

# Genetic and Antigenic Evolution of Influenza A (H3N2) Virus Neuraminidase



Kim B. Westgeest



# **Genetic and Antigenic Evolution of Influenza A (H3N2) Virus Neuraminidase**

Kim B. Westgeest

The research described in this thesis was conducted at the ErasmusMC Viroscience department, Rotterdam, the Netherlands and was performed within the framework of the Erasmus Postgraduate School Molecular Medicine. The research presented in this thesis was financially supported by an NWO-VICI grant and NIH contracts no. HHSN266200700010C and HHSN272201400008C.

Financial support for the printing of this thesis was provided by: Viroclinics Biosciences BV, GR instruments, ZEISS, Greiner Bio-One, QIAGEN

ISBN:	978-94-6375-287-9
Lay-out:	David de Groot (persoonlijkproefschrift.nl)
Cover design:	Corné van Leuveren & Kim Westgeest
Chapter design:	Corné van Leuveren & Kim Westgeest
Print:	Ridderprint ( <a href="http://www.ridderprint.nl">www.ridderprint.nl</a> )

© K.B. Westgeest, 2019

All rights reserved. No part of this thesis may be reproduced or transmitted, in any form or by any means, without the permission of the author.

# Genetic and Antigenic Evolution of Influenza A (H3N2) Virus Neuraminidase

Genetische en antigene evolutie van het  
neuraminidase van A(H3N2) influenzavirussen

Proefschrift

ter verkrijging van de graad van doctor aan de  
Erasmus Universiteit Rotterdam  
op gezag van de  
rector magnificus

Prof.dr. R.C.M.E. Engels

en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op  
donderdag 7 maart 2019 om 15:30 uur  
door

Kim Birgitta Westgeest  
geboren te Den Haag

**Promotiecommissie:**

Promotor: Prof.dr. R.A.M. Fouchier

Overige leden: Prof.dr. M.P.G. Koopmans  
Prof.dr. G.F. Rimmelzwaan  
Prof.dr. R.W. Hendriks

Copromotor: Dr. M. de Graaf

'The greatest thing you'll ever learn,  
is just to love and be loved in return'

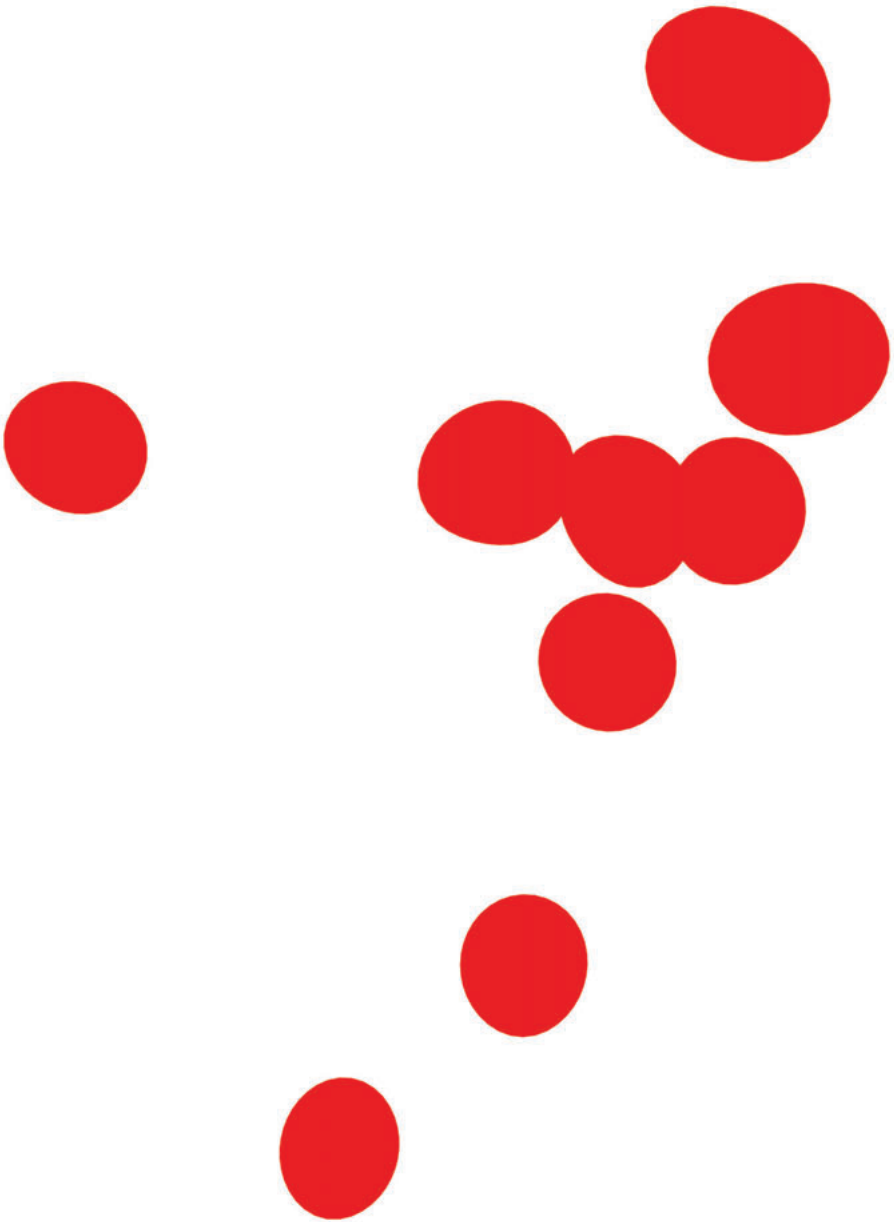
*George Alexander Aberle, 1947*





## TABLE OF CONTENTS

<b>Chapter 1</b>	Introduction	9
<b>Chapter 2</b>	Genetic evolution of the neuraminidase of influenza A (H3N2) viruses from 1968 to 2009 and its correspondence to haemagglutinin evolution. <i>Journal of General Virology 2012</i>	25
<b>Chapter 3</b>	Genomewide analysis of reassortment and evolution of human influenza A(H3N2) viruses circulating between 1968 and 2011. <i>Journal of Virology 2014</i>	47
<b>Chapter 4</b>	Discordant antigenic drift of neuraminidase and haemagglutinin in H1N1 and H3N2 influenza viruses. <i>Proceedings of the National Academy of Sciences 2011</i>	77
<b>Chapter 5</b>	An optimized enzyme-linked lectin assay to measure influenza A virus neuraminidase inhibition antibody titres in human sera. <i>Journal of Virological Methods 2014</i>	93
<b>Chapter 6</b>	Optimization of an enzyme-linked lectin assay suitable for rapid antigenic characterization of the neuraminidase of human influenza A(H3N2) viruses. <i>Journal of Virological Methods 2015</i>	113
<b>Chapter 7</b>	Mapping the Antigenic Evolution of the N2 Neuraminidase of Human Influenza A Viruses from 1957 to 2012. <i>In preparation</i>	133
<b>Chapter 8</b>	Summarizing discussion	165
<b>Chapter 9</b>	References	181
<b>Chapter 10</b>	Nederlandse samenvatting	223
<b>Chapter 11</b>	Dankwoord + About the author	233
	<i>Curriculum vitae</i>	242
	<i>PhD portfolio</i>	243
	<i>List of publications</i>	246



## **CHAPTER 1**

# **Introduction**

## INTRODUCTION

### Influenza virus history

The first documented influenza virus was isolated from swine in 1931<sup>1</sup>, and the first human isolate was reported in 1933<sup>2</sup>. However, in retrospect, the first verified description of influenza dates back to 1878, where it was described as a contagious disease of poultry associated with high mortality. The disease was initially confused with fowl cholera, and it was not until 1955 that the causative agent of this disease was shown to be an avian influenza virus<sup>3</sup>.

### Influenza A virus taxonomy

Influenza viruses are members of the *Orthomyxoviridae* family comprising enveloped, negative-sense, single-stranded, segmented RNA viruses<sup>4</sup>. The family of *Orthomyxoviridae* is divided over seven genera<sup>5</sup>: Alphainfluenzavirus, Betainfluenzavirus, Deltainfluenzavirus, Gammmainfluenzavirus, Isavirus, Quaranjavirus, and Thogotovirus. All infect humans, except Deltainfluenzavirus and Isavirus. Infectious salmon anaemia virus, of the Isavirus genus, is found exclusively in fish<sup>6</sup>, and influenza D virus (IDV), of the Deltainfluenzavirus genus, was isolated from swine and their primary natural reservoir, cattle<sup>7</sup>. The genus Thogotovirus are a group of arthropod-borne viruses which may periodically infect humans<sup>8,9</sup>. Quaranjavirus members appear to predominantly occur in soft ticks and aquatic birds<sup>10</sup>. Influenza A virus (IAV) and influenza B virus (IBV), the only species of the Alphainfluenzavirus and Betainfluenzavirus genera, respectively, are responsible for epidemic influenza outbreaks in humans. Influenza C virus (ICV), the only species of the Gammmainfluenzavirus genus, is associated with sporadic and mild common cold-like illness, also in humans<sup>4</sup>. In addition to humans, IAV can infect a broad range of mammalian and avian species, unlike IBV and ICV, although an IBV has been identified in ray-finned fish<sup>11</sup>. Moreover, IAV is responsible for the major influenza pandemics in humans throughout history<sup>6</sup>.

Division into subtypes of the IAV species is based on the antigenic properties of the major glycoproteins, 'haemagglutinin' (H or HA) and 'neuraminidase' (N or NA)<sup>4</sup>; the subtypes are labelled with an H number and an N number<sup>12</sup>. To date, eighteen HA<sup>13-15</sup> and eleven NA<sup>13,16,17</sup> subtypes have been found in nature, of which the bat IAV surface glycoproteins—H17, H18, N10, and N11—are referred to as HA-like and NA-like<sup>18</sup>. In addition to the subtype information, IAV nomenclature includes the host of origin (with the exception of human-origin viruses where no host of origin designation is given), geographic location of initial isolation, strain number, and year of isolation<sup>19</sup>. For example,

A/duck/Alberta/35/76 (H1N1) is an IAV of duck origin of the H1N1 subtype, the 35th strain isolated in Alberta in 1976. The main reservoir of IAV is wild aquatic birds, and the H1-H16 and N1-N9 subtypes are found within these species. Although several different subtypes have caused zoonotic infections, such as H5N1<sup>20,21</sup>, H5N2<sup>22</sup>, H7N2<sup>23,24</sup>, H7N3<sup>25</sup>, H7N7<sup>26,27</sup>, H9N2<sup>28-30</sup>, and H10N7<sup>4</sup>, only subtypes H1N1, H2N2, and H3N2 have caused pandemics and epidemics in the last century<sup>4,31</sup>.

### **Influenza A virus in humans**

Approximately, 5–15 % of the world population is infected by influenza each year, resulting in an estimated three to five million hospitalizations and 290,000 to 650,000 deaths annually<sup>32</sup>. This yearly phenomenon is referred to as 'seasonal flu' or 'epidemic influenza'. The clinical presentation of epidemic influenza is not straightforward and depends on the nature of the virus and the condition of the patient. It may vary from asymptomatic to a primary viral pneumonia or acute respiratory distress syndrome. In general, the manifestations of uncomplicated influenza during an epidemic are relatively mild. Common symptoms include sore throat, rhinorrhoea, myalgia, headache, feverishness, loss of appetite, malaise, and/or weakness. Nonetheless, the appearance of fever and dry cough within 48 h of symptom onset significantly distinguish IAV infection from other viral or bacterial pathogens causing influenza-like illness. In risk groups on the other hand, influenza may lead to hospitalization or even death (roughly nineteen hundred and eight hundred annual cases, respectively, in the Netherlands<sup>33,34</sup>). People in these risk groups include children below the age of 2, adults older than 65 years of age, and adults with medical conditions such as diabetes, respiratory disease, or cardiovascular disease.

The worldwide spread of a new influenza virus subtype, or 'pandemic influenza', is declared once there is increased and sustained human-to-human transmission within the general population<sup>35</sup>. A large portion of the population will be immunologically naive towards the new subtype, which can lead to serious health problems and even fatal outcomes in the general population including among healthy, young adults<sup>36</sup>. There were four major influenza pandemics in the past century<sup>37</sup> (Figure 1): the 1918 pandemic caused by influenza A H1N1 (A(H1N1)) virus, the 1957-1958 pandemic caused by influenza A H2N2 (A(H2N2)) virus, the 1968 pandemic caused by influenza A H3N2 (A(H3N2)) virus, and the 2009 H1N1 pandemic caused by influenza A H1N1pdm09 (A(H1N1) pdm09) virus. After the 1918 pandemic, the A(H1N1) virus caused annual epidemics until the 1957-1958 pandemic where it was replaced by the A(H2N2) virus which then

became the seasonal influenza until the 1968 pandemic. A(H2N2) and A(H3N2) viruses cocirculated from 1968 until 1971, after which A(H2N2) viruses became extinct in the human population<sup>38,39</sup>. A(H3N2) viruses have been a serious cause of influenza epidemics since their introduction in 1968, with significant morbidity and mortality<sup>32,40</sup>. The 1977 A(H1N1) virus outbreak, which closely resembled viruses that had circulated in the early 1950s<sup>41</sup>, did not cause a pandemic formally, because it was not a novel virus subtype, nor did it replace A(H3N2) viruses. Instead, it cocirculated with A(H3N2) viruses until the 2009 H1N1 pandemic, when it was replaced by the A(H1N1)pdm09 virus which, together with A(H3N2) viruses, continues to cause epidemics<sup>32</sup>.

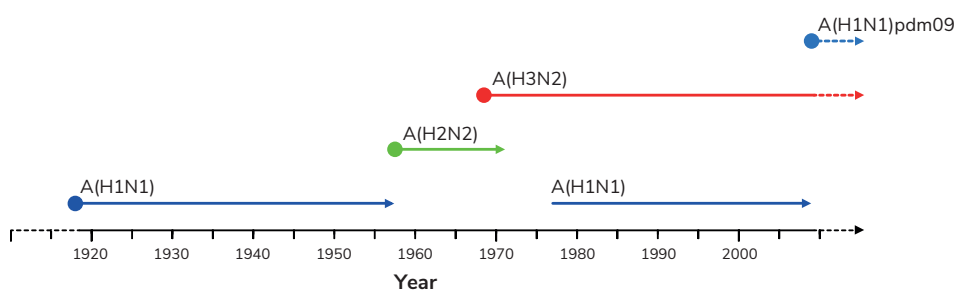


Figure 1 | **Recorded pandemic and epidemic influenza A viruses of the past century.** Horizontal arrows show the period of circulation of the specific influenza A virus subtype after the pandemic onset (solid circles). The blue arrows show the 1918 pandemic caused by influenza A H1N1 (A(H1N1)) virus and 1977 A(H1N1) virus outbreak. The latter was not a pandemic as these viruses closely resembled viruses that had circulated in the early 1950s. The green arrow shows the 1957-1958 pandemic caused by influenza A H2N2 (A(H2N2)) virus and the red arrow illustrates the 1968 pandemic caused by influenza A H3N2 (A(H3N2)) virus. The 2009 H1N1 pandemic caused by influenza A H1N1 (A(H1N1)pdm09) virus is indicated by the light blue arrow. Currently, A(H3N2) and A(H1N1)pdm09 viruses continue to cause epidemics.

### Influenza A virus genome and its encoded proteins

The IAV genome is divided into eight gene segments varying in length and consisting of: basic polymerase 2 (PB2), basic polymerase 1 (PB1), acidic polymerase (PA), HA, nucleocapsid protein (NP), NA, matrix (M), and nonstructural protein (NS). The gene segment that encodes PB1 also encodes an N-truncated form of PB1 starting at codon 40 (PB1-N40)<sup>42</sup>, and PB1-frame 2 (PB1-F2) expressed through an alternative open reading frame (ORF)<sup>43</sup>. The gene segment that encodes PA also encodes PA-X through

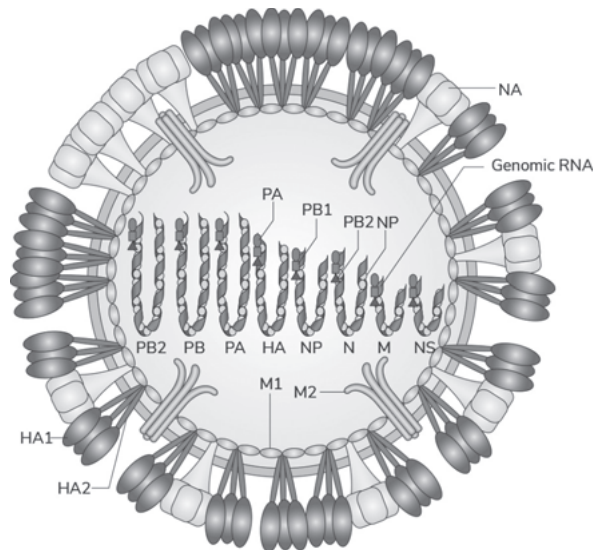


Figure 2 | **Schematic diagram of influenza A virus.** The influenza A virus particle consists of a viral envelope and a viral core containing the eight-segmented genome. Three viral proteins are embedded in the host-derived lipid bilayer membrane: trimeric haemagglutinin (HA), tetrameric neuraminidase (NA), and tetrameric matrix protein M2 (ion-channel). Beneath the viral envelope resides the matrix, containing matrix protein M1, which provides rigidity and surrounds the eight viral ribonucleoproteins (vRNPs). The vRNPs are strands of negative sense single-stranded viral RNA (vRNA) molecules coated with nucleocapsid protein (NP) and attached to a heterotrimeric polymerase complex at the 5' and 3' termini. This polymerase complex consists of PB2, PB1, and PA. Since the vRNA in the vRNPs has complementary sequences at the 5' and 3' termini, the vRNA segments perform base pairing, forming a panhandle structure. Adapted from<sup>47</sup>.

a second ORF accessed via ribosomal frame-shifting<sup>44</sup> and two N-truncated versions of PA; PA-N155 and PA-N182<sup>45</sup>. The M and NS gene segments each encode a protein from the collinear transcript (M1 and NS1). Additionally, they each encode a second protein through mRNA splicing (M2 and NS2), the latter being better known as the nuclear export protein (NEP)<sup>46</sup>.

Within the virus particle (Figure 2), each segment of negative sense single-stranded viral RNA (vRNA) forms a panhandle structure through complementarity of the 5' and 3' termini. The RNA is tightly packed as a viral ribonucleoprotein (vRNP) by NP and the heterotrimeric RNA-dependent RNA polymerase complex for the purposes of RNA

transcription, replication, and packaging. This polymerase complex consists of PB2, PB1, and PA. PB1-N40 lacks transcriptase function but does interact with the polymerase complex in the cellular environment<sup>42</sup>, and although PA-N155 and PA-N182 do not show polymerase activity, mutant viruses lacking the N-truncated PAs replicate with lower efficiency in cell culture and have reduced pathogenicity in mice compared to the wild-type virus<sup>45</sup>. PB1-F2, PA-X, and dimeric NS1 are undetectable in the viral particle and are only found in infected cells where they induce apoptosis (PB1-F2), modulate host response and viral virulence (PA-X), allow escape from antiviral mechanisms (NS1), or regulate gene expression (NS1)<sup>48</sup>. Besides the eight vRNPs, NEP is found within the viral particle and, along with M1, directs the nuclear export of the vRNPs. M1 proteins form a layer providing rigidity to the virus particle, which underlies the host-derived lipid bilayer membrane. Three viral proteins—HA, NA, and M2—are embedded in the lipid bilayer membrane. HA is a homotrimeric type I integral membrane protein of which each monomer contains a receptor binding site (RBS). The precursor HA molecule—HA0—is proteolytically cleaved into HA1 and HA2 subunits<sup>49</sup> to activate virus infectivity<sup>50,51</sup>. HA1 is the immunogenic section of the HA protein<sup>52,53</sup>, and HA2 contains a fusion domain<sup>54</sup>. NA is a tetrameric type II integral membrane protein possessing sialidase activity. M2—a tetrameric type III integral membrane protein—has ion channel activity allowing protons from the acidified endosomes into the virus particle to separate the vRNPs from M1.

### **Replication cycle**

Respiratory droplets or aerosols loaded with virus are released during breathing, coughing, and sneezing by an infected individual<sup>55–57</sup>. The droplets or aerosols enter the respiratory tract of another individual where the IAV binds to terminal sialic acid-containing receptors on the epithelial cell surface through the HA RBS<sup>58</sup>, enabling entrance through internalization into endosomes by receptor mediated endocytosis. The ion channel activity of M2 is activated, by which ions flow from the acidified endosome to the virion interior<sup>59</sup>. The acidic environment triggers a conformational change of the cleaved HA, exposing the fusion peptide at the N-terminus of the HA2 subunit<sup>60</sup>. This allows fusion of the viral membrane with the cellular membrane resulting in the release of the eight vRNPs into the cytoplasm. The intact vRNPs are actively imported into the nucleus through the nuclear pore, since all proteins of the vRNP—PB2, PB1, PA, and, particularly, NP—contain nuclear localization signals<sup>61</sup>. In the nucleus, the polymerase complex transcribes the vRNAs into mRNAs. Initiation of influenza mRNA requires capped primers generated from the host cellular mRNA



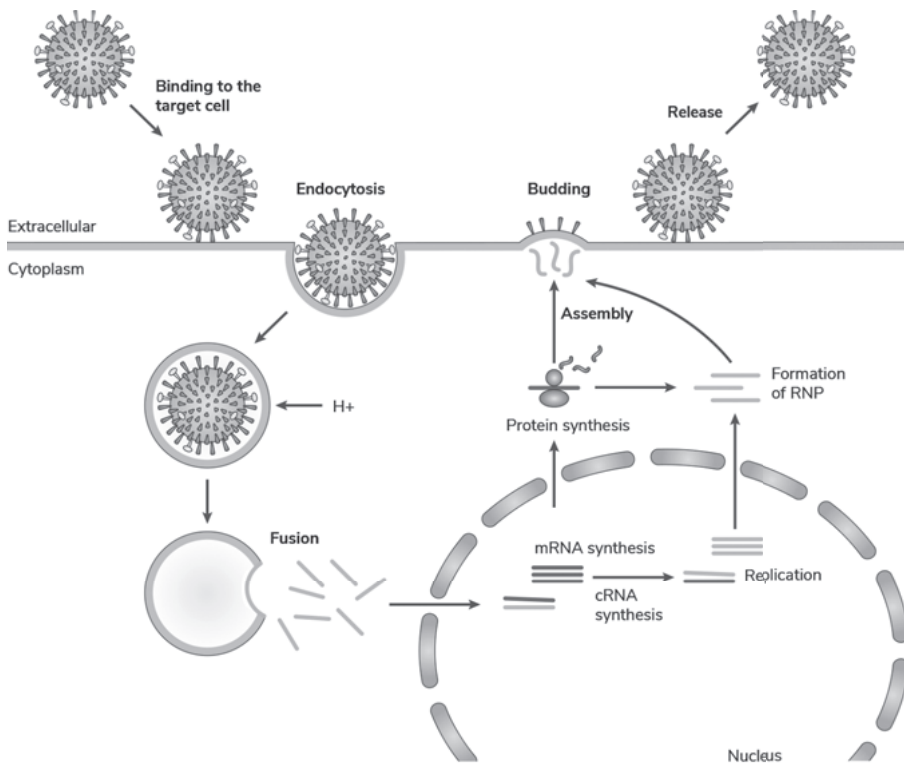


Figure 3 | **The replication cycle of influenza A virus.** Upon binding to sialic acid receptors on the epithelial cell surface via the haemagglutinin (HA) glycoprotein, the influenza A virus is internalized into endosomes by receptor mediated endocytosis. Hereafter, the ion channel activity of matrix protein 2 (M2) is activated, by which ions flow from the endosome to the virion interior. The low pH in the endosome triggers fusion of the viral and endosomal membranes, releasing the viral ribonucleoproteins (vRNPs) into the cytoplasm. The intact vRNPs are actively imported into the nucleus through the nuclear pore. In the nucleus, the polymerase complex transcribes the viral RNAs (vRNAs) into mRNAs. After synthesis of the mRNAs in the nucleus, the mRNAs are translated within the cytoplasm, yielding viral proteins. Newly formed proteins are either transported back into the nucleus or transported to the cell membrane. The vRNA serves additionally as a template for synthesis of complementary RNA (cRNA). This cRNA is positive stranded and needed as a template for the synthesis of vRNA. The vRNA is exported from the nucleus, together with newly synthesized viral proteins, as vRNP complexes. The vRNPs are then packaged, leading to the assembly of progeny virus. This is realized by means of 'budding', in which the lipid membrane of the host cell serves as the viral envelope. The sialic acids are enzymatically removed by neuraminidase (NA), resulting in the release of the virion from the cell. Adapted from<sup>70</sup>.

by the cap-dependent endonuclease activity of PB2<sup>62</sup>. After synthesis in the nucleus, the mRNAs are translated within the cytoplasm, yielding viral proteins. Newly formed proteins are either transported back into the nucleus (PB2, PB1, PA, NP, M1, and NEP) or transported to the cell membrane (HA, NA, and M2)<sup>4,6</sup>. The vRNA additionally serves as a template for synthesis of complementary RNA (cRNA)<sup>6</sup>. This cRNA is positive stranded and needed as a template for the synthesis of vRNA<sup>6</sup>. NEP and M1 enable the export of vRNA from the nucleus, together with newly synthesized viral proteins, as vRNP complexes<sup>46,63,64</sup>. The vRNPs are then packaged, leading to the assembly of progeny virus. This is realized by means of 'budding', in which the lipid membrane of the host cell serves as the viral envelope<sup>65</sup>. The terminal sialic acids are enzymatically removed by NA, resulting in the release of the virion from the cell<sup>66-69</sup> (Figure 7a). This activity also prevents aggregation of virions<sup>68</sup>.

### Reassortment

The segmented nature of the influenza virus genome allows for the exchange of entire genes between different influenza viruses during simultaneous co-infection of a host, in a process called reassortment<sup>4</sup> (Figure 4). At least one copy of each gene segment is needed for an influenza virus particle to be infectious<sup>4</sup>, and this incorporation mechanism is selective<sup>71,72</sup>. Electron microscopy studies show a highly organized structure within influenza A virions, comprising one central vRNP surrounded by seven others of varying length<sup>73-75</sup>, and segment-specific packaging signals located at the 5' and 3' end of the coding regions have been defined<sup>71,72,76-83</sup>.

Intersubtypic reassortments between swine, avian, and/or human IAV have led to several pandemics. The 1957-1958 pandemic A(H2N2) virus was the result of reassortment between the at the time-circulating seasonal influenza A(H1N1) virus and an avian influenza A(H2N2) virus<sup>84</sup> from which it derived the HA, NA and PB1. The 1968 pandemic A(H3N2) virus was the result of reassortment between a human A(H2N2) virus with an avian H3 virus<sup>84</sup> of which the latter provided the HA and PB1 proteins. Between 1977 and 2009, occasional reassortants between contemporary human A(H1N1) and A(H3N2) viruses have been detected<sup>31,85-93</sup>, but these have never persisted as seasonal influenza. The 2009 H1N1 pandemic was the product of reassortment in which the NA and M genes were derived from a Eurasian avian-like swine virus, and the remaining six gene segments were derived from triple reassortant swine viruses that possessed genes originating from classical A(H1N1) swine, North American avian, and human A(H3N2) viruses<sup>94</sup>. Reassortment between IAV of the same subtype, intrasubtypic reassortment, is

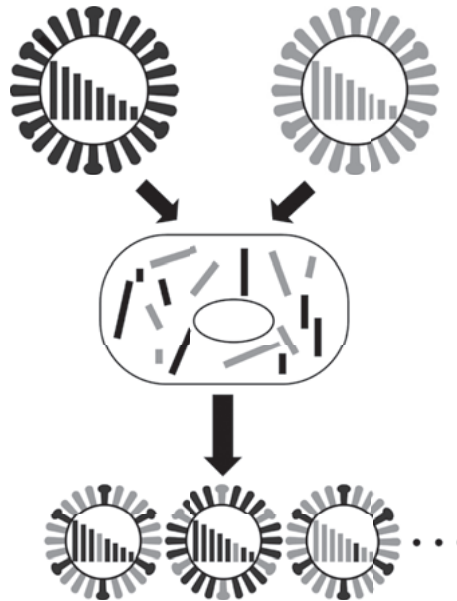


Figure 4 | **Influenza A virus reassortment.** Two virus particles are shown, each with a full complement of eight viral genome segments. Following reassortment, hybrid progeny can be formed that contain segments derived from both parents. Adapted from<sup>4</sup>.

most likely one of many ways for the virus to increase diversity<sup>95–100</sup>, including producing variants with enhanced transmissibility<sup>101</sup>, and those that are resistant to antiviral drugs<sup>102</sup>.

### Genetic and antigenic evolution

A hallmark of RNA viruses is their high genome mutation frequency. The influenza virus polymerase lacks proofreading capacity, leading to an enormous variety of progeny viruses. The majority of these mutations contribute to the neutral sequence evolution<sup>99</sup>, also known as genetic drift. However, some of the progeny mutant viruses, in particular under selective conditions, have advantageous characteristics and are thus rapidly amplified. This process of natural selection acts on B-cell<sup>103</sup> and T-cell epitopes<sup>104</sup>, aids in drug resistance, and enables propagation under altered growth conditions<sup>103</sup>.

Mutations that arise during antibody mediated selection, leading to alterations in the major surface glycoproteins, can result in 'antigenic drift'<sup>105</sup>. Antibodies elicited to the 'original' virus may lose the ability to recognize the virus or have reduced affinity, resulting in infection of the host. Antigenic drift increases diversity but, more importantly,

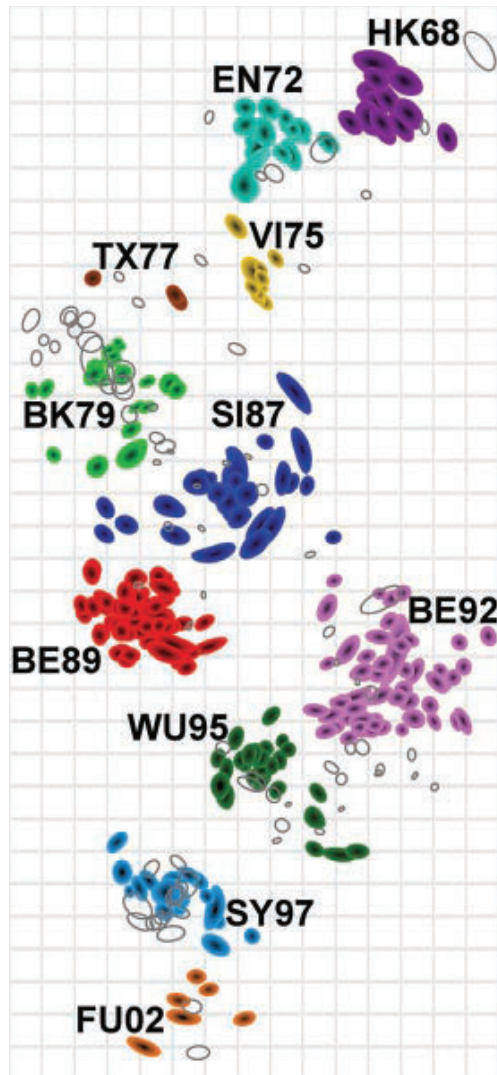


Figure 5 | **Antigenic map of influenza A (H3N2) virus from 1968 to 2003.** The distance between antiserum (uncoloured open shapes) point S and virus (antigen, coloured shapes) point A corresponds to the difference between the base 2 logarithm ( $\log_2$ ) of the maximum titre for antiserum S against any antigen and the  $\log_2$  of the titre for antiserum S against antigen A. Consequently, each titre in an haemagglutination inhibition (HI) table can be thought of as specifying the ideal target distance between the antiserum and antigen points in the resulting antigenic map. Modified multidimensional scaling (MDS) methods are used to arrange the antigen and antiserum points in an antigenic map to best satisfy the target distances specified by the HI data. Taken from<sup>107</sup>.

is a way for the virus to effectively evade the host's immune system and is responsible for the recurrence of influenza epidemics almost every winter.

Hirst was the first to discover the agglutination of red blood cells by the influenza virus<sup>106</sup>. This agglutination is the basis of the haemagglutination inhibition (HI) assay which can be used to detect antigenic differences in HA. In the HI assay, the presence of HA specific antibodies in antisera prevent the binding of HA to sialic acid containing receptors present on red blood cells, thus preventing the agglutination of red blood cells. The amount of inhibition by antisera is converted into an HI titre. In 2004, Smith et al. designed 'antigenic cartography', a tool to visualize and quantify these HI data<sup>107</sup>. With this tool the antigenic evolution of the HA proteins of A(H3N2) viruses was mapped, from their introduction into humans in 1968 until 2003 (Figure 5). The study revealed a punctuated pattern of eleven antigenic clusters, named after the first vaccine strain in the cluster; Hong Kong 1968, England 1972, Victoria 1975, Texas 1977, Bangkok 1979, Sichuan 1987, Beijing 1989, Beijing 1992, Wuhan 1995, Sydney 1997, and Fujian 2002. Each of these antigenic clusters contains viruses that are antigenically similar for some time, after which these viruses change antigenically, resulting in a "cluster transition". These cluster transitions warrant at least one vaccine update. Antigenic clusters remain dominant for on average 3.3 years, and strains appeared up to 2 years before and until 2 years after the period in which that cluster is dominant. These clusters were somewhat linearly aligned, and it was hypothesized that new clusters tend to move away as far as possible from the previous cluster, thereby most effectively escaping existing population-level immunity<sup>107</sup>.

