten Human Health, p 59-66. *In* Curtis N, Finn A, Pollard AJ (ed), Hot Topics in Infection and Immunity in Children VIII doi:10.1007/978-1-4614-0204-6_6. Springer New York, New York, NY.

207. Ma MJ, Wang GL, Anderson BD, Bi ZQ, Lu B, Wang XJ, Wang CX, Chen SH, Qian YH, Song SX, Li M, Lednicky JA, Zhao T, Wu MN, Cao WC, Gray GC. 2018. Evidence for Cross-species Influenza A Virus Transmission Within Swine Farms, China: A One Health, Prospective Cohort Study. Clin Infect Dis 66:533-540.

208. Nelson MI, Stratton J, Killian ML, Janas-Martindale A, Vincent AL. 2015. Continual Reintroduction of Human Pandemic H1N1 Influenza A Viruses into Swine in the United States, 2009 to 2014. Journal of virology 89:6218-6226.

209. VISAR-MAGA. P. 2021. Swine census report data base 2021.

210. Anderson TK, Macken CA, Lewis NS, Scheuermann RH, Van Reeth K, Brown IH, Swenson SL, Simon G, Saito T, Berhane Y, Ciacci-Zanella J, Pereda A, Davis CT, Donis RO, Webby RJ, Vincent AL. 2016. A Phylogeny-Based Global Nomenclature System and Automated Annotation Tool for H1 Hemagglutinin Genes from Swine Influenza A Viruses. mSphere 1:e00275-16.

211. Russell RJ, Haire LF, Stevens DJ, Collins PJ, Lin YP, Blackburn GM, Hay AJ, Gamblin SJ, Skehel JJ. 2006. The structure of H5N1 avian influenza neuraminidase suggests new opportunities for drug design. Nature 443:45-49.

212. Hai R, Schmolke M, Varga ZT, Manicassamy B, Wang TT, Belser JA, Pearce MB, García-Sastre A, Tumpey TM, Palese P. 2010. PB1-F2 expression by the 2009 pandemic H1N1 influenza virus has minimal impact on virulence in animal models. Journal of virology 84:4442-4450.

213. Shaw M, Cooper L, Xu X, Thompson W, Krauss S, Guan Y, Zhou N, Klimov A, Cox N, Webster R, Lim W, Shortridge K, Subbarao K. 2002. Molecular changes associated with the transmission of avian influenza a H5N1 and H9N2 viruses to humans. J Med Virol 66:107-14.

214. Saito R, Sakai T, Sato I, Sano Y, Oshitani H, Sato M, Suzuki H. 2003. Frequency of amantadine-resistant influenza A viruses during two seasons featuring cocirculation of H1N1 and H3N2. J Clin Microbiol 41:2164-5.

215. Baudon E, Peyre M, Peiris M, Cowling BJ. 2017. Epidemiological features of influenza circulation in swine populations: A systematic review and meta-analysis. PLOS ONE 12:e0179044.

216. Descalzo MA, Clara W, Guzmán G, Mena R, Armero J, Lara B, Saenz C, Aragón A, Chacón R, El-Omeiri N, Méndez-Rico J, Cerpa M, Palekar R, Jara J, Azziz-Baumgartner E. 2016. Estimating the burden of influenza-associated hospitalizations and deaths in Central America. Influenza and other respiratory viruses 10:340-345.

217. Hutchinson EC, von Kirchbach JC, Gog JR, Digard P. 2010. Genome packaging in influenza A virus. Journal of General Virology 91:313-328.

218. Davis AR, Hiti AL, Nayak DP. 1980. Influenza defective interfering viral RNA is formed by internal deletion of genomic RNA. Proceedings of the National Academy of Sciences 77:215-219.

219. Noble S, Dimmock NJ. 1995. Characterization of Putative Defective Interfering (DI) A/WSN RNAs Isolated from the Lungs of Mice Protected from an Otherwise Lethal Respiratory Infection with Influenza Virus A/WSN (H1N1): A Subset of the Inoculum DI RNAs. Virology 210:9-19.

220. Dimmock NJ, Easton AJ. 2014. Defective Interfering Influenza Virus RNAs: Time To Reevaluate Their Clinical Potential as Broad-Spectrum Antivirals? Journal of Virology 88:5217.

APPENDIX

LIST OF REAGENTS/SUPPLIES

The following table contains a comprehensive list of reagents and supplies used to perform the different laboratory assays of this research.

Table A.1. List of reagents/supplies used for this research work

Reagent/supply	Brand	Catalog	Procedure
Sodium bicarbonate	Merck	1370132500	sample collection
Sterile foam swabs	Puritan	25-1506-1PF	sample collection
Gentamicine	Merck	345814-1GM	sample collection
Sterilization filters (0.12 um)	ThermoFisher Scientific	450-0020	sample collection
Albumin from bovine serum	Sigma Aldrich	A2153	sample collection
L-Glutamine	Sigma Aldrich	G7513	sample collection
Medium 199	Sigma Aldrich	M9163	sample collection
Nystatin	Sigma Aldrich	N6261	sample collection
Ofloxacin	Sigma Aldrich	O8757	sample collection
Polymyxin B	Sigma Aldrich	P4932	sample collection
Penicillin	Sigma Aldrich	P7794	sample collection
Sulfamethoxazole	Sigma Aldrich	S7507 FLUKA	sample collection
Streptomicine	Sigma Aldrich	S9137	sample collection
TRIzol LS Reagent	ThermoFisher Scientific	10296-028	RNA extraction
Glycogen	ThermoFisher Scientific	AM9510	RNA extraction
Chloroform	Sigma Aldrich	C2432-500ML	RNA extraction
Isopropanol (molecular grade)	Sigma Aldrich	9516-500mL	RNA extraction
MagMAX AI/ND Viral RNA extraction Kit	ThermoFisher Scientific	AM1835	RNA extraction

