Genetic and evolutionary dynamics of avian influenza A virus in wild birds

by Jessica Benkaroun

A thesis submitted to the School of Graduate Studies

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Biology, Faculty of Science

Memorial University of Newfoundland

St. John's, Newfoundland and Labrador, Canada

September 2018

Abstract

Influenza A virus (IAV) is the prototype of the family *Orthomyxoviridae*, a group of segmented, negative-sense, single-stranded RNA viruses. The virus circulates in wild bird species but does not usually cause severe disease in these hosts. However, highly pathogenic forms exist and have caused numerous deaths in wild and farmed birds. The eastern coast of Canada represents an interesting location for the study of IAVs in their natural reservoir as it has a large number of bird breeding colonies and migratory bird connections with the mainland of North America and Eurasia. Previous research on IAV ecology and transmission has shown that migratory birds in this region move the virus around the globe and contribute an important facet to IAV dynamics. My thesis focuses on the study of the virus genetics and evolutionary dynamics in different wild bird species. By applying high-throughput next-generation sequencing technologies, I characterized complete IAV genomes from different wild bird species from Newfoundland and Labrador and conducted in-depth analyses of the virus genomic structure. My study revealed that the structure of the virus genome is conserved among similar avian hosts. I also demonstrated though experimental mutation studies that a change of host can causes major changes in the viral genome. I also explored evolutionary patterns in the viral genomic non-coding regions (NCRs), and found that variation in the NCR sequences is correlated with the original host species and geographic origin. Finally, I analyzed IAVs from Laridae family hosts (gulls and terns) and demonstrated that

these hosts are important for the transmission of IAVs around the globe and to other hosts and participate in the generation of pandemic viruses. Overall, my results contribute to give a better understanding on the evolution and geographic patterns of influenza A viruses in their natural hosts.

Acknowledgments

First of all, I would like to extend my gratitude to my supervisors, Dr. Andrew Lang and Dr. Hugh Whitney, for giving me the opportunity to take part and guide me through this thesis project. I thank you for your support that allowed me to complete this work.

I would like to thank the members of my thesis committee, Dr. Sherri Christian, and Dr. Kensuke Hirasawa, for their valuable help in bringing this project to fruition.

Many thanks to all members of Andrew's lab who gave me precious help and support. A deep thank you to Marc Grüll, Dr. Marta Canuti, and Purvikalyan Pallegar.

I am grateful to all MUN students I met during my thesis: Esteban, Elena, Paula, Boni, Nathalia, Abdou, Daigo, Damiano, Anastasia, Oihane, Rosa, Pramod, Amrita, Carlos, Fabio, Fritz, and Lizbeth.

Thank you to my friends from home: Magali, Ananda, Delphine, Olivier, Mako, Julien, Stephane, and Yan.

An eternal gratitude to my parents and my little sisters for their listening and advice.

Thank you to my husband's family for their support: Jesus, Irene, Amparo, Rosa, Nubia, David, and Viviana. *Gracias por estar en nuestros vidas.*

At last, a deep and grateful thank you to my love, my companion, you have been admirable and deeply supportive throughout this journey. I love you John Ceballos.

Je dédie cette thèse à mes parents, Christine Waille et Didier Benkaroun, à mes sœurs Julie Waille et Rachel Waille. Je vous remercie de tout mon cœur pour votre soutien inconditionnel tout au long de ce séjour.

My thesis project was supported by funding from the Newfoundland and Labrador Research & Development Corporation, Environment and Climate Change Canada, and the Newfoundland and Labrador Forestry and Agrifoods Agency. I received a fellowship from the Memorial University School of Graduate Studies and the Newfoundland and Labrador Forestry and Agrifoods Agency.

Table of Contents

List of Tables

List of Figures

List of Abbreviations and Symbols

AIV: avian influenza virus

PB1: polymerase basic 1 protein

PB2: polymerase basic 2 protein

CDS: coding sequence

cRNA: complementary RNA

dN: non-synonymous substitution frequency

d_N/d_S: Ratio of the number of non-synonymous to synonymous substitutions

dS: synonymous substitution frequency

ER: endoplasmic reticulum

H': Shannon diversity index

HA: hemagglutinin A

HEF: hemagglutinin-esterase-fusion glycoprotein

HPAI: highly pathogenic avian influenza virus

IAV: influenza A virus

IFN: interferon

ISP : ion sphere particles

LPAI: low pathogenic avian influenza virus

M1: matrix 1 protein

M2: matrix 2 protein

MDS: classical multidimensional analysis

mRNA: messenger RNA

NA: neuraminidase A

NCR: non-coding region

NEP: nuclear export protein

NP: nucleoprotein

NS: non-structural gene

NS1: non-structural protein 1

PA: polymerase acidic protein

PCR: polymerase chain reaction

RACE: rapid amplification of cDNA ends

RBD: receptor binding site

RIG-I: retinoic acid-inducible gene I

RNP: ribonucleoprotein

RT-PCR: Reverse transcription polymerase chain reaction

TMRCA: time to most recent common ancestor

vRNA: viral RNA

List of Appendices

Supporting documents from chapter 2

Supporting documents from chapter 3

Supporting documents from chapter 4

Other scientific publications

Structural insights into the HIV-1 minus-strand strong-stop DNA.

Yingying Chen, Ouerdia Maskri, Françoise Chaminade, Brigitte René, Jessica Benkaroun, Julien Godet, Yves Mély, Olivier Mauffret, Philippe Fossé. The Journal of Biological Chemistry, 291(7):3468-82, 2016

Perpetuation and reassortment of gull influenza A viruses in Atlantic North America.

Huang Y., M. Wille, J. Benkaroun, H. Munro, A.L. Bond, D.A. Fifield D, G.J. Robertson, D. Ojkic, H. Whitney, and A.S. Lang. Virology 456-457: 353-363, 2014

Emerging complexities of APOBEC3G action on immunity and viral fitness during HIV infection and treatment.

Monajemi M, Woodworth CF, Benkaroun J, Grant M, Larijani M. Retrovirology, 30;9:35, 2012

Co-authorship Statement

This thesis work was completed with the collaboration of: Dr. Andrew Lang from the Department of Biology, at Memorial University of Newfoundland, Canada; Dr. Hugh Whitney from the Newfoundland and Labrador Forestry and Agrifoods Agency, Canada; Dr. Gregory Robertson from the Wildlife Research Division at Environment and Climate Change Canada; and Dr. Dany Shoham from Begin-Sadat Center for Strategic Studies in Bar-Ilan University, Israel.

Chapters 1, 2, and 5 were drafted by myself and revised in response to comments from my supervisors, Dr. Andrew Lang and Dr. Hugh Whitney, and supervisory committee members, Drs. Sherri Christian and Kensuke Hirasawa.

Chapter 3 is a version of a research article published in the journal *Cogent Biology* (Jessica Benkaroun, Dany Shoham, Ashley N.K. Kroyer, Hugh Whitney and Andrew S. Lang. Analysis of influenza A viruses from gulls: An evaluation of inter-regional movements and interactions with other avian and mammalian influenza A viruses. *Cogent Biology* (2016), 2: 1234957). I conducted all analyses in this research paper with the help of Ashley N.K. Kroyer for retrieving nucleotide sequences and performing phylogenetic analyses (M.Sc. student). Dr. Dany Shoham and I participated in the writing of the article with the help of Drs. Andrew Lang and Hugh Whitney.

Chapter 4 is a version of a research article published in the journal *Veterinary Sciences* (Jessica Benkaroun, Gregory J. Robertson, Hugh Whitney, and Andrew S.

Lang. Analysis of the variability in the non-coding regions of influenza A viruses. *Veterinary Sciences* (2018), 5(3): 76). I was responsible for all analyses performed in this article. I wrote the article with the help of Dr. Andrew Lang and with feedback from co-authors Dr. Gregory J. Robertson and Dr. Hugh Whitney prior to publication.

Chapter 1: Introduction and Overview

1.1 The viral family *Orthomyxoviridae*

Influenza A virus (IAV) belongs to the family *Orthomyxoviridae*, a family of viruses with segmented, negative-sense, single-stranded RNA genomes, that contains six genera, *Influenzavirus A*, *B*, and *C*, *Thogotovirus*, *Isavirus*, and *Quaranjavirus* (Figure 1.1) (Baltimore 1971, Lefkowitz, Dempsey et al. 2018).

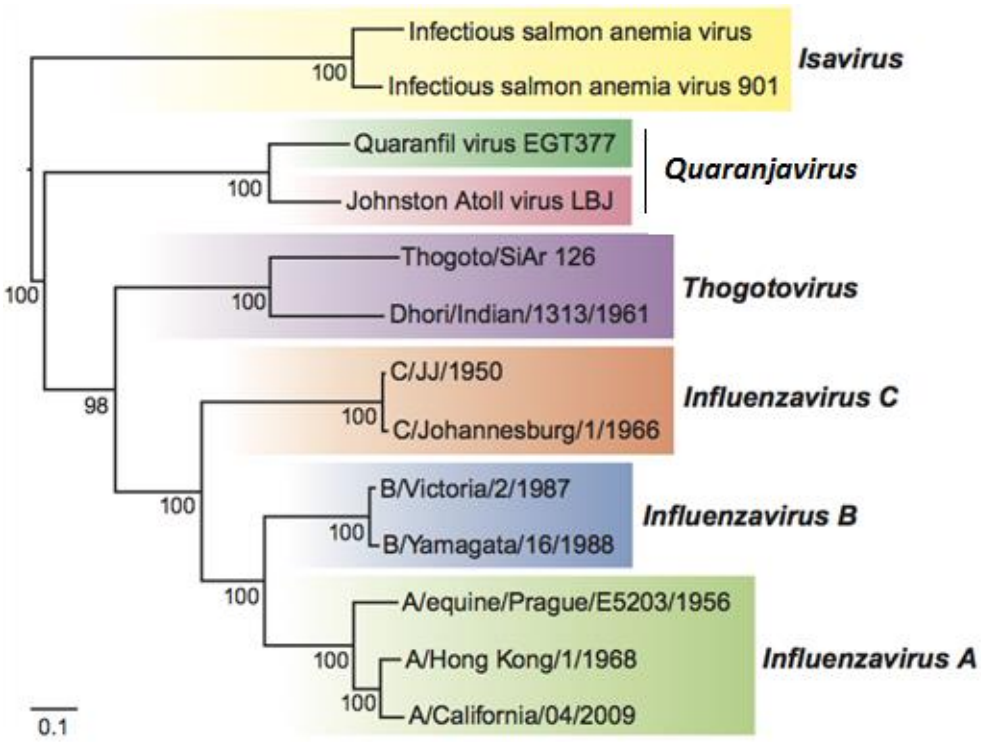


Figure 1.1: *Orthomyxoviridae* genera phylogeny. The phylogenetic tree was made based on PB1 nucleotide sequences and shows the relationships among genera within the family (Adapted from Lefkowitz, Dempsey et al. 2018, material free to use).

The genus *Isavirus* only contains the virus species, *Infectious salmon anemia virus*, which is restricted to fish and causes major disease outbreaks in farmed salmon production (Mjaaland, Rimstad et al. 1997).

The genus *Quaranjavirus* contains two virus species, *Quaranfil virus* and *Johnston Atoll virus*. These viruses were first isolated from ticks in Egypt and the Johnston Atoll island in the Pacific ocean, respectively (Presti, Zhao et al. 2009). They have subsequently been isolated from birds (Presti, Zhao et al. 2009). Only the *Quaranfil virus* infects humans according to serological studies (Clifford, Thomas et al. 1968, Baskerville and Lloyd 1976). No signs of disease have been observed in humans or birds for either virus, but infection by those viruses causes high mortality in mice (Clifford, Thomas et al. 1968, Baskerville and Lloyd 1976).

The genus *Thogotovirus* contains two virus species, *Thogoto virus* and *Dhori virus*, which are both tick-borne viruses that can infect humans and other mammals. Both viruses have been found in Africa, Europe, India, and North America (Williams, Hoogstraal et al. 1973).

The genus *Influenzavirus C* contains viruses that infect humans, dogs, and pigs (Guo, Jin et al. 1983, Moriuchi, Katsushima et al. 1991, Manuguerra and Hannoun 1992). The virus was isolated in the late 1940s from a man with mild clinical signs of disease (Taylor 1949). Only mild respiratory symptoms are associated with this virus in humans (Moriuchi, Katsushima et al. 1991). Frequent interspecies transmission occurs between human and pigs (Kimura, Abiko et al. 1997).

The genus *Influenzavirus B* contains viruses that have been isolated from humans and seals (Hiromoto, Saito et al. 2000, Osterhaus, Rimmelzwaan et al. 2000).

Two distinct virus lineages also co-circulate in the human population. In humans, the virus is associated with mild to severe respiratory disease, but rarely causes mortality (Glezen, Schmier et al. 2013).

The genus *Influenzavirus A* harbours viruses that are known to have caused the deadliest pandemic in humans which occurred in 1918 and is also called the Spanish flu. Following this pandemic, the virus was isolated a decade later from pigs in the early 1930s in North America (Shope 1931) and some years later from a human patient (Smith, 1933). Multiple subtypes of the virus exist and it is classified by two of its genes, HA and NA, which are important in infection and transmission. There are 18 different HA types and 11 NA types, 16 HA and 9 NA have been identified in birds, and 2 HA and 2 NA in bats (Webster, Bean et al. 1992, Tong, Li et al. 2012). The virus is known to cause yearly outbreaks with more rare severe pandemics in the human population and frequently causes outbreaks in the swine and the poultry industries (Simonsen, Clarke et al. 1998, Capua and Marangon 2000, Horimoto and Kawaoka 2001, Girard, Tam et al. 2010, Vincent, Awada et al. 2014).

In comparison to influenza A virus, both influenza B and C have unique subtypes and evolve relatively slowly (Suzuki and Nei 2002, Bedford, Suchard et al. 2014). Their narrow genetic diversity likely accounts for their more limited host range. Based on evolutionary studies, influenza C virus is more distant to A and B (Suzuki and Nei 2002). Moreover, the influenza C genome contains 7 segments whereas influenza A and B genomes are made of 8 segments. Both influenza A and B viruses possess genes encoding both surface glycoproteins, hemagglutinin and neuraminidase, while influenza C viruses possess one gene encoding for a protein

that harbours both hemagglutinin and neuraminidase functions, the hemagglutinin-esterase-fusion glycoprotein (HEF) (Wang and Veit 2016).

1.2 Influenza A virus structure and genomic organization

1.2.1 The virus structure

IAV particles range from 80 to 120 nm in diameter and have a spherical capsid surrounded by an envelope derived from the host cell cytoplasmic membrane (Figure 1.2).

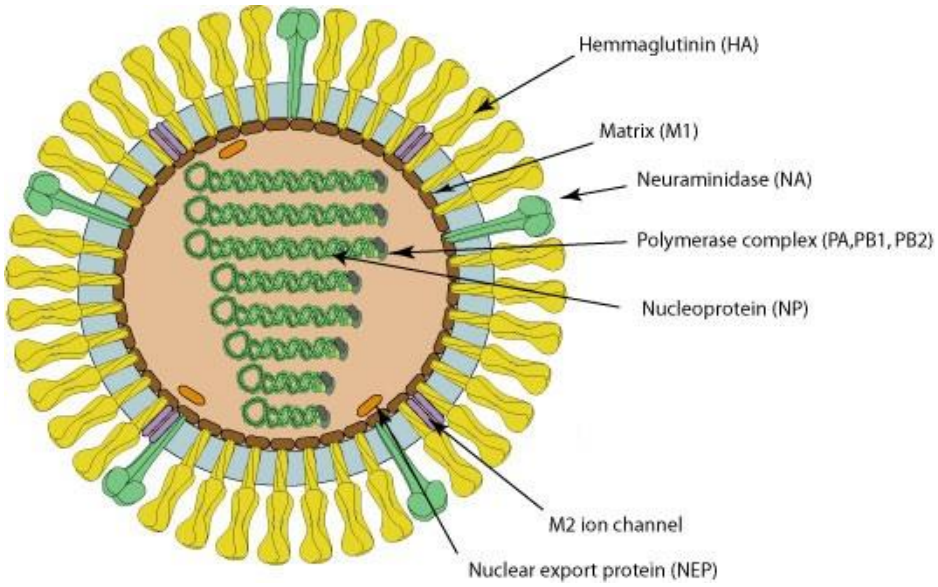


Figure 1.2: Influenza A virus structure (From Hulo, de Castro et al. 2011, permission to use this material is granted by the Swiss Institute of Bioinformatics). The virus has an outer membrane where the glycoproteins HA (yellow) and NA (green) and matrix protein M2 (purple) are embedded. The matrix protein M1 underlays the membrane (brown). Within the capsid is the nuclear export protein (NEP; orange) and the viral genome made of 8 gene segments, with each bound to a polymerase complex and nucleoproteins (NP).

The virus envelope harbours two types of surface glycoproteins: the hemagglutinin A (HA) and neuraminidase A (NA). The HA protein (~60 kDa) is encoded by one gene (~1.8 Kb), contains two domains, HA1 and HA2, and exists as a trimer in the virus envelope (Wilson, Skehel, & Wiley, 1981). The HA1 domain is located in the N-terminal region of the protein and is formed by a globular head that contains receptor binding sites that interact with cellular sialic acid receptors to allow the entry of the virus into the host cell (Edinger, Pohl et al. 2014). The HA2

domain is in the C-terminal part of the protein and contains a pH-sensitive fusion domain that allows the release of the viral contents into the host cell following conformational changes. A hydrophobic transmembrane region is also located in the C-terminal region that anchors the protein in the membrane (Wilson, Skehel et al. 1981).

The NA protein (~60 kDa) is involved in virus exit from the host cell (Gamblin and Skehel 2010). This glycoprotein is present as a tetramer on the virus envelope (Varghese, Laver et al. 1983). The protein is encoded by one gene (~1.6 Kb) and contains in the N-terminal portion a conserved hydrophobic transmembrane region responsible for the protein anchoring into the viral envelope, and a stalk region that varies in size that is thought to be related to the virus pathogenicity. A conserved catalytic domain is located in the C-terminal region that forms the globular head of the protein. This catalytic domain is responsible for the enzymatic activity of the protein, which catalyzes the cleavage of glycosidic bonds from cellular sialic acid receptors that are bound to HA proteins, to prevent the aggregation of new viral particles on the surface of the host cell (Palese and Compans 1976).

Within the inner part of the virion are embedded matrix 1 protein (M1) dimers and matrix 2 protein (M2) tetramers. The M1 and M2 proteins are encoded on the same segment (~1 Kb) with their mRNAs generated by differential splicing (Dubois, Terrier et al. 2014). The M1 protein (~28 kDa) has multiple functions. It is responsible for maintaining the integrity of the virus particle, for importing the viral genome into the cell nucleus, and for the virus budding from the host cell membrane (Burleigh, Calder et al. 2005). The M2 protein (~11 Da) is a proton channel that is

responsible for altering the pH inside the virion once the virus is inside an endocytic vesicle, which causes a conformational change in the HA protein that exposes its fusion domain and results in fusion of the viral envelope with the vesicle membrane, thereby releasing the virion's contents within the host cell (Pinto, Holsinger et al. 1992).

The virion contains the nuclear export protein (NEP) (~13 kDa) that is encoded by the non-structural (NS) segment (~0.9 Kb) and responsible for the viral genome export from the nucleus of the host cell cytoplasm (Boulo, Akarsu et al. 2007). The NS segment also encodes an additional protein, produced by differential splicing, the non-structural protein 1 (NS1) (~26 kDa) (Dubois, Terrier et al. 2014). The NS1 protein has been described as an antagonist to the host type I interferon (IFN) immune responses (Lin, Lan et al. 2007). Type I IFN responses correspond to the production of IFNα and β proteins by the host and these display antiviral properties (Weber and Haller 2007). Deletion studies on the NS segment have confirmed its crucial role in preventing host IFN responses upon virus infection (Garcia-Sastre, Egorov et al. 1998). Inhibition of the cellular retinoic acid-inducible gene I (RIG-I) pathway, also involved in initiation of immune responses, by NS1 has also been documented (Hale, Albrecht et al. 2010).

1.2.2 The virus genome

The influenza A virus genome contained in the virion is formed by 8 negative-sense RNA segments that are coated by the nucleoprotein (NP) (~60 kDa) and bound with a polymerase complex. The polymerase complex (~260 kDa) is formed by three

subunits, the polymerase acidic protein (PA), the polymerase basic 1 protein (PB1), and polymerase basic 2 protein (PB2) (Figure 1.3) (Pflug, Guilligay et al. 2014).

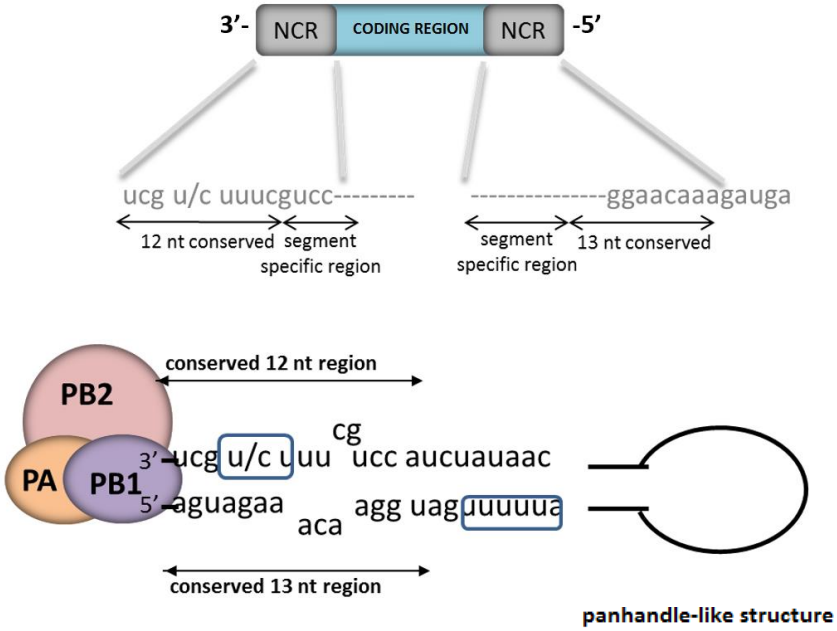


Figure 1.3: Influenza A segment organization. The influenza A genome is composed of 8 negative-sense gene segments, which vary in size from 2.3 Kb to 0.9 Kb. Each coding region is flanked by non-coding regions (NCRs) at both extremities. The 3' NCR is composed of 12 conserved nucleotides followed by a segment-specific region of various lengths. The 5' NCR is composed of 13 conserved nucleotides followed as well by a segment-specific region of various lengths. Both NCRs have inverted partial complementary sequences that allow formation of a panhandle-like structure and that acts as the promoter for transcription. The polymerase complex made of PB2, PB1, and PA proteins binds the promoter region to initiate the transcription and replication of the virus.

This entire structure of RNA, NP and polymerase complex form the ribonucleoprotein complex (RNP). Each genomic segment contains non-coding regions (NCRs) at its extremities that are partially complementary and form a panhandle structure (Hsu, Parvin et al. 1987). This structure constitutes the promoter where the polymerase complex binds and initiates transcription and replication. Additional accessory proteins with various functions have been characterised that are encoded by alternative splicing or leaky scanning mechanisms (Chen, Calvo et al. 2001, Wise, Foeglein et al. 2009, Wise, Hutchinson et al. 2012, Muramoto, Noda et al. 2013, Yamayoshi, Watanabe et al. 2016).

1.3 Influenza A virus life cycle

The first step of the virus life cycle consists of the virus entry into the host cell (Figure 1.4). The virus entry into the host cell is mediated by the viral glycoprotein HA, which recognizes sialic acid receptors present on the host cell surface in a specific manner. After receptor recognition, the interaction leads to the internalization of the virus within the cytoplasm of the host cell in endosomal vesicles. The next step consists of the release of the virus genome from endosomes into the cytoplasm of the host cell. This step is triggered by the M2 protein, which allows the influx of protons into the virion from the endosomes that contain the virus, which acidify the virus capsid environment. This acidification activates a conformational change in the HA protein that exposes its fusion domain, leading to fusion of the endosome and envelope membranes to release the RNP complexes into the cytoplasm (Stegmann 2000).

RNP complexes are imported into the nucleus with the help of the NEP protein. In the nucleus, each segment is transcribed and replicated by the viral polymerase complex. The virus genome is transcribed into positive-sense complementary RNA (cRNA) followed by replication of negative-sense copies that are packaged into new virus particles. In a concomitant manner, the viral gene segments are transcribed into messenger RNA (mRNA) and converted into mature mRNA by a process of cap snatching from cellular mRNA (Dias, Bouvier et al. 2009).

Viral mRNAs are exported from the nucleus to the cytoplasm to be translated into proteins, with some processed through the host endoplasmic reticulum (ER) and the golgi systems. New virions are formed by neo-synthesized structural viral proteins M1, M2, HA, and NA. HA proteins are subjected to post-translational modifications at the ER including glycosylation and proteolytic cleavage, which leads to mature and functional HA proteins. HA proteins are cleaved into two domains, HA1 and HA2, by specific cellular enzymes (Chen, Lee et al. 1998). The viral polymerase complex along with the NP proteins and viral RNAs (vRNAs) are packaged into new virus particles.

Mature infectious particles are released from the host cell by budding from the cytoplasmic membrane. The viral glycoprotein NA is known to facilitate the exit of the virus from the host cell by cleaving sialic acids bound to the HA proteins on newly assembled virions.

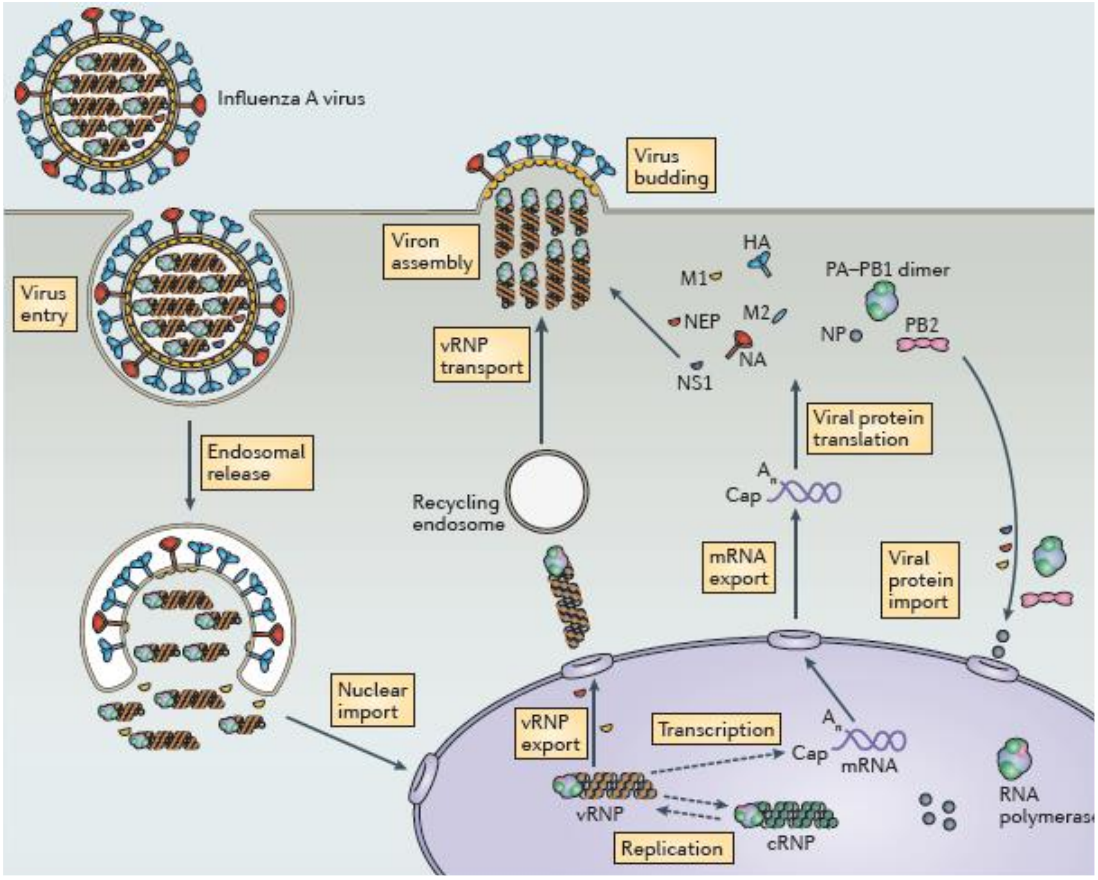


Figure 1.4: Influenza virus life cycle (From Te Velthuis and Fodor 2016, permission to use this material is granted). The virus is internalized in cellular endosomes after receptor recognition with a host cell. The virus is released into the cell's cytoplasm and the viral ribonucleoproteins (vRNP) are imported into the nucleus. The viral genome is replicated (via cRNP) and transcribed into mRNA. Both mRNA and vRNP are exported to the cytoplasm. mRNAs are translated to produce the viral proteins. The new vRNP are packaged into new virions along with viral proteins at the plasma membrane.

1.4 Influenza A virus reservoir and host transmission

1.4.1 IAV in birds

Aquatic wild birds constitute the main reservoir of IAV. The avian orders Anseriformes (waterfowl) and Charadriiformes (gull and shorebirds) form the primary reservoir of the virus and carry the majority of IAV subtypes (Webster, Bean et al. 1992). To date, 18 HA and 11 NA subtypes have been characterized. Among those subtypes, H1 to H16 and N1 to N9 subtypes are known to circulate in birds, (Webster, Bean et al. 1992, Olsen, Munster et al. 2006), while H17, H18, N10, and N11 subtypes have been identified solely in bats (Figure 1.5) (Tong, Li et al. 2012, Tong, Zhu et al. 2013).

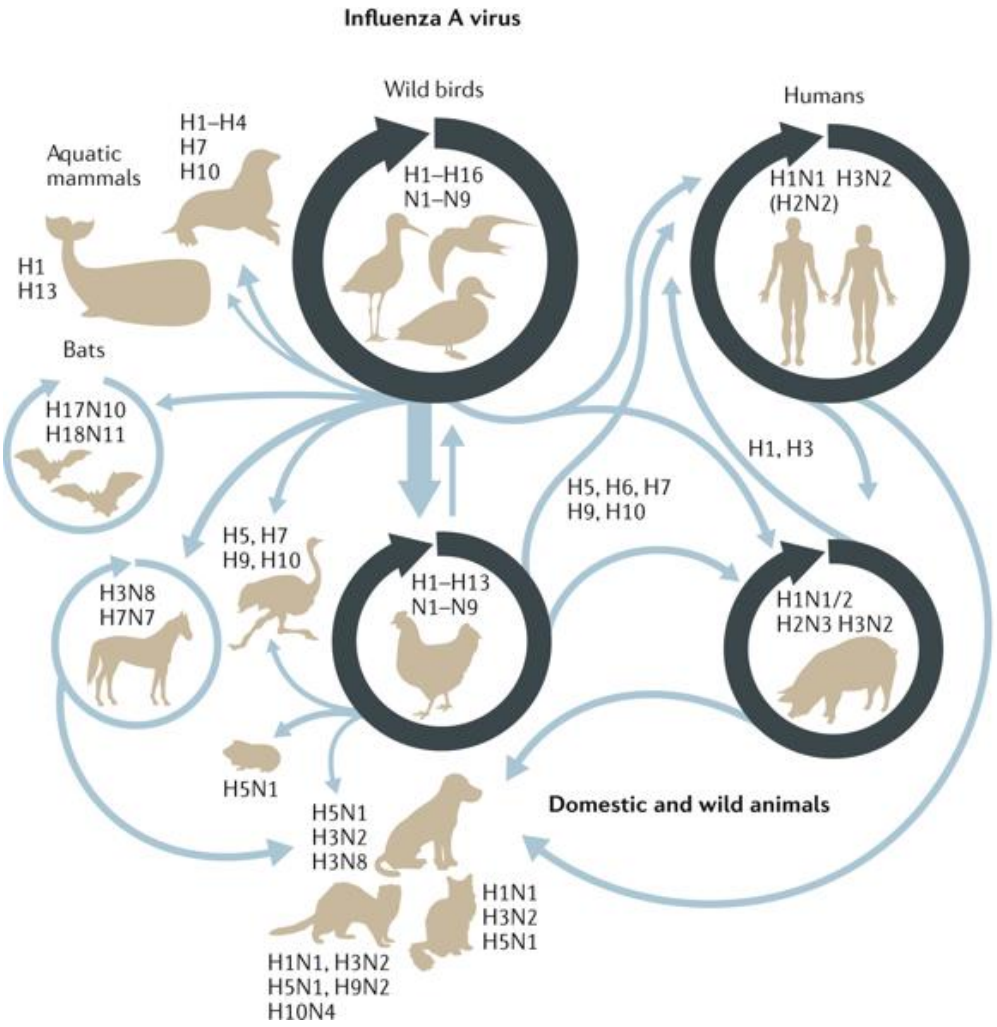


Figure 1.5: Influenza A subtype distribution across the virus' reservoir and other hosts (Adapted from Long, Mistry et al. 2019, permission to use this material is granted). 16 HA and 9 NA subtypes have been identified and circulate in wild birds. Some subtypes are restricted and only identified in certain hosts such as the H3N8 and H7N7 subtypes that only circulate in horses or the H1-H3 subtypes in pigs.

IAV transmission in birds occurs through the fecal oral route via feces or from the environment (Hinshaw, Webster et al. 1979, Webster, Bean et al. 1992). Studies have revealed that IAV can reside in water for long periods of time, over 6 months,

and this could contribute to virus transmission and perpetuation (Stallknecht, Kearney et al. 1990, Stallknecht, Shane et al. 1990). In wild birds, the virus replicates mostly in the epithelial cells of the intestinal tract and infection is usually asymptomatic or causes mild clinical symptoms, while in domestic Galliformes (chicken, quail, and turkey) the virus is known to be more frequently associated with high mortality rates due to systemic infections (Naeem and Hussain 1995, Shortridge 1999, Latorre-Margalef, Gunnarsson et al. 2009, Jourdain, Gunnarsson et al. 2010, Pasick, Berhane et al. 2015). Systemic infection is thought to occur due to the presence of multiple basic amino acids in the HA protein's cleavage site (Nao, Yamagishi et al. 2017). This causes a variety of enzymes to recognize and cleave HA proteins at the cleavage site, which allows viral replication to spread to multiple tissues and organs. To date, only the H5 and H7 subtypes have caused severe mortality in poultry, while the other subtypes are mostly asymptomatic (Spalding 2009).

1.4.2 IAV inter-species transmission

Avian influenza virus (AIV) strains usually do not infect humans efficiently. Strong barriers (viral and host factors) usually prevent the transmission of AIVs to humans and other mammals (Ito and Kawaoka 2000). The host cell surface receptor that allows the entry of the virus (sialic acid receptors) into the host cell is one of the factors responsible for the species barrier. The HA protein of avian strains bind preferentially to Siaα2,3Gal receptors that are mostly found on avian epithelial cells while the HA protein of human strains have affinity for Siaα2,6Gal receptors found on human epithelial cells (Couceiro, Paulson et al. 1993, Matrosovich, Matrosovich

et al. 2004). The host-virus specificity is also influenced by a difference in the distribution of these cellular receptors. Humans also possess Siaα2,3Gal receptors recognized by avian strains, but these are located in the lower respiratory tract and in smaller quantities (Kumlin, Olofsson et al. 2008).

Historically, avian strains have been involved in pandemic outbreaks in humans. The human H2N2 virus that caused a deadly pandemic in 1957 in Asia is believed to have originated from the exchange of genes between an avian H2N2 virus and a circulating human H1N1 virus. Phylogenetic studies of gene sequences from H2N2 viruses isolated from humans during the outbreak showed that three genes, HA, NA, and PB1, were related to avian strains from Eurasia. In contrast, the other genes were associated with human H1N1 viruses circulating in the human population (Kawaoka, Krauss et al. 1989, Schafer, Kawaoka et al. 1993).

AIV inter-species transmission also commonly occurs with swine. Swine harbour both Siaα2,3Gal and Siaα2,6Gal receptors on their epithelial cells, so they support the replication of avian and human strains and are considered as an intermediate host involved in the generation of recombinant viruses through reassortment (Ito, Couceiro et al. 1998, Brown 2001). Indeed, the human pandemic H1N1 virus from 2009 contained avian, swine and human virus genes (de Silva and Yasunaga 2011). Characterization and phylogenetic analyses of IAVs isolated from whales have shown that some of the virus genes are closely related to those from gulls, suggesting IAV inter-species transmission (Groth, Lange et al. 2014).

1.4.3 IAV in swine

To date, three major subtypes are currently circulating in the swine population: H1N1, H3N2, and H1N2. The first identification of IAV in swine occurred in North America in the 1930s and is thought to have originated from the human pandemic H1N1 virus from 1918 (Shope 1931, Reid and Taubenberger 2003). Later in the 1930s, the same virus was isolated in Europe in swine (Blakemore F 1941). This virus, also called classical swine virus, circulated in the swine population until it got replaced by another H1N1 virus in the 1990s that has an Eurasian and avian origin (Schultz, Fitch et al. 1991). The H3N2 subtype was first detected in the 1970s in the swine population and was introduced from humans (Harkness, Schild et al. 1972). The H1N2 virus subtype is the result of mixed infections between co-circulating H1N1 and H3N2 viruses (Brown, Harris et al. 1998).

The role of swine in the generation of human pandemic outbreaks is thought to be facilitated by the presence of two types of cellular sialic acid receptors, Siaα2,6Gal and Siaα2,3Gal, present on the epithelial cell surface of the swine trachea (Ito, Couceiro et al. 1998). Siaα2,6Gal are mostly found in mammalian host epithelial tissues while Siaα2,3Gal are present in avian species' tissues. The presence of both receptors allows co-infections of viruses that have avian and mammalian origins, which could result in the generation of new viruses with mixed-species origins and this is known to have potential for pandemic virus generation.

1.4.4 IAV in humans

In humans, the virus mostly replicates in the respiratory tract (Taubenberger and Morens 2008) and transmission usually occurs through direct contact with

infected people by aerosol droplet inhalation. Two major virus subtypes currently co-circulate in the human population, H1N1 and H3N2 (Finkelman, Viboud et al. 2007). The virus does not efficiently infect immunocompetent individuals, with pre-existing immunity. It is typically mostly young children, elderly persons, and immunocompromised individuals that are at risk of mortality from IAV infection (Simonsen 1999).

1.4.5 IAV in marine mammals

Serological studies revealed the occurrence of IAV infections in marine mammals such as seals and whales (Kida, Brown et al. 1982, Nielsen, Clavijo et al. 2001). Several mortality events associated with IAV in seals have also been reported (Geraci, Staubin et al. 1982, Hinshaw, Bean et al. 1984). In all cases, the virus had an avian origin and marine birds such as gulls have been associated with IAV transmission to whales (Hinshaw, Bean et al. 1986, Groth, Lange et al. 2014).

1.5 Influenza A virus diversity and evolutionary mechanisms

IAVs evolve rapidly due to mutational and gene exchange mechanisms that contribute to the generation of modifications within the virus genome. Two major events create diversity and are responsible for the fast accumulation of changes within the virus genome: antigenic drift, which corresponds to the introduction of random mutations, and antigenic shift, which corresponds with the exchange of virus genes between virus strains (Figure 1.6).

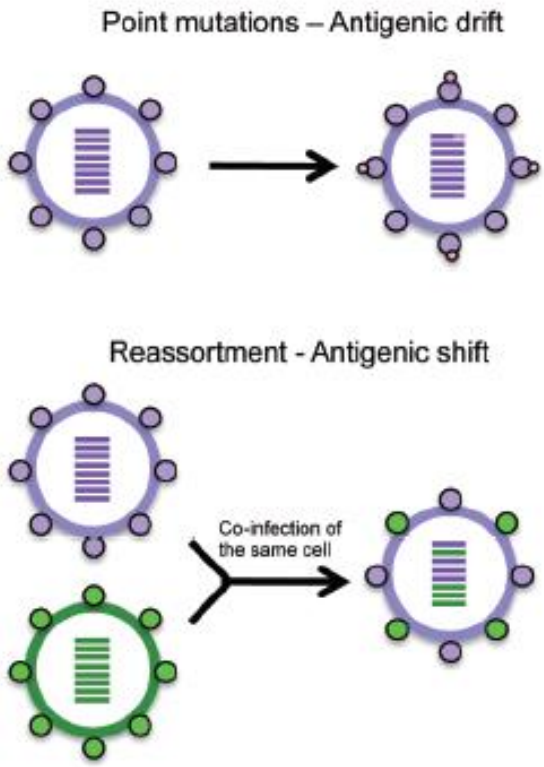


Figure 1.6: IAV antigenic shift and drift mechanisms (From Thi H. O. Nguyen 2016, permission to use this material is granted). Antigenic drift corresponds to random nucleotide mutations occurring within the virus gene segments. Antigenic shift corresponds to the shuffling of gene segments of two viruses that could have different subtypes. This leads to the formation of virus progenies with genes from both parental viruses.

1.5.1 Random mutations and gene reassortments

Random point mutations appear because the viral RNA polymerase replicase enzyme is error-prone and lacks proof-reading activity during the replication of the viral RNA. The viral polymerase has a high error rate, estimated at around 10^{-5}

mutations per genome replication (Drake, 1993), which creates diversity within the virus genome. Gene exchange by reassortment occurs when two different viruses infect the same cell. The co-infection can lead to an exchange of one or several gene segments during assembly. The combination of these two events can enable the virus to evolve and potentially gain the ability to infect new hosts. In some cases, the infection of a new host can be responsible for deadly outbreaks. For instance, the human H3N2 virus that caused a pandemic in humans in 1968 had acquired two genes, HA and PB1, from an avian H3 virus strain circulating in birds (Wendel, Rubbenstroth et al. 2015). Similarly, the human pandemic H2N2 virus, as mentioned earlier, was a product of reassortment of avian and human viruses (Udayan Joseph 2015). The introduction of an HA gene segment from an avian host in humans seems to have an important impact for inter-species transmissions of the virus. However, the introduction of genes does not seem to be sufficient to explain the efficient adaptation of new viruses. The H2N2 virus disappeared from the human population after the pandemic, while other pandemic human viruses, such as the H3N2 virus, are still endemic in the human population (Westgeest, Russell et al. 2014, Joseph, Linster et al. 2015). Subsequent adaptive mutations in the IAV genome are certainly required for viral establishment and circulation in a new host. Experimental infections in ferrets, which are used as an animal model to mimic IAV infection in humans, with H5 subtype viruses revealed the appearance of mutations within the HA segment that enabled the virus to be efficiently transmitted among ferrets (Imai, Watanabe et al. 2012).

1.5.2 Other mutational mechanisms

Homologous recombination occurs between similar gene segments of closely related IAVs. For most research studies, it is considered a rare event in negative-sense RNA viruses (Boni, Zhou et al. 2008) but other studies have shown its involvement in the evolutionary dynamics of IAV (He, Xie et al. 2009) (He, Han et al. 2008, Hao 2011).

Heterologous recombination occurs between non-related gene segments, and was observed with the NP segment of the virus A/seal/Mass/1/80 (H7N7) that contains a region corresponding to the HA segment (Orlich, Gottwald et al. 1994). This mutational insertion also increased the pathogenicity of the recombinant virus in chickens.

1.5.3 The role of the host immune system

Selective pressure from the host immune system also contributes to the virus' evolution (Shao, Li et al. 2017). It can trigger the selection of pre-existing mutants from a pool of related viruses, called quasispecies, and lead to the emergence of viruses with different phenotypes such as a better replicative fitness, or being able to recognise new entry cell receptors (Figure 1.7) (Domingo, 1998).

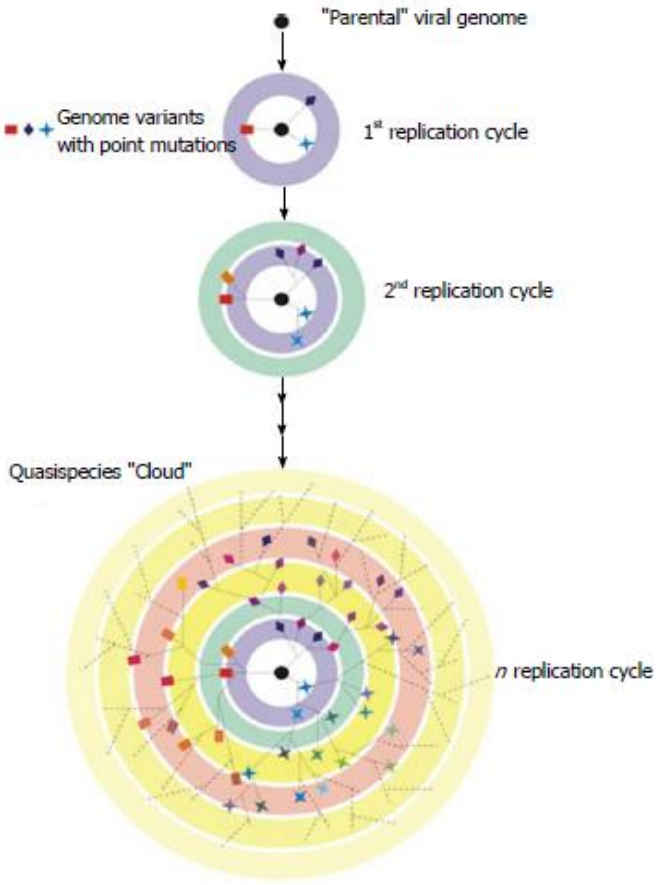


Figure 1.7: Virus quasispecies formation. A pool of related viruses is formed at each round of replication that harbour different mutations within their genomes (From Echeverria, 2015, permission to use this material is granted).

1.6 IAV in Newfoundland and Labrador, Canada

Newfoundland and Labrador is the most eastern province of Canada. The island of Newfoundland is surrounded by the Atlantic Ocean in the east and the Gulf of St. Lawrence in the west. In particular, the island of Newfoundland contains a large number of ponds and lakes which are common breeding grounds of different duck species such as American black ducks (*Anas rubripes*), mallards (*Anas*

platyrhynchos), and northern pintails (*Anas acuta*). During the summer period, many different migratory birds inhabit the island for breeding. The Atlantic flyway overlaps the province and brings migratory birds from the Americas and Eurasia. This creates the potential for vast circulation of IAV by migratory birds around the globe and facilitates the perpetuation and transmission of the virus among different species. However, IAV is not homogenized across the globe. Geographic isolation of the virus has caused it to evolve into distinct genetic lineages based on geographic origin. These genetic lineages correspond to groups of viruses that share common genetic composition. Because of this, a virus' origin (host and geographic location) can be retraced with evolutionary trees (Penny, Hendy et al. 1992). Evolutionary reconstructions for IAV have shown that it groups into two major geographic genetic lineages, a North American lineage and a Eurasian lineage (Olsen, Munster et al. 2006). Similarly, two major bird host group-specific viral lineages have evolved, avian and gull (Olsen, Munster et al. 2006).

Previous research on the island of Newfoundland has shown that the virus circulates yearly in the duck population (Huang, Wille et al. 2014, Long, Mistry et al. 2019). During a 4-year epidemiological study, a virus prevalence of 7.2% was detected, principally in American black ducks in the autumn period. Genetic studies of those viruses have revealed that they largely originate from a North American lineage and from waterfowl hosts (Long, Mistry et al. 2019). Seabird species found in Newfoundland and Labrador were also previously investigated (Wille, Huang et al. 2014, Thi H. O. Nguyen 2016). From the different seabirds investigated for the presence of IAVs, most viruses were identified in common murres. These viruses

originated mostly from waterfowl and displayed a higher rate of gene exchanges with viruses from Eurasia in comparison to the viruses that were found in ducks. This could be attributed to their migration across continents. Viruses isolated from gull species from Newfoundland and Labrador were also investigated and found to have a low active infection prevalence of 1.8% but a high seroprevalence of 50% (Wille, Huang et al. 2014). The role of gulls in moving the virus over long distances was also shown.

1.7 Thesis Aim and Outline

In this thesis, I explore the genetic and evolutionary dynamics of IAVs across different natural reservoir bird species. In chapter 2, I characterized and analysed the genomic population structure and evolution of AIVs isolated from different wild birds in Newfoundland and Labrador. To highlight variation and specificity that could be involved in AIV evolution and transmission, I characterized and investigated the genomic non-coding regions of different wild bird viruses in Chapter 3. Among wild birds carrying AIVs, the Laridae (gulls and terns) are one important host group that contributes to the dynamics of AIV transmission and evolution. In Chapter 4, I analyzed the relatedness of a variety of viruses isolated from gulls and terns to those from other hosts to highlight their role in virus transmission among host groups and across large distances, and for being potential contributors to the generation of pandemic viruses. In Chapter 5, I summarize my findings and suggest new research perspectives.

1.8 References

Baltimore, D. (1971). "Expression of animal virus genomes." Bacteriological Reviews 35(3): 235-241.

Baskerville, A. and G. Lloyd (1976). "The pathogenesis and pathology of experimental Quaranfil virus infection." British journal of experimental pathology 57(2): 152-156.

Bedford, T., M. A. Suchard, P. Lemey, G. Dudas, V. Gregory, A. J. Hay, J. W. McCauley, C. A. Russell, D. J. Smith and A. Rambaut (2014). "Integrating influenza antigenic dynamics with molecular evolution." Elife 3: e01914.

Blakemore F, G. A. (1941). "Swine Influenza in the British Isles." Proceedings of the Royal Society of Medicine 34(9): 611-618.

Boni, M. F., Y. Zhou, J. K. Taubenberger and E. C. Holmes (2008). "Homologous recombination is very rare or absent in human influenza A virus." Journal of Virology 82(10): 4807-4811.

Boulo, S., H. Akarsu, R. W. Ruigrok and F. Baudin (2007). "Nuclear traffic of influenza virus proteins and ribonucleoprotein complexes." Virus Res 124(1-2): 12-21.

Brown, I. H. (2001). "The pig as an intermediate host for influenza A viruses between birds and humans." International Congress Series 1219: 173-178.

Brown, I. H., P. A. Harris, J. W. McCauley and D. J. Alexander (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.

Burleigh, L. M., L. J. Calder, J. J. Skehel and D. A. Steinhauer (2005). "Influenza a viruses with mutations in the m1 helix six domain display a wide variety of morphological phenotypes." Journal of Virology 79(2): 1262-1270.

Capua, I. and S. Marangon (2000). "The avian influenza epidemic in Italy, 1999-2000: a review." Avian Pathology 29(4): 289-294.

Chen, J., K. H. Lee, D. A. Steinhauer, D. J. Stevens, J. J. Skehel and D. C. Wiley (1998). "Structure of the hemagglutinin precursor cleavage site, a determinant of influenza pathogenicity and the origin of the labile conformation." Cell 95(3): 409-417.

Chen, W., P. A. Calvo, D. Malide, J. Gibbs, U. Schubert, I. Bacik, S. Basta, R. O'Neill, J. Schickli, P. Palese, P. Henklein, J. R. Bennink and J. W. Yewdell (2001). "A novel influenza A virus mitochondrial protein that induces cell death." Nat Med 7(12): 1306-1312.

Clifford, C. M., L. A. Thomas, L. E. Hughes, G. M. Kohls and C. B. Philip (1968). "Identification and comparison of two viruses isolated from ticks of the genus Ornithodoros." The American Journal of Tropical Medicine and Hygiene 17(6): 881-885.

Couceiro, J. N., J. C. Paulson and L. G. Baum (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(2): 155-165.

de Silva, U. C. and T. Yasunaga (2011). "Swine to human transmission of reassortants of pandemic (H1N1) 2009 and endemic swine influenza viruses: Abstract." PLoS Currents 3: RRN1285.

Dias, A., D. Bouvier, T. Crepin, A. A. McCarthy, D. J. Hart, F. Baudin, S. Cusack and R. W. H. Ruigrok (2009). "The cap-snatching endonuclease of influenza virus polymerase resides in the PA subunit." Nature 458(7240): 914-918.