### Neuraminidase

The receptor-destroying activity of influenza virus was first observed by Hirst<sup>106</sup> in 1941, and the product that was cleaved was identified by Gottschalk in 1956 as N-acetylneuraminic acid, one of the sialic acids<sup>108</sup>. This led to the discovery of the second major surface glycoprotein 'NA' in 1957 by Gottschalk, who showed that NA split off the terminal neuraminic acid unit from inhibitory mucoproteins and inhibitory glycolipids<sup>69</sup>. In 1964, it was shown that the NA and HA activities reside on different proteins<sup>109</sup>. The NA of most IAV subtypes and of IBV has been crystallized<sup>110-118</sup>, allowing study of its three-dimensional structure. The three-dimensional structure of NA consists of four domains: the cytoplasmic, transmembrane, stalk, and head. NA is a mushroom-shaped tetrameric protein<sup>119</sup>, and one viral particle has approximately thirty to forty NA tetramers, formed into clusters, among 290-300 HA spikes on an average-sized virion<sup>73</sup>. Each identical subunit of the tetrameric protein carries the

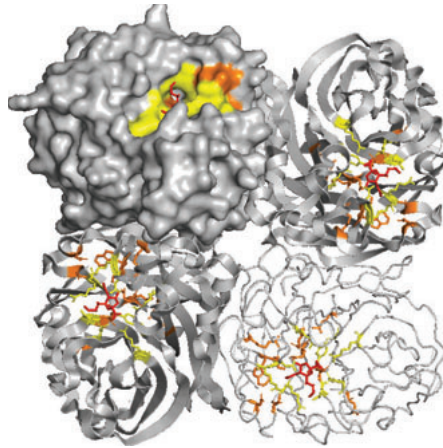


Figure 6 | **The NA globular head.** Illustration of the tetrameric NA structure of subtype N2 (PDB code 2BAT<sup>115</sup>) generated with MacPyMOL. The viral receptor sialic acid, shown as red sticks, is docked into the active site, represented in yellow, surrounded by the framework sites, depicted in orange. NA monomers are shown in surface filled-space (top left), cartoon ribbons (top right and bottom left), and smoothed cartoon loops (bottom right).

enzyme's active site. However, the monomeric form of NA has no enzyme activity<sup>120,121</sup>. The NA active site is a shallow pocket lined by conserved amino acid residues, some of which bind directly to the viral receptor sialic acid and participate in catalysis, while others provide a structural framework<sup>119</sup> (Figure 6).

Antigenic drift has been studied most considerably for HA, even though NA has also been observed to undergo antigenic drift<sup>122–124</sup>. NA-specific antibodies can reduce viral replication and disease severity in mice<sup>125</sup> and chickens<sup>126</sup>, and have similarly been associated with resistance against influenza in humans<sup>127,128</sup>. Although antibodies against NA do not prevent infection, preclinical and clinical studies showed that NA-specific immunity can reduce the severity of disease<sup>125,128–132</sup>. Antibodies directed towards NA inhibit release and spread of newly formed virus particles from infected cells<sup>133</sup>. Early crystallographic studies of NA have shown that antigenic regions<sup>134</sup> surround the enzyme's highly conserved active site<sup>135,136</sup> and are highly variable, most likely due to immune pressure<sup>123,124</sup>. Studies comparing the antigenic evolution of HA and NA using a limited number of viruses have discovered that their evolution differs and is often asynchronous<sup>122,137</sup>.

Despite the correlation of NA with immunity and the asynchrony in antigenic evolution between HA and NA, NA is not routinely examined in seroepidemiology or

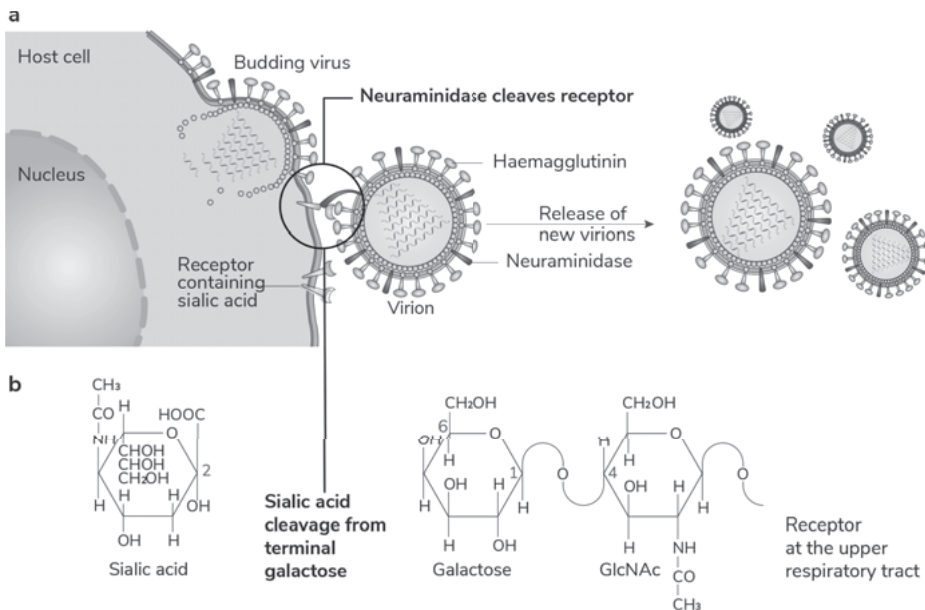


Figure 7 | **Receptor-destroying activity of neuraminidase.** **a** | Neuraminidase cleaves sialic acid (N-acetylneuraminic acid) from the cell receptor, resulting in the release of newly formed influenza virus particles from the cells. **b** | Sialic acid linked to galactose is cleaved, exposing the galactose. Galactose is linked to N-acetylglucosamine (GlcNAc). Adapted from<sup>145</sup>.

vaccine studies. Antigenic characterization of NA can be performed using NA inhibition (NI) assays which determine the extent of antibody-mediated interference with enzyme activity<sup>138</sup>. These assays rely on the enzymatic sialidase activity of NA, which is determined by measuring cleavage of sialic acid from highly glycosylated proteins such as fetuin. The NI thiobarbituric acid (TBA) assay<sup>139,140</sup> is based on the detection of free sialic acid. This assay is recommended by the World Health Organization (WHO)<sup>141</sup>, but the use of large volumes of hazardous chemicals and laborious procedures has impeded antigenic characterization of NA during influenza surveillance. Alternative assays that have been developed include an enzyme-linked lectin assay (ELLA)<sup>142</sup>, which also relies on the sialidase activity of NA, but instead of measuring free sialic acid, it detects the terminal galactose that becomes exposed after sialic acid cleavage by NA (Figure 7b). Both the NI TBA and the NI ELLA measure functional inhibition of NA activity by antibodies, and consequently have clear relevance to immunity *in vivo*<sup>130,143,144</sup>.

A complication of NI assays is the interference of HA-specific antibodies that block NA activity non-specifically when they bind to HA<sup>137</sup>. H6Nx viruses generated by classical

reassortment were used in the traditional TBA method<sup>122,138</sup>. However, the generation of recombinant viruses is time-consuming for large numbers of viruses and therefore this method is not suitable for large-scale surveillance of antigenic evolution of circulating influenza viruses. For analysis of the antigenic evolution of NA it would be ideal to have the possibility of using wild type viruses as antigen in assays that are not impacted by non-specific inhibitors, including antibodies to HA.

### **Vaccine**

The effect of antigenic drift made it necessary to reformulate vaccines after only 2 years of use<sup>146</sup>, and the WHO established the WHO Global Influenza Surveillance Network (GISN) in 1952 for the early detection of drifted strains<sup>147,148</sup> renamed to the Global Influenza Surveillance and Response System (GISRS) in 2011<sup>149</sup>. The 1968 pandemic led to the development of the intramuscular trivalent inactivated vaccines (TIVs) against IAV and IBV<sup>150</sup>. Influenza vaccines reduce the morbidity and mortality associated with annual influenza epidemics but need to be updated frequently due to antigenic drift. The current seasonal influenza vaccine is designed to protect against circulating A(H1N1)pdm09 viruses, A(H3N2) viruses, and IBV. Thorough surveillance by the National Influenza Centres and WHO Collaborating Centres within the WHO's GISRS is needed to identify the most suitable strains to use in vaccines for the next epidemic<sup>151</sup>. Representatives of the predominant circulating viruses are sent to the WHO Collaborating Centres, which produce vaccine seed strains that can be shared with vaccine manufacturers. Vaccine strain selection depends on three aspects: epidemiological information, genetic analyses of the HA and NA genes, and serological analysis of HA. The main focus of influenza surveillance is on HA, and official influenza vaccines are required to contain 15 µg of each HA subtype<sup>152</sup> while the amount of NA is not standardized. Vaccine effectiveness (VE) is dependent on how well the vaccine strains match the circulating influenza strains, and estimates showed a VE of 10 % in 2004–2005, 21 % in 2005–2006, and 52 % in 2006–2007<sup>153</sup>.

### **Scope of this thesis**

There is still much uncertainty about the underlying mechanisms that govern antigenic drift of influenza viruses, and several theories have been proposed. These theories consider HA to be the primary driving force, while the second key surface glycoprotein (NA) and the other viral proteins have largely been ignored. This thesis focuses on all



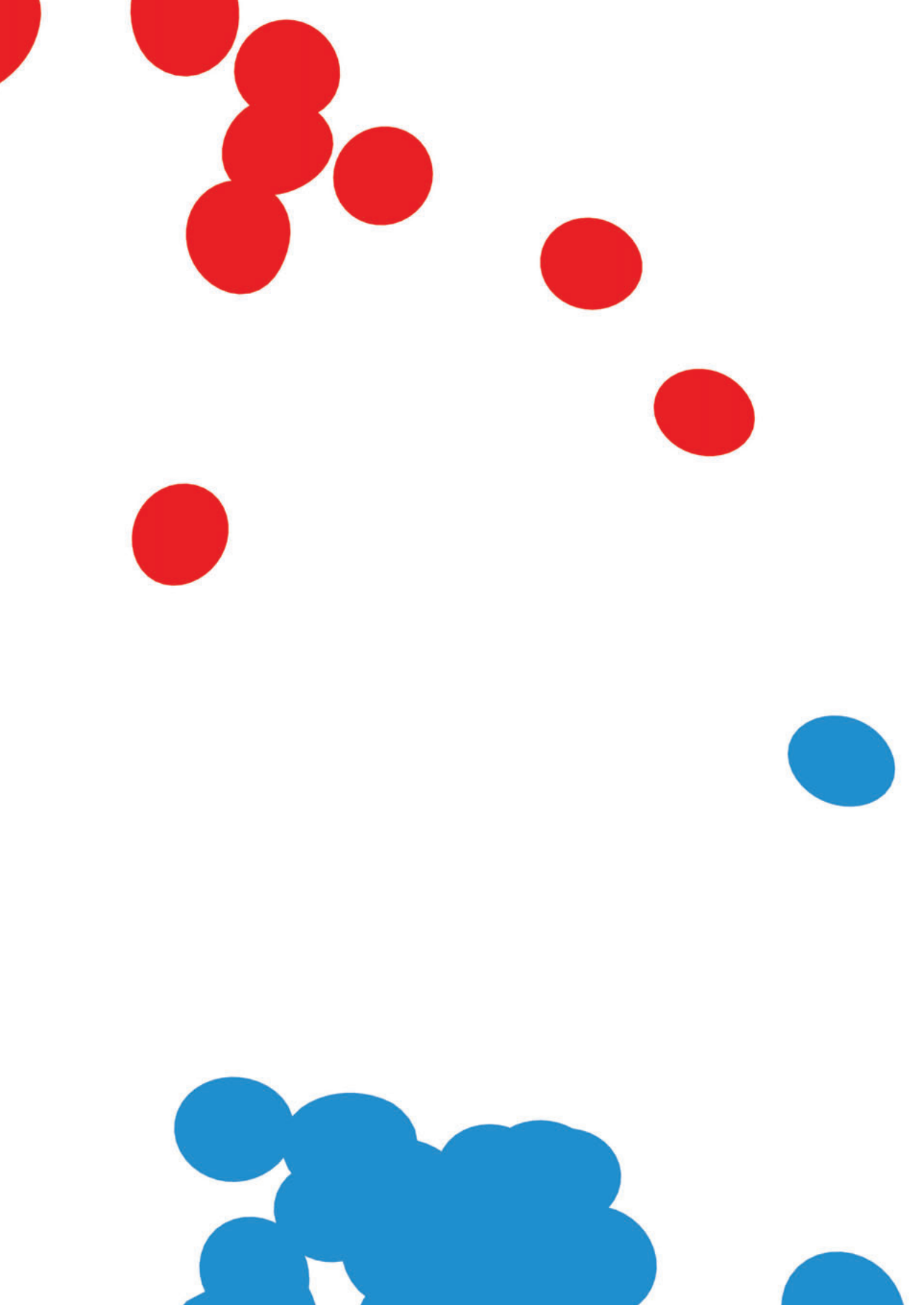
influenza proteins, with special emphasis on NA, to better understand—and ultimately predict—the complex evolution of A(H3N2) viruses.

In **chapter 2** the genetic evolution of NA is compared with HA1 using an extensive data set containing virus isolates obtained from consecutive influenza seasons from 1968 till the 2009-2010 season. The antigenic properties as measured by the HI assay were known, and with this information, reassortment events, rates of evolutionary change, and selection pressures were analysed over 40 years of influenza A(H3N2) virus evolution. In **chapter 3**, a similar approach is used to study the full influenza virus genomes and their relation to antigenic change.

**Chapter 4** utilizes the TBA NI assay to characterize antigenic evolution of NA in human A(H1N1) and A(H3N2) viruses recommended for United States of America influenza vaccines between 1995 and 2010. The TBA NI assays are performed using ferret antisera directed against wild-type A(H1N1) and A(H3N2) viruses, and as antigens virus reassortants with the targeted NA and a mismatched HA of the H6 subtype. Use of these reassortants prevents false NI signals due to interfering HA antibodies. Antigenic maps were constructed from the data sets using multidimensional scaling, to position the antigens and antisera on the map, as previously described<sup>107</sup>.

Because the conventional NI assays are cumbersome and laborious, critical steps in the NI ELLA were modified and optimized to support the use of the NI ELLA in human serology in **chapter 5**. **Chapter 6** describes the steps that are important for obtaining consistent results with the NI ELLA in a format suitable for rapid antigenic characterization of NA, with wild type viruses as antigen. Interference by antibodies directed against HA of wild type A(H3N2) viruses is prevented by raising ferret sera against recombinant influenza A H7N2 (A(H7N2)) viruses that contain the NA of various A(H3N2) viruses. Through this approach it is possible to screen wild type viruses, thus preventing the time-consuming generation of recombinant viruses or proteins for each epidemic virus of interest. Finally, in **chapter 7**, the optimized NI ELLA is used to analyse roughly three hundred viruses from 1968 till the 2010-2011 season. The resulting antigenic map of NA is compared to the antigenic map of HA.

The results of all studies, and the implications they have on future research, are discussed in **chapter 8**.



## CHAPTER 2

# Genetic Evolution of Neuraminidase of Influenza A (H3N2) Viruses from 1968 to 2009 and its Correspondence to Haemagglutinin

Kim B. Westgeest, Miranda de Graaf, Mathieu Fourment, Theo M. Bestebroer, Ruud van Beek, Monique I.J. Spronken, Jan C. de Jong, Guus F. Rimmelzwaan, Colin A Russell, Albert D.M.E. Osterhaus, Gavin J.D. Smith, Derek J. Smith, Ron A.M. Fouchier

*Journal of General Virology*, volume 93(9), 1 September 2012, pages 1996-2007



## **ABSTRACT**

Each year, influenza viruses cause epidemics by evading pre-existing humoral immunity through mutations in the major glycoproteins: the haemagglutinin (HA) and the neuraminidase (NA). In 2004, the antigenic evolution of HA of human influenza A (H3N2) viruses was mapped from its introduction in humans in 1968 until 2003<sup>107</sup>. The current study focused on the genetic evolution of NA and compared it with HA using the data set of Smith and colleagues, updated to the epidemic of the 2009/2010 season. Phylogenetic trees and genetic maps were constructed to visualize the genetic evolution of NA and HA. The results revealed multiple reassortment events over the years. Overall rates of evolutionary change were lower for NA than for HA1 at the nucleotide level. Selection pressures were estimated, revealing an abundance of negatively selected sites and sparse positively selected sites. The differences found between the evolution of NA and HA1 warrant further analysis of the evolution of NA at the phenotypic level, as has been done previously for HA.

## INTRODUCTION

Influenza epidemics affect approximately 5–15 % of the world population resulting in an estimated 3–5 million hospitalizations and between 250,000 and 500,000 deaths annually<sup>40,154</sup>. Influenza type A viruses can be divided into subtypes based on the antigenic properties of the major surface glycoproteins haemagglutinin (HA) and neuraminidase (NA). To date, 17 HA<sup>14,15</sup> and nine NA<sup>17</sup> subtypes have been found in nature. Subtypes A (H1N1) and A (H3N2) are currently the causative agents of influenza A virus epidemics in humans, of which the influenza A (H3N2) viruses are the most recurring and virulent<sup>155,156</sup> and show the strongest antigenic drift<sup>95</sup>.

HA, a homotrimeric type I integral membrane protein, mediates viral entry into the cell by binding to sialic acids<sup>58</sup>. NA, a tetrameric type II integral membrane protein with sialidase activity, allows virus release from the cell<sup>66–68</sup>. Both HA and NA are located on the surface of the viral membrane and are the main targets for antibodies. Antibodies against influenza viruses result in protective immunity, but mutations in HA and NA allow the virus to escape host immunity. This process, known as antigenic drift<sup>157</sup>, is responsible for the recurrence of influenza epidemics almost every winter.

Vaccines have proven effective, but need to be updated frequently due to antigenic drift. Since 1999, the influenza A (H3N2) virus component has been updated six times<sup>155</sup>. Thorough surveillance by the National Influenza Centres and WHO Collaborating Centres within the World Health Organization's Global Influenza Surveillance Network is required to identify the most suitable strains to use in vaccines for the next epidemic<sup>151,155</sup>. Vaccine strain selection depends on three aspects: epidemiological information, HA and NA gene sequence phylogeny and serological analysis using an HA inhibition assay. The main focus of genetic and antigenic surveillance is on HA, and official influenza vaccine formulations prescribe the amount of HA<sup>158</sup>.

Although antibodies against NA do not prevent infection, numerous preclinical and clinical studies indicate a role of NA immunity in reducing the severity of influenza virus infection<sup>125,128–132</sup>. Early crystallographic studies of NA have shown that antigenic regions surround the enzyme's highly conserved active site<sup>135,136</sup>. Antigenic sites A, B and C<sup>134</sup> were shown to be highly variable probably due to antigenic drift<sup>123,124</sup>. Studies comparing the antigenic drift of HA and NA using limited numbers of viruses have revealed that their evolution differs and is often asynchronous<sup>122,137,159</sup>. Given these factors, understanding the patterns of evolution in NA is important.

In 2004, Smith et al. mapped the antigenic evolution of HA of human influenza A (H3N2) virus from its introduction in humans in 1968 until 2003<sup>107</sup>. The study was

based on an extensive data set, comprising influenza virus isolates obtained within each influenza season. Here, we focused on the genetic evolution of NA and compared it with HA1 (the immunogenic section of the HA), using the data set of Smith et al., updated to the epidemic of the 2009/2010 season. Reassortment events, rates of evolutionary change, and selection pressures were analysed over 40 years of influenza A (H3N2) virus evolution. The differences found between the evolution of NA and HA1 warrant further analysis of the antigenic properties of NA.

## **MATERIALS AND METHODS**

### **Viruses**

Human influenza A (H3N2) viruses, isolated over 35 years of influenza virus surveillance between 1968 and 2003 that were used in the study of Smith et al. in 2004<sup>107</sup> served as the basis for this study. Two viruses within this data set were no longer available (A/Victoria/7/87 and A/Netherlands/440/93) and were thus excluded from the data set. The data set was updated with 19 influenza A (H3N2) viruses circulating between 2003 and 2009, including seven vaccine or reference strains and 12 isolates from epidemics in the Netherlands. The latter 12 isolates were chosen based on divergent placement in the updated A (H3N2) antigenic map<sup>160</sup>. This led to a total of 291 human influenza A (H3N2) viruses. Human influenza A (H3N2) viruses were propagated in Madin–Darby canine kidney (MDCK) cells or, if unsuccessful, in 11-day-old embryonated chicken eggs.

### **Sequence analysis**

MDCK supernatant or allantoic fluid was used for RNA extraction, using a High Pure RNA Isolation Kit (Roche Applied Science) and cDNA was synthesized with SuperScript<sup>®</sup> III reverse transcriptase (Invitrogen). The complete NA and HA1 gene segments were amplified by PCR using AmpliTaq Gold<sup>®</sup> DNA polymerase (Applied Biosystems) and purified by gel extraction with the QIAquick Gel Extraction Kit (Qiagen). Sequencing was performed with NA or HA specific primers using a BigDye Terminator version 3.1 cycle sequencing kit (Applied Biosystems) and a 3100 Genetic Analyzer (Applied Biosystems), according to the instructions of the manufacturer.

Nucleotide sequences of the NA open reading frames (ORF) and HA1 coding regions were aligned using the ClustalW program running within the BioEdit software package, version 7.0.9.0<sup>184</sup>. Seven HA1 and 14 NA sequences contained one or more degenerate nucleotide positions. In all HA1 and ten of the 14 NA sequences, the

degenerate nucleotide led to degenerate amino acid positions. One NA sequence contained an insertion of 3 nucleotides between codons 221 and 222, and one sequence had a 3-nucleotide deletion at codon 154. Newly sequenced and previously published NA and HA1 domain accession numbers are provided in the Supplementary data.

### Phylogeny

Nucleotide sequence alignments with or without an additional outgroup (A/Duck/Hokkaido/33/80 for HA1 and A/Japan/305/1957 for NA) were generated. With these nucleotide sequence alignments, the best-fit models of nucleotide substitution were determined by jModelTest<sup>185</sup>. In all cases, the preferred maximum likelihood (ML) optimized model of nucleotide substitution was TVM+I+ $\Gamma$ 4 (transversion model with the proportion of invariant sites and the gamma distribution of among-site rate variation with four categories estimated from the empirical data) with GTR+I+ $\Gamma$ 4 (general time reversible model) as second best based on the Akaike information criterion (AIC).

As seven HA1 and ten NA nucleotide sequences contained degenerate amino acid positions, either a B (aspartate or asparagine) or an X (any amino acid) was applied at these positions. The best-fit models of protein evolution were determined with ProtTest<sup>186</sup>. The preferred ML optimized model was HIVw+ $\Gamma$ 4 (human immunodeficiency virus within) for HA1 and HIVw+I+ $\Gamma$ 4+F (equilibrium amino acid frequencies estimated from the empirical data) for NA based on the AIC.

### Phylogenetic analysis

With the nucleotide sequence alignments, initial ML trees were inferred using the PhyML package version 3.0<sup>187</sup>, by means of a full heuristic search and the subtree pruning and regrafting (SPR) method. As the TVM+I+ $\Gamma$ 4 model of nucleotide substitution was not an option within the PhyML package, the GTR+I+ $\Gamma$ 4 model was chosen for all data sets. Garli version 0.951<sup>188</sup> was run on the best nucleotide tree from PhyML for 2 million generations to optimize tree topology and branch lengths. Additionally, ML trees were estimated using the TVM+I+ $\Gamma$ 4 model of base substitution in combination with tree bisection-reconnection (TBR) searches using the PAUP\* (version 4.0b10) package<sup>189</sup>. For each of the trees, the reliability of all phylogenetic groupings was determined through a non-parametric bootstrap resampling analysis; either 500 replicates of ML trees using the GTR+I+ $\Gamma$ 4 model and the SPR method using PhyML, or 1,000 replicates of neighbour-joining trees estimated under the ML substitution model using PAUP\*. All trees are available from the authors upon request. Trees were visualized with

the FigTree program version 1.3.1. (<http://www.tree.bio.ed.ac.uk/software/figtree>). Trees were rooted on the outgroup strains (A/Duck/Hokkaido/33/80 for HA1 and A/Japan/305/1957 for NA) or on the 1968 pandemic strain (A/Hong Kong/1/1968).

Amino acid alignments were used in combination with the HIVw+r4 (HA1) or HIVw+l+r4+F (NA) model of protein evolution to infer ML trees using the PhyML package version 3.0<sup>187</sup>, by means of a full heuristic search and the SPR method.

### **Genetic maps**

Amino acid sequence alignments were used to calculate a distance matrix with the number of amino acid substitutions between pairs of strains to produce 'genetic maps', as described previously<sup>107</sup>. Genetic mapping is a way to facilitate a quantitative analysis and visualization of genetic data. In a genetic map, the distance between isolates A and B corresponds to the number of amino acid substitutions between the amino acid sequences of a particular protein-coding region of isolates A and B. Thus, each difference in an amino acid alignment can be thought of as specifying a target distance for the points in a genetic map. Modified multidimensional scaling (MDS) methods are then used to arrange the points between two isolates in a genetic map to best satisfy the target distances specified by the amino acid alignment distance matrix. The result is a map in which the distance between points represents the number of amino acid substitutions in a particular protein-coding region between isolates. To avoid underestimating genetic distances due to sequential mutations at the same location, a threshold needs to be applied. For example, when a threshold of 30 is applied to the alignment, the software is free to use a distance of at least 30 aa between two points that have a 30 aa difference in order to retrieve the lowest error function. For HA1, the previously determined threshold of 30 was used<sup>107</sup>. Threshold considerations for NA were based on adequate correlation between observed distances and genetic map distances and sufficient correlation between ML amino acid tree distances and genetic map distances. The threshold that approached both considerations best for NA was 20 (data not shown). The correlation between the number of amino acid substitutions and the corresponding distances between strains in the genetic map was 0.93 (data not shown), indicating that the two-dimensional genetic map is a reasonable representation of the target amino acid distance matrix. For software, see <http://www.antigenic-cartography.org>.



### Detection of reassortment

ML trees of NA and HA1 were used in TreeMap version 1.0 (<http://taxonomy.zoology.gla.ac.uk/rod/treemap.html>). TreeMap, originally designed for comparing host and parasite trees, was applied solely to display a tanglegram between the HA and NA phylogenies. The twines were colour coded according to the HA antigenic clusters<sup>107</sup>.

Reassortment events were identified by the graph-incompatibility-based reassortment finder (GiRaF) program<sup>162</sup>. Nucleotide alignments of HA1 or NA were used as input for MrBayes<sup>190,191</sup> to sample 1,000 unrooted candidate trees with the GTR+I+Γ4 substitution model, a burn-in of 100,000 and sampling every 200 iterations. These trees were subsequently used to model the phylogenetic uncertainty for each segment with the GiRaF program under default settings. This procedure was repeated 100 times with ten independent MrBayes runs for NA and ten independent MrBayes runs for HA1.

### Estimation of nucleotide substitution rates and times of divergence

To identify potential errors in sequence data annotation that might affect the clock estimation, the ML nucleotide HA1 and NA trees were exported to Path-O-Gen version 1.3 (<http://tree.bio.ed.ac.uk/software/pathogen>)<sup>192</sup> and a linear regression plot for the years of sampling versus root-to-tip distance was generated (see Figure S3 in the supplemental material). No anomalies were seen in the NA and HA1 data sets which both behaved in a clock-like manner ( $R^2 = 0.977$  and  $R^2 = 0.967$ , respectively).

Overall rates of evolutionary change (nucleotide substitutions per site per year), most recent common ancestor (MRCA) in years and relative genetic diversity were estimated using the BEAST program version 1.6.2 (<http://beast.bio.ed.ac.uk/>)<sup>193</sup>. For all analyses, the uncorrelated log-normal relaxed molecular clock was used to accommodate variation in molecular evolutionary rate among lineages in combination with the SRD06 codon position model, with a different rate of nucleotide substitution for the first plus second versus the third codon position, and the HKY85 substitution model<sup>194</sup>.

Isolation dates were added to calibrate the molecular clock. This analysis was conducted with a time-aware linear Bayesian skyride coalescent tree prior<sup>195</sup> over the unknown tree space, with relatively uninformative priors on all model parameters. Two independent Bayesian Markov chain Monte Carlo (MCMC)-analyses, performed for HA1 and NA for 100 million states, sampling every 1,000 states, were performed. Convergences and effective sample sizes of the estimates were checked using Tracer version 1.5 (<http://www.evolve.zoo.ox.ac.uk/software.html?id=tracer>) and the first 10

% of each chain was discarded as burn-in. Uncertainty in parameter estimates is reported as values of the 95 % highest probability density (HPD).

### **Selection pressures**

To determine the degree of natural selection acting on HA1 and NA, the mean number of  $d_N$  and  $d_S$  substitutions per site ( $d_N/d_S$  ratio or  $\omega$ ) was estimated using the 'one-ratio' model of the Codeml program implemented in the PAML package<sup>196</sup>. Codeml uses the codon substitution model of Goldman & Yang for protein coding DNA sequences<sup>197</sup> and was used in combination with the ML tree rooted on A/Hong Kong/1/68. Selection pressures on the internal and external branches of the HA1 and NA ML tree were estimated with the 'two-ratio' model of the Codeml program.

To determine the positively and negatively selected sites along internal branches, the internal fixed effects likelihood (IFEL) method was used<sup>164</sup>. Sites subject to episodic diversifying selection were identified using Mixed Effects Model of Evolution (MEME)<sup>165</sup>. Both methods were accessed through the Datamonkey interface (<http://www.datamonkey.org>) and the GTR (HA1) or TVM (NA) nucleotide substitution model was used. Sites with a p-value < 0.05 were reported as positively selected sites.

Directional positive selection was detected with the Directional Evolution in Protein Sequences (DEPS) test implemented in the HyPhy package<sup>166</sup>. For this procedure, the ML trees with the A/Hong Kong/1/68 pandemic strain as outgroup were used in combination with the amino acid alignment and the HIVw model of protein evolution.

Positively selected sites were shown on the monomeric and tetrameric NA structure constructed using MacPyMOL (PyMOL Molecular Graphics System, version 13) on subtype N2 (pdb code 2BAT<sup>115</sup>).

## **RESULTS**

### **Dataset background**

Virus isolates obtained between 1968 and 2003 and used in the study of Smith et al.<sup>107</sup> were combined with 19 influenza A (H3N2) viruses isolated between 2003 and 2009 into a single updated data set consisting of 291 virus isolates. The new isolates included seven vaccine or reference strains and 12 isolates from epidemics in the Netherlands, of which the latter were chosen based on divergent placement in the updated influenza A (H3N2) antigenic map<sup>160</sup>. All NA gene segments of these 291 viruses were sequenced. For the study by Smith et al. only the HA1 coding regions were sequenced, as this is the immunogenic section of the HA protein<sup>52,53</sup>. To obtain a similar data set for HA and NA,

the HA1 coding regions of the 19 recent influenza viruses were sequenced. Deduced amino acid sequence alignments showing only positions with at least one mutation are available in Figure S1 and S2 for NA and HA1, respectively (available in JGV Online). Both alignments were colour coded according to the antigenic clusters of HA<sup>107</sup>.

### Genetic evolution of NA and HA1

ML phylogenetic trees were constructed to observe the genetic evolution of NA and HA1, updated from 2004, at the nucleotide level (Figure 1a). The general topology of the NA tree was similar to that of the HA1 tree, showing the typical 'ladder-like' gradual evolution with rapid replacement of old strains by newer ones. NA had fewer nucleotide substitutions over 40 years of evolution compared with HA1; overall, the genetic distance from the root of the tree to the most recent cluster of CA04 strains was roughly 1.5-fold greater for HA1 than for NA. Although we did see that variants on the trunk were the ancestors of all variants in future years<sup>161</sup>, the HA1 phylogeny did have some variants that clustered away from the trunk in an evolutionary terminal or dead-end clade, most notably for the VI75- and the BE89-like viruses. Compared with HA1, such dead-end clades were less obvious in the NA tree. For NA, the most obvious—yet relatively small—evolutionary dead-end clades included a number of the SY97- and BE92-like viruses. Whilst the HA genes of the VI75-like and BE92-like viruses clustered continuously in the ML tree, the NA genes appeared as two separate lineages for both groups of viruses. Notably, the NA genes of HK68-like viruses were more genetically divergent than HA1 (0.044403 vs 0.03556 nucleotide substitutions per site, respectively). Larger genetic distances were found between BK79- and SI87-like viruses and within the SY97-like viruses for NA compared with HA1. In contrast, the distances between EN72- and VI75-, between TX77- and BK79-, and between SY97- and FU02-like viruses were greater for HA1.