Reagent/supply	Brand	Catalog	Procedure
QIAamp Viral RNA Mini Kit	Qiagen	52906	RNA extraction
Ethanol absolute for analysis	MERCK	1009835000	RNA extraction
QuantiTect Probe RT-PCR kit	Qiagen	204445	RRT-PCR
MicroAmp optical adhesive film	ThermoFisher Scientific	4311971	RRT-PCR
M+25 primer (STD purification)	IDT	---	RRT-PCR
M-124 primer (STD purification)	IDT	---	RRT-PCR
M-124*SIV primer (STD purification)	IDT	---	RRT-PCR
M+64 FAM labeled probe (HPLC purification)	IDT	---	RRT-PCR
Ribonuclease inhibitor	Promega	N2515	RRT-PCR
MicroAmp optical 96-well rxn plate	ThermoFisher Scientific	N8010560	RRT-PCR
SuperScript III One-Step RT-PCR system	ThermoFisher Scientific	12574035	MS-RTPCR
OptiF1 primer (HPLC purification)	IDT	---	MS-RTPCR
OptiF2 primer (HPLC purification)	IDT	---	MS-RTPCR
OptiR2 primer (HPLC purification)	IDT	---	MS-RTPCR
SYBR Safe	ThermoFisher Scientific	S33102	Agarose gels
GeneRuler 1kb Plus ladder	ThermoFisher Scientific	SM0311	Agarose gels
TBE Buffer 10X (molecular biology grade)	Calbiochem	US1574795-1L	Agarose gels
Agarose (analytical grade)	Promega	V3125	Agarose gels
9-day-old SPF eggs	---	---	Viral isolation
Syringe filters	ThermoFisher Scientific	09-720-000	Viral isolation
Syringes (1mL)	ThermoFisher Scientific	14-826-88	Viral isolation
Antibiotic/antimycotic solution	VWR	45000-616	Viral isolation

Reagent/supply	Brand	Catalog	Procedure
PBS (ultra-pure grade)	VWR	10010023	Viral isolation/HA
96 well V-bottom plates	VWR	29442-402	HA
Alsevers solution	Sigma Aldrich	A3551	HA
Buffer EB (elution buffer)	Qiagen	19086	sequencing
Index adapter replacement caps (orange)	Illumina	15026585	sequencing
Index adapter replacement caps (white)	Illumina	15026586	sequencing
Magnetic stand for tubes	ThermoFisher Scientific	12321D	sequencing
Bioanalyzer high sensitivity DNA Analysis	Agilent	5067-4626	sequencing
Sodium hydroxide solution	Sigma Aldrich	72068-100ML	sequencing
Agencourt AMPure XP beads	Agencourt	A63880	sequencing
EDTA	Sigma Aldrich	E6635-500G	sequencing
PhiX control v3 (10nM)	Illumina	FC-110-3001	sequencing
Nextera XT DNA Library Prep	Illumina	FC-131-1096	sequencing
Nextera XT Index Kit v2 Set A	Illumina	FC-131-2001	sequencing
MiSeq Reagent Nano Kit v2 (300-cycles)	Illumina	MS-103-1001	sequencing
Qubit dsDNA HS assay Kit	ThermoFisher Scientific	Q32854	sequencing
Qubi assay tubes	ThermoFisher Scientific	Q32856	sequencing
Trizma hydrochloride	Sigma Aldrich	T1535	sequencing
Tween 20	J. T. Baker	X251-07	sequencing
Gloves	MidSci	78886	general
Ethanol (for 70% ethanol solution)	ThermoFisher Scientific	22032601	general
Distilled water	---	---	general
Cryogenic vial storage boxes	ThermoFisher Scientific	07-200-615	general
10uL filter tips	USA Scientific	1180-3810	general
200 uL filter tips	USA Scientific	1180-8810	general

Reagent/supply	Brand	Catalog	Procedure
1000 uL filter tips	USA Scientific	1182-1830	general
20 uL filter tips	USA Scientific	1183-1810	general
100 uL filter tips	USA Scientific	1183-1840	general
0.2mL strip tubes	USA Scientific	1402-4708	general
Microcentrifuge tubes (2mL)	USA Scientific	1420-2799	general
Microcentrifuge tubes (1.5mL)	USA Scientific	1615-5599	general
Screw-cap tubes (1.5mL)	Genesse	21-256	general
Biohazard Autoclave bags (mini)	Genesse	30-162R	general
Bleach	Grainger	41H893	general
Storage boxes and dividers	VWR	82007-154	general
5mL serological pipettes	VWR	82050-478	general
10mL serological pipettes	VWR	82050-482	general
25 ml serological pipettes	VWR	82051-182	general
50 ml serological pipettes	VWR	82051-212	general
Centrifuge tubes	VWR	89039-668	general
Large label for laser printers	USA Scientific	9187-2016	general
RNaseZAP	Sigma Aldrich	R2020-250ML	general
Water (molecular biology grade)	VWR	82007-332	general