Dubois, J., O. Terrier and M. Rosa-Calatrava (2014). "Influenza viruses and mRNA splicing: doing more with less." MBio 5(3): e00070-00014.

Edinger, T. O., M. O. Pohl and S. Stertz (2014). "Entry of influenza A virus: host factors and antiviral targets." Journal of General Virology 95: 263-277.

Finkelman, B. S., C. Viboud, K. Koelle, M. J. Ferrari, N. Bharti and B. T. Grenfell (2007). "Global Patterns in Seasonal Activity of Influenza A/H3N2, A/H1N1, and B from 1997 to 2005: Viral Coexistence and Latitudinal Gradients." PLoS One 2(12): e1296.

Fodor, A. J. W. t. V. a. E. (2016). "Influenza virus RNA polymerase: insights into the mechanisms of viral RNA synthesis." Nature reviews. Microbiology 14(8): 479-493.

Gamblin, S. J. and J. J. Skehel (2010). "Influenza Hemagglutinin and Neuraminidase Membrane Glycoproteins." The Journal of Biological Chemistry 285(37): 28403-28409.

Garcia-Sastre, A., A. Egorov, D. Matassov, S. Brandt, D. E. Levy, J. E. Durbin, P. Palese and T. Muster (1998). "Influenza A virus lacking the NS1 gene replicates in interferon-deficient systems." Virology 252(2): 324-330.

Geraci, J. R., D. J. Staubin, I. K. Barker, R. G. Webster, V. S. Hinshaw, W. J. Bean, H. L. Ruhnke, J. H. Prescott, G. Early, A. S. Baker, S. Madoff and R. T. Schooley (1982). "Mass mortality of harbor seals: pneumonia associated with influenza A virus." Science 215(4536): 1129-1131.

Girard, M. P., J. S. Tam, O. M. Assossou and M. P. Kieny (2010). "The 2009 A (H1N1) influenza virus pandemic: A review." Vaccine 28(31): 4895-4902.

Glezen, W. P., J. K. Schmier, C. M. Kuehn, K. J. Ryan and J. Oxford (2013). "The Burden of Influenza B: A Structured Literature Review." American Journal of Public Health 103(3): E43-E51.

Groth, M., J. Lange, P. Kanrai, S. Pleschka, C. Scholtissek, A. Krumbholz, M. Platzer, A. Sauerbrei and R. Zell (2014). "The genome of an influenza virus from a pilot whale: Relation to influenza viruses of gulls and marine mammals." Infection Genetics and Evolution 24: 183-186.

Guo, Y. J., F. G. Jin, P. Wang, M. Wang and J. M. Zhu (1983). "Isolation of influenza C virus from pigs and experimental infection of pigs with influenza C virus." The Journal of General Virology 64 (Pt 1): 177-182.

Hale, B. G., R. A. Albrecht and A. García-Sastre (2010). "Innate immune evasion strategies of influenza viruses." Future microbiology 5: 23.

Hao, W. (2011). "Evidence of intra-segmental homologous recombination in influenza A virus." Gene 481(2): 57-64.

Harkness, J. W., G. C. Schild, P. H. Lamont and C. M. Brand (1972). "Studies on relationships between human and porcine influenza: 1. Serological evidence of infection in swine in Great Britain with an influenza A virus antigenically like human Hong Kong/68 virus." Bulletin of the World Health Organization 46(6): 709-719.

He, C. Q., G. Z. Han, D. Wang, W. Liu, G. R. Li, X. P. Liu and N. Z. Ding (2008). "Homologous recombination evidence in human and swine influenza A viruses." Virology 380(1): 12-20.

He, C. Q., Z. X. Xie, G. Z. Han, J. B. Dong, D. Wang, J. B. Liu, L. Y. Ma, X. F. Tang, X. P. Liu, Y. S. Pang and G. R. Li (2009). "Homologous recombination as an evolutionary force in the avian influenza A virus." Molecular Biology and Evolution 26(1): 177-187.

Hinshaw, V. S., W. J. Bean, J. Geraci, P. Fiorelli, G. Early and R. G. Webster (1986). "Characterization of two influenza A viruses from a pilot whale." Journal of Virology 58(2): 655-656.

Hinshaw, V. S., W. J. Bean, R. G. Webster, J. E. Rehg, P. Fiorelli, G. Early, J. R. Geraci and D. J. St Aubin (1984). "Are seals frequently infected with avian influenza viruses?" Journal of Virology 51(3): 863-865.

Hinshaw, V. S., R. G. Webster and B. Turner (1979). "Water-bone transmission of influenza A viruses?" Intervirology 11(1): 66-68.

Hiromoto, Y., T. Saito, S. E. Lindstrom, Y. Li, R. Nerome, S. Sugita, M. Shinjoh and K. Nerome (2000). "Phylogenetic analysis of the three polymerase genes (PB1, PB2 and PA) of influenza B virus." The Journal of General Virology 81(Pt 4): 929-937.

Horimoto, T. and Y. Kawaoka (2001). "Pandemic threat posed by avian influenza A viruses." Clinical Microbiology Reviews 14(1): 129-+.

Hsu, M. T., J. D. Parvin, S. Gupta, M. Krystal and P. Palese (1987). "Genomic RNAs of influenza viruses are held in a circular conformation in virions and in infected cells by a terminal panhandle." Proceedings of the National Academy of Sciences of the United States of America 84(22): 8140-8144.

Huang, Y., G. J. Robertson, D. Ojkic, H. Whitney and A. S. Lang (2014). "Diverse inter-continental and host lineage reassortant avian influenza A viruses in pelagic seabirds." Infection, Genetics and Evolution : Journal of molecular epidemiology and evolutionary genetics in infectious diseases 22: 103-111.

Huang, Y., M. Wille, J. Benkaroun, H. Munro, A. L. Bond, D. A. Fifield, G. J. Robertson, D. Ojkic, H. Whitney and A. S. Lang (2014). "Perpetuation and reassortment of gull influenza A viruses in Atlantic North America." Virology 456-457: 353-363.

Huang, Y., M. Wille, A. Dobbin, G. J. Robertson, P. Ryan, D. Ojkic, H. Whitney and A. S. Lang (2013). "A 4-year study of avian influenza virus prevalence and subtype

diversity in ducks of Newfoundland, Canada." Canadian Journal of Microbiology 59(10): 701-708.

Huang, Y., M. Wille, A. Dobbin, N. M. Walzthoni, G. J. Robertson, D. Ojkic, H. Whitney and A. S. Lang (2014). "Genetic structure of avian influenza viruses from ducks of the Atlantic flyway of North America." PLoS One 9(1): e86999.

Hulo, C., E. de Castro, P. Masson, L. Bougueleret, A. Bairoch, I. Xenarios and P. Le Mercier (2011). "ViralZone: a knowledge resource to understand virus diversity." Nucleic Acids Research 39(Database issue): D576-582.

Imai, M., T. Watanabe, M. Hatta, S. C. Das, M. Ozawa, K. Shinya, G. Zhong, A. Hanson, H. Katsura, S. Watanabe, C. Li, E. Kawakami, S. Yamada, M. Kiso, Y. Suzuki, E. A. Maher, G. Neumann and Y. Kawaoka (2012). "Experimental adaptation of an influenza H5 haemagglutinin (HA) confers respiratory droplet transmission to a reassortant H5 HA/H1N1 virus in ferrets." Nature 486(7403): 420-428.

Ito, T., J. N. Couceiro, S. Kelm, L. G. Baum, S. Krauss, M. R. Castrucci, I. Donatelli, H. Kida, J. C. Paulson, R. G. Webster and Y. Kawaoka (1998). "Molecular basis for the generation in pigs of influenza A viruses with pandemic potential." Journal of Virology 72(9): 7367-7373.

Ito, T. and Y. Kawaoka (2000). "Host-range barrier of influenza A viruses." Vet Microbiol 74(1-2): 71-75.

Joseph, U., M. Linster, Y. Suzuki, S. Krauss, R. A. Halpin, D. Vijaykrishna, T. P. Fabrizio, T. M. Bestebroer, S. Maurer-Stroh, R. J. Webby, D. E. Wentworth, R. A. Fouchier and J. Bahl (2015). "Adaptation of pandemic H2N2 influenza A viruses in humans." Journal of Virology 89(4): 2442-2447.

Jourdain, E., G. Gunnarsson, J. Wahlgren, N. Latorre-Margalef, C. Brojer, S. Sahlin, L. Svensson, J. Waldenstrom, A. Lundkvist and B. Olsen (2010). "Influenza virus in a natural host, the mallard: experimental infection data." PLoS One 5(1): e8935.

Kawaoka, Y., S. Krauss and R. G. Webster (1989). "Avian-to-human transmission of the PB1 gene of influenza A viruses in the 1957 and 1968 pandemics." Journal of Virology 63(11): 4603-4608.

Kida, H., L. E. Brown and R. G. Webster (1982). "Biological activity of monoclonal antibodies to operationally defined antigenic regions on the hemagglutinin molecule of A/Seal/Massachusetts/1 /80 (H7N7) influenza virus." Virology 122(1): 38-47.

Kimura, H., C. Abiko, G. Peng, Y. Muraki, K. Sugawara, S. Hongo, F. Kitame, K. Mizuta, Y. Numazaki, H. Suzuki and K. Nakamura (1997). "Interspecies transmission of influenza C virus between humans and pigs." Virus Res 48(1): 71-79.

Kumlin, U., S. Olofsson, K. Dimock and N. Arnberg (2008). "Sialic acid tissue distribution and influenza virus tropism." Influenza and Other Respiratory Viruses 2(5): 147-154.

Latorre-Margalef, N., G. Gunnarsson, V. J. Munster, R. A. Fouchier, A. D. Osterhaus, J. Elmberg, B. Olsen, A. Wallensten, P. D. Haemig, T. Fransson, L. Brudin and J. Waldenstrom (2009). "Effects of influenza A virus infection on migrating mallard ducks." Proceedings. Biological Sciences 276(1659): 1029-1036.

Lefkowitz, E. J., D. M. Dempsey, R. C. Hendrickson, R. J. Orton, S. G. Siddell and D. B. Smith (2018). "Virus taxonomy: the database of the International Committee on Taxonomy of Viruses (ICTV)." Nucleic Acids Research 46(D1): D708-d717.

Lin, D., J. Lan and Z. Zhang (2007). "Structure and function of the NS1 protein of influenza A virus." Acta Biochimica et Biophysica Sinica 39(3): 155-162.

Long, J. S., Mistry, B., Haslam, S. M., Barclay, W. S. "Host and viral determinants of influenza A virus species specificity." Nat Rev Microbiol 17(2): 67-81

Manuguerra, J. C. and C. Hannoun (1992). "Natural infection of dogs by influenza C virus." Research in Virology 143(3): 199-204.

Matrosovich, M. N., T. Y. Matrosovich, T. Gray, N. A. Roberts and H. D. Klenk (2004). "Human and avian influenza viruses target different cell types in cultures of human airway epithelium." Proc Natl Acad Sci U S A 101(13): 4620-4624.

Mjaaland, S., E. Rimstad, K. Falk and B. H. Dannevig (1997). "Genomic characterization of the virus causing infectious salmon anemia in Atlantic salmon (Salmo salar L.): an orthomyxo-like virus in a teleost." Journal of Virology 71(10): 7681-7686.

Moriuchi, H., N. Katsushima, H. Nishimura, K. Nakamura and Y. Numazaki (1991). "Community-acquired influenza C virus infection in children." The Journal of Pediatrics 118(2): 235-238.

Muramoto, Y., T. Noda, E. Kawakami, R. Akkina and Y. Kawaoka (2013). "Identification of Novel Influenza A Virus Proteins Translated from PA mRNA." Journal of Virology 87(5): 2455.

Naeem, K. and M. Hussain (1995). "An outbreak of avian influenza in poultry in Pakistan." Vet Rec 137(17): 439.

Nao, N., J. Yamagishi, H. Miyamoto, M. Igarashi, R. Manzoor, A. Ohnuma, Y. Tsuda, W. Furuyama, A. Shigeno, M. Kajihara, N. Kishida, R. Yoshida and A. Takada (2017). "Genetic Predisposition To Acquire a Polybasic Cleavage Site for Highly Pathogenic Avian Influenza Virus Hemagglutinin." MBio 8(1).

Nielsen, O., A. Clavijo and J. A. Boughen (2001). "Serologic evidence of influenza A infection in marine mammals of Arctic Canada." Journal of Wildlife Diseases 37(4): 820-825.

Olsen, B., V. J. Munster, A. Wallensten, J. Waldenstrom, A. Osterhaus and R. A. M. Fouchier (2006). "Global patterns of influenza A virus in wild birds." Science 312(5772): 384-388.

Orlich, M., H. Gottwald and R. Rott (1994). "Nonhomologous recombination between the hemagglutinin gene and the nucleoprotein gene of an influenza virus." Virology 204(1): 462-465.

Osterhaus, A. D., G. F. Rimmelzwaan, B. E. Martina, T. M. Bestebroer and R. A. Fouchier (2000). "Influenza B virus in seals." Science 288(5468): 1051-1053.

Palese, P. and R. W. Compans (1976). "Inhibition of influenza virus replication in tissue culture by 2-deoxy-2,3-dehydro-N-trifluoroacetylneuraminic acid (FANA): mechanism of action." The Journal of General Virology 33(1): 159-163.

Pasick, J., Y. Berhane, T. Joseph, V. Bowes, T. Hisanaga, K. Handel and S. Alexandersen (2015). "Reassortant Highly Pathogenic Influenza A H5N2 Virus Containing Gene Segments Related to Eurasian H5N8 in British Columbia, Canada, 2014." Scientific Reports 5: 9484.

Penny, D., M. D. Hendy and M. A. Steel (1992). "Progress with methods for constructing evolutionary trees." Trends in Ecology & Evolution 7(3): 73-79.

Pflug, A., D. Guilligay, S. Reich and S. Cusack (2014). "Structure of influenza A polymerase bound to the viral RNA promoter." Nature 516(7531): 355-360.

Pinto, L. H., L. J. Holsinger and R. A. Lamb (1992). "Influenza virus M2 protein has ion channel activity." Cell 69(3): 517-528.

Presti, R. M., G. Zhao, W. L. Beatty, K. A. Mihindukulasuriya, A. P. da Rosa, V. L. Popov, R. B. Tesh, H. W. Virgin and D. Wang (2009). "Quaranfil, Johnston Atoll, and

Lake Chad viruses are novel members of the family Orthomyxoviridae." Journal of Virology 83(22): 11599-11606.

Presti, R. M., G. Y. Zhao, W. L. Beatty, K. A. Mihindukulasuriya, A. da Rosa, V. L. Popov, R. B. Tesh, H. W. Virgin and D. Wang (2009). "Quaranfil, Johnston Atoll, and Lake Chad Viruses Are Novel Members of the Family Orthomyxoviridae." Journal of Virology 83(22): 11599-11606.

Reid, A. H. and J. K. Taubenberger (2003). "The origin of the 1918 pandemic influenza virus: a continuing enigma." The Journal of General Virology 84(Pt 9): 2285-2292.

Schafer, J. R., Y. Kawaoka, W. J. Bean, J. Suss, D. Senne and R. G. Webster (1993). "Origin of the pandemic 1957 H2 influenza A virus and the persistence of its possible progenitors in the avian reservoir." Virology 194(2): 781-788.

Schultz, U., W. M. Fitch, S. Ludwig, J. Mandler and C. Scholtissek (1991). "Evolution of pig influenza viruses." Virology 183(1): 61-73.

Shao, W., X. Li, M. U. Goraya, S. Wang and J.-L. Chen (2017). "Evolution of Influenza A Virus by Mutation and Re-Assortment." International Journal of Molecular Sciences 18(8): 1650.

Shope, R. E. (1931). "Swine Influenza. I. Experimental Transmission and Pathology." The Journal of Experimental Medicine 54(3): 349-359.

Shors, T. (2013, 2018/8/28/). "H5N1 virus (bird flu) controversy." from https://www.accessscience.com:443/content/h5n1-virus-bird-flu-controversy/YB130019.

Shortridge, K. F. (1999). "Poultry and the influenza H5N1 outbreak in Hong Kong, 1997: abridged chronology and virus isolation." Vaccine 17 Suppl 1: S26-29.

Simonsen, L. (1999). "The global impact of influenza on morbidity and mortality." Vaccine 17 Suppl 1: S3-10.

Simonsen, L., M. J. Clarke, L. B. Schonberger, N. H. Arden, N. J. Cox and K. Fukuda (1998). "Pandemic versus epidemic influenza mortality: A pattern of changing age distribution." Journal of Infectious Diseases 178(1): 53-60.

Spalding, M. (2009). "Diseases of Poultry, 12th Edition." Journal of Wildlife Diseases 45(1): 251-256.

Stallknecht, D. E., M. T. Kearney, S. M. Shane and P. J. Zwank (1990). "Effects of pH, temperature, and salinity on persistence of avian influenza viruses in water" Avian Diseases 34(2): 412-418.

Stallknecht, D. E., S. M. Shane, M. T. Kearney and P. J. Zwank (1990). "Persistence of avian influenza viruses in water." Avian Diseases 34(2): 406-411.

Stegmann, T. (2000). "Membrane fusion mechanisms: the influenza hemagglutinin paradigm and its implications for intracellular fusion." Traffic 1(8): 598-604.

Suzuki, Y. and M. Nei (2002). "Origin and evolution of influenza virus hemagglutinin genes." Mol Biol Evol 19(4): 501-509.

Suzuki, Y. and M. Nei (2002). "Origin and Evolution of Influenza Virus Hemagglutinin Genes." Molecular Biology and Evolution 19(4): 501-509.

Taubenberger, J. K. and D. M. Morens (2008). "The Pathology of Influenza Virus Infections." Annual review of pathology 3: 499-522.

Taylor, R. M. (1949). "Studies on Survival of Influenza Virus Between Epidemics and Antigenic Variants of the Virus." American Journal of Public Health and the Nations Health 39(2): 171-178.

Thi H. O. Nguyen, Marios Koutsakos, Emma J. Grant, Peter C. Doherty, Katherine Kedzierska. "Towards Future T Cell-Mediated Influenza Vaccines".Infectious Diseases and Translational Medicine 2(1): 20-29

Tong, S. X., Y. Li, P. Rivailler, C. Conrardy, D. A. A. Castillo, L. M. Chen, S. Recuenco, J. A. Ellison, C. T. Davis, I. A. York, A. S. Turmelle, D. Moran, S. Rogers, M. Shi, Y. Tao, M. R. Weil, K. Tang, L. A. Rowe, S. Sammons, X. Y. Xu, M. Frace, K. A. Lindblade, N. J. Cox, L. J. Anderson, C. E. Rupprecht and R. O. Donis (2012). "A distinct lineage of influenza A virus from bats." Proceedings of the National Academy of Sciences of the United States of America 109(11): 4269-4274.

Tong, S. X., X. Y. Zhu, Y. Li, M. Shi, J. Zhang, M. Bourgeois, H. Yang, X. F. Chen, S. Recuenco, J. Gomez, L. M. Chen, A. Johnson, Y. Tao, C. Dreyfus, W. L. Yu, R. McBride,

P. J. Carney, A. T. Gilbert, J. Chang, Z. Guo, C. T. Davis, J. C. Paulson, J. Stevens, C. E. Rupprecht, E. C. Holmes, I. A. Wilson and R. O. Donis (2013). "New World Bats Harbor Diverse Influenza A Viruses." Plos Pathogens 9(10).

Udayan Joseph, M. L., Yuka Suzuki, Scott Krauss, Rebecca A. Halpin, Dhanasekaran Vijaykrishna, Thomas P. Fabrizio, Theo M. Bestebroer, Sebastian Maurer-Stroh, Richard J. Webby, David E. Wentworth, Ron A. M. Fouchier, Justin Bahl, Gavin J. D. Smith and members of the CEIRS H2N2 Working Group (2015). "Adaptation of Pandemic H2N2 Influenza A Viruses in Humans." Journal of Virology 89(4).

Varghese, J. N., W. G. Laver and P. M. Colman (1983). "Structure of the influenza virus glycoprotein antigen neuraminidase at 2.9 A resolution." Nature 303(5912): 35-40.

Vincent, A., L. Awada, I. Brown, H. Chen, F. Claes, G. Dauphin, R. Donis, M. Culhane, K. Hamilton, N. Lewis, E. Mumford, T. Nguyen, S. Parchariyanon, J. Pasick, G. Pavade, A. Pereda, M. Peiris, T. Saito, S. Swenson, K. Van Reeth, R. Webby, F. Wong and J. Ciacci-Zanella (2014). "Review of Influenza A Virus in Swine Worldwide: A Call for Increased Surveillance and Research." Zoonoses and Public Health 61(1): 4-17.

Wang, M. and M. Veit (2016). "Hemagglutinin-esterase-fusion (HEF) protein of influenza C virus." Protein & Cell 7(1): 28-45.

Weber, F. and O. Haller (2007). "Viral suppression of the interferon system." Biochimie 89(6-7): 836-842.

Webster, R. G., W. J. Bean, O. T. Gorman, T. M. Chambers and Y. Kawaoka (1992). "Evolution and ecology of influenza A viruses." Microbiological Reviews 56(1): 152-179.

Wendel, I., D. Rubbenstroth, J. Doedt, G. Kochs, J. Wilhelm, P. Staeheli, H. D. Klenk and M. Matrosovich (2015). "The avian-origin PB1 gene segment facilitated replication and transmissibility of the H3N2/1968 pandemic influenza virus." Journal of Virology 89(8): 4170-4179.

Westgeest, K. B., C. A. Russell, X. Lin, M. I. Spronken, T. M. Bestebroer, J. Bahl, R. van Beek, E. Skepner, R. A. Halpin, J. C. de Jong, G. F. Rimmelzwaan, A. D. Osterhaus, D. J. Smith, D. E. Wentworth, R. A. Fouchier and M. de Graaf (2014). "Genomewide analysis of reassortment and evolution of human influenza A(H3N2) viruses circulating between 1968 and 2011." Journal of Virology 88(5): 2844-2857.

Wille, M., Y. Huang, G. J. Robertson, P. Ryan, S. I. Wilhelm, D. Fifield, A. L. Bond, A. Granter, H. Munro, R. Buxton, I. L. Jones, M. G. Fitzsimmons, C. Burke, L. M. Tranquilla, M. Rector, L. Takahashi, A. L. Kouwenberg, A. Storey, C. Walsh, A. Hedd, W. A. Montevecchi, J. A. Runstadler, D. Ojkic, H. Whitney and A. S. Lang (2014). "Evaluation of seabirds in Newfoundland and Labrador, Canada, as hosts of influenza A viruses." J Wildl Dis 50(1): 98-103.

Williams, R. E., H. Hoogstraal, J. Casals, M. N. Kaiser and M. I. Moussa (1973). "Isolation of Wanowrie, Thogoto, and Dhori viruses from Hyalomma ticks infesting camels in Egypt." Journal of Medical Entomology 10(2): 143-146.

Wilson, I. A., J. J. Skehel and D. C. Wiley (1981). "Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 A resolution." Nature 289(5796): 366-373.

Wise, H. M., A. Foeglein, J. Sun, R. M. Dalton, S. Patel, W. Howard, E. C. Anderson, W. S. Barclay and P. Digard (2009). "A complicated message: Identification of a novel PB1-related protein translated from influenza A virus segment 2 mRNA." J Virol 83(16): 8021-8031.

Wise, H. M., E. C. Hutchinson, B. W. Jagger, A. D. Stuart, Z. H. Kang, N. Robb, L. M. Schwartzman, J. C. Kash, E. Fodor, A. E. Firth, J. R. Gog, J. K. Taubenberger and P. Digard (2012). "Identification of a novel splice variant form of the influenza A virus M2 ion channel with an antigenically distinct ectodomain." PLoS Pathog 8(11): e1002998.

Yamayoshi, S., M. Watanabe, H. Goto and Y. Kawaoka (2016). "Identification of a Novel Viral Protein Expressed from the PB2 Segment of Influenza A Virus." Journal of Virology 90(1): 444.

Chapter 2: Evaluation of influenza A virus quasispecies populations in different avian reservoir hosts

Abstract

Influenza A viruses (IAVs) evolve rapidly through mutations, due to an error-prone viral replicase enzyme, and through the process of reassortment during co-infections. Mutations generate new, closely related virus variants that can have different capacities of infection and replication. The pool of closely related variants that are produced during an infection is known as a quasispecies. Wild birds are the natural reservoir of IAVs and are known to carry the majority of IAV subtypes and overall genetic diversity. Most avian IAVs have low pathogenicity in their natural hosts and circulate without causing symptoms, but they are still of broad importance due to their involvement in the generation of highly pathogenic viruses in domestic poultry and viruses infecting humans. To better understand the genetic structure and evolutionary dynamics of IAVs in their natural hosts, we performed a comparative analysis of quasispecies populations shed by different wild birds. This was done by performing high-throughput sequencing of IAVs present in paired oropharyngeal/cloacal swab samples collected from three different groups of wild bird hosts: seabirds, gulls and ducks. My analysis shows that most of the virus population structure is conserved within the same bird host (i.e., among ducks, gulls, or murres). However, comparison of the virus population structure for two different hosts, such as ducks and gulls, shows more dissimilarities. Deep sequencing was also

applied during the serial passage of a virus isolated from a wild duck in Atlantic Puffin embryos to investigate the effect of change in host species on viral evolution. I found an overall increase of viral variants within the virus quasispecies in the fifth and last virus passages analyzed compared to the initial pre-passage virus swab sample. I also observed a rapid accumulation of non-synonymous mutations in the HA and NA segments that could alter their function, and both of these genes showed high dN/dS ratios, which indicates that they were under positive selection.

2.1 Introduction

Influenza A viruses (IAVs) are members of the viral family *Orthomyxoviridae*, a group of RNA viruses with segmented, negative-sense, single-stranded RNA genomes. Wild aquatic birds (waterfowl, gulls, and shorebirds) are considered the predominant natural reservoir hosts of IAVs but other bird groups, such as seabirds, also carry the virus (Webster, Bean et al. 1992). The virus evolves rapidly through the exchange of genes (reassortment) during co-infections and via the accumulation of mutations during the replication of the virus (Holland, Spindler et al. 1982, Yoon, Webby et al. 2014). The evolution of the virus is affected by host immune pressure, which can select for specific mutants within the virus quasispecies that escape host immune responses. A viral quasispecies is formed by a “cloud” of related viruses generated during the replication of the virus due to the error-prone nature of the viral RNA polymerase (Domingo, Baranowski et al. 1998, Andino and Domingo 2015). During infection of a new host, new immune pressures can then restructure the dynamics of the virus quasispecies. This can lead to dramatic changes in the virus

population by causing extinction of the virus or mediating its maintenance if adequate virus mutants are present and able to replicate within the new host (Kuiken, Holmes et al. 2006). Evolutionary studies in avian hosts have mostly been conducted with viruses isolated from ducks and chickens and these analyses were mostly restricted to the HA and NA gene segments (Li, 2010, Dlugolenski, 2011). These studies were also mostly focused on the main genomic composition of the virus genome (i.e. the genome's consensus sequence), not at the quasispecies level considering all the virus mutants.

To date, limited information exists on IAV genetic diversity and evolution in the natural host reservoirs at the quasispecies level. I have performed a comparative in-depth analysis of the quasispecies population of IAVs shed by reservoir bird hosts. The structure of the IAV populations was investigated in paired oropharyngeal/cloacal swab samples from three avian host groups: seabirds, represented by common murre (*Uria aalge*), a seabird species found in the Northern Hemisphere; ducks, represented by American black duck (*Anas rubripes*), a duck species commonly found in parts of North America, and a feral domestic duck; and gulls, represented by American herring gull (*Larus smithsonianus*), which is widely distributed in North America. The evolution and genetic structure of a duck virus was also investigated experimentally by 10 serial passages of the virus in Atlantic puffin (*Fratercula arctica*) egg embryos.

2.2 Materials and Methods

2.2.1 Viruses and RNA isolation

The bird swab samples used in this research were collected according to guidelines of the Canadian Council on Animal Care (approved protocols 10-01-AL and 11-01-AL) from the Memorial University Institutional Animal Care Committee). Lab work was performed under Memorial University Biosafety Certificate S-103. Puffin eggs were collected from the Witless Bay Ecological Reserve under permits from the Newfoundland and Labrador Department of Environment and Conservation.

The viruses used in this study originated from three different host groups identified in Newfoundland and Labrador and are described in the Supplementary table 2.1. Virus genotypes were previously determined based on phylogenetic analyses of the segments (Huang, Wille et al. 2013, Huang, Robertson et al. 2014, Huang, Wille et al. 2014). Viral quasispecies were characterized from paired oropharyngeal and cloacal swabs. A duck virus (MW668) contained in a swab sample was serially passaged 10 times via the allantoic route in embryonated bird eggs from Atlantic puffin. Embryonated puffin eggs were chosen for this analysis due to their availability in the province over the summer period. The presence and titer of the virus was determined by hemagglutination assay. A 2-fold serial dilution of virus sample was performed in 1X phosphate-buffered saline in 96-well plates. An equal volume of 0.25 % chicken red blood cells was added to each well containing the diluted virus sample, and the HA titer corresponding to the last dilution factor where hemagglutination occurred was used to determine the virus titer. RNA was extracted directly from 200 µL of the viral transport medium of the swab sample and from 200

µL of allantoic fluid containing viruses from the 10 serial passages, using the Trizol LS reagent (Thermo Fisher Scientific). The swab sample had been initially tested for the presence of IAV by QRT-PCR targeting the M gene (Huang, Wille et al. 2014).

2.2.2 Deep sequencing workflow of IAV genomes

Prior to sequencing, the viral RNA genome was converted into cDNA and amplified in a one-step reaction by RT-PCR (SuperScript III One-Step RT-PCR System with Platinum *Taq* High Fidelity DNA Polymerase, Thermo Fisher Scientific) with a pair of universal primers that bind specifically to the non-coding regions of the virus, MBTuni-12: 5'-ACGCGTGATCAGCAAAAGCAGG-3' and MBTuni-13: 5'-ACGCGTGATCAGTAGAAACAAGG-3' (Hoffmann, Stech et al. 2001). This allows the amplification of the complete viral genome. The construction of the sequencing libraries was first performed with the Ion Xpress™ Plus gDNA Fragment Library Preparation according to the manufacturer's instructions (Thermo Fisher Scientific). For cost reasons, an in-house method was later developed and used for the virus genomic library preparation (Verhoeven, Canuti et al. 2018). For both methods, the quality (presence of primer-dimers, size of the library) and quantity of each library was assessed and determined on a High Sensitivity DNA chip with the 2100 Bioanalyzer instrument (Agilent). Libraries were pooled in equimolar concentrations at 100 pM. The pooled libraries were amplified onto Ion Sphere Particles (ISPs) by emulsion PCR (emPCR). Coated ISPs were enriched with the Ion OneTouch 2 instrument to eliminate uncoated ISPs and loaded on a 314 chip and sequenced using the Ion Torrent Personal Genome Machine (Thermo Fisher Scientific). The virus

genomes from GR256 and MW668 passage 5 were prepared with the in-house method, while the rest of the viruses were prepared with the commercial kit (Supplementary table 2.1).

2.2.3 IAV genome assembly and variant calling

Adapters and barcodes were removed from the sequencing reads prior to analysis with the Torrent Suite software (Thermo Fisher Scientific). The reads from each genomic library were quality-trimmed at the 5' and 3' extremities and *de novo* assembled and mapped to an IAV reference genome with Geneious version 8 (Biomatters, New Zealand). Read quality was checked based on the PHRED score at each base of the sequence, determined by the sequencer, which gives information on the quality of each nucleotide sequenced. All IAV genomes showed an average PHRED quality score of 30, which means that the accuracy of the sequencing at each base is 99.9 %. Variant calling for each IAV genome was performed with the Geneious software) with a minimum variant frequency of 25% within the total number of reads and a maximum p-value cut-off of 10^{-6} to decrease the chance of selecting variants that could result from sequencing errors. In this analysis, a lower p-value means that the variation observed is less likely due to a sequencing error. Homopolymer quality reduction was applied to account for possible issues from the Ion Torrent sequencing because this instrument is known to be prone to errors in homopolymeric regions. The sequencing coverage for each virus is provided in Supplementary figure 2.1. Viruses with similar average coverages were chosen to avoid potential bias due to coverage differences in the identification of virus variants.

2.2.4 Statistical analyses

The dN/dS ratios were calculated according to the following formulas: $d_N = -3Ln(1 - ((\frac{4N}{3})/4)$ and $d_S = -3Ln(1 - ((\frac{4S}{3})/4)$ where N and S represent the proportion of non-synonymous and synonymous substitutions, respectively. The Shannon diversity index (H') was calculated according to the formula: $H' = -\sum_{i=0}^{n-1} P_i \times \log_2 P_i$, where P_i represents the frequency of each variant and n the 4 bases A, T, C, and G. GraphPad Prism 7 was used to generate data plots. A Kruskal-Wallis test was performed to compare the distribution of variants among the same bird host.

2.3 Results

2.3.1 Analysis of IAV quasispecies populations shed by different avian reservoir hosts

2.3.1.1 Gull IAV quasispecies structure

In order to assess whether the quasispecies of IAVs were conserved within the same host, I first compared the genomic structure of three H13N6 gull viruses, two of which have an identical genotype; their genotypes were previously determined based on phylogenetic analyses of the segments (Huang, Wille et al. 2014) (Table 2.1).

Table 2.1: Genomic comparison of three gull H13N6 viruses.

	American herring gull viruses (H13N6)		
	GR654	GR658	GR848
	Genotype		
	I	I	K
PB2	EA GULL*	EA[+] GULL	EA GULL
PB1	NA[$] GULL	NA GULL	NA AVIAN
PA	NA GULL	NA GULL	NA GULL
HA	NA GULL	NA GULL	NA GULL
NA	NA AVIAN	NA AVIAN	NA AVIAN
NP	NA GULL	NA GULL	EA GULL
M	EA GULL	EA GULL	EA GULL
NS	NA GULL	NA GULL	NA GULL

***Similar hosts or geographic origins are indicated by the same colours**

[+]EA: Eurasian origin; [$]NA: North American origin

Among the viruses, GR654 and GR658 are genotype I while GR848 is genotype K. The sequence identities among the three viruses for each segment are given in Supplementary table 2.2. Even between the non-identical viruses, some of the segments are the same in terms of host and geographic lineages. For instance, the PB2 segment falls within a Eurasian gull clade for all three viruses, whereas the PB1 segment of GR848 belongs to a North American avian lineage.

I first investigated the overall distribution of variant sites present within each segment of the three gull viruses (Figure 2.1). Each variant corresponds to a single mutation that causes a genetic change in the virus genome.

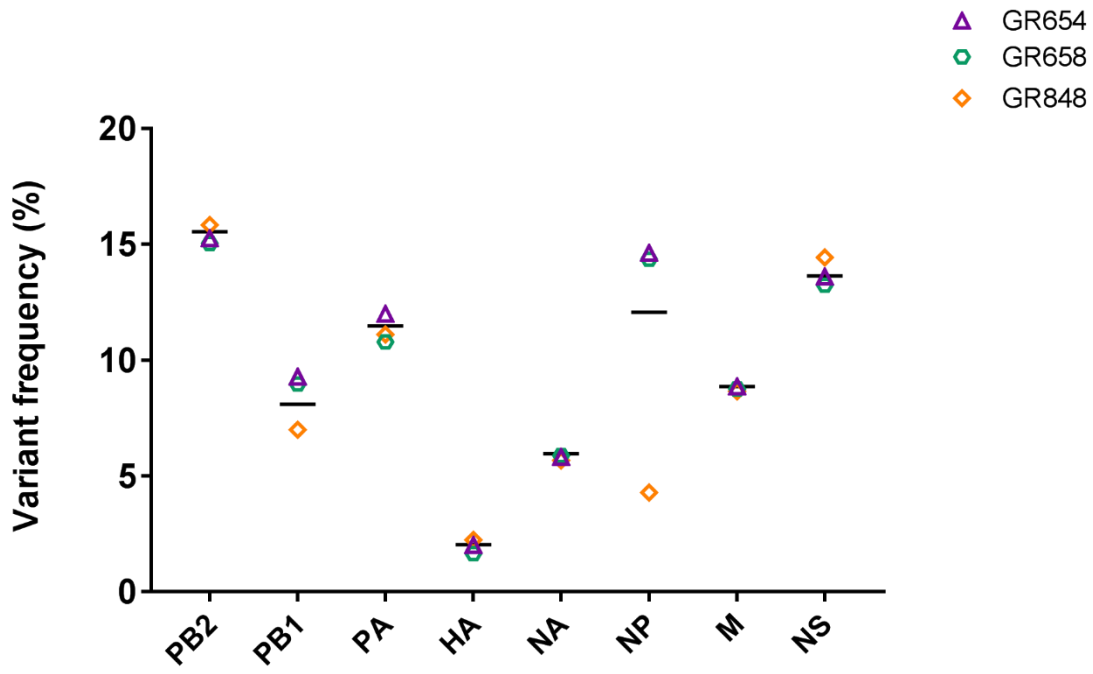


Figure 2.1: Overall variant frequencies for the quasispecies of gull viruses. Each symbol represents the proportion of variants found in gull genomic segments. Horizontal bars represent the mean of the proportion of variants present in the three gull viruses.

Most of the segments had very similar variant proportions, except for PB1 and NP. This difference in proportions seems to be correlated with the phylogenetic relationships of the segments such that similar variant frequencies are detected in more closely related segments. For instance, the PB1 and NP segments of virus GR848 are from different phylogenetic lineages than the other two viruses (Table 2.1).

Analysis of the proportions of shared variants for all segments of the different gull viruses revealed that higher proportions of identical variants were detected in more closely related segments. Indeed, the overall shared variant proportion of the two viruses with identical genotypes, GR654 and GR658, was proportionally high. In contrast, the PB1 and NP segments from virus GR848 have much lower shared variants compare to the other two viruses. This could be explained by the fact that the PB1 and NP genes of this virus have different host and geographic origins, respectively, compared to GR654 and GR658 (Figure 2.2, Table 2.1).

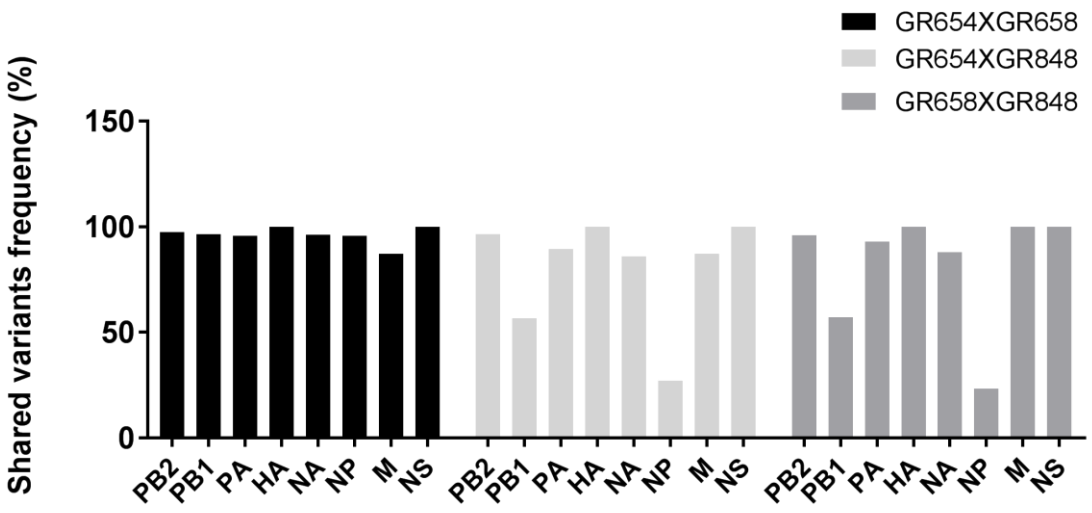


Figure 2.2: Shared variant frequencies for pairwise comparisons of the three gull viruses. The proportion of identical virus variants present in the segments of the gull viruses were compared pairwise, as indicated by the legend.

The types of variants within each gull virus quasispecies were investigated

Figure 2.3).

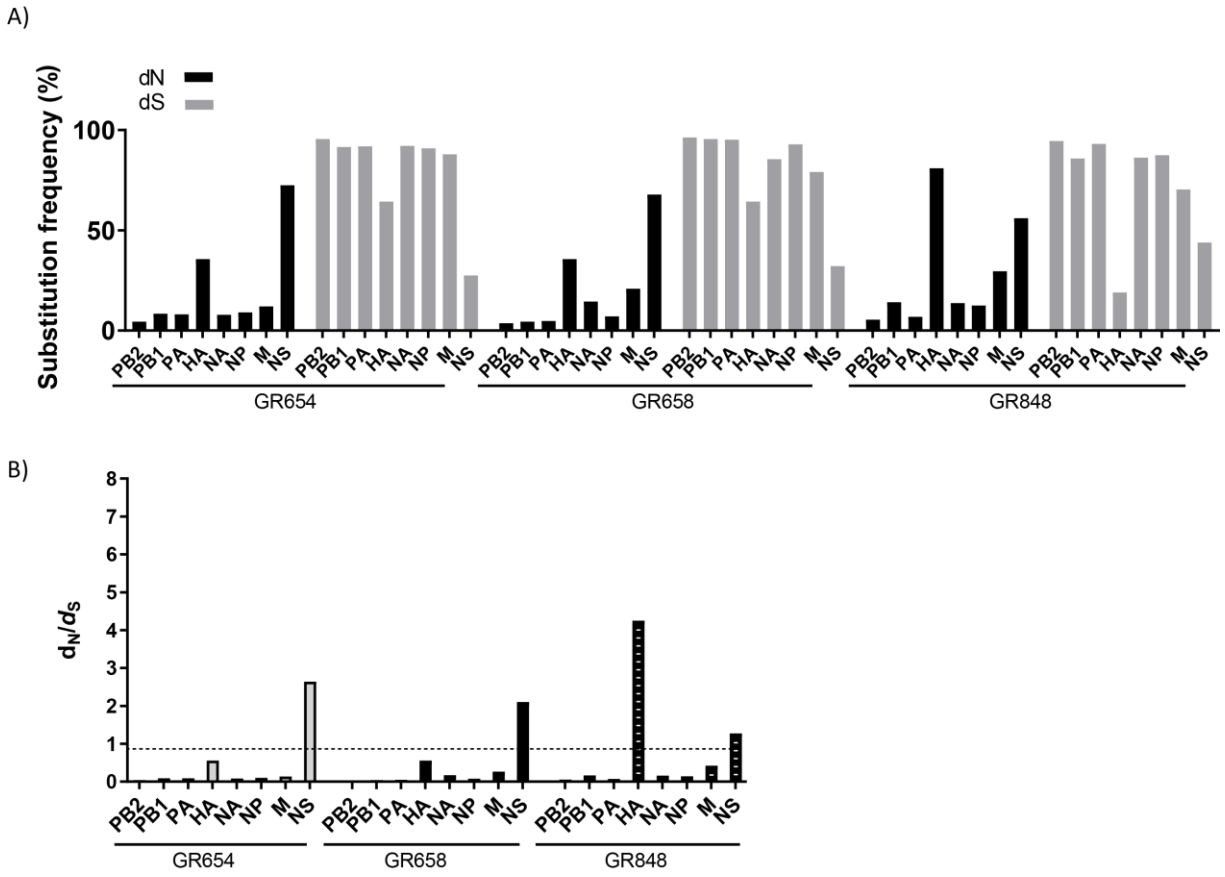


Figure 2.3: Synonymous and non-synonymous substitution frequencies for all segments of the gull viruses. A) The proportions of non-synonymous substitutions (dN) and synonymous substitutions (dS) were determined in each virus quasispecies for all segments. B) The dN/dS ratio was determined for all segments.

I identified the synonymous (silent) and non-synonymous (resulting in change to the encoded amino acid sequence) substitutions found in each segment and quantified their proportions). The proportion of non-synonymous substitutions (dN) for each of

the segments of each virus shows that the highest numbers are detected for the HA and NS segments (Figure 2.3A, Supplementary table 2.2). The NS segment of the three gull viruses have a dN/dS ratio >1 which means that the sequence is under positive selection. Positive selection favours the selection of variants, within the virus quasipecies, that confer a replicative advantage for the virus. A ratio <1 implies that the sequence is under negative selection. Negative selection causes the elimination of variants within the virus quasipecies that are deleterious for the virus replication. A ratio equal to 1 indicates that the gene is under neutral pressure (Jukes and Cantor 1969) (Figure 2.3B).

Lastly, I compared the variants' diversities, in terms of frequencies within the virus quasispecies, of the two most similar gull viruses, GR654 and GR658 (Figure 2.4).

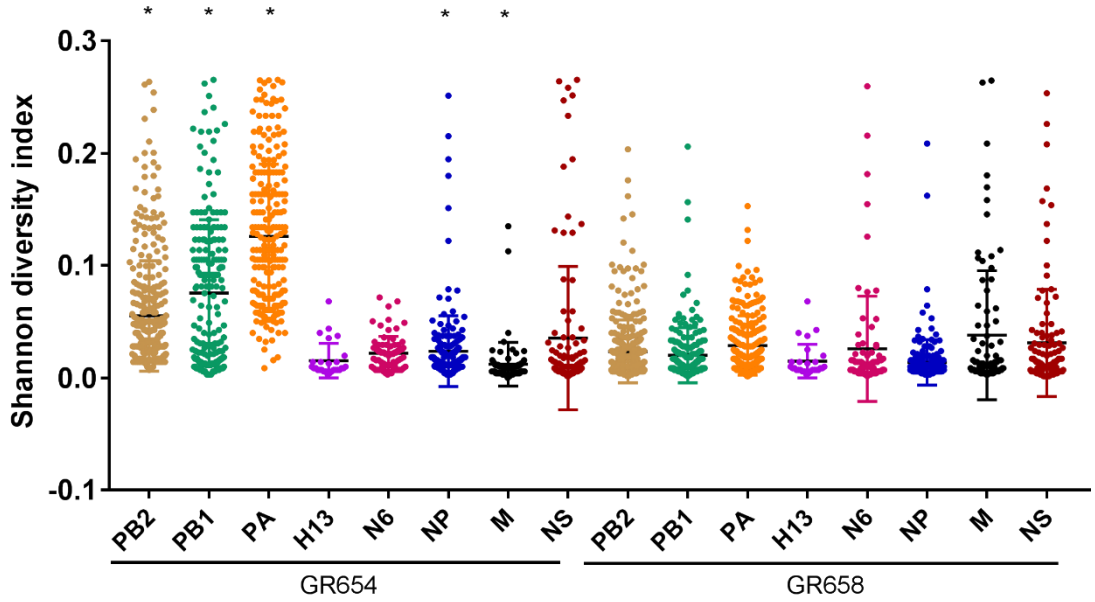


Figure 2.4: Comparison of the quasispecies structure of two gull viruses with the same genotype. Shannon diversity index was applied to all variants present for each segment of the two gull viruses, GR654 and GR658. Each dot represents how frequent the variant is within the segment quasispecies. Horizontal bars represent the mean of the variant diversity within each segment. The distributions of the variants within each gull quasispecies were compared with a Kruskal-Wallis test ($P<0.0001$). Dissimilar distribution in terms of variant frequencies between compared segments are indicated by an asterisk above the plots (P value ≥ 0.0001). No asterisk means the variant distributions are very similar for the gene compared.

This analysis was done using the Shannon diversity index, which is based on variant frequencies.This gives a measure of how diverse the virus quasipecies is in terms of variant frequencies. A Kruskal-Wallis test was used to test whether each virus' population, per segment, is distributed in a similar manner. This showed that the

degree of variant diversity in terms of frequency is different for some segments (Figure 2.4). Indeed, only the H13, N6, and NS segments had similar degrees of diversity. This is in accordance with my earlier analysis indicated that they shared a high number of identical variants (Figure 2.2). Therefore, two viruses with identical genotypes can harbour the same variants but with differences in their frequencies within the individual quasispecies populations.

2.3.1.2 Murre IAV quasispecies structure

Similar to what was done for the gull viruses, I then compared a set of three murre viruses in which two, AB341 and AB319, have identical genotypes (Table 2.2).

Table 2.2: Genomic comparison of three murre H1N2 viruses

	Common murre viruses (H1N2)		
	AB327	AB341	AB319
	Genotype		
	D	A	A
PB2	NA[$] AVIAN	EA[+] AVIAN	EA AVIAN
PB1	NA AVIAN*	NA AVIAN	NA AVIAN
PA	NA AVIAN	NA AVIAN	NA AVIAN
HA	NA AVIAN	NA AVIAN	NA AVIAN
NA	NA AVIAN	NA AVIAN	NA AVIAN
NP	NA AVIAN	NA AVIAN	NA AVIAN
M	NA AVIAN	EA GULL	EA GULL
NS	NA AVIAN	NA AVIAN	NA AVIAN

***Similar hosts or geographic origins are indicated by the same colours**

[+]EA: Eurasian origin; [$]NA: North American origin

Analysis of the quasispecies structure of these viruses showed that they all shared a similar proportion of variants for most segments, except for the PB2 and M segments of AB327. Indeed, both of these genes have different phylogenetic affiliations compared to the M segment of AB341 AND ab319 (Figure 2.5, Table 2.2).

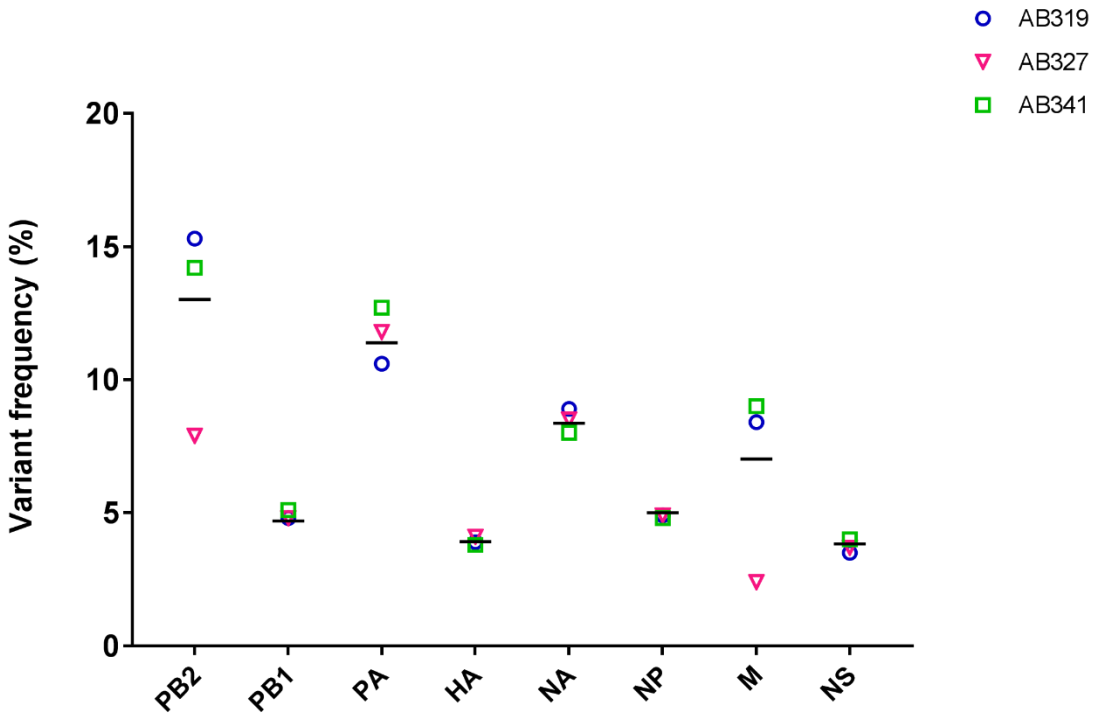


Figure 2.5: Overall variant frequencies of the quasispecies of murre viruses. Each symbol represents the proportion of variants found in each genomic segment. Horizontal bars represent the mean of the proportions of variants present in the three viruses.