Genetic maps were generated to visualize the genetic evolution of NA at the amino acid level (Figure 2a). The HA1 genetic map was updated with the newly added isolates (Figure 2b). In Figure 2c, the two maps are compared with arrows that point from the position of a virus in the HA1 map to the position of the corresponding virus in the NA map. Overall, the genetic map of the NA appeared less clustered and more gradual compared with the HA1 genetic map. Although the genetic clustering of strains based on HA1 amino acid sequences was in good agreement with the antigenic properties of the HA proteins as described previously<sup>107</sup>, the NA amino acid sequences did not strictly follow the same pattern of clustering. Other features of the NA genetic map, including the

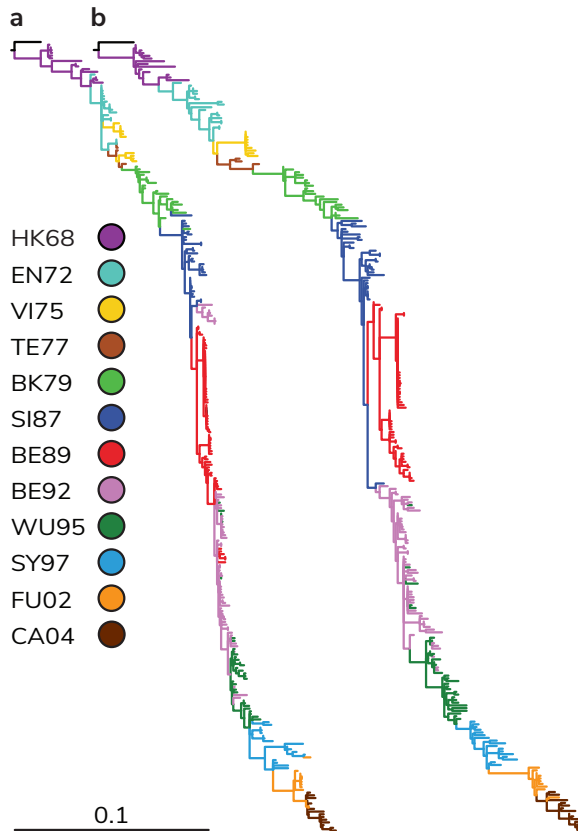


Figure 1 | **Phylogenetic ML trees of the major glycoproteins of human influenza A (H3N2) viruses.** The phylogenetic ML trees of NA (a) and HA1 (b) were generated with 292 nucleotide sequences. The bar represents ~10 % of nucleotide substitutions between close relatives. The colour coding of viruses is based on the antigenic clusters of HA<sup>107</sup> and is consistent between both trees.

less pronounced evolutionary dead ends (yellow and red clusters in Figure 2b compared with 2a) and differences in genetic distance within and between clusters as compared with HA1, were in agreement with the ML trees.

### Reassortment events

In 2005, Holmes et al. demonstrated for a limited number of influenza seasons that multiple lineages of influenza A (H3N2) viruses cocirculate, persist, and reassort in epidemiologically significant ways<sup>96</sup>. By applying the antigenic cluster colours of HA<sup>107</sup>

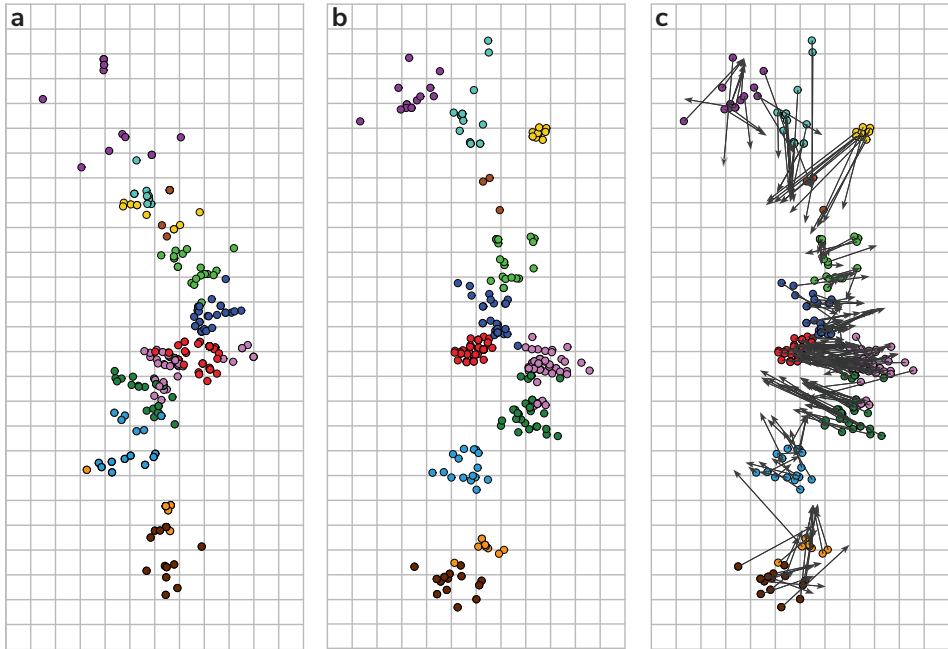


Figure 2 | Genetic maps of the major glycoproteins of human influenza A (H3N2) viruses. Genetic maps were generated with 291 aa sequences and MDS algorithms for NA (a) and HA1 (b). The vertical and horizontal axes correspond to the number of amino acid substitutions; one square represents five amino acid substitutions. The orientation of the HA1 and NA maps was chosen to match the orientation of the antigenic map of human influenza A (H3N2) viruses<sup>107</sup>. The colour coding of viruses is based on the antigenic clusters of HA and is consistent among all maps (see legend to Figure 1). The right map (c) depicts the HA genetic map with superimposed arrows pointing towards the corresponding data points in the NA genetic map.

to both the HA1 and NA trees, it was noted that the clusters of NA sequences did not strictly coincide with those of HA (Figure 1). The HA1 and NA trees shown in Figure 1 were used in TreeMap to generate a tanglegram. A tanglegram enables visualization of the location of particular isolates within both ML trees. In the absence of reassortment, the twines should connect both trees, in theory, in a seamlessly horizontal way. This was not the case for the isolates used in this study, suggesting frequent reassortment between the NA and HA gene segments. However, this was based on only one phylogeny per gene segment, with clades and branches ordered by branch length in a fixed fashion. To improve on this, the computational software package GiRaF was used to identify reassortments in multiple trees per gene segment<sup>162</sup>. This method compares large collections of MCMC-sampled trees for groups of incompatible splits

to identify sets of taxa with differential phylogenetic placement, whilst accounting for uncertainties in the inferred phylogenies. We performed 100 independent GiRaF analyses on ten independent MrBayes HA runs compared with ten independent MrBayes NA runs. Each run consisted of two independent tree files each with 1,000 MCMC-sampled trees. GiRaF removed the first 500 of these trees as burn-in, and the remaining 500 were used for the actual analysis. All reassortment events are reported in Table S1 in the supplemental material and those with a support value of  $\geq 50\%$  of the GiRaF runs are depicted on the tanglegram made from the ML trees as bold twines (Figure 3). A 50% cut-off was chosen arbitrarily.

For the HK68-like viruses, two possible reassortment events were reported with a 58 and 52% support of the GiRaF runs (Figure 3 and Table S1 in the supplemental material; events 1 and 2, respectively). There was no evidence for reassortment during the circulation of the EN72-like viruses with support from GiRaF of  $\geq 50\%$ . The VI75-like viruses in the HA1 tree formed one clade, which descended from EN72-like viruses, and this clade formed the common ancestor to TX77-like viruses. In the NA tree, the VI75-like viruses were divided into two clades. The two different clades represented two different influenza seasons. The first clade, comprising viruses isolated in 1975 and early 1976, was directly descended from EN72-like viruses, similar to the HA tree. In contrast, the second clade, consisting of strains isolated in early 1977, originated, together with the TX77-like viruses, from late EN72-like viruses. The viruses containing the NA of the second clade were reassortants, with a reported support of 71% of the GiRaF runs (event 3). GiRaF provided no support for reassortment between HA and NA during circulation of the TX77-like viruses. The BK79-like strain A/Philippines/2/1982 was a reported reassortment event with 93% support (event 4). Two strains with very different HAs, one BK79- and the other SI78-like, both contained a closely related NA and were identified as reassortants in 80% of the runs (event 5). A/Hong Kong/1/1989, with an intermediate SI87-like HA and an NA that clustered with early SI87-like viruses, was also reported as a reassortant (event 6, 90%). Multiple reassortment events included BE92-like viruses and five out of 14 were supported by  $\geq 50\%$  of the GiRaF reports (events 8 and 10–13). The most notable observation was that the BE92-like viruses were divided into two clades in the NA tree, one descending from SI87-like viruses and the second one from BE89-like viruses. The upper clade included viruses circulating from late 1991 until mid-1993, whilst the second clade contained viruses circulating from late 1992 until early 1996.

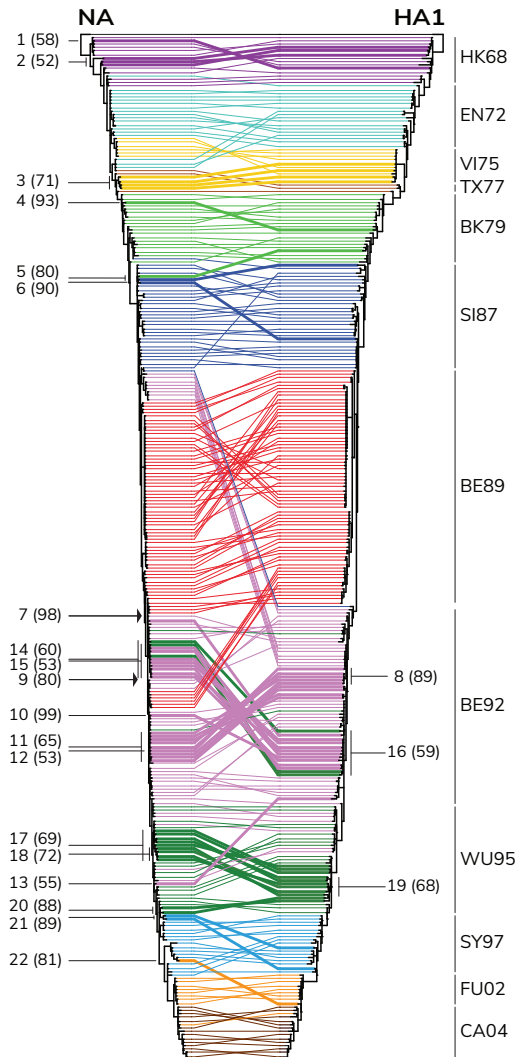


Figure 3 | **Reassortment events between NA and HA1 during 42 years of influenza A (H3N2) evolution.** TreeMap version 1.0 was used to generate a tanglegram with the phylogenetic ML trees shown in Figure 1. Twines between both trees were colour coded according to the antigenic clusters of HA<sup>107</sup> (see legend to Figure 1). GiRaF version 1.01<sup>162</sup> was used to detect reassortment events between HA and NA. Reassortment events (see Table S1 in the supplemental material) reported in  $\geq 50\%$  of the cases, are indicated in bold, with percentages shown in parentheses. The arrows indicate the introduction of an NA that represents the common ancestor of all more recent NAs.

The first isolate of the second BE92 clade and everything descending from it were all reported as a single reassortment event supported by 98 % of the GiRaF runs (event 7). A similar event was reported for a more recent BE92-like virus and all its descendants, although with weaker support (event 9, 80 %). Both reports suggest that, at some point during the influenza season of 1992/1993, an NA was introduced by reassortment that represents the common ancestor of the more recent NAs.

Multiple reassortment events were reported by GiRaF within and between the BE92 and WU95 clusters. Clustering of BE92- and WU95- like viruses appeared scattered in both trees, but when comparing these scattered clades between the NA and HA tree, a number of clades positioned differently. Three reassortment events that were supported by  $\geq 50$  % of the runs included a mixture of BE92- and WU95-like viruses (events 14–16). During circulation of the more recent WU95-like viruses, four reassortment events were supported by  $\geq 50$  % of the GiRaF runs (events 17–20). One reassortment event involving SY97-like viruses was supported by GiRaF  $\geq 50$  % (event 21, 89 %). For FU02-like viruses, only one reassortant was reported (event 22, 81 %). Reassortment of HA and NA of CA04-like viruses was not supported at  $\geq 50$  % by GiRaF.

Thus, visual inspection of a single NA tree and a single HA1 tree within the tanglegram suggested numerous reassortment events. Inference from multiple MCMC-sampled trees provided support for a portion of these events, using an arbitrary cut-off of 50 %. Overall, these data indicated that, although mostly singular, reassortment events have occurred throughout the evolution of human influenza A (H3N2) viruses, especially during the periods of circulation of BE92-, and WU95-like viruses. Most reassortment events were within antigenic clusters, rather than between antigenic clusters. There were 13 reassortment events within the same HA antigenic group (events 1, 2, 4, 6, 8, 10–13 and 17–20). There were six reassortment events involving strains from different HA groups (events 5, 7, 9 and 14–16). Event 3 included only VI75-like strains, but the NAs originated from late EN72-like viruses instead of descending from VI75-like viruses. Event 21 contained an NA derived from WU95-like viruses and an HA descending from SY97-like viruses. Reassortment event 22 was a FU02-like strain, although the NA of this virus originated from SY97-like viruses. Whereas most reassortants were not detected during prolonged periods, possibly suggesting that the reassortment events were neutral or detrimental, two reassortment events (events 7 and 9) persisted in the population.



Table 1 | Mean rate of nucleotide substitutions and time of MRCA of NA and HA1.

Segment	Mean rate of nucleotide substitution ( $\times 10^{-3}$ substitutions/site/year)			Time of MRCA		
	Mean	95 % HPD		Mean	95 % HPD	
		Lower	Upper		Lower	Upper
NA	3.15	2.81	3.49	1965.10	1962.82	1967.10
HA1	5.15	4.62	5.70	1967.00	1966.18	1967.73

HPD, highest probability density  
MRCA, most recent common ancestor

### Evolutionary rates

Rates of nucleotide substitution and time of the MRCA for NA and HA1 were estimated using BEAST version 1.6.2 with the relaxed log-normal clock and the Bayesian skyride time-aware model (Table 1). The mean rates of nucleotide substitution for NA and HA1 were  $3.15 \times 10^{-3}$  (HPD  $2.81 \times 10^{-3}$ – $3.49 \times 10^{-3}$ ) and  $5.15 \times 10^{-3}$  (HPD  $4.62 \times 10^{-3}$ – $5.70 \times 10^{-3}$ ) nucleotide substitutions per site per year, respectively. The time of MRCA was 1965.10 (HPD, 1962.82–1967.10) for NA and 1967.00 (HPD, 1966.18–1967.73) for HA1.

### Selection pressures

The degree of natural selection acting on NA and HA1 was estimated by looking at the mean number of amino acid-changing (non-synonymous or  $d_N$ ) and silent (synonymous or  $d_S$ ) substitutions per site ( $d_N/d_S$  ratio) using Codeml. The overall  $d_N/d_S$  and the mean  $d_N/d_S$  ratios for the internal and external branches were estimated (Table 2). Both genes were under strong selection (NA  $d_N/d_S = 0.249$  and HA1  $d_N/d_S = 0.362$ ) however, HA1 was less constrained. The  $d_N/d_S$  values for the internal and external branches were similar for NA (0.250 and 0.248) and HA1 (0.350 and 0.371), suggesting no difference between selection pressures on NA and HA1. In theory, a higher  $d_N/d_S$  value for the external relative to the internal branches indicates an excess of non-synonymous mutations that are eventually removed from the virus population by purifying selection, suggesting that these mutations are deleterious<sup>163</sup>. This was not the case for NA and HA1.

To determine the positively and negatively selected sites along internal branches only, where advantageous mutations are more likely to fall, the internal fixed effects

Table 2 | Global, inner and outer  $d_N/d_S$  values for NA and HA1.

Segment	Sequence length (codons)	$d_N/d_S$		
		Overall	Internal	External
NA	470	0.249	0.250	0.248
HA1	329	0.362	0.350	0.371

likelihood (IFEL) method was used<sup>164</sup>. IFEL detected three positively selected sites in NA: codons 43, 267 and 370 (see Table S2 in the supplemental material). In HA1, nine positively selected sites were reported, all located in antigenic sites (see Table S3 in the supplemental material).

Sites subject to episodic diversifying selection were identified using mixed effects model of evolution (MEME)<sup>165</sup>. MEME detected five sites (codons 43, 148, 199, 338 and 465) in NA (see Table S2 in the supplemental material) and 14 in HA1 (see Table S3 in the supplemental material) of which all but two (codons 31 and 185) were found in antigenic sites.

Directional positive selection was detected with the directional evolution in protein sequences (DEPS) test implemented in the HyPhy package<sup>166</sup>. The DEPS test identifies directional evolution towards residues at sites within an amino acid alignment in combination with a nucleotide tree and is particularly useful for the detection of selective sweeps<sup>167</sup>. For NA, 11 sites (see Tables S2 and S4 in the supplemental material) were identified to be involved in this directional evolution (see Figure S4 in the supplemental material for the inferred amino acid substitution patterns). For HA1, nine sites were involved (see Tables S3 and S5 in the supplemental material). In contrast, this method detected more sites under directional selection in NA compared with HA1 (11 and nine sites, respectively).

The positively selected sites within HA1 (see Table S3 in the supplemental material) were found mainly in the antigenic sites of HA1 (with two exceptions). One site, namely 145, was detected by all three methods. Residues 31, 133, 135, 137, 159, 186, 193, 226, and 262 were found by two of the methods. For the positively selected sites found by the IFEL, MEME and DEPS analyses it was noted that only a few were within antigenic sites of NA (see Table S2 in the supplemental material; residues 199, 328, 334, 338, 367, and 370). All positively selected sites are indicated in the NA monomeric and tetrameric protein diagrams shown in Figures 4a and 4b, respectively. Figure S4 in the

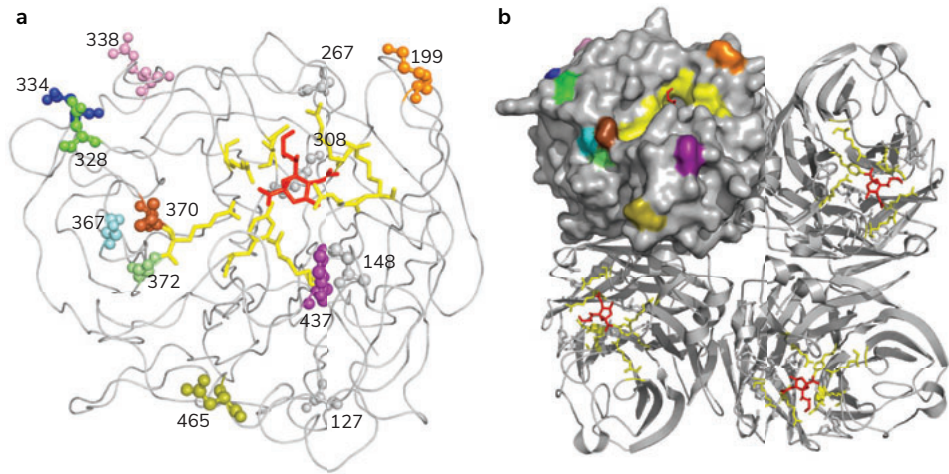


Figure 4 | Sites identified as being positively selected, depicted on the NA globular head. Positively selected sites are shown on wire and filled space models of the monomeric (a) and tetrameric (b) NA structure constructed using MacPyMOL on subtype N2 (pdb code 2BAT<sup>115</sup>). The viral receptor sialic acid, represented as red sticks, is docked into the active site, depicted as yellow sticks. All sites located in the globular head found to be positively selected are illustrated as spheres of which the residues visible on the tetrameric structure were given a colour: orange, 199; green, 328; blue, 334; magenta, 338; cyan, 367; brown, 370; lime, 372; purple, 437; olive, 465. Residues 127, 148, 267 and 308 (grey) were also found to be positively selected but were not surface exposed. Residues 43, 46 and 52 are not shown, as they are in the NA stalk domain. Numbers correlate to codon positions identified as under positive selection found within this study (see Table S2 in the supplemental material).

supplemental material depicts ML trees coloured by each positively selected site found by at least two methods.

The transition of codon 43 from an aspartic acid to a serine occurred around the time of the cluster transition from TX77 to BK79. The change of serine to an asparagine at position 43 corresponded with the emergence of the BE89 cluster and all descending viruses. The change of arginine to leucine at position 338 correlated roughly with the transition from BE92 to WU95. The changes at position 370 from leucine to serine and from serine back to leucine did not correspond with antigenic cluster transitions for HA, as they occurred during the continued circulation of the HK68-like and BK79-like strains, respectively. The late leucine to serine substitution correlated roughly with the FU02 to CA04 transition.

## DISCUSSION

In this study, we compared the genetic evolution of NA of 291 human influenza A (H3N2) viruses sampled between 1968 and 2009 with that of HA1. Although similarities were apparent, the main observation was that HA and NA exhibited clear differences in evolutionary genetics.

The general topology of the NA and HA1 trees were similar, showing the typical 'ladder-like' gradual evolution with rapid replacement of old strains by newer ones as described previously for HA1<sup>161</sup>. NA evolved more slowly at the nucleotide level and more gradually, with overall shorter branches and fewer and smaller evolutionary dead-end clades. The genetic distances between clades in the ML trees and in the genetic maps for NA and HA1 were often discordant, suggesting an asynchronous genetic evolution of the two genes. This observation is in agreement with the reported asynchronous evolution at the phenotypic level, with discordant patterns of antigenic drift of HA and NA as measured by serological assays<sup>122,137,159</sup>.

Reassortment is a common feature in influenza viruses<sup>168</sup>. The Asian, Hong Kong, and 2009 swine flu pandemics of 1957, 1968 and 2009, respectively, all emerged following reassortment of swine, avian and/or human influenza viruses<sup>84,169</sup>. In addition to the emergence of pandemic strains, it has been suggested that the emergence of new virus lineages during seasonal epidemics of the past decade may have been associated with reassortment events<sup>96</sup>, perhaps resulting from improved matching of the activity of HA and NA in relation to HA antigenic change<sup>170-172</sup>. However, when amino acid substitutions in the active sites of HA and NA were analysed, no evidence for a 'matched' evolution of HA and NA was found, as there were only isolated cases of substitutions in the catalytic site of NA (specifically at position 151) that seemed to provide no increased fitness as their occurrences were transient.

For our influenza A virus data set spanning 40 years of influenza A (H3N2) virus epidemics, differences in clustering of HA and NA gene segments within phylogenetic trees indeed suggested that reassortment events had occurred relatively frequently over time. The large genetic distances between NA sequences of the HK68-like viruses, shown in both the phylogenetic ML tree (Figure 1a) and the genetic map (Figure 2a), could be explained by a reassortment event involving an early influenza A (H3N2) virus, in which a heterologous N2 gene of H2N2 virus origin was reintroduced<sup>38</sup>. Indeed, in an ML tree generated with NA genes of all early influenza A (H3N2) viruses and additional H2N2 strains, the NA of the late H2N2 viruses clustered together with HK68-like viruses of 1969 and 1970 (data not shown). GiRaF analyses provided further support

that several HK68-like strains (A/Bilthoven/17938/1969, A/Bilthoven/2668/1970, and A/Bilthoven/93/1970) were reassortants. Apart from these late HK68-like strains, GiRaF provided evidence for reassortment events between 1971 and 1977. Co-circulation of antigenically distinct viruses has been reported in this period<sup>173–175</sup>, providing opportunities for such reassortment events. In agreement with a previous study<sup>176</sup>, there was also support for reassortment events between BE89- and BE92-like viruses. GiRaF analyses further suggested that reassortment events between HA and NA of influenza A (H3N2) viruses occurred particularly frequently during the circulation of BE92- and WU95-like viruses. It remains unclear whether the frequency of reassortment events over time is significant and whether these reassortment events have been of epidemiological significance. Analyses with all genomic sequences would provide us with a better insight into reassortment patterns.

Some concerns arise from the use of GiRaF in the intra-H3N2-evolution context. The evolutionary distances sampled in the present data set are not the same as those GiRaF was originally benchmarked against, and the performance of GiRaF was shown to depend on the distance distribution. Moreover, GiRaF seemed to perform slightly better for single-taxa reassortments compared with larger events. It should further be noted that the low false-positive rate of GiRaF was only estimated on synthetic data sets assuming neutral evolution models<sup>162</sup>. Here, we used a tanglegram based on robust HA and NA ML trees to visually inspect the results reported by GiRaF as an independent test. The reassortment events reported by GiRaF with > 50 % support could all be confirmed by visual inspection of the tanglegram. To test whether the small genetic distances between the sets of HA and NA sequences were problematic for GiRaF, we also ran GiRaF on HA trees only (comparing one set of HA trees with another set of HA trees) and on NA trees only, and using alignments in which the HA and NA sequences were split in half. One would not expect to detect any reassortment events in these runs. GiRaF analysis reported only three events for the 5'HA–3'HA analysis, and no reassortment events for the analysis of 5'NA–3'NA, HA–HA, and NA–NA trees. Thus, we concluded from this that the 22 HA–NA reassortment events reported in the present work are supported by GiRaF and visual inspection.

The evolutionary rates of NA and HA1 were high<sup>177,178</sup> and of the expected level for influenza viruses<sup>179</sup>. The evolutionary rate of NA of  $3.15 \times 10^{-3}$  nucleotide substitutions per site per year was slightly higher than previously reported rate ( $2.3 \times 10^{-3}$ <sup>176</sup>). Although the rate of nucleotide substitution of NA was lower than that of HA ( $5.15 \times 10^{-3}$  nucleotide substitutions per site per year), it is important to note that only HA1 was analysed here.

As HA1 is the most variable part of the HA gene, this part would probably yield higher evolutionary rates than the full HA gene<sup>179</sup>.

The coalescent analyses revealed that the time of circulation of the MRCA of both HA1 and NA of influenza A (H3N2) viruses was around 1965 and 1967, respectively (Table 1). As the 95 % HPDs for the MRCAs of both NA and HA1 overlapped, it is likely that there was a single seeding event for the 1968 pandemic, possibly prior to the first recognition in 1968.

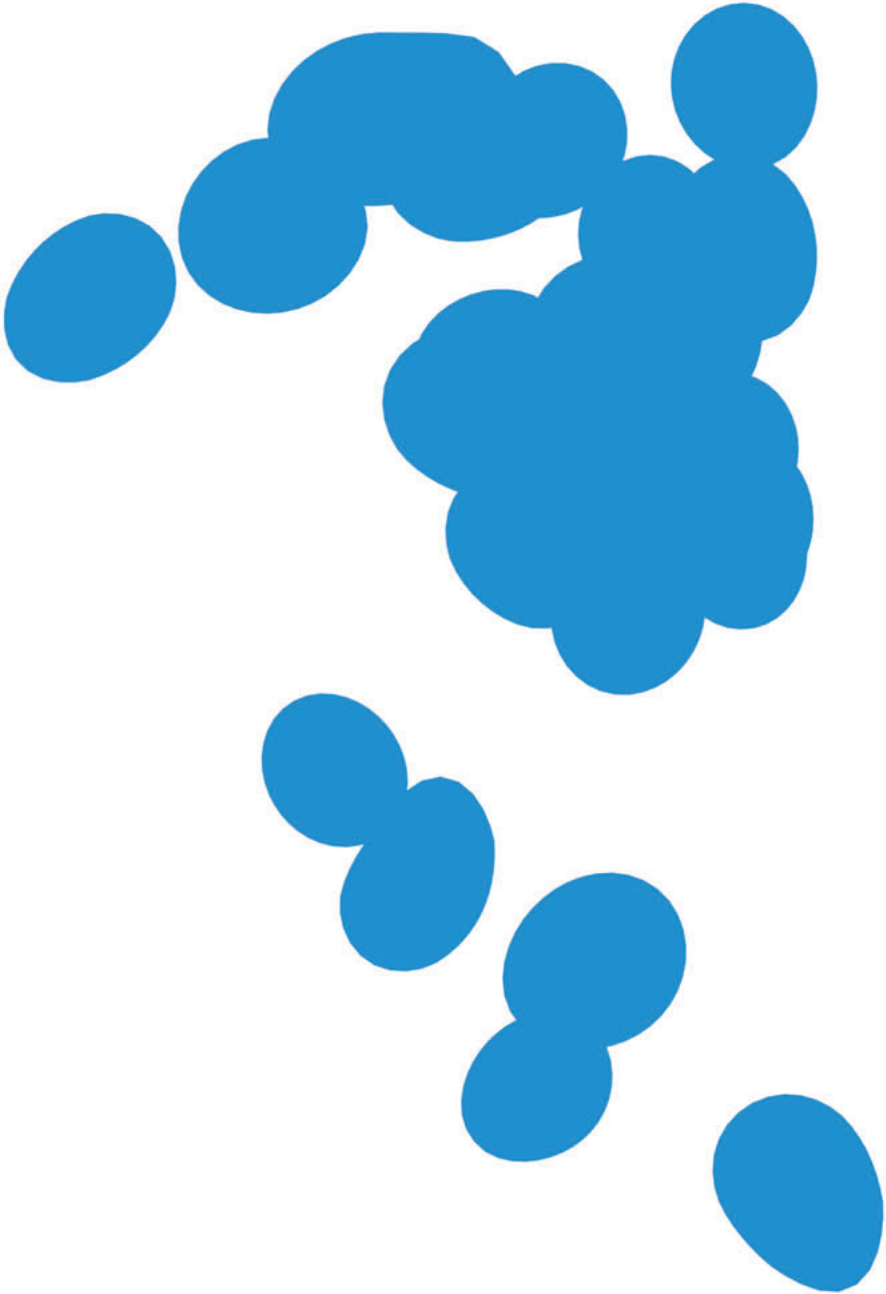
The overall mean  $d_N/d_S$  values of 0.249 (NA) and 0.362 (HA1) estimated by the PAML program (Table 2) were in agreement with previously reported values<sup>180</sup>. There was an abundance of negatively selected sites and a limited number of positively selected sites in both HA1 and NA (Table 3; see Tables S2 and S3 in the supplemental material). Generally, the number of positively selected sites was lower in NA compared with HA1. In contrast, DEPS detected more sites under directional selection in NA compared with HA1, which suggests that positive selection in HA1 is more random or it could reflect toggling of antigenic sites due to antibody pressure.

The positively selected codons 43, 46, and 52, are all located in the highly variable stalk region<sup>181</sup>. Residue 199, detected by MEME, was reported previously as an antigenic site<sup>182</sup>. Residue 267, detected by IFEL, is not an antigenic site but has been reported previously as being positively selected<sup>167</sup>. Two documented antigenic sites, 328<sup>134</sup> and 334<sup>123,134</sup>, were detected by DEPS. Residue 370, a reported antigenic residue<sup>123,134,183</sup>, was found by IFEL and DEPS as being positively selected. By IFEL, MEME and DEPS, the number of positively selected sites found to be within antigenic sites in NA was lower than in HA1.

In summary, we have analysed the evolutionary genetics of NA of influenza A (H3N2) viruses isolated from 1968 to 2009 and compared them with those of HA. High-level similarities were observed between the evolution of HA and NA, although notable differences were apparent. Future research on the evolution of NA should focus at the phenotypic level, using serological tests and antigenic cartography methods, as described recently<sup>122,137,159</sup>, on extensive virus data sets. Such analysis would allow not only the side-by-side comparison of HA and NA evolution at the genetic level as carried out here, but also its relationship with immune escape, the major driver of evolution of the surface glycoproteins of influenza A virus. Such analyses, along with full virus genome data, ultimately may lead to a better understanding and increased predictability of the evolution of influenza A (H3N2) virus.

## **ACKNOWLEDGMENTS**

The authors would gratefully like to thank Dr. Mathilde Richard for constructive discussions, Stefan van der Vliet and Oanh Vuong for excellent technical assistance and Ton Marzec for providing part of the influenza A (H3N2) virus isolates. We are grateful for the outstanding training received at the 15<sup>th</sup> and 16<sup>th</sup> International Bioinformatics Workshop on Virus Evolution and Molecular Epidemiology, September 2009, Rotterdam and 2010, Baltimore (<http://www.rega.kuleuven.be/cev/workshop/>). This work was supported by an NWO-VICI grant and NIH contract no. HHSN266200700010C.