Therefore, as seen with the gull viruses, the similarities in variant proportions in murre viruses are correlated with the phylogenetic relatedness of the segments.

Investigation of the proportion of shared variants in the murre viruses showed that a high number of identical variants were also present in genes that are closely related phylogenetically In fact, the PB2 and M segments for AB327, which are phylogenetically different, again showed a different pattern in this regard, and had the least number of shared variants with the corresponding segments from the other viruses (Figure 2.6).

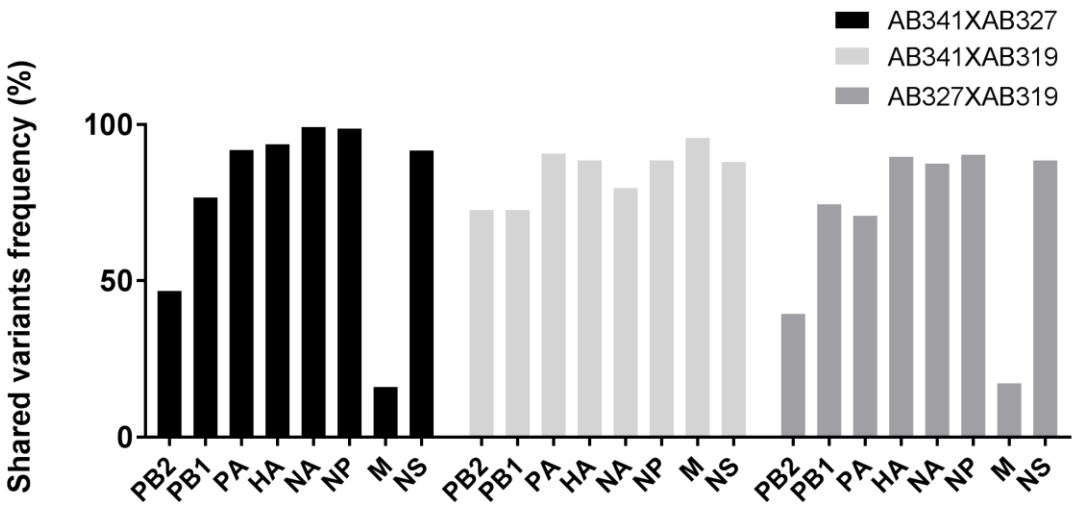


Figure 2.6: Shared variant frequencies for pairwise comparisons of the three murre viruses. The proportion of identical virus variants present in the segments of the murre viruses were compared pairwise, as indicated by the legend.

A higher proportion of non-synonymous substitutions (dN) was observed in the HA, NA, and NS segments of all murre viruses, and in the PB1 segment of only the AB341 virus (Figure 2.7).

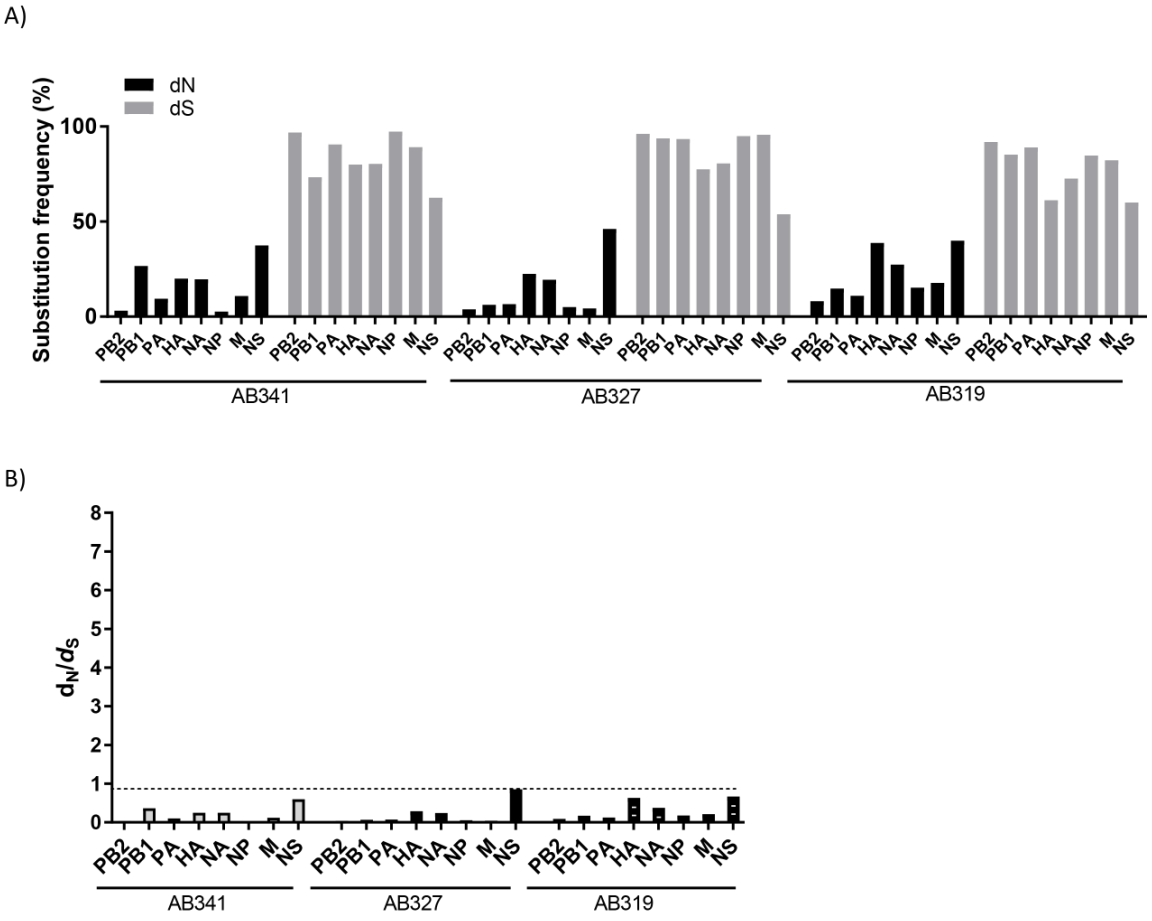


Figure 2.7: Synonymous and non-synonymous substitution frequencies for all segments of the murre viruses. A) The proportions of non-synonymous substitutions (dN) and synonymous substitutions (dS) were determined in each virus quasispecies for all segments. B) The dN/dS ratio was determined for all segments.

Overall, synonymous substitutions (dS) were more common for all segments of the three viruses (Figure 2.7A, Supplementary table 2.3). Contrary to the gull viruses, the dN/dS ratios of the NS segment of all murre viruses are <1, which means that the segment is under negative selection (Figure 2.7B).

The distribution of variants, in terms of frequencies within the quasispecies of the two murre viruses with an identical genotype, AB341 and AB319, were similar for some segments but not for all. Indeed, only the variant distributions of the H1, M, and NS segments are conserved among the two murre viruses compared (Figure 2.8).

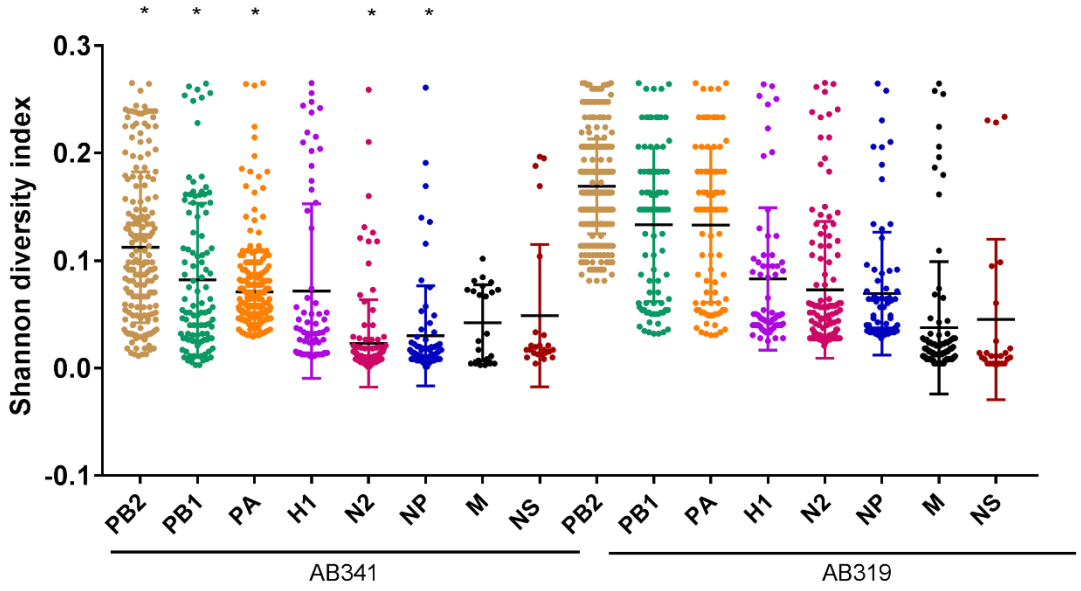


Figure 2.8: Comparison of the quasispecies structure of two murre viruses with an identical genotype. Shannon diversity index was applied to all variants present for each segment of the two murre viruses, AB341 and AB319. Each dot represents how frequent the variant is within the segment quasispecies. Horizontal bars represent the mean of the variant diversity within each segment. The distributions of the variants within each murre quasispecies were compared with a Kruskal-Wallis test ($P<0.0001$). Segments indicated with an asterisk do not have the same variant distribution as in the other virus. No asterisk means the genes compared have very similar variant distribution.

2.3.1.3 Duck IAV quasispecies structure

Finally, I investigated a set of duck viruses that have different genotypes but display similarities in terms of host and geographic origins (Table 2.3).

Table 2.3: Genomic comparison of three duck viruses.

	Duck viruses		
	MW668 H1N1	GR679 H3N2	GR256 H3N2
PB2	C-2.1*	C-2.5	C-2.2
PB1	F-3.1	F-4.1	F-2.1
PA	E-1.1	H-1.5	E-2.1
HA	1D-1.1	3C-1.1	3C-1.1
NA	1E-1.1	2D	2D-1.2
NP	H-4.1	H-5.1	H-5.1
M	E-1.1	E-1.6	E-1.4
NS	2B-1.1	2B-1.2	2B-1.2

***Identical phylogenetic clades are indicated by the same colours.**

Duck viruses with identical subtypes and genotypes were not available to include in the analysis. Two duck viruses have the same H3N2 subtype, GR679 and GR256, and were isolated from American black duck. The third virus, MW668, was isolated from a feral domestic duck and has the H1N1 subtype. All three viruses have gene segments that originate from the North American waterfowl/avian lineage (Huang, Wille et al. 2014). Genotypes with identical first letters but followed with different numbers indicates that the segments are within the same phylogenetic clade but within different subclades (Huang, Wille et al. 2014).

Analysis of the proportion of variants for the three duck viruses showed that similar proportions were found in segments with similar genotypes. For instance, the PA segment of the virus GR679 has a different genotype compared to the other two viruses and showed a different (higher) proportion of variants. Similarly, the MW668 virus that has a different HA subtype than the other two viruses also had a different (lower) proportion of variants (Figure 2.9).

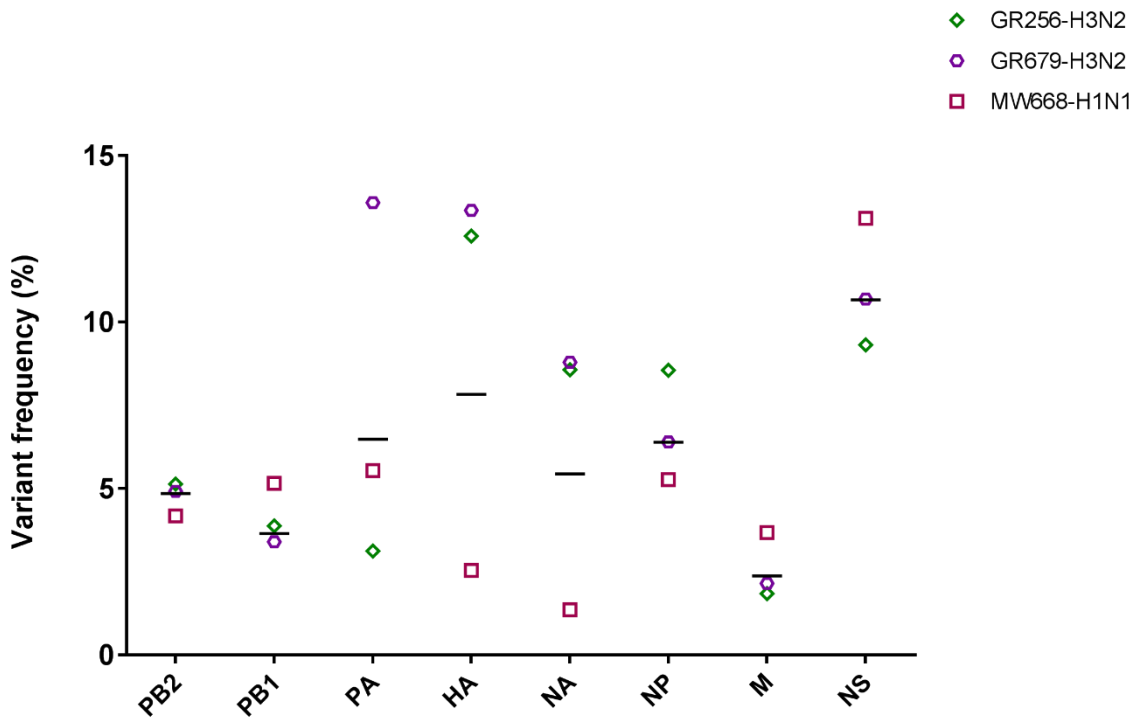


Figure 2.9: Overall variant frequencies of the quasispecies of duck viruses. Each symbol represents the proportion of variants found in each genomic segment. Horizontal bars represent the mean of the proportions of variants present in the three viruses.

Identical variants were also shared between virus quasispecies that had genes with similar phylogenies. In fact, the highest proportion of identical variants among all the three viruses was found for the NS segment. The GR679 and GR256 viruses, which have identical HA and NA subtypes and similar phylogenies, had a high proportion of shared variants in these segments (Figure 2.10).

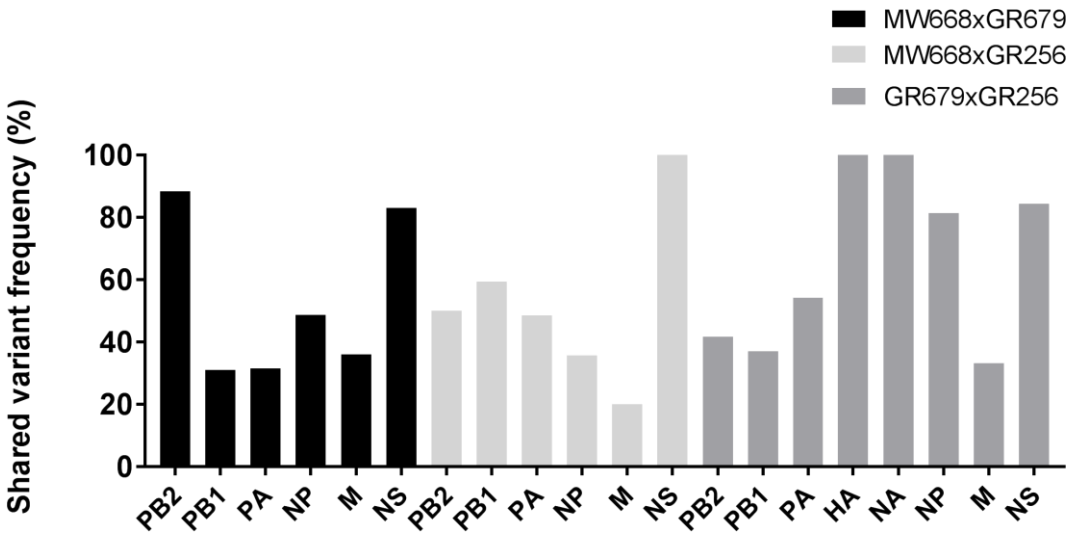


Figure 2.10: Shared variant frequencies for pairwise comparisons of the three duck viruses. The proportion of identical virus variants present in the segments of the duck viruses were compared pairwise, as indicated by the legend.

Analysis of the proportions of synonymous and non-synonymous substitutions in the duck viruses showed that the gene segments mostly harboured synonymous substitutions. Indeed, the NS segment showed the highest amount of non-synonymous substitutions in all three duck viruses (Figure 2.11A). Similar to the

murre viruses, none of the NS segment of all duck viruses have a dN/dS ratio>1 (Figure 2.11B).

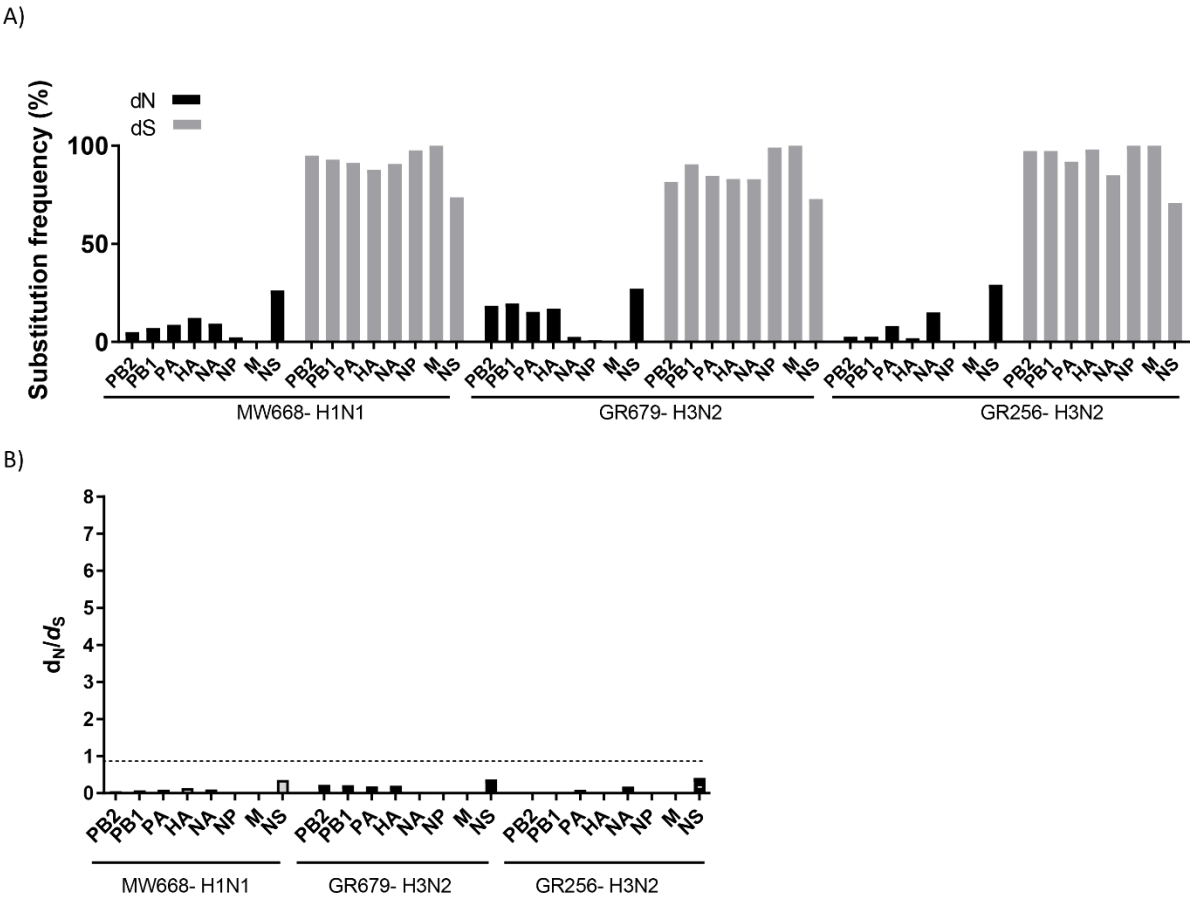


Figure 2.11: Synonymous and non-synonymous substitution frequencies for all segments of the duck viruses. A) The proportions of non-synonymous substitutions (dN) and synonymous substitutions (dS) were determined in each virus quasispecies for all segments. B) The dN/dS ratio was determined for all segments.

Variants present in the quasispecies of the two similar duck viruses, GR679 and GR256, shared some similar distributions in terms of frequency but only for the H3, M, and NS segments (Figure 2.12).

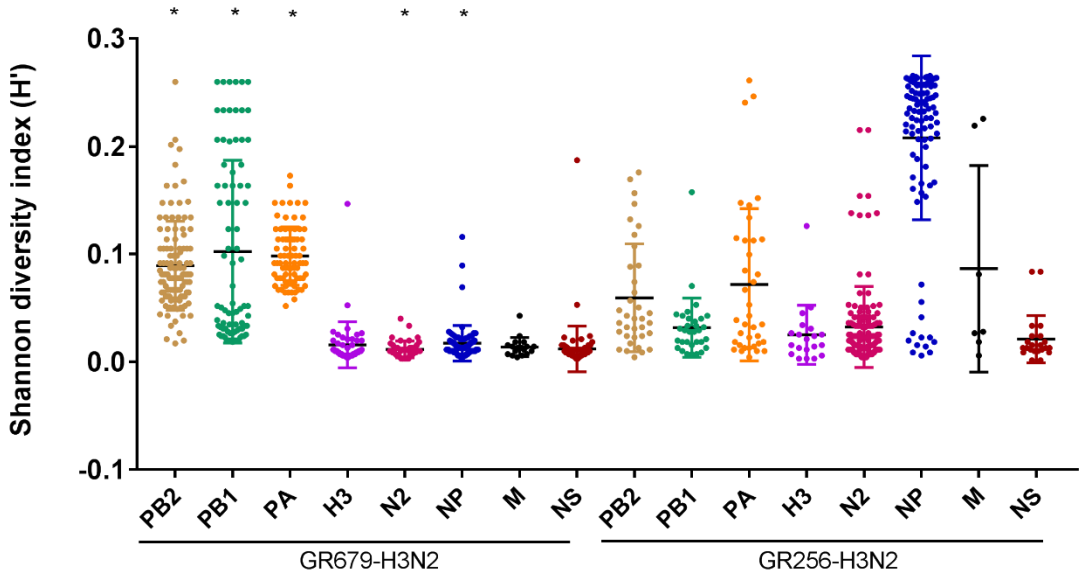


Figure 2.12: Comparison of the quasispecies structure of two duck viruses with similar genotypes. Shannon diversity index was applied to all variants present for each segment of the two duck viruses, GR679 and GR256. Each dot represents how frequent the variant is within the segment quasispecies. Horizontal bars represent the overall variant diversity within each segment. The distributions of the variants within each duck quasispecies were compared with a Kruskal-Wallis test ($P<0.0001$). An asterisk above the plot represents significantly dissimilar distribution compared to the corresponding segment of the other virus. No asterisk means the variant distributions are very similar for the segments.

Overall, my analyses show that gull, murre or duck viruses with similar genotypes display similar quasispecies structure in terms of shared variants and proportions, but the amount of diversity among the variants, in terms of

frequencies, differ since some segments contain shared variants in higher proportions than others.

2.3.2 Effect of replication in a different host on the virus quasispecies

Adaptation experiments were performed by infecting Atlantic puffin embryos with a single IAV isolate from a wild duck through 10 consecutive passages. Due to the limited availability of such eggs from wild birds, the serial passage was performed over three consecutive years due to the availability of wild bird eggs and allowance of the permits. The virus genome was deep-sequenced from the initial swab sample (P0), the first (P1), the fifth (P5) and last passage (P10). Mutation analyses were performed by comparing the viral genome deep sequenced and translated computational protein sequences. The viral titer, determined by hemagglutination assays for the 10 passages, decreased over the passages and became undetectable in the last two passages (Figure 2.13).

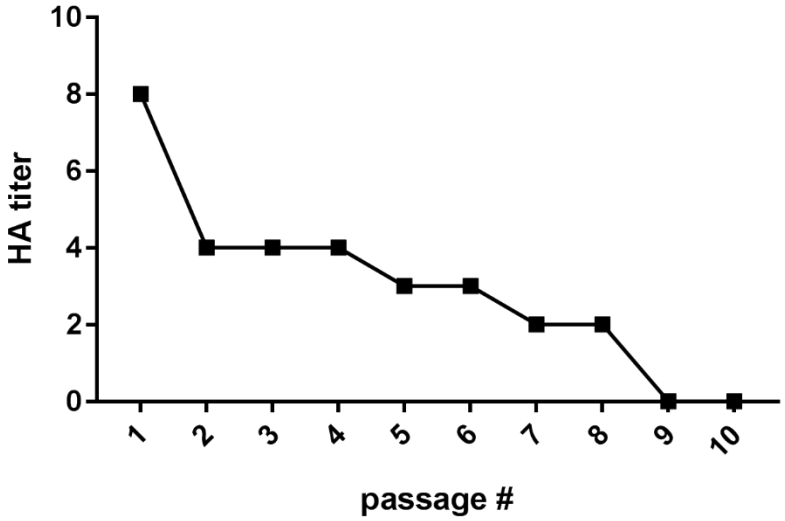


Figure 2.13: Viral titer of the duck virus over the serial passages in puffin embryos. The virus titer was determined by hemagglutination assay in the 10 samples collected after the serial virus passages.

2.3.2.1 Mutational analysis of the duck viral genome

Analysis of the number of variants present in the initial swab sample (P0), the fifth passage (P5), and the last passage in puffin embryos (P10) shows an overall increase of the number of variants in the last virus passage for all gene segments except for NS compared to the initial swab (P0) (Figure 2.14).

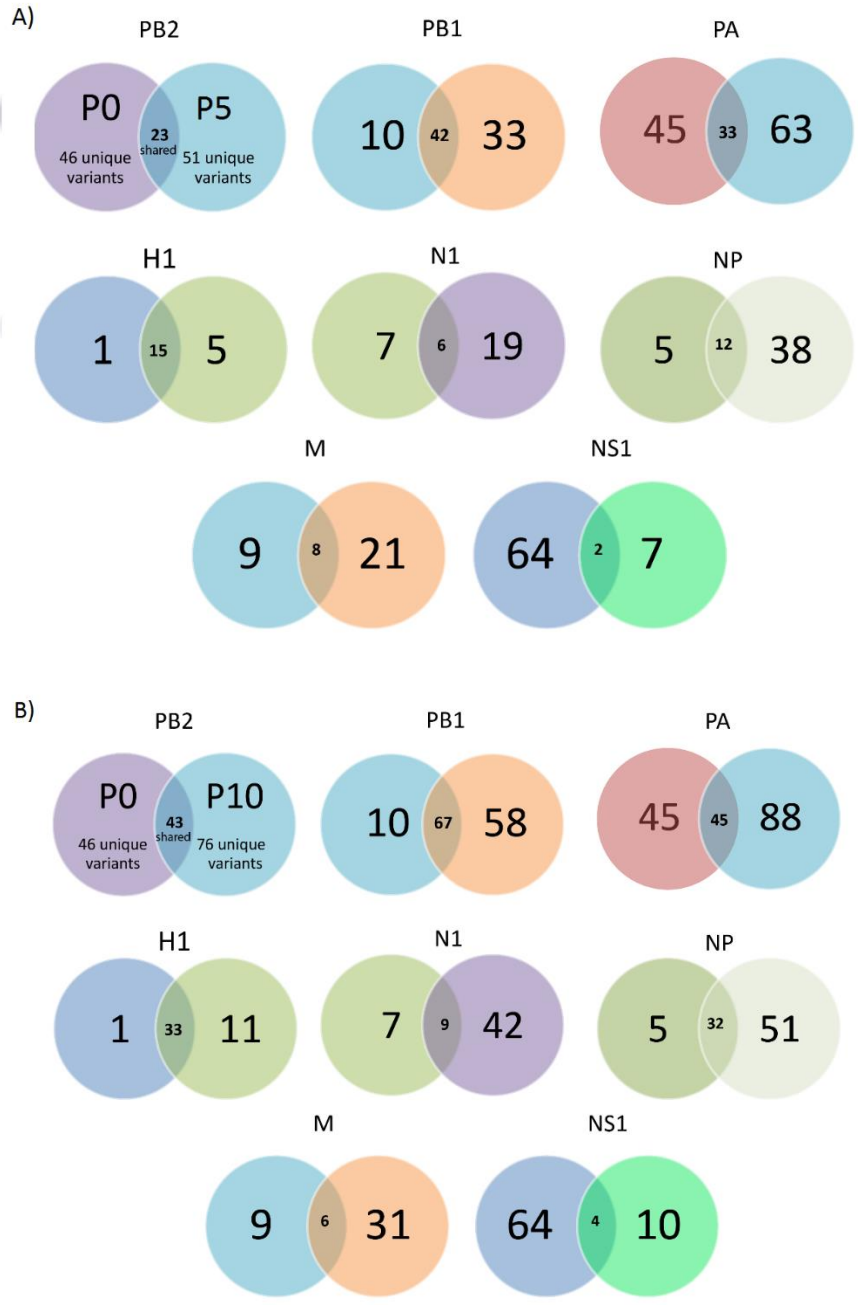


Figure 2.14: Venn diagrams of the overall number of virus variants in the initial swab sample (P0), the fifth passage (P5), and the last passage (P10). A) The total number of variants was compared between the initial swab sample (P0) and the fifth passage (P5). B) The total number of variants was compared between the initial swab sample (P0) and the last passage (P10).

From the overall number of virus variants presented in Figure 2.14, I then quantified the proportions of synonymous (dS) and non-synonymous (dN) substitutions in all segments of the initial swab sample (P0), the fifth passage (P5), and the last virus passage (P10) were analyzed (Figure 2.15).

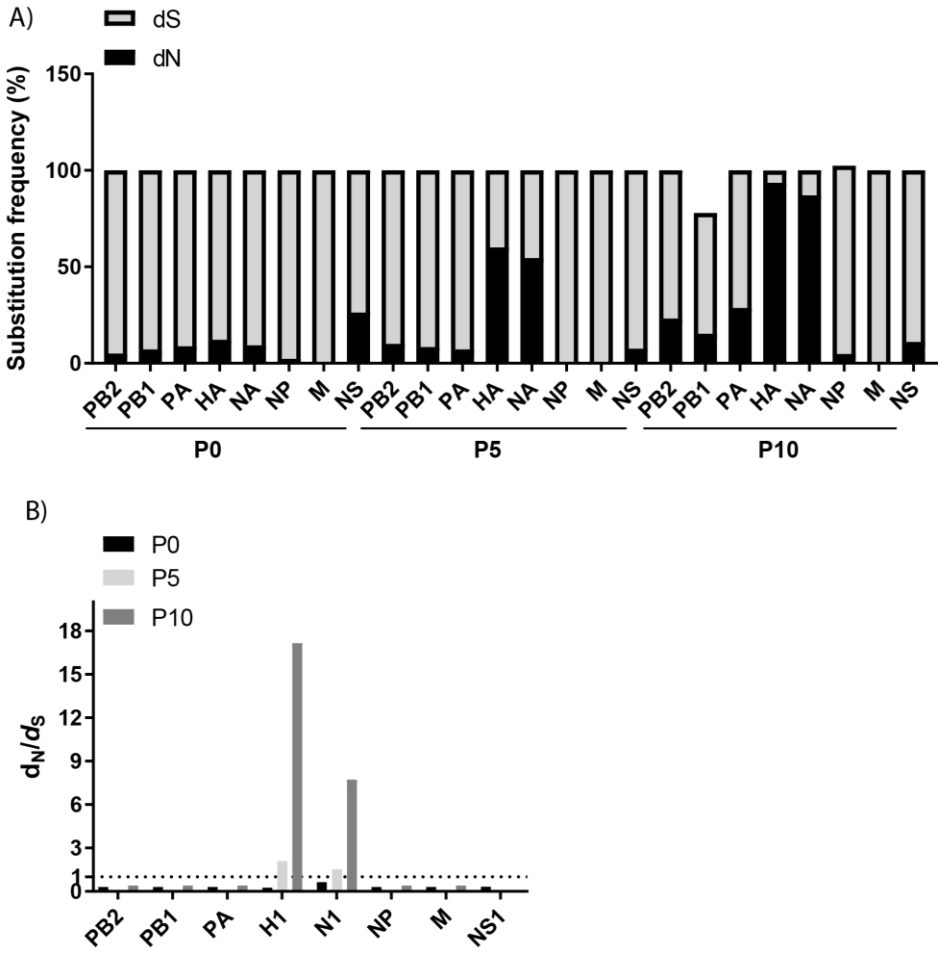


Figure 2.15: Proportions of non-synonymous (dN) and synonymous (dS) substitutions in the duck virus segments. A) Both dN and dS were determined in the initial swab sample (P0), the fifth passage (P5), and the last virus passage (P10) for all segments. B) The dN/dS ratio was determined for all segments in P0, P5 and P10.

There was an increase in the proportion of non-synonymous substitutions in the fifth and last virus passages for the HA and NA gene segments compared to P0 (Figure 2.15A). The proportion of non-synonymous substitutions was higher in P10 compared to P5. I next examined the direction of the evolution of the different gene segments with the use of dN/dS ratios we previously described in section 2.3.1.1. To resume, a dN/dS ratio >1 means that the sequence is under positive selection while a ratio <1 implies that the sequence is under negative selection. A ratio equal to 1 indicates that the gene is under neutral pressure. Both the HA and NA genes had a dN/dS ratio >1 in P5 and P10 meaning that both genes were under positive selective pressure over the passage of the virus (Figure 2.15, B).

2.3.2.2 Mutational analysis of the duck protein sequences

To investigate the effects of non-synonymous mutations introduced in the HA and NA proteins during the fifth and last virus passages, I analyzed their amino acid sequences. I performed protein alignments with the sequences from P5 and P10 and additional HA and NA proteins retrieved from the NCBI Influenza Virus Database in order to identify more- and less-conserved regions. Analysis of the HA protein during the different passages showed that amino acid changes were mostly being accumulated, for both P5 and P10, in conserved regions of the cleavage site and fusion domain of the protein. Also, some truncations and substitutions also appeared in the central region of HA2 which could negatively interfere with the protein functions (Figure 2.16).

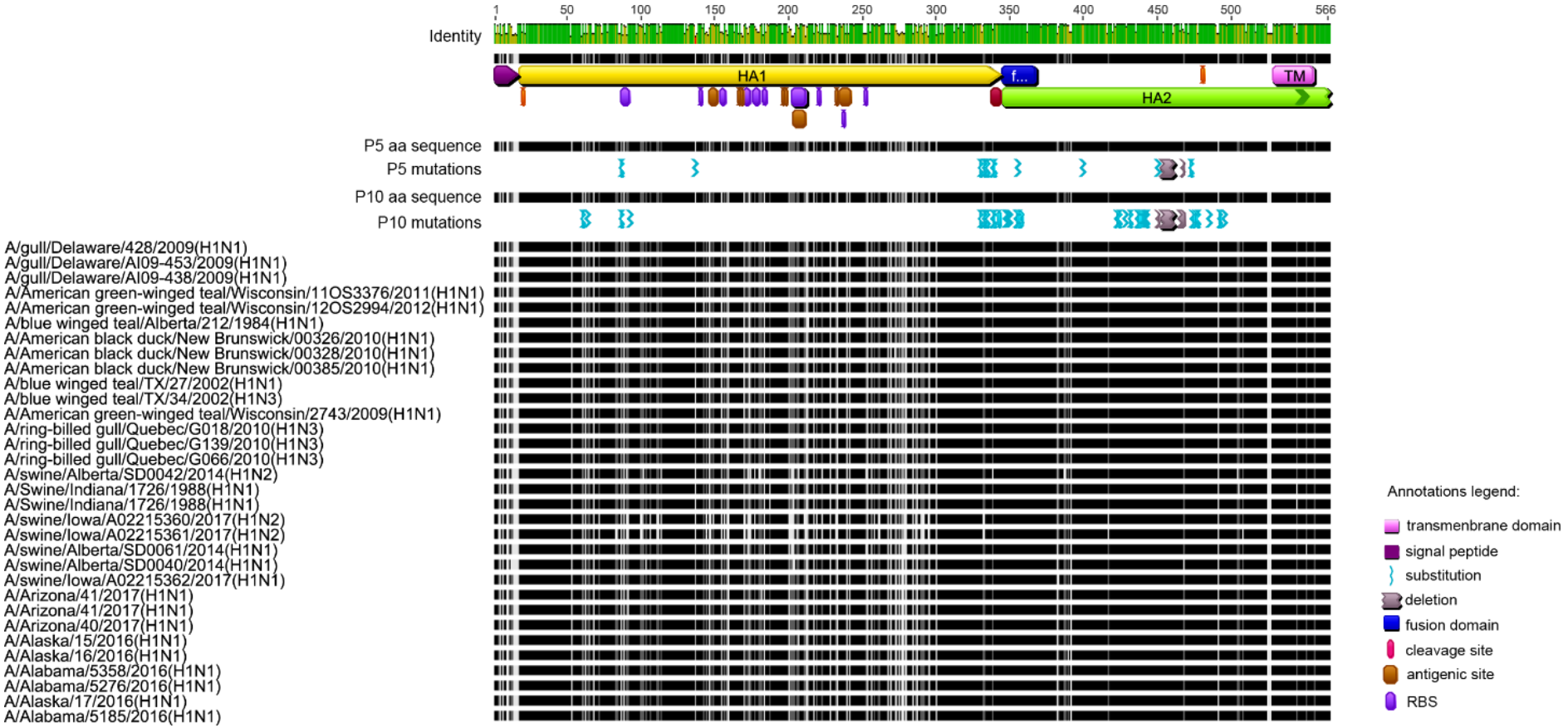

1
50
100
150
200
250
300
350
400
450
500
566
Identity
HA1
f...
TM
HA2
P5 aa sequence
P5 mutations
P10 aa sequence
P10 mutations
A/gull/Delaware/428/2009(H1N1)
A/gull/Delaware/AI09-453/2009(H1N1)
A/gull/Delaware/AI09-438/2009(H1N1)
A/American green-winged teal/Wisconsin/11OS3376/2011(H1N1)
A/American green-winged teal/Wisconsin/12OS2994/2012(H1N1)
A/blue winged teal/Alberta/212/1984(H1N1)
A/American black duck/New Brunswick/00326/2010(H1N1)
A/American black duck/New Brunswick/00328/2010(H1N1)
A/American black duck/New Brunswick/00385/2010(H1N1)
A/blue winged teal/TX/27/2002(H1N1)
A/blue winged teal/TX/34/2002(H1N3)
A/American green-winged teal/Wisconsin/2743/2009(H1N1)
A/ring-billed gull/Quebec/G018/2010(H1N3)
A/ring-billed gull/Quebec/G139/2010(H1N3)
A/ring-billed gull/Quebec/G066/2010(H1N3)
A/swine/Alberta/SD0042/2014(H1N2)
A/Swine/Indiana/1726/1988(H1N1)
A/Swine/Indiana/1726/1988(H1N1)
A/swine/Iowa/A02215360/2017(H1N2)
A/swine/Iowa/A02215361/2017(H1N2)
A/swine/Alberta/SD0061/2014(H1N1)
A/swine/Alberta/SD0040/2014(H1N1)
A/swine/Iowa/A02215362/2017(H1N1)
A/Arizona/41/2017(H1N1)
A/Arizona/41/2017(H1N1)
A/Arizona/40/2017(H1N1)
A/Alaska/15/2016(H1N1)
A/Alaska/16/2016(H1N1)
A/Alabama/5358/2016(H1N1)
A/Alabama/5276/2016(H1N1)
A/Alaska/17/2016(H1N1)
A/Alabama/5185/2016(H1N1)
Annotations legend:
transmenbrane domain
signal peptide
substitution
deletion
fusion domain
cleavage site
antigenic site
RBS

Figure 2.16 Comparison of the HA amino acid changes from the fifth and last virus passages. The fifth and last virus passages were aligned with other HA protein sequences with MUSCLE implemented in Geneious 8. The HA amino acid (aa) sequence is annotated above the alignment. The N-terminal part of the protein contains a peptide signal (17 aa) responsible for membrane anchoring at the cell surface. The remainder of the protein is divided into two domains, HA1 and HA2, with a cleavage site (7 aa) located between the two domains. A fusion domain (23 aa) is located in the HA2 domain. A transmembrane domain (28 aa) is located in the C-terminal part of the protein. Receptor binding sites (RBS) and antigenic sites are also displayed on the annotation. The amino acid changes (substitutions and deletions) on the fifth (P5 mutations) and last virus passage (P10 mutations) are indicated according to the legend.

Analysis of the NA amino acid sequences from P5 and P10 showed an accumulation of mostly substitutions in similar locations within the stalk and catalytic domains of the protein (Figure 2.17).

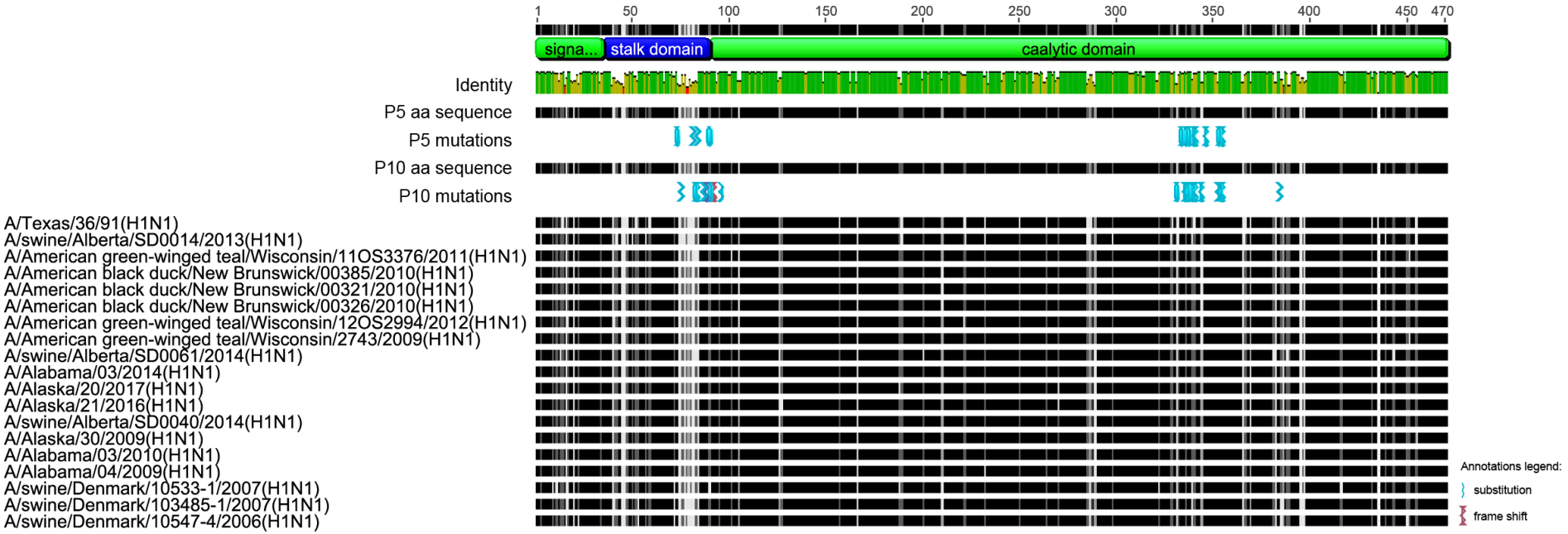

1
50
100
150
200
250
300
350
400
450
470
signa...
stalk domain
caalytic domain
Identity
P5 aa sequence
P5 mutations
P10 aa sequence
P10 mutations
A/Texas/36/91(H1N1)
A/swine/Alberta/SD0014/2013(H1N1)
A/American green-winged teal/Wisconsin/11OS3376/2011(H1N1)
A/American black duck/New Brunswick/00385/2010(H1N1)
A/American black duck/New Brunswick/00321/2010(H1N1)
A/American black duck/New Brunswick/00326/2010(H1N1)
A/American green-winged teal/Wisconsin/12OS2994/2012(H1N1)
A/American green-winged teal/Wisconsin/2743/2009(H1N1)
A/swine/Alberta/SD0061/2014(H1N1)
A/Alabama/03/2014(H1N1)
A/Alaska/20/2017(H1N1)
A/Alaska/21/2016(H1N1)
A/swine/Alberta/SD0040/2014(H1N1)
A/Alaska/30/2009(H1N1)
A/Alabama/03/2010(H1N1)
A/Alabama/04/2009(H1N1)
A/swine/Denmark/10533-1/2007(H1N1)
A/swine/Denmark/103485-1/2007(H1N1)
A/swine/Denmark/10547-4/2006(H1N1)
Annotations legend:
substitution
frame shift

Figure 2.17: Comparison of the NA amino acid changes from the fifth and last virus passages. The fifth and last virus passages were aligned with other NA protein sequences with MUSCLE implemented in Geneious. The NA amino acid (aa) sequence is annotated and displayed above the alignment. The N-terminal part of the protein contains a signal-anchoring domain (35 aa) followed by a stalk domain (55 aa), and a catalytic domain (380 aa). The amino acid changes (substitutions and frameshift) from P5 and P10 are indicated.

2.4 Discussion

Virus quasispecies are spectra of non-identical, but closely related viruses that are released after a viral infection (Andino and Domingo 2015). The virus quasispecies can change depending on the host environment (Schneider and Roossinck 2001). IAV genetic studies are often focused on the consensus nucleotide composition of the viral genome. Few research studies have undertaken investigations on IAV genomic structure at the quasispecies level. Thus, I performed an analysis of the genomic structure of viruses shed by three different bird hosts, murres, gulls, and ducks to investigate whether the virus quasispecies could be conserved within similar hosts. My analyses showed that viruses from the same bird host display similar quasispecies patterns. This was also demonstrated in one study when comparing the quasispecies of hepatitis C viruses within similar human recipients (Laskus, Wilkinson et al. 2004, Asghar, Pettersson et al. 2017). In my study, all viruses showed variant frequency levels that matched with other viruses

according to the corresponding phylogenetic relatedness of the viral segments. The proportions of shared variants were also correlated with the phylogenetic relationships of the segments such that similar genes in term of host and geographic origins share similar proportions of variants. Analysing the association of shared variant frequencies with the sequence homologies between segments (Supplementary figure 2.2) showed a general correlation between the proportion of shared variants and nucleotide sequence similarities. The pattern was most striking for the gull viruses, with slightly broader distributions for the murre and duck viruses. When comparing the variant diversities, as quantified by the Shannon diversity index, of two viruses with similar genotypes I saw that some segments harbour similar variant distributions but not all. This could be explained by the fact that viruses that harbour similar genotypes have and maintain similar variants within the same host. Although the quasispecies of viruses with similar genotypes share a large number of variants, the proportion of those shared variants within each quasispecies varied. The NS segments from the gull and murre viruses had a higher proportion of non-synonymous substitutions compared to the duck viruses. These differences might be explained by the fact that different bird host species carry different allelic versions of the NS segment. The gull and murre viruses contained NS allele A, which is found in bird and mammalian hosts, while the duck viruses in this study contained allele B, which is typically only found in bird species (Ludwig, Schultz et al. 1991). Differences in host genetic and phenotypic makeup could explain why certain viral genes are transmitted or not in certain hosts. The host immunity could act as a barrier for allowing the transmission of viral genes.

The effect of replication in a different host on the evolutionary adaptation of a duck virus in Atlantic puffin embryos was studied experimentally. A decrease of viral titer was observed during serial passage of the virus. An overall increase of variants within the virus quasispecies was observed in the fifth and last virus passages examined by deep sequencing, except for the NS segment. I would have expected an increase in viral titer once the virus adapted to a new host. Hence, additional passages could be required to see an increase of viral titers. Also, the technique I used to titer the virus, hemagglutination, might lack sensitivity to detect the virus at low concentrations. A sensitive QPCR assay to quantify the number of viral genome copies could be performed in parallel. Lastly, the possible presence of maternal antibodies in embryos from wild birds could restrict the virus replication and also affect the viral titer I measured.

I analyzed the accumulated mutations over the different passages within the amino acid sequences of both surface glycoproteins, HA and NA.

For the HA protein, most of the mutations were nucleotide substitutions that changed the amino acid sequence at the cleavage site or fusion domain. The fusion domain contains conserved amino acids that are necessary for the virus to be released into the cell's cytoplasm (Smrt, Draney et al. 2015). The HA cleavage site allows maturation of HA protein to produce functional and replicative viruses (Garten and Klenk 1983, Steinhauer 1999). A large number of mutations, nucleotide substitutions and deletions, were also found in the central domain of HA2, and these mutations could have an impact on the conformational changes during the fusion process (Armstrong, Kushnir et al. 2000). Surprisingly, few mutations were found in

RBS regions that are involved in cellular receptor recognition (Sriwilaijaroen and Suzuki 2012, Lazniewski, Dawson et al. 2017). A previous investigation of the adaptation of wild viruses to new avian hosts showed that new mutations accumulated in the HA1 domain of the HA protein and some of these mutations were located in RBS regions (Li and Cardona 2010, Li, zu Dohna et al. 2010). Other studies found similar findings to what I observed, with mutations in the HA cleavage site and HA2 domain (Hossain, Hickman et al. 2008, Dlugolenski, Jones et al. 2011). We could not directly compare mutations with those studies because the viruses do not have the same HA subtype.

For the NA protein, several mutations, mostly substitutions but also one frameshift, were found in the stalk and catalytic domains. A deletion within the stalk region has been previously reported during experimental infections of wild bird viruses in some poultry hosts, but not in wild bird hosts (Li and Cardona 2010, Dlugolenski, Jones et al. 2011). The NA stalk deletion conferred a replicative advantage of the virus in poultry hosts (Dlugolenski, 2011, Hossain, 2008, Pei, 2009). Deep sequencing analysis of a wild bird virus genome before introduction into poultry revealed the NA stalk deletion preexisted in the virus quasispecies (Croville, Soubies et al. 2012). Thus, it seems that the introduction of mutations within the stalk region of the NA protein I analyzed during the experimental assay could be the result of an adaption process of the duck virus in the Atlantic puffin host.

To conclude, my analysis of the of different IAVs shed by wild bird hosts shows a general conservation of overall quasispecies genomic structure for viruses replicating in the same host groups. This suggests that similar virus variants are

favored within each host group, and these are likely those best adapted to that host and which allow their effective transmission and preservation within the same host species. The study of the adaptation of a wild duck virus to a different avian host, conducted experimentally in bird embryos, showed major genetic changes in the HA and NA proteins despite of an overall decrease of viral titers. In our study, regions associated with the maturation of the HA and NA proteins appeared to be the most frequent regions where mutations accumulate, and this is similar to observations from other studies with different viruses and host groups. Despite giving insights on the origin of viruses in term of host and geographic location, applying high-throughput next-generation sequencing technologies also allows the dissection of the virus genomic structure which gives a better understanding on how the virus could evolve and be transmitted in different hosts. It allows us to foresee mutations that could have occurred and have an impact of the virus function and evolution in different hosts. This could also be a useful method to predict and detect emerging pathogenic viruses in their natural reservoir that could be deadly to other hosts.

2.5 References

Andino, R. and E. Domingo (2015). "Viral quasispecies." Virology 479-480: 46-51.

Armstrong, R. T., A. S. Kushnir and J. M. White (2000). "The Transmembrane Domain of Influenza Hemagglutinin Exhibits a Stringent Length Requirement to Support the Hemifusion to Fusion Transition." The Journal of Cell Biology 151(2): 425-438.

Asghar, N., J. H. Pettersson, P. Dinnetz, A. Andreassen and M. Johansson (2017). "Deep sequencing analysis of tick-borne encephalitis virus from questing ticks at natural foci reveals similarities between quasispecies pools of the virus." The Journal of General Virology 98(3): 413-421.

Croville, G., S. M. Soubies, J. Barbieri, C. Klopp, J. Mariette, O. Bouchez, C. Camus-Bouclainville and J. L. Guerin (2012). "Field monitoring of avian influenza viruses: whole-genome sequencing and tracking of neuraminidase evolution using 454 pyrosequencing." Journal of Clinical Microbiology 50(9): 2881-2887.

Dlugolenski, D., L. Jones, G. Saavedra, S. M. Tompkins, R. A. Tripp and E. Mundt (2011). "Passage of low-pathogenic avian influenza (LPAI) viruses mediates rapid genetic adaptation of a wild-bird isolate in poultry." Archives of Virology 156(4): 565-576.

Domingo, E., E. Baranowski, C. M. Ruiz-Jarabo, A. M. Martín-Hernández, J. C. Sáiz and C. Escarmís (1998). "Quasispecies structure and persistence of RNA viruses." Emerging Infectious Diseases 4(4): 521-527.

Garten, W. and H. D. Klenk (1983). "Characterization of the carboxypeptidase involved in the proteolytic cleavage of the influenza haemagglutinin." The Journal of General Virology 64 (Pt 10): 2127-2137.

Hoffmann, E., J. Stech, Y. Guan, R. G. Webster and D. R. Perez (2001). "Universal primer set for the full-length amplification of all influenza A viruses." Arch Virol 146(12): 2275-2289.

Holland, J., K. Spindler, F. Horodyski, E. Grabau, S. Nichol and S. VandePol (1982). "Rapid evolution of RNA genomes." Science 215(4540): 1577-1585.

Hossain, M. J., D. Hickman and D. R. Perez (2008). "Evidence of Expanded Host Range and Mammalian-Associated Genetic Changes in a Duck H9N2 Influenza Virus Following Adaptation in Quail and Chickens." PLoS One 3(9): e3170.

Huang, Y., M. Wille, J. Benkaroun, H. Munro, A. L. Bond, D. A. Fifield, G. J. Robertson, D. Ojkic, H. Whitney and A. S. Lang (2014). "Perpetuation and reassortment of gull influenza A viruses in Atlantic North America." Virology 456-457: 353-363.

Huang, Y., M. Wille, A. Dobbin, N. M. Walzthoni, G. J. Robertson, D. Ojkic, H. Whitney and A. S. Lang (2014). "Genetic structure of avian influenza viruses from ducks of the Atlantic flyway of North America." PLoS One 9(1): e86999.

Jukes, T. H. and C. R. Cantor (1969). CHAPTER 24 - Evolution of Protein Molecules. Mammalian Protein Metabolism. H. N. Munro, Academic Press: 21-132.

Kuiken, T., E. C. Holmes, J. McCauley, G. F. Rimmelzwaan, C. S. Williams and B. T. Grenfell (2006). "Host species barriers to influenza virus infections." Science 312(5772): 394-397.

Laskus, T., J. Wilkinson, J. F. Gallegos-Orozco, M. Radkowski, D. M. Adair, M. Nowicki, E. Operskalski, Z. Buskell, L. B. Seeff, H. Vargas and J. Rakela (2004). "Analysis of hepatitis C virus quasispecies transmission and evolution in patients infected through blood transfusion." Gastroenterology 127(3): 764-776.

Lazniewski, M., W. K. Dawson, T. Szczepinska and D. Plewczynski (2017). "The structural variability of the influenza A hemagglutinin receptor-binding site." Brief Funct Genomics.

Li, J. and C. J. Cardona (2010). "Adaptation and transmission of a wild duck avian influenza isolate in chickens." Avian Diseases 54(1 Suppl): 586-590.

Li, J., H. zu Dohna, N. L. Anchell, S. C. Adams, N. T. Dao, Z. Xing and C. J. Cardona (2010). "Adaptation and transmission of a duck-origin avian influenza virus in poultry species." Virus Research 147(1): 40-46.

Ludwig, S., U. Schultz, J. Mandler, W. M. Fitch and C. Scholtissek (1991). "Phylogenetic relationship of the nonstructural (NS) genes of influenza A viruses." Virology 183(2): 566-577.

Schneider, W. L. and M. J. Roossinck (2001). "Genetic Diversity in RNA Virus Quasispecies Is Controlled by Host-Virus Interactions." Journal of Virology 75(14): 6566-6571.

Smrt, S. T., A. W. Draney and J. L. Lorieau (2015). "The Influenza Hemagglutinin Fusion Domain Is an Amphipathic Helical Hairpin That Functions by Inducing Membrane Curvature." The Journal of Biological Chemistry 290(1): 228-238.

Sriwilaijaroen, N. and Y. Suzuki (2012). "Molecular basis of the structure and function of H1 hemagglutinin of influenza virus." Proceedings of the Japan Academy. Series B, Physical and Biological Sciences 88(6): 226-249.

Steinhauer, D. A. (1999). "Role of hemagglutinin cleavage for the pathogenicity of influenza virus." Virology 258(1): 1-20.

Verhoeven, J. T. P., M. Canuti, H. J. Munro, S. C. Dufour and A. S. Lang (2018). "ViDiT-CACTUS: an inexpensive and versatile library preparation and sequence analysis method for virus discovery and other microbiology applications." Canadian Journal of Microbiology: 1-13.

Webster, R. G., W. J. Bean, O. T. Gorman, T. M. Chambers and Y. Kawaoka (1992). "Evolution and ecology of influenza A viruses." Microbiological Reviews 56(1): 152-179.

Yoon, S. W., R. J. Webby and R. G. Webster (2014). "Evolution and ecology of influenza A viruses." Current topics in Microbiology and Immunology 385: 359-375.

Chapter 3: Analysis of the Variability in the Non-Coding Regions of Influenza A Viruses

Abstract

The genomes of influenza A viruses (IAVs) comprise eight negative-sense single-stranded RNA segments. In addition to the protein-coding region, each segment possesses 5' and 3' non-coding regions (NCR) that are important for transcription, replication and packaging. The NCRs contain both conserved and segment-specific sequences, and the impacts of variability in the NCRs are not completely understood. Full NCRs have been determined from some viruses, but a detailed analysis of potential variability in these regions among viruses from different host groups and locations has not been performed. To evaluate the degree of conservation in NCRs among different viruses, we sequenced the NCRs of IAVs isolated from different wild bird host groups (ducks, gulls and seabirds). We then extended our study to include NCRs available from the National Center for Biotechnology Information (NCBI) Influenza Virus Database, which allowed us to analyze a wider variety of host species and more HA and NA subtypes. We found that the amount of variability within the NCRs varies among segments, with the greatest variation found in the HA and NA and the least in the M and NS segments. Overall, variability in NCR sequences was correlated with the coding region phylogeny, suggesting vertical coevolution of the (coding sequence) CDS and NCR regions.

3.1 Introduction

Influenza A viruses (IAVs) are most well known for their circulation in humans, causing yearly seasonal epidemics and occasional pandemics, and for outbreaks in poultry. However, the viral strains responsible for these infections represent only a small portion of the total known IAV diversity, most of which is maintained in the natural wild bird reservoir.

IAVs are members of the family Orthomyxoviridae and have segmented negative-sense single-stranded RNA genomes. The viral genome is composed of 8 segments that encode at least 11 proteins, depending on the strain. Each segment has non-coding regions (NCRs) at both ends. The 3' NCR contains 12 conserved nucleotides followed by a segment-specific region of variable length. The 5' NCR contains 13 conserved nucleotides and a segment-specific region of variable length. The two conserved regions are partially complementary (Skehel and Hay 1978, Robertson 1979, Desselberger, Racaniello et al. 1980) and form a panhandle-like structure (Hsu, Parvin et al. 1987, Bergmann and Muster 1996, Kim, Fodor et al. 1997) involved in the transcription and replication steps of the viral life cycle (Seong and Brownlee 1992, Fodor, Pritlove et al. 1994, Flick, Neumann et al. 1996). Mutagenesis studies identified residues within the NCRs that are required for the viral polymerase complex to bind and initiate transcription (Li and Palese 1992, Fodor, Seong et al. 1993). However, a binding preference for the 5' NCR was demonstrated (Pritlove, Fodor et al. 1995). There is also a stretch of uridines at the 5' end that is involved in the polyadenylation of viral mRNAs (Luo, Luytjes et al. 1991, Li and Palese 1994, Pritlove, Poon et al. 1998, Poon, Pritlove et al. 1999). Further

studies have revealed the presence of some variability within the two NCRs at different positions. There is variability in the fourth nucleotide (C or U) in the 3' end of all RNA segments (Robertson 1979), which affects the promoter and could play a role in differential regulation of transcription and replication (Lee, Bae et al. 2003). Non-conserved nucleotides located in the segment-specific regions have also been shown to play an important role in viral replication (Zheng, Palese et al. 1996). Mutagenesis studies have revealed the importance of NCRs for the incorporation of the eight different gene segments during viral assembly (Fujii, Fujii et al. 2005, Gao and Palese 2009, Zhao, Peng et al. 2014, Crescenzo-Chaigne, Barbezange et al. 2017). These packaging signals can influence the reassortment of genes during co-infections of different strains by preventing their incorporation during the virus genome packaging (Gao and Palese 2009, White, Steel et al. 2017).

Amplification of complete IAV genomes with universal primers was initially designed based on the conserved nature of NCRs (Hoffmann, Stech et al. 2001), but amplification biases have been reported (Zhou, Donnelly et al. 2009, Widjaja, de Vries et al. 2012). To date, a thorough analysis of IAV NCRs from different hosts and across segments has not been performed. We hypothesized that the NCRs within both the conserved and the segment-specific regions may vary according to the original host species. To test our hypothesis, we characterized the true NCRs of IAV genomes from different wild bird species by the rapid amplification of cDNA ends (RACE) method and then expanded our analyses to include additional viral subtypes, hosts and geographic origins.

3.2. Materials and Methods

3.2.1. NCR sequence determination

NCR sequences from previously identified and sequenced viruses originating from three different wild bird host groups (ducks, gulls and seabirds) were determined: A/domestic duck/Newfoundland/MW668/2010(H1N1), A/American black duck/Newfoundland/GR679/2011(H3N2), A/herring gull/Newfoundland/GR578/2011(H13N6), A/herring gull/Newfoundland/GR848/2011(H13N6), A/common murre/Newfoundland/AB318/2011(H1N2), and A/common murre/Newfoundland/AB324/2011(H1N2). From these six viruses, we determined 92 NCR sequences, comprising 46 3' and 46 5' NCRs. When possible, viruses with the same subtype were chosen for NCR characterization.

Viral genomic RNA was extracted from allantoic fluids of specific pathogen free eggs inoculated with the different viruses using the Trizol reagent (Thermo Fisher Scientific). Primers were designed to efficiently amplify the 3' and 5' NCRs of the targeted viruses based on the full genomic sequences of viruses that were available from previous studies (Supplementary table 3.1).

For the 3' NCRs, genomic RNAs were polyadenylated with poly(A) polymerase (2 U µL-1) (Thermo Fisher Scientific) in the presence of $MnCl_2$ (2.5 mM) and ATP (1 mM) in 1X reaction buffer and a final volume of 50 µL. Reactions were incubated at 37°C for 30 min and the resulting RNA used as the input for the 3' RACE System for Rapid Amplification of cDNA Ends (Thermo Fisher Scientific). For the 5' NCRs, the 5' RACE System for Rapid Amplification of cDNA Ends (Thermo Fisher Scientific) was used

according to the manufacturer's recommendations. The resulting amplicons were purified and sequenced using Sanger chemistry at The Center for Applied Genomics (Toronto, Canada).

3.2.2 Sequence analyses

The 92 NCRs determined from our samples were compared with NCRs retrieved from the influenza resource database (http://www.ncbi.nlm.nih.gov/genomes/FLU/FLU.html), chosen to represent different viral subtypes, host species, geographic origins, and collection dates. When possible, all available or a high number of nucleotide sequences of complete gene segments were downloaded from the database. The number of NCRs analyzed are indicated in Figure 3.3. Sequence alignments were performed using MUSCLE, implemented in MEGA version 6 (Tamura, Stecher et al. 2013). Multiple alignments were then edited in Geneious version 8 (Kearse, Moir et al. 2012). The CDS regions of segments that had complete NCRs were removed from the sequence, with only the NCRs kept for further analysis. Complete segments without NCRs were removed from the analysis.

Pairwise genetic distance matrices determined with the Maximum Composite Likelihood method, generated by MEGA, were used for classical multidimensional analysis (MDS) in R (R Development Core Team 2008). The viral segments were assigned geographic origins based on nucleotide BLAST searches (Altschul, Gish et al. 1990) and neighbor-joining phylogenetic analyses in MEGA. High bootstrap values are shown at the nodes on the trees. A small number of sequences was added in

each multiple alignment figures due to space limitation, but the exact number of NCRs originally aligned with MUSCLE are indicated in section 3.3.2.

3.3 Results and Discussion

3.3.1 Characterization of NCRs from different wild bird viruses

Comparison of the 3’ and 5’ NCRs we determined with the RACE method showed that the 12- and 13-nt conserved regions, respectively, contained no novel substitutions in any of the different wild bird viruses (Figures 3.1 and 3.2), which agrees with previous studies (Skehel and Hay 1978, Robertson 1979, Desselberger, Racaniello et al. 1980).

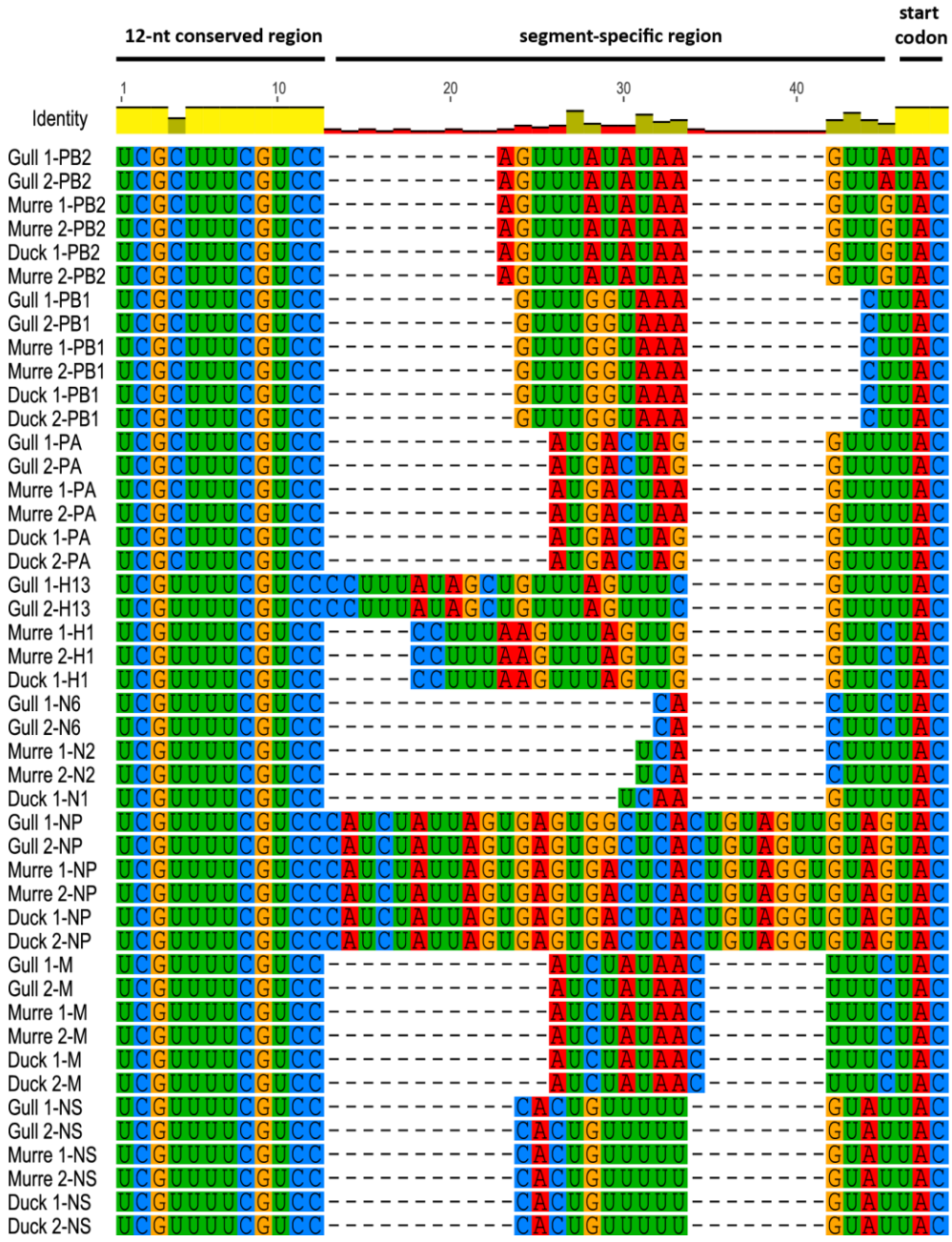


Figure 3.1: 3' NCRs of wild bird virus segments determined in this study. The multiple alignment of 3' NCRs was generated with MUSCLE, implemented in Geneious version 8. The RNA sequences are displayed in the 3' to 5' orientation corresponding to the packaged RNAs.

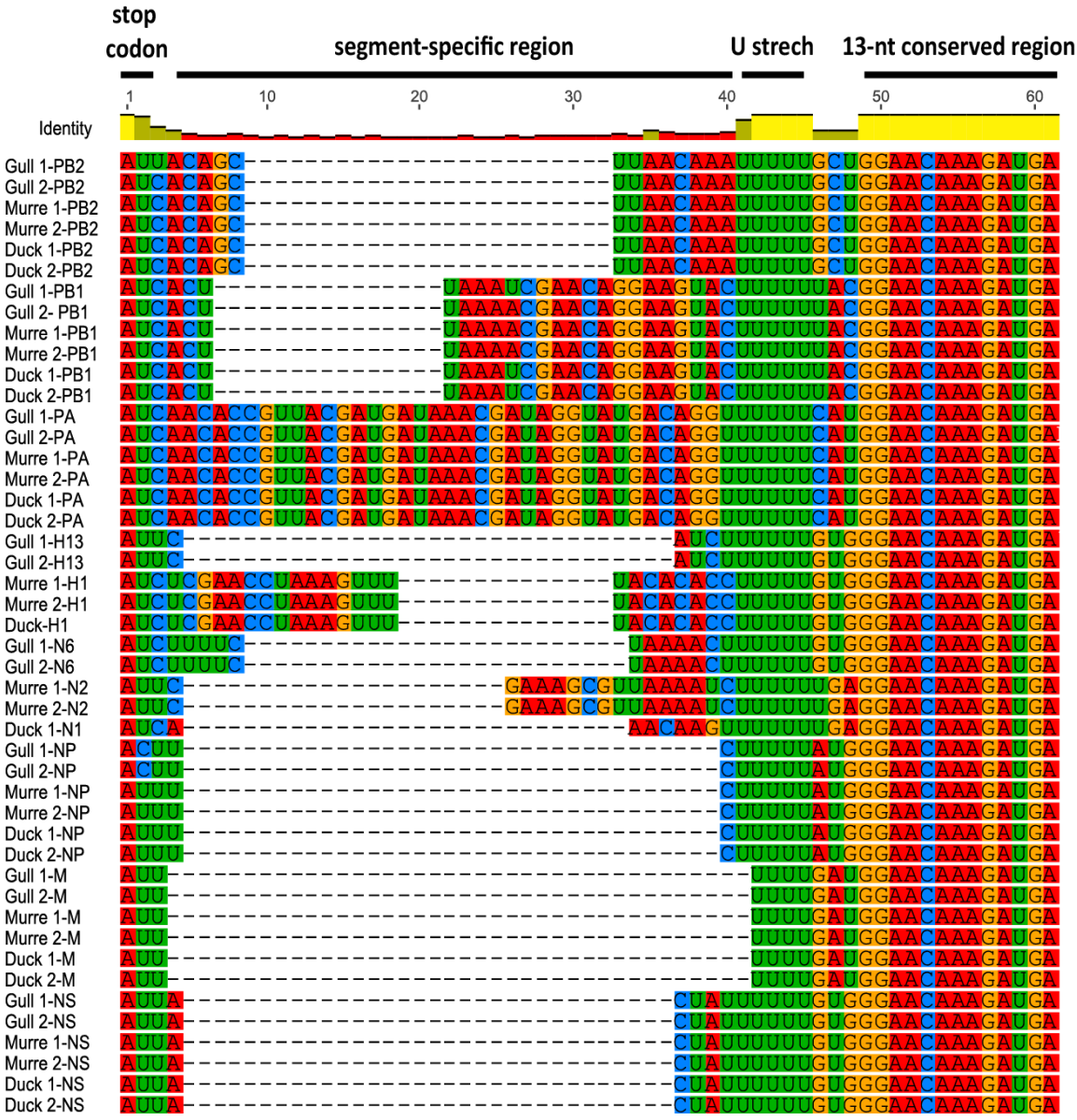


Figure 3.2: 5' NCRs of wild bird virus segments determined in this study. The multiple alignment of 5' NCRs was generated with MUSCLE implemented in Geneious version 8. The RNA sequences are displayed in the 3' to 5' orientation corresponding to the packaged RNAs.

Comparison of the 3' NCRs (Figure 3.1) showed that the 4th position in the 12-nucleotide conserved region displays the expected variability, with the polymerase

gene segments, PB2, PB1 and PA, containing a cytosine residue while the other segments contain a uridine residue, as previously observed (Robertson 1979, Lee, Bae et al. 2003).

For the 3' segment-specific regions (Figure 3.1), some of the segments (PB1, M and NS) had identical sequences across all six viruses, but the others showed some differences. The murre and gull PB2 and NP segments were identical whereas the gull viruses differed from the others for these segments. The PA segment was identical in the gull and duck viruses but differed from these in the murre viruses. The HA sequences were identical within subtypes, even for the H1 viruses that came from murres and a duck. Similarly, the NA sequences were identical within the same subtype.

Comparison of the 5' NCRs showed the conservation of a stretch of 5 or 6 uridine residues upstream of the 13-nt conserved region among the different wild bird viruses (Figure 3.2). Similar to the 3' NCRs, the segment-specific regions varied among segments, with some of the individual segments' NCRs conserved among the six viruses and others varying among the viruses (Figure 3.2). All PB2, PA, M, and NS segments' NCRs were identical among the viruses. There was a single variation in the PB1 segment, with the murre and one gull virus differing from the other gull virus and the duck viruses. As observed for the 3' NCRs, the HA and NA segments were the same within subtypes, and the H1 and H2 segments were also identical.

3.3.2 The variability in segment-specific NCR sequences differs among segments

To further investigate the differences observed in the NCRs of our limited set of six viruses, we expanded our analysis of NCR genetic diversity to include viral

sequences from a variety of host species and geographic origins retrieved from the NCBI Influenza Virus Database (Supplementary Table 3.2). There is no apparent relationship between the number of sequences analyzed and the NCR sequence diversity (Figure 3.3).

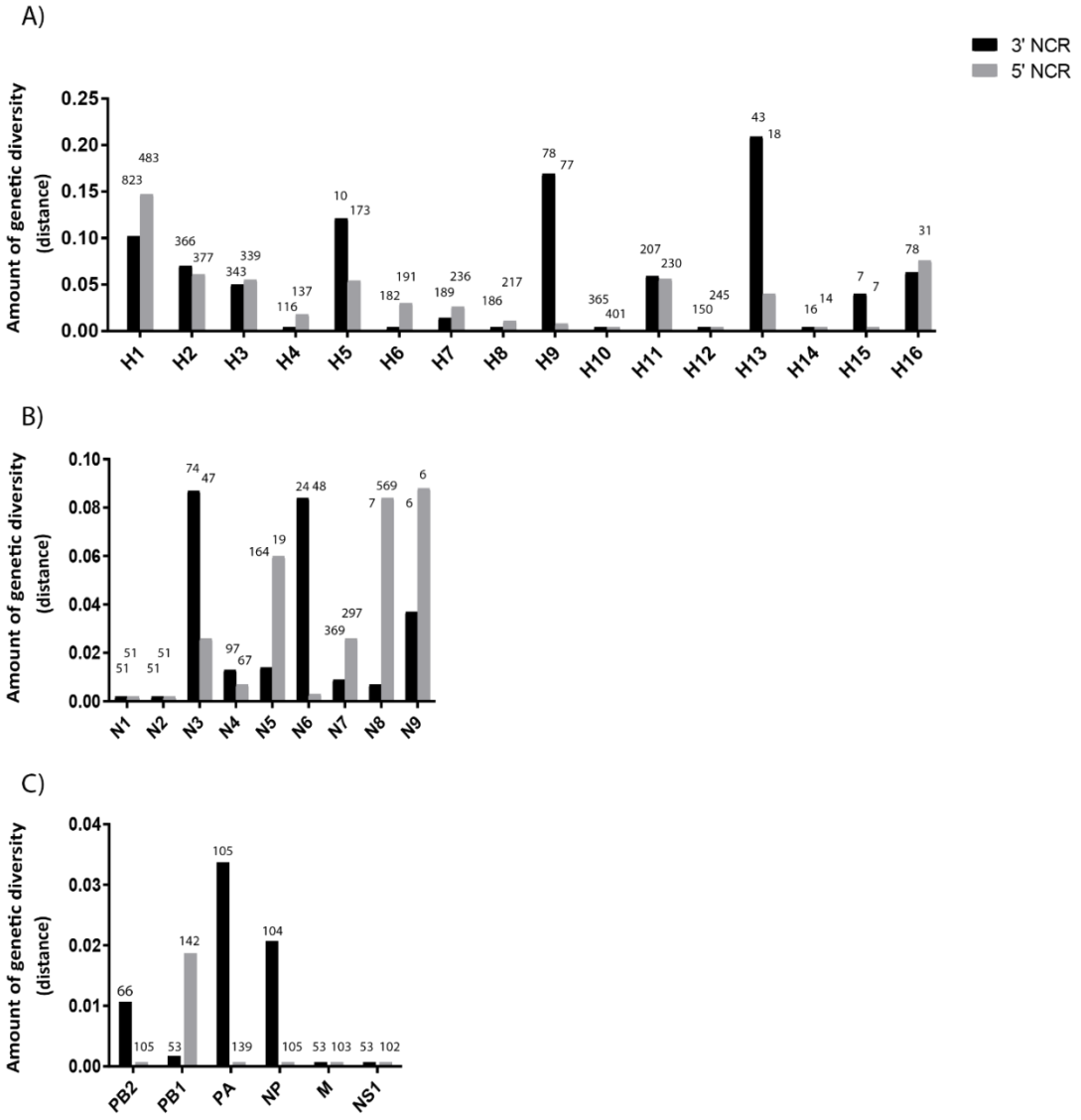


Figure 3.3: Genetic diversity of 3' and 5' NCRs for different segments. (A) HA segments of the different subtypes. (B) NA segments of the different subtypes. (C) "Internal" protein-coding segments. The genetic diversity for each of the NCR regions for the different segments was calculated as mean genetic distance for the NCRs included in our analyses. The genetic distance was determined with the Maximum Composite Likelihood method implemented in MEGA. The number above each bar represents the number of sequences used to calculate the amount of genetic diversity for each segment.

The greatest amount of variability in NCRs was observed in the HA segments (Figure 3.3A), but there is variability among the segments and between 3' and 5' NCRs for some subtypes. The H4, H6, H7, H8, H10, H12, and H14 NCRs showed very little variability overall. The H5, H9 and H13 subtypes show greater variability within their 3 'NCRs relative to their 5' NCRs, whereas the H1, H2, H3, H11 and H16 were equally variable at the two ends. The lower variability at the NCRs could be attributed to a lack of host diversity or a bias of sampling.

The genetic diversity of different NA subtypes' NCRs (Figure 3.3B), which originated from a large number of host species and geographic locations, showed that the N3 and N6 subtypes have more diversity at their 3' ends while the N5, N8, and N9 are more diverse at the 5' ends. The N1 and N2 viruses showed very little diversity, despite the large assortment of hosts from which viruses were included.

Matching the general pattern observed for IAV segment coding regions, the NCRs of the remaining segments are overall less variable than found for HA and NA (Figure 3.3C). The M and NS segments showed the lowest diversity, in agreement with the slower evolution rates for these segments' CDS [e.g. (Bahl, Vijaykrishna et al. 2009, Qu, Zhang et al. 2011)]. These differences we have observed across many viruses may not be too surprising given that a previous analysis of NCRs from a limited dataset of human H3N2 viruses circulating worldwide since 1968 found that NCR variability differed among the segments (Furuse and Oshitani 2011).

3.3.3 Patterns of variability within NCRs can be explained by viral host species and geographic origins

To better understand the origin of the segment-specific NCR diversity, we assessed whether the NCRs of viruses from different hosts and geographic origins showed similar patterns of relationships as the respective coding regions (CDS). We used a multidimensional scaling (MDS) analysis based on pairwise genetic distance matrices for the NCR sequences, which allowed us to visualize the relationships among NCRs relative to the viral hosts and geographic origins. The results of these analyses were then compared to MDS analyses performed on the respective coding regions. We focused this analysis on segments that showed greater NCR diversity: H1, H9 and H13 for the hemagglutinin, N6 for the neuraminidase, and the NP segment.

The relationships among the 3' NCRs and CDS regions of a set of H9 segments were compared (Figure 3.4).

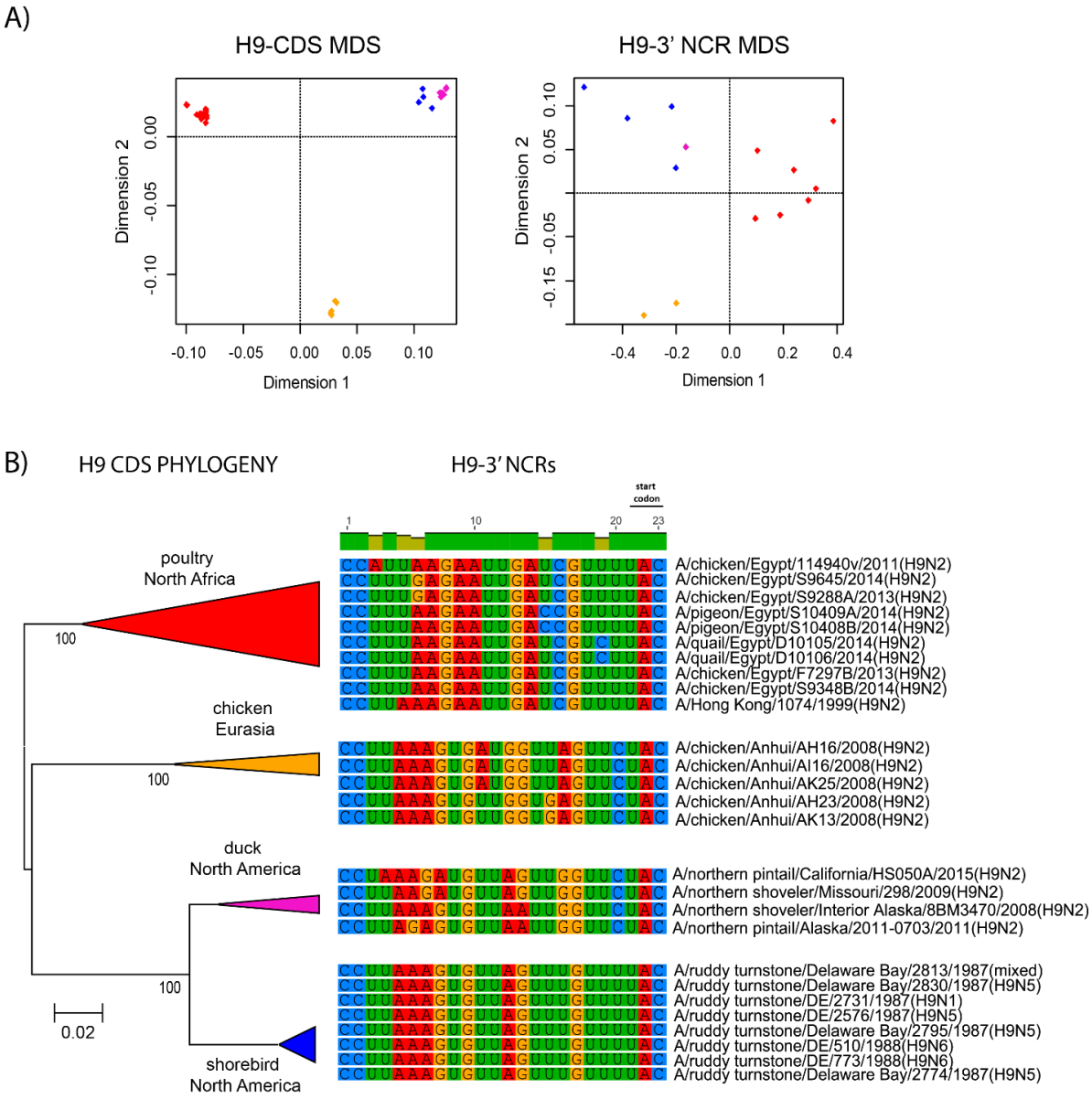


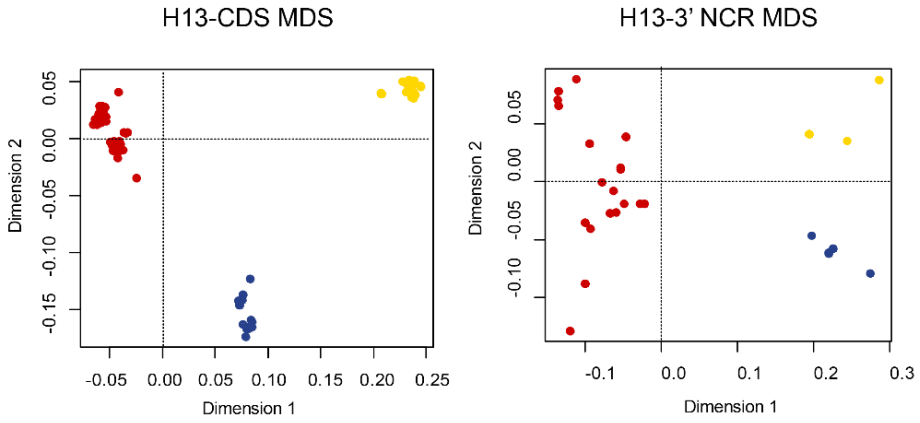
Figure 3.4: Relationships among H9 3' NCR and CDS regions for viruses from different host species and geographic locations. The MDS scatter plots (A) are based on the pairwise genetic distance matrices from the multiple alignments of complete CDS regions and 3'NCRs. Each dot represents a virus, colored according to their origins as indicated on the CDS-based phylogenetic tree (B). Alignments of the 3'

NCRs, shown corresponding to the genomic RNAs, from the viruses used in the analyses are next to the CDS-based neighbor-joining tree for these viruses.

The 3’ NCRs and CDS regions showed similar clustering in the MDS analyses (Figure 3.4A), corresponding to their CDS phylogenetic clades (Figure 3.4B). The two North American clade sequences grouped more closely and are more distant from the Eurasian and North African sequences (Figure 3.4A). This matches the patterns for the CDS phylogeny and 3’ NCR alignments (Figure 3.4B).

The same patterns were observed with the H13 subtype 3’ NCRs in the MDS analyses (Figure 3.5A), and there was again a clear correspondence between the CDS phylogeny and the NCR sequence relationships (Figure 3.5B).

A)



B) H13 CDS PHYLOGENY

H13-3' NCRs

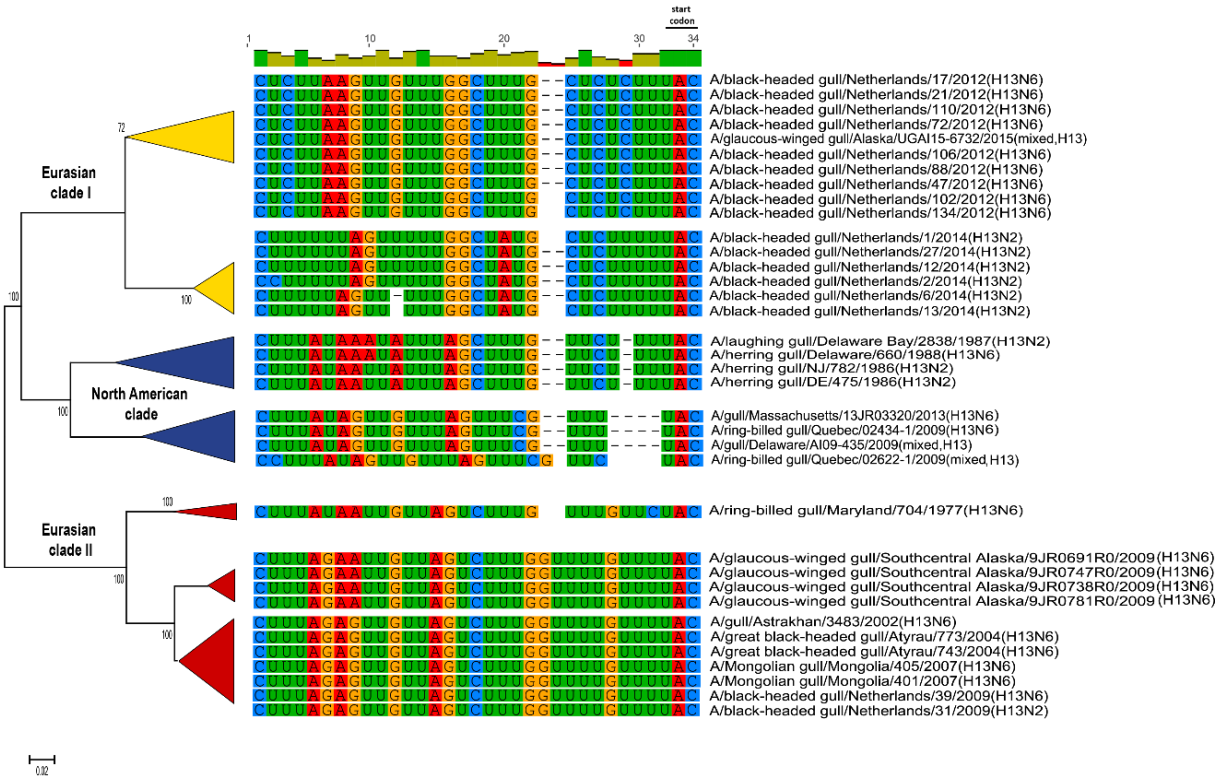


Figure 3.5: Relationships among H13 3' NCR and CDS regions for viruses from different host species and geographic locations. The MDS scatter plots (A) are based on the pairwise genetic distance matrices from the multiple alignments of complete CDS regions and 3'NCRs. Each dot represents a virus, colored according to their origins as indicated on the CDS-based phylogenetic tree (B). Alignments of the 3' NCRs, shown corresponding to the genomic RNAs, from the viruses used in the analyses are next to the CDS-based neighbor-joining tree for these viruses.

For the H1 subtype, the CDS phylogeny contains larger host-specific (swine, human and avian) clades that separate into geographic-specific (Eurasian and North American) clades (Figure 3.6).

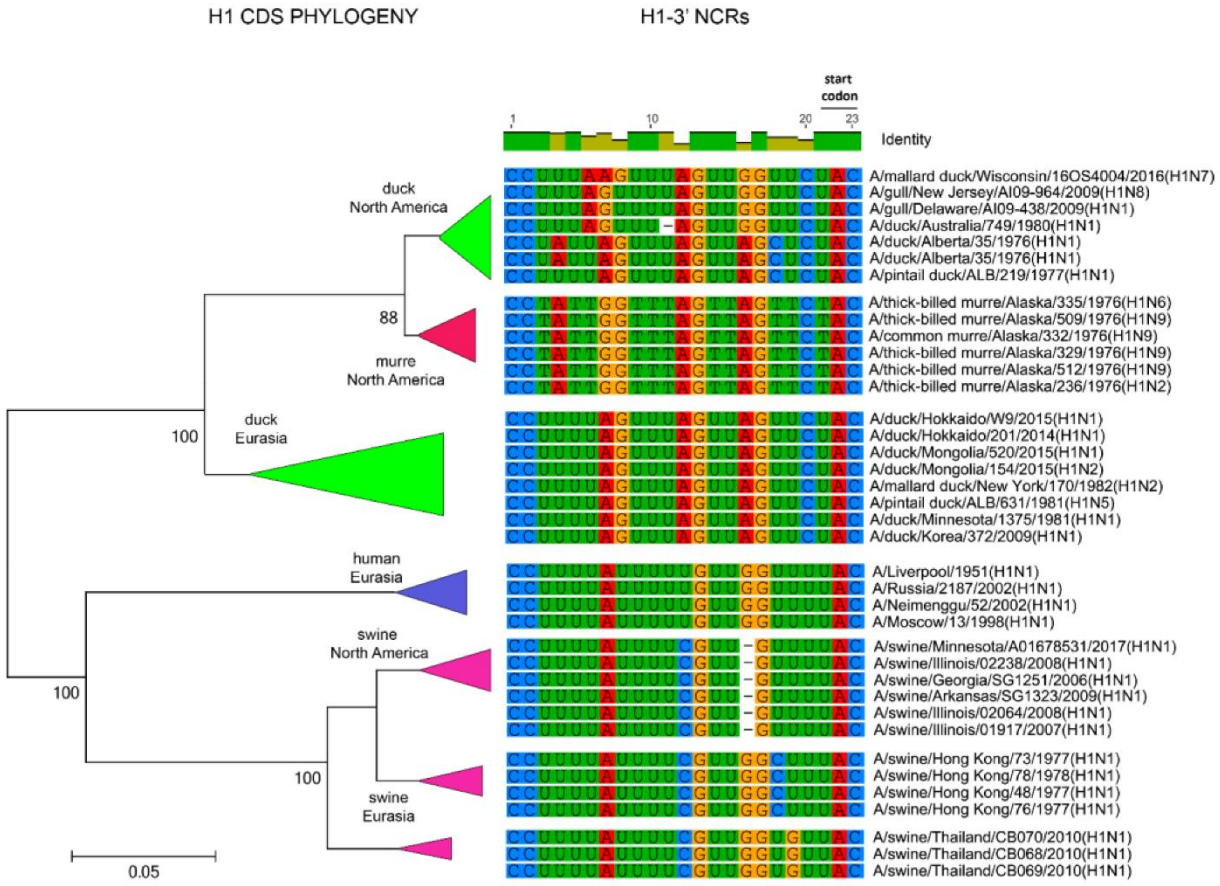


Figure 3.6: Relationships among H1 3' NCR and CDS regions for viruses from different host species and geographic locations. Alignments of the 3' NCRs, shown corresponding to the genomic RNAs, from the viruses used in the analyses are next to the CDS-based neighbor-joining tree for these viruses.

The correspondence between CDS phylogeny and 3' NCR similarities was again found for these H1 sequences (Figure 3.6). A similar pattern was also found for the N6 NA (Figure 3.7) and NP sequences (Figure 3.8), where the CDS phylogenies also determined the 3' NCR relationships. This was true even when the viruses falling within specific clades originated from different hosts (e.g. gulls and ducks) or

different continental origins (Eurasia and North America). Overall, these patterns suggest vertical coevolution of the CDS and NCRs.

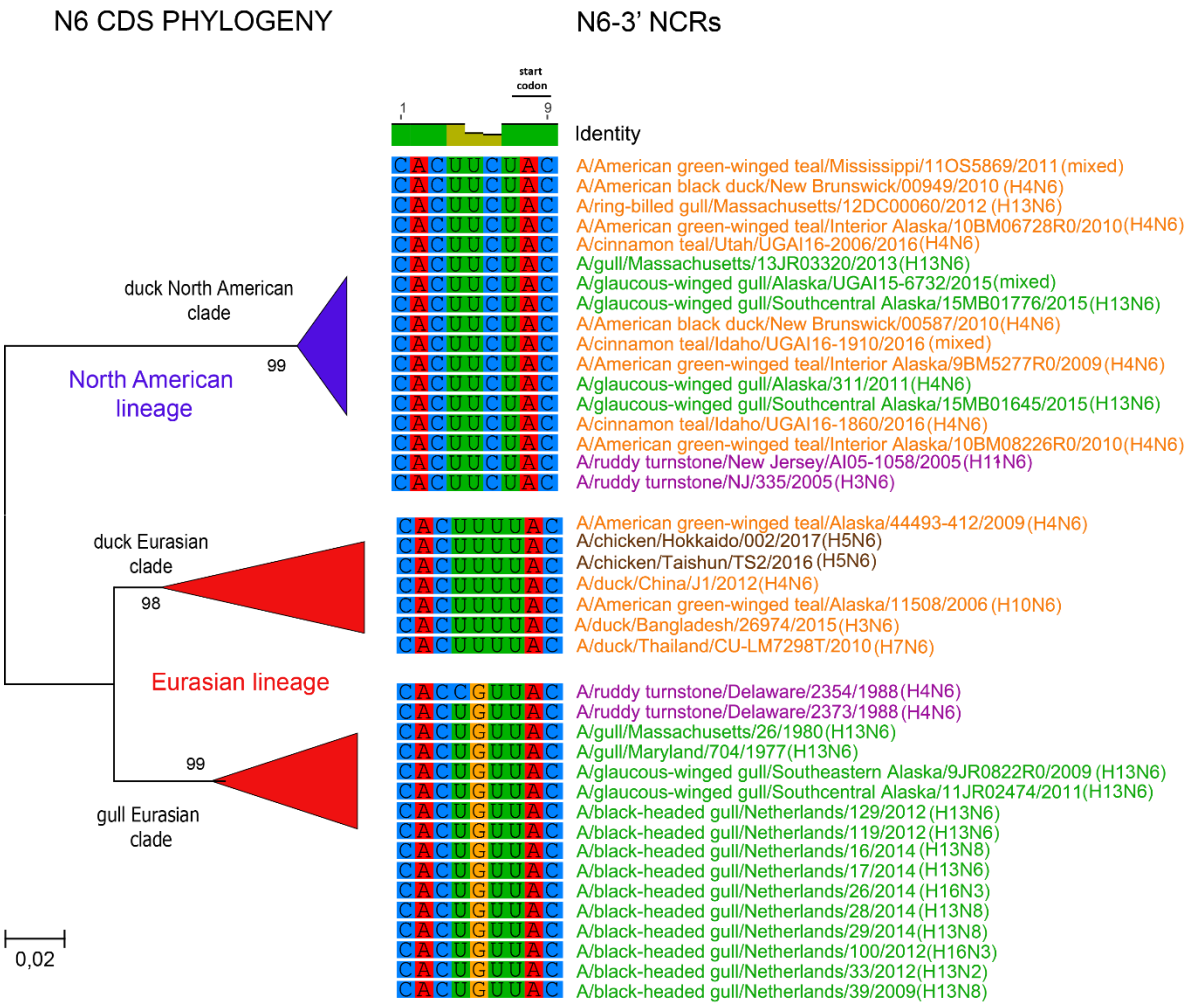


Figure 3.7: Relationships among N6 3' NCR and CDS regions for viruses from different host species and geographic locations. Alignments of the 3' NCRs, shown corresponding to the genomic RNAs, from the viruses used in the analyses are next to the CDS-based neighbor-joining tree for these viruses. To facilitate the distinction of viruses from different hosts, gull viruses are in green, duck viruses are in orange, shorebird viruses are in purple, and chicken viruses are in brown.

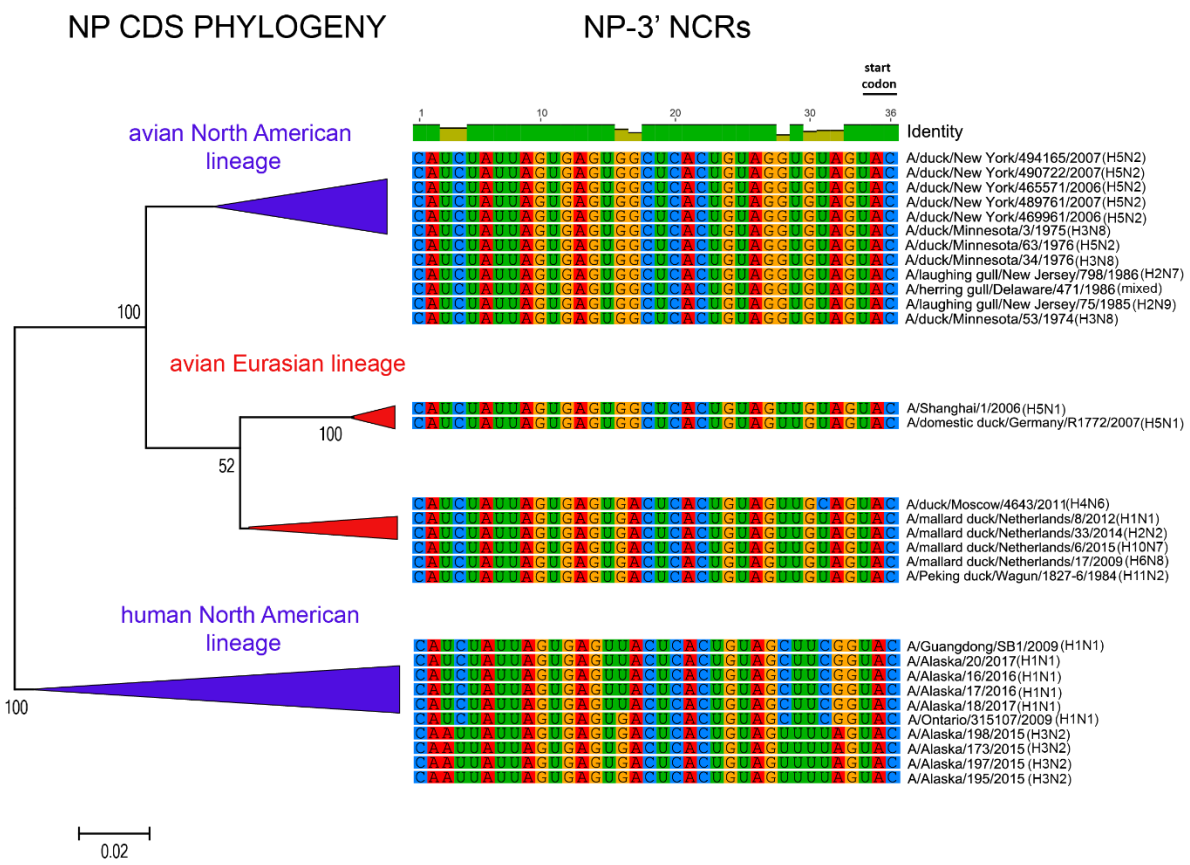


Figure 3.8: Relationships among NP 3' NCR and CDS regions for viruses from different host species and geographic locations. Alignments of the 3' NCRs, shown corresponding to the genomic RNAs, from the viruses used in the analyses are next to the CDS-based neighbor-joining tree for these viruses.

Lastly, we looked at the M and NS segments, which showed very low overall NCR diversity (Figure 3.3C). These two segments share very similar 3' NCRs among different viruses, regardless of the host species or geographic origins (Figure 3.9, Supplementary figure 3.1).