## CHAPTER 3

# Genome-wide Analysis of Reassortment and Genomic Evolution of Human Influenza A(H3N2) Viruses Circulating between 1968 and 2011

Kim B. Westgeest, Colin A. Russell, Xudong Lin, Justin Bahl, Monique I. J. Spronken,  
Theo M. Bestebroer, Ruud van Beek, Eugene Skepner, Rebecca A. Halpin,  
Jan C. de Jong, Guus F. Rimmelzwaan, Albert D.M.E. Osterhaus, Derek J. Smith,  
David E. Wentworth, Ron A.M. Fouchier, Miranda de Graaf

*Journal of Virology*, volume 88(5), 6 February 2014, pages 2844–2857



## **ABSTRACT**

Influenza A(H3N2) viruses became widespread in humans during the 1968 H3N2 virus pandemic and have been a major cause of influenza epidemics ever since. These viruses evolve continuously by reassortment and genomic evolution. Antigenic drift is the cause for the need to update influenza vaccines frequently. Using two data sets that span the entire period of circulation of human influenza A(H3N2) viruses, it was shown that influenza A(H3N2) virus evolution can be mapped to 13 antigenic clusters. Here we analysed the full genomes of 286 influenza A(H3N2) viruses from these two data sets to investigate the genomic evolution and reassortment patterns. Numerous reassortment events were found, scattered over the entire period of virus circulation, but most prominently in viruses circulating between 1991 and 1998. Some of these reassortment events persisted over time, and one of these coincided with an antigenic cluster transition. Furthermore, selection pressures and nucleotide and amino acid substitution rates of all proteins were studied, including those of the recently discovered PB1-N40, PA-X, PA-N155, and PA-N182 proteins. Rates of nucleotide and amino acid substitutions were most pronounced for the haemagglutinin, neuraminidase, and PB1-F2 proteins. Selection pressures were highest in haemagglutinin, neuraminidase, matrix 1, and nonstructural protein 1. This study of genotype in relation to antigenic phenotype throughout the period of circulation of human influenza A(H3N2) viruses leads to a better understanding of the evolution of these viruses.

## INTRODUCTION

Influenza A viruses are members of the *Orthomyxoviridae* family, which comprises enveloped, negative-sense, single-stranded RNA viruses containing a genome divided over eight segments. The eight segments include: basic polymerase 2 (PB2), basic polymerase 1 (PB1), acidic polymerase (PA), haemagglutinin (HA), nucleoprotein (NP), neuraminidase (NA), matrix (M), and nonstructural protein (NS).

The segmented nature of the genome allows for the exchange of entire genes between different influenza viruses during simultaneous infection of a host, a process called reassortment. Intersubtypic reassortments between swine, avian, and/or human influenza A viruses have led to several pandemics. These pandemic viruses later became established as seasonal influenza viruses causing annual epidemics. Influenza virus type A is subdivided based on the antigenic properties of the major surface glycoproteins: HA and NA. To date, 17 HA<sup>14,15</sup> and 10 NA<sup>16,17</sup> subtypes have been found in nature.

The 1957 H2N2 influenza pandemic emerged as a consequence of reassortment between the then circulating seasonal influenza A(H1N1) virus and an avian influenza A(H2N2) virus<sup>84</sup>. The A(H2N2) virus caused annual epidemics until the 1968 H3N2 influenza pandemic emerged. This pandemic was the result of reassortment of a human A(H2N2) virus with an avian influenza A(H3N2) virus<sup>84</sup>. A(H3N2) viruses have been a major cause of influenza epidemics ever since with significant morbidity and mortality<sup>40,154</sup>. A(H2N2) and A(H3N2) viruses cocirculated until 1971<sup>38</sup> after which A(H2N2) viruses became extinct in the human population.

Reassortment between influenza viruses of the same subtype (intrasubtypic reassortment) together with genomic evolution, is presumably one of many ways for the virus to increase diversity<sup>95,97,99</sup>, shaping the short-term evolution of influenza A viruses<sup>97</sup>. Mutations in the major surface glycoproteins HA and NA resulting from antibody pressure, a process known as antigenic drift<sup>157</sup>, increase diversity but, more importantly, are a way for the virus to effectively evade the host's immune system.

In 2004, Smith et al. mapped the antigenic evolution of the HA proteins of A(H3N2) viruses from their introduction in 1968 until 2003<sup>107</sup>. Their study was based on an extensive data set of A(H3N2) virus isolates obtained from each consecutive influenza season. The study revealed 11 antigenic clusters named after the first vaccine strain of each cluster: A/Hong Kong/1/1968 (HK68), A/England/42/1972 (EN72), A/Victoria/3/1975 (VI75), A/Texas/1/1977 (TX77), A/Bangkok/1/1979 (BK79), A/Sichuan/2/1987 (SI87), A/Beijing/353/1989 (BE89), A/Beijing/32/1992 (BE92), A/Wuhan/359/1995 (WU95), A/Sydney/5/1997 (SY97), and A/Fujian/411/2002 (FU02).

Each of these antigenic clusters contains viruses that are antigenically similar for some time, after which a “cluster transition” warrants a vaccine update. Recently two additional major antigenic clusters were described by de Jong et al.: A/California/7/2004 (CA04) and A/Perth/16/2009 (PE09)<sup>160</sup>.

Genetic analysis of the NA and HA1 proteins of A(H3N2) viruses from these two studies revealed a large number of reassortment events and an asynchronous and slightly slower evolution of NA than that of HA1<sup>198</sup>. Here, we extended this analysis to the full genome. In total 284 full A(H3N2) virus genomes comprising viruses described by Smith et al.<sup>107</sup> and de Jong et al.<sup>160</sup> were submitted to GenBank. In addition, two vaccine/reference strains from GenBank were included in our data set, resulting in 286 full genomes spanning more than 40 years of A(H3N2) virus evolution. These full genomes were analysed with a focus on reassortment, rates of evolutionary change, and selection pressures of the eight segments and 15 proteins. The results presented in this study contribute to a better understanding of the complex evolution of A(H3N2) viruses.

## **MATERIALS AND METHODS**

### **Viruses**

A total of 284 A(H3N2) viruses isolated between 1968 and 2011<sup>107,160,198</sup> were subjected to full genome sequencing. Viruses were propagated as described previously<sup>198</sup> and full genome sequencing was performed as described previously<sup>198</sup> or by the J. Craig Venter Institute (JCVI). Sequences of two of the vaccine strains that were included in this study were retrieved from GenBank, as well as a number of sequences from segments that were not fully sequenced.

Of the 284 A(H3N2) viruses, 281 viruses were sequenced at JCVI. Viral RNA was purified using a Zymo Research Corporation 96 Viral RNA kit, and the complete genome was simultaneously amplified from 3 µl of purified RNA by using a multisegment reverse transcription-PCR (M-RT-PCR) strategy<sup>199,200</sup>. The majority of the samples were sequenced using a PCR/Sanger high-throughput sequencing pipeline<sup>201</sup>. Primer sequences are available upon request. Thirteen samples [A/Bangkok/1/1979, A/Auckland/4382/1982(mixed), A/Netherlands/233/1982, A/Leningrad/360/1986, A/Victoria/1/1988, A/Oklahoma/5/1988, A/Beijing/353/1989, A/United Kingdom/261/1991, A/Brisbane/8/1996, A/Netherlands/5/1998, A/Netherlands/301/1999, A/Netherlands/009/2010, and A/Netherlands/034/2010] were sequenced by M-RT-PCR<sup>199,200</sup>, followed by sequence-independent single-primer

amplification (SISPA)<sup>202</sup> and then combined next-generation sequencing using a 454/Roche GS-FLX instrument and an Illumina Genome Analyzer II instrument.

### Sequence preparation

Nucleotide sequences of 286 complete A(H3N2) virus genomes (details on the segments can be found in Table S1 in the supplemental material, available in JVI Online) were aligned using the ClustalW program running within the BioEdit software package, version 7.0.9.0<sup>184</sup>. Nucleotides before the first start codon and after the last stop codon of each segment were removed. The sequences were manually edited and translated into amino acids by using BioEdit. Ambiguous amino acids were assigned either a “B” (indicating aspartate or asparagine) or an “X” (indicating any amino acid). The stop codons of the amino acid alignments of PB1-F2 and PA-X were replaced by “X.” Both the nucleotide and amino acid NS1 alignments were C-terminally truncated to a length of 220 codons/amino acids (aa) due to stop codons at codons/aa 221 and 231 in some of the NS1 sequences.

### Phylogeny

We inferred maximum likelihood (ML) phylogenetic trees by using the GTR+ $\Gamma$ 4+I (the general time reversible model with the proportion of invariant sites and the gamma distribution of among-site rate variation with four categories estimated from the empirical data) model of nucleotide substitution and the PhyML package, version 3.0<sup>187</sup> performing a full heuristic search and subtree pruning and regrafting (SPR) searches. Garli version 0.951<sup>188</sup> was used to perform 1 million generations of the best nucleotide tree from PhyML to optimize tree topology and branch lengths. The reliability of all phylogenetic groupings of each tree was determined through a nonparametric bootstrap resampling analysis with Garli: 1,000 replicates of ML trees were analysed by applying the GTR+ $\Gamma$ 4+I model of nucleotide substitution. Detailed trees, including bootstrap values, are shown in Figure S1 in the supplemental material. Trees were visualized through the FigTree program version 1.3.1. (<http://www.tree.bio.ed.ac.uk/software/figtree>).

### Rates of amino acid evolution

To determine the ML-optimized best-fit models of protein evolution, amino acid sequence alignments were subjected to analysis by ProtTest<sup>186</sup> (see Table S2 in the supplemental material). ML trees were inferred by applying the PhyML package, version

3.0<sup>187</sup> in combination with the best-fit model of protein evolution based on the Akaike information criterion (AIC), and performing a full heuristic search and SPR searches.

For all A(H3N2) viruses, the genetic distance of each open reading frame (ORF), HA1, or HA2 to A/Hong Kong/1/68 was calculated from the phylogenetic tree and plotted as a function of time.

### **Detection of reassortment**

TreeMap version 1.0 (<http://taxonomy.zoology.gla.ac.uk/rod/treemap.html>) was used to generate tanglegrams with the phylogenetic ML trees.

Reassortment events were identified by the graph-incompatibility-based reassortment finder (GiRaF) program<sup>162</sup>, as described previously<sup>198</sup>. Nucleotide alignments of the full segments were used as input for MrBayes<sup>190,191</sup>. The best-fit models of nucleotide substitution were determined using jModelTest<sup>185</sup>. As input for the GiRaF program, 1,000 unrooted candidate trees were inferred by using the GTR+ $\Gamma$ 4+I (PB1, HA, and NP) or GTR+ $\Gamma$ 4 (PB2, PA, NA, M, and NS) substitution model, a burn-in of 50 % (100,000 iterations), and sampling every 200 iterations. These trees were subsequently used to model the phylogenetic uncertainty for each segment, using the GiRaF program with default settings. The default confidence threshold was 0.7; for our data set, all events reported by GiRaF had confidence levels of > 0.83, with almost all falling above 0.9. This procedure was repeated 50 times with 50 independent MrBayes tree files per segment. Rather than looking at only one pair of segments, GiRaF was applied to all 28 combinations of segments. This was done to more comprehensively catalog reassortment events, making the analysis even more stringent<sup>162</sup>.

### **Estimation of nucleotide substitution rates and times of divergence**

Overall rates of evolutionary change (number of nucleotide [nt] substitutions per site per year) and times of circulation of the most recent common ancestor (MRCA) (in years) were estimated using the BEAST program, version 1.7.0 (<http://beast.bio.ed.ac.uk/>)<sup>193</sup>. For all analyses, the uncorrelated log-normal relaxed molecular clock was used to accommodate variation in the molecular evolutionary rate among lineages, in combination with the GTR+ $\Gamma$ 4+I (PB1, HA, HA1, and NP), GTR+ $\Gamma$ 4 (PB2, PB1-F2, PB1-N40, PA, PA-N155, PA-N182, HA2, NA, M, M2, NS, NS1, and NEP), or GTR+I (PA-X and M1) substitution model. Isolation dates were used to calibrate the molecular clock. This analysis was conducted using a time-aware linear Bayesian skyride coalescent tree prior<sup>195</sup> over the unknown tree space, relatively uninformative priors on

all model parameters, and a normal prior on the mean skyride size (log units) of 11.0 (standard deviation [SD], 1.8)<sup>203</sup>.

Sequences were grouped into the following 13 taxon sets based on the antigenic clusters of HA<sup>107,160</sup>: HK68, EN72, VI75, TX77, BK79, SI87, BE89, BE92, WU95, SY97, FU02, CA04, and PE09. Three independent Bayesian Markov chain Monte Carlo (MCMC) analyses were performed on all segments, coding regions, HA1, and HA2 for 50 million states, with sampling every 2,000 states. Convergences and effective sample sizes of the estimates were checked with Tracer, version 1.5 (<http://tree.bio.ed.ac.uk/software/tracer/>). These analyses were combined with LogCombiner, version 1.7.0, with a burn-in of 10 to 20 % to reach full convergence. Uncertainty in parameter estimates was reported as the 95 % highest posterior density (HPD). A Bayes factor test was employed to evaluate if the times of circulation of the MRCAs of the gene segments of a given antigenic cluster were significantly different<sup>39,95,204</sup> (see Table S3 in the supplemental material).

### **Selection pressures**

Selection pressures were measured for all codon alignments in combination with the ML trees of each segment. To determine the degree of natural selection acting on all protein coding regions, the average numbers of nonsynonymous ( $d_N$ ) and synonymous ( $d_S$ ) substitutions per codon ( $d_N/d_S$  ratio) were estimated for the entire tree, the internal branches, and the external branches by using single-likelihood ancestor counting (SLAC)<sup>205</sup>.

Positively selected codons along internal branches were detected with the internal fixed-effects likelihood (IFEL) method<sup>164</sup>. Codons subject to episodic diversifying selection were identified with the mixed-effects model of evolution (MEME) method<sup>165</sup>. Codons with P values of < 0.05 were reported as positively selected codons.

Fast unconstrained Bayesian approximation (FUBAR) was used to rapidly detect negative and positive selection by Bayesian MCMC analyses to robustly account for parameter estimation errors<sup>206</sup>. Codons with posterior probabilities of  $\geq 0.9$  were reported as being either negatively or positively selected.

Directional positive selection was detected using the directional evolution in protein sequences (DEPS) test<sup>167</sup>. For this procedure, the ML trees were combined with the amino acid alignment. Residues with a Bayes factor of > 20 were reported as positively selected.

All methods were implemented in the HyPhy package<sup>166</sup>, combined with the best-fit model of nucleotide substitution or, in the case of the DEPS test, the best-fit

model of protein evolution (see Table S2 in the supplemental material) based on the AIC. The FUBAR analyses were accessed through the Datamonkey webserver<sup>207,208</sup>.

Because of overlapping ORFs in the PB1, PA, M, and NS segments, these analyses were restricted to the nonoverlapping ORFs of PB1, PA, M1, M2, NS1, and NEP. The  $d_N/d_S$  ratios for PB1-F2 and PA-X were biased by the +1 ORF from PB1 or PA and were therefore excluded in the SLAC, IFEL, MEME, and FUBAR analyses.

#### **Nucleotide sequence accession numbers.**

Nucleotide sequences presented in this study are available from GenBank under the following accession numbers: CY012104 to CY012111, CY031812, CY033638, CY033639, CY033641 to CY033645, CY034108, CY034109, CY034111 to CY034115, CY035025, CY035027, CY039093, CY039094, CY077825, CY112233 to CY112318, CY112320 to CY112368, CY112396 to CY112411, CY112420 to CY112564, CY112566 to CY112630, CY112632 to CY112972, CY112981 to CY113004, CY113013 to CY113028, CY113037 to CY113109, CY113111 to CY113476, CY113485 to CY113820, CY113829 to CY114348, CY114357 to CY114516, CY116573 to CY116575, CY116587 to CY116591, CY116594, CY116596, CY116600, CY121181, DQ508846 to DQ508850, DQ508852, DQ508853, GQ293081, GQ293082, EU283414, EU597800 to EU597805, GQ293081, GQ293082, HQ166049 to HQ166056, and KC296462 to KC296481. Sequences for PB2, PB1, PA, NP, M, and NS of A/Perth/16/2009 are available from the Global Initiative on Sharing Avian Influenza Data (GiSAID) database under accession numbers EPI272741 to EPI272746.

## **RESULTS**

### **Dataset background**

To study the genotype in relation to the antigenic phenotype throughout the period of circulation of human A(H3N2) influenza viruses, we used a data set comprising 286 full genomes of viruses sampled between 1968 and 2011. The antigenic phenotype of each of these viruses was previously determined, and this data set represents all 13 antigenic clusters of viruses<sup>107,160</sup>. Thus far, only the HA1 and NA sequences of these viruses have been studied<sup>107,198</sup>. To extend these studies to the full genome, 284 isolates were sequenced. Sequences of two vaccine/reference strains were obtained from GenBank. Deduced amino acid sequences were aligned, together with the sequences of the two vaccine/reference strains (see Figure S2 in the supplemental material).



### Genomic evolution

ML trees were inferred to study the nucleotide evolution of all segments (Figure 1). The ML trees of each segment were colour coded based on antigenic clusters<sup>107,160</sup>. All trees displayed a similar topological structure with a typical ladder-like gradual evolution, with rapid replacement of old viruses by newer ones. For the PB2, PA, NP, and NA segments, the early (1968 to 1970) HK68-like viruses were separated from the late (1970 to 1972) HK68-like viruses by a long branch. This was most likely due to cocirculation and reassortment with A(H2N2) viruses in 1969 to 1971 cocirculation<sup>38</sup>.

For all segments, the viruses on the trunk were the ancestors of all variants in future years. However, in some cases, viruses belonging to an antigenic cluster were not seeded by the viruses from the antigenic cluster preceding it. For example, the two TX77-like viruses (1976 to 1977) were seeded by the previously circulating EN72-like (1972 to 1975) viruses rather than by the VI75-like viruses. This was more prominent for the BE89-like, BE92-like, and WU95-like viruses circulating between 1992 and 1998. There were, however, differences in the origins of the different segments. In the PB2, PA, HA, NP, M, and NS trees, the SI87-like viruses were the progenitors of the BE92-like viruses instead of the BE89-like viruses. As a result, the BE89-like viruses clustered away from the trunk in an evolutionarily terminal clade. In contrast, in the PB1 and NA trees, the BE89-like viruses were the progenitors of the BE92-like viruses.

By applying the antigenic cluster colours of HA<sup>107,160</sup> to all trees of the individual segments, it became clear that the topologies, although similar, were not identical to the clustering observed for HA. These observations strongly suggest frequent reassortment between segments.

### Reassortment of A(H3N2) viruses

To investigate reassortment over the entire genome, a combined approach of tanglegrams and GiRaF<sup>162</sup> analyses<sup>198</sup> was applied (Table 1 and Figure 2; see Table S4 in the supplemental material). In short, tanglegrams were made to enable visualization of the locations of particular isolates within the HA tree and each of the trees of the other segments. In the absence of reassortment, the twines should connect both trees, in theory, in a seamlessly horizontal way. This was the case for some of the isolates used in this study, but not for all, thus suggesting reassortment between the segments. To explore whether these were indeed reassortment events, multiple trees per segment were analysed for incompatible splits in each tree pair by using the GiRaF program. We performed 50 independent GiRaF analyses on 50 independent MrBayes runs per

segment. Reassortment events found in  $\geq 50\%$  of the 50 independent GiRaF runs are indicated in Table 1. The 50% cut-off was chosen arbitrarily. The events found in 100% of the 50 independent GiRaF runs and that included the HA segment are indicated with asterisks in Table 1 and depicted on the tanglegrams as bold twines in Figure 2. Numbers in Figure 2 correspond to the numbers in Table 1.

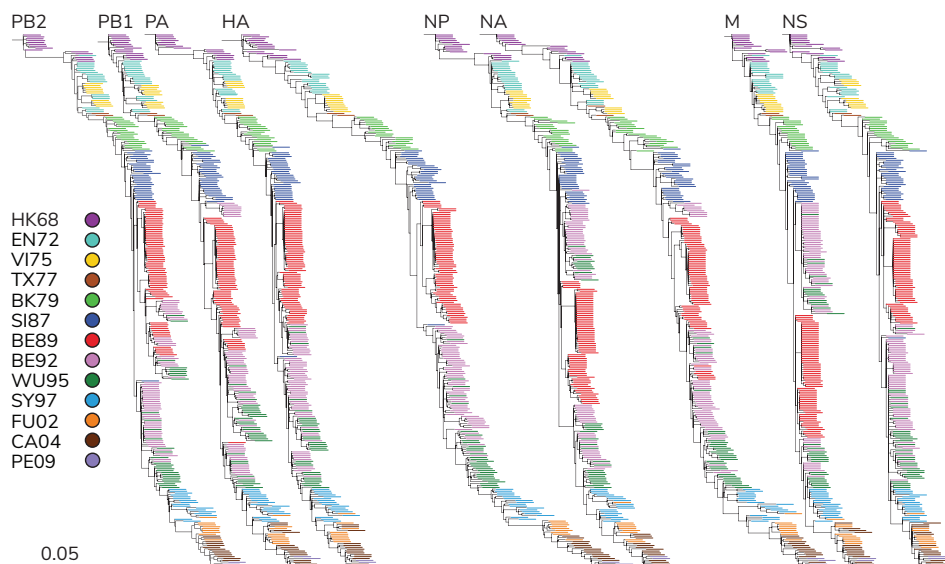


Figure 1 | ML trees of all segments of A(H3N2) viruses circulating between 1968 and 2011. The ML trees of PB2, PB1, PA, HA, NP, NA, M, and NS were generated with 286 nucleotide sequences per segment. Scale bars roughly represent 5% of nucleotide substitutions between close relatives. The colour coding of viruses is based on the antigenic clusters of HA HA<sup>107,160</sup> and is consistent between all trees. Trees were rooted on A/Hong Kong/1/1968.

Table 1 | Reassortment events between segments of A(H3N2) viruses circulating between 1968 and 2011 analysed with GiRaF<sup>a,b</sup>.

Event <sup>c</sup>	% <sup>d</sup>	Year of isolation	Cluster	Segments involved <sup>e</sup>	
				Virus 1	Virus 2
2	68	1970	HK68	HA PB2	PA
3	56	1977	VI75	HA (PA)	M (NA)
4	74	1982	BK79	NA	NS
5*	100	1992-1993	BE92	HA M NP PA	PB2

Table 1 | Reassortment events between segments of A(H3N2) viruses circulating between 1968 and 2011 analysed with GiRaF<sup>a,b</sup>.

Event <sup>c</sup>	% <sup>d</sup>	Year of isolation	Cluster	Segments involved <sup>e</sup>	
				Virus 1	Virus 2
6	90	1993	BE92	NA (NS) PB1	NP
7	88	1993	BE92	HA NS PB2	PA
8	78	1993	BE92	HA	PA
9	78	1993	BE92	HA PA PB2	NA
10	56	1994-1996	BE92	HA	M PA
11*	100	1994, 1996	BE92	HA PA PB2	NP
12	98	1995	BE92	(M) NS	PB2
13	68	1995	BE92	HA NA	M (NP)
14*	100	1996	BE92	HA	M NP PB1 PB2
<b>15*</b>	<b>100</b>	<b>1992-2011</b>	<b>BE92-end</b>	<b>HA M NP</b>	<b>NA PB1</b>
<b>16*</b>	<b>100</b>	<b>1992-2011</b>	<b>BE92-end</b>	<b>HA NP NS</b>	<b>NA</b>
<b>17</b>	<b>84</b>	<b>1993-2011</b>	<b>BE92-end</b>	<b>HA M NP</b>	<b>PB1</b>
<b>18*</b>	<b>100</b>	<b>1994-2011</b>	<b>BE92-end</b>	<b>HA</b>	<b>M</b>
19	90	1992-1997	BE92/WU95	HA NP PA	PB1
20*	100	1993	BE92/WU95	HA M NP	NS (PB2)
21*	100	1993	BE92/WU95	HA (PB2)	(NA) NP (PA)
22*	100	1993-1994	BE92/WU95	HA M NP	PA
23*	100	1993-1994	BE92/WU95	HA NS PB2	NA
24*	100	1993-1994	BE92/WU95	HA NS PB2	NA
25	54	1993-1994	BE92/WU95	HA NS PB2	PB1
26	86	1993-1996	BE92/WU95	NA PA PB1	NP
27	84	1993-1996	BE92/WU95	NP	(NS) (PB1) PB2
28	72	1993-1996	BE92/WU95	HA NS PB2	NP
29*	100	1993-1997	BE92/WU95	HA M NP	PB2
30*	100	1993-1997	BE92/WU95	HA M NP	PB2
31	92	1994-1997	BE92/WU95	M NS PA	PB2
32*	100	1995-1996	BE92/WU95	(HA) M NA PA PB1 PB2	NP
33*	100	1995-1996	BE92/WU95	HA NS PB1	M NP PB2
34*	100	1995-1997	BE92/WU95	HA M NS PA	PB2
35	100	1993	WU95	-	-
36	96	1995	WU95	HA M	NP PB1 PB2

Table 1 | Reassortment events between segments of A(H3N2) viruses circulating between 1968 and 2011 analysed with GiRaF<sup>a,b</sup>.

Event <sup>c</sup>	% <sup>d</sup>	Year of isolation	Cluster	Segments involved <sup>e</sup>	
				Virus 1	Virus 2
37*	100	1995-1996	WU95	M NP (PA)	NA NS
38	98	1996	WU95	HA (NA)	M NP PA
39	50	1996	WU95	HA	(M) NP PA
40*	100	1996-1997	WU95	M	PB1
41	100	1996-1997	WU95	NA PB1	(NP) M (PA)
42*	100	1997	WU95	(HA) PA PB2	M NA NP PB1
43*	100	1997	WU95	(HA) M (NA) NP PB2	PA
44*	100	1997	WU95	HA (M)	NP (PB1)
45	100	1997-1998	WU95	M NP PA	PB2
46*	100	1997-1998	WU95	HA PB2	NA NP PB1
47*	100	1998	WU95	HA (M)	(NA) NP PA
<b>48</b>	<b>100</b>	<b>1995-2011</b>	<b>WU95-end</b>	<b>M NA</b>	<b>PB1</b>
<b>49</b>	<b>72</b>	<b>1996-2011</b>	<b>WU95-end</b>	<b>(HA)</b>	<b>(PA) (PB1)</b>
<b>50</b>	<b>68</b>	<b>1996-2011</b>	<b>WU95-end</b>	<b>NA PB1 PB2</b>	<b>PA</b>
51*	100	1999	SY97	HA	NA
52	54	2001	SY97	(M) NP	PB2
53	76	2000-2001	SY97	-	-
54*	100	2003	FU02	HA	M NA NP NS PA PB1 PB2
<b>55*</b>	<b>100</b>	<b>2002-2011</b>	<b>FU02-end</b>	<b>HA</b>	<b>M NP PA PB1</b>
56	90	2005	CA04	HA	NP
57*	100	2006	CA04	HA PA	M PB1
<b>58*</b>	<b>100</b>	<b>2005-2011</b>	<b>CA04-end</b>	<b>HA PB2</b>	<b>M PB1</b>
<b>59</b>	<b>100</b>	<b>2007-2011</b>	<b>CA04-end</b>	<b>M PB1</b>	<b>(NA) (NP) PB2</b>

<sup>a</sup>Persistent reassortment events are shown in bold. Details on isolates involved in the reassortment events can be found in Table S4 in the supplemental material.

<sup>b</sup>events with a support level of 100 % and involving HA are depicted as bold twines in the tanglegrams (Figure 2).

<sup>c</sup>Event numbering corresponds to that in Figure 2.

<sup>d</sup>Percentage of reassortment events observed in 50 independent GiRaF analyses of 50 independent MrBayes runs per segment.

<sup>e</sup>Involved segments with a support level of  $\geq 75$  % of the runs are indicated in parentheses.

Reassortment events during circulation of the HK68-like, VI75-like, BK79-like, BE92-like, WU95-like, SY97-like, FU02-like, CA04-like, and PE09-like viruses were identified by GiRaF analysis, although most did not persist over time. Reassortment was found for all eight segments, although the frequency of reassortment of the individual segments varied. The majority of these reassortment events occurred between viruses from one antigenic cluster (Table 1, events 1 to 14, 35 to 47, 51 to 54, 56, and 57), whereas between 1992 and 1997, reassortment events comprised viruses from multiple antigenic clusters, i.e., BE92-like and WU95-like viruses (events 19 to 34). Apart from nonpersistent reassortment events, several persistent reassortment events were established during the circulation of BE92-like (events 15 to 18), WU95-like (events 48 to 50), FU02-like (event 55), and CA04-like (events 58 and 59) viruses. No support for reassortment was found during circulation of the EN72-like, SI87-like, and BE89-like viruses. It should be noted that reassortment between highly similar sequences is likely to be missed. Similarly, discordance in phylogenies due to high similarity of some sequences can result in an under- or overestimation of reassortment events.

We investigated whether the timing of persistent reassortment coincided with antigenic cluster transitions. From our GiRaF analysis, the only cluster transition that was linked to a persistent reassortment event was the SY97-to-FU02 antigenic cluster transition (event 55) (Figure 2 and Table 1). Persistent reassortment events during circulation of BE92-like viruses were detected in late 1992, while the first BE92-like viruses were isolated in the beginning of 1992. The first WU95-like viruses were isolated in 1993, while the persistent reassortment events were not detected until late 1995. Similarly, persistent reassortment events for CA04-like viruses were only detected in 2005. In agreement with the GiRaF results, visual inspection of the tanglegrams revealed that the early viruses (and their ancestors) of each antigenic cluster were similarly positioned in each of the ML trees, with an exception for the reported reassortment event of SY95-FU02.

### **Time of circulation of the MRCA of each antigenic cluster**

We further studied reassortment in correlation with antigenic change in HA by estimating the time of circulation of the MRCA of each segment for viruses from each antigenic cluster (Figure 3). For the HA segment, the time of circulation of the MRCA roughly corresponds to the sample date of the first isolated virus of each antigenic cluster.

At the onset of the 1968 H3N2 influenza pandemic, the PB1 and HA segments were newly acquired through reassortment<sup>84,209–211</sup>. Similarly, our BEAST analyses showed that the MRCA of the HA and PB1 segments circulated at the onset of the 1968 H3N2 influenza pandemic. The MRCA of the PA segment circulated 2 years before this time, most likely because PA was introduced from the A(H2N2) virus lineage that circulated at that time.

The MRCA of each segment of the EN72-like viruses originated in 1972, with the onset of the antigenic cluster. This was indicative of a genomewide selective sweep (Figure 3). GiRaF found no support for reassortment during circulation of EN72-like viruses (Figure 2 and Table 1).

During circulation of the VI75-like viruses, the MRCA of the HA and M segments originated in 1975, while the other segments had MRCAs that circulated in 1973 and 1974 (Figure 3). This signifies a selective sweep for the HA and M segments, while the genetic diversity of the other segments was likely maintained by reassortment.

Strikingly, the EN72-like viruses, not the VI75-like viruses, seeded the TX77-like viruses (Figure 1). Neither GiRaF nor the time of circulation of the MRCAs gave an indication of reassortment for TX77-like viruses; however, the data set used comprised only two TX77-like viruses. The transition to BK79-like viruses coincided with a genomewide selective sweep (Figure 3).

Little information was available on the A(H3N2) viruses circulating between 1987 and 1992 (SI87-like viruses). The MRCA of the HA segment of SI87-like viruses circulated around 1986 to 1987, while the NA, M, and NS segments had MRCAs that circulated before that time (1984), indicative of reassorted NA, M, and NS segments.

In agreement with previous studies<sup>176,212</sup>, there was evidence for reassortment between BE89- and BE92-like viruses. The PB2, PB1, PA, NA, M, and NS segments of the BE92-like viruses shared roughly the same times of their MRCAs as those of the BE89-like viruses, indicating that these segments originated from BE89-like viruses through reassortment. This is in contrast to the HA segment, which underwent a selective sweep (Figure 3).

For the WU95-like viruses the times of MRCAs of the PB2, PA, M, and NS segments remained the same as those of the BE89-like viruses, also indicative of reassortment, while the PB1, HA, and NA segments underwent a selective sweep (Figure 3). Three persistent reassortment events were reported by GiRaF, all involving the HA segment. GiRaF analyses further suggested that reassortment events of A(H3N2) viruses occurred particularly frequently during the circulation of BE92- and WU95-like viruses (Figure 2

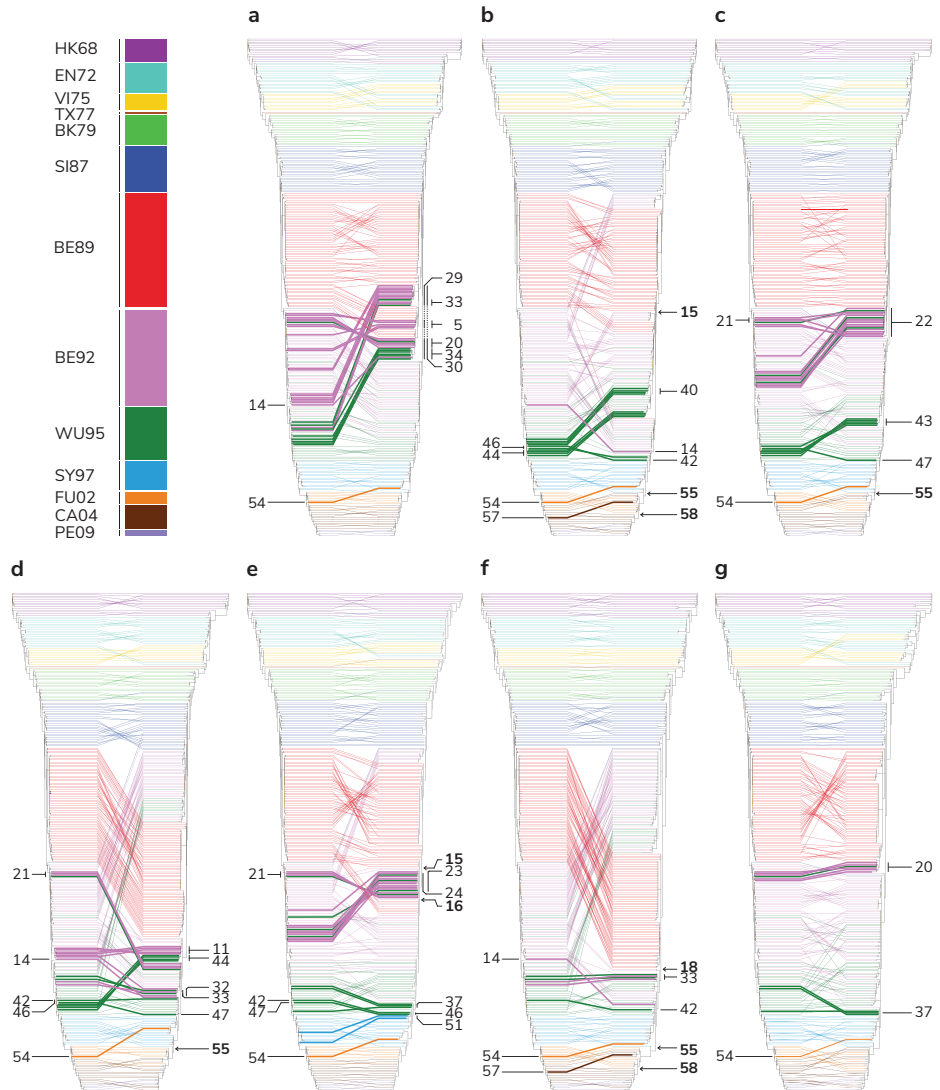


Figure 2 | **Reassortment events between segments of A(H3N2) viruses circulating between 1968 and 2011.** Tanglegrams are displayed with the ML HA tree on the left side and the mirrored trees of PB2 (A), PB1 (B), PA (C), NP (D), NA (E), M (F), and NS (G) on the right side. Twines between both trees are colour coded according to the antigenic clusters of HA (see the legend to Figure 1). GiRaF was used to detect reassortment events between all segments. Reassortment events with a support of 100 % between HA and any other segment are depicted as bold twines, and numbers correspond to reassortment events in Table 1. The arrows indicate the introduction of new segments by reassortment into the population that persisted until the last sampled isolate.

and Table 1). The increase in reassortment rates during the circulation of the BE89-like, BE92-like, and WU95-like viruses could be explained partially by cocirculation: BE89-like viruses circulated from mid-1989 until early-1993, BE92-like viruses circulated from late 1991 until late 1996, and WU95-like viruses circulated from early 1993 until early 1998.

The SY97-like-to-FU02-like virus antigenic cluster transition has been well described and coincided with reassortment<sup>95,96</sup>. Our BEAST analyses showed an MRCA of the HA segment that circulated around 2002, while the other segments had MRCAs that circulated before that time (1998 to 1999) (Figure 3). This is indicative of a selective sweep for the HA segment, while the genetic diversity of the other segments was retained. Next to one nonpersistent reassortment event, GiRaF found a persistent reassortment event involving the HA segment of a different virus lineage compared to that of the PB1, PA, NP, and M segments. Moreover, a large genetic distance was seen between SY97-like and FU02-like viruses, which was represented by a long branch separating the two clades in the HA ML tree, and this is in line with a previous study<sup>213</sup>.

The transition of FU02-like to CA04-like viruses coincided with a selective sweep for the PB2, HA, NA, and NS segments according to the BEAST analyses. GiRaF provided evidence of two persistent reassortment events, in which the PB1 and M segments were derived from a virus lineage different from that of the PB2 segment. The PB2 and NS segments have an MRCA that circulated before the MRCA of HA for the PE09-like viruses, suggesting reassortment of the PB2 and NS segments and a selective sweep of the HA segment.

### **Nucleotide substitution rates of all A(H3N2) open reading frames**

Rates of nucleotide substitution (number of nt substitutions/site/year) for all segments, and the proteins they encode, were estimated using BEAST (Table 2). The PB2, HA, NP, and NA segments each encode single proteins: PB2, HA, NP, and NA. The PB1 segment encodes three proteins: PB1; an N-terminally truncated version of PB1 starting at codon 40 (PB1-N40)<sup>42</sup> and PB1-frame 2 (PB1-F2), expressed through an alternative ORF<sup>43</sup>. The PA segment encodes four proteins: PA; PA-X, through a second ORF accessed via ribosomal frameshifting<sup>44</sup>; and two N-terminally truncated versions of PA, i.e., PA-N155 and PA-N182<sup>45</sup>. The M and NS segments each individually encode proteins from the colinear transcript: M1 and NS1. Additionally, they encode a second protein through mRNA splicing: M2 and NS2, known as the nuclear export protein (NEP)<sup>46</sup>. The HA protein is composed of two subunits that are cleaved by host proteases from their



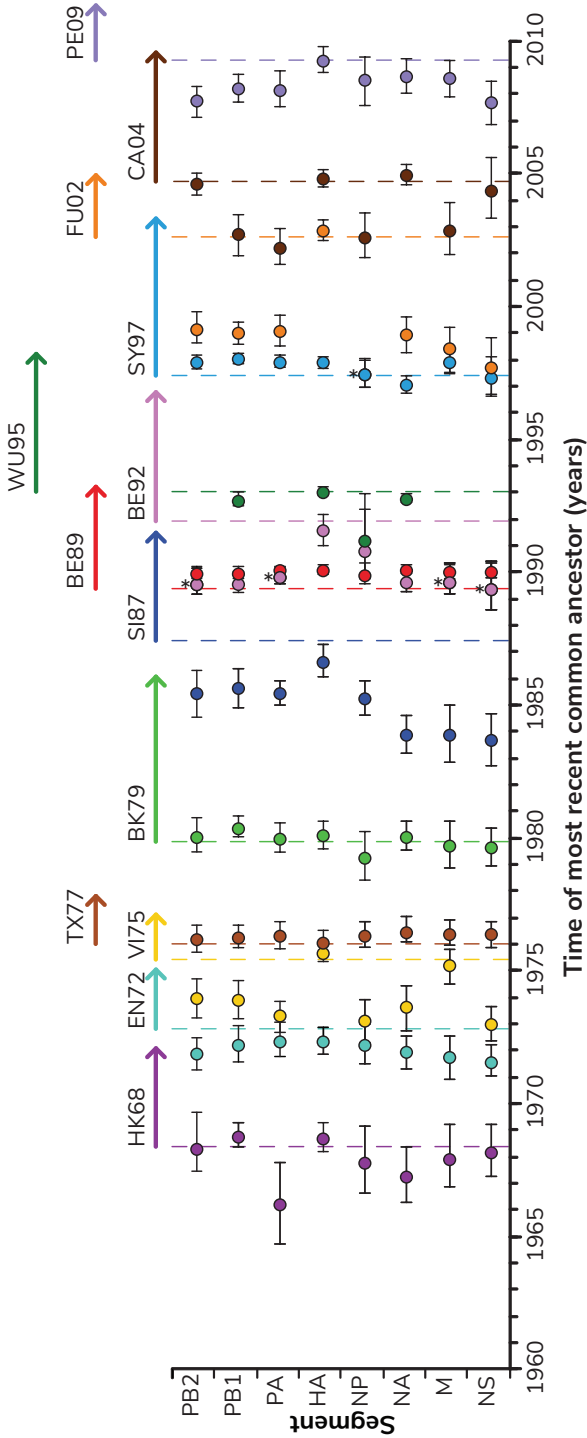


Figure 3 | Time of circulation of the MRCA of A(H3N2) viruses of each antigenic cluster. The time of circulation of the MRCA of each genomic segment is shown. The values shown represent the means (spheres) and 95 % highest posterior density intervals (error bars) for the times of circulation of the MRCAs estimated across trees sampled using Bayesian MCMC analyses. Dashed lines show the date of the first sampled isolate of each antigenic cluster. Arrows show the time span from the first sampled isolate until the last sampled isolate of the data set for each antigenic cluster. Asterisks represent two data points on top of each other (i.e., BE89/WU95 and SY97/FU02). Colour coding was done according to the antigenic clusters of HA (see the legend to Figure 1).

precursor, HA0<sup>49</sup>: HA1, which is the immunogenic section of the HA protein<sup>52,53</sup>, and HA2.

The mean rates of nucleotide substitution of the individual segments varied from  $2.07 \times 10^{-3}$  to  $3.99 \times 10^{-3}$  nt substitutions/site/year, with the highest rates of nucleotide substitution for the major glycoproteins HA and NA and the lowest rate for the M segment. Multiple ORFs carried by the same segment showed different rates of nucleotide substitution. The PB1 segment and the PB1-N40 ORF shared the same rate ( $2.54 \times 10^{-3}$  nt substitutions/site/year), in contrast to PB1-F2, which showed a rate comparable to those of the HA and NA segments ( $3.67 \times 10^{-3}$  nt substitutions/site/year). PA-X displayed a much lower rate ( $1.58 \times 10^{-3}$  nt substitutions/site/year) than that of the PA segment ( $2.28 \times 10^{-3}$  nt substitutions/site/year). Both PA-N155 and PA-N182 had a rate comparable to that of the PA segment. Upon comparing the two HA subunits, we noted that HA1 displayed a higher rate than HA2 ( $4.84 \times 10^{-3}$  and  $3.12 \times 10^{-3}$  nt substitutions/site/year).

#### **Amino acid substitution rates of all A(H3N2) proteins**

Linear regression plots for the year of sampling versus amino acid distance to A/Hong Kong/1/1968 in the ML tree were generated (Figure 4), and the rates of amino acid substitution were estimated based on the slopes (number of aa substitutions/site/year) (Table 2). The HA1 subunit showed the highest rate of amino acid substitution, whereas the rate for the HA2 subunit was much lower ( $14.9 \times 10^{-3}$  compared to  $1.4 \times 10^{-3}$  aa substitutions/site/year). This was likely due to the fact that HA1 is the major immunogenic region of HA<sup>52,53</sup>. PB1-F2 and NA also displayed high rates of amino acid substitution:  $9.5 \times 10^{-3}$  and  $9.1 \times 10^{-3}$  aa substitutions/site/year. For all other proteins, rates of amino acid substitution were at least three times lower. Although the differences were minimal, PA had the highest rate of amino acid substitutions of the polymerase complex proteins. Of the M proteins, M2 displayed a higher rate of amino acid substitution than M1 ( $1.4 \times 10^{-3}$  and  $0.9 \times 10^{-3}$  aa substitutions/site/year). The difference between the NS proteins was even larger, with a much higher rate of amino acid substitution for NS1 than for NEP ( $2.2 \times 10^{-3}$  and  $0.5 \times 10^{-3}$  aa substitutions/site/year, respectively). The recently discovered PB1-N40, PA-X, PA-N155, and PA-N182 proteins all showed low rates of amino acid substitution ( $0.8 \times 10^{-3}$  to  $1.2 \times 10^{-3}$  aa substitutions/site/year).

Most plots illustrated a linear trend, reflecting a constant rate of amino acid substitutions over time. For the HK68-like viruses, a sudden increase in genetic distance

**Table 2 | Mean rates of substitution for all segments and coding regions of A(H3N2) viruses circulating between 1968 and 2011.**

Segment	Region	Mean rate of substitution ( $\times 10^{-3}$ substitutions/site/year)				Amino acid <sup>b</sup>
		Nucleotide <sup>a</sup>			Amino acid <sup>b</sup>	
		Mean	95 % HPD			
			Lower	Upper		
PB2	Full	2.41	2.17	2.67	1.4	
PB1	Full	2.54	2.31	2.78	0.8	
	PB1-N40	2.54	2.30	2.78	0.8	
	PB1-F2	3.67	2.81	4.58	9.5	
PA	Full	2.28	2.05	2.52	1.8	
	PA-N155	2.39	2.12	2.68	1.1	
	PA-N182	2.41	2.12	2.71	1.2	
	PA-X	1.58	0.98	2.20	1.6	
HA	HA0	3.99	3.63	4.38	10.8	
	HA1	4.84	4.32	5.38	14.9	
	HA2	3.12	2.63	3.64	1.4	
NP	Full	2.51	2.21	2.82	2.9	
NA	Full	3.27	2.93	3.63	9.1	
M	Full	2.07	1.75	2.40	-	
	M1	1.88	1.55	2.22	0.9	
	M2	2.63	1.94	3.33	1.4	
NS	Full	2.49	2.13	2.85	-	
	NS1	2.48	2.10	2.88	2.2	
	NEP	2.32	1.79	2.86	0.5	

<sup>a</sup>The mean rate of nucleotide substitution per site per year was estimated using BEAST.

<sup>b</sup>Genetic distance to A/Hong Kong/1/1968 was calculated from the phylogenetic tree and was plotted as a function of time. The mean rate of amino acid substitution per site per year was inferred from the linear slope (Figure 4).

to A/Hong Kong/1/1968 for PB2, PA, NP, and, NA was seen, reflecting the long branches also observed in the nucleotide ML trees. Such an increase was also observed for PA at the transition of the BK79-like to SI87-like viruses. For NP, some of the BE89-like, BE92-like, and WU95-like viruses showed a linear trend; however, there was also a large group of BE89-like, BE92-like, and WU95-like viruses that did not accumulate amino



acid mutations and shared the same amino acid distance to A/Hong Kong/1/1968. In the NP nucleotide ML tree, these groups reflect two clades, and the viruses that accumulate amino acid mutations are the ancestors of the SY97-like viruses. The M1, M2, NEP, and PA-X proteins did not accumulate amino acid mutations in a constant manner. For NS1, the rate of amino acid substitution decelerated after 1991. This was roughly around the same time that the NS1 gene acquired a stop codon at position 231, resulting in a C-terminally truncated NS1 protein. This premature stop codon, observed in 218 (76 %) NS1 sequences, first appeared at the end of 1971 and became fully fixed in late 1991 (see Table S1 in the supplemental material). In addition, two Dutch viruses circulating in 1970 harboured a stop codon at position 221 of NS1, but this stop codon did not become fixed. Premature stop codons were also observed in the PB1-F2 amino acid alignments. Twenty-four (8 %) of the PB1-F2 sequences contained an additional stop codon, at position 9, 12, 26, 35, 58, 64, 80, or 88. Of the C-terminally truncated proteins, 19 (7 %) were more than 78 codons long and still functional<sup>214,215</sup>. However, none of these stop codons became fixed over time. One PA-X sequence contained a stop codon at position 42.

### Selection pressures

The degree of natural selection acting on all ORFs was estimated by calculating the overall  $d_N/d_S$  and the mean  $d_N/d_S$  for the internal and external branches, as estimated using SLAC<sup>205</sup> (Table 3).

The membrane proteins; HA, NA, M2, and NS1 showed relatively high overall  $d_N/d_S$  values (0.341, 0.292, 0.391, and 0.352, respectively) compared to those of the other proteins which ranged from 0.058 to 0.098. As expected, the HA1 subunit revealed a relatively higher overall  $d_N/d_S$  than that of HA2 (0.466 for HA1 and 0.123 for HA2). The PB1-N40, PA-N155, and PA-N182 proteins had relatively low overall  $d_N/d_S$  values that were comparable to those of the full PB1 and PA ORFs.

Comparing the mean internal and mean external  $d_N/d_S$  values provides a ratio (internal  $d_N/d_S$  to external  $d_N/d_S$ ) that gives an indication of whether there is an excess of nonsynonymous mutations that are eventually removed from the virus population by purifying selection (ratio of < 1) or kept in the population by positive selection (ratio of > 1). The full and N-terminally truncated polymerases, HA, M2, and NEP, showed ratios of < 1, indicative of purifying selection. NS1 had comparable external and internal  $d_N/d_S$  values. The NP, NA, and M1 proteins showed ratios that were > 1, suggestive of positive selection.

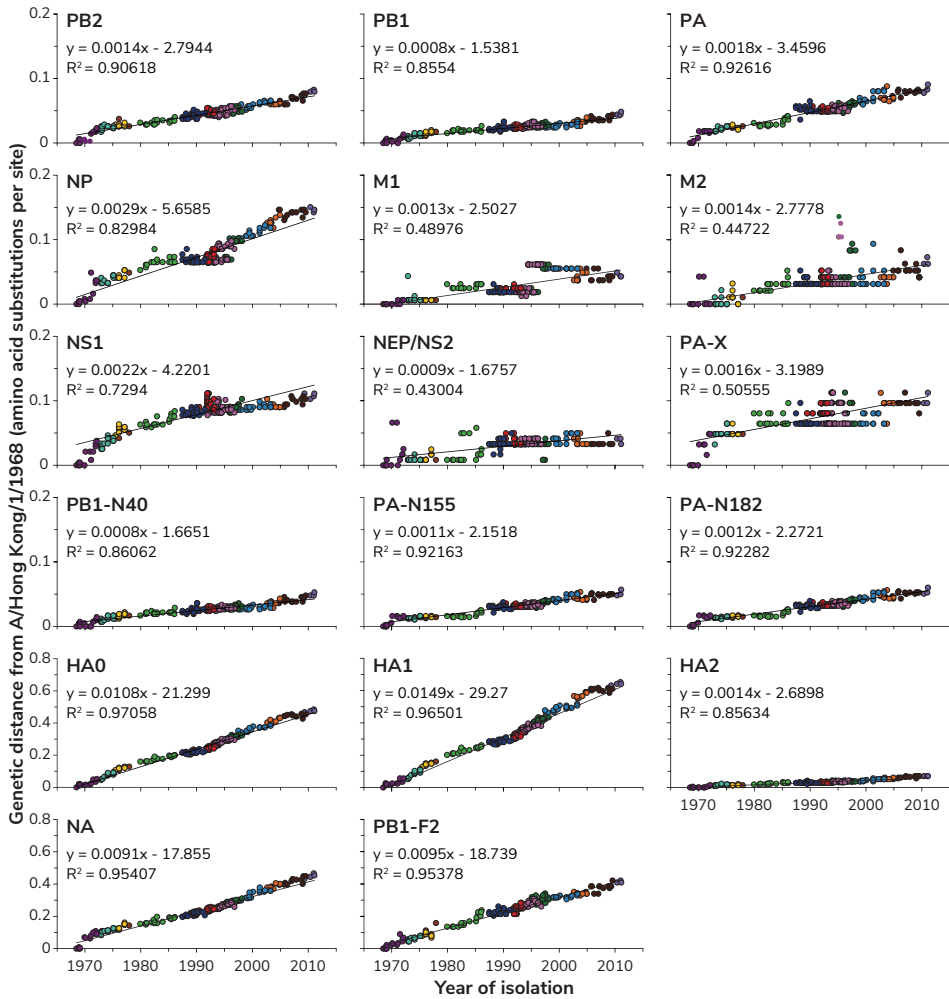


Figure 4 | Rates of amino acid evolution of all segments and ORFs of A(H3N2) viruses. Phylogenetic trees were generated with 286 amino acid sequences for PB2, PB1, PB1-N40, PB1-F2, PA, PA-N155, PA-N182, PA-X, HA0, HA1, HA2, NP, NA, M1, M2, NS1, and NEP. For all A(H3N2) viruses, the amino acid distance of each ORF to A/Hong Kong/1/68 was calculated from the phylogenetic tree and was plotted as a function of time. The colour coding of viruses is based on the antigenic clusters of HA and is consistent between all plots (see the legend to Figure 1). Note that the vertical axes differ between proteins with lower rates of amino acid substitution (four upper rows) and proteins with higher rates of amino acid substitution (two bottom rows).

**Detection of positive and negative selection by FUBAR**

Negatively and positively selected codons were estimated by the FUBAR method, which detects pervasive diversifying or purifying selection by Bayesian MCMC analyses<sup>164</sup> (Table 3; see Table S5 in the supplemental material). All ORFs contained an abundance of negatively selected codons. For PB1, PA, and NP, 63 to 69 % of the codons were negatively selected. This percentage was higher for PB2 (83 %) and slightly lower for HA, NA, and M1 (48 %, 50 %, and 58 %, respectively). M2, NS1, and NEP contained the lowest percentages of negatively selected codons (10 %, 21 %, and 28 %, respectively). Positive selection was found only for PB2, HA, and NA. All eight positively selected codons of HA were located within antigenic sites<sup>52,53</sup>. Two codons within NA were positively selected: codon 267, which is not surface exposed, and a codon in antigenic site C<sup>134</sup> (codon 370).

**Detection of positive and negative selection by MEME**

MEME was used to identify not only fixed but also more sporadic positively selected codons<sup>165</sup> (Table 3; see Table S5 in the supplemental material). MEME found positively selected codons in PB1, PA, HA, NP, NA, M1, and NS1. Four codons were selected in PA, among which one (codon 87) is located in the endonuclease activity region<sup>216</sup>. The other three (codons 278, 487, and 550) are located in the region involved in PB1 binding<sup>216</sup>. Of the 13 codons that were selected in HA, 12 were selected in HA1 and almost solely located within antigenic sites<sup>52,53</sup>. One codon was selected in NP (codon 54). In NA, six codons were selected, among which three are located in antigenic sites<sup>134</sup> (codons 199, 338, and 401). MEME identified codon 166 of M1, which is located in the C-terminal domain<sup>216</sup>, and codon 65 of NS1.

**Directional positive selection**

Directional positive selection was detected with the DEPS program, which helps in detecting selective sweeps<sup>167</sup> (Table 3; see Table S5 in the supplemental material). DEPS identified codons in PB2, PB1-F2, PA, HA, NA, and NS1. In the PB2 segment, codon 553, involved in cap binding<sup>217,218</sup> and located in the RNA binding domain<sup>216</sup>, and codon 679, interacting with PB1<sup>217,218</sup> and located in the importin binding domain<sup>216</sup>, were positively selected. Both codons interact with NP<sup>217,218</sup>. In PB1-F2, codons 23 and 90 were positively selected. Codon 311, associated with PB1 binding<sup>216</sup>, was positively selected in PA. There were 17 positively selected codons in HA, of which 15 were located in HA1. Twelve of the codons in HA1 are located in antigenic sites<sup>52,53</sup>. The

Table 3 | Selection pressures, positively, and negatively selected codons of all coding regions of A(H3N2) viruses circulating between 1968 and 2011.

Coding region	Codons	$d_N/d_S^a$ Overall	Internal	External	Internal/ external	Selected sites			
						FUBAR <sup>b</sup>	MEME <sup>b</sup>	DEPS <sup>c</sup>	FUBAR <sup>b</sup>
PB2	759	0.077	0.063	0.088	0.716	1	0	2	628 (83)
PB1	666	0.075	0.058	0.086	0.674	0	1	0	459 (69)
PB1-N40	635	0.074	0.059	0.084	0.702	0	1	0	431 (68)
PB1-F2 <sup>f</sup>	Nd	Nd	Nd	Nd	Nd	Nd	Nd	2	Nd
PA	654	0.085	0.070	0.095	0.737	0	4	1	435 (67)
PA-N155	500	0.094	0.084	0.100	0.840	0	3	1	327 (65)
PA-N182	473	0.097	0.090	0.102	0.882	0	3	1	306 (65)
PA-X <sup>f</sup>	Nd	Nd	Nd	Nd	Nd	Nd	Nd	0	Nd
HA0	566	0.341	0.301	0.370	0.814	10	13	17	274 (48)
HA1	329	0.466	0.406	0.515	0.788	10	12	15	149 (45)
HA2	221	0.123	0.093	0.142	0.655	0	0	1	120 (54)
NP	498	0.098	0.109	0.091	1.198	0	1	0	315 (63)
NA	469	0.292	0.311	0.277	1.123	2	6	11	233 (50)
M1	229	0.058	0.078	0.043	1.814	0	1	0	132 (58)
M2	73	0.391	0.264	0.481	0.549	0	0	0	7 (10)
NS1	157	0.352	0.350	0.353	0.992	0	1	3	33 (21)
NEP	40	0.074	0.055	0.087	0.632	0	0	0	11 (28)

<sup>a</sup>Calculated by SLAC.

<sup>b</sup>Posterior probability of  $\geq 0.9$ .

<sup>c</sup>P value of  $< 0.05$ .

<sup>d</sup>Bayes factor of  $> 20$ .

<sup>e</sup>Information on the specific positively selected residues and their functions can be found in Table S5 in the supplemental material.

<sup>f</sup>ND, coding regions of PB1-F2 and PA-X were excluded from the codon-based analyses (SLAC, FUBAR, and MEME).

positively selected codon in HA2 (codon 538) is located in the transmembrane anchor domain. Eleven positively selected codons were found in NA: three in the stalk region (codons 46, 52, and 56) and eight in the globular head, of which three are located within antigenic site C<sup>134</sup> (codons 328, 334, and 370). NS1 had three codons that were positively selected (codons 56, 84, and 129). In addition to FUBAR, MEME, and DEPS analyses, the data sets were analysed with IFEL (see Table S5 in the supplemental material). IFEL detects negatively and positively selected codons along internal branches<sup>164</sup>.

## DISCUSSION

In this study, the relationship of genetic evolution to antigenic change spanning the entire period of A(H3N2) virus circulation was studied for the first time. We analysed 286 full genomes scattered over 43 years of A(H3N2) virus evolution. The antigenic evolution based on haemagglutination inhibition data has been studied extensively for these viruses<sup>107,160</sup>. These antigenic properties were used as a basis for looking at genomic evolution and reassortment patterns of the whole genome between viruses belonging to the same antigenic cluster as well as between viruses belonging to different antigenic clusters.

The general topologies of the ML trees were similar for each segment, displaying the typical ladder-like gradual evolution previously described for HA1<sup>161</sup>. However, differences could be seen in reassortment patterns, clustering of viruses, evolutionary rates, and MRCA upon comparison of their phylogenies. An asynchronous and slightly slower evolution of NA than of HA was observed, which is consistent with previous studies<sup>198</sup>. In comparing HA with PB2, PB1, PA, NP, M, and NS, even larger differences were seen in the rate of nucleotide evolution.

Intrasubtypic reassortment of the eight segments of human A(H3N2) viruses has been studied for New York State viruses circulating between 1999 and 2004<sup>96</sup>, 1992 and 2005<sup>95</sup>, and 1997 and 2005<sup>97</sup> and for German isolates sampled between 1998 and 2005<sup>98</sup>. The studies showed that the evolution of the A(H3N2) virus was not so much determined by adaptive processes<sup>97</sup> but was shaped by frequent reassortment<sup>95-98</sup>, cocirculation<sup>96,97</sup> and persistence<sup>95,96</sup> of clades, virus migration<sup>97</sup>, and selective sweeps<sup>95</sup>. The data sets of these studies contain data mostly for recent A(H3N2) viruses, and studies that included earlier isolates are sparse. Reassortment of early A(H3N2) viruses and A(H2N2) viruses circulating between 1957 and 1972 was investigated, providing evidence of cocirculation of A(H2N2) and A(H3N2) viruses<sup>38</sup>, but no full genome studies



on A(H3N2) viruses circulating between 1972 and 1992 have been published. The data set used in the present study contains isolates, mostly of European origin, sampled from each consecutive season between 1968 and 2011 whose antigenic properties were determined previously<sup>107,160</sup>. Although North America and Europe are both in the Northern Hemisphere and therefore use the same vaccine composition, the viruses between these regions could still potentially evolve independently. Consequently, this data set may not be comparable to the American data sets, nor does it represent the A(H3N2) virus worldwide. However, this data set does offer the opportunity to study reassortment in relation to antigenic change spanning the entire period of A(H3N2) virus circulation.

A combined approach of tanglegrams colour coded by antigenic cluster and GiRaF analysis<sup>162</sup> was applied to map reassortment events. In addition, a methodology similar to that of Rambaut et al.<sup>95</sup> was used, but now investigating the time of circulation of the MRCA of each segment per antigenic cluster<sup>107,160</sup>. The tanglegrams showed abundant clustering differences suggestive of frequent reassortment over time. This was confirmed by GiRaF analyses, which revealed reassortment events particularly during the circulation of BE92-like and WU95-like viruses, in agreement with our previous study<sup>198</sup>. A lower incidence of reassortment than that in our previous study was detected. This analysis was more stringent because GiRaF was applied to all combinations of segments rather than to HA and NA alone<sup>162</sup>. Several reassortment events persisted, mostly in more recent years, suggesting that these reassortment events aided in increasing genetic diversity of the virus to improve fitness or to evade population immunity. However, most reassortment events were nonpersistent, indicating that reassortment is rarely beneficial enough to reach fixation.

Rates of nucleotide substitution estimated for A(H3N2) viruses were lower overall than rates previously reported by others<sup>95,97</sup>. However, the previous data sets spanned shorter time frames (1997 to 2005 and 1992 to 2005, respectively) and contained viruses sampled from North America, whereas our data set contained mostly European samples. Compared to a study of the polymerases of A(H3N2) viruses isolated between 1968 and 1997<sup>219</sup> or the HA proteins of A(H3N2) viruses isolated between 1968 and 1986<sup>220</sup>, rates of amino acid substitution reported here were substantially higher. Estimated rates of amino acid substitution for PB1 and PB1-F2 were comparable to those in a different study, in which the authors describe a substantially elevated rate of amino acid substitution for PB1-F2 compared to PB1<sup>221</sup>. However, in contrast to our results, they found only a slightly elevated rate of nucleotide substitution for PB1-F2 compared to PB1<sup>221</sup>. The authors explained the increase in rate of amino acid substitution in PB1-F2

via the frameshift relative to PB1<sup>221</sup>. A study that investigated different influenza virus strains [including human A(H3N2) viruses] showed that genetic and functional diversity of PB1-F2 is needed to support replication efficiency and virulence<sup>222</sup>. Similar to PB1-F2, PA-X is located in the frame +1 relative to PA, but no increase of amino acid substitutions was detected. The rate of nucleotide substitution of PA-X was the lowest of all segments/ORFs. It was described for human influenza viruses that there is an increase in nonsynonymous substitutions in the PA ORF spanning the +1 PA-X ORF. This is indicative of selective constraints in the PA-X protein and hence of functional importance<sup>223</sup>.

The PB1-F2 alignment revealed that 8 % of all sequences exhibited a stop codon, although none of these stop codons were fixed over time. In NS1, 76 % of all sequences acquired a stop codon at position 231, which is a lower percentage than that found in other studies (91 %) <sup>224</sup>. This stop codon started to appear at the end of 1971 and became fully fixed by late 1991, leaving the NS1 protein functional, since its N-terminal nuclear localization signal was retained<sup>215</sup> and can still interact with one of the essential components of the machinery for the 3'-end processing of cellular pre-mRNAs<sup>214</sup>.

M2, NS1, HA and NA had the highest  $d_N/d_S$  values (0.391, 0.352, 0.341, and 0.292, respectively). M2, HA, and NA are the surface glycoproteins and are thus accessible to antibodies, which could partially explain the relatively higher overall  $d_N/d_S$  values than those of the other proteins. Although it is surface exposed, M2 was shown to be almost nonimmunogenic<sup>225</sup>. Five amino acid residues in the M2 ion channel have been linked to adamantane resistance in A(H3N2) viruses<sup>102,226</sup>. Although not detected as being positively selected for, there was a serine-to-asparagine change in one of these amino acid residues (codon 31) that was starting to appear in 2005 and became fixed in the population. NS1 plays a role in evading the innate and adaptive immune responses<sup>227-229</sup>, which may explain its relatively high  $d_N/d_S$  value. The overall  $d_N/d_S$  results were in agreement with those of previous studies<sup>180,230</sup>.

HA contained the largest number of positively selected codons (29 codons), most prominently located in the HA1 subunit (26 codons), 22 of which are located in antigenic sites<sup>52,53</sup>. For NA, 19 sites were positively selected, 7 of which are located within antigenic sites<sup>134</sup>. Both HA and NA are under antibody-mediated selection<sup>53,123,124,231</sup>, and at least some of the positively selected sites were likely due to immune evasion. It is likewise probable that some of the mutations, while not directly needed for evasion of the immune system, restore possible loss of viral fitness due to mutations that cause antigenic change. Although it is not surface exposed, PA has a large number of positively selected sites.

While little is known for most mutations in the polymerases, these mutations in PA may possibly be responsible for fine-tuning replication. In contrast to the predominant humoral immune response-related drift of HA and NA, drift in internal proteins such as NP is linked to escape from cytotoxic T-lymphocyte immunity<sup>104,232,233</sup>. Codon 146 of NP was found to be positively selected and has changed the T-cell epitope NP146–154 (TTYQRTRAL)<sup>234</sup> with the T146A substitution, which was fully fixed in the population in 2001.