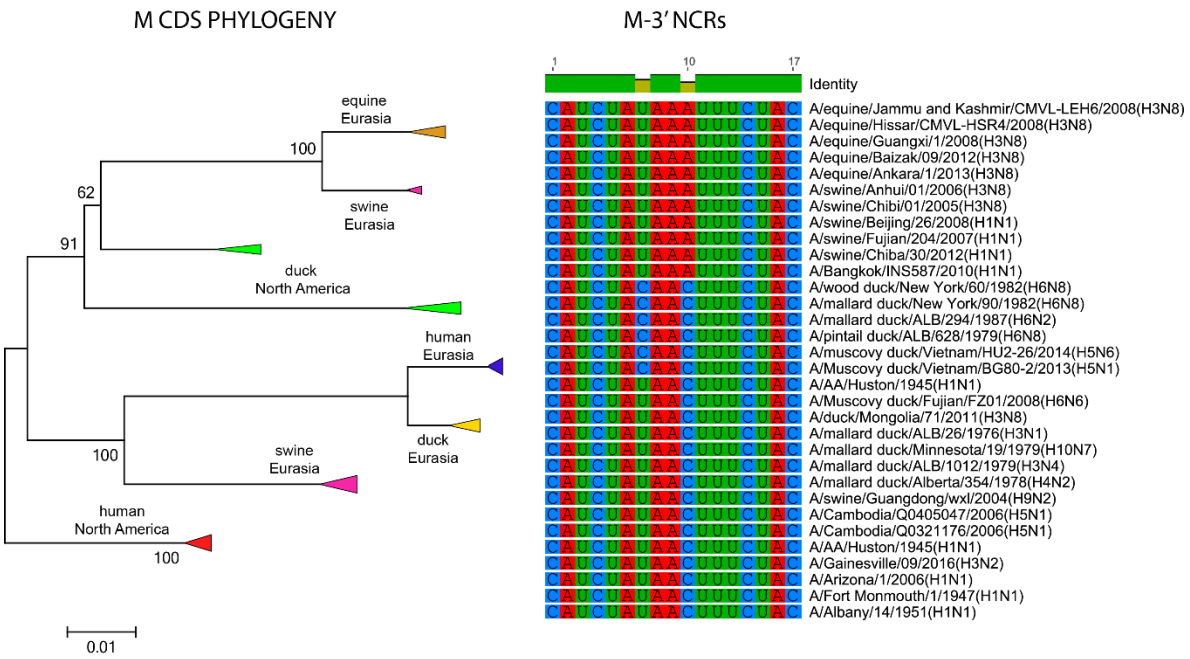


Figure 3.9: Relationships among M 3' NCR and CDS regions for viruses from different host species and geographic locations. Alignments of the 3' NCRs, shown corresponding to the genomic RNAs, from the viruses used in the analyses are next to the CDS-based neighbor-joining tree for these viruses.

Some variations were found for the M segment, for a total of three different M 3' NCR sequences (Figure 3.9), but they do not show the same correspondence to the phylogeny of their CDS as observed for the other segments. Comparison of a subset of NS NCRs from varied hosts and geographic locations found these all to be identical (Supplementary figure 3.1). The conservation of the M and NS NCRs may be linked to the slower evolutionary rate of the coding regions of these segments, as mentioned earlier.

3.4 General conclusions

We first hypothesized that NCRs may vary according to the host species. Our initial evaluation of the 3’ and 5’ NCRs from viruses isolated from different wild bird hosts showed that the conserved regions are identical among different hosts, but the segment-specific regions showed some variability among the viruses for some segments. This prompted a larger analysis including more sequences, which showed that the amount of variability in the NCRs varies among the different segments and can also vary between the 3’ and 5’ ends, depending on the segment. The overall variability is highest for the HA and NA segments. Examination of 3’ NCRs from HA and NA subtypes with higher variability indicates that the relationships among NCR sequences matches the coding region phylogenies, which generally follow patterns based on the viral host species and geographic origins, and possibly reflects a lack of recombination in these segments. The most conserved NCRs are found on the M and NS segments, which also show the slowest coding sequence evolutionary rates. The exact evolutionary pressures acting on the NCRs and consequences of sequence changes in these regions remain to be fully defined.

3.5 References

Altschul, S. F., W. Gish, W. Miller, E. W. Myers and D. J. Lipman (1990). "Basic local alignment search tool." Journal of Molecular Biology. 215(3): 403-410.

Bahl, J., D. Vijaykrishna, E. C. Holmes, G. J. Smith and Y. Guan (2009). "Gene flow and competitive exclusion of avian influenza A virus in natural reservoir hosts." Virology 390(2): 289-297.

Bergmann, M. and T. Muster (1996). "Mutations in the nonconserved noncoding sequences of the influenza A virus segments affect viral vRNA formation." Virus Research 44(1): 23-31.

Crescenzo-Chaigne, B., C. V. Barbezange, S. Leandri, C. Roquin, C. Berthault and S. van der Werf (2017). "Incorporation of the influenza A virus NA segment into virions does not require cognate non-coding sequences." Scientific reports 7: 43462.

Desselberger, U., V. R. Racaniello, J. J. Zazra and P. Palese (1980). "The 3' and 5'-terminal sequences of influenza A, B and C virus RNA segments are highly conserved and show partial inverted complementarity." Gene 8(3): 315-328.

Flick, R., G. Neumann, E. Hoffmann, E. Neumeier and G. Hobom (1996). "Promoter elements in the influenza vRNA terminal structure." Rna 2(10): 1046-1057.

Fodor, E., D. C. Pritlove and G. G. Brownlee (1994). "The influenza virus panhandle is involved in the initiation of transcription." Journal of Virology 68(6): 4092-4096.

Fodor, E., B. L. Seong and G. G. Brownlee (1993). "Photochemical cross-linking of influenza A polymerase to its virion RNA promoter defines a polymerase binding site at residues 9 to 12 of the promoter." The Journal of General Virology 74 (Pt 7): 1327-1333.

Fujii, K., Y. Fujii, T. Noda, Y. Muramoto, T. Watanabe, A. Takada, H. Goto, T. Horimoto and Y. Kawaoka (2005). "Importance of both the coding and the segment-specific noncoding regions of the influenza A virus NS segment for its efficient incorporation into virions." Journal of Virology 79(6): 3766-3774.

Furuse, Y. and H. Oshitani (2011). "Evolution of the influenza A virus untranslated regions." Infection, Genetics and Evolution 11(5): 1150-1154.

Gao, Q. and P. Palese (2009). "Rewiring the RNAs of influenza virus to prevent reassortment." PNAS 106(37): 15891.

Hoffmann, E., J. Stech, Y. Guan, R. G. Webster and D. R. Perez (2001). "Universal primer set for the full-length amplification of all influenza A viruses." Archives of Virology 146(12): 2275-2289.

Hsu, M. T., J. D. Parvin, S. Gupta, M. Krystal and P. Palese (1987). "Genomic RNAs of influenza viruses are held in a circular conformation in virions and in infected cells by a terminal panhandle." PNAS 84(22): 8140-8144.

Kearse, M., R. Moir, A. Wilson, S. Stones-Havas, M. Cheung, S. Sturrock, S. Buxton, A. Cooper, S. Markowitz, C. Duran, T. Thierer, B. Ashton, P. Meintjes and A. Drummond (2012). "Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data." Bioinformatics 28(12): 1647-1649.

Kim, H. J., E. Fodor, G. G. Brownlee and B. L. Seong (1997). "Mutational analysis of the RNA-fork model of the influenza A virus vRNA promoter in vivo." The Journal of General Virology 78 (Pt 2): 353-357.

Lee, M. K., S. H. Bae, C. J. Park, H. K. Cheong, C. Cheong and B. S. Choi (2003). "A single-nucleotide natural variation (U4 to C4) in an influenza A virus promoter exhibits a large structural change: implications for differential viral RNA synthesis by RNA-dependent RNA polymerase." Nucleic Acids Research 31(4): 1216-1223.

Li, X. and P. Palese (1992). "Mutational analysis of the promoter required for influenza virus virion RNA synthesis." Journal of Virology 66(7): 4331-4338.

Li, X. and P. Palese (1994). "Characterization of the polyadenylation signal of influenza virus RNA." Journal of Virology 68(2): 1245-1249.

Luo, G. X., W. Luytjes, M. Enami and P. Palese (1991). "The polyadenylation signal of influenza virus RNA involves a stretch of uridines followed by the RNA duplex of the panhandle structure." Journal of Virology 65(6): 2861-2867.

Poon, L. L., D. C. Pritlove, E. Fodor and G. G. Brownlee (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(4): 3473-3476.

Pritlove, D. C., E. Fodor, B. L. Seong and G. G. Brownlee (1995). "In vitro transcription and polymerase binding studies of the termini of influenza A virus cRNA: evidence for a cRNA panhandle." The Journal of General Virology 76 (Pt 9): 2205-2213.

Pritlove, D. C., L. L. Poon, E. Fodor, J. Sharps and G. G. Brownlee (1998). "Polyadenylation of influenza virus mRNA transcribed in vitro from model virion RNA templates: requirement for 5' conserved sequences." Journal of Virology 72(2): 1280-1286.

Qu, Y., R. Zhang, P. Cui, G. Song, Z. Duan and F. Lei (2011). "Evolutionary genomics of the pandemic 2009 H1N1 influenza viruses (pH1N 1v)." Virology Journal 8: 250.

R Development Core Team (2008). R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing.

Robertson, J. S. (1979). "5' and 3' terminal nucleotide sequences of the RNA genome segments of influenza virus." Nucleic Acids Research 6(12): 3745-3757.

Seong, B. L. and G. G. Brownlee (1992). "Nucleotides 9 to 11 of the influenza A virion RNA promoter are crucial for activity in vitro." The Journal of General Virology 73 (Pt 12): 3115-3124.

Skehel, J. J. and A. J. Hay (1978). "Nucleotide sequences at the 5' termini of influenza virus RNAs and their transcripts." Nucleic Acids Research 5(4): 1207-1219.

Tamura, K., G. Stecher, D. Peterson, A. Filipski and S. Kumar (2013). "MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0." Molecular Biology and Evolution 30(12): 2725-2729.

White, M. C., J. Steel and A. C. Lowen (2017). "Heterologous Packaging Signals on Segment 4, but Not Segment 6 or Segment 8, Limit Influenza A Virus Reassortment." Journal of Virology 91(11).

Widjaja, I., E. de Vries, P. J. M. Rottier and C. A. M. de Haan (2012). "Competition between Influenza A Virus Genome Segments." PLOS One 7(10): e47529.

Zhao, L., Y. Peng, K. Zhou, M. Cao, J. Wang, X. Wang, T. Jiang and T. Deng (2014). "New Insights into the Nonconserved Noncoding Region of the Subtype-Determinant Hemagglutinin and Neuraminidase Segments of Influenza A Viruses." Journal of Virology 88(19): 11493-11503.

Zheng, H., P. Palese and A. Garcia-Sastre (1996). "Nonconserved nucleotides at the 3' and 5' ends of an influenza A virus RNA play an important role in viral RNA replication." Virology 217(1): 242-251.

Zhou, B., M. E. Donnelly, D. T. Scholes, K. St George, M. Hatta, Y. Kawaoka and D. E. Wentworth (2009). "Single-reaction genomic amplification accelerates sequencing and vaccine production for classical and Swine origin human influenza a viruses." Journal of Virology 83(19): 10309-10313.

Chapter 4: Analysis of influenza A viruses from gulls: An evaluation of inter-regional movements and interactions with other avian and mammalian influenza A viruses

In this chapter, the results and discussion sections are combined in one section and the methodology of the different analyses are described in their corresponding figure legends.

Abstract

Birds, including members of the families Anatidae (waterfowl) and Laridae (gulls and terns), serve as the major reservoir of influenza A viruses (IAVs). The ecogeographic contributions of gulls to global IAV dynamics, in terms of geographic scale and virus movements, are important and are distinct from those of waterfowl. Gulls primarily carry the H13 and H16 subtypes, yet can be infected by additional subtypes. Also, gulls are frequently infected by IAVs that contain mixtures of genes from different geographic phylogenetic lineages (e.g. North American and Eurasian). The present analysis examines a variety of viruses isolated from gulls and terns across the world that exhibit particularly high phylogenetic affinities to viruses found in other hosts. This illustrates the potential for gulls to act as highly pathogenic virus carriers, disseminators of viruses over long distances, and contributors in the genesis of pandemic strains. The historical evolution of an entirely Eurasian gull virus isolated in North America was also traced and indicates the Caspian Sea, in Southwestern

Asia, was an important area for the generation of this virus, and analysis of IAVs from terns also points to this region as relevant for the generation of novel strains.

4.1 Introduction

Wild birds, including waterfowl, shorebirds and gulls, harbor the vast majority of known influenza A virus (IAV) genetic diversity. Within these avian influenza IAVs, there is clear distinction between Eurasian and North American gene pools (Olsen, Munster et al. 2006) and there is increasing evidence of additional phylogeographic lineages (Pereda, Uhart et al. 2008, Gonzalez-Reiche and Perez 2012, Hurt, Vijaykrishna et al. 2014). However, bird movements lead to inter-regional transmission of IAV genes and their establishment in new regions. Intercontinental transmission of IAV genes from Eurasia to America has been frequently observed, and isolation of reassortant viruses that contain gene segments of different continental origins is particularly common from gulls in these regions, which is presumably facilitated by gull movements across the oceans. The North Atlantic Ocean is one location where gull movements between North America and Europe have been documented (Wille, Robertson et al. 2011), and intercontinental IAV reassortants have also been identified frequently (Echeverría, Moratorio et al. 2015). Additionally, both entirely Eurasian and North American gull IAVs have been found in Iceland (Dusek, Hallgrimsson et al. 2014), suggesting this location as important for the mixing of viruses between regions. An equivalent position has also been proposed for Greenland (Shoham and Rogers 2006). In North America, only two wholly Eurasian avian influenza A viruses have been found to date. One was an H16N3 virus from an American herring gull (*Larus smithsonianus*) in eastern Canada,

(Echeverría, Moratorio et al. 2015). The other was the highly pathogenic avian influenza (HPAI) H5N8 virus that recently emerged in China, spread to South Korea, Japan and Siberia, and is proposed to have been transmitted through the Beringia region by migratory waterfowl into Pacific North America (Lee, Swayne et al. 2015). This H5N8 arrival ended the long speculation about whether or not virulent Asian avian strains, such as H5N1, could be transmitted into the Americas by migrating birds. There is also a third case, involving H9N2 viruses from Alaska that were nearly identical (>99%) to viruses in China and South Korea, which were proposed to be carried to Alaska by migratory waterfowl (Ramey, Reeves et al. 2015). In this case, however, the gene segments of the viruses showed mixed geographic origins.

In addition to geographic divisions in IAV genes, distinct host group lineages are found. These include mainly the avian, gull, human, swine and equine lineages (Olsen, Munster et al. 2006). These distinctions make it possible to recognize that inter-species transmissions of IAVs between gulls and other hosts have occurred. Although most IAVs from gulls contain gull-lineage gene segments, and most are of the hemagglutinin (HA) subtypes H13 and H16, various viruses found in gulls contain HA genes and other gene segments that do not appear to be gull-specific, but rather fall within avian clades and are closely related to waterfowl viruses in phylogenetic analyses (Wille, Robertson et al. 2011, Van Borm, Rosseel et al. 2012, Hall, TeSlaa et al. 2013). The opportunities for these transmissions certainly exist, as gulls share terrestrial, freshwater and marine habitats with other birds and mammals, and therefore gulls could be important for moving IAVs amongst ecosystems and for contributing to the transmission of highly pathogenic IAVs to other hosts.

The present analysis is an investigation of the potential for gulls to act as carriers of highly pathogenic viruses, disseminators of viruses over long distances, and contributors in the genesis of pandemic strains. This is done through examination a variety of gull and tern viruses isolated from different locations that exhibit particularly high phylogenetic affinities to viruses found in other hosts.

4.2 Results and Discussion

4.2.1 Ecogenetic aspects of gull IAVs

Avian IAVs constitute the major, and likely the primal, component within the total IAV gene pool worldwide (Webster, Bean et al. 1992, Olsen, Munster et al. 2006). Distinct phylogeographic lineages, such as those representing Eurasia and North America, can be identified within avian IAVs, mainly as a result of partial segregation of migratory birds (Olsen, Munster et al. 2006). The first indication that gulls might represent a distinct reservoir for IAVs came from surveillance work on the Atlantic coast of North America, with the identification of a new HA antigen, H13, distinct from the previous 12 HA antigens identified in other avian species (Hinshaw, Air et al. 1982). Two decades later, an additional novel HA subtype, H16, was identified and characterized from gulls in Sweden (Fouchier, Munster et al. 2005). Continued surveillance and research have shown that gulls are predominantly infected by these H13 and H16 subtypes, and these subtypes are not often found outside of gulls (other than in shorebirds) (Arnal, Vittecoq et al. 2015), (Figure 4.1).

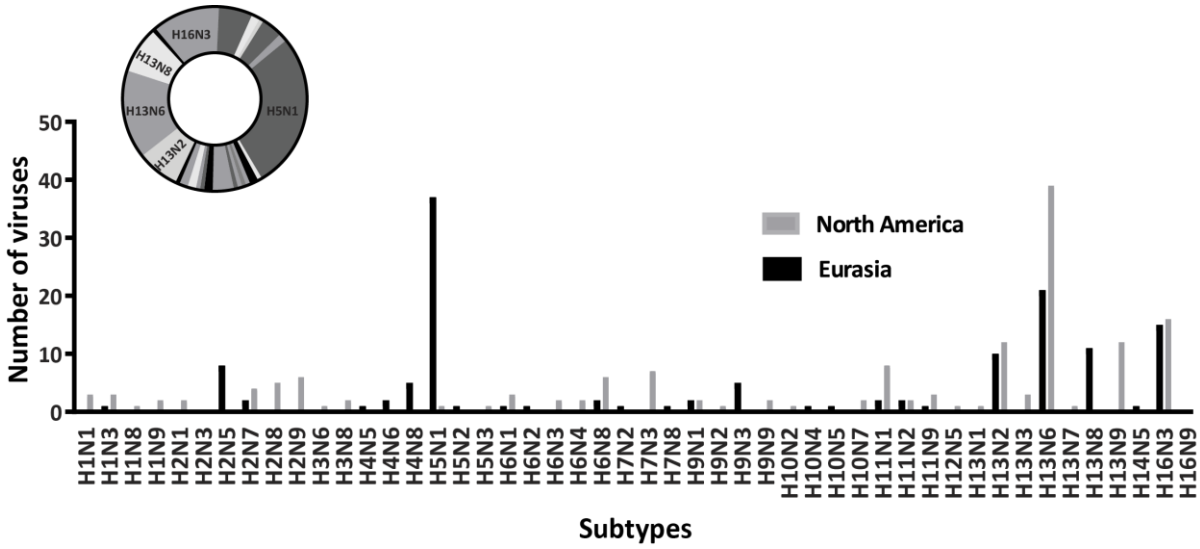


Figure 4.1: Distribution of IAV subtypes isolated from gulls. The numbers of viruses for which sequence data are available with the indicated subtypes isolated from gulls in North America and Eurasia are shown. The circular representation shows the total proportions when both regions are combined. Gull viruses were located by presence of M gene segment sequences in the NCBI Influenza Virus Resource (http://www.ncbi.nlm.nih.gov/genomes/FLU/FLU.html), as of September 2015.

We performed phylogenetic and time of most recent common ancestor (TMRCA) analyses of H13 and H16 nucleotide sequences. These showed three distinct clades for each subtype that were estimated to originate from a common ancestor in the 1930s for the H13 lineage (Figure 4.2) and in the 1920s for the H16 lineage (Figure 4.3), similar to previous estimates (Worobey, Han et al. 2014). Both trees show some segregation of sequences according to geographic origin as well as clades that contain mixtures of viruses from both Eurasia and North America. These gull lineage HA genes have been transferred to waterfowl, poultry, and marine

mammals but there is no evidence of persistence of these sequences in these other host groups.

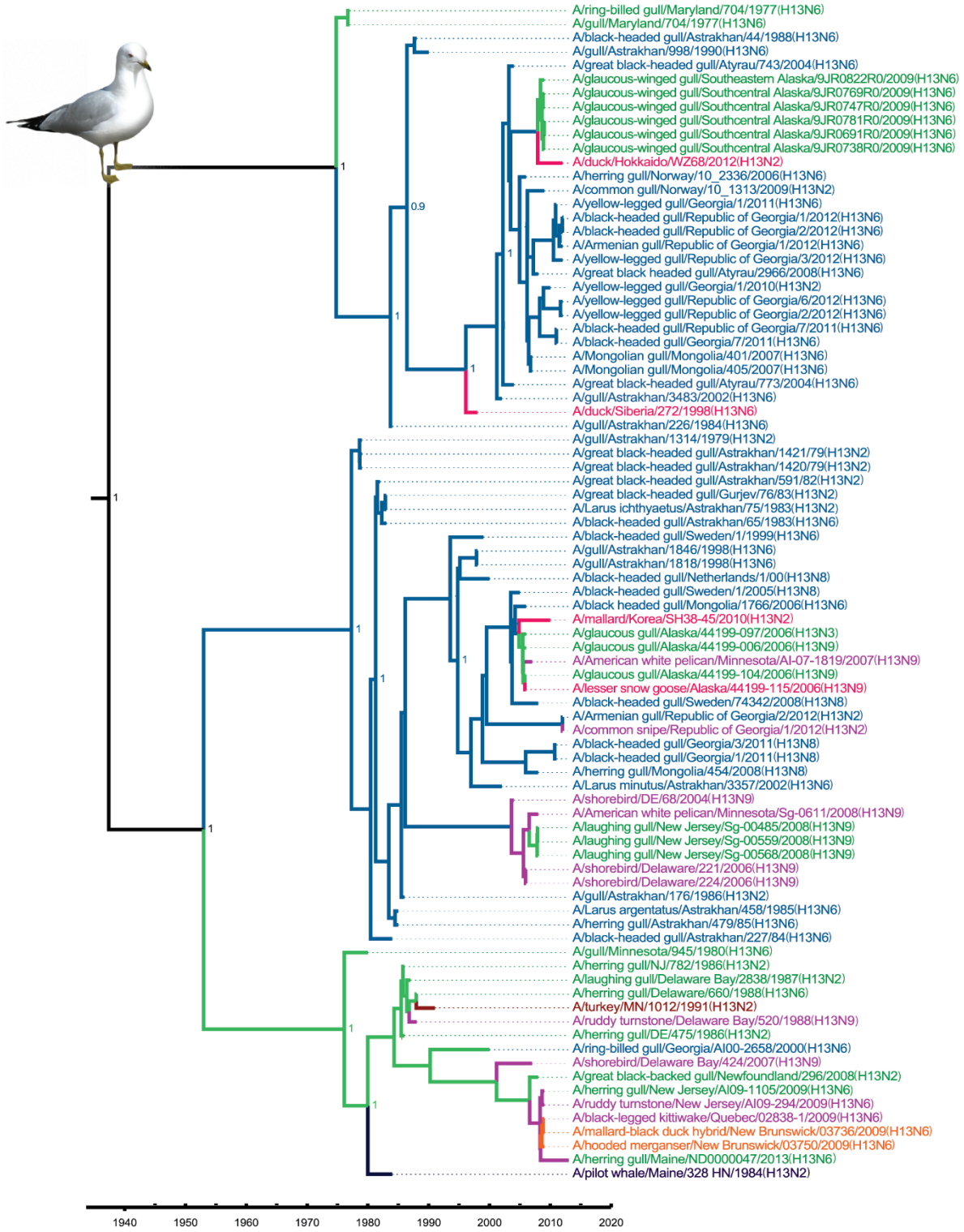


Figure 4.2: Evolutionary history of H13 nucleotide sequences from North America and Eurasia. All H13 nucleotide sequences were downloaded from the NCBI Influenza Virus Resource (Bao, Bolotov et al. 2008) and aligned using MUSCLE integrated in MEGA6 (Tamura, Stecher et al. 2013). The aligned sequences were quality-trimmed, resulting in an alignment of 1133 nts (corresponding to nt positions 465-1597). The maximum credibility tree was generated from the trimmed alignment by the Bayesian inference method implemented in BEAST v1.8.0 (Drummond, Suchard et al. 2012) using the SRD06 substitution model with a strict molecular clock and with the use of the GMRF Bayesian skyride coalescent prior distribution. These parameters were chosen because they gave the best distribution of posterior probabilities. A strict molecular clock was chosen as appropriate to estimate the time of the most recent ancestors at each node because the sequences are almost all from the same host species. The branches are colored according to the host group: green, North American gulls; blue, Eurasian gulls; brown, poultry; dark blue, marine mammals; dark pink, Eurasian waterfowl; orange, North American waterfowl; purple, shorebirds and other wild birds excluding gulls, waterfowls, and poultry. The posterior probabilities for the support of the branches are given at major nodes. Photo credit: http://www.freeimages.com.

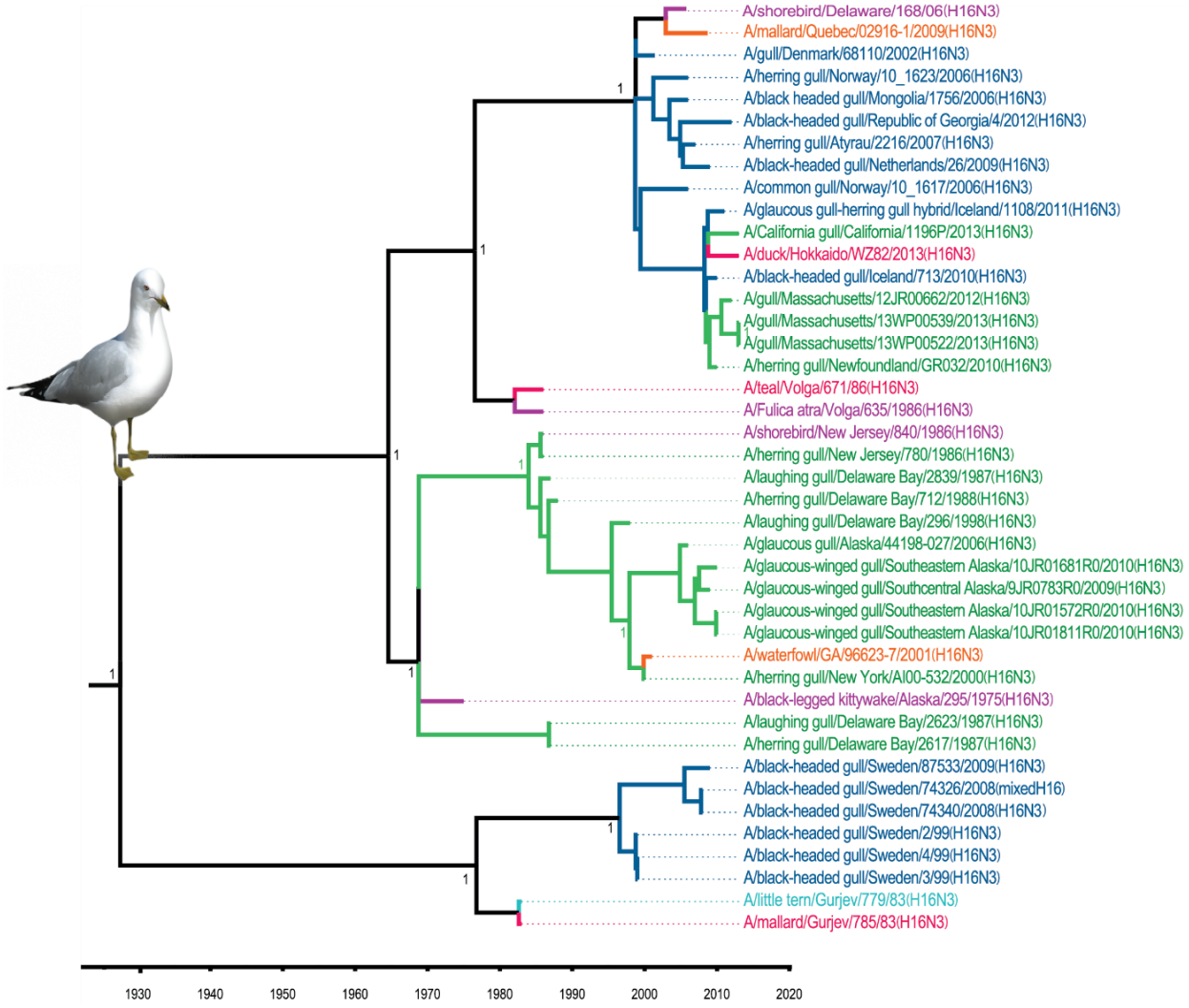

A/shorebird/Delaware/168/06(H16N3)
A/mallard/Quebec/02916-1/2009(H16N3)
A/gull/Denmark/68110/2002(H16N3)
A/herring gull/Norway/10_1623/2006(H16N3)
A/black headed gull/Mongolia/1756/2006(H16N3)
A/black-headed gull/Republic of Georgia/4/2012(H16N3)
A/herring gull/Atyrau/2216/2007(H16N3)
A/black-headed gull/Netherlands/26/2009(H16N3)
A/common gull/Norway/10_1617/2006(H16N3)
A/glaucous gull-herring gull hybrid/Iceland/1108/2011(H16N3)
A/California gull/California/1196P/2013(H16N3)
A/duck/Hokkaido/WZ82/2013(H16N3)
A/black-headed gull/Iceland/713/2010(H16N3)
A/gull/Massachusetts/12JR00662/2012(H16N3)
A/gull/Massachusetts/13WP00539/2013(H16N3)
A/gull/Massachusetts/13WP00522/2013(H16N3)
A/herring gull/Newfoundland/GR032/2010(H16N3)
A/teal/Volga/671/86(H16N3)
A/Fulica atra/Volga/635/1986(H16N3)
A/shorebird/New Jersey/840/1986(H16N3)
A/herring gull/New Jersey/780/1986(H16N3)
A/laughing gull/Delaware Bay/2839/1987(H16N3)
A/herring gull/Delaware Bay/712/1988(H16N3)
A/laughing gull/Delaware Bay/296/1998(H16N3)
A/glaucous gull/Alaska/44198-027/2006(H16N3)
A/glaucous-winged gull/Southeastern Alaska/10JR01681R0/2010(H16N3)
A/glaucous-winged gull/Southcentral Alaska/9JR0783R0/2009(H16N3)
A/glaucous-winged gull/Southeastern Alaska/10JR01572R0/2010(H16N3)
A/glaucous-winged gull/Southeastern Alaska/10JR01811R0/2010(H16N3)
A/waterfowl/GA/96623-7/2001(H16N3)
A/herring gull/New York/AI00-532/2000(H16N3)
A/black-legged kittywake/Alaska/295/1975(H16N3)
A/laughing gull/Delaware Bay/2623/1987(H16N3)
A/herring gull/Delaware Bay/2617/1987(H16N3)
A/black-headed gull/Sweden/87533/2009(H16N3)
A/black-headed gull/Sweden/74326/2008(mixedH16)
A/black-headed gull/Sweden/74340/2008(H16N3)
A/black-headed gull/Sweden/2/99(H16N3)
A/black-headed gull/Sweden/4/99(H16N3)
A/black-headed gull/Sweden/3/99(H16N3)
A/little tern/Gurjev/779/83(H16N3)
A/mallard/Gurjev/785/83(H16N3)
1930
1940
1950
1960
1970
1980
1990
2000
2010
2020

Figure 4.3: Evolutionary history of H16 nucleotide sequences from North America and Eurasia. All complete H16 nucleotide sequences were downloaded from the NCBI Influenza Virus Resource (Bao, Bolotov et al. 2008) and aligned using MUSCLE integrated in MEGA6 (Tamura, Stecher et al. 2013). The aligned sequences were quality-trimmed, resulting in an alignment of 1562 nts (corresponding to nt positions 94-1655). The maximum credibility tree was generated from the trimmed alignment by the Bayesian inference method implemented in BEAST v1.8.0 (Drummond, Suchard et al. 2012) using the SRD06 substitution model with a strict molecular clock and with the use of the GMRF Bayesian skyride coalescent prior distribution. The branches are colored according to the host group: green, North American gulls; blue, Eurasian gulls; brown, poultry; dark pink, Eurasian waterfowl; orange, North American waterfowl; purple, shorebirds and other wild birds excluding gulls and waterfowl. The posterior probabilities for the support of the branches are given at major nodes. Photo credit: http://www.freeimages.com.

Inter-hemispheric movements of avian IAV genes from Eurasia to America have repeatedly been documented (Ip, Flint et al. 2008, Fries, Nolting et al. 2013, Dusek, Hallgrimsson et al. 2014). However, less common observations of IAV movements from America to Eurasia have been observed. Possible explanations for this include a higher rate of infection in birds during movements from Eurasia to America, differences in net avian biomass traveling in the two directions, incompatibility of the Eurasian IAV gene pool with North American genes introduced into Eurasia, or uneven virological monitoring. Regardless, the North American avian

IAV gene pool appears appreciably permissive for Eurasian genes, which are frequently assimilated (Bahl, Vijaykrishna et al. 2009). This is exemplified by the case of the H6 subtype, where it was shown that IAV movement from Eurasia to North America led to the replacement of the North American H6 subtype, continent-wide, within a decade of its introduction (Zu Dohna, Li et al. 2009). However, this is an extreme case, and Eurasian gene introductions do not usually appear to be so dramatic and successful. As mentioned above, there have been only two discoveries of entirely Eurasian IAVs in North America, the H16N3 gull virus and the highly pathogenic H5N8 virus. Examining the evolutionary histories of the entirely Eurasian H16N3 gull virus' genes reveals that the Caspian Sea appears to be a source for the generation of this virus and others. Indeed, genetic analyses illustrate the spatiotemporal distribution of this virus that traces its gene segment origins from Astrakhan, in Russia, in 1983 to Ukraine in 2005, Scandinavia between 2006-2009, then to Iceland and Eastern Canada in 2010 (Figure 4.4 and Supplementary table 4.1).

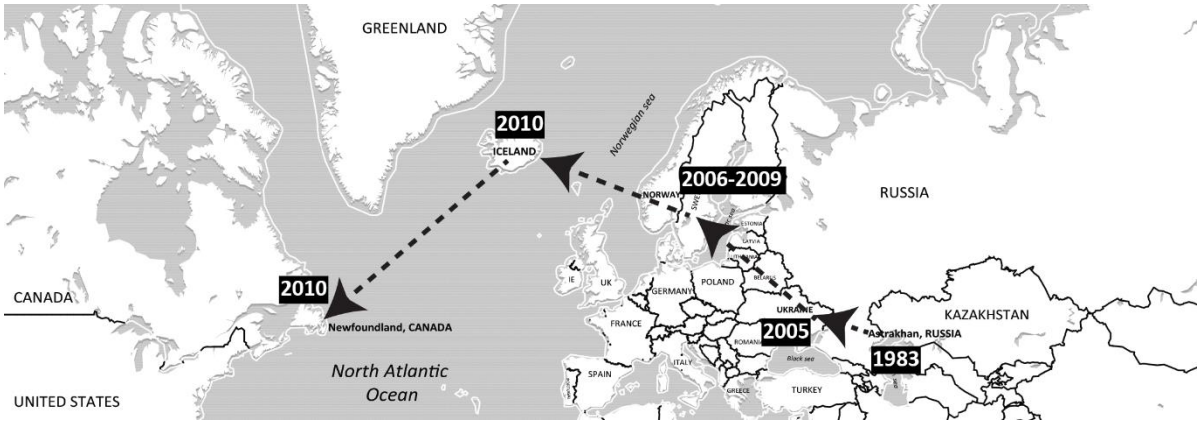


Figure 4.4: Putative geographic progression of gull virus genes that contributed to the genesis of the Eurasian H16N3 gull virus, identified in Newfoundland, Canada. The background map was downloaded from Mapbox Editor online tool (https://www.mapbox.com/editor). The evolutionary history of the virus was retraced based on the relationships of the virus genomic sequences with other viruses by BLAST searches of the NCBI database. The pairwise genetic distances of the virus genomic sequences to other related viruses are provided in Supplementary Table 4.1.

Therefore, it appears that the Eurasian isolate from Eastern Canada has been generated through the gradual accumulation of different genes (through reassortment events) at different locations from 1983 to 2010, although we recognize that there are limitations in the consistency of availability of data through time and from all regions. This was presumably facilitated through the migration of gulls across these regions and the spatiotemporal pattern suggests a predominantly east to west movement and evolutionary trend. While only single gene segments were found closely related to the Eurasian gull virus preceding Iceland, this recent virus from Iceland showed a close relationship to the virus from Eastern Canada for four gene segments (HA, PB2, PA, and NP). It is also notable that an H16 segment similar to the North American virus was detected three years later in a duck in Japan, which points towards an active inter-hemispheric gull-duck interface (Supplementary table 4.1). Globally, the H16N3 North American gull virus is presently the first and only whole IAV of which a trans-Atlantic ecophylogenetic course has been concretely

traced. Its origin and spatiotemporal phylogenetic history are analogous to the apparent natural history of the H14 viruses that were originally isolated from a gull and mallard (*Anas platyrhynchos*) in Astrakhan, and subsequently recovered 28 years later in ducks in America (Fries, Nolting et al. 2013). The H14 subtype was originally isolated in 1982 from both European herring gulls (*Larus argentatus*) and mallards sampled in Central Asia along the northern shore of the Caspian Sea. Subsequent identifications of H14 viruses have occurred solely in North and Central America (USA and Guatemala) during 2010 to 2012 (Boyce, Schobel et al. 2013, Ramey, Poulson et al. 2014). These viruses originated from sea ducks and dabbling ducks, which are commonly associated with marine and freshwater habitats, respectively. This unusual spatiotemporal profile might indicate a role played by gulls in inter-hemispheric and marine-freshwater avifauna IAV transmission. The gap in H14 virus isolation from 1982 to 2010 could be due to a lack of fortuitous sampling during those periods and the same applies for the large gap between the identification of gene segments similar to the North American gull virus H16N3. Regardless, the overall movements of the involved host species are compatible with the viral migration routes. Broadly, such analyses, to retrace the origin and evolution of a virus, might be useful for detailed reconstruction of evolutionary processes in other IAVs.

4.2.2 Ecogeographic aspects of gull IAVs

The global distribution of gulls is vast and nearly unlimited worldwide (Olsen and Larsson 2004). They breed on every continent, from the margins of Antarctica to the high Arctic. Many species breed in coastal colonies and although they are less

common on tropical islands, where terns are often very common, a few species do inhabit islands such as the Galapagos and New Caledonia. There is considerable variety within the group, and species may breed and feed in marine, freshwater and/or terrestrial habitats. Most gull species migrate to warmer habitats during the winter, but the extent to which they migrate varies among species. Some species migrate long distances, such as Franklin's gull (*Leucophaeus pipixcan*) that migrates from Canada to wintering grounds in southern South America. Two examples illustrating inter-hemispheric distribution are black-headed gull (*Chroicocephalus ridibundus*) and laughing gull (*Leucophaeus atricilla*). The black-headed gull mostly breeds in Europe and Asia, and most of the population is migratory, wintering further south, but some birds in the milder westernmost areas of Europe are non-migratory. Some birds also spend the winter in northeastern North America. Additionally, black-headed gulls have an impressive ability to adapt to man-made environments. The laughing gull is largely a species of North and South America although it also occurs as a rare vagrant to Western Europe. Herring gulls exhibit intercontinental distribution and connectivity and there is overlap in the ranges of closely related species/subspecies within a larger species complex that includes American herring gull, European herring gull and East Siberian gull (*Larus vegae*) as members of a larger species complex (Sangster, Collinson et al. 2007). However, it is likely that migratory connectivity exists among these related American, East Eurasian and European populations, and this may serve as a platform for virus movements across both the Pacific and Atlantic Oceans.

4.2.3 Clinical and subclinical IAV infections of gulls

Historically, IAVs have been classified as virulent in gulls and also in terns, in only two natural episodes. These are the first wild bird IAV isolate, A/Tern/South Africa/61(H5N3) (Becker 1966), and the Asian highly pathogenic H5N1 lineage that infects many species of mammals and birds (Claas, Osterhaus et al. 1998, Subbarao, Klimov et al. 1998, Chen, Smith et al. 2005, Amonsin, Songserm et al. 2007, Mushtaq, Juan et al. 2008). However, an H13 virus circulating in a Canadian colony of dense-breeding ring-billed gulls (*Larus delawarensis*) caused inflammation of the heart, kidney, pancreas, and liver in chicks but no clinical symptoms (Velarde, Calvin et al. 2010). Details of experimental infection studies involving gulls or other hosts infected with avian strains or with gull viruses are summarized in Tables 4.1-4.3, and some details are expanded below.

Table 4.1. Experimental infections of gulls with avian strains.

Challenged species	Virus strain	Pathogenicity	Route of inoculation	Inoculum titre	Clinical and pathological findings	Virus shedding	Reference
Laughing gull (*Larus atricilla*)	A/tern/South Africa/61/H5N3	HPAI	Intranasal	10^6 ELD_{50}	Mild lesions	Oropharynx and cloaca	(Perkins and Swayne 2002)
	A/chicken/Hong Kong/220/97/H5N1	HPAI	Intranasal	10^6 ELD_{50}	Mild lesions	Oropharynx and cloaca	
Laughing gull	A/Duck Meat/Anyang/01/H5N1	HPAI	Intranasal	10^6 EID_{50}	Severe disease, death (2/3), histological lesions	Oropharynx > cloaca	(Brown, Stallknecht et al. 2006)
Herring gull (*Larus argentatus*)	A/whooper swan/Mongolia/244/05/H5N1	HPAI	Intranasal	10^6 EID_{50}	Sign of disease, death (2/3)	Oropharynx > cloaca	(Brown, Stallknecht et al. 2008)
	A/duck meat/Anyang/AVL-1/01/H5N1	HPAI	Intranasal	10^6 EID_{50}	Sign of disease, no death		
	Infected chicken meat (A/whooper swan/Mongolia/244/05(H5N1))	HPAI	Ingestion	10^6 EID_{50}	Sign of disease, death (1/3)		
Black-headed gull (*Chroicocephalus ridibundus*)	A/turkey/Turkey 1/2005/H5N1	HPAI	Intratracheal and intraoesophageal	10^4 $TCID_{50}$	Neurological disorder; spontaneous mortality; loss of body weight; no histological lesions	Pharynx and cloaca	(Ramis, van Amerongen et al. 2014)
Ring-billed gull (*Larus delawarensis*)	A/Chicken/Pennsylvania/1370/83/H5N2	HPAI	Nasal and ocular	10^8 EID_{50}	Few or no clinical sign of disease	Upper respiratory tract	(Wood, Webster et al. 1985)
Franklin's gull (*Larus pipixcan*)	A/turkey/Minn/BF/72/Hav6Neq2/H6N8	LPAI	Intratracheal	$3.1x10^3$ EID_{50}	None, 1/5 died with no histological lesions observed	Trachea and cloaca	(Bahl and Pomeroy 1977)

Ring-billed gull	A/laughing gull/NJ/AI08-1460/2008/H13N9	LPAI	Intranasal and intratracheal	10^4 EID_{50}	No sign of disease	Oropharynx and cloaca	(Brown, Poulson et al. 2012)
Silver gull (*Chroicocephalus novaehollandiae*)	A/Eurasian coot/WA/2727/79/H6N2	LPAI	Oropharyngeal	$10^{7.95}$ EID_{50}	None	Trachea and cloaca	(Curran, Robertson et al. 2013)

[a] ELD50 and EID50: amount of infectious virus that will cause the death of 50 % of inoculated embryonated eggs; TCID50: amount of infectious virus that will cause the infection of 50% of tissue culture.

Table 4.2. Experimental infections of avian and mammalian hosts with gull strains.

Challenged species	Virus strain	Pathogenicity	Route of inoculation	Inoculum titre	Clinical and pathological findings	Virus shedding	Reference
White Leghorn chicken	A/black-tailed gull/Tottori/61/80/ Hav1Neq1/H7N7	LPAI	Intratracheal	$10^{7.8}$-$10^{8.3}$ EID_{50}	No sign of disease	Brain, lungs, liver, spleen,	(Otsuki, Kawaoka et al. 1982)
			Intraperitoneal		1/9 died	Kidneys, jejunum, and rectum	
Duck	A/ring-billed gull/Maryland/704/ 77/H13N6	LPAI	Oral and intratracheal	10^{8} EID_{50}	No sign of disease	No tracheal and cloacal shedding	(Hinshaw, Air et al. 1982)
Chicken	A/ring-billed gull/Maryland/704/ 77/H13N6	LPAI	Oral and intratracheal	10^{8} EID_{50}	No sign of disease	No tracheal and cloacal shedding	
Ferret	A/gull/Massachusetts/26/1980/H13N6	LPAI	Intranasal	10^{6} EID_{50}	No sign of disease	Nasal shedding	
Domestic Lohmann white chicken	A/herring gull/Norway/10_1623/2006/H16N3	LPAI	Intranasal	10^{6} EID_{50}	No sign of disease	Oropharynx (2/19) and nasal (1/19)	(Tonnessen, Valheim et al. 2011)
Turkey	A/ laughing gull /NJ/AI08-1460/2008/H13N9	LPAI	Intranasal and intratracheal	10^{7} EID_{50}	No sign of disease	Oropharynx (1/8) and cloaca	(Brown, Poulson et al. 2012)
	A/ring-billed gull/MN/AI10-1708/2010/H13N6	LPAI	Intranasal and intratracheal	10^{7} EID_{50}	No sign of disease	None	
Mallard (*Anas platyrhynchos*)	A/Gull/Ontario/680-6/2001/H13N6	LPAI	Intratracheally and intraesophageally	10^{8} EID_{50}	No sign of disease, some lung tissue lesions observed	Limited pharyngeal, cloacal, and respiratory tissue	(Daoust, van de Bildt et al. 2013)
Mallard	A/herring gull/Germany/R3309/07/H16N3	LPAI	Oculo-nasal-oral	10^{5} $TCID_{50}$	No sign of disease	None	(Fereidouni, Harder et al. 2014)
Ferret	A/gull/Delaware/428/2009(H1N1)	LPAI	Airborne	10^{6} EID_{50}	Lethargy, fever, sneezing, coughing	Nasal	(Kocer, Krauss et al. 2015)

[a] ELD50 and EID50: amount of infectious virus that will cause the death of 50 % of inoculated embryonated eggs; TCID50: amount of infectious virus that will cause the infection of 50% of tissue culture

Table 4.3. Gull virus binding assays on avian and human tissues (Lindskog, Ellström et al. 2013).

Species' tissue tested	Virus strain	Tissues bound
Franklin's gull (*Leucophaeus pipixcan*)	A/black-headed gull/Sweden/2/99/H16N3	Trachea, duodenum, ileum, ileocecal junction, colon
Herring gull (*Larus argentatus*)	A/black-headed gull/Sweden/2/99/H16N3	Duodenum, ileum, ileocecal junction, colon
Mallard (*Anas platyrhynchos*)	A/black-headed gull/Sweden/2/99/H16N3	Trachea
Human	A/black-headed gull/Sweden/2/99/H16N3	Cornea, conjunctiva, nasopharynx, bronchus, pulmonary alveolus, salivary gland

Experimental infection with the A/chicken/Hong Kong/220/97(H5N1) strain produced mild disease in laughing gulls (Perkins and Swayne 2002) whereas severe disease and mortality were reported in laughing gull after infection with highly pathogenic H5N1 strains from 2001 and 2006 (Brown, Stallknecht et al. 2006). Another highly pathogenic H5 strain, A/Chicken/Penn./1370/83(H5N2), which caused 80% mortality in chickens in a Pennsylvania outbreak, produced little to no clinical signs in experimentally infected ring-billed gulls (Wood, Webster et al. 1985). Franklin's gulls experimentally infected with a turkey virus that was pathogenic for turkeys did not develop signs of disease, yet were capable of shedding the virus (Bahl and Pomeroy 1977). These studies suggest that gulls could be infected with virulent strains from poultry and serve as asymptomatic carriers for subsequent virus transmission. In another study, the virus A/laughing gull/NJ/AI08-1460/2008(H13N9) was used to experimentally infect ring-billed gulls by the respiratory route, and high susceptibility to the virus was observed without developing any clinical symptoms.

Most importantly, the virus was excreted via the oropharynx and cloaca for several days (Brown, Poulson et al. 2012). Collectively, *in vivo* experimental infection studies show fairly consistent clinical sensitivity of gulls to different highly pathogenic viruses demonstrating signs of disease, and subclinical sensitivity to low pathogenicity viruses with no observable signs of disease, with consistent evidence of virus shedding.

Studies involving experimental infections of mammals with gull viruses have been limited but several are of particular interest. In one study (Hinshaw, Air et al. 1982), the H13 gull virus A/gull/Massachusetts/26/1980(H13N6) was used in experimental infection of ferrets, which are generally considered as the best animal model for human influenza infection. When challenged intranasally, the ferrets contracted the virus and shed it in nasal secretions. Additionally, attachment of the strain A/black-headed gull/Sweden/2/99(H16N3) to human, gull and mallard tissues was investigated using tissue microarrays and virus histochemistry, using mallard and human viruses as references (Lindskog, Ellström et al. 2013). The H16N3 virus attached more readily to the human respiratory tract than did the other avian IAVs, and it could also bind to the human cornea and conjunctiva. The human virus, A/Netherlands/213/03(H3N2), also attached to the trachea of both gull species investigated, which demonstrates the presence of the first necessary characteristic for respiratory transmission. It was consequently suggested that H16 viruses might present a different cell tropism in humans and mallards than do other bird-origin IAVs (Lindskog, Ellström et al. 2013). Elsewhere it was shown that several gull species display α2,6-linked sialic acid receptors, to which human IAVs usually bind, on the

surface of their tracheal epithelium (Jourdain, van Riel et al. 2011). Those findings indicate the potential for either direct or indirect gull-human IAV interface. Viruses from gulls were also among those identified in wild birds as encoding proteins that differed at relatively few amino acid positions from the 1918 pandemic strain (Watanabe, Zhong et al. 2014), with gull viruses found that had <15 amino acid differences in their PB2, PB1, PA, NP, M1 or NS1 proteins. A virus that contained such segments identified in wild bird viruses showed higher pathogenicity than typical avian viruses and was able to gain mammalian respiratory drop transmission after acquiring as few as three substitutions in the HA protein (Watanabe, Zhong et al. 2014). This supports the notion that the potential to generate a pandemic virus exists within wild birds and, although the sequences used in the experimental study were not from gull viruses, gulls could be a relevant part of this scenario.

4.2.4 Inter-species transmission of gull IAVs

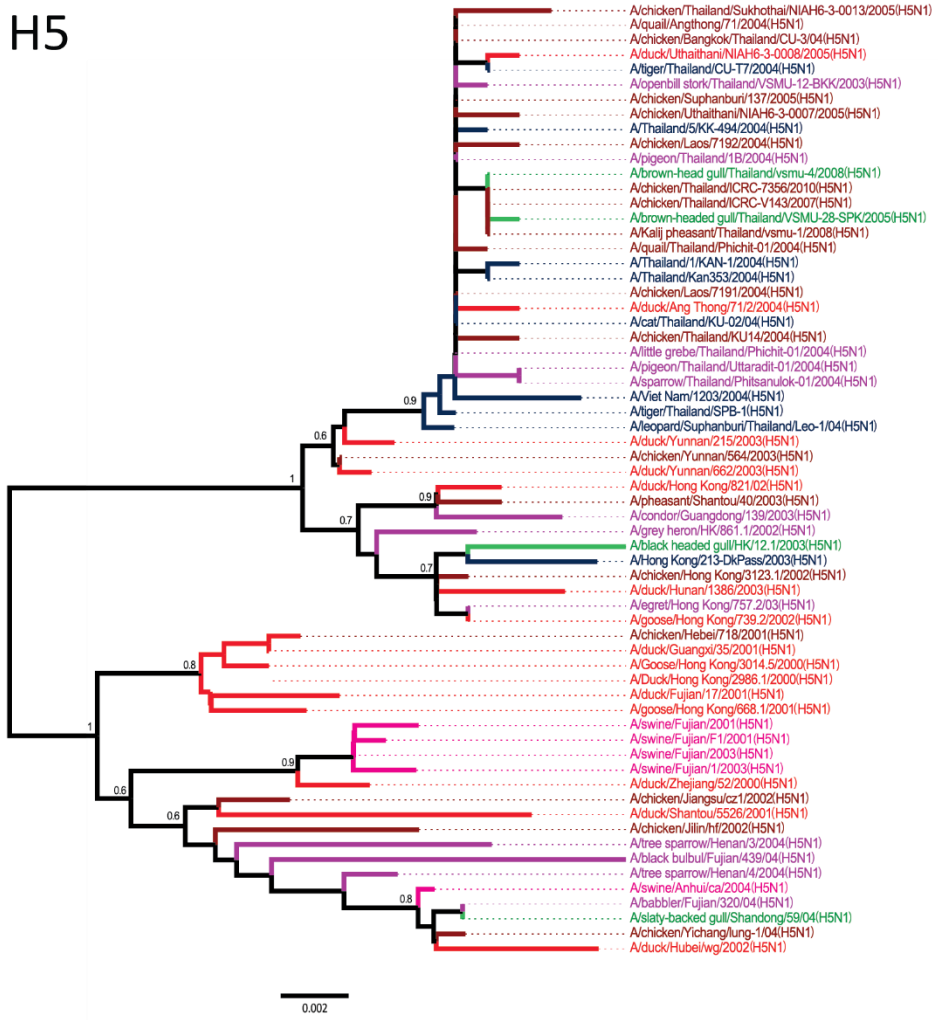
Although the level of active infection found is much lower, seroprevalence for IAV in gulls is high, reaching 45-92% (Brown, Luttrell et al. 2010, Velarde, Calvin et al. 2010, Toennessen, Germundsson et al. 2011, Echeverría, Moratorio et al. 2015). As mentioned above, gulls also exhibit notable capacity in terms of inter-hemispheric virus transmission (Perkins and Swayne 2002, Wille, Robertson et al. 2011, Wille, Robertson et al. 2011, Ratanakorn, Wiratsudakul et al. 2012, Van Borm, Rosseel et al. 2012, Hall, TeSlaa et al. 2013, Echeverría, Moratorio et al. 2015). If gull-imported Eurasian gene segments persist within the American avian gene pool, they could subsequently be transmitted to other domestic hosts such as chickens, turkeys and pigs. In consideration of experimental studies that frequently show sub-clinical

infections of gulls by highly pathogenic avian viruses, these hosts should be regarded as candidates for inter-hemispheric transmission of highly pathogenic virus. The following sections will evaluate whether gulls could play a major role in inter-species and long-range transmission of IAVs and specific IAV genes and whether they could be considered as contributors to the transmission of HPAI viruses.

4.2.4.1 The gull-poultry IAV interface

The first evidence of poultry infection by a gull IAV came from a study in which H13N2 viruses were isolated in North America from subclinically infected domestic turkey flocks living near accessible pond water and gulls (Sivanandan, Halvorson et al. 1991). The H13 sequence of one of the resultant viruses, A/turkey/MN/1012/1991(H13N2), is included within the North American gull AIV lineage (Figure 4.2). These turkey H13N2 viruses did exhibit marked pathogenicity in experimentally infected turkeys (Laudert, Sivanandan et al. 1993). However, various H13 gull viruses used for experimental challenges of various poultry species infrequently resulted in infection and appeared only subclinical (Table 4.2). There have been no reports of H16 viruses infecting poultry. A low pathogenic H7N7 gull strain was widely invasive and caused systemic, though sub-clinical, infection in experimental studies with chickens (Otsuki, Kawaoka et al. 1982). Another low pathogenic gull virus of the H7 subtype, A/Laughing gull/DE/42/06(H7N3), is believed to have been naturally transmitted to chickens, with subsequent evolution to a highly pathogenic form therein (Krauss, Obert et al. 2007). This virus was also fully infectious, although mostly subclinically, in experimentally challenged chickens (Krauss, Obert et al. 2007). Inter-species transmission between gulls and chickens

was also observed with A/brown-head gull/Thailand/vsmu-4/2008(H5N1), being highly similar to two highly pathogenic chicken isolates identified from the same region (Figure 4.5).


H5
A/chicken/Thailand/Sukhothai/NIAH6-3-0013/2005(H5N1)
A/quail/Angthong/71/2004(H5N1)
A/chicken/Bangkok/Thailand/CU-3/04(H5N1)
A/duck/Uthaithani/NIAH6-3-0008/2005(H5N1)
A/tiger/Thailand/CU-T7/2004(H5N1)
A/openbill stork/Thailand/VSMU-12-BKK/2003(H5N1)
A/chicken/Suphanburi/137/2005(H5N1)
A/chicken/Uthaithani/NIAH6-3-0007/2005(H5N1)
A/Thailand/5/KK-494/2004(H5N1)
A/chicken/Laos/7192/2004(H5N1)
A/pigeon/Thailand/1B/2004(H5N1)
A/brown-head gull/Thailand/vsmu-4/2008(H5N1)
A/chicken/Thailand/ICRC-7356/2010(H5N1)
A/chicken/Thailand/ICRC-V143/2007(H5N1)
A/brown-headed gull/Thailand/VSMU-28-SPK/2005(H5N1)
A/Kalij pheasant/Thailand/vsmu-1/2008(H5N1)
A/quail/Thailand/Phichit-01/2004(H5N1)
A/Thailand/1/KAN-1/2004(H5N1)
A/Thailand/Kan353/2004(H5N1)
A/chicken/Laos/7191/2004(H5N1)
A/duck/Ang Thong/71/2/2004(H5N1)
A/cat/Thailand/KU-02/04(H5N1)
A/chicken/Thailand/KU14/2004(H5N1)
A/little grebe/Thailand/Phichit-01/2004(H5N1)
A/pigeon/Thailand/Uttaradit-01/2004(H5N1)
A/sparrow/Thailand/Phitsanulok-01/2004(H5N1)
A/Viet Nam/1203/2004(H5N1)
A/tiger/Thailand/SPB-1(H5N1)
A/leopard/Suphanburi/Thailand/Leo-1/04(H5N1)
A/duck/Yunnan/215/2003(H5N1)
A/chicken/Yunnan/564/2003(H5N1)
A/duck/Yunnan/662/2003(H5N1)
A/duck/Hong Kong/821/02(H5N1)
A/pheasant/Shantou/40/2003(H5N1)
A/condor/Guangdong/139/2003(H5N1)
A/grey heron/HK/861.1/2002(H5N1)
A/black headed gull/HK/12.1/2003(H5N1)
A/Hong Kong/213-DkPass/2003(H5N1)
A/chicken/Hong Kong/3123.1/2002(H5N1)
A/duck/Hunan/1386/2003(H5N1)
A/egret/Hong Kong/757.2/03(H5N1)
A/goose/Hong Kong/739.2/2002(H5N1)
A/chicken/Hebei/718/2001(H5N1)
A/duck/Guangxi/35/2001(H5N1)
A/Goose/Hong Kong/3014.5/2000(H5N1)
A/Duck/Hong Kong/2986.1/2000(H5N1)
A/duck/Fujian/17/2001(H5N1)
A/goose/Hong Kong/668.1/2001(H5N1)
A/swine/Fujian/2001(H5N1)
A/swine/Fujian/F1/2001(H5N1)
A/swine/Fujian/2003(H5N1)
A/swine/Fujian/1/2003(H5N1)
A/duck/Zhejiang/52/2000(H5N1)
A/chicken/Jiangsu/cz1/2002(H5N1)
A/duck/Shantou/5526/2001(H5N1)
A/chicken/Jilin/hf/2002(H5N1)
A/tree sparrow/Henan/3/2004(H5N1)
A/black bulbul/Fujian/439/04(H5N1)
A/tree sparrow/Henan/4/2004(H5N1)
A/swine/Anhui/ca/2004(H5N1)
A/babbler/Fujian/320/04(H5N1)
A/slaty-backed gull/Shandong/59/04(H5N1)
A/chicken/Yichang/lung-1/04(H5N1)
A/duck/Hubei/wg/2002(H5N1)
0.9
0.6
1
0.9
0.7
0.7
0.8
1
0.9
0.6
0.6
0.8
0.002

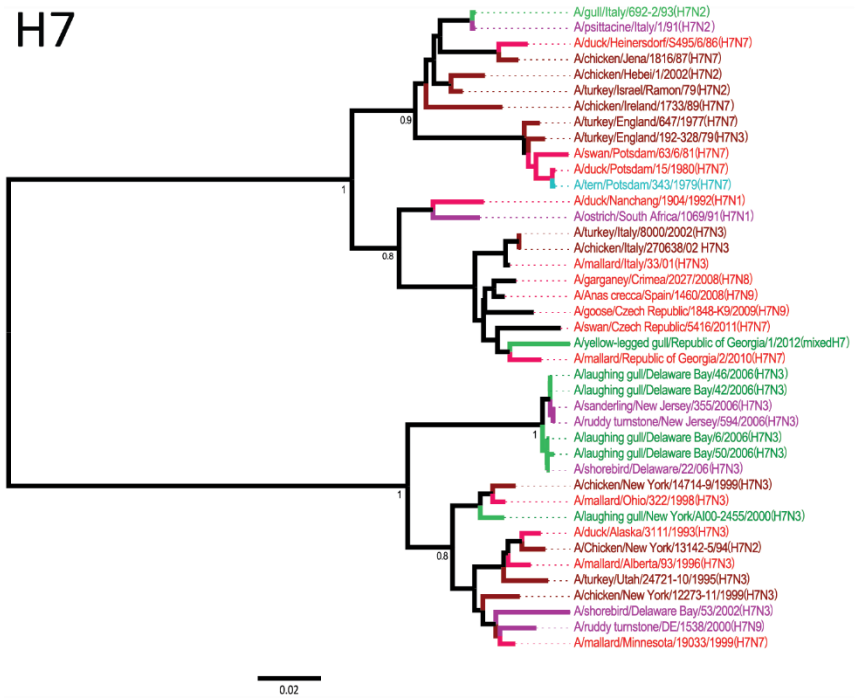

H7
A/gull/Italy/692-2/93(H7N2)
A/psittacine/Italy/1/91(H7N2)
A/duck/Heinersdorf/S495/6/86(H7N7)
A/chicken/Jena/1816/87(H7N7)
A/chicken/Hebei/1/2002(H7N2)
A/turkey/Israel/Ramon/79(H7N2)
A/chicken/Ireland/1733/89(H7N7)
A/turkey/England/647/1977(H7N7)
A/turkey/England/192-328/79(H7N3)
A/swan/Potsdam/63/6/81(H7N7)
A/duck/Potsdam/15/1980(H7N7)
A/tern/Potsdam/343/1979(H7N7)
A/duck/Nanchang/1904/1992(H7N1)
A/ostrich/South Africa/1069/91(H7N1)
A/turkey/Italy/8000/2002(H7N3)
A/chicken/Italy/270638/02 H7N3
A/mallard/Italy/33/01(H7N3)
A/garganey/Crimea/2027/2008(H7N8)
A/Anas crecca/Spain/1460/2008(H7N9)
A/goose/Czech Republic/1848-K9/2009(H7N9)
A/swan/Czech Republic/5416/2011(H7N7)
A/yellow-legged gull/Republic of Georgia/1/2012(mixedH7)
A/mallard/Republic of Georgia/2/2010(H7N7)
A/laughing gull/Delaware Bay/46/2006(H7N3)
A/laughing gull/Delaware Bay/42/2006(H7N3)
A/sanderling/New Jersey/355/2006(H7N3)
A/ruddy turnstone/New Jersey/594/2006(H7N3)
A/laughing gull/Delaware Bay/6/2006(H7N3)
A/laughing gull/Delaware Bay/50/2006(H7N3)
A/shorebird/Delaware/22/06(H7N3)
A/chicken/New York/14714-9/1999(H7N3)
A/mallard/Ohio/322/1998(H7N3)
A/laughing gull/New York/AI00-2455/2000(H7N3)
A/duck/Alaska/3111/1993(H7N3)
A/Chicken/New York/13142-5/94(H7N2)
A/mallard/Alberta/93/1996(H7N3)
A/turkey/Utah/24721-10/1995(H7N3)
A/chicken/New York/12273-11/1999(H7N3)
A/shorebird/Delaware Bay/53/2002(H7N3)
A/ruddy turnstone/DE/1538/2000(H7N9)
A/mallard/Minnesota/19033/1999(H7N7)
0.9
1
0.8
1
1
0.8
0.02

Figure 4.5: Phylogenetic analysis of gull H5 and H7 nucleotide sequences.

Sequences of H5 and H7 genes that showed the highest similarity in BLAST (Altschul, Gish et al. 1990) searches to gull gene sequences were downloaded from NCBI and aligned by MUSCLE with MEGA6 (Tamura, Stecher et al. 2013). The aligned sequences were quality-trimmed, resulting in alignments of 1073 nts (corresponding to nt positions 172-1242) and 1054 nts (corresponding to nt positions 32-1085) for the H5 and H7 gene segments, respectively. The phylogenetic tree was inferred in MEGA6 with the neighbor-joining method with 1000 bootstrap replications. High bootstrap values are shown at the nodes on the trees. Evolutionary distances were computed using the Maximum Composite Likelihood method and are represented by the number of base substitutions per site (scale bar). The branches are colored according to the host group: green, gulls; red, waterfowl from Eurasia; orange, waterfowl from North America; pink, swine; brown, poultry; purple, shorebirds and other wild birds excluding gulls and waterfowl; blue, humans and other mammals.

Further sampling in turkey and chicken flocks, especially where there are potential interactions with gulls, would help confirm whether gulls could play a significant role for IAV transmission to poultry.

4.2.4.2 The gull-waterfowl IAV interface

Sympatry exists between many gull species and waterfowl across freshwater and, secondarily, marine habitats. In addition to the H14 viruses discussed above, particularly high gene sequence identity can be found between various gull and duck viruses (Figure 4.4 and 4.5, Supplementary table 4.2). These include both wild

(mostly mallard) and domestic duck viruses in America and Eurasia. These diverse affinities substantiate an ongoing, reciprocal gull-waterfowl interface, worldwide, which spans both natural ecosystems and farms containing domestic ducks. Additionally, one study showed that H16N3 gull viruses isolated in Russia readily replicated in respiratory and intestinal tissues in challenged ducks (Iamnikova, Gambarian et al. 2009).

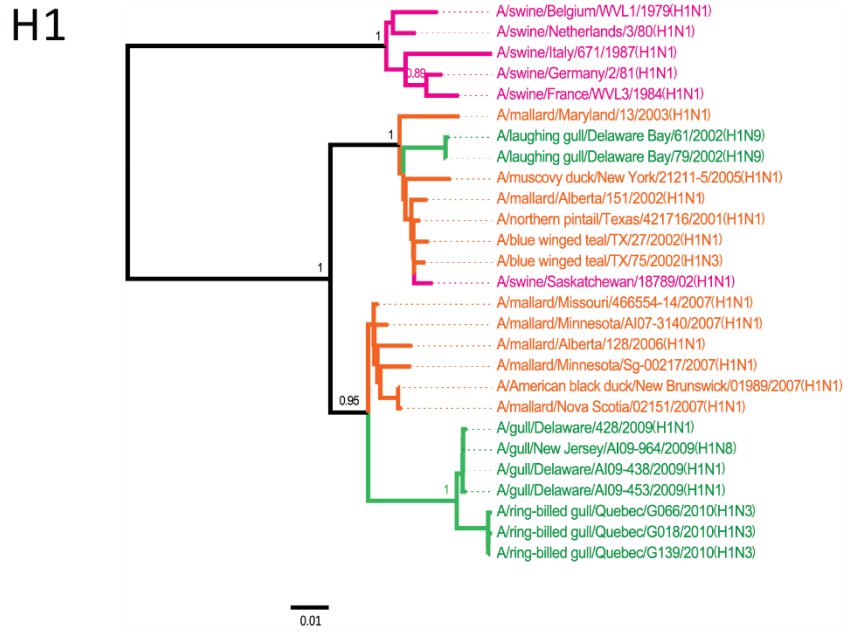
4.2.4.3 The gull-marine mammal IAV interface

The movement of IAV between gulls and mammals was identified for the first time in 1981, when it was shown that all genomic segments from a seal H7N7 virus were highly similar to viral gene segments from avian species, including gulls (Webster, Hinshaw et al. 1981). Interestingly, the seal isolate replicated poorly in avian species but well in mammals such as ferrets, pigs, and cats, which indicated this isolate had probably accumulated mammalian-adaptive mutations. Phylogenetic analysis of some seal and whale viruses has shown that these are frequently closely related to gull and shorebird viruses (Groth, Lange et al. 2014). More specifically, relatedness between NP gene segments of the strains A/gull/Maryland/704/77(H13N6) and A/whale/Maine/328/84(H13N2) from the Atlantic Coast of the USA was found (Mandler, Gorman et al. 1990). Similarly, the HA gene segment from the same whale isolate falls within a North American gull H13 lineage, and presumably originated from gulls (Figure 4.2, Supplementary table 4.2). High sequence identity was also found between the H5 gene segments from a gull isolate, A/black_headed_gull/HK/12.1/2003(H5N1), and a human isolate,

A/Hong_Kong/213/2003(H5N1), both obtained in Asia (Figure 4.5, Supplementary table 4.2).

4.2.4.4 The gull-swine IAV interface

Swine are considered one of the most important mixing vessels of IAV, as genetic reassortment between avian and mammalian viruses occurs in this host. Many of the viruses found in swine also contain genes that originated from human and bird viruses. H1N1, H1N2, and H3N2 viruses are endemic in the swine population worldwide (Brown 2000), which also represent the subtypes circulating in the human population and the presence of those viruses or their genes within gulls might be important in terms of transmission of viruses across distances and host species. To date, several H1 viruses have been found in the gull population in North America and they all fall within a North American lineage that mainly comprises waterfowl viruses. Although no direct evidence of AIV transmission between gulls and swine has been found to date, we identified gull viruses that do exhibit phylogenetic affinities to swine viruses, in conjunction with related waterfowl and poultry viruses. One swine virus from 2002, A/swine/Saskatchewan/18789/02(H1N1), which was closely related to waterfowl viruses (Karasin, West et al. 2004), was also close to two gull viruses obtained in 2002 from the United States for the HA gene segment (Figure 4.6).


H1
A/swine/Belgium/WVL1/1979(H1N1)
A/swine/Netherlands/3/80(H1N1)
A/swine/Italy/671/1987(H1N1)
A/swine/Germany/2/81(H1N1)
A/swine/France/WVL3/1984(H1N1)
A/mallard/Maryland/13/2003(H1N1)
A/laughing gull/Delaware Bay/61/2002(H1N9)
A/laughing gull/Delaware Bay/79/2002(H1N9)
A/muscovy duck/New York/21211-5/2005(H1N1)
A/mallard/Alberta/151/2002(H1N1)
A/northern pintail/Texas/421716/2001(H1N1)
A/blue winged teal/TX/27/2002(H1N1)
A/blue winged teal/TX/75/2002(H1N3)
A/swine/Saskatchewan/18789/02(H1N1)
A/mallard/Missouri/466554-14/2007(H1N1)
A/mallard/Minnesota/AI07-3140/2007(H1N1)
A/mallard/Alberta/128/2006(H1N1)
A/mallard/Minnesota/Sg-00217/2007(H1N1)
A/American black duck/New Brunswick/01989/2007(H1N1)
A/mallard/Nova Scotia/02151/2007(H1N1)
A/gull/Delaware/428/2009(H1N1)
A/gull/New Jersey/AI09-964/2009(H1N8)
A/gull/Delaware/AI09-438/2009(H1N1)
A/gull/Delaware/AI09-453/2009(H1N1)
A/ring-billed gull/Quebec/G066/2010(H1N3)
A/ring-billed gull/Quebec/G018/2010(H1N3)
A/ring-billed gull/Quebec/G139/2010(H1N3)
1
1
0.89
1
1
0.95
1
0.01

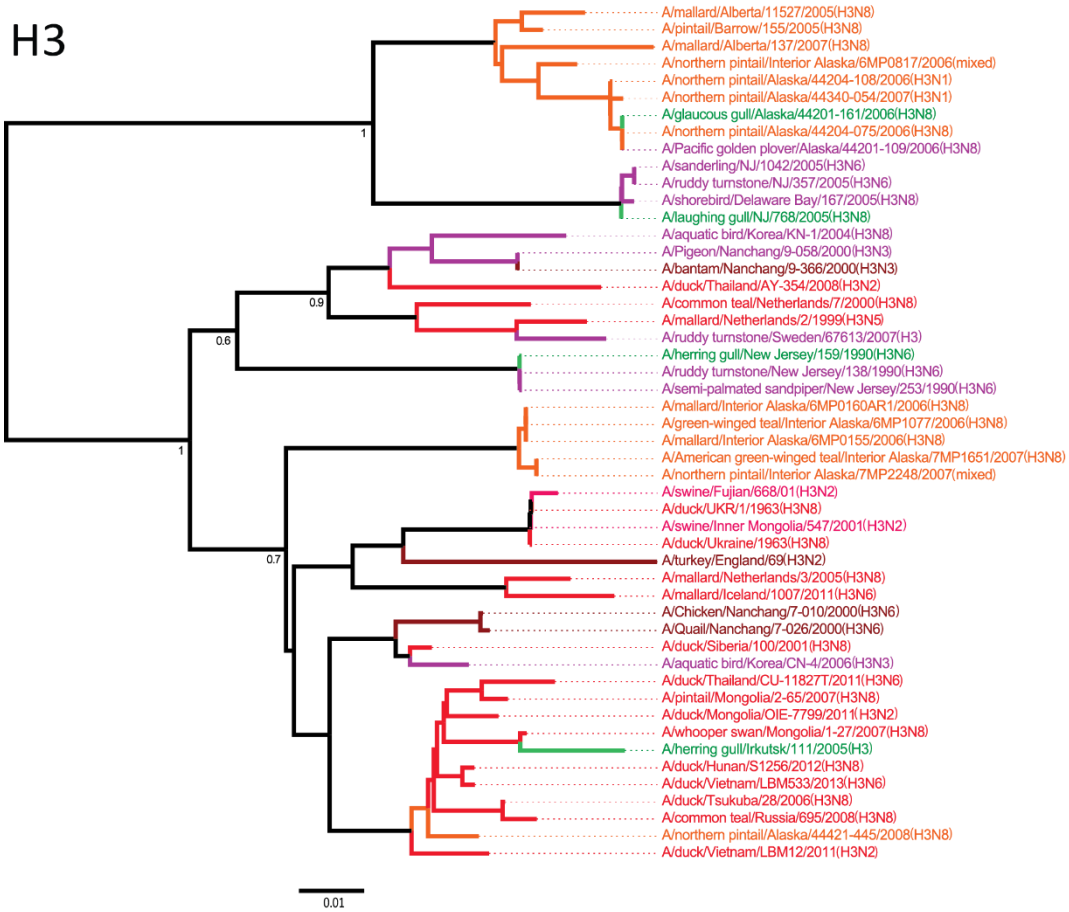

H3
A/mallard/Alberta/11527/2005(H3N8)
A/pintail/Barrow/155/2005(H3N8)
A/mallard/Alberta/137/2007(H3N8)
A/northern pintail/Interior Alaska/6MP0817/2006(mixed)
A/northern pintail/Alaska/44204-108/2006(H3N1)
A/northern pintail/Alaska/44340-054/2007(H3N1)
A/glaucous gull/Alaska/44201-161/2006(H3N8)
A/northern pintail/Alaska/44204-075/2006(H3N8)
A/Pacific golden plover/Alaska/44201-109/2006(H3N8)
A/sanderling/NJ/1042/2005(H3N6)
A/ruddy turnstone/NJ/357/2005(H3N6)
A/shorebird/Delaware Bay/167/2005(H3N8)
A/laughing gull/NJ/768/2005(H3N8)
A/aquatic bird/Korea/KN-1/2004(H3N8)
A/Pigeon/Nanchang/9-058/2000(H3N3)
A/bantam/Nanchang/9-366/2000(H3N3)
A/duck/Thailand/AY-354/2008(H3N2)
A/common teal/Netherlands/7/2000(H3N8)
A/mallard/Netherlands/2/1999(H3N5)
A/ruddy turnstone/Sweden/67613/2007(H3)
A/herring gull/New Jersey/159/1990(H3N6)
A/ruddy turnstone/New Jersey/138/1990(H3N6)
A/semi-palmated sandpiper/New Jersey/253/1990(H3N6)
A/mallard/Interior Alaska/6MP0160AR1/2006(H3N8)
A/green-winged teal/Interior Alaska/6MP1077/2006(H3N8)
A/mallard/Interior Alaska/6MP0155/2006(H3N8)
A/American green-winged teal/Interior Alaska/7MP1651/2007(H3N8)
A/northern pintail/Interior Alaska/7MP2248/2007(mixed)
A/swine/Fujian/668/01(H3N2)
A/duck/UKR/1/1963(H3N8)
A/swine/Inner Mongolia/547/2001(H3N2)
A/duck/Ukraine/1963(H3N8)
A/turkey/England/69(H3N2)
A/mallard/Netherlands/3/2005(H3N8)
A/mallard/Iceland/1007/2011(H3N6)
A/Chicken/Nanchang/7-010/2000(H3N6)
A/Quail/Nanchang/7-026/2000(H3N6)
A/duck/Siberia/100/2001(H3N8)
A/aquatic bird/Korea/CN-4/2006(H3N3)
A/duck/Thailand/CU-11827T/2011(H3N6)
A/pintail/Mongolia/2-65/2007(H3N8)
A/duck/Mongolia/OIE-7799/2011(H3N2)
A/whooper swan/Mongolia/1-27/2007(H3N8)
A/herring gull/Irkutsk/111/2005(H3)
A/duck/Hunan/S1256/2012(H3N8)
A/duck/Vietnam/LBM533/2013(H3N6)
A/duck/Tsukuba/28/2006(H3N8)
A/common teal/Russia/695/2008(H3N8)
A/northern pintail/Alaska/44421-445/2008(H3N8)
A/duck/Vietnam/LBM12/2011(H3N2)
1
0.9
0.6
1
0.7
0.01

Figure 4.6: Phylogenetic analysis of gull H1 and H3 nucleotide sequences. Sequences of H1 and H3 genes that showed the highest similarity in BLAST (Altschul, Gish et al. 1990) searches to gull gene segments were downloaded from NCBI and aligned by MUSCLE with MEGA6 (Tamura, Stecher et al. 2013). The aligned sequences were quality-trimmed, resulting in alignments of 1028 nts (corresponding to nt positions 10-1037) and 641 nts (corresponding to nt positions 1049-1689) for the H1 and H3 gene segments, respectively. The phylogenetic tree was inferred in MEGA6 with the neighbor-joining method with 1000 bootstrap replications. High bootstrap values are shown at the nodes on the trees. Evolutionary distances were computed using the Maximum Composite Likelihood method and are represented by the number of base substitutions per site (scale bar). The branches are colored according to the host group: gulls, green; red, waterfowl from Eurasia; orange, waterfowl from North America; swine, pink; brown, poultry; purple, shorebirds and other wild birds excluding gulls and waterfowl.

The H3N2 viruses currently found in swine are believed to have originated mostly from reassortment events involving avian and swine viruses or avian, swine and human viruses (Brown 2000) and the triple reassortant virus has now become predominant in the swine population (Richt, Lager et al. 2003). Sequence analysis of H3 genes from gull viruses showed that two viruses from America and Asia, A/herring gull/New Jersey/159/1990(H3N6) and A/herring gull/Irkutsk/111/2005(H3), which belong to a Eurasian avian lineage, were related to swine and poultry viruses from England and China (Figure 4.6). Additionally, H3 gene

segments from shorebirds related to A/herring gull/New Jersey/159/1990(H3N6) and found in North America are falling within a Eurasian lineage, which further illustrate the potential role of gulls in inter-hemispheric movement of IAVs (Bahl, Krauss et al. 2013).

Investigating the evolutionary history and origin of some gull viruses has shown that they exhibit close relationships to viruses from other hosts such as poultry, swine and humans. These analyses also show that pathogenic strains are closely related to some gull viruses, which shows their potential for being carriers and probably disseminators of such strains.

4.2.5 The role of gulls as contributors to the transmission of HPAI viruses

Transmission of IAV from wild birds to poultry and swine is a major concern for the agriculture industry and must also be considered with respect to the possible generation of novel pandemic viruses. H5 and H7 subtypes have been commonly associated with high mortality in poultry (Stech and Mettenleiter 2013). Analysis of H5 sequences from gull viruses showed that most of them are closely related to highly pathogenic H5N1 virus isolated from waterfowl, poultry, and swine (Figure 4.5). Specifically, the virus A/slaty-backed gull/Shandong/59/04(H5N1) is identical to the avian virus A/babbler/Fujian/320/04(H5N1) and groups closely with a swine strain, A/swine/Anhui/ca/2004(H5N1), with all these isolates from 2004. This shows that there has been frequent inter-species transmission of this virus between avian species, including gulls, and swine. Similarly, phylogenetic analysis of H7 sequences

from gulls showed that gull viruses from both North America and Eurasia are closely related to viruses associated with lethal H7 outbreaks in poultry (Figure 4.5).

The movements of viruses amongst gulls, waterfowl, poultry, and mammals indicate that gulls need to be considered as important contributors to IAV evolutionary dynamics and ecology. It is also important to recognize the potential contribution of gulls to the generation of pandemic viruses, which always carry gene segments of bird origin, and could contribute to the emergence of pathogenic strains.

4.2.6 The role of terns in global IAV transmission, and relationships with gull IAVs

Terns are the most closely related taxon to gulls; they exhibit shared ecological features and can also be sympatric with gulls. It is also noteworthy that the first isolate of IAV from wild birds was from terns, with A/tern/South Africa/1961(H5N3) representing a highly pathogenic H5 virus that caused many deaths among terns in South Africa (Becker 1966). Close relationships of viruses from terns to those from other hosts can also be found. A low pathogenicity tern isolate, A/tern/Potsdam/342-6/1979(H7N7), is similar to a virus that infected chickens in Germany and subsequently accrued several mutations to become a highly pathogenic virus that caused a severe outbreak in chickens (Rohm, Suss et al. 1996). A connection from terns to marine mammals has also been previously documented with the NP gene segment of the tern isolate A/tern/Turkmenistan/18/1972(H3N3) closely related to the whale isolate A/whale/Pacific ocean/19/1976(H1N3) (Mandler, Gorman et al. 1990).

Phylogenetic analyses with tern virus sequences showed high relatedness of these viruses to viruses identified in gulls, ducks, turkeys and whales from around the world (Supplementary figure 4.1 and 4.2, Supplementary table 4.3). Broadly, these sequence comparisons reveal a series of IAVs isolated from two partially sympatric and highly migratory tern species, the common tern (*Sterna hirundo*) and the Arctic tern (*Sterna paradisaea*). In particular, the common tern isolate from South Africa showed inter-hemispheric reassortment, as well as affinities to isolates from turkey, seal, whale and another common tern strain (Supplementary figure 4.2). This isolate from Turkmenistan is one of a group of common tern viruses from a region of Central Asia that show affinities to North American gull isolates, a European mallard isolate, a tufted duck (*Aythya fuligula*) isolate from southern Siberia and some duck isolates from Taiwan. The phylogenetic analysis of M gene segments also showed significant homology between an Arctic tern virus (A/arctic tern/Alaska/300/1975(H5N3)) and gull and turkey viruses from North America (Supplementary figure 4.1). Finally, the three polymerase gene segments of the same A/arctic tern/Alaska/300/1975(H5N3) isolate also exhibit affinities to these same North American gull isolates, and the PB2 gene segment is also highly related to gull isolates from Astrakhan, Asia (Supplementary figure 4.1). Overall, these analyses show that terns also constitute an important IAV host. They share many characteristics and display migratory connectivity with gulls, and they should also be considered as contributors to the IAV dynamics, diversity and transmission to different hosts.

4.2.7 Ecobiological aspects of gull IAVs

Two aspects of the ecobiology of gulls are particularly relevant for their role in IAV dynamics. These concern their utilization of both marine and freshwater habitats and their potential role in zoonotic transmission pathways. Gulls are generalist feeders and eat both live-captured prey and opportunistically scavenge, including on carcasses of birds and mammals. Presumably, one means by which gulls could potentially contract IAVs from other species in aquatic habitats is by eating IAV-infected dead birds or mammals. The utilization of both marine and freshwater habitats, and interactions with different species in both environments, raises the possibility for gulls to be exposed to viruses from both freshwater and marine sources. The ecology of IAVs is fairly well understood in freshwater habitats, but this has been explored less in marine environments. However, alongside gulls, additional marine birds such as murres and puffins (members of the family Alcidae) and sea ducks (family Anatidae, tribe Mergini) are known to host IAVs and be sympatric with gulls (Ip, Flint et al. 2008). Although some IAVs do not appear very stable in elevated salinity (Stallknecht, Kearney et al. 1990), marine birds drink seawater and it must be presumed that effective fecal-oral transmission can take place in seawater. Viral circulation within marine environments also involves seals and whales, which are known to host avian-derived IAVs and apparently mostly those originating from gulls and terns as our phylogenetic and phylogeographic findings concerning both gull and tern viruses indicate relationships with viruses from marine mammals globally.

There is also an important zoonotic dimension for IAV circulation in gulls. While the important role played by waterfowl as facilitators of the emergence and

dispersal of epizootic and pandemic strains is soundly established, gulls are both cosmopolitan migrants and salient synanthropic birds adjacent to the duck-poultry-swine-human IAV interface. It has previously been articulated that there is a crucial epidemiological position occupied by gulls among wildlife, domestic birds as well as humans (Arnal, Vittecoq et al. 2015), and the potential for gulls to act in transmission of microorganisms from sewage plants or outflows to domestic animals has also been described (Teale 2002). Natural interfaces between gull viruses and those from ducks, turkeys, chickens and swine have been identified in the present analysis, either in terms of interchangeability of certain gene segments, or whole viruses, and IAV-related interactions of gulls with ducks, turkeys, chickens and mammals, phylogenetically and ecologically, might contribute to emergence and spread of epizootic and pandemic strains. It is also of note that a connection was found between marine mammal and human infections (Lvov, Zhdanov et al. 1983, Ohishi, Kishida et al. 2004, Belser, Bridges et al. 2009), and that the 2009 H1N1 human pandemic strain naturally infected and spread, sub-clinically, among seals in California (Goldstein, Mena et al. 2013). However, it is not known if gulls or terns were directly involved in transmission of this virus.

4.3 Concluding remarks

Within the enormous spectrum of IAVs hosted by avian species, gull IAVs constitute an important component, somewhat parallel to waterfowl viruses, of the global ecophylogenetic system defining this protean virus. Gulls affect, both directly and indirectly, the evolutionary dynamics of low and high pathogenicity avian viruses

and various mammalian IAVs worldwide. Analysis of the evolutionary history of an entirely Eurasian virus isolated in North America allowed the reconstruction of the gradual evolutionary process that generated this virus, starting in Central Asia, moving northwest, and ending in Iceland, with subsequent transmission to Canada. The origin of this virus appears to be associated with a region of Central Asia that also appears to have given rise to an important H14 strain that involved both gull and duck viruses. Such reverse ecophylogenetic analysis is useful for attaining a broadened perspective of the evolutionary history of a given virus and can allow identification of the presumed strains, genes, and hosts involved in the evolutionary and spatiotemporal pathways that gave rise to a strain in question. Additional gull viruses that were identified as phylogenetically interesting, in terms of their relationships to pathogenic strains and viruses from other hosts, in this analysis were also from the same Central Asian geographic region. A similar scenario was found when analyzing a series of common tern viruses that were obtained in the same general region of Central Asia, where those viruses exhibited clear affinities to gull, duck and whale viruses. Our analyses of the tern viruses reveal a broad phylogeographic contribution for these viruses beyond Central Asia, with gene transactions extending to Europe, South Africa, North America and the Pacific Ocean.

The traits contributing to the importance of gulls as IAV hosts include their cosmopolitan distributions, migratory behaviors, synanthropic associations with humans, use of both marine and freshwater habitats, prevalence as a host of many different IAVs of a large variety of host group origins (including farm animals), and predominance of subclinical infections. As shown in the present study, the summary

of these properties defines an important IAV host group that needs to be considered in the context of the perpetuation and emergence of enzootic, epizootic and pandemic strain. A routine surveillance of the virus in migratory birds such as gulls and terns could be conducted worldwide to foresee potential zoonotic transmission of pathogenic strains to humans and domestic livestock.

4.4 References

Adams, S., Z. Xing, J. Li, K. Mendoza, D. Perez, K. Reed and C. Cardona (2013). "The effect of avian influenza virus NS1 allele on virus replication and innate gene expression in avian cells." Mol Immunol 56(4): 358-368.

Altschul, S. F., W. Gish, W. Miller, E. W. Myers and D. J. Lipman (1990). "Basic local alignment search tool." Journal of Molecular Biology. 215(3): 403-410.

Altschul, S. F., W. Gish, W. Miller, E. W. Myers and D. J. Lipman (1990). "Basic local alignment search tool." Journal of Molecular Biology 215: 403-410.

Amonsin, A., T. Songserm, S. Chutinimitkul, R. Jam-On, N. Sae-Heng, N. Pariyothorn, S. Payungporn, A. Theamboonlers and Y. Poovorawan (2007). "Genetic analysis of influenza A virus (H5N1) derived from domestic cat and dog in Thailand." Archives of Virology 152: 1925-1933.

Andino, R. and E. Domingo (2015). "Viral quasispecies." Virology 479-480: 46-51.

Arnal, A., M. Vittecoq, J. Pearce-Duvet, M. Gauthier-Clerc, T. Boulinier and E. Jourdain (2015). "Laridae: A neglected reservoir that could play a major role in avian influenza virus epidemiological dynamics." Critical Reviews in Microbiology 41: 508-519.

Asghar, N., J. H. Pettersson, P. Dinnetz, A. Andreassen and M. Johansson (2017). "Deep sequencing analysis of tick-borne encephalitis virus from questing ticks at natural foci reveals similarities between quasispecies pools of the virus." J Gen Virol 98(3): 413-421.

Bahl, A. K. and B. S. Pomeroy (1977). "Experimental exposure of Franklin's gulls (*Larvus pipixcan*) and mallards (*Anas platyrhynchos*) to a turkey influenza A virus A/Turkey/Minn/BF/72(Hv6Neq2)." Journal of Wildlife Diseases 13: 420-426.

Bahl, J., S. Krauss, D. Kuhnert, M. Fourment, G. Raven, S. P. Pryor, L. J. Niles, A. Danner, D. Walker, I. H. Mendenhall, Y. C. Su, V. G. Dugan, R. A. Halpin, T. B. Stockwell, R. J. Webby, D. E. Wentworth, A. J. Drummond, G. J. Smith and R. G. Webster (2013). "Influenza a virus migration and persistence in North American wild birds." PLoS Pathogens 9: e1003570.

Bahl, J., D. Vijaykrishna, E. C. Holmes, G. J. Smith and Y. Guan (2009). "Gene flow and competitive exclusion of avian influenza A virus in natural reservoir hosts." Virology 390(2): 289-297.

Baltimore, D. (1971). "Expression of animal virus genomes." Bacteriological Reviews 35(3): 235-241.

Bao, Y., P. Bolotov, D. Dernovoy, B. Kiryutin, L. Zaslavsky, T. Tatusova, J. Ostell and D. Lipman (2008). "The influenza Virus Resource at the National Center for Biotechnology Information." Journal of Virology 82: 596-601.

Baskerville, A. and G. Lloyd (1976). "The pathogenesis and pathology of experimental Quaranfil virus infection." British journal of experimental pathology 57(2): 152-156.

Becker, W. B. (1966). "The isolation and classification of Tern virus: influenza A-Tern South Africa--1961." Journal of Hygiene 64: 309-320.

Bedford, T., M. A. Suchard, P. Lemey, G. Dudas, V. Gregory, A. J. Hay, J. W. McCauley, C. A. Russell, D. J. Smith and A. Rambaut (2014). "Integrating influenza antigenic dynamics with molecular evolution." Elife 3: e01914.

Belser, J. A., C. B. Bridges, J. M. Katz and T. M. Tumpey (2009). "Past, present, and possible future human infection with influenza virus A subtype H7." Emerging Infectious Diseases 15: 859-865.

Bergmann, M. and T. Muster (1996). "Mutations in the nonconserved noncoding sequences of the influenza A virus segments affect viral vRNA formation." Virus Research 44(1): 23-31.

Blakemore F, G. A. (1941). "Swine Influenza in the British Isles." Proceedings of the Royal Society of Medicine 34(9): 611-618.

Boni, M. F., Y. Zhou, J. K. Taubenberger and E. C. Holmes (2008). "Homologous recombination is very rare or absent in human influenza A virus." Journal of Virology 82(10): 4807-4811.

Boulo, S., H. Akarsu, R. W. Ruigrok and F. Baudin (2007). "Nuclear traffic of influenza virus proteins and ribonucleoprotein complexes." Virus Res 124(1-2): 12-21.

Boyce, W. M., S. Schobel, V. G. Dugan, R. A. Halpin, X. Lin, D. E. Wentworth, L. L. Lindsay, E. Mertens and M. Plancarte (2013). "Complete genome sequence of a reassortant H14N2 avian influenza virus from California." Genome Announcements 1: e00543-00513.

Brown, I. H. (2000). "The epidemiology and evolution of influenza viruses in pigs." Veterinary Microbiology 74: 29-46.

Brown, I. H. (2001). "The pig as an intermediate host for influenza A viruses between birds and humans." International Congress Series 1219: 173-178.

Brown, I. H., P. A. Harris, J. W. McCauley and D. J. Alexander (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.

Brown, J., R. Poulson, D. Carter, C. Lebarbenchon, M. Pantin-Jackwood, E. Spackman, E. Shepherd, M. Killian and D. Stallknecht (2012). "Susceptibility of avian species to

North American H13 low pathogenic avian influenza viruses." Avian Diseases 56: 969-975.

Brown, J. D., M. P. Luttrell, R. D. Berghaus, W. Kistler, S. P. Keeler, A. Howey, B. Wilcox, J. Hall, L. Niles, A. Dey, G. Knutsen, K. Fritz and D. E. Stallknecht (2010). "Prevalence of antibodies to type a influenza virus in wild avian species using two serologic assays." Journal of Wildlife Diseases 46: 896-911.

Brown, J. D., D. E. Stallknecht, J. R. Beck, D. L. Suarez and D. E. Swayne (2006). "Susceptibility of North American ducks and gulls to H5N1 highly pathogenic avian influenza viruses." Emerging Infectious Diseases 12: 1663-1670.

Burleigh, L. M., L. J. Calder, J. J. Skehel and D. A. Steinhauer (2005). "Influenza a viruses with mutations in the m1 helix six domain display a wide variety of morphological phenotypes." Journal of Virology. 79(2): 1262-1270.

Capua, I. and S. Marangon (2000). "The avian influenza epidemic in Italy, 1999-2000: a review." Avian Pathology 29(4): 289-294.

Chen, H., G. J. Smith, S. Y. Zhang, K. Qin, J. Wang, K. S. Li, R. G. Webster, J. S. Peiris and Y. Guan (2005). "Avian flu: H5N1 virus outbreak in migratory waterfowl." Nature 436: 191-192.

Chen, J., K. H. Lee, D. A. Steinhauer, D. J. Stevens, J. J. Skehel and D. C. Wiley (1998). "Structure of the hemagglutinin precursor cleavage site, a determinant of influenza pathogenicity and the origin of the labile conformation." Cell 95(3): 409-417.

Claas, E. C., A. D. Osterhaus, R. van Beek, J. C. De Jong, G. F. Rimmelzwaan, D. A. Senne, S. Krauss, K. F. Shortridge and R. G. Webster (1998). "Human influenza A H5N1 virus related to a highly pathogenic avian influenza virus." Lancet 351: 472-477.

Clifford, C. M., L. A. Thomas, L. E. Hughes, G. M. Kohls and C. B. Philip (1968). "Identification and comparison of two viruses isolated from ticks of the genus Ornithodoros." The American Journal of Tropical Medicine and Hygiene. 17(6): 881-885.

Couceiro, J. N., J. C. Paulson and L. G. Baum (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(2): 155-165.

Crescenzo-Chaigne, B., C. V. Barbezange, S. Leandri, C. Roquin, C. Berthault and S. van der Werf (2017). "Incorporation of the influenza A virus NA segment into virions does not require cognate non-coding sequences." Scientific reports 7: 43462.

Croville, G., S. M. Soubies, J. Barbieri, C. Klopp, J. Mariette, O. Bouchez, C. Camus-Bouclainville and J. L. Guerin (2012). "Field monitoring of avian influenza viruses: whole-genome sequencing and tracking of neuraminidase evolution using 454 pyrosequencing." J Clin Microbiol 50(9): 2881-2887.

De Silva, U. C. and T. Yasunaga (2011). "Swine to human transmission of reassortants of pandemic (H1N1) 2009 and endemic swine influenza viruses: Abstract." PLoS Currents 3: RRN1285.

Desselberger, U., V. R. Racaniello, J. J. Zazra and P. Palese (1980). "The 3' and 5'-terminal sequences of influenza A, B and C virus RNA segments are highly conserved and show partial inverted complementarity." Gene 8(3): 315-328.

Dias, A., D. Bouvier, T. Crepin, A. A. McCarthy, D. J. Hart, F. Baudin, S. Cusack and R. W. H. Ruigrok (2009). "The cap-snatching endonuclease of influenza virus polymerase resides in the PA subunit." Nature 458(7240): 914-918.

Dlugolenski, D., L. Jones, G. Saavedra, S. M. Tompkins, R. A. Tripp and E. Mundt (2011). "Passage of low-pathogenic avian influenza (LPAI) viruses mediates rapid genetic adaptation of a wild-bird isolate in poultry." Arch Virol 156(4): 565-576.

Domingo, E., E. Baranowski, C. M. Ruiz-Jarabo, A. M. Martín-Hernández, J. C. Sáiz and C. Escarmís (1998). "Quasispecies structure and persistence of RNA viruses." Emerging Infectious Diseases 4(4): 521-527.

Drummond, A. J., M. A. Suchard, D. Xie and A. Rambaut (2012). "Bayesian Phylogenetics with BEAUti and the BEAST 1.7." Molecular Biology and Evolution 29: 1969-1973.

Dubois, J., O. Terrier and M. Rosa-Calatrava (2014). "Influenza viruses and mRNA splicing: doing more with less." MBio 5(3): e00070-00014.

Dusek, R. J., G. T. Hallgrimsson, H. S. Ip, J. E. Jónsson, S. Sreevatsan, S. W. Nashold, J. L. Teslaa, S. Enomoto, R. A. Halpin, X. Lin, N. Fedorova, T. B. Stockwell, V. G. Dugan, D. E. Wentworth and J. S. Hall (2014). "North Atlantic migratory bird flyways provide routes for intercontinental movement of avian influenza viruses." PLoS One 9: e92075.

Edinger, T. O., M. O. Pohl and S. Stertz (2014). "Entry of influenza A virus: host factors and antiviral targets." Journal of General Virology 95: 263-277.

Finkelman, B. S., C. Viboud, K. Koelle, M. J. Ferrari, N. Bharti and B. T. Grenfell (2007). "Global Patterns in Seasonal Activity of Influenza A/H3N2, A/H1N1, and B from 1997 to 2005: Viral Coexistence and Latitudinal Gradients." PLoS One 2(12): e1296.

Flick, R., G. Neumann, E. Hoffmann, E. Neumeier and G. Hobom (1996). "Promoter elements in the influenza vRNA terminal structure." Rna 2(10): 1046-1057.

Fodor, A. J. W. t. V. a. E. (2016). "Influenza virus RNA polymerase: insights into the mechanisms of viral RNA synthesis." Nature reviews. Microbiology 14(8): 479-493.

Fodor, E., D. C. Pritlove and G. G. Brownlee (1994). "The influenza virus panhandle is involved in the initiation of transcription." Journal of Virology 68(6): 4092-4096.

Fodor, E., B. L. Seong and G. G. Brownlee (1993). "Photochemical cross-linking of influenza A polymerase to its virion RNA promoter defines a polymerase binding site at residues 9 to 12 of the promoter." The Journal of General Virology 74 (Pt 7): 1327-1333.

Fouchier, R. A. M., V. Munster, A. Wallensten, T. M. Bestebroer, S. Herfst, D. Smith, G. F. Rimmelzwaan, B. Olsen and A. D. M. E. Osterhaus (2005). "Characterization of a novel influenza A virus hemagglutinin subtype (H16) obtained from black-headed gulls." Journal of Virology 79: 2814-2822.

Fries, A. C., J. M. Nolting, A. Danner, R. G. Webster, A. S. Bowman, S. Krauss, R. D. Slemons and Y. Wang (2013). "Evidence for the circulation and inter-hemispheric movement of the H14 subtype influenza A virus." PLoS ONE 8: e59216.

Fujii, K., Y. Fujii, T. Noda, Y. Muramoto, T. Watanabe, A. Takada, H. Goto, T. Horimoto and Y. Kawaoka (2005). "Importance of both the coding and the segment-specific noncoding regions of the influenza A virus NS segment for its efficient incorporation into virions." Journal of Virology 79(6): 3766-3774.

Furuse, Y. and H. Oshitani (2011). "Evolution of the influenza A virus untranslated regions." Infection, Genetics and Evolution 11(5): 1150-1154.

Gamblin, S. J. and J. J. Skehel (2010). "Influenza Hemagglutinin and Neuraminidase Membrane Glycoproteins." The Journal of Biological Chemistry 285(37): 28403-28409.

Gao, Q. and P. Palese (2009). "Rewiring the RNAs of influenza virus to prevent reassortment." PNAS 106(37): 15891.

Garcia-Sastre, A., A. Egorov, D. Matassov, S. Brandt, D. E. Levy, J. E. Durbin, P. Palese and T. Muster (1998). "Influenza A virus lacking the NS1 gene replicates in interferon-deficient systems." Virology 252(2): 324-330.

Garten, W. and H. D. Klenk (1983). "Characterization of the carboxypeptidase involved in the proteolytic cleavage of the influenza haemagglutinin." J Gen Virol 64 (Pt 10): 2127-2137.

Geraci, J. R., D. J. Staubin, I. K. Barker, R. G. Webster, V. S. Hinshaw, W. J. Bean, H. L. Ruhnke, J. H. Prescott, G. Early, A. S. Baker, S. Madoff and R. T. Schooley (1982). "Mass mortality of harbor seals: pneumonia associated with influenza A virus." Science 215(4536): 1129-1131.

Girard, M. P., J. S. Tam, O. M. Assossou and M. P. Kieny (2010). "The 2009 A (H1N1) influenza virus pandemic: A review." Vaccine 28(31): 4895-4902.

Glezen, W. P., J. K. Schmier, C. M. Kuehn, K. J. Ryan and J. Oxford (2013). "The Burden of Influenza B: A Structured Literature Review." American Journal of Public Health 103(3): E43-E51.

Goldstein, T., I. Mena, S. J. Anthony, R. Medina, P. W. Robinson, D. J. Greig, D. P. Costa, W. I. Lipkin, A. Garcia-Sastre and W. M. Boyce (2013). "Pandemic H1N1 influenza isolated from free-ranging Northern Elephant Seals in 2010 off the central California coast." PLoS ONE 8: e62259.

Gonzalez-Reiche, A. S. and D. R. Perez (2012). "Where do avian influenza viruses meet in the Americas?" Avian Diseases 56: 1025-1033.

Groth, M., J. Lange, P. Kanrai, S. Pleschka, C. Scholtissek, A. Krumbholz, M. Platzer, A. Sauerbrei and R. Zell (2014). "The genome of an influenza virus from a pilot whale: Relation to influenza viruses of gulls and marine mammals." Infection Genetics and Evolution 24: 183-186.

Guo, Y. J., F. G. Jin, P. Wang, M. Wang and J. M. Zhu (1983). "Isolation of influenza C virus from pigs and experimental infection of pigs with influenza C virus." The Journal of General Virology. 64 (Pt 1): 177-182.

Hale, B. G., R. A. Albrecht and A. García-Sastre (2010). "Innate immune evasion strategies of influenza viruses." Future microbiology 5: 23.