There are several hypotheses to explain the mechanism that governs antigenic cluster transitions. One theory involves neutral networks in which phenotypically neutral mutations occur in the HA. Although these mutations do not have a direct impact on the antigenic phenotype, they might allow for subsequent mutations that do influence the antigenic properties of the virus<sup>235</sup>. Other studies showed that a new HA lineage was acquired at the time of the SY97 to FU02 antigenic change<sup>95,96</sup>. Also, according to our GiRaF analysis, this cluster transition was linked to a persistent reassortment event (event 55) (Figure 2 and Table 1). Based on our data set, there was no support for other persistent reassortment events linked to any of the other antigenic cluster transitions. Just recently, the molecular basis of antigenic drift was determined for human A(H3N2) viruses. From 1968 to 2003, antigenic change was caused mainly by single amino acid substitutions at seven positions around the receptor binding site<sup>236</sup>. No clear correlation between antigenic cluster transition and positively selected codons in proteins other than the HA was found. Some positively selected codons in PB2 (codon 697), PB1-F2 (codon 90), PA (codons 277, 311, and 437), NP (codon 146), and NA (codons 43, 46, 52, 56, 127, 199, 253, 267, 308, 328, 334, 338, 339, 370, 372, 401, and 437) were fixed (for a certain time) in the population. Codon 146 in NP is located in a T-cell epitope<sup>234</sup> and codon 199, 328, 334, 338, 339, 370, and 401 in NA are all located within antigenic sites<sup>134</sup>. Future functional studies that investigate whether these mutations result in antigenic change due to immune pressure or change functional compatibility with HA, such as the HA-NA balance<sup>171</sup>, will be interesting.

In summary, we have analysed the evolutionary genetics of the full genomes of A(H3N2) viruses isolated between 1968 and 2011 and compared it to the antigenic evolution of HA. Reassortment of all eight segments occurred throughout almost the entire period of influenza A(H3N2) virus circulation, although persistent reassortment events were mostly found in recent years. Between 1987 and 1998 there was an increase in reassortment events, an increase or decrease in accumulation of amino acid substitutions of certain genes, and the acquisition of a stop codon in NS1. In addition, different antigenic lineages cocirculated during this time. These findings imply that the

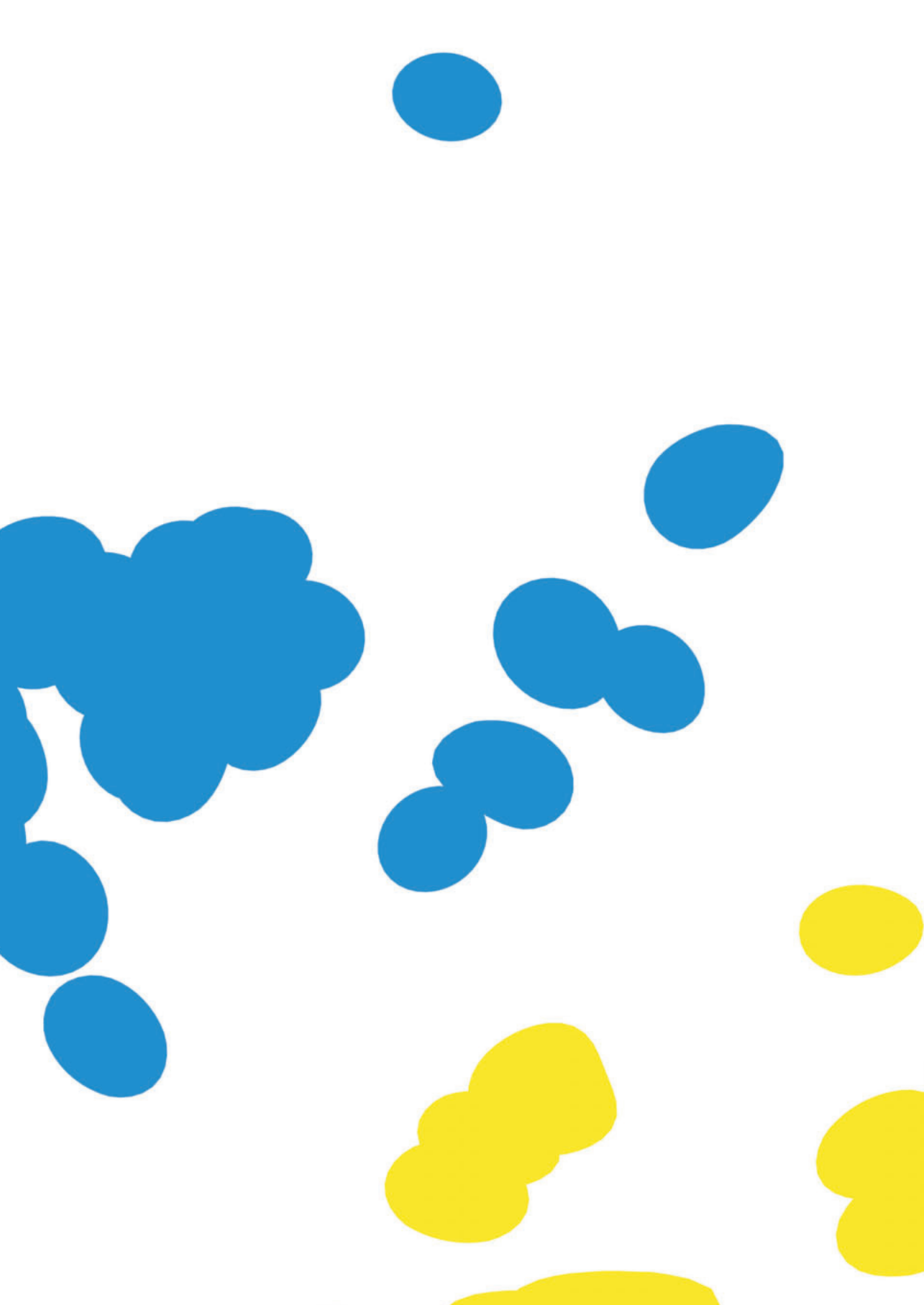
viruses used multiple mechanisms to increase virus diversity, presumably to improve fitness and/or to evade the host immune responses. This collection of viruses, together with the genetic and antigenic data, can now be used to carefully study genotypic and phenotypic relationships throughout the period of circulation of A(H3N2) viruses. To fully understand the dynamics of A(H3N2) virus evolution, future research of such extensive data sets should also focus on the functional and antigenic properties of proteins other than HA, using serological assays and antigenic cartography methods<sup>159</sup>.

## **ACKNOWLEDGMENTS**

This work was supported by an NWO-VICI grant and by NIH contract HHSN266200700010C. This project was also funded in part by federal funds from the National Institute of Allergy and Infectious Diseases, National Institutes of Health, Department of Health and Human Services, under contract HHSN272200900007C.

We gratefully thank G. J. D. Smith for stimulating discussions and technical assistance, N. Nagarajan and C. Kingsford for technical assistance with GiRaF, S. van der Vliet and O. Vuong for technical assistance in the lab, and B. Kalverda for critically reading the manuscript. We also thank the following members of the JCVI viral sequencing and assembly team for their technical assistance: T. Stockwell, E. Wester, K. Geer, D. Busam, and N. Fedorova.



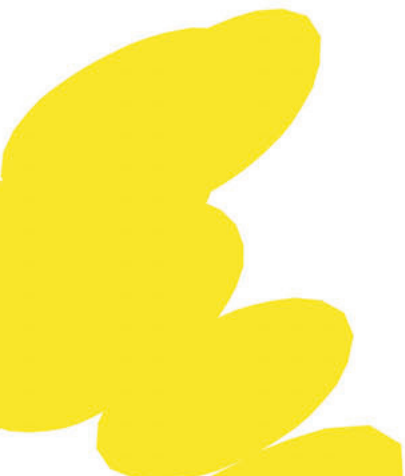


## **CHAPTER 4**

# **Discordant antigenic drift of neuraminidase and haemagglutinin in H1N1 and H3N2 influenza viruses**

Matthew R. Sandbulte, Kim B. Westgeest, Jin Gao, Xiyan Xu, Alexander I. Klimov,  
Colin A. Russell, David F. Burke, Derek J. Smith, Ron A. M. Fouchier, and  
Maryna C. Eichelberger

*Proceedings of the National Academy of Sciences U S A*, volume 108(51), 20 December 2011, pages 20748-20753



## **ABSTRACT**

Seasonal epidemics caused by influenza virus are driven by antigenic changes (drift) in viral surface glycoproteins that allow evasion from pre-existing humoral immunity. Antigenic drift is a feature of not only the haemagglutinin (HA), but also of neuraminidase (NA). We have evaluated the antigenic evolution of each protein in H1N1 and H3N2 viruses used in vaccine formulations during the last 15 y by analysis of HA and NA inhibition titres and antigenic cartography. As previously shown for HA, genetic changes in NA did not always lead to an antigenic change. The noncontinuous pattern of NA drift did not correspond closely with HA drift in either subtype. Although NA drift was demonstrated using ferret sera, we show that these changes also impact recognition by NA-inhibiting antibodies in human sera. Remarkably, a single point mutation in the NA of A/Brisbane/59/2007 was primarily responsible for the lack of inhibition by polyclonal antibodies specific for earlier strains. These data underscore the importance of NA inhibition testing to define antigenic drift when there are sequence changes in NA.



## INTRODUCTION

Susceptibility to infection with circulating influenza viruses is determined to a large degree by the presence or absence of strain-specific functional antibodies elicited by prior infection or vaccination. Influenza viruses constantly evade antibody-mediated inhibition of replication by antigenic drift, an accumulation of mutations in epitopes of major surface proteins, HA and neuraminidase (NA)<sup>237</sup>. Antigenic drift has been studied most extensively for HA, although NA has also been observed to undergo antigenic drift<sup>122–124</sup>. NA-specific antibodies can reduce viral replication and disease severity in mice<sup>125</sup> and chickens<sup>126</sup>, and have similarly been associated with resistance against influenza in humans<sup>127,128</sup>. Despite this correlation with immunity, antigenic drift of NA is not routinely examined. Early studies with a few virus strains demonstrated discordant antigenic drift of HA and NA<sup>122</sup>, suggesting the virus can overcome host antibody resistance by modifying either antigen. Although it is likely that NA's drift is most often the result of antibody selection, antigenic change may on occasion be a consequence of a functional change in HA<sup>238</sup>.

Vigilant surveillance by public health agencies is required to maximize the match between seasonal vaccine antigens and predominant circulating viruses<sup>151,155</sup>. At present, vaccine strain selection decisions are based on antigenic characterization of HA coupled with HA and NA genetic data, also taking into consideration epidemiologic and human serologic data<sup>155</sup>. Antigenic cartography using data from HA inhibition (HI) assays provides a tool to visualize and quantitate antigenic relatedness of the HAs of circulating viruses<sup>107</sup> in relation to vaccine viruses. Antigenic characterization of NA can be performed using NA inhibition (NI) assays to determine the extent of antibody-mediated interference with enzyme activity<sup>138</sup>, but the cumbersome nature of the standard NI assay using large volumes of hazardous chemicals has precluded routine analysis of NA. We recently developed a miniaturized format of this assay and confirmed its accuracy and sensitivity for analysis of NI antibody titres in human and animal sera<sup>239</sup>. In the present study, we use this assay to characterize the antigenic drift of NA in human H1N1 and H3N2 viruses recommended for United States influenza vaccines over the past ~15 y (Table S1 in the supporting information (SI), available in PNAS Online). In the NI assay, we used panels of ferret antisera against each wild-type H1N1 and H3N2 virus and virus reassortants generated by reverse genetics to combine the targeted NA and a mismatched HA of the H6 subtype. Use of these reassortants prevented false NI signals because of interfering HA antibodies. We then constructed antigenic maps from the data

sets, using multidimensional scaling to position the antigens and antisera on the map, as previously described<sup>107</sup>.

## **MATERIALS AND METHODS**

### **Viruses**

The sources of viruses and plasmids are provided in SI Materials and Methods. NA genes were cloned from H1N1 and H3N2 vaccine seed strains listed in Table S1 (SI) and reassortant viruses generated by eight-plasmid reverse genetics for the NI assays, as described previously<sup>239</sup>. Single amino acid mutations were introduced to the plasmid-encoded BR/07 NA gene using the QuikChange II site-directed mutagenesis kit (Agilent Technologies). Virus stocks were amplified in 10-d embryonated chicken eggs for use in NI and HI assays. Viruses grown in Madin-Darby canine kidney cells obtained from the National Influenza Centre, Erasmus MC, Rotterdam, The Netherlands, were also used in HI testing.

### **Genetic analyses**

Phylogenetic trees and genetic maps based on the coding region of NA or HA1 were constructed to analyse the genetic evolution of human H1N1 and H3N2 viruses, as described in SI Materials and Methods.

### **Sera and serological assays**

The sources of ferret and human antisera used in NI and HI analyses are provided in SI Materials and Methods. Titration of serum NI antibodies was performed by analysing NA activity of the HA mismatched reassortant viruses in a 96-well plate format of the conventional thiobarbituric acid assay<sup>239</sup>, as described in SI Materials and Methods. A conventional assay was used to determine HI titres<sup>252</sup>, as described in SI Materials and Methods.

### **Antigenic cartography**

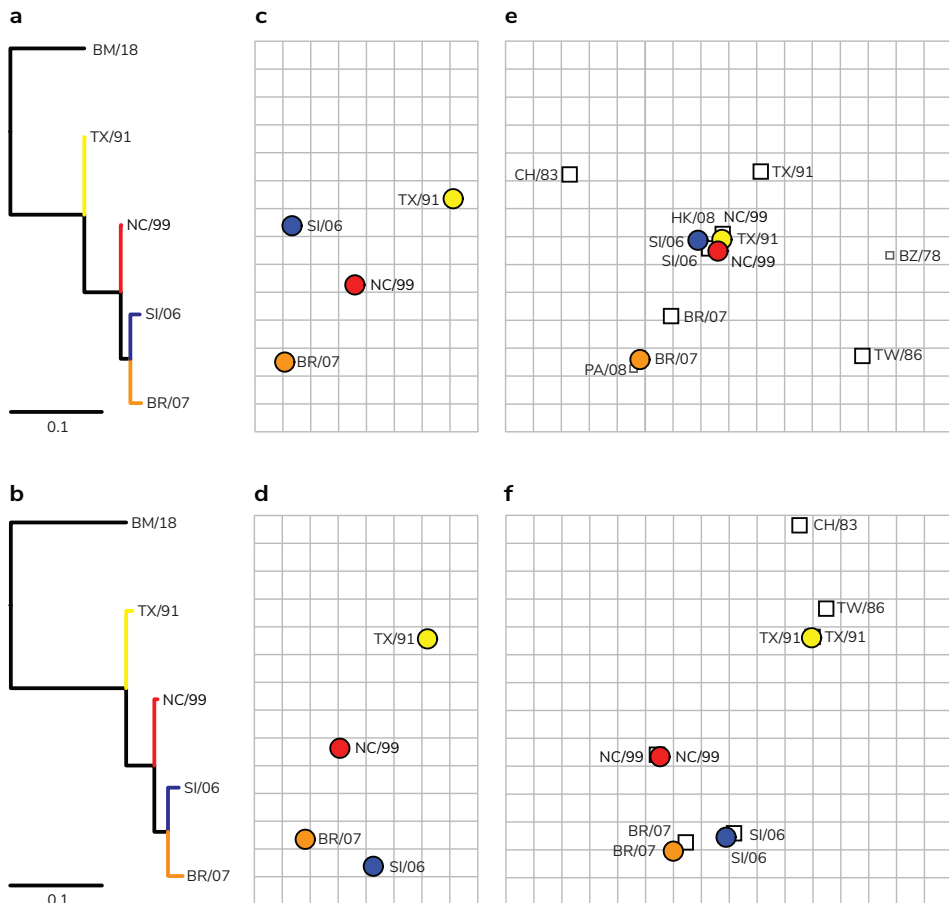
The antigenic properties of human influenza A (H1N1) and human influenza A (H3N2) viruses were characterized using NI and HI data and antigenic cartography methods as described previously for HI data of human influenza A (H3N2) viruses<sup>107</sup> and in SI Materials and Methods.

## RESULTS AND DISCUSSION

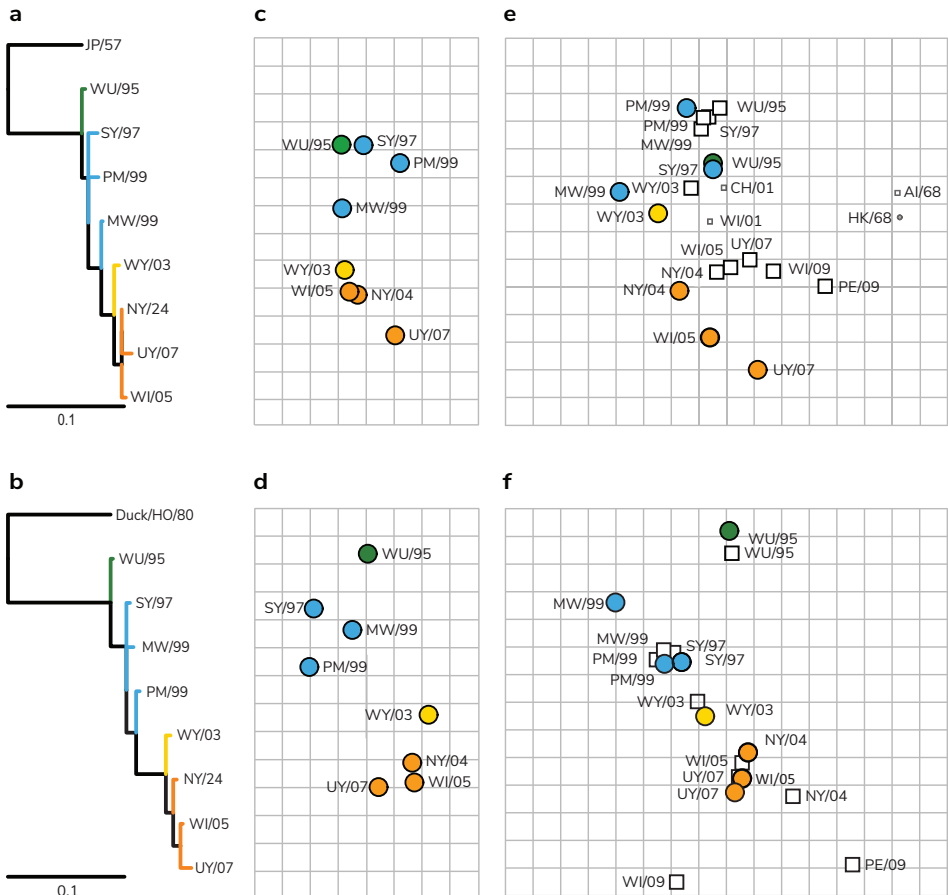
### Antigenic characterization of the HA and NA of H1N1 and H3N2 viruses.

H6 reassortant viruses containing NA of historical as well as recent H1N1 and H3N2 vaccine viruses (see Table S1 in the SI) were used to measure NI titres of strain-specific ferret antisera. There was minimal NI cross-reactivity between the phylogenetically distant early NAs and antisera raised against recent seasonal strains of the classical H1N1 (see Table S2 in the SI) and H3N2 (see Table S3 in the SI) human lineages, demonstrating extensive antigenic drift since introduction of these subtypes. Ferret serum raised against A/California/7/2009 (CA/09), a representative 2009 H1N1 pandemic (H1N1pdm) virus, demonstrated a robust homologous NI antibody titre with very low cross-reactive titres against NAs of the long-established human seasonal H1N1 lineage (see Table S2 in the SI).

Phylogenetic trees, genetic maps based on amino acid sequences, and antigenic maps were generated for HA and NA of H1N1 (Figure 1) and H3N2 (Figure 2) viruses recommended for inclusion in 1996–2009 seasonal influenza vaccines. Since NI data had not previously been used to generate an NA antigenic map, we first established that end-point NI titres provided sufficient sensitivity for the analysis. This process was done by comparative analysis of end-point titres (inverse of the serum dilution that inhibits NA activity  $\geq 50\%$ ) and the precise 50% inhibition titre ( $IC_{50}$ ), the inverse of the serum dilution that results in exactly 50% inhibition determined by nonlinear regression analysis<sup>239</sup> of ferret antisera generated in response to H3N2 infection. The maps generated using these data sets were very similar, with an excellent correlation of distances between antigens on the map,  $R^2 = 0.99$  (see Figure S1 in the SI). End-point NI titres were therefore used in subsequent analyses. We also showed that there was good correlation between antigenic distance determined using NI data in tables and the antigenic distance determined from map location (see Figure S2 in the SI) ( $R^2 = 0.95$  for N1 and  $R^2 = 0.86$  for N2 antigens), providing confidence that the antigenic maps are representative of the raw data. To confirm the location of antigens on the map, NI assays were repeated using a larger number of antisera. There was excellent correlation between the location of antigens generated by analyses of the first and second data sets (see Figure S3 in the SI) ( $R^2 = 0.98$  for both N1 and N2 antigens).



**Figure 1 | Comparison of the antigenic and genetic evolution of NA and HA of influenza A (H1N1) virus. a and b | Phylogenetic trees of the coding region of NA (a) and HA1 (b) nucleotide sequences. A/Brevig Mission/1918 was chosen as outgroup for both trees. Scale bars roughly represent 10 % of nucleotide substitutions between close relatives. c and d | NA (c) and HA (d) genetic maps. The vertical and horizontal axes represent genetic distance, in this case the number of amino acid substitutions between strains; the spacing between grid lines is 3 amino acid substitutions. e and f | NA (e) and HA (f) antigenic maps based on NI and HI data, in which the viruses are shown as circles and antisera as squares. The spacing between grid lines is one unit of antigenic distance, corresponding to a 2-fold dilution of antisera in the NI and HI assays. The orientations of all maps were chosen to roughly match the orientation of the HI antigenic map in f, and the colour coding of viruses is consistent among all panels.**



**Figure 2 | Comparison of the antigenic and genetic evolution of NA and HA of influenza A (H3N2) virus. a and b | Phylogenetic trees of the coding region of NA (a) and HA1 (b) nucleotide sequences. A/Japan/305/1957 was chosen as outgroup for the NA tree and A/Duck/Hokkaido/33/80 for the HA1 tree. Scale bars represent approximately 10 % of nucleotide substitutions between close relatives. c and d | Genetic maps of the amino acid sequences of NA (c) and HA1 (d). The vertical and horizontal axes represent genetic distance, in this case the number of amino acid substitutions between strains; the spacing between grid lines is 4 amino acid substitutions. e and f | NA (e) and HA (f) antigenic maps of influenza A (H3N2) virus based on NI and HI data, in which the viruses are shown as circles and antisera as squares. The spacing between grid lines is one unit of antigenic distance, corresponding to a 2-fold dilution of antisera in the NI and HI assays. The orientations of all maps were chosen to roughly match the orientation of the H3 antigenic map in f, and the colour coding of viruses is consistent among all panels.**

The raw NI data as well as the antigenic map show that the NA of human H1N1 viruses was in antigenic stasis for over a decade, as antisera to vaccine strains A/Texas/36/1991 (TX/91), A/New Caledonia/20/1999 (NC/99), and A/Solomon Islands/3/2006 (SI/06) each had similar NI titres against all three corresponding NAs, even though there had been substantial genetic diversity between them. In contrast, an abrupt antigenic change was observed for the NA of A/Brisbane/59/2007 (BR/07), the H1N1 virus recommended for use in Northern Hemisphere trivalent vaccines during the 2008/09 and 2009/10 seasons (Figures 1a, 1c, and 1e; see Table S2 in the SI). Antisera to the older strains had at least 16-fold lower NI titres against BR/07 than against the homologous antigen. Antigenic differences in the HA for these strains were quite the opposite: HAs of TX/91, NC/99 and SI/06 were distinct from one another, but the antigenic difference between HAs of SI/06 and BR/07 was less pronounced (Figure 1f; see Table S2 in the SI). The antigenic distance between the NA of BR/07 and the previous vaccine strain, SI/06, is 4.6 units (Figure 1e), which is greater than the antigenic distance between the HAs of these viruses (2.0 units). From this observation, we speculate that the abrupt antigenic change of N1 NA may have been a factor contributing to the dominance of H1N1 in many parts of the Northern Hemisphere during the 2007/08 season<sup>240</sup>. The antigenic drift of NA was also not proportional to number of amino acid changes for H3N2 viruses isolated from the mid-1990s [A/Wuhan/359/95 (WU/95)] to the 2009/10 Northern hemisphere formulation [A/Uruguay/716/2007 (UY/07), an A/ Brisbane/10/2007-like virus], even though phylogenetic analysis of the NAs of these viruses shows a chronologic progressive evolution at the nucleotide level (Figure 2a). The antigenic analyses (see Table S3 in the SI) and maps (Figure 2e and 2f) of H3N2 vaccine viruses show that the antigenic drift of HA and NA of this subtype is also discordant. A major step in HA antigenic drift between WU/95 and A/Sydney/5/1997 (SY/97) was well-recognized previously<sup>151,241</sup>. However, there is almost no genetic or antigenic difference between the NAs of WU/95 and SY/97. Despite this antigenic similarity in the NAs, effectiveness of the 1997/98 inactivated vaccine was poor for healthy adults<sup>242</sup>, suggesting that NA in the vaccine did not induce levels of antibodies with capacity to protect against laboratory-confirmed influenza-like illness.

The recommended H3N2 component for seasonal vaccine formulations from 2000/01–2003/04 was A/Moscow/10/1999 (MW/99). Because of difficulty in generating a high-growth MW/99 reassortant and based on HI data, A/Panama/2007/1999 (PM/99) was widely used for vaccine production as a MW/99-like virus. At the time, the difference in MW/99 and PM/99 NA sequences was of some concern<sup>243</sup>, but it was not known

whether this corresponded with a difference in antigenic properties. Our results show the NAs of MW/99 and PM/99 are antigenically distinct, with the NA of MW/99 positioned closer to the subsequent A/Fujian/411/2002-like vaccine strain, A/Wyoming/3/2004 (WY/03), than to PM/99 (Figure 2e).

There tends to be an asymmetrical aspect to the reactivity of antisera with NA; antisera raised against old strains reacted weakly with more recent viruses, whereas reactivity of sera prepared against more recent viruses was often retained in NI tests against older strains. For example, antisera to BR/07 had good inhibition of TX/91 NA, and antisera to WY/03 had excellent inhibition of WU/95 NA (see Tables S2 and S3 in the SI). An exception to this trend was the antigenic profile of PM/99's NA. This virus had weak reactivity with all sera generated against subsequent viruses.

#### **Sera from vaccinated individuals discriminate between viruses with drifted NAs.**

Whether NA antigenic properties determined from reactivity with ferret antisera are highly relevant predictors of human antibody titres is an important question. Ferret sera differ from human sera in the complexity of past history with influenza virus antigens, particularly because ferrets are screened to ensure seronegative status before inoculation with reference viruses, making their influenza immunity monospecific. In a previous study, we analysed human sera collected before and after immunization with 2006–07 seasonal influenza vaccines for NI titres against the homologous NC/99 N1 and A/Wisconsin/67/2005 (WI/05) N2 antigens<sup>239</sup>. From the small cohort in that study, we identified subsets of vaccinees with increases in serum NI titre against N1 or N2 that were robust enough to be consistently measured across repeated assays. Serum pairs that showed distinct increases in NI titre against N1 or N2 [in each case, three volunteers vaccinated with live attenuated influenza vaccine (LAIV) and three with trivalent inactivated vaccine (TIV)] were reassayed against homologous NA antigens plus NA antigens of the H3N2 or H1N1 vaccine viruses of subsequent vaccine campaigns.

There were two new H1N1 vaccine strain recommendations after the 2006–07 clinical trial: SI/06 and BR/07. Pre-vaccine NI titres against NC/99 and SI/06 NAs were similar in magnitude, and for most serum pairs, NI titres against NA of the vaccine strain, NC/99, and SI/06 increased to similar levels postvaccination (Figure 3a). Both observations are consistent with the data from NI tests with use of monospecific ferret reference sera (Figure 1c), which showed high antigenic similarity between these two NAs. In contrast, most volunteers' pre-vaccine sera showed significantly less reactivity against the more

recent BR/07 NA. After vaccination, NI titres against the BR/07 NA remained significantly lower than titres against the two earlier strains. The mean fold-increase in titre against BR/07 NA (1.2) was less than that against SI/06 NA (1.75) and homologous NC/99 NA (2.2), although differences were not statistically significant ( $P = 0.16$ ). These data from human sera mirror the trend observed in assays with monospecific ferret sera, where BR/07 NA marked a substantial antigenic drift in the human H1N1 influenza lineage after great similarity among NAs of TX/91, NC/99, and SI/06.

The 2006 recommended H3N2 vaccine strain, WI/05, was updated in 2008 to include UY/07. All of the prevaccine serum samples possessed detectable NI activity against the NA of WI/05 and UY/07 (Figure 3b). After vaccination with LAIV or TIV, there were significantly greater increases in NI titre against WI/05 NA than against UY/07 NA (Figure 3d) ( $P < 0.05$ ). Thus, the modest antigenic distinction between two N2 NAs that was indicated by assaying ferret reference antisera was confirmed with sera of human vaccinees, showing that it is reasonable to correlate ferret serologic responses with human responses.

Single Amino Acid Change Is Largely Responsible for the Antigenic Drift of the SI/06 NA. Studies of NA's antigenic structure with monoclonal antibodies have been used together with molecular analysis of escape variants to identify epitopes in N2 NA<sup>182,183</sup>, providing clues as to which sequence changes in circulating viruses signify antigenic variation. After identifying globular head-region amino acids that differ between N1 NA proteins of BR/07 and both NC/99 and SI/06, we used site-specific mutagenesis to generate constructs containing the NA of BR/07 in which specific amino acids from the corresponding sequence of SI/06 and NC/99 were introduced (Figure 4a). All of the amino acid changes are surface residues at the opposite side to the tetramer interface (Figure 5) and are thus potentially accessible to antibodies. Amino acid changes at positions 222, 329, and 344 (using amino acid numbering for the N1 sequence) correspond to antibody-binding determinants previously identified for N2 proteins<sup>182,244</sup>. Three of the variant residues (222, 249, and 344) are located on the periphery of the enzyme active site, and another (329) is located on a loop more distant from the active site. None of these changes is predicted to add or remove a glycosylation site. The reactivity of ferret anti-BR/07, anti-SI/06, and anti-NC/99 sera was tested against the mutated antigens (Figure 4b–d). Antisera generated against NC/99 and SI/06 could inhibit the BR/07 NA when amino acid 329 was changed from Glu to Lys. Some reactivity with antiserum to NC/99 was restored when amino acid 222 (Gln) was changed to Arg, 249 (Lys) was changed to Gly, or 287 (Ile) was changed to Thr. Of note, none of the mutants lost reactivity



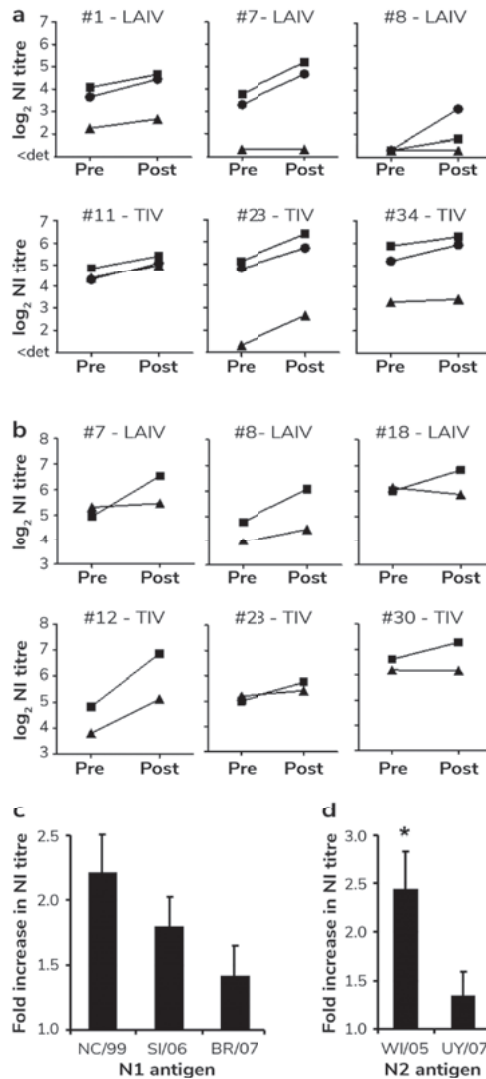


Figure 3 | Human pre- and postvaccination NI titres to homologous NA antigens and NAs of influenza virus strains in subsequent vaccine formulations. Volunteers received LAIV or TIV vaccines, as noted in graphs. Both vaccine types contained 2006/07 seasonal vaccine strains, including NC/99 (H1N1) and WI/05 (H3N2). Serum specimens were collected at day 0 (pre) and day 28 postvaccination. Pre and postvaccination NI titres of three volunteers in each vaccine group determined (a) against H1N1-derived NC/99 NA (●), SI/06 NA (■), and BR/07 NA (▲) and (b) against H3N2-derived WI/05 NA (●) and UY/07 NA (▲). The mean fold-increase in NI titre after vaccination against (c) the three N1 antigens, and (d) the two NA antigens. Statistical significance, \*P < 0.05.