Hall, J. S., J. L. TeSlaa, S. W. Nashold, R. A. Halpin, T. Stockwell, E. Wentworth, V. G. Dugan and H. S. Ip (2013). "Evolution of a reassortant North American gull influenza virus lineage: drift, shift and stability." Virology Journal 10: 179.

Hao, W. (2011). "Evidence of intra-segmental homologous recombination in influenza A virus." Gene 481(2): 57-64.

Harkness, J. W., G. C. Schild, P. H. Lamont and C. M. Brand (1972). "Studies on relationships between human and porcine influenza: 1. Serological evidence of infection in swine in Great Britain with an influenza A virus antigenically like human Hong Kong/68 virus." Bulletin of the World Health Organization 46(6): 709-719.

He, C. Q., G. Z. Han, D. Wang, W. Liu, G. R. Li, X. P. Liu and N. Z. Ding (2008). "Homologous recombination evidence in human and swine influenza A viruses." Virology 380(1): 12-20.

He, C. Q., Z. X. Xie, G. Z. Han, J. B. Dong, D. Wang, J. B. Liu, L. Y. Ma, X. F. Tang, X. P. Liu, Y. S. Pang and G. R. Li (2009). "Homologous recombination as an evolutionary force in the avian influenza A virus." Molecular Biology and Evolution. 26(1): 177-187.

Hinshaw, V. S., G. M. Air, A. J. Gibbs, L. Graves, B. Prescott and D. Karunakaran (1982). "Antigenic and genetic characterization of a novel hemagglutinin subtype of influenza A viruses from gulls." Journal of Virology 42: 865-872.

Hinshaw, V. S., W. J. Bean, J. Geraci, P. Fiorelli, G. Early and R. G. Webster (1986). "Characterization of two influenza A viruses from a pilot whale." Journal of Virology 58(2): 655-656.

Hinshaw, V. S., W. J. Bean, R. G. Webster, J. E. Rehg, P. Fiorelli, G. Early, J. R. Geraci and D. J. St Aubin (1984). "Are seals frequently infected with avian influenza viruses?" Journal of Virology. 51(3): 863-865.

Hinshaw, V. S., R. G. Webster and B. Turner (1979). "Water-bone transmission of influenza A viruses?" Intervirology 11(1): 66-68.

Hiromoto, Y., T. Saito, S. E. Lindstrom, Y. Li, R. Nerome, S. Sugita, M. Shinjoh and K. Nerome (2000). "Phylogenetic analysis of the three polymerase genes (PB1, PB2 and PA) of influenza B virus." The Journal of General Virology. 81(Pt 4): 929-937.

Hoffmann, E., J. Stech, Y. Guan, R. G. Webster and D. R. Perez (2001). "Universal primer set for the full-length amplification of all influenza A viruses." Arch Virol 146(12): 2275-2289.

Hoffmann, E., J. Stech, Y. Guan, R. G. Webster and D. R. Perez (2001). "Universal primer set for the full-length amplification of all influenza A viruses." Archives of Virology 146(12): 2275-2289.

Holland, J., K. Spindler, F. Horodyski, E. Grabau, S. Nichol and S. VandePol (1982). "Rapid evolution of RNA genomes." Science 215(4540): 1577-1585.

Horimoto, T. and Y. Kawaoka (2001). "Pandemic threat posed by avian influenza A viruses." Clinical Microbiology Reviews 14(1): 129-+.

Hossain, M. J., D. Hickman and D. R. Perez (2008). "Evidence of Expanded Host Range and Mammalian-Associated Genetic Changes in a Duck H9N2 Influenza Virus Following Adaptation in Quail and Chickens." PLoS ONE 3(9): e3170.

Hsu, M. T., J. D. Parvin, S. Gupta, M. Krystal and P. Palese (1987). "Genomic RNAs of influenza viruses are held in a circular conformation in virions and in infected cells by a terminal panhandle." Proceedings of the National Academy of Sciences of the United States of America 84(22): 8140-8144.

Hsu, M. T., J. D. Parvin, S. Gupta, M. Krystal and P. Palese (1987). "Genomic RNAs of influenza viruses are held in a circular conformation in virions and in infected cells by a terminal panhandle." PNAS 84(22): 8140-8144.

Huang, Y., G. J. Robertson, D. Ojkic, H. Whitney and A. S. Lang (2014). "Diverse inter-continental and host lineage reassortant avian influenza A viruses in pelagic seabirds." Infection, Genetics and Evolution : Journal of molecular epidemiology and evolutionary genetics in infectious diseases. 22: 103-111.

Huang, Y., M. Wille, J. Benkaroun, H. Munro, A. L. Bond, D. A. Fifield, G. J. Robertson, D. Ojkic, H. Whitney and A. S. Lang (2014). "Perpetuation and reassortment of gull influenza A viruses in Atlantic North America." Virology 456-457: 353-363.

Huang, Y., M. Wille, A. Dobbin, G. J. Robertson, P. Ryan, D. Ojkic, H. Whitney and A. S. Lang (2013). "A 4-year study of avian influenza virus prevalence and subtype diversity in ducks of Newfoundland, Canada." Canadian Journal of Microbiology. 59(10): 701-708.

Huang, Y., M. Wille, A. Dobbin, N. M. Walzthoni, G. J. Robertson, D. Ojkic, H. Whitney and A. S. Lang (2014). "Genetic structure of avian influenza viruses from ducks of the Atlantic flyway of North America." PLoS One 9(1): e86999.

Hulo, C., E. de Castro, P. Masson, L. Bougueleret, A. Bairoch, I. Xenarios and P. Le Mercier (2011). "ViralZone: a knowledge resource to understand virus diversity." Nucleic Acids Research. 39(Database issue): D576-582.

Hurt, A. C., D. Vijaykrishna, J. Butler, C. Baas, S. Maurer-Stroh, M. C. Silva-de-la-Fuente, G. Medina-Vogel, B. Olsen, A. Kelso, I. G. Barr and D. Gonzalez-Acuna (2014). "Detection of evolutionarily distinct avian influenza A viruses in Antarctica." mBio 5: e01098-01014.

Iamnikova, S. S., A. S. Gambarian, V. A. Aristova, K. L'Vov D, N. F. Lomakina, V. Munster, P. Lexmond and R. A. Foucher (2009). "[A/H13 and A/H16 influenza viruses: different lines of one precursors]." Voprosy virusologii 54: 10-18.

Imai, M., T. Watanabe, M. Hatta, S. C. Das, M. Ozawa, K. Shinya, G. Zhong, A. Hanson, H. Katsura, S. Watanabe, C. Li, E. Kawakami, S. Yamada, M. Kiso, Y. Suzuki, E. A. Maher, G. Neumann and Y. Kawaoka (2012). "Experimental adaptation of an influenza H5 haemagglutinin (HA) confers respiratory droplet transmission to a reassortant H5 HA/H1N1 virus in ferrets." Nature 486(7403): 420-428.

Ip, H. S., P. L. Flint, J. C. Franson, R. J. Dusek, D. V. Derksen, R. E. Gill, Jr., C. R. Ely, J. M. Pearce, R. B. Lanctot, S. M. Matsuoka, D. B. Irons, J. B. Fischer, R. M. Oates, M. R. Petersen, T. F. Fondell, D. A. Rocque, J. C. Pedersen and T. C. Rothe (2008). "Prevalence of Influenza A viruses in wild migratory birds in Alaska: patterns of variation in detection at a crossroads of intercontinental flyways." Virology Journal 5: 71.

Ito, T., J. N. Couceiro, S. Kelm, L. G. Baum, S. Krauss, M. R. Castrucci, I. Donatelli, H. Kida, J. C. Paulson, R. G. Webster and Y. Kawaoka (1998). "Molecular basis for the

generation in pigs of influenza A viruses with pandemic potential." Journal of Virology. 72(9): 7367-7373.

Ito, T. and Y. Kawaoka (2000). "Host-range barrier of influenza A viruses." Vet Microbiol 74(1-2): 71-75.

Joseph, U., M. Linster, Y. Suzuki, S. Krauss, R. A. Halpin, D. Vijaykrishna, T. P. Fabrizio, T. M. Bestebroer, S. Maurer-Stroh, R. J. Webby, D. E. Wentworth, R. A. Fouchier and J. Bahl (2015). "Adaptation of pandemic H2N2 influenza A viruses in humans." Journal of Virology 89(4): 2442-2447.

Jourdain, E., G. Gunnarsson, J. Wahlgren, N. Latorre-Margalef, C. Brojer, S. Sahlin, L. Svensson, J. Waldenstrom, A. Lundkvist and B. Olsen (2010). "Influenza virus in a natural host, the mallard: experimental infection data." PLoS One 5(1): e8935.

Jourdain, E., D. van Riel, V. J. Munster, T. Kuiken, J. Waldenstrom, B. Olsen and P. Ellstrom (2011). "The pattern of influenza virus attachment varies among wild bird species." PLoS ONE 6: e24155.

Karasin, A. I., K. West, S. Carman and C. W. Olsen (2004). "Characterization of avian H3N3 and H1N1 influenza A viruses isolated from pigs in Canada." Journal of Clinical Microbiology 42: 4349-4354.

Kawaoka, Y., S. Krauss and R. G. Webster (1989). "Avian-to-human transmission of the PB1 gene of influenza A viruses in the 1957 and 1968 pandemics." Journal of Virology. 63(11): 4603-4608.

Kearse, M., R. Moir, A. Wilson, S. Stones-Havas, M. Cheung, S. Sturrock, S. Buxton, A. Cooper, S. Markowitz, C. Duran, T. Thierer, B. Ashton, P. Meintjes and A. Drummond (2012). "Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data." Bioinformatics 28(12): 1647-1649.

Kida, H., L. E. Brown and R. G. Webster (1982). "Biological activity of monoclonal antibodies to operationally defined antigenic regions on the hemagglutinin molecule of A/Seal/Massachusetts/1 /80 (H7N7) influenza virus." Virology 122(1): 38-47.

Kim, H. J., E. Fodor, G. G. Brownlee and B. L. Seong (1997). "Mutational analysis of the RNA-fork model of the influenza A virus vRNA promoter in vivo." The Journal of General Virology 78 (Pt 2): 353-357.

Kimura, H., C. Abiko, G. Peng, Y. Muraki, K. Sugawara, S. Hongo, F. Kitame, K. Mizuta, Y. Numazaki, H. Suzuki and K. Nakamura (1997). "Interspecies transmission of influenza C virus between humans and pigs." Virus Res 48(1): 71-79.

Krauss, S., C. A. Obert, J. Franks, D. Walker, K. Jones, P. Seiler, L. Niles, S. P. Pryor, J. C. Obenauer, C. W. Naeve, L. Widjaja, R. J. Webby and R. G. Webster (2007). "Influenza in migratory birds and evidence of limited intercontinental virus exchange." PLoS Pathogens 3: e167.

Kuiken, T., E. C. Holmes, J. McCauley, G. F. Rimmelzwaan, C. S. Williams and B. T. Grenfell (2006). "Host species barriers to influenza virus infections." Science 312(5772): 394-397.

Kumlin, U., S. Olofsson, K. Dimock and N. Arnberg (2008). "Sialic acid tissue distribution and influenza virus tropism." Influenza and Other Respiratory Viruses 2(5): 147-154.

Lang, A. S., C. Lebarbenchon, A. M. Ramey, G. J. Robertson, J. Waldenstrom and M. Wille (2016). "Assessing the Role of Seabirds in the Ecology of Influenza A Viruses." Avian Dis 60(1 Suppl): 378-386.

Latorre-Margalef, N., G. Gunnarsson, V. J. Munster, R. A. Fouchier, A. D. Osterhaus, J. Elmberg, B. Olsen, A. Wallensten, P. D. Haemig, T. Fransson, L. Brudin and J. Waldenstrom (2009). "Effects of influenza A virus infection on migrating mallard ducks." Proceedings. Biological Sciences. 276(1659): 1029-1036.

Laudert, E., V. Sivanandan, D. Halvorson, D. Shaw and R. G. Webster (1993). "Biological and molecular characterization of H13N2 influenza type A viruses isolated from turkeys and surface water." Avian Diseases 37: 793-799.

Lee, D.-H., D. E. Swayne, M. K. Torchetti, K. Winker, H. S. Ip and C.-S. Song (2015). "Intercontinental spread of asian-origin H5N8 to North America through Beringia by migratory birds." Journal of Virology 89: 6521-6524.

Lee, M. K., S. H. Bae, C. J. Park, H. K. Cheong, C. Cheong and B. S. Choi (2003). "A single-nucleotide natural variation (U4 to C4) in an influenza A virus promoter exhibits a large structural change: implications for differential viral RNA synthesis by RNA-dependent RNA polymerase." Nucleic Acids Research 31(4): 1216-1223.

Lefkowitz, E. J., D. M. Dempsey, R. C. Hendrickson, R. J. Orton, S. G. Siddell and D. B. Smith (2018). "Virus taxonomy: the database of the International Committee on Taxonomy of Viruses (ICTV)." Nucleic Acids Research. 46(D1): D708-d717.

Li, J. and C. J. Cardona (2010). "Adaptation and transmission of a wild duck avian influenza isolate in chickens." Avian Dis 54(1 Suppl): 586-590.

Li, J., H. zu Dohna, N. L. Anchell, S. C. Adams, N. T. Dao, Z. Xing and C. J. Cardona (2010). "Adaptation and transmission of a duck-origin avian influenza virus in poultry species." Virus Res 147(1): 40-46.

Li, X. and P. Palese (1992). "Mutational analysis of the promoter required for influenza virus virion RNA synthesis." Journal of Virology 66(7): 4331-4338.

Li, X. and P. Palese (1994). "Characterization of the polyadenylation signal of influenza virus RNA." Journal of Virology 68(2): 1245-1249.

Lin, D., J. Lan and Z. Zhang (2007). "Structure and function of the NS1 protein of influenza A virus." Acta Biochimica et Biophysica Sinica. 39(3): 155-162.

Lindskog, C., P. Ellström, B. Olsen, F. Pontén, D. van Riel, V. J. Munster, D. González-Acuña, T. Kuiken and E. Jourdain (2013). "European H16N3 gull influenza virus attaches to the human respiratory tract and eye." PLoS ONE 8: e60757.

Ludwig, S., U. Schultz, J. Mandler, W. M. Fitch and C. Scholtissek (1991). "Phylogenetic relationship of the nonstructural (NS) genes of influenza A viruses." Virology 183(2): 566-577.

Luo, G. X., W. Luytjes, M. Enami and P. Palese (1991). "The polyadenylation signal of influenza virus RNA involves a stretch of uridines followed by the RNA duplex of the panhandle structure." Journal of Virology 65(6): 2861-2867.

Lvov, D. K., V. M. Zhdanov, F. Sadykhova, S. S. Yamnikova, V. A. Isachenko and E. A. Vladimirtseva (1983). "Isolation of a natural recombinant of influenza A virus (H1N3) from a sick child." Lancet 2: 1246-1247.

Mandler, J., O. T. Gorman, S. Ludwig, E. Schroeder, W. M. Fitch, R. G. Webster and C. Scholtissek (1990). "Derivation of the nucleoproteins (NP) of influenza A viruses isolated from marine mammals." Virology 176: 255-261.

Manuguerra, J. C. and C. Hannoun (1992). "Natural infection of dogs by influenza C virus." Research in Virology. 143(3): 199-204.

Matrosovich, M. N., T. Y. Matrosovich, T. Gray, N. A. Roberts and H. D. Klenk (2004). "Human and avian influenza viruses target different cell types in cultures of human airway epithelium." Proc Natl Acad Sci U S A 101(13): 4620-4624.

Mjaaland, S., E. Rimstad, K. Falk and B. H. Dannevig (1997). "Genomic characterization of the virus causing infectious salmon anemia in Atlantic salmon (Salmo salar L.): an orthomyxo-like virus in a teleost." Journal of Virology 71(10): 7681-7686.

Moncla, L. H., G. Zhong, C. W. Nelson, J. M. Dinis, J. Mutschler, A. L. Hughes, T. Watanabe, Y. Kawaoka and T. C. Friedrich (2016). "Selective bottlenecks shape

evolutionary pathways taken during mammalian adaptation of a 1918-like avian influenza virus." Cell host & microbe 19(2): 169-180.

Moriuchi, H., N. Katsushima, H. Nishimura, K. Nakamura and Y. Numazaki (1991). "Community-acquired influenza C virus infection in children." The Journal of Pediatrics 118(2): 235-238.

Munir, M., S. Zohari, G. Metreveli, C. Baule, S. Belak and M. Berg (2011). "Alleles A and B of non-structural protein 1 of avian influenza A viruses differentially inhibit beta interferon production in human and mink lung cells." J Gen Virol 92(Pt 9): 2111-2121.

Mushtaq, M. H., H. Juan, P. Jiang, Y. Li, T. Li, Y. Du and M. M. Mukhtar (2008). "Complete genome analysis of a highly pathogenic H5N1 influenza A virus isolated from a tiger in China." Archives of Virology 153: 1569-1574.

Naeem, K. and M. Hussain (1995). "An outbreak of avian influenza in poultry in Pakistan." Vet Rec 137(17): 439.

Nao, N., J. Yamagishi, H. Miyamoto, M. Igarashi, R. Manzoor, A. Ohnuma, Y. Tsuda, W. Furuyama, A. Shigeno, M. Kajihara, N. Kishida, R. Yoshida and A. Takada (2017). "Genetic Predisposition To Acquire a Polybasic Cleavage Site for Highly Pathogenic Avian Influenza Virus Hemagglutinin." MBio 8(1).

Nielsen, O., A. Clavijo and J. A. Boughen (2001). "Serologic evidence of influenza A infection in marine mammals of Arctic Canada." Journal of Wildlife Diseases 37(4): 820-825.

Ohishi, K., N. Kishida, A. Ninomiya, H. Kida, Y. Takada, N. Miyazaki, A. N. Boltunov and T. Maruyama (2004). "Antibodies to human-related H3 influenza A virus in Baikal seals (*Phoca sibirica*) and ringed seals (*Phoca hispida*) in Russia." Microbiology and Immunology 48: 905-909.

Olsen, B., V. J. Munster, A. Wallensten, J. Waldenstrom, A. Osterhaus and R. A. M. Fouchier (2006). "Global patterns of influenza A virus in wild birds." Science 312(5772): 384-388.

Olsen, B., V. J. Munster, A. Wallensten, J. Waldenstrom, A. D. Osterhaus and R. A. Fouchier (2006). "Global patterns of influenza A virus in wild birds." Science 312: 384-388.

Olsen, K. M. and H. Larsson (2004). Gulls of North America, Europe, and Asia. Princeton, Princeton University Press.

Orlich, M., H. Gottwald and R. Rott (1994). "Nonhomologous recombination between the hemagglutinin gene and the nucleoprotein gene of an influenza virus." Virology 204(1): 462-465.

Osterhaus, A. D., G. F. Rimmelzwaan, B. E. Martina, T. M. Bestebroer and R. A. Fouchier (2000). "Influenza B virus in seals." Science 288(5468): 1051-1053.

Otsuki, K., Y. Kawaoka, T. Nakamura and M. Tsubokura (1982). "Pathogenicity for chickens of avian influenza viruses isolated from whistling swans and a black-tailed gull in Japan." Avian Diseases 26: 314-320.

Palese, P. and R. W. Compans (1976). "Inhibition of influenza virus replication in tissue culture by 2-deoxy-2,3-dehydro-N-trifluoroacetylneuraminic acid (FANA): mechanism of action." The Journal of General Virology. 33(1): 159-163.

Pasick, J., Y. Berhane, T. Joseph, V. Bowes, T. Hisanaga, K. Handel and S. Alexandersen (2015). "Reassortant Highly Pathogenic Influenza A H5N2 Virus Containing Gene Segments Related to Eurasian H5N8 in British Columbia, Canada, 2014." Scientific Reports 5: 9484.

Pereda, A. J., M. Uhart, A. A. Perez, M. E. Zaccagnini, L. La Sala, J. Decarre, A. Goijman, L. Solari, R. Suarez, M. I. Craig, A. Vagnozzi, A. Rimondi, G. Konig, M. V. Terrera, A. Kaloghlian, H. Song, E. M. Sorrell and D. R. Perez (2008). "Avian influenza virus isolated in wild waterfowl in Argentina: evidence of a potentially unique phylogenetic lineage in South America." Virology 378: 363-370.

Perkins, L. E. L. and D. E. Swayne (2002). "Susceptibility of laughing gulls (*Larus atricilla*) to H5N1 and H5N3 highly pathogenic avian influenza viruses." Avian Diseases 46: 877-885.

Pflug, A., D. Guilligay, S. Reich and S. Cusack (2014). "Structure of influenza A polymerase bound to the viral RNA promoter." Nature 516(7531): 355-360.

Pinto, L. H., L. J. Holsinger and R. A. Lamb (1992). "Influenza virus M2 protein has ion channel activity." Cell 69(3): 517-528.

Poon, L. L., D. C. Pritlove, E. Fodor and G. G. Brownlee (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(4): 3473-3476.

Presti, R. M., G. Zhao, W. L. Beatty, K. A. Mihindukulasuriya, A. P. da Rosa, V. L. Popov, R. B. Tesh, H. W. Virgin and D. Wang (2009). "Quaranfil, Johnston Atoll, and Lake Chad viruses are novel members of the family Orthomyxoviridae." Journal of Virology. 83(22): 11599-11606.

Presti, R. M., G. Y. Zhao, W. L. Beatty, K. A. Mihindukulasuriya, A. da Rosa, V. L. Popov, R. B. Tesh, H. W. Virgin and D. Wang (2009). "Quaranfil, Johnston Atoll, and Lake Chad Viruses Are Novel Members of the Family Orthomyxoviridae." Journal of Virology 83(22): 11599-11606.

Pritlove, D. C., E. Fodor, B. L. Seong and G. G. Brownlee (1995). "In vitro transcription and polymerase binding studies of the termini of influenza A virus cRNA: evidence for a cRNA panhandle." The Journal of General Virology 76 (Pt 9): 2205-2213.

Pritlove, D. C., L. L. Poon, E. Fodor, J. Sharps and G. G. Brownlee (1998). "Polyadenylation of influenza virus mRNA transcribed in vitro from model virion RNA templates: requirement for 5' conserved sequences." Journal of Virology 72(2): 1280-1286.

Qu, Y., R. Zhang, P. Cui, G. Song, Z. Duan and F. Lei (2011). "Evolutionary genomics of the pandemic 2009 H1N1 influenza viruses (pH1N 1v)." Virology Journal 8: 250.

R Development Core Team (2008). R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing.

Ramey, A. M., R. L. Poulson, A. S. Gonzalez-Reiche, D. R. Perez, D. E. Stallknecht and J. D. Brown (2014). "Genomic characterization of H14 subtype influenza A viruses in new world waterfowl and experimental infectivity in mallards (*Anas platyrhynchos*)." PLoS ONE 9: e95620.

Ramey, A. M., A. B. Reeves, S. A. Sonsthagen, J. L. TeSlaa, S. Nashold, T. Donnelly, B. Casler and J. S. Hall (2015). "Dispersal of H9N2 influenza A viruses between East Asia and North America by wild birds." Virology 482: 79-83.

Ratanakorn, P., A. Wiratsudakul, W. Wiriyarat, K. Eiamampai, A. H. Farmer, R. G. Webster, K. Chaichoune, S. Suwanpakdee, D. Pothieng and P. Puthavathana (2012). "Satellite tracking on the flyways of Brown-Headed Gulls and their potential role in the spread of highly pathogenic avian influenza H5N1 virus " PLoS ONE 7: e49939.

Reid, A. H. and J. K. Taubenberger (2003). "The origin of the 1918 pandemic influenza virus: a continuing enigma." The Journal of General Virology. 84(Pt 9): 2285-2292.

Richt, J. A., K. M. Lager, B. H. Janke, R. D. Woods, R. G. Webster and R. J. Webby (2003). "Pathogenic and antigenic properties of phylogenetically distinct reassortant

H3N2 swine influenza viruses cocirculating in the United States." Journal of Clinical Microbiology 41: 3198-3205.

Robertson, J. S. (1979). "5' and 3' terminal nucleotide sequences of the RNA genome segments of influenza virus." Nucleic Acids Research 6(12): 3745-3757.

Rohm, C., J. Suss, V. Pohle and R. G. Webster (1996). "Different hemagglutinin cleavage site variants of H7N7 in an influenza outbreak in chickens in Leipzig, Germany." Virology 218: 253-257.

Sangster, G., J. M. Collinson, A. G. Knox, D. T. Parkin and L. Svensson (2007). "Taxonomic recommendations for British birds: Fourth report." Ibis 149: 853-857.

Schafer, J. R., Y. Kawaoka, W. J. Bean, J. Suss, D. Senne and R. G. Webster (1993). "Origin of the pandemic 1957 H2 influenza A virus and the persistence of its possible progenitors in the avian reservoir." Virology 194(2): 781-788.

Schneider, W. L. and M. J. Roossinck (2001). "Genetic Diversity in RNA Virus Quasispecies Is Controlled by Host-Virus Interactions." Journal of Virology 75(14): 6566-6571.

Schultz, U., W. M. Fitch, S. Ludwig, J. Mandler and C. Scholtissek (1991). "Evolution of pig influenza viruses." Virology 183(1): 61-73.

Seong, B. L. and G. G. Brownlee (1992). "Nucleotides 9 to 11 of the influenza A virion RNA promoter are crucial for activity in vitro." The Journal of General Virology 73 (Pt 12): 3115-3124.

Shao, W., X. Li, M. U. Goraya, S. Wang and J.-L. Chen (2017). "Evolution of Influenza A Virus by Mutation and Re-Assortment." International Journal of Molecular Sciences 18(8): 1650.

Shoham, D. and S. Rogers (2006). "Greenland as a plausible springboard for trans-Atlantic avian influenza spread." Medical Hypotheses 67: 1460-1461.

Shope, R. E. (1931). "Swine Influenza. I. Experimental Transmission and Pathology." The Journal of Experimental Medicine. 54(3): 349-359.

Shors, T. (2013, 2018/8/28/). "H5N1 virus (bird flu) controversy." from https://www.accessscience.com:443/content/h5n1-virus-bird-flu-controversy/YB130019.

Shortridge, K. F. (1999). "Poultry and the influenza H5N1 outbreak in Hong Kong, 1997: abridged chronology and virus isolation." Vaccine 17 Suppl 1: S26-29.

Simonsen, L. (1999). "The global impact of influenza on morbidity and mortality." Vaccine 17 Suppl 1: S3-10.

Simonsen, L., M. J. Clarke, L. B. Schonberger, N. H. Arden, N. J. Cox and K. Fukuda (1998). "Pandemic versus epidemic influenza mortality: A pattern of changing age distribution." Journal of Infectious Diseases 178(1): 53-60.

Sivanandan, V., D. A. Halvorson, E. Laudert, D. A. Senne and M. C. Kumar (1991). "Isolation of H13N2 influenza A virus from turkeys and surface water." Avian Diseases 35: 974-977.

Skehel, J. J. and A. J. Hay (1978). "Nucleotide sequences at the 5' termini of influenza virus RNAs and their transcripts." Nucleic Acids Research 5(4): 1207-1219.

Smrt, S. T., A. W. Draney and J. L. Lorieau (2015). "The Influenza Hemagglutinin Fusion Domain Is an Amphipathic Helical Hairpin That Functions by Inducing Membrane Curvature." The Journal of Biological Chemistry 290(1): 228-238.

Sorrell, E. M. and D. R. Perez (2007). "Adaptation of influenza A/Mallard/Potsdam/178-4/83 H2N2 virus in Japanese quail leads to infection and transmission in chickens." Avian Dis 51(1 Suppl): 264-268.

Spalding, M. (2009). "Diseases of Poultry, 12th Edition." Journal of Wildlife Diseases 45(1): 251-256.

Stallknecht, D. E., M. T. Kearney, S. M. Shane and P. J. Zwank (1990). "Effects of pH, temperature, and salinity on persistence of avian influenza viruses in water" Avian Diseases 34(2): 412-418.

Stallknecht, D. E., M. T. Kearney, S. M. Shane and P. J. Zwank (1990). "Effects of pH, temperature, and salinity on persistence of avian influenza viruses in water." Avian Diseases 34: 412-418.

Stallknecht, D. E., S. M. Shane, M. T. Kearney and P. J. Zwank (1990). "Persistence of avian influenza viruses in water." Avian Diseases 34(2): 406-411.

Stech, J. and T. C. Mettenleiter (2013). "Virulence determinants of high-pathogenic avian influenza viruses in gallinaceous poultry." Future Virology 8: 459-468.

Stegmann, T. (2000). "Membrane fusion mechanisms: the influenza hemagglutinin paradigm and its implications for intracellular fusion." Traffic 1(8): 598-604.

Steinhauer, D. A. (1999). "Role of hemagglutinin cleavage for the pathogenicity of influenza virus." Virology 258(1): 1-20.

Subbarao, K., A. Klimov, J. Katz, H. Regnery, W. Lim, H. Hall, M. Perdue, D. Swayne, C. Bender, J. Huang, M. Hemphill, T. Rowe, M. Shaw, X. Xu, K. Fukuda and N. Cox (1998). "Characterization of an avian influenza A (H5N1) virus isolated from a child with a fatal respiratory illness." Science 279: 393-396.

Suzuki, Y. (2006). "Natural selection on the influenza virus genome." Mol Biol Evol 23(10): 1902-1911.

Suzuki, Y. and M. Nei (2002). "Origin and Evolution of Influenza Virus Hemagglutinin Genes." Molecular Biology and Evolution 19(4): 501-509.

Tamura, K., G. Stecher, D. Peterson, A. Filipski and S. Kumar (2013). "MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0." Molecular Biology and Evolution 30(12): 2725-2729.

Taylor, R. M. (1949). "Studies on Survival of Influenza Virus Between Epidemics and Antigenic Variants of the Virus." American Journal of Public Health and the Nations Health 39(2): 171-178.

Teale, C. J. (2002). "Antimicrobial resistance and the food chain." Journal of Applied Microbiology 92 Suppl: 85s-89s.

Toennessen, R., A. Germundsson, C. M. Jonassen, I. Haugen, K. Berg, R. T. Barrett and E. Rimstad (2011). "Virological and serological surveillance for type A influenza in the black-legged kittiwake (*Rissa tridactyla*)." Virology Journal 8: 21.

Tong, S. X., Y. Li, P. Rivailler, C. Conrardy, D. A. A. Castillo, L. M. Chen, S. Recuenco, J. A. Ellison, C. T. Davis, I. A. York, A. S. Turmelle, D. Moran, S. Rogers, M. Shi, Y. Tao, M. R. Weil, K. Tang, L. A. Rowe, S. Sammons, X. Y. Xu, M. Frace, K. A. Lindblade, N. J. Cox, L. J. Anderson, C. E. Rupprecht and R. O. Donis (2012). "A distinct lineage of influenza A virus from bats." Proceedings of the National Academy of Sciences of the United States of America 109(11): 4269-4274.

Tong, S. X., X. Y. Zhu, Y. Li, M. Shi, J. Zhang, M. Bourgeois, H. Yang, X. F. Chen, S. Recuenco, J. Gomez, L. M. Chen, A. Johnson, Y. Tao, C. Dreyfus, W. L. Yu, R. McBride, P. J. Carney, A. T. Gilbert, J. Chang, Z. Guo, C. T. Davis, J. C. Paulson, J. Stevens, C. E. Rupprecht, E. C. Holmes, I. A. Wilson and R. O. Donis (2013). "New World Bats Harbor Diverse Influenza A Viruses." Plos Pathogens 9(10).

Van Borm, S., T. Rosseel, D. Vangeluwe, F. Vandenbussche, T. van den Berg and B. Lambrecht (2012). "Phylogeographic analysis of avian influenza viruses isolated from Charadriiformes in Belgium confirms intercontinental reassortment in gulls." Archives of Virology 157: 1509-1522.

Varghese, J. N., W. G. Laver and P. M. Colman (1983). "Structure of the influenza virus glycoprotein antigen neuraminidase at 2.9 A resolution." Nature 303(5912): 35-40.

Velarde, R., S. E. Calvin, D. Ojkic, I. K. Barker and E. Nagy (2010). "Avian influenza virus H13 circulating in ring-billed gulls (*Larus delawarensis*) in southern Ontario, Canada." Avian Diseases 54: 411-419.

Verhoeven, J. T. P., M. Canuti, H. J. Munro, S. C. Dufour and A. S. Lang (2018). "ViDiT-CACTUS: an inexpensive and versatile library preparation and sequence analysis method for virus discovery and other microbiology applications." Canadian journal of microbiology: 1-13.

Vincent, A., L. Awada, I. Brown, H. Chen, F. Claes, G. Dauphin, R. Donis, M. Culhane, K. Hamilton, N. Lewis, E. Mumford, T. Nguyen, S. Parchariyanon, J. Pasick, G. Pavade, A. Pereda, M. Peiris, T. Saito, S. Swenson, K. Van Reeth, R. Webby, F. Wong and J. Ciacci-Zanella (2014). "Review of Influenza A Virus in Swine Worldwide: A Call for Increased Surveillance and Research." Zoonoses and Public Health 61(1): 4-17.

Wang, M. and M. Veit (2016). "Hemagglutinin-esterase-fusion (HEF) protein of influenza C virus." Protein & Cell 7(1): 28-45.

Watanabe, T., G. Zhong, C. A. Russell, N. Nakajima, M. Hatta, A. Hanson, R. McBride, D. F. Burke, K. Takahashi, S. Fukuyama, Y. Tomita, E. A. Maher, S. Watanabe, M. Imai, G. Neumann, H. Hasegawa, J. C. Paulson, D. J. Smith and Y. Kawaoka (2014). "Circulating avian influenza viruses closely related to the 1918 virus have pandemic potential." Cell Host & Microbe 15: 692-705.

Weber, F. and O. Haller (2007). "Viral suppression of the interferon system." Biochimie 89(6-7): 836-842.

Webster, R. G., W. J. Bean, O. T. Gorman, T. M. Chambers and Y. Kawaoka (1992). "Evolution and ecology of influenza A viruses." Microbiol Rev 56(1): 152-179.

Webster, R. G., V. S. Hinshaw, W. J. Bean, K. L. Van Wyke, J. R. Geraci, D. J. St. Aubin and G. Petursson (1981). "Characterization of an influenza A virus from seals." Virology 113: 712-724.

Wendel, I., D. Rubbenstroth, J. Doedt, G. Kochs, J. Wilhelm, P. Staeheli, H. D. Klenk and M. Matrosovich (2015). "The avian-origin PB1 gene segment facilitated replication and transmissibility of the H3N2/1968 pandemic influenza virus." Journal of Virology. 89(8): 4170-4179.

Westgeest, K. B., C. A. Russell, X. Lin, M. I. Spronken, T. M. Bestebroer, J. Bahl, R. van Beek, E. Skepner, R. A. Halpin, J. C. de Jong, G. F. Rimmelzwaan, A. D. Osterhaus, D. J. Smith, D. E. Wentworth, R. A. Fouchier and M. de Graaf (2014). "Genomewide analysis of reassortment and evolution of human influenza A(H3N2) viruses circulating between 1968 and 2011." Journal of Virology. 88(5): 2844-2857.

White, M. C., J. Steel and A. C. Lowen (2017). "Heterologous Packaging Signals on Segment 4, but Not Segment 6 or Segment 8, Limit Influenza A Virus Reassortment." Journal of Virology 91(11).

Widjaja, I., E. de Vries, P. J. M. Rottier and C. A. M. de Haan (2012). "Competition between Influenza A Virus Genome Segments." PLOS One 7(10): e47529.

Wille, M., G. J. Robertson, H. Whitney, M. A. Bishop, J. A. Runstadler and A. S. Lang (2011). "Extensive geographic mosaicism in avian influenza viruses from gulls in the northern hemisphere." PLoS ONE 6: e20664.

Wille, M., G. J. Robertson, H. Whitney, D. Ojkic and A. S. Lang (2011). "Reassortment of American and Eurasian genes in an influenza A virus isolated from a great black-backed gull (*Larus marinus*), a species demonstrated to move between these regions." Archives of Virology 156: 107-115.

Williams, R. E., H. Hoogstraal, J. Casals, M. N. Kaiser and M. I. Moussa (1973). "Isolation of Wanowrie, Thogoto, and Dhori viruses from Hyalomma ticks infesting camels in Egypt." Journal of Medical Entomology. 10(2): 143-146.

Wilson, I. A., J. J. Skehel and D. C. Wiley (1981). "Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 A resolution." Nature 289(5796): 366-373.

Wood, J. M., R. G. Webster and V. F. Nettles (1985). "Host range of A/Chicken/Pennsylvania/83 (H5N2) influenza virus." Avian Diseases 29: 198-207.

Worobey, M., G. Z. Han and A. Rambaut (2014). "A synchronized global sweep of the internal genes of modern avian influenza virus." Nature 508: 254-257.

Yoon, S. W., R. J. Webby and R. G. Webster (2014). "Evolution and ecology of influenza A viruses." Curr Top Microbiol Immunol 385: 359-375.

Zhao, L., Y. Peng, K. Zhou, M. Cao, J. Wang, X. Wang, T. Jiang and T. Deng (2014). "New Insights into the Nonconserved Noncoding Region of the Subtype-Determinant Hemagglutinin and Neuraminidase Segments of Influenza A Viruses." Journal of Virology 88(19): 11493-11503.

Zheng, H., P. Palese and A. Garcia-Sastre (1996). "Nonconserved nucleotides at the 3' and 5' ends of an influenza A virus RNA play an important role in viral RNA replication." Virology 217(1): 242-251.

Zhou, B., M. E. Donnelly, D. T. Scholes, K. St George, M. Hatta, Y. Kawaoka and D. E. Wentworth (2009). "Single-reaction genomic amplification accelerates sequencing and vaccine production for classical and Swine origin human influenza a viruses." Journal of Virology 83(19): 10309-10313.

Zu Dohna, H., J. Li, C. J. Cardona, J. Miller and T. E. Carpenter (2009). "Invasions by Eurasian avian influenza virus H6 genes and replacement of the virus' North American clade." Emerging Infectious Diseases 15: 1040-1045.

Chapter 5: Summary and conclusions

The province of Newfoundland and Labrador, located on the Atlantic coast of Canada, represents an interesting location for the study of avian influenza A viruses (IAVs) in their natural reservoir because it lies at an intersection for birds from the Arctic, the southern mainland of North America, and Eurasia. Initial investigations into IAVs in the province have shown that different bird species, ranging from seabirds to gulls to ducks, actively participate in the maintenance and transmission of these viruses within and among different wild bird populations (Huang, Wille et al. 2013, Huang, Robertson et al. 2014, Huang, Wille et al. 2014). My thesis work involved employing new technologies to characterize the full genome of viruses from different wild bird species and perform in-depth analyses on the genomic structure and evolution of the virus.

5.1 IAVs quasispecies population structure in wild birds

I employed high-throughput sequencing technologies in order to characterize the full genome and quasispecies population of IAVs isolated from three different groups of wild bird hosts: seabirds, gulls and ducks (Chapter 2). My study revealed that a high proportion of shared variants exists within quasispecies of similar viruses and from similar hosts. Research on quasispecies conservation among related viruses is limited, but some comparison studies have shown that virus variants are also conserved in high proportions between similar viruses and within the same host (Laskus, Wilkinson et al. 2004, Asghar, Pettersson et al. 2017). The presence of non-

synonymous substitutions within all virus quasispecies analyzed was proportionally high for the HA and NS gene segments while synonymous substitutions predominate for the rest of the segments. Lastly, analysis of the variant diversity within each virus quasispecies shows that viruses with identical genotypes display similar quasispecies in terms of the numbers of shared variants, but the amount of diversity differs since some segments contain a higher proportion of shared variants than others. Overall, my results show that the virus quasispecies patterns observed in similar viruses is conserved during the virus transmission in similar individual bird hosts. This shows that the viral genomic composition is closely related to a certain host environment since similar virus quasispecies are found in similar hosts in our study. Similar studies could be applied to a broader variety of hosts to investigate the evolution pattern of the virus at the quasispecies level.

5.2 Effect of replication in a different host on the IAV genome

To understand how IAVs change genetically in order to replicate and adapt in different bird hosts, I tracked experimentally the evolution of a virus isolated from a wild duck in a different bird host, the Atlantic puffin (Chapter 2). After 10 passages of the virus in puffin embryos, the virus genome displayed genetic changes in all segments with an overall increase of virus variants in the different virus passages analyzed. Non-synonymous substitutions were found in high proportion in the HA and NA gene segments and both evolved under positive selection as indicated by the dN/dS ratio. Mutational analysis of the HA proteins showed that non-synonymous substitutions were introduced within the cleavage site and fusion domain. The NA proteins showed mutations within the stalk and catalytic domains. The preferential

occurrence of non-synonymous mutations in the HA and NA segments was also found in other mutation studies with bird hosts (Hossain, Hickman et al. 2008, Dlugolenski, Jones et al. 2011). To conclude, my study shows that the replication of a wild duck virus in a different bird host results in changes mostly within the HA and NA proteins, which is in accordance with previous studies.

5.3 Analysis of IAV genomic non-coding regions

I characterized the 3' and 5' NCRs of IAVs isolated from different wild bird hosts (ducks, gulls and seabirds) by the RACE method (Chapter 3). This analysis showed that the conserved region at both extremities is identical among the different hosts analyzed, but the segment-specific regions showed some variability, which corresponds to previous research findings (Robertson 1979, Desselberger, Racaniello et al. 1980, Benkaroun, Robertson et al. 2018). I extended the analysis to a greater number of NCRs and showed that the amount of variability in the segment-specific regions is highest for the HA and NA NCRs. Therefore, these two segments show the greatest variability among the eight segments for both coding and non-coding regions. Examination of some 3' NCRs from HA and NA subtypes showed that the relationships among NCR sequences matches their coding region phylogenies, in terms of viral host species and geographic origins. I found that the M and NS segments have the most conserved NCRs, which also correlates with their slowest coding region evolutionary rates (Bahl, Vijaykrishna et al. 2009, Qu, Zhang et al. 2011). Overall, I showed that the variability of the genomic NCRs of IAVs are closely related to the viral host and geographic origins. It appears that NCRs evolve closely

with the coding region since their clustering is similar to the phylogeny of the coding regions.

5.4 The role of gulls and terns in IAV transmission

I undertook investigations of a variety of IAVs isolated from gull and tern species and demonstrated that they carry and disseminate viruses from other host species between continents, and they could also play an important role for the generation of highly pathogenic viruses (Chapter 4). The evolutionary history of an entirely Eurasian gull virus isolated in Newfoundland was retraced. The Caspian Sea, in southwestern Asia, was determined to be the potential origin of this virus, a region which also appears to have given rise to other H14 gull and duck viruses (Boyce, Schobel et al. 2013). It was concluded that different factors related to gull ecology, such as their cosmopolitan distributions, migratory behaviors, their associations with human habitats, and their use of both marine and freshwater environments, could significantly contribute to the perpetuation and emergence of enzootic, epizootic and pandemic IAVs.

5.5 Conclusion and future perspectives

Overall, my thesis findings provided a better comprehension of the genomic structure of IAVs at the quasispecies level from different wild bird hosts. It also brought further insight into the evolutionary mechanisms of the virus in its reservoir hosts. I highlighted the importance of certain bird hosts in IAV ecology for the emergence of potential pathogenic viruses and for the perpetuation of the virus in different hosts. The use of next-generation sequencing methods to fully characterize

the virus genome at the quasispecies level should be implemented in future monitoring research programs to screen IAVs in different hosts to be able to give a better insight on the viral dynamics in term of evolution and transmission. Additional evolutionary experiments performed *in vivo* in different birds with the application of high throughput sequencing methods could be beneficial to understand how the virus is able to evolve and be transmitted to other hosts. Our in-depth genomic analysis could also be extended to transcriptomic studies (characterizing all RNA molecules that are transcribed) for different bird hosts in the presence and absence of the virus. Different viruses could be tested in different hosts. This could highlight factors that could be specifically involved in some hosts during virus infection.

5.6 References

Asghar, N., J. H. Pettersson, P. Dinnetz, A. Andreassen and M. Johansson (2017). "Deep sequencing analysis of tick-borne encephalitis virus from questing ticks at natural foci reveals similarities between quasispecies pools of the virus." The Journal of General Virology. 98(3): 413-421.

Bahl, J., D. Vijaykrishna, E. C. Holmes, G. J. Smith and Y. Guan (2009). "Gene flow and competitive exclusion of avian influenza A virus in natural reservoir hosts." Virology 390(2): 289-297.

Benkaroun, J., G. J. Robertson, H. Whitney and A. S. Lang (2018). "Analysis of the Variability in the Non-Coding Regions of Influenza A Viruses." 5(3).

Benkaroun, J., D. Shoham, A. N. K. Kroyer, H. Whitney and A. S. Lang (2016). "Analysis of influenza A viruses from gulls: An evaluation of inter-regional movements and interactions with other avian and mammalian influenza A viruses." Cogent Biology 2(1): 1234957.

Boyce, W. M., S. Schobel, V. G. Dugan, R. Halpin, X. Lin, D. E. Wentworth, L. L. Lindsay, E. Mertens and M. Plancarte (2013). "Complete Genome Sequence of a Reassortant H14N2 Avian Influenza Virus from California." Genome Announcements 1(4): e00543-00513.

Croville, G., S. M. Soubies, J. Barbieri, C. Klopp, J. Mariette, O. Bouchez, C. Camus-Bouclainville and J. L. Guerin (2012). "Field monitoring of avian influenza viruses: whole-genome sequencing and tracking of neuraminidase evolution using 454 pyrosequencing." Journal of Clinical Microbiology. 50(9): 2881-2887.

Desselberger, U., V. R. Racaniello, J. J. Zazra and P. Palese (1980). "The 3' and 5'-terminal sequences of influenza A, B and C virus RNA segments are highly conserved and show partial inverted complementarity." Gene 8(3): 315-328.

Dlugolenski, D., L. Jones, G. Saavedra, S. M. Tompkins, R. A. Tripp and E. Mundt (2011). "Passage of low-pathogenic avian influenza (LPAI) viruses mediates rapid genetic adaptation of a wild-bird isolate in poultry." Archives of Virology. 156(4): 565-576.

Hossain, M. J., D. Hickman and D. R. Perez (2008). "Evidence of Expanded Host Range and Mammalian-Associated Genetic Changes in a Duck H9N2 Influenza Virus Following Adaptation in Quail and Chickens." PLoS One 3(9): e3170.

Huang, Y., G. J. Robertson, D. Ojkic, H. Whitney and A. S. Lang (2014). "Diverse inter-continental and host lineage reassortant avian influenza A viruses in pelagic seabirds." Infect Genet Evol 22: 103-111.

Huang, Y., M. Wille, J. Benkaroun, H. Munro, A. L. Bond, D. A. Fifield, G. J. Robertson, D. Ojkic, H. Whitney and A. S. Lang (2014). "Perpetuation and reassortment of gull influenza A viruses in Atlantic North America." Virology 456-457: 353-363.

Huang, Y., M. Wille, A. Dobbin, G. J. Robertson, P. Ryan, D. Ojkic, H. Whitney and A. S. Lang (2013). "A 4-year study of avian influenza virus prevalence and subtype diversity in ducks of Newfoundland, Canada." Can J Microbiol 59(10): 701-708.

Laskus, T., J. Wilkinson, J. F. Gallegos-Orozco, M. Radkowski, D. M. Adair, M. Nowicki, E. Operskalski, Z. Buskell, L. B. Seeff, H. Vargas and J. Rakela (2004). "Analysis of hepatitis C virus quasispecies transmission and evolution in patients infected through blood transfusion." Gastroenterology 127(3): 764-776.

Li, J. and C. J. Cardona (2010). "Adaptation and transmission of a wild duck avian influenza isolate in chickens." Avian Diseases. 54(1 Suppl): 586-590.

Li, J., H. zu Dohna, N. L. Anchell, S. C. Adams, N. T. Dao, Z. Xing and C. J. Cardona (2010). "Adaptation and transmission of a duck-origin avian influenza virus in poultry species." Virus Research. 147(1): 40-46.

Qu, Y., R. Zhang, P. Cui, G. Song, Z. Duan and F. Lei (2011). "Evolutionary genomics of the pandemic 2009 H1N1 influenza viruses (pH1N 1v)." Virology Journal 8: 250.

Robertson, J. S. (1979). "5' and 3' terminal nucleotide sequences of the RNA genome segments of influenza virus." Nucleic Acids Research 6(12): 3745-3757.

Appendices

Supplementary documents in chapter 2

Supplementary table 2.1: Viruses sequenced and analyzed.

Virus*	Bird host
A/American black duck/Newfoundland/GR256/2011(H3N2) A/domestic duck/Newfoundland/MW668/2010(H1N1) A/American black duck/Newfoundland/GR679/2011(H3N2)	Duck
A/herring gull/Newfoundland/GR658/2011(H13N6) A/herring gull/Newfoundland/GR848/2011(H13N6) A/herring gull/Newfoundland/GR657/2011(H13N6)	Gull
A/common murre/Newfoundland/AB324/2011(H1N2) A/common murre/Newfoundland/AB319/2011(H1N2) A/common murre/Newfoundland/AB327/2011(H1N2)	Murre

***: All viruses were initially identified in paired of oropharyngeal and cloacal swabs in Newfoundland and Labrador, Canada by q-PCR screening for IAVs.**

Supplementary table 2.2: Pairwise nucleotide sequence identities between the different viruses.

		Gull viruses		
		GR654	GR848	GR658
PB2	GR654		99.211	99.956
	GR848	99.211		99.167
	GR658	99.956	99.167	
PB1	GR654		92.216	99.604
	GR848	92.216		91.821
	GR658	99.604	91.821	
PA	GR654		97.977	98.897
	GR848	97.977		99.070
	GR658	98.897	99.070	
HA	GR654		97.977	99.941
	GR848	97.977		97.941
	GR658	99.941	97.941	
NA	GR654		98.868	99.929
	GR848	98.868		98.338
	GR658	99.929	98.338	
NP	GR654		91.427	99.672
	GR848	91.427		91.015
	GR658	99.672	91.015	
M	GR654		99.084	99.896
	GR848	99.084		99.084
	GR658	99.896	99.084	
NS	GR654		99.761	99.881
	GR848	99.761		99.642
	GR658	99.881	99.642	
		Murre viruses		
		AB327	AB341	AB319
PB2	AB327		81.123	83.991
	AB341	81.123		99.079
	AB319	83.991	99.079	
PB1	AB327		96.336	95.954
	AB341	96.336		99.120
	AB319	95.954	99.120	
PA	AB327		99.535	94.747
	AB341	99.535		94.654

	AB319	94.747	94.654	
HA	AB327		99.118	99.059
	AB341	99.118		99.558
	AB319	99.059	99.558	
NA	AB327		99.362	99.078
	AB341	99.362		99.291
	AB319	99.078	99.291	
NP	AB327		99.733	99.332
	AB341	99.733		99.529
	AB319	99.332	99.529	
M	AB327		90.937	90.733
	AB341	90.937		99.491
	AB319	90.733	99.491	
NS	AB327		99.761	100.000
	AB341	99.761		99.761
	AB319	100.000	99.761	
		Duck viruses		
		MW668	GR679	GR256
PB2	MW668		97.065	97.379
	GR679	97.065		96.960
	GR256	97.379	96.960	
PB1	MW668		95.492	95.427
	GR679	95.492		95.765
	GR256	95.427	95.765	
PA	MW668		88.300	92.413
	GR679	88.300		90.037
	GR256	92.413	90.037	
HA	MW668		N/A	N/A
	GR679	N/A*		99.530
	GR256	N/A	99.530	
NA	MW668		N/A	N/A
	GR679	N/A		99.823
	GR256	N/A	99.823	
NP	MW668		92.184	93.387
	GR679	92.184		97.195
	GR256	93.387	97.195	
M	MW668		96.068	96.281
	GR679	96.068		97.556
	GR256	96.281	97.556	
NS	MW668		97.744	97.870
	GR679	97.744		99.875
	GR256	97.870	99.875	

* N/A: Comparison not applicable since both viruses have a different subtype.