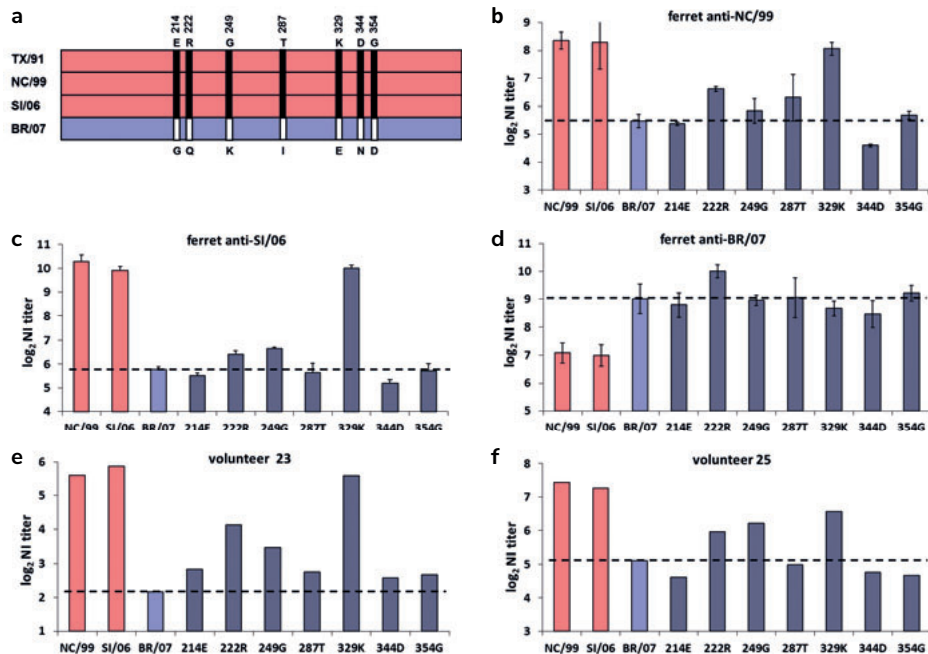


Figure 4 | One amino acid in BR/07 accounts for most of the NA drift variation from preceding H1N1 strains. **a** | Alignment of amino acid differences in the NA globular head between the consensus of three antigenically similar NA proteins (TX/91, NC/99, and SI/06) and the antigenically divergent BR/07 NA. **b - d** | Reactivity of H6 reassortant viruses with NA of BR/07 containing point mutations with ferret antisera raised against NC/99 (**b**), SI/06 (**c**), and BR/07 (**d**). Each data point represents the mean result of two independent assays, and error bars represent SD. **e** and **f** | NI titres of human sera against viruses containing NA of NC/99, SI/06, BR/07, and BR/07 mutants.

with antiserum against BR/07, reflecting the retention of reactivity with previous viruses noted earlier in this report and supporting the idea that other subdominant conserved epitopes are present. Most importantly, sera from human volunteers previously noted as responding to the vaccine antigen NC/99 regained full or partial reactivity with BR/07 that included the single E329K mutation (Figure 4e and 4f). Although the precise structural determinants required for antibody binding are unclear, the impact of the E329K substitution suggests that a charge interaction is important for this NA–antibody interaction. Our results demonstrate the impact of a single point mutation on NA drift and its consequence on recognition by human sera.

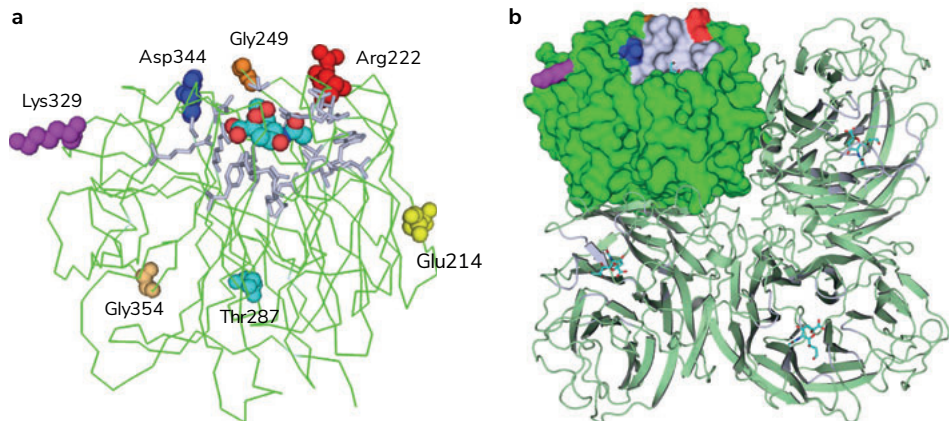


Figure 5 | Sites of amino acid differences between the NA globular head of Solomon Islands/3/2006 and A/Brisbane/59/2007. Sites of amino acid differences between the NA globular head of SI/06 and BR/07 are shown on wire and filled space models of the monomeric (a) and tetrameric (b) NA, constructed using Modeler<sup>253</sup> on subtype N1 (pdb code 2HTY). Sialic acid, docked into the enzyme active site (grey), is represented as sticks with coloured carbon (cyan), oxygen (red), and nitrogen (blue) atoms. Amino acids of SI/06 are highlighted: Glu214 (yellow), Arg222 (red), Gly249 (orange), Thr287 (cyan), Lys329 (magenta), Asp344 (blue), and Gly354 (pale orange).

In summary, our data provide confirmatory evidence of antigenic drift of NA that is often discordant with the antigenic drift of HA across contemporary H1N1 and H3N2 vaccine viruses. The finding that a single point mutation in NA is predominantly responsible for the antigenic drift between SI/06 and BR/07 highlights the importance of evaluating antigenic properties in addition to sequence changes when characterizing the NAs of emerging influenza viruses. In addition, our results show that for NC/99 (H1N1) and WI/05 (H3N2) components of the 2006/07 vaccine, human NA-specific antibodies reacted in a similar manner as monospecific ferret antisera, suggesting the antigenic drift identified by analysis of ferret antisera is relevant to human serologic responses. Although HA is the most abundant and important antigen for inducing neutralizing antibody responses, the contribution of NA-specific antibodies to efficacy of vaccines under development that contain significant amounts of immunogenic NA, such as virus-like particles, recombinant proteins, and high-dose vaccines<sup>129,245,246</sup>, should not be overlooked. Previous studies have pointed out the importance of NA-based immunity in curbing the severity or spread of emerging strains with novel HAs. Epidemiologic analyses<sup>247,248</sup> support the concept that NA antibodies contributed to protection against

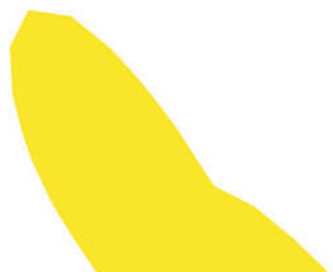
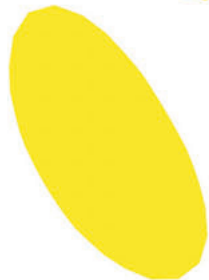
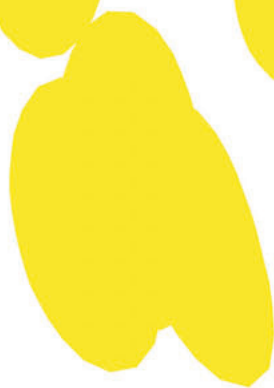
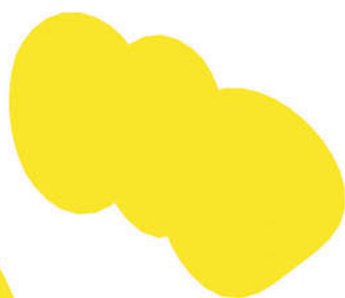
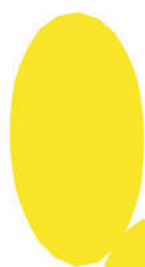
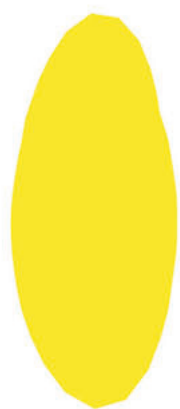
the 1968 pandemic H3N2 virus that possessed an NA antigenically similar to that of the previously circulating H2N2 lineage<sup>137,249</sup>.

In addition, antibodies directed to conserved regions of NA may contribute a significant level of protection against disease: the presence of such heterologous immunity has been demonstrated in H3N2<sup>250</sup> and H5N1<sup>251</sup> challenge studies in mice. Studies to identify the NI titre that correlates with protection in man, the vaccine dose needed to elicit a protective response, and the breadth of this protection in man, are certainly warranted to facilitate the development of NA-containing influenza vaccines.

## **ACKNOWLEDGMENTS**

We thank Zhiping Ye and Vladimir Lugovtsev for viruses; Robert Webster for plasmids; Olga Zoueva and Arash Hassantoufighi for technical support in preparing reassortant viruses; Theo Bestebroer and Stefan van der Vliet for excellent technical assistance with haemagglutination inhibition assays; and Timothy Straight for provision of human sera. This study used the CamGrid distributed computing resource and was supported by Centre for Biologics Evaluation and Research Pandemic Influenza funds (to M.C.E.); a VICI grant of the Netherlands Organization for Scientific Research (to R.A.M.F.); a National Institutes of Health Director's Pioneer Award (DP1-OD000490-01 to D.J.S.), European Union FP7 programs EMPERIE (223498 to D.J.S.) and ANTIGONE (278976 to D.J.S.), a Human Frontier Science Program grant (P0050/2008 to D.J.S.), and a University Research Fellowship from the Royal Society (to C.A.R.).





## CHAPTER 5

# **An optimized enzyme-linked lectin assay to measure influenza A virus neuraminidase inhibition antibody titres in human sera**

Laura Couzens, Jin Gao, Kim Westgeest, Matthew Sandbulte, Vladimir Lugovtsev,  
Ron Fouchier, Maryna Eichelberger

*Journal of Virological Methods*, volume 210, 15 December 2014, pages 7-14



## **ABSTRACT**

Antibodies to neuraminidase (NA), the second most abundant surface protein on influenza virus, contribute toward protection against influenza. The traditional thiobarbituric acid (TBA) method to quantify NA inhibiting antibodies is cumbersome and not suitable for routine serology. An enzyme-linked lectin assay (ELLA) described by Lambre et al.<sup>142</sup> is a practical alternative method for measuring NA inhibition (NI) titres. This report describes optimization of the ELLA for measuring NI titres in human sera against influenza A viruses, using H6N1 and H6N2 viruses as antigens. The optimized ELLA is subtype-specific and reproducible. While the titres measured by ELLA are somewhat greater than those measured by a miniaturized TBA method, seroconversion rates are the same, suggesting similarity in assay sensitivity under these optimized conditions. The ELLA described in this report provides a practical format for routine evaluation of human antibody responses to NA.



## INTRODUCTION

Neuraminidase (NA) inhibiting antibodies are associated with protection against influenza<sup>128</sup> and correlate with reduced viral shedding and disease symptoms in a human challenge study<sup>127</sup>. These antibodies contribute to immunity by inhibiting release and spread of newly formed virus particles from infected cells<sup>133</sup>. Despite the established importance of NA inhibiting (NI) antibodies, these titres are rarely measured in seroepidemiologic or vaccine studies because the traditional thiobarbituric acid (TBA) assay used to quantify these antibodies is impractical for large numbers of samples and employs hazardous chemicals. Alternative assays that have been developed include a miniaturization of the TBA method<sup>239</sup> and an enzyme-linked lectin assay (ELLA)<sup>142</sup>. The read-out of each method is different – while both quantify products of enzyme activity, the TBA method measures the amount of free sialic acid, the soluble product of NA activity, whereas ELLA measures the amount of penultimate galactose that becomes available after the terminal sialic acid is cleaved from substrate. Despite this difference, the assay principals are the same, usually employing whole virus as a source of antigen and using fetuin, a highly glycosylated protein, as substrate.

Since HA-specific antibodies in human sera can block access of substrate to NA, it is essential to use viruses with a HA subtype that is not in circulation when performing either of these assays for human serology. The assay described in this manuscript uses reverse genetics-derived H6 reassortant viruses that contain the targeted NA<sup>159</sup>. This strategy follows the original approach to measure NA inhibition antibody titres in which H6 reassortant viruses generated by classical reassortment were used in the traditional TBA method<sup>122,138</sup>.

The substrate for NA, fetuin, is coated onto the surface of 96 well plates used in the ELLA. Virus is incubated in the wells in the presence or absence of serial serum dilutions. NA cleaves terminal sialic acid moieties from glycoprotein complexes, and therefore enzyme activity can be quantified by measuring the amount of galactose that is consequently at the terminus of each carbohydrate complex. This is accomplished using peroxidase-conjugated peanut-agglutinin (PNA), a lectin with specificity for terminal galactose. Following incubation with PNA-peroxidase, a substrate for peroxidase is added, leading to a colorimetric change proportional to NA activity. The ELLA has recently been used to measure NI antibody titres in a number of studies<sup>246,254–257</sup>, but assay optimization and details of the assay procedure to measure titres against the NA of human seasonal viruses have not been reported. This report describes steps that are

important for obtaining consistent results and provide data to support the use of ELLA in human serology.

## **MATERIALS AND METHODS**

### **Viruses**

Reassortant influenza viruses were generated by reverse genetics as described previously<sup>159,258</sup>. These viruses contain the HA (H6) gene from A/turkey/Massachusetts/3740/1965, gene segments encoding internal proteins from A/Puerto Rico/8/1934 (PR/34), and one of the following NA gene segments: N1 of A/Texas/36/1991 (TX/91), A/New Caledonia/20/1999 (NC/99), A/Solomon Islands/3/2006 (SI/06), A/Brisbane/59/2007 (BR/07), A/California/07/2009 (CA/09); and N2 of A/Wisconsin/67/2005 (WI/05) or A/Uruguay/716/2007 (UR/07). The following wild type influenza B viruses were used: B/Florida/4/2006 (B/FL/06) as representative of the B/Yamagata lineage and B/Brisbane/60/2008 (B/BR/08) as representative of the B/Victoria lineage. Viruses were cultured in the allantoic cavity of 9–12 day old embryonated chicken eggs at 33 °C, harvested 72 h post-inoculation and stored in aliquots at –80 °C.

### **Serum samples**

The following animal sera were used: ferret antisera against NC/99, UR/07, B/FL/06 and B/BR/08 generated by infecting ferrets with the respective wild-type influenza viruses, and cotton rat sera from naïve as well as PR/34 (H1N1)-immune animals. Ferret and cotton rat inoculations were performed following federal guidelines under a protocol approved by the institutional Animal Care and Use Committee. Pooled rabbit antisera (Capralogics, Hardwick, MA, USA) collected before and after immunization with purified NA were also used in this study. The NA was purified from WI/05 by cellulose acetate electrophoresis<sup>259</sup>. De-identified human sera were obtained from a clinical vaccine study in which groups of young, healthy adults were immunized with either a live or inactivated trivalent seasonal influenza vaccine. The study was approved by the Institutional Research Involving Human Subjects Committee. Antibody and cellular immune responses were measured before and 4 weeks after vaccination; these results were reported previously<sup>260</sup>. Unless otherwise noted, antisera were heat-inactivated at 56 °C for 45 min before conducting assays.

### ELLA procedure

The principles of the ELLA described by Lambre et al.<sup>142</sup> and Cate et al.<sup>246</sup> were followed to optimize and validate the method. The standard operating procedure (SOP) for this method is included in this manuscript as supplementary information. Fetuin (Sigma, St. Louis, MO, USA) was diluted to 25 µg/ml in 0.1 M phosphate buffered saline (PBS) and 100 µl added to each well to coat high-binding 96-well plates (Nalge Nunc, Rochester, NY, USA). Plates were stored at 4 °C and used 24 h to 2 months after coating. To determine the amount of antigen (virus) to use in ELLA, serial dilutions of the targeted H6 reassortant virus were prepared in Dulbecco's PBS (pH 7.4) – 0.9 mM CaCl<sub>2</sub> - 0.5 mM MgCl<sub>2</sub> containing 1 % bovine serum albumin (BSA) and 0.5 % Tween and then dispensed (50 µl/well) into fetuin-coated plates containing an equal volume of PBS. The plates were incubated for 16-18 h at 37 °C, then washed 6 times with PBS-0.05 % Tween 20 (PBST) before adding 100 µl peanut agglutinin (PNA) conjugated to horse-radish peroxidase (HRPO, Sigma). PNA–HRPO was used at the highest dilution that gave the maximum signal when titrated on fully digested fetuin. Plates were incubated at room temperature for 2 h and washed 3 times with PBST before adding o-phenylenediamine dihydrochloride (OPD, Sigma) to the plate. The colour reaction was stopped after 10 min by the addition of 1 N H<sub>2</sub>SO<sub>4</sub>. The plates were read at 490 nm for 0.1 s using a Victor V 96-well plate reader (PerkinElmer, Waltham, MA, USA). The dilution of virus (antigen) that resulted in 90–95 % maximum signal was elected for use in serology.

To measure the NI titres, each serum sample was heat treated (56 °C for 45 min) and then diluted serially in PBS–BSA. Fifty microliters of each dilution was added to duplicate wells of a fetuin-coated plate. An equal volume (50 µl) of the selected virus dilution was added to all serum-containing wells in addition to at least 4 wells containing diluent without serum that served as a positive (virus only) control. At least 4 wells were retained as a background control (PBS only). The plates were incubated for 16–18 h at 37 °C. As described for the virus titration, the plates were washed and PNA–HRPO was added to all wells. After a 2 h incubation period, the plates were washed and peroxidase substrate (OPD) was added. The colour reaction was stopped after 10 min and absorbance read. The mean absorbance of the background ( $A_{\text{bkg}}$ ) was subtracted from the test wells and positive control ( $A_{\text{pos}}$ ) wells. The percent NA activity was calculated by dividing the mean absorbance of duplicate test wells ( $A_{\text{test}}$ ) by the mean absorbance of virus only wells and multiplied by 100, i.e.  $(A_{\text{test}} - A_{\text{bkg}}) / (A_{\text{pos}} - A_{\text{bkg}}) \times 100$ . To determine percent NA inhibition, the percent activity was subtracted from 100. The NI titres were

defined as the reciprocal of the last dilution that resulted in at least 50 % inhibition. An alternative way to report results is to calculate the titres of replicate wells independently and then report the geometric mean of the duplicates as the 50 % end-point NI titre. In some instances, the exact 50 % inhibition ( $IC_{50}$ ) was determined by 4 parameter logistics regression analysis (GraphPad Prism software). An assay was considered valid if the background absorbance was less than 10 % of the virus only control, control sera had a similar NI titre to the median established in previous assays ( $\leq 2$ -fold difference), and the raw A490 values of the duplicates did not vary more than 20 %.

### **Miniaturized TBA method**

The miniaturized TBA method was followed as described previously<sup>239</sup>.

### **Statistical analysis**

Microsoft Office Excel was used to calculate standard deviations (SDs) of NI titres, percent coefficient of variation (% CV) for repeat assays and Pearson's correlation coefficient. Bland–Altman analysis<sup>261</sup> was performed using GraphPad Prism to assess the agreement between ELLA and TBA results.

## **RESULTS**

### **Assay optimization**

The published ELLA method<sup>142</sup> was optimized for routine analysis of human sera. This assay uses reassortant viruses with a mismatched HA as antigen (source of NA enzyme) to avoid non-specific inhibition by H1 and H3-specific antibodies in human sera. The H6 reassortant viruses containing the targeted NA were generated by reverse genetics as described previously<sup>159,258</sup>. Fetuin was used as substrate in the assay.

The assay requires the use of a defined amount of antigen (enzyme). Titration of different batches of H6N1 and H6N2 viruses on fetuin-coated plates resulted in expected sigmoidal regression curves, with maximum signal at A490 nm  $\sim 3.0$  and background  $< 0.2$ . Bacterial NA purified from *Vibrio cholerae* (Sigma) resulted in a similar titration curve. The linear range of absorbance values corresponded to enzyme activity units of 1–15  $\mu\text{U}$  NA/ml (Figure 1). Three different human sera and normal rabbit serum were titrated at each virus concentration, showing dependence of NI titre on the amount of antigen, with low titres at high virus concentrations and high titres when low amounts of antigen were added to the wells (Table 1). A minimal difference in NI titre was recorded when dilutions of virus that resulted in 50–90 % of the maximum signal were used. This corresponded

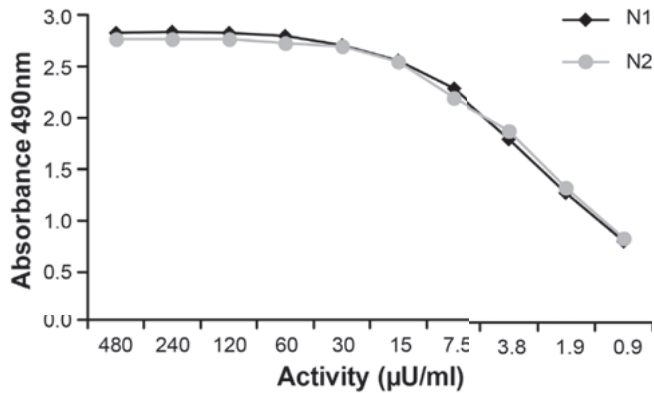


Figure 1 | **Examples of H6N1 and H6N2 virus titrations.** Serial dilutions of H6N1<sub>NC/99</sub> (N1 shown in black symbols) and H6N2<sub>WI/05</sub> (N2 shown in grey symbols) were incubated for 18 h in fetuin-coated plates and the reactivity with PNA determined as described in Section 2. Average absorbance of 2 wells is plotted against enzyme activity of N1 and N2 antigens. The enzyme activity of each antigen stock was determined previously using bacterial neuraminidase as the standard.

to the amount of virus within the linear region of the virus titration curve and resulted in a signal 5–10 times greater than background. Under optimized conditions this corresponded to absorbance signals (A<sub>490nm</sub>) of 1.5–2.5. To exploit the full range of the assay, the dilution of virus that was at the top of the linear range (90–95 % of maximal signal) was elected for routine serology because this amount of virus resulted in greatest signal:noise ratio and allowed maximum assay range (16-fold increases in titre could be measured). In the examples shown in Table 1, the amounts of virus selected for serology were dilutions of 1:640 for CA/09 and 1:60 for UR/07.

Since human sera usually contain NA-specific antibodies, initial experiments used sera from naïve animals to establish conditions that minimize the effect of non-specific inhibitors. Naïve cotton rat serum tested in ELLA against different dilutions of PR/34 showed considerable inhibition of NA activity when the recommended amount of virus was used (Figure 2). Freeze-thawing (F/T) of the serum sample did not reduce the non-specific inhibition significantly. However, heat-treatment at 56 °C for 45–60 min was sufficient to reduce non-specific inhibition so that titres of naïve serum against either H6N1 or H6N2 reassortant viruses were < 5. Non-specific inhibition of NA activity is dependent on HA subtype and therefore needs to be considered when assays use antigens that contain HA subtypes other than H6. In instances when heat-inactivation is

**Table 1 | NA inhibition titres of serum samples incubated with different amounts of antigen.**

Serum sample	Dilution of H6N1 <sub>CA/09</sub> (percent of maximum signal)							
	160 (100)	320 (98)	640 (89)	1,280 (80)	2,560 (46)	5,120 (30)	10,240 (18)	20,480 (11)
1	< 10	< 10	< 10	< 10	< 10	< 10	10	10
2	40	40	80	160	160	160	320	320
3	320	640	1,280	2,560	2,560	5,120	5,120	5,120
4	160	320	640	640	1,280	2,560	5,120	5,120

Serum sample	Dilution of H6N1 <sub>UR/07</sub> (percent of maximum signal)							
	20 (100)	40 (99)	80 (82)	160 (67)	320 (43)	640 (23)	1,280 (15)	2,560 (8)
1	10	< 10	< 10	< 10	10	10	20	20
2	20	40	40	80	80	80	160	160
3	80	160	320	320	640	640	2,560	2,560
4	40	80	80	160	320	320	640	640

insufficient to eliminate inhibitors, serum samples can be treated with receptor destroying enzyme (RDE) prior to heat treatment (results not shown).

Following heat-treatment, 2-fold dilutions of serum samples were added to duplicate wells of a washed, fetuin-coated plate. An equal volume of virus (antigen) was then added. Initial experiments included an incubation step of serum and virus in a 'dilution' plate. This resulted in inconsistent results, most likely reflecting adherence of virus to the dilution plate wells. The assay was therefore simplified to add serum dilutions and virus directly to the fetuin-coated plate.

Our initial experiments hoped to reduce assay time by optimizing conditions to allow short virus and substrate incubation times. However, the overall signal obtained after short incubations was low (data not shown) and therefore to preserve assay robustness, the assay was optimized for overnight (16–18 hr) incubation of virus/serum on the fetuin-coated plate. After this incubation, the lowest dilution of PNA-HRPO that resulted in maximum signal was added and the plates held at room temperature for 2 h.

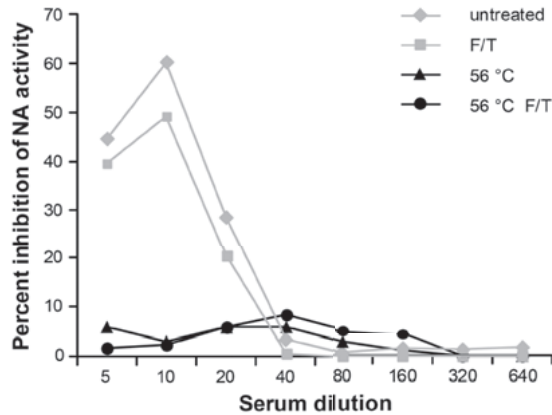


Figure 2 | **Heat treatment of sera is required to reduce non-specific inhibition of NA activity.** Serum from naïve cotton rats was either not treated, heated at 56 °C for 60 min, frozen and thawed 3 times (F/T), or heat-treated in addition to 3 freeze–thaws.

OPD was then added as substrate to all wells. OPD was used because it is used for other assays in our laboratory, however, it can be replaced by 3,3',5,5'-tetramethylbenzidine (TMB).

### Assay specificity

The specificity of ELLA was examined with monospecific sera. These included antisera from ferrets previously infected with NC/99 (H1N1), UR/07 (H3N2), B/FL/06 and B/BR/08 and antiserum from rabbits immunized with NA purified from WI/05 (H3N2). NI titres were measured against H6N1<sub>NC/99</sub>, H6N2<sub>WI/05</sub>, and H6N2<sub>UR/07</sub>. B/FL/06 and B/BR/08. Ferret anti-NC/99 inhibited enzyme activity of H6N1<sub>NC/99</sub> by a 64-fold higher titre than H6N2 viruses (Table 2). Similarly, rabbit antiserum specific for WI/05, had a 32-fold higher inhibition titre against H6N2<sub>WI/05</sub> than H6N1<sub>NC/99</sub>. Control sera did not inhibit enzyme activity of H6N1<sub>NC/99</sub> or H6N2<sub>WI/05</sub>. Ferret anti-B/FL/06 (B/Yamagata lineage) and anti-B/BR/08 (B/Victoria lineage) antisera did not react with the NA of influenza A viruses demonstrating specificity for the B antigens. Since H6 reassortants cannot be generated with the NA of B strains, the antibody titres measured against B viruses may not accurately quantify NA-specific antibodies because HA-specific antibodies may hinder access of substrate to NA's active site and thereby contribute to the reduced signal; nevertheless, the NI titre of the B/FL/06-specific serum was greater against the homologous virus than the heterologous B/Victoria-lineage and

Table 2 | Specificity of ELLA.

NA antigen <sup>a</sup>	NI titre of antiserum against					
	None	H1N1	H3N2		B	
		NC/99	WI/05 <sup>b</sup>	UR/07	B/FL/06	B/BR/08
NC/99 (N1)	< 10	1,280	80	< 10	< 10	< 10
WI/05 (N2)	< 10	10	2,560	80	< 10	< 10
UR/07 (N2)	10	20	2,560	160	< 10	< 10
B/FL/06	< 10	< 10	40	< 10	640	40
B/BR/08	< 10	< 10	< 10	< 10	40	160

<sup>a</sup>The N1 and N2 antigens were in the form of H6N1 and H6N2 reassortant viruses, respectively. Wild type B viruses were used in this assay.

<sup>b</sup>Antisera against NC/99, UR/07, B/FL/06 and B/BR/08 were generated by infection of ferrets. The anti-WI/05 serum was from rabbits immunized with purified NA purified mixed with an equal volume of Freund's complete adjuvant. The purification and mixing may have resulted in denaturation of some of the antigen.

vice versa, suggesting that these NAs are indeed antigenically distinct. Overall, these results demonstrate that virus-specific antisera show specificity for the homologous NA, although a small degree of cross-reactivity with a different subtype is observed when sera have very high homologous NI titres.

### Assay linearity

NA inhibition titration curves were established for human serum samples having a range of titres against N1<sub>NC/99</sub> and N2<sub>WI/05</sub>. This was accomplished by adding serial dilutions of serum samples to a fetuin-coated plate in duplicate wells, followed by addition of standardized amounts of H6N1<sub>NC/99</sub> and H6N2<sub>WI/05</sub> antigens. After incubation overnight, the assay was completed as outlined in the materials and methods section (the detailed method is provided as supplementary information). Absorbance increased as the amount of serum decreased, indicating the presence of increasing amounts of active NA. Percent inhibition was proportional to the amount of serum and followed a sigmoidal curve. Examples of human serum titrations are shown in Figure 3; one serum has a relatively high titre against the N1 antigen but lower titre against the N2 antigen whereas the second serum has similar titres against N1 and N2 antigens. To assess the linearity of the assay, NI antibody titres of different dilutions of 4 different samples with known titres were measured. The results demonstrated that the IC<sub>50</sub>



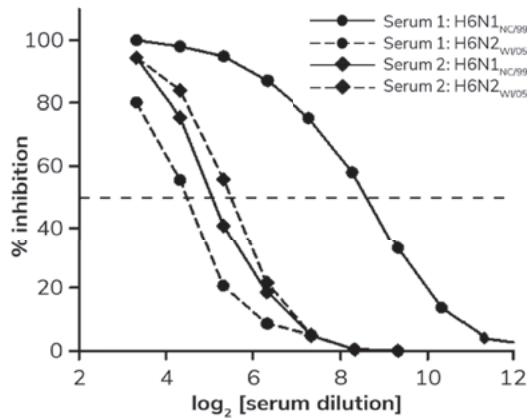


Figure 3 | Titration of sera against H6N1<sub>NC/99</sub> (solid lines) and H6N2<sub>WI/05</sub> (dashed lines) viruses. Percent enzyme inhibition for serial dilutions of serum 1 (filled circle) and serum 2 (filled diamond) is shown on each graph with the dashed horizontal line indicating 50 % inhibition. The NA inhibition titres of sera 1 and 2 against the NA of NC/99 were 320 and 20 respectively; the NA inhibition titres of sera 1 and 2 against the NA of WI/05 were 20 and 40 respectively.

calculated for each dilution varied < 10 % from the assigned value (results not shown) and supported relative accuracy of the method for measuring NI antibody titres over a range of 5–2,560.

### Assay precision

Repeatability of the assay was determined by measuring geometric mean NA inhibition titres (GMT) of 5 sera. One operator performed the assays in which each sample was tested 8 times on 3 separate 96-well plates. The results obtained for titrations within the same plate were highly reproducible (CV < 10 %). Plate-to-plate variability was also minimal (Table 3). In some instances (e.g. sample 2 tested against H6N1 virus), the percent inhibition calculated for a particular dilution was close to 50 % inhibition in all tests but because the titre is defined as the reciprocal of the last dilution that resulted in at least 50 % inhibition, a lower titre was assigned even when percent inhibition was 49 %. As a result, there were some instances when a 2-fold difference was reported for replicate serum titrations even though the assay variability was minimal (the average relative standard deviation calculated at approximately 50 % inhibition for assays against either N1 or N2 antigens was 7 %). This was the largest difference observed for sample 2 shown in Table 3 (H6N1<sub>NC/99</sub> antigen), even when assays were run on different

Table 3 | Assay precision: titration of samples within the same and different plates. Four replicates of each titration were conducted on each plate.

Plate	Sample 1		Sample 2		Sample 3		Sample 4		Sample 5	
	Titre ± SD <sup>a</sup>	%CV	Titre ± SD	%CV	Titre ± SD	%CV	Titre ± SD	%CV	Titre ± SD	%CV
Antigen 1: H6N1NC/99										
1	< 2.3	0	5.3 ± 0	0	6.3 ± 0	0	7.3 ± 0	0	6.3 ± 0	0
2	< 2.3	0	6.1 ± 0.5	8.2	6.3 ± 0	0	7.3 ± 0	0	6.3 ± 0	0
3	< 2.3	0	5.6 ± 0.5	8.9	6.3 ± 0	0	7.3 ± 0	0	6.3 ± 0	0
All	< 2.3	0	5.6 ± 0.5	8.7	6.3 ± 0	0	7.3 ± 0	0	6.3 ± 0	0
Antigen 2: H6N2W/05										
1	< 2.3	0	12.3 ± 0	0	4.3 ± 0	0	4.3 ± 0	0	5.7 ± 0.5	8.9
2	< 2.3	0	12.3 ± 0	0	4.3 ± 0	0	4.3 ± 0	0	5.3 ± 0	0
3	< 2.3	0	12.3 ± 0	0	4.3 ± 0	0	4.3 ± 0	0	6.3 ± 0	0
All	< 2.3	0	12.3 ± 0	0	4.3 ± 0	0	4.3 ± 0	0	5.7 ± 0.5	8.7

<sup>a</sup>The titre is shown as log 2.

days (total of 24 titrations). The titres for other samples shown in Table 3 were the same in all 24 tests, reflecting the unambiguous assignment of titre when percent inhibition was well below 50 % (e.g. sample 1 against either the N1 or N2 antigen) and when the percent inhibition was clearly greater than 50 % at the assigned titre (e.g. sample 2 against the N2 antigen). Therefore, these results show that when all samples are considered for repeatability, a 4-fold difference in 50 % end point titre is statistically significant ( $p \leq 0.05$ ). It is likely that smaller differences in titres reported as  $IC_{50}$  or geometric means from repeat assays would be considered significant.

Intermediate precision of the assay was also assessed, with 2 technicians running the same samples. Both operators were experienced and trained in the same standard operating procedure. Comparable results were obtained for each sample, with measurement of no more than a 2-fold difference in titre.

Changes to the source of fetuin, storage conditions of fetuin (4 °C instead of -20 °C), use of coated plates stored for different lengths of time at 4 °C and use of different plate readers were evaluated for assay robustness. The changes were acceptable because the modifications did not result in NI titres that differed from the titres obtained with the written SOP by more than 2-fold. As described earlier, the amount of antigen (virus) in the assay was critical and therefore the dilution of virus used for the purpose of serology was defined as the dilution that results in 90–95 % of maximum signal (Table 1). Although assay sensitivity may be increased by using less antigen (allowing elucidation of smaller differences in antigenic structure), an amount of antigen that gives a signal  $\geq 10$ -fold background and within the linear range of the virus titration resulted in reproducible sample titres.

### **Comparability between ELLA and mini-TBA methods**

Comparison of NI antibody titres of ferret antisera measured by the mini-TBA and ELLA against a number of N1 (NA of TX/91, NC/99, SI/06 and BR/07) and N2 (NA of AI/68, WU/95, SY/97, PA/99, WY/03, NY/04, UR/07) antigens demonstrated excellent correlation between NI titres, although the absolute titre was somewhat greater when measured by ELLA (Figures 4a and 4b; see Figures S1a and S1b in the supplemental material, available in *J. Virol. Methods Online*). This study also compared absolute titres and response rates measured by mini-TBA and ELLA methods of 32 serum pairs reported previously for a clinical study comparing responses to live, attenuated and inactivated split trivalent influenza vaccines<sup>262</sup>. As for the ferret sera and reported by Fritz et al.<sup>257</sup>, the human antibody titres measured by mini-TBA and ELLA methods

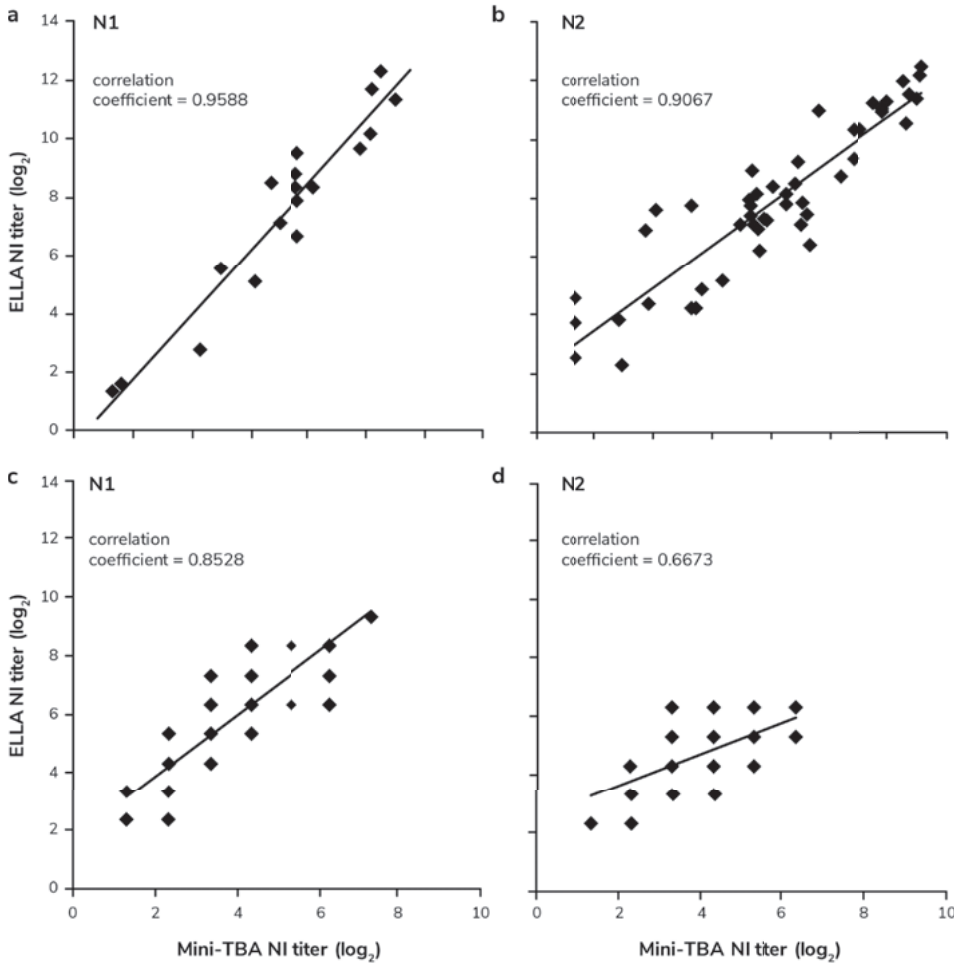


Figure 4 | **Correlation between NI titres measured by TBA and ELLA.** Titres are shown for ferret (a, b) and human (c, d) sera tested against H6N1 (a, c) and H6N2 (b, d) antigens. The number of titres compared (n) and Pearson’s correlation coefficient are indicated on each figure.

showed some correlation (Figures 4c and 4d), however NI titres generated by ELLA were often greater than the titres measured by the mini-TBA method (see Figures S1c and S1d in the supplemental material). Bland–Altman analysis showed better agreement of ELLA and mini-TBA NI antibody titres for ferret sera than human sera (see Figure S1 in the supplemental material), with the mini-TBA assay generating lower NI antibody titres against both H6N1 and H6N2 antigens for the majority of human sera tested. The bias ( $\pm$ SD) for agreement between TBA and ELLA titres against the

H6N1 antigen was  $-2.2 (\pm 15.7)$  for ferret sera but  $-45.0 (\pm 19.3)$  for human sera; and against the H6N2 antigen was  $-2.8 (\pm 30.2)$  for ferret sera but  $-22.1 (\pm 24.2)$  for human sera.

There was similar sensitivity in determining seroconversion by ELLA and mini-TBA methods. For paired samples taken before vaccination and 4 weeks later, increases in NI titres were evident in a similar proportion of volunteers; of 32 volunteers, the mini-TBA method identified 12 volunteers with increased serum NA inhibition titres against the N1 component (NC/99) and 13 volunteers with increased NA inhibition titres against the N2 component (WI/05) after vaccination<sup>262</sup>. The ELLA identified a similar number of volunteers that had increased NI titres following vaccination: the 12 volunteers identified by mini-TBA as well as two additional subjects had increased NI titres against the N1 component (14 seroconversions when considering a 2-fold increase as significant); 12 of the 13 volunteers who had increased NI titres identified by mini-TBA against the N2 component were identified by ELLA in addition to 2 volunteers who did not have increased titres measured in the mini-TBA assay (Table 4).

Table 4 | Assay sensitivity: NA inhibition titres against A/WI/05 (H3N2) following seasonal influenza vaccination of healthy adults.

Volunteer #	Mini-TBA			ELLA		
	Pre	Post	Fold increase	Pre	Post	Fold increase
1	10	10		20	20	
3	40	40		40	40	
4	10	10		20	20	
5	40	20		20	20	
6	5	10	2	10	20	2
7	10	20	2	20	40	2
8	5	20	4	10	40	4
9	< 5	< 5		< 10	< 10	
10	< 5	5	2	< 10	10	2
11	10	10		10	20	2
12	20	40	2	10	40	4



**Table 4 | Assay sensitivity: NA inhibition titres against A/WI/05 (H3N2) following seasonal influenza vaccination of healthy adults.**

Volunteer #	Mini-TBA			ELLA		
	Pre	Post	Fold increase	Pre	Post	Fold increase
13	80	80		40	40	
14	5	5		20	20	
15	20	20		20	20	
17	10	5		20	20	
18	10	20	2	10	20	2
19	< 5	5	2	< 10	10	2
20	10	10		40	40	
21	20	20		40	80	2
22	10	10		20	20	
23	10	10		40	20	
24	10	10		40	40	
25	10	10		20	20	
26	40	40		40	40	
28	10	20	2	20	40	2
29	10	20	2	20	40	2
30	10	40	4	20	80	4
31	10	10		40	40	
32	20	40	2	40	80	2
33	10	20	2	20	40	2
34	40	80	2	80	80	
35	80	40		40	20	

## DISCUSSION

NA inhibiting antibodies correlate with reduction in influenza disease<sup>128</sup> and therefore serology that routinely includes measurement of NI antibody titres is warranted. In the past, this was difficult to accomplish with the standard TBA method, but development of a miniaturized TBA method<sup>239</sup> and use of the enzyme-linked lectin assay<sup>142</sup> have resulted in more practical platforms that have allowed NI titres to be measured in a number of studies. This report describes steps that are important for obtaining consistent results and demonstrates reproducibility of the ELLA method. An SOP that can be used by investigators interested in performing this assay routinely is provided as supplementary information.

This study is not the first to use ELLA to measure human serum NI antibody titres – others have adapted the method published by Lambre et al.<sup>142</sup> and measured responses to NA in clinical studies that demonstrate a number of important points. Couch et al.<sup>255</sup> demonstrated that current influenza vaccines contain immunogenic amounts of NA and that NI antibody seroconversion rates are greater when vaccine dose is increased<sup>246</sup>. Fritz et al.<sup>257</sup> compared the TBA and ELLA methods in a study measuring NI titres after vaccination with a cell grown inactivated whole virus H5N1 vaccine, demonstrating increased sensitivity of ELLA compared to the TBA method and showing excellent increases in NI titre following a single dose of vaccine. NI titres were not increased following a boost with this vaccine, possibly reflecting competition between HA and NA antigens when they are presented on the same whole virus<sup>263</sup>. Since antigenic competition can be circumvented by immunizing with dissociated HA and NA antigens in mice<sup>264</sup>, it would be of interest to determine whether NI titres are boosted following vaccination with split, inactivated influenza vaccines. NI titres measured after vaccination with VLPs suggest the requirement for dissociated HA and NA antigens may not be applicable to all vaccine types – responses to both HA and NA were boosted following receipt of a second dose of adjuvanted H7N9 VLPs<sup>256</sup>. Interestingly, significant increases in NI but not HAI titres were evident after the first dose of both unadjuvanted and adjuvanted VLPs in this clinical study, suggesting that the N9 component of the VLP vaccine is more immunogenic than H7 or that the NI assay has greater sensitivity than the HAI assay.

Although NI titres measured by ELLA correlate with the titres measured in the mini-TBA assay, the values measured by ELLA are often higher than measured by the more traditional assay. This difference was more noticeable for titration of human sera than ferret sera. It is not known whether the difference in assay bias reflects the relatively low NI antibody titres of samples used in the human analysis or whether it is due to

qualitative differences in antibodies or the presence of additional non-specific inhibitors in human sera. Further studies to identify the factors that contribute to this difference will be helpful in understanding how best to interpret results and may lead to assay improvements.

As for other serologic assays, non-specific inhibitors present in serum samples need to be removed to avoid inaccurate measurement of antibody titres. The non-specific inhibition observed in the ELLA when animal (mouse and cotton rat) sera were tested against H6N1 and H6N2 viruses was removed by heat-treatment (56 °C for 45 min), indicating the presence of thermolabile  $\beta$ -inhibitors.  $\beta$ -inhibitors are  $\text{Ca}^{2+}$ -dependent (C-type) lectins that bind to mannose-rich glycans on glycoproteins<sup>265</sup>. Surfactant protein D is an example of a  $\beta$ -inhibitor in the lung; both infectivity and NA activity are inhibited in the presence of this inhibitor<sup>266</sup>, indicating that SP-D has capacity to bind to glycans on HA and NA, blocking their functional activity. Conglutinin is a  $\beta$ -inhibitor present in human serum<sup>267</sup> that may contribute to non-specific inhibition of NA activity in the ELLA. Other non-specific inhibitors ( $\alpha$  and  $\gamma$ -class) of influenza HA have been described, particularly in relation to H3N2 virus haemagglutination and infectivity<sup>268–270</sup>, however, these do not appear to result in non-specific inhibition of NA activity in assays that use H6 reassortant viruses as antigen.

Various forms of NA have been used in ELLAs – purified recombinant NA<sup>257</sup>, reassortant H6N1 and H6N2 viruses (this report), detergent split wild-type virus<sup>246</sup> and virus like particles<sup>254,256</sup>. Although it is advantageous to use reassortant whole viruses that contain NA in the “natural” form in ELLA, it has not been possible to generate reassortant viruses expressing the NA of influenza B viruses and mismatched HAs. Purified recombinant NA protein or VLPs are therefore needed to measure antibody responses to the NA of influenza B viruses. Regardless of antigen source, all assays show a correlation between TBA and ELLA titres (our data as well as that reported by Fritz et al.<sup>257</sup>), with NI antibody titres measured by ELLA somewhat greater than measured by the TBA method.

The results of assays that examine repeatability show that a 4-fold rise in NA inhibition titre is indicative of seroconversion. However, when consecutive serum samples are run on the same plate and the raw data examined to verify that differences in titre are not ambiguous (as may be the case when percent inhibition is close to 50 %), assay variability may be less, allowing a 2-fold increase in NI titre to be used as evidence of a NA-specific antibody response. Seroconversion rates that are based on 2-fold differences in NI titre should be verified by performing repeat assays.



In conclusion, the steps taken to optimize and validate the ELLA for measuring NA inhibiting antibody titres in human sera are described in this report. The data show that this very practical assay is subtype specific, and yields reproducible results. Since routine measurement of NI antibody titres is now possible, the immunogenicity of NA in influenza vaccines can be examined more routinely and NI antibody titres that correlate with protection from clinical disease can be identified.

## **ACKNOWLEDGEMENTS**

This project was funded by intramural PanFlu funds (CBER); LC was supported by funds administered through an ORISE training fellowship; KW was supported by NOW VICI grant 91896613 (awarded to RAMF). We are indebted to Miranda de Graaf, Wei Wang and Freya Lynn for critical reading of the manuscript, and Lev Sirota for guidance in statistical analysis.



## CHAPTER 6

# Optimization of an Enzyme-Linked Lectin Assay suitable for Rapid Antigenic Characterization of the Neuraminidase of Human Influenza A(H3N2) Viruses

Kim Westgeest, Theo M. Bestebroer, Monique I. J. Spronken, Jin Gao, Laura Couzens, Albert D.M.E. Osterhaus, Maryna Eichelberger, Ron A.M. Fouchier, Miranda de Graaf

*Journal of Virological Methods*, volume 217, 1 June 2015, pages 55-63



## **ABSTRACT**

Antibodies to neuraminidase (NA), the second most abundant surface protein of the influenza virus, contribute to protection against influenza virus infection. Although traditional and miniaturized thiobarbituric acid (TBA) neuraminidase inhibition (NI) assays have been successfully used to characterize the antigenic properties of NA, these methods are cumbersome and not easily amendable to rapid screening. An additional difficulty of the NI assay is the interference by haemagglutinin (HA)-specific antibodies. To prevent interference of HA-specific antibodies, most NI assays are performed with recombinant viruses containing a mismatched HA. However, generation of these viruses is time consuming and unsuitable for large-scale surveillance. The feasibility of using the recently developed enzyme-linked lectin assay (ELLA) to evaluate the antigenic relatedness of NA of wild type A(H3N2) viruses was assessed. Rather than using recombinant viruses, wild type A(H3N2) viruses were used as antigen with ferret sera elicited against recombinant viruses with a mismatched HA. In this study, details of the critical steps that are needed to modify and optimize the NI ELLA in a format that is reproducible, highly sensitive, and useful for influenza virus surveillance to monitor antigenic drift of NA are provided.

## INTRODUCTION

Influenza vaccines reduce the morbidity and mortality associated with annual influenza epidemics. The seasonal influenza vaccine is designed to protect against circulating influenza A H1N1 viruses (A(H1N1)), influenza A H3N2 viruses (A(H3N2)), and influenza B viruses. The influenza virus escapes host immunity through mutations in the major surface glycoproteins haemagglutinin (HA) and neuraminidase (NA). This process is known as antigenic drift<sup>157,237</sup> and as a result of this drift, the influenza vaccine has to be updated frequently. In the period from 1999 to 2010, the A(H3N2) virus component was updated 6 times<sup>155</sup>. Recently, it has been shown that only few mutations near the receptor-binding site of HA are responsible for antigenic drift of A(H3N2) viruses circulating between 1968 and 2003<sup>236</sup>. For NA a number of antigenic sites have been described<sup>134</sup>. These antigenic regions surround the enzyme's active site<sup>135,136</sup> and are highly variable, most likely due to immune pressure<sup>123,124</sup>. Influenza virus surveillance by national influenza centres is done year-round<sup>151,155</sup>. Representatives of the predominant circulating viruses are sent to the World Health Organization (WHO) Collaborating Centres. These centres characterize the viruses by sequencing the HA and NA genes and performing haemagglutination inhibition (HI) assays<sup>155</sup>. During vaccine strain selection, the main focus is on the genetic and antigenic characterization of HA<sup>158</sup>.

Influenza viruses attach to the host cell surface via binding of the HA to sialic acid-containing receptors<sup>58</sup>. The enzymatic activity of NA allows virus release from the cell<sup>66-68</sup> by cleaving the sialic acid residues from the newly formed virus particles and from the host cell<sup>70</sup>. Preclinical and clinical studies showed that NA-specific immunity can reduce the severity of disease<sup>125,128-132</sup>. Antibodies directed toward NA inhibit release and spread of newly formed virus particles from infected cells<sup>133</sup>. The antigenic drift of NA does not closely correspond to that of HA<sup>122,137,159</sup>. Considering these findings, investigating options to include routine analysis of NA during vaccine strain selection next to HA seems to be warranted.

Antigenic characterization of NA can be performed using NA inhibition (NI) assays to determine the extent of antibody-mediated interference with enzyme activity<sup>138</sup>. These assays rely on the enzymatic sialidase activity by measuring cleavage of sialic acid from highly glycosylated proteins such as fetuin. The NI thiobarbituric acid (TBA) assay<sup>139,140</sup> is based on the detection of free sialic acid. This assay is recommended by the WHO<sup>271</sup>, but the use of large volumes of hazardous chemicals and laborious procedures has impeded antigenic characterization of NA during influenza virus surveillance. A simplified and miniaturized version of the TBA was developed<sup>239</sup>, but this assay still remains

cumbersome. The recently developed enzyme linked lectin assay (ELLA)<sup>246,254,255,257,272,273</sup> also relies on the sialidase activity of NA, but instead of measuring free sialic acid, it detects the terminal galactose that becomes exposed after sialic acid cleavage by NA.

A complication of NI assays, is the interference of HA-specific antibodies that block NA activity non-specifically when they bind to HA<sup>137</sup>. Recombinant influenza viruses with a heterologous HA are commonly used for NI assays. Antibodies directed toward the H1 or H3 HA of A(H1N1) and A(H3N2) viruses do not cross-react with a heterologous HA (e.g. H6), and hence only the interaction between NA and NA-specific antibodies is measured<sup>273</sup>. However, the generation of recombinant viruses is time-consuming for large numbers of viruses and therefore this method is not suitable for large-scale surveillance of antigenic drift of circulating influenza viruses. For analysis of the antigenic drift of NA, it would be ideal to have the capability of using wild type viruses as antigen in assays that are not impacted by non-specific inhibitors, including antibodies to HA.

In this study, optimized methods to enable rapid antigenic characterization of NA, with wild type viruses as antigen, are described. Since the ELLA is less laborious and shows a good correlation to the miniaturized TBA<sup>257</sup>, this assay was selected as a platform. To prevent interference by antibodies directed against HA of wild type A(H3N2) viruses, ferret sera were raised against recombinant influenza A H7N2 viruses (A(H7N2)) viruses that contain the NA of various A(H3N2) viruses. Through this approach it is possible to screen wild type viruses, thus preventing the time-consuming generation of recombinant viruses or proteins for each virus of interest. Reproducibility and sensitivity of the NI assay were highest using virus concentrations that resulted in ~50 % of total NA activity of that virus. Non-specific inhibition of ferret sera was observed for some wild type viruses, especially A(H3N2) viruses, but the critical steps to overcome this non-specific inhibition and obtain reproducible and highly sensitive results are also described.

## **MATERIALS AND METHODS**

### **Cells**

Madin-Darby canine kidney (MDCK) cells were cultured in Eagle's minimal essential medium (EMEM, Lonza, Breda, The Netherlands) supplemented with 10 % fetal bovine serum (FBS, Sigma-Aldrich, St. Louis, MO, USA), 100U/ml penicillin (Lonza), 100U/ml streptomycin (Lonza), 2mM glutamine (Lonza), 1.5mg/ml sodium bicarbonate (Lonza), 10mM HEPES (Lonza), and non-essential amino acids (MP Biomedicals, Europe, Illkirch, France).

293T cells were cultured in Dulbecco's modified Eagle's medium (DMEM, Lonza) supplemented with 10 % FBS, 100U/ml penicillin, 100U/ml streptomycin, 2mM glutamine, 1mM sodium pyruvate (Life Technologies, Bleiswijk, The Netherlands), 500 ug/ml geneticin (Life Technologies) and non-essential amino acids.

### Plasmids

The NA gene segments of A(H3N2) viruses A/Bilthoven/21793/1972 (BI/72); A/Bilthoven/1761/1976 (BI/76); A/Netherlands/233/1982 (NL/82); A/Netherlands/823/1992 (NL/92); A/Netherlands/178/1995 (NL/95); A/Netherlands/69/2009 (NL/09) and the HA and NA gene segments of A(H2N2) viruses A/Netherlands/M1/1957 (NL/57) and A/Netherlands/B1/1968 (NL/68) were amplified by reverse transcription polymerase chain reaction (RT-PCR) and cloned in a modified version of the bidirectional reverse genetics plasmid pHW2000<sup>258,274</sup>. To reduce pathogenicity, the multibasic cleavage site (MBCS) was removed from the bidirectional reverse genetics HA plasmid of the highly pathogenic avian A(H7N7) virus (A/Netherlands/33/03)<sup>275</sup>. This was done using the QuikChange multi-site-directed mutagenesis kit (Stratagene, Huissen, The Netherlands) according to the instructions of the manufacturer with specific primers (available from the authors upon request). The plasmids encoding the internal genes of A/Netherlands/219/03 (H7N7) have been described previously<sup>275</sup>.

The accession numbers that were used are as follows: for NA, CY112307, CY113199, CY114439, CY113735, CY116590, CY113023, KM402803, and KM402811; for HA, KM402801 and KM402809.

### Viruses

All human and recombinant influenza A viruses were propagated at 37 °C and influenza B viruses at 33 °C in MDCK cells in EMEM supplemented with, 100U/ml penicillin, 100U/ml streptomycin, 2mM glutamine, 1.5 mg/ml sodium bicarbonate, 10mM HEPES, non-essential amino acids and 1 µg/mL L-1-tosylamide-2-phenylethyl chloromethyl ketone-treated trypsin (Sigma-Aldrich, Zwijndrecht, The Netherlands). Avian influenza viruses were propagated in 11-day old embryonated chicken eggs. All viruses were harvested 48 h post-inoculation and cell debris was removed by centrifugation for 15 min at 3,000 rpm. Supernatant was either immediately stored in suitable aliquots at -80 °C or, if needed, after concentration using an Amicon Ultra-15 Centrifugal Filter (Millipore, Amsterdam, The Netherlands).

Recombinant A(H2N2) and A(H7N2) viruses were generated by reverse genetics using transient calcium phosphate-mediated transfections of 293T cells as described previously<sup>274</sup>. A(H2N2) viruses were generated with plasmids carrying the HA and NA gene segments of A(H2N2) viruses (NL/57 or NL/68) and the six remaining gene segments of A/Puerto Rico/8/1934 (H1N1) under biosafety level 2 (BSL-2) conditions. A(H7N2) viruses were generated under biosafety level 3 (BSL-3) conditions with plasmids carrying the HA gene segment of A/Netherlands/33/03 (H7N7) without the MBCS, the internal gene segments of A/Netherlands/219/03 (H7N7)<sup>275</sup>, and one of the following NA gene segments of A(H3N2) viruses: A/Bilthoven/16190/1968 (BI/68)<sup>276</sup>; BI/72; BI/76; NL/82; NL/92; NL/95; A/Netherlands/213/2003 (NL/03)<sup>277</sup>; NL/09; and A(H2N2) viruses: NL/57 or NL/68. The supernatant of the transfected cells was harvested 48 h post-transfection and was passaged twice on MDCK cells. MDCK passage 2 supernatant was harvested 48 h post-inoculation and cell debris was removed by centrifugation for 15 min at 3,000 rpm and immediately stored in suitable aliquots at  $-80^{\circ}\text{C}$ .

### **Generation of ferret sera**

Male ferrets (1-year-old, 0.8–1 kg) were inoculated intranasally with 1 ml of A(H7N2) virus containing MDCK passage 2 supernatant under animal BSL-3 conditions. Sera were collected two weeks post-inoculation. RNA was isolated from ferret sera and RT-PCR was performed as described previously<sup>278</sup> to detect A(H7N2) virus. Once absence of viral RNA was confirmed, sera were stored in suitable aliquots at  $-20^{\circ}\text{C}$ . The independent animal experimentation ethical review committee “Stichting DEC Consult” approved all animal protocols (Erasmus MC permit number EMC 2617), and experiments were performed in compliance with Dutch and European guidelines for animal experimentation.

Unless noted otherwise, sera for the NI ELLA were treated with Burnet’s receptor destroying enzyme (RDE) filtrated from cultures of *Vibrio cholerae*<sup>279</sup> (Burnet and Stone, 1947). Sera were incubated overnight at  $37^{\circ}\text{C}$  in a ratio of 1:6 with a 10-fold dilution of RDE in Dulbecco’s phosphate-buffered saline (DPBS) containing 0.133 g/L  $\text{CaCl}_2$  and 0.1 g/L  $\text{MgCl}_2$  with 1 % Bovine Serum Albumin (BSA) solution (Sigma-Aldrich) and 0.05 % Tween 20 (Sigma-Aldrich, DPBST<sub>BSA</sub>), followed by heat inactivation at  $56^{\circ}\text{C}$  for 8 h.

### **ELLA and NI ELLA**

The principals of the NI ELLA described by Lambre et al.<sup>272</sup> and Cate et al.<sup>246</sup> were optimized for rapid antigenic analysis of wild type viruses. Stock solutions of fetuin were



made by dissolving fetuin from FBS (Sigma-Aldrich) in Coating Solution Concentrate (tebu-bio, Heerhugowaard, The Netherlands) to a concentration of 25 mg/ml. Fetuin was further diluted to a concentration of 25 µg/ml in Coating Solution Concentrate and used to coat Nunc-Immuno™ MicroWell™ 96 well solid plates (100 µl/well; Sanbio, Uden, The Netherlands) at 4 °C for at least 24 h. Aliquoted fetuin (25 mg/ml) could be stored for at least 6 months at -20 °C and coated plates could be stored for at least 2 months at 4 °C without a significant decrease in reactivity.

The ELLA was used to determine the optimal amount of each virus (antigen) for routine serology. Two-fold serial dilutions of the antigens were made in DPBST<sub>BSA</sub>. Fetuin-coated plates were washed 3 times in 0.1 M phosphate buffered saline (PBS, pH 7.4) with 0.05 % Tween 20 (PBST, Sigma-Aldrich). Fifty microlitre of the serial dilutions was transferred to duplicate fetuin-coated plates containing 50 µl DPBST<sub>BSA</sub>. Plates were sealed and incubated at 37 °C for 16-18 h. The plates were subsequently washed six times with PBST, and 100 µl/well of an optimized dilution of horseradish peroxidase-conjugated peanut agglutinin lectin (PNA-HRPO, Sigma-Aldrich) in DPBS containing 0.133 g/L CaCl<sub>2</sub> and 0.1 g/L MgCl<sub>2</sub> with 1 % BSA (DPBS<sub>BSA</sub>) was added. The optimal PNA-HRPO concentration, is the concentration that gives the maximal signal in ELLA assays that allowed complete digestion of sialic acid from fetuin. PNA-HRPO could be stored at -20 °C in DPBS<sub>BSA</sub> for a maximum of one month, without loss of signal intensity. Plates were incubated at room temperature for 2 h followed by washing three times with PBST. O-Phenylenediamine dihydrochloride (OPD, Sigma-Aldrich) substrate was freshly prepared following the manufacturer's instruction and added to the plate (100 µl/well). The reaction was stopped after 10 min by the addition of stop solution (0.5 M H<sub>2</sub>SO<sub>4</sub>, 100 µl/well). The plates were read at 490 nm (OD<sub>490</sub>) for 0.25 s using an Infinite 200 96-well plate reader (Tecan, Giessen, The Netherlands). The mean background (A<sub>bkg</sub>) value was established for each plate using at least four wells that were treated identically to test wells, but in the absence of antigen. The NA activity was determined by subtracting the background value from the test well values (A<sub>test</sub>) i.e. (A<sub>test</sub> - A<sub>bkg</sub>). Next the NA activity was plotted against the antigen dilutions. Unless noted otherwise, the dilution of antigen that resulted in 50 % of total NA activity but with an OD<sub>490</sub> > 0.75 was selected to subsequently perform the NI ELLA.

For the NI ELLA, serum samples were serially diluted (2-fold) in DPBST<sub>BSA</sub> and incubated in duplicate fetuin-coated plates with an equal volume (50 µl) of the selected antigen dilution in DPBST<sub>BSA</sub>. These plates were subsequently sealed and incubated for 16-18 h at 37 °C. Incubation with PNA-HRPO, the addition of substrate, stop buffer and

subsequent absorbance measurements were performed as described above for the ELLA. The NI titres were calculated as follows, first the mean background was subtracted from the absorbance of the test wells and the wells that contained antigen but not ferret serum (positive control,  $A_{\text{pos}}$ ). Next, the percentage of NA activity was calculated by dividing the mean absorbance of test wells by the mean absorbance of positive control wells and multiplied by 100, i.e.  $(A_{\text{test}} - A_{\text{bkg}}) / (A_{\text{pos}} - A_{\text{bkg}}) \times 100$ . Finally, to determine the percentage of NA inhibition, the percentage of activity was subtracted from 100. The NI titres were defined as the reciprocal of the last dilution that resulted in at least 50 % inhibition. When the duplicate titres were different, the geometric mean titres were calculated. An assay was considered valid if the NI titres of control sera run in separate plates and assays using the same conditions yielded similar titres (titre  $\leq$  2-fold difference). The OD<sub>490</sub> values of the duplicates did not vary more than 20 %.

### **Haemagglutination inhibition assay**

HI titres were determined using standard procedures<sup>280</sup> (Hirst, 1943). In brief, antisera were pre-treated overnight with RDE at 37 °C and heat inactivated at 56 °C for 1 h. Twofold serial dilutions of RDE-treated antisera (in 50  $\mu$ l), starting at a 1:20 dilution, were incubated with 25  $\mu$ l of four HA units of each virus, and incubated at 37 °C for 30 min. Next, 25  $\mu$ l of 1 % turkey erythrocytes were added, followed by an 1 h incubation at 4 °C. The reciprocal of the last serum dilution that was able to block the agglutination of the turkey erythrocytes was recorded as the HI titre.

## **RESULTS**

### **Principles of the ELLA and NI ELLA**

The ELLA measures the enzymatic activity of NA by detecting desialylation of the highly glycosylated fetuin by NA. PNA-HRPO is used to detect the terminal galactoses that become exposed after desialylation of fetuin by NA. The intensity of the signal after addition of the substrate is dependent on the level of desialylation and thus NA activity. In the NI ELLA, binding of NA by specific antibodies will inhibit the enzymatic function of NA and result in a reduction of desialylation and, hence, in a reduction of the signal. An illustration of the ELLA and NI ELLA principles is shown in Figure 1. To enable rapid antigenic characterization of the NA of wild-type human A(H3N2) viruses, monospecific ferret antisera raised against recombinant viruses with a mismatched HA were generated. These recombinant A(H7N2) viruses contain the NA of A(H3N2) viruses and the HA and internal gene segments of an A(H7N7) virus. Therefore, the HA-specific