Supplementary table 2.3: Type of mutations present in the HA of the duck virus from passage 5.

Nucleotide change	Protein effect	Amino acid change	Amino acid position	Length (amino acid)
TG -> AA	Substitution	...C...	492	1
TC -> CA	Substitution	...H...	486	1
GG -> CA	Substitution	L...	477	1
	Deletion	...RV	465	2
	Deletion	...HDSNVR...	455	6
GTTC -> ACAG	Substitution	V...	444	1
GCT -> CGA	Substitution	...L...	443	1
CTGAAC -> ACACTG	Substitution	...E...	441	1
AT -> CA	Substitution	...N	439	439
AC -> CT	Substitution	...Y	438	438
TTTGG -> CAAAC	Substitution	...W	436	436
CTGG -> TACC	Substitution	L...	433	433
TGAGAAT -> CACTGTA	Substitution	...K	426	426
GG -> CT	Substitution	...W	358	358
GG -> CC	Substitution	G...	357	357
TT -> AA	Substitution	...I	354	354
TG -> CA	Substitution	...A...	351	351
AG -> CA	Substitution	...A...	349	349
GGACT -> TTCAA	Substitution	G...	348	348
TCCAGA -> CTGTTT	Substitution	S	343	343
TCCATTC -> GAAGCAA	Substitution	S...	340	340
CT -> TA	Substitution	L...	88	88
GGATAGC -> CTTCGAG	Substitution	...I...	64	64
AA -> GT	Substitution	N...	55	55

Supplementary table 2.4: Type of mutations present in the HA of the duck virus from passage 10.

Nucleotide change	Protein effect	Amino acid change	Amino acid position	Length (amino acid)
TG -> AA	Substitution	...V...	496	1
AT -> TG	Substitution	M...	493	1
GT -> AC	Substitution	...C	492	1
TC -> CA	Substitution	...H...	486	1
GG -> CA	Substitution	G...	478	1
AA -> CT	Substitution	...E	476	1
CC -> GG	Substitution	...A	474	1
	Deletion	...ERVKSQ	464	6
	Deletion	...LDFHDSNVRNL...	452	11
	Deletion	...R	450	1
GTTC -> ACAG	Substitution	V...	444	1
GCT -> CGA	Substitution	...L...	443	1
CTGAAC -> ACACTG	Substitution	...E...	441	1
AT -> CA	Substitution	...N	439	1
AC -> CT	Substitution	...Y	438	1
CA -> AC	Substitution	...T	437	1
TTTGG -> CAAAC	Substitution	...W	436	1
CTGG -> TACC	Substitution	L...	433	1
CA -> GG	Substitution	...N...	429	1
AA -> CT	Substitution	N...	425	1
TGAGAAT -> CACTGTA	Substitution	...EN	422	2
GG -> CT	Substitution	...W	358	1
GG -> CC	Substitution	G...	357	1
AA -> CT	Substitution	...E	355	1
TT -> AA	Substitution	...I	354	1
TG -> CA	Substitution	...A...	351	1
AG -> CA	Substitution	...A...	349	1
TG -> CT	Substitution	...G...	348	1
TCCAGA -> CTGTTT	Substitution	SR	343	2
TCCATTC -> GAAGCAA	Substitution	SI...	340	2
TCCA -> GATG	Substitution	...P	339	1
AG -> TC	Substitution	R...	336	1
GGACT -> TTCAA	Substitution	G...	334	1
AT -> CA	Substitution	M...	331	1
CAT -> GTG	Substitution	...W...	93	1
CT -> TA	Substitution	L...	88	1
GGATAGC -> CTTCGAG	Substitution	...I...	64	1

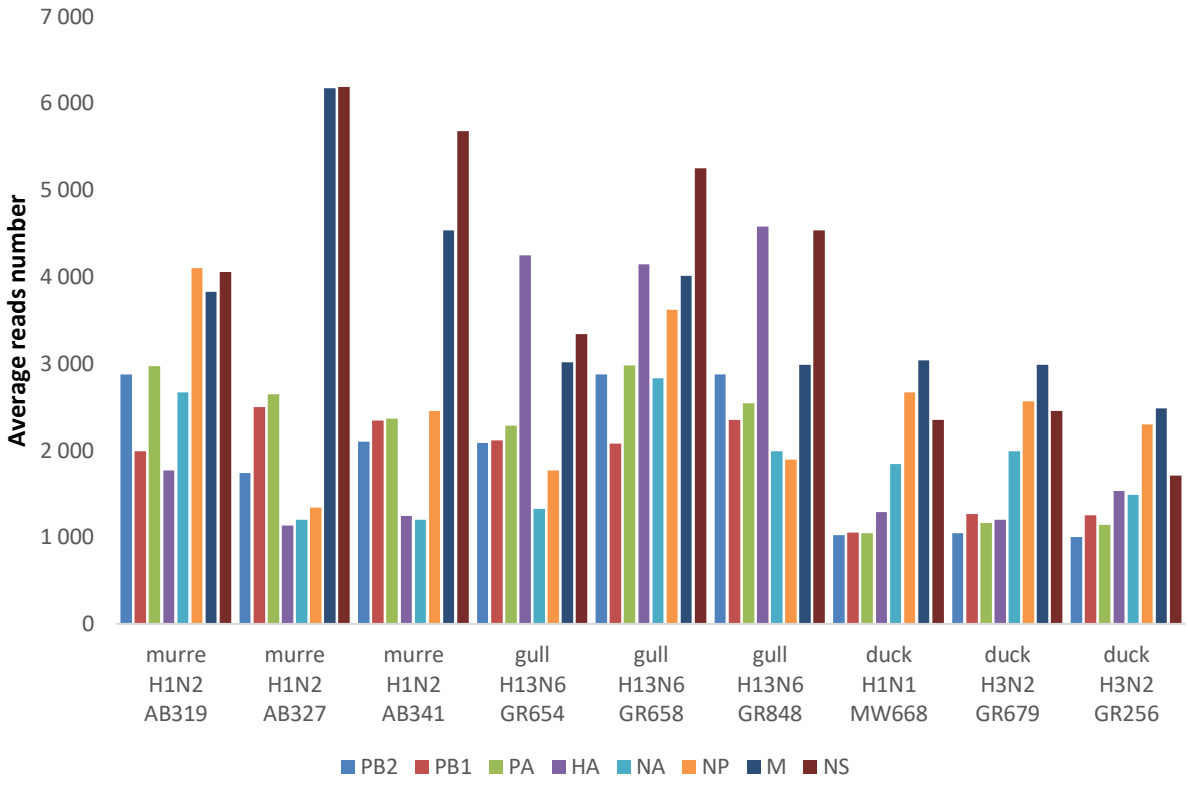
AA -> GT	Substitution	N...	62	1

Supplementary table 2.5: Type of mutations present in the NA of the duck virus from passage 5.

Nucleotide change	Amino acid position	Length (aa)	Amino Acid change	CDS Position	Codon Change	Protein Effect
GGTGT -> CAAAC	356	1	GV -> QT	1,066	GGT,GTT -> CAA,ACT	Substitution
GGT -> TCA	353	1	G -> S	1,060	GGT -> TCA	Substitution
GG -> AA	347	1	G -> K	1,033	GGG -> AAG	Substitution
GGGGCA -> ACAACT	342	1	GA -> TT	1,024	GGG,GCA -> ACA,ACT	Substitution
TC -> GG	341	1	S -> G	1,018	TCT -> GGT	Substitution
GTA -> AAG	338	1	V -> K	1,012	GTA -> AAG	Substitution
CCA -> AAT	336	1	P -> N	1,009	CCA -> AAT	Substitution
ACA -> CAT	334	1	T -> H	994	ACA -> CAT	Substitution
TCTCTC -> GGAAGT	89	3	SL -> GS	268	TCT,CTC -> GGA,AGT	Substitution
CA -> AT	85	1	GN -> GY	261	GGC,AAT -> GGA,TAT	Substitution
TGCAG -> GAAGC	82	1	IAE -> MKQ	225	ATT,GCA,GAA -> ATG,AAG,CAA	Substitution
GTAACA -> ACTGAT	74	1	VT -> TD	247	GTA,ACA -> ACT,GAT	Substitution

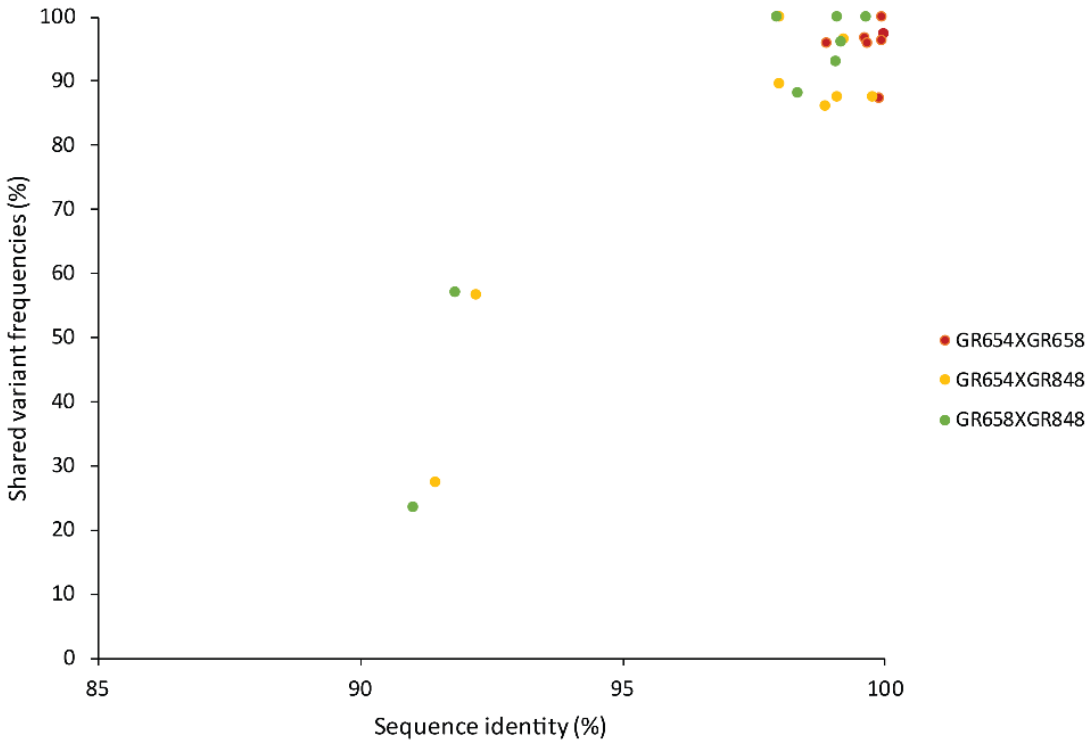
Supplementary table 2.6: Type of mutations present in the NA of the duck virus from passage 10.

Nucleotide change	Amino acid position	Length (aa)	Amino Acid change	CDS Position	Codon Change	Protein Effect
TA -> CG	385	1	DN -> DD	1,152	GAT,AAC -> GAC,GAC	Substitution
GGTGT -> CAAAC	356	1	GV -> QT	1,066	GGT,GTT -> CAA,ACT	Substitution
GGT -> TCA	354	1	G -> S	1,060	GGT -> TCA	Substitution
ATAC -> TCAT	353	1	RY -> SH	1,056	AGA,TAC -> AGT,CAT	Substitution
GG -> AA	345	1	G -> K	1,033	GGG -> AAG	Substitution
GGGGCA -> ACAACT	342	1	GA -> TT	1,024	GGG,GCA -> ACA,ACT	Substitution
AA -> GT	341	1	N -> V	1,021	AAT -> GTT	Substitution
TC -> GG	340	1	S -> G	1,018	TCT -> GGT	Substitution
GTA -> AAG	338	1	V -> K	1,012	GTA -> AAG	Substitution
CCA -> AAT	337	1	P -> N	1,009	CCA -> AAT	Substitution
GG -> AA	336	1	G -> N	1,006	GGT -> AAT	Substitution
ACA -> CAT	332	1	T -> H	994	ACA -> CAT	Substitution
GTGGA -> CCCAG	96	1	SG -> TQ	284	AGT,GGA -> ACC,CAG	Substitution
CC -> GTT	93	1		277		Frame Shift
TCTCTC -> GGAAGT	90	2	SL -> GS	268	TCT,CTC -> GGA,AGT	Substitution
TC -> AG	89	1		265	TCC -> AGC	None
CA -> AT	88	1	GN -> GY	261	GGC,AAT -> GGA,TAT	Substitution
CTAG -> TGGT	85	1	LA -> WS	253	CTA,GCA -> TGG,TCA	Substitution
GTAACA -> ACTGAT	83	2	VT -> TD	247	GTA,ACA -> ACT,GAT	Substitution
TGCAG -> GAAGC	76	1	IAE -> MKQ	225	ATT,GCA,GAA -> ATG,AAG,CAA	Substitution

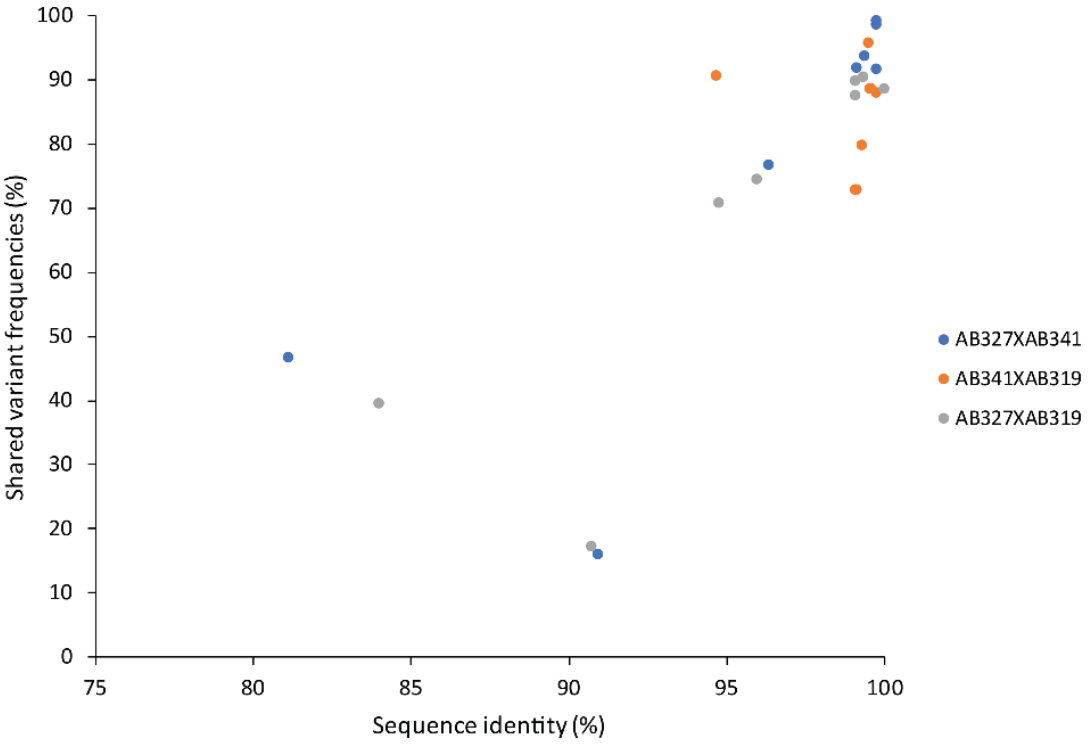


Supplementary figure 2.1: Average coverage for each segment from the IAV genomic sequencing.

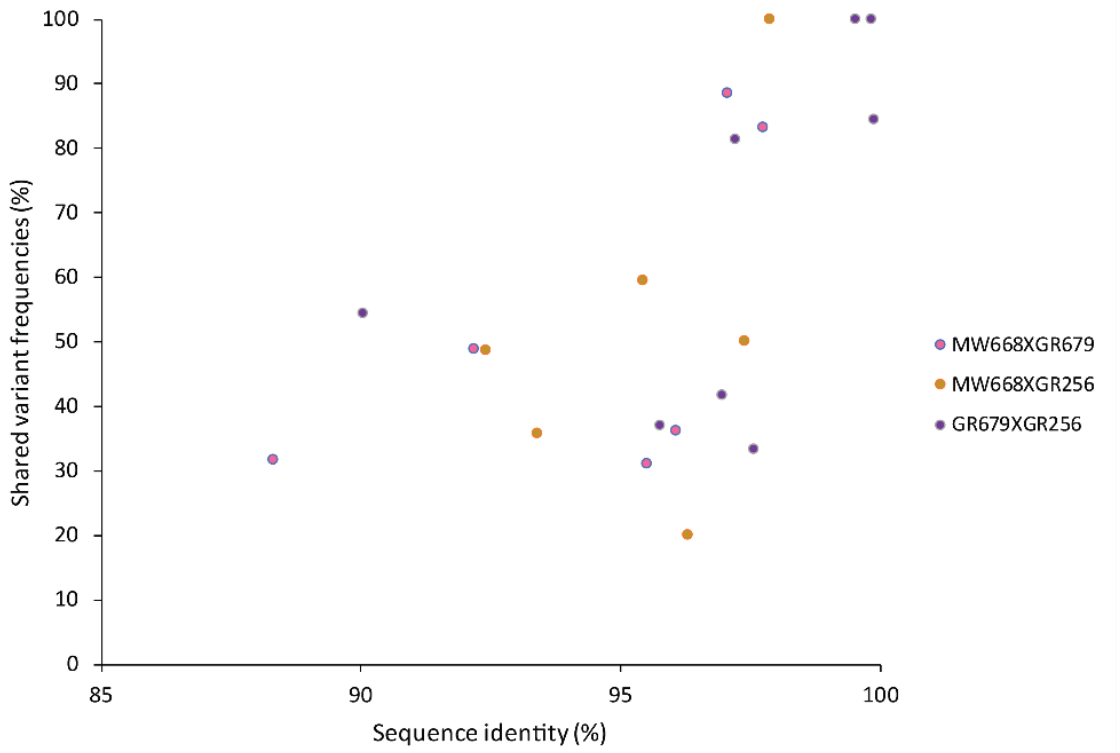
A)


Shared variant frequencies (%)
100
90
80
70
60
50
40
30
20
10
0
85
90
95
100
Sequence identity (%)
GR654XGR658
GR654XGR848
GR658XGR848

B)


Shared variant frequencies (%)
100
90
80
70
60
50
40
30
20
10
0
75
80
85
90
95
100
Sequence identity (%)
AB327XAB341
AB341XAB319
AB327XAB319

C)



Supplementary figure 2.2: Relationships between segment shared variants and sequence homology. The scatter plot shows the two values for each pair of segments compared within the different hosts: A) gull viruses, B) murre viruses, and C) duck viruses.

Supplementary documents in chapter 3

Supplementary table 3.1: List of RACE primers used.

Primer	Primer sequence (5' to 3')
3’ RACE	
3CM-PB2-1	CATTAACGGAGCAATCTTGCAG
3CM-PB2-2	GTGATACCATTACTCTGTCTGATCC
3HG-PB2-1	CCGGTAGGAATCTGGTCTTGC
3HG-PB2-2	ACACCATCACCCTATCAGAACC
3DD-PB2-1	ACTGCTTGTTCCACCAGCC
3DD-PB2-2	CACGGCCAGAGGTGATACC
3-CM-PB1-1	CCGGCTGATTCCTGTTTAGTG
3-PB1-2	ATATCCACTTGGCTCATTGTCC
3-HGDD-PB1-1	CCGGTTGATTCCTGTTTAATGTC
3-PA-1	CAATGCAGCCGTTCGGTT
3-CM-PA-2	ATTATCGATTCACCTCGTTCATC
3-HG-PA-2	CAGATGCTATTCACCACAGTCC
3-DD-PA-2	GATTCCACAATTATTGATTCGCC
3-CM-HA-1	GTGTCCAATAATAGTTCATCCTGCC
3-HG-HA-1	CGGGTGGAAACTAGTGTATAAGG
3-DD-HA-1	GTGTCCAGTAATAGTTCATCCTGCC
3-CMDD-HA-2	TTCAGGCTGCAGAGCTTCC

3-HG-HA-2	CTCCATTCAAAGAACAGTATGTTCC
3-CM-NA-1	CATGGCAACTTGAGCTGGAC
3-HG-NA-1	ATGACATGAGGTGCTGGACC
3-DD-NA-1	CAATTGTCAACCAGCTAATGCC
3-CM-NA-2	CGATGTGGAATTCTATCATGTATTG
3-HG-NA-2	CTCTCCTATTCTAATTGCATTGTCC
3-DD-NA-2	GCCTGCTAATGTTATTGGAGCC
3-CMDD-NP-1	GTGGCATCATTCAGATTGGAAT
3-HG-NP-1	GTGGCGTCATTCAAGTTAGAATG
3-HG-NP-2	CTGAGCTTGAGTTCGGTGCAC
3-DD-NP-2	GATGTTCCTCAAGGTATCTGTTCC
3-CM-NP-2	CCAGATACTTGTTTCTCCTCTCATC
3-M-1	GCTCACAAGTGGCACATACTAGAC
3-M-2	CTACGCTGCAGTCCTCGCT
3-NS-1	AGTATGTCCTGGAAGAGAAGG
3-CMDD-NS-2	CAGTTAGGTAGCGTGAAGTAGGC
3-HG-NS-2	CTCTTGACATCTCCTCTGGAGAC

5’ RACE

5-CM-PB2-1	GTAAGGACGCTATTCCAGCAGAT
5-HG-PB2-1	CAAGGCTGCTAGAGGCCAATAC
5-DD-PB2-1	CGTGAGGACACTATTCCAGCAG
5-CM-PB2-2	GGAGTGGAATCTGCGGTATTAAG

5-HG-PB2-2	GGATGCAGGTTCATTGACAGAAG
5-DD-PB2-2	GCATTGACTGAAGATCCAGATGA
5-CM-PB2-3	GATATGGACCAGCATTGAGCAT
5-HG-PB2-3	GCATCAACGAACTGAGCAATC
5-DD-PB2-3	GGAGAGAAGGCTAATGTGCTAATTG
5-PB1-1	CATGAGCATTGGAGTAACAG
5-PB1-2	GACACACAAATTCAGACAAG
5-PB1-3	CCACAAAGAGATAGAGTCTGT
5-CM-PA-1	GTATGTGAGGACCAATGGAACTTC
5-HG-PA-1	GAGAGCATGATTGAAGCTGAATC
5-DD-PA-1	GACCGCAATAGGCCAAGTATC
5-CM-PA-2	GGAGTAGAGGAAGGTTCCATTG
5-HG-PA-2	GGCTCCATCGGTAAAGTGTG
5-DD-PA-2	GGAAGGCTCTATTGGGAAGGTAT
5-PA-3	CTCATTGTTCAGGCACTTAG
5-CMDD-HA-1	GAATTCAACAACCTAGAGAG
5-HG-HA-1	GAGTTCAGCCAGGTAGAA
5-CMDD-HA-2	GAGTTCTATCACAAGTGTGAT
5-HG-HA-2	GAGCTCCTTCATAAATGTAAT
5-CM-HA-3	CGGAGTGAAACTAGAGTCGATGG
5-HG-HA-3	GCTGAATATGCAGAAGAATCG
5-DD-HA-3	CGGCACATATGACTATCCCAAATAT

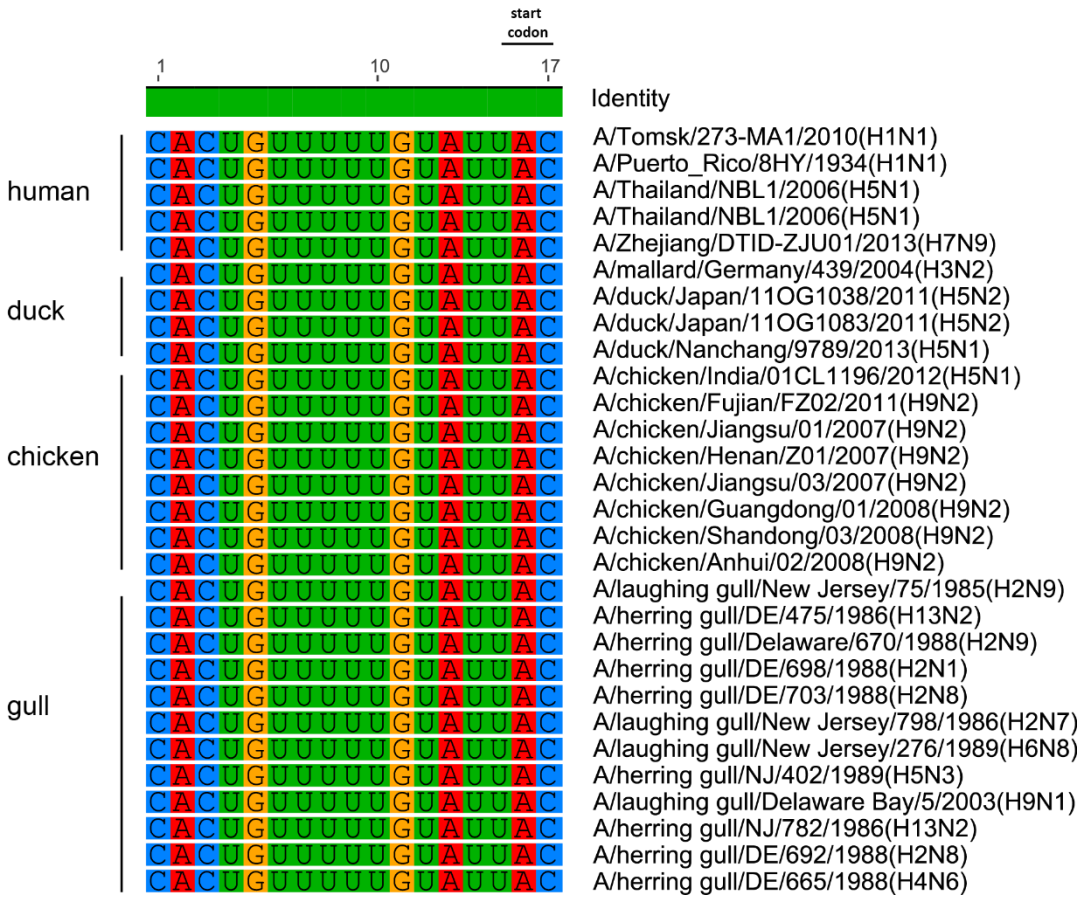
5-CM-NA-1	GGAACGATGATAGCTCTAG
5-HG-NA-1	GCAAATACTTATGTTCAAAG
5-DD-NA-1	GTACTAGCTCAAGGAGTGG
5-CM-NA-2	CGGCTAATTCCAAGTCACAGG
5-HG-NA-2	GCAGAAACAGACAATCAGTCCG
5-DD-NA-2	GGATGGACAGAGACGGACG
5-CM-NA-3	GGTGGACTTCAAATAGTATCGTC
5-HG-NA-3	GGAATTAATCAGAGGCAGGCC
5-DD-NA-3	GAGCTAATCAGAGGAAGACCCAAG
5-NP-1	GTGTCAAGCTTCATCAGA
5-CM-NP-2	GAGAGAGCAACCATTATGGCA
5-HG-NP-2	GGAGGAAACACCAACCAAC
5-DD-NP-2	GGAAACACTGAAGGCAGAACTTC
5-NP-3	CTGAGATCATAAGGATGATGG
5-M-1	GTGCCACTTGTGAGCAGATTG
5-M-2	GGCTAGACAGATGGTGCA
5-M-3	GCAGCGATTCAAGTGATCCTC
5-CMDD-NS-1	GCCTTCTCTTCCAGGAC
5-HG-NS-1	GACATGTCTCCAGAGGAG
5-CMDD-NS-2	GGACTTGAATGGAATGATAACACAG
5-HG-NS-2	CTCACCATTACCTTCTCTTCCAGG
5-CMDD-NS-3	CGCTTGGAGAAGCAGTGATG

5-HG-NS-3	GGAACAAGAACTGGAGAGAACAAT

The abbreviations HG, CM, and DD refer to the avian viruses: herring gull, common murre, and duck respectively.

Supplementary table 3.2: Number of NCRs analyzed by host.

	Swine	Duck	Human	Poultry	Murre	Shorebird	Swan/Goose	Gull	Equine
H1	191	200	380	11	26	36			
H9		38		33			4	3	
H13		2						41	
N6			10					14	
NP		29	16	23				36	
M	16	12	18						7
NS		16	10	12				15	



Supplementary figure 3.1: NS 3' NCRs from different host species and locations.

Hosts of the viruses are indicated on the left.

Supplementary documents in chapter 4

Supplementary table 4.1: Pairwise genetic distance[a] matrices comparing gene segments of A/herring gull/Newfoundland/YH019/2010(H16N3) to segments from other related viruses.

	Virus[b]/Segment	Location	Virus number[c]		
	PB2		**1**	**2**	**3**
1	A/herring_gull/Newfoundland/YH019/2010(H16N3)	Canada	-		
2	A/black-headed_gull/Sweden/74340/2008(H16N3)	Sweden	0.034	-	
3	A/glaucous_gull-herring_gull_hybrid/Iceland/1108/2011(H16N3)	Iceland	0.042	0.022	-
	PB1		**1**	**2**	
1	A/herring_gull/Newfoundland/YH019/2010(H16N3)	Canada	-		
2	A/common_gull/Norway/10_1313/2009(H13N2)	Norway	0.019	-	
	PA		**1**	**2**	
1	A/herring gull/Newfoundland/YH019/2010(H16N3)	Canada	-		
2	A/black-headed gull/Iceland/713/2010(H16N3)	Iceland	0.009	-	
	HA		**1**	**2**	**3**

1	A/herring_gull/Newfoundland/YH019/2010(H16N3)	Canada	-		
2	A/black-headed_gull/Iceland/713/2010(H16N3)	Iceland	0.013	-	
3	A/duck/Hokkaido/WZ82/2013(H16N3)	Japan	0.021	0.011	-
4	A/Herring_Gull/Norway/10_2298/2007(H16)	Norway	0.029	0.019	0.027

NP			**1**	**2**
1	A/herring_gull/Newfoundland/YH019/2010(H16N3)	Canada	-	
2	A/black-headed_gull/Iceland/713/2010(H16N3)	Iceland	0.005	-

NA			**1**	**2**
1	A/herring_gull/Newfoundland/YH019/2010(H16N3)	Canada	-	
2	A/common_gull/Norway/10_1617/2006(H16N3)	Norway	0.067	-

M			**1**	**2**
1	A/blackhead_gull/Astrakhan/65/1983(H13N6)	Russia	-	
2	A/herring_gull/Newfoundland/YH019/2010(H16N3)	Canada	0.031	-

NS			**1**	**2**
1	A/herring_gull/Newfoundland/YH019/2010(H16N3)	Canada	-	
2	A/yellow-legged_gull/Ukraine/912306/2005(H13)	Ukraine	0.02	-

[a] Genetic distance was calculated as 1 - % identity from nucleotide alignments.

[b] Gull sequences from North America are in green, gull sequences from Eurasia are in dark blue, and waterfowl sequences from Eurasia are in red.
[c] Also corresponds to numbers in leftmost column.

Supplementary table 4.2: Pairwise genetic distances[a] for comparison of gull virus HA gene segments to related segments from different hosts.

HA	Viruses[b]	Distance[c]
	Gull-waterfowl	
H1	A/laughing_gull/Delaware_Bay/61/2002(H1N9)	
	A/muscovy_duck/New_York/21211-5/2005(H1N1)	0.025
	A/mallard/Maryland/13/2003(H1N1)	0.030
	A/gull/Delaware/428/2009(H1N1)	
	A/mallard/Nova Scotia/02151/2007(H1N1)	0.033
H3	A/glaucous_gull/Alaska/44201-161/2006(H3N8)	
	A/northern_pintail/Alaska/44204-075/2006(H3N8)	0.000
H5	A/slaty-backed gull/Shandong/59/04(H5N1)	
	A/duck/Hubei/wg/2002(H5N1)	0.006
		1
	A/Black Headed Gull/HK/12.1/2003(H5N1)	
	A/duck/Hunan/1386/2003(H5N1)	0.009

	A/gull/Italy/692-2/93(H7N2)	
	A/duck/Heinersdorf/S495/6/86(H7N7)	0.038
H7	A/yellow-legged_gull/Republic_of_Georgia/1/2012/mixed(H7)	
	A/mallard/Republic_of_Georgia/2/2010(H7N7)	0.029
	A/laughing_gull/New_York/AI00-2455/2000(H7N3)	
	A/mallard/Ohio/322/1998(H7N3)	0.012
	A/glaucous-winged_gull/Southeastern_Alaska/9JR0822R0/2009(H13N6)	0.031
	A/duck/Hokkaido/WZ68/2012(H13N2)	
	A/lesser_snow_goose/Alaska/44199-115/2006(H13N9)	
	A/glaucous_gull/Alaska/44199-097/2006(H13N3)	0.003
H13	A/black_headed_gull/Mongolia/1766/2006(H13N6)	0.019
	A/black-headed_gull/Sweden/74342/2008(H13N8)	0.030
	A/mallard/Korea/SH38-45/2010(H13N2)	
	A/black_headed_gull/Mongolia/1766/2006(H13N6)	0.045
	A/black-headed_gull/Sweden/74342/2008(H13N8)	0.046

	A/herring_gull/New_Jersey/AI09-1105/2009(H13N6)	
	A/hooded_merganser/New_Brunswick/03750/2009(H13N6)	0.007
	A/mallard-black_duck_hybrid/New_Brunswick/03736/2009(H13N6)	0.005
H16	A/gull/Denmark/68110/2002(H16N3)	0.048
	A/mallard/Quebec/02916-1/2009(H16N3)	
	A/California gull/California/1196P/2013(H16N3)	
	A/duck/Hokkaido/WZ82/2013(H16N3)	0.032
	A/black-headed gull/Iceland/713/2010(H16N3)	0.021
	A/waterfowl/GA/96623-7/2001(H16N3)	
	A/herring gull/New York/AI00-532/2000(H16N3)	0.001
	A/mallard/Gurjev/785/83(H16N3)	
	A/black-headed gull/Sweden/2/99(H16N3)	0.070
	Gull-poultry	
H5	A/slaty-backed_gull/Shandong/38/04(H5N1)	0.002
	A/chicken/Yichang/lung-1/04(H5N1)	

	A/brown-head_gull/Thailand/vsmu-4/2008(H5N1)	
	A/chicken/Thailand/ICRC-V143/2007(H5N1)	0.000
	A/chicken/Thailand/ICRC-213/2007(H5N1)	0.000
	A/black_headed_gull/HK/12.1/2003(H5N1)	
	A/chicken/Hong_Kong/3123.1/2002(H5N1)	0.007
	A/gull/Italy/692-2/93(H7N2)	
	A/turkey/Israel/Ramon/79(H7N2)	0.016
	A/chicken/Hebei/1/2002(H7N2)	0.031
	A/chicken/Jena/1816/87(H7N7)	0.037
H7		
	A/yellow-legged_gull/Republic_of_Georgia/1/2012/mixed(H7)	
	A/turkey/Italy/8000/2002(H7N3)	0.046
	A/laughing_gull/New_York/AI00-2455/2000(H7N3)	
	A/chicken/New_York/14714-9/1999(H7N3)	0.021
	A/herring_gull/Delaware/660/1988(H13N6)	
H13	A/laughing_gull/Delaware_Bay/2838/1987(H13N2)	0.019
	A/turkey/MN/1012/1991(H13N2)	0.020

	Gull-swine	
	A/laughing_gull/Delaware_Bay/61/2002(H1N9)	
H1		
	A/swine/Saskatchewan/18789/02(H1N1)	0.020
	A/slaty-backed_gull/Shandong/59/04(H5N1)	
H5		
	A/swine/Anhui/ca/2004(H5N1)	0.001
	Gull-other mammals	
	A/black_headed_gull/HK/12.1/2003(H5N1)	
H5		
	A/Hong_Kong/213/2003(H5N1)	0.009
	A/gull/Minnesota/945/1980(H13N6)	
H13		
	A/pilot_whale/Maine/328_HN/1984(H13N2)	0.042
	Gull-gull	
	A/gull/Astrakhan/3483/2002(H13N6)	
	A/glaucous-winged_gull/Southcentral_Alaska/9JR0738R0/2009(H13N6)	0.036
H13		
	A/glaucous-winged_gull/Southcentral_Alaska/9JR0738R0/2009(H13N6)	
	A/great_black-headed_gull/Atyrau/773/2004(H13N6)	0.032

	A/black_headed_gull/Mongolia/1766/2006(H13N6)	
	A/glaucous_gull/Alaska/44199-097/2006(H13N3)	0.022
	A/black-headed_gull/Sweden/74342/2008(H13N8)	
	A/glaucous_gull/Alaska/44199-097/2006(H13N3)	0.033
	A/gull/Astrakhan/3483/2002(H13N6)	
	A/glaucous-winged_gull/Southcentral_Alaska/9JR0738R0/2009(H13N6)	0.036
	A/great_black-headed_gull/Atyrau/773/2004(H13N6)	
	A/glaucous-winged_gull/Southcentral_Alaska/9JR0738R0/2009(H13N6)	0.032
	A/glaucous gull-herring gull hybrid/Iceland/1108/2011(H16N3)	
H16		
	A/California gull/California/1196P/2013(H16N3)	0.023

[a] Genetic distance was calculated as 1 - % identity from nucleotide alignments.
[b] Gull sequences from North America are in green, gull sequences from Eurasia are in blue, waterfowl sequences from Eurasia are in red, waterfowl sequences from North America are in orange, swine sequences are in pink, other mammalian sequences are in dark blue, and poultry sequences are in brown.
[c] The distance values for each virus relative to the top listed virus sequence within groups is given.

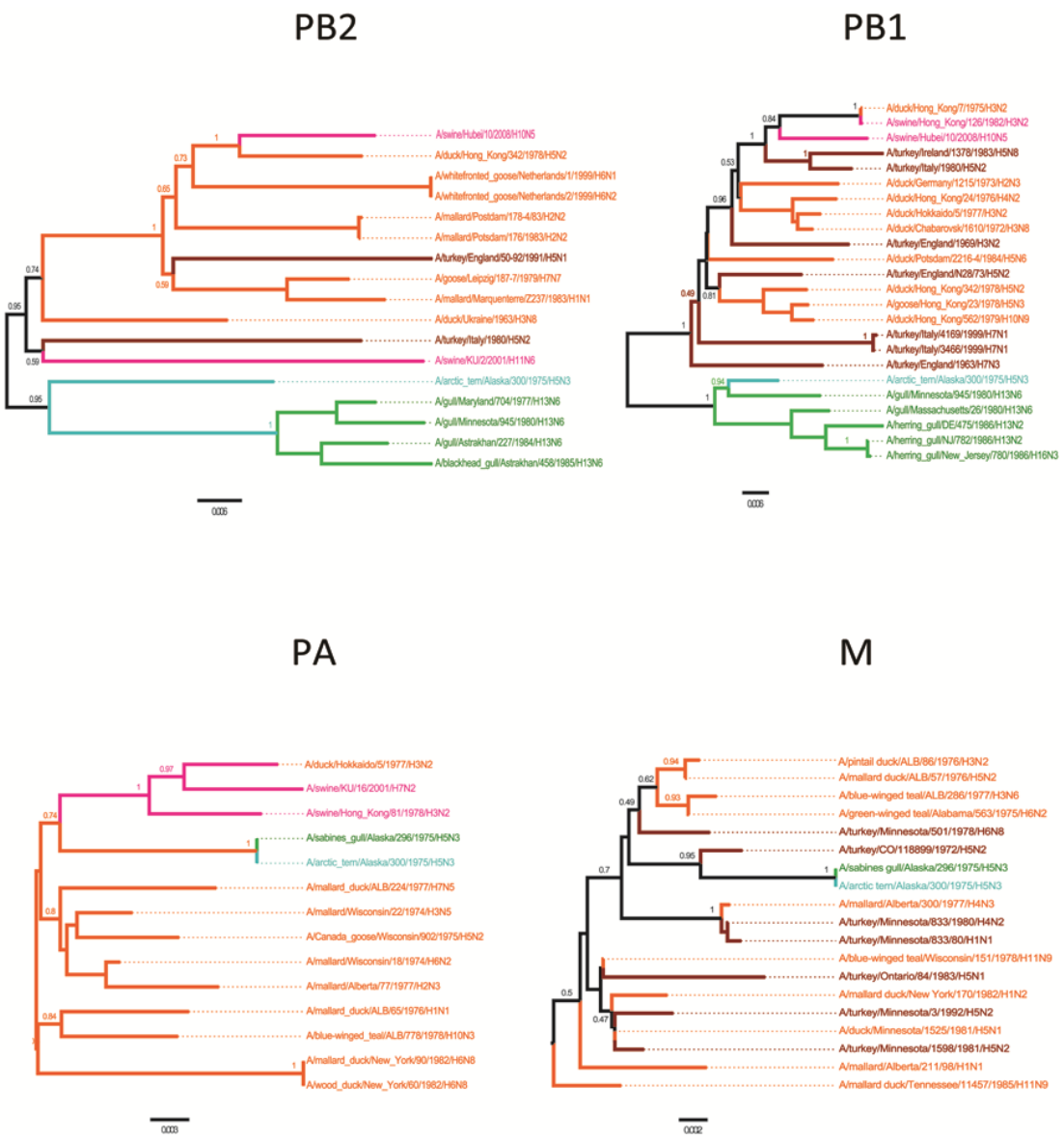
Supplementary table 4.3: Pairwise distances[a] between tern virus sequences and viruses from various hosts.

Segment	Viruses[b]	Distance[c]
	A/arctic_tern/Alaska/300/1975(H5N3)	
	A/gull/Astrakhan/227/1984(H13N6)	0.074
PB2	A/blackhead_gull/Astrakhan/458/1985(H13N6)	0.084
	A/gull/Minnesota/945/1980(H13N6)	0.078
	A/gull/Maryland/704/1977(H13N6)	0.075
	A/arctic_tern/Alaska/300/1975(H5N3)	
	A/gull/Minnesota/945/1980(H13N6)	0.032
	A/tern/South_Africa/1961(H5N3)	
PB1	A/turkey/England/1963(H7N3)	0.047
	A/turkey/England/N28/1973(H5N2)	0.051
	A/tern/South_Africa/1961(H5N3)	
	A/seal/Massachusetts/3911/1992(H3N3)	0.070
PA	A/arctic_tern/Alaska/300/1975(H5N3)	

	A/sabines_gull/Alaska/296/1975(H5N3)	0.000
HA	A/common_tern/Buryatiya/1901/2000(H4N6)	
	A/tufted_duck/Buryatuya/1905/2000(H4N6)	0.003
	A/wild_duck/Taiwan/WB2/1998(H4N6)	0.013
NP	A/tern/Turkmenia/18/1972(H3N3)	
	A/whale/Pacific ocean/19/1976(H1N3)	0.053
	A/tern/South_Africa/1961(H5N3)	0.054
NA	A/tern/Astrakan/775/83(H13N3)	
	A/sabines_gull/Alaska/296/1975(H5N3)	0.123
	A/herring_gull/New_Jersey/780/1986(H16N3)	0.142
	A/herring_gull/Delaware_Bay/2617/1987(H16N3)	0.142
M	A/common_tern/Buryatiya/1901/2000(H4N6)	
	A/tufted_duck/Buryatiya/1905/2000(H4N6)	0.001
	A/arctic_tern/Alaska/300/1975(H5N3)	
	A/sabines_gull/Alaska/296/1975(H5N3)	0
	A/turkey/CO/118899/1972(H5N2)	0.012
NS	A/common_tern/Korgalzhyn/847/2004(H3N6)	

	A/mallard/Sweden/101993/2009(H2N9)	0.010

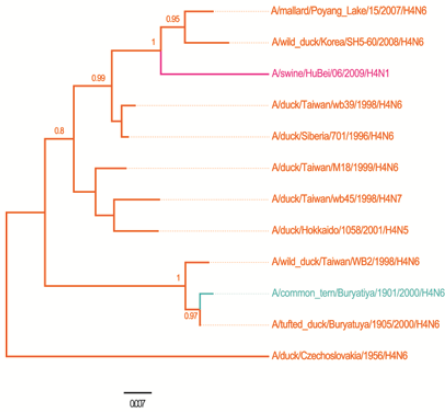
[a] Genetic distance was calculated as 1 - % identity from nucleotide alignments.
[b] Tern sequences are in light blue, gull sequences from North America are in green, gull sequences from Eurasia are in blue, waterfowl sequences from Eurasia are in red, waterfowl sequences from North America are in orange, mammalian sequences are in dark blue, and poultry sequences are in brown.
[c] The distance values for each virus relative to the top listed virus sequence within groups is given.



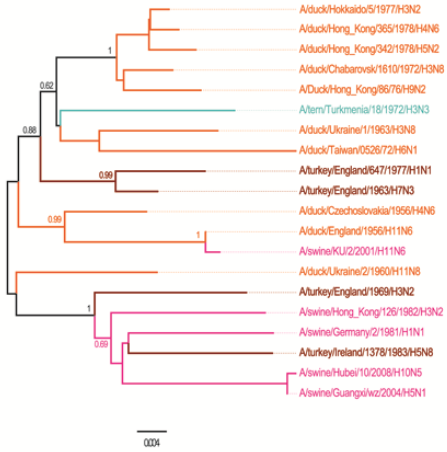
Supplementary figure 4.1. Relationship of Alaskan Arctic tern viruses to gull viruses from North America. Sequences showing the highest similarity in BLAST searches to tern gene segments were aligned by MUSCLE with MEGA6. Phylogenetic trees were inferred with the neighbor-joining method with 1000 bootstrap replications. Evolutionary distances were computed using the Maximum Composite Likelihood method and are represented by the number of base substitutions per site (scale bar).

The branches are coloured according to the host group: green, gulls; brown, poultry; light blue, terns; orange, waterfowl; pink, swine.

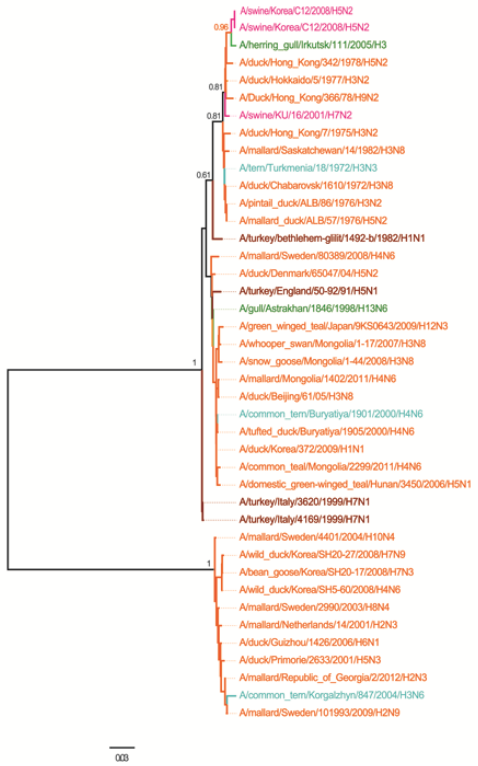
HA


A/mallard/Poyang_Lake/15/2007/H4N6
A/wild_duck/Korea/SH5-60/2008/H4N6
A/swine/HuBei/06/2009/H4N1
A/duck/Taiwan/wb39/1998/H4N6
A/duck/Siberia/701/1996/H4N6
A/duck/Taiwan/M18/1999/H4N6
A/duck/Taiwan/wb45/1998/H4N7
A/duck/Hokkaido/1058/2001/H4N5
A/wild_duck/Taiwan/WB2/1998/H4N6
A/common_tern/Buryatiya/1901/2000/H4N6
A/tufted_duck/Buryatuya/1905/2000/H4N6
A/duck/Czechoslovakia/1956/H4N6
0.95
1
0.99
0.8
1
0.97
0.07

NP


A/duck/Hokkaido/5/1977/H3N2
A/duck/Hong_Kong/365/1978/H4N6
A/duck/Hong_Kong/342/1978/H5N2
A/duck/Chabarovsk/1610/1972/H3N8
A/Duck/Hong_Kong/86/76/H9N2
A/tern/Turkmenia/18/1972/H3N3
A/duck/Ukraine/1/1963/H3N8
A/duck/Taiwan/0526/72/H6N1
A/turkey/England/647/1977/H1N1
A/turkey/England/1963/H7N3
A/duck/Czechoslovakia/1956/H4N6
A/duck/England/1956/H11N6
A/swine/KU/2/2001/H11N6
A/duck/Ukraine/2/1960/H11N8
A/turkey/England/1969/H3N2
A/swine/Hong_Kong/126/1982/H3N2
A/swine/Germany/2/1981/H1N1
A/turkey/Ireland/1378/1983/H5N8
A/swine/Hubei/10/2008/H10N5
A/swine/Guangxi/wz/2004/H5N1
1
0.62
0.88
0.99
0.99
1
1
0.69
0.04

NS


A/swine/Korea/C12/2008/H5N2
A/swine/Korea/C12/2008/H5N2
A/herring_gull/Irkutsk/111/2005/H3
A/duck/Hong_Kong/342/1978/H5N2
A/duck/Hokkaido/5/1977/H3N2
A/Duck/Hong_Kong/366/78/H9N2
A/swine/KU/16/2001/H7N2
A/duck/Hong_Kong/7/1975/H3N2
A/mallard/Saskatchewan/14/1982/H3N8
A/tern/Turkmenia/18/1972/H3N3
A/duck/Chabarovsk/1610/1972/H3N8
A/pintail_duck/ALB/86/1976/H3N2
A/mallard_duck/ALB/57/1976/H5N2
A/turkey/bethlehem-gillit/1492-b/1982/H1N1
A/mallard/Sweden/80389/2008/H4N6
A/duck/Denmark/65047/04/H5N2
A/turkey/England/50-92/91/H5N1
A/gull/Astrakhan/1846/1998/H13N6
A/green_winged_teal/Japan/9KS0643/2009/H12N3
A/whooper_swan/Mongolia/1-17/2007/H3N8
A/snow_goose/Mongolia/1-44/2008/H3N8
A/mallard/Mongolia/1402/2011/H4N6
A/duck/Beijing/61/05/H3N8
A/common_tern/Buryatiya/1901/2000/H4N6
A/tufted_duck/Buryatiya/1905/2000/H4N6
A/duck/Korea/372/2009/H1N1
A/common_teal/Mongolia/2299/2011/H4N6
A/domestic_green-winged_teal/Hunan/3450/2006/H5N1
A/turkey/Italy/3620/1999/H7N1
A/turkey/Italy/4169/1999/H7N1
A/mallard/Sweden/4401/2004/H10N4
A/wild_duck/Korea/SH20-27/2008/H7N9
A/bean_goose/Korea/SH20-17/2008/H7N3
A/wild_duck/Korea/SH5-60/2008/H4N6
A/mallard/Sweden/2990/2003/H8N4
A/mallard/Netherlands/14/2001/H2N3
A/duck/Guizhou/1426/2006/H6N1
A/duck/Primorie/2633/2001/H5N3
A/mallard/Republic_of_Georgia/2/2012/H2N3
A/common_tern/Korgalzhyn/847/2004/H3N6
A/mallard/Sweden/101993/2009/H2N9
0.96
0.81
0.81
0.61
1
1
0.03

Supplementary figure 4.2. Relationships of tern viruses to waterfowl viruses.

Sequences showing the highest similarity in BLAST searches to tern gene segments were aligned by MUSCLE with MEGA6. The phylogenetic trees were inferred with the neighbor-joining method with 1000 bootstrap replications. Evolutionary distances were computed using the Maximum Composite Likelihood method and are represented by the number of base substitutions per site (scale bar). The branches are coloured according to the host group: green, gulls; brown, poultry; light blue, terns; orange, waterfowl; pink, swine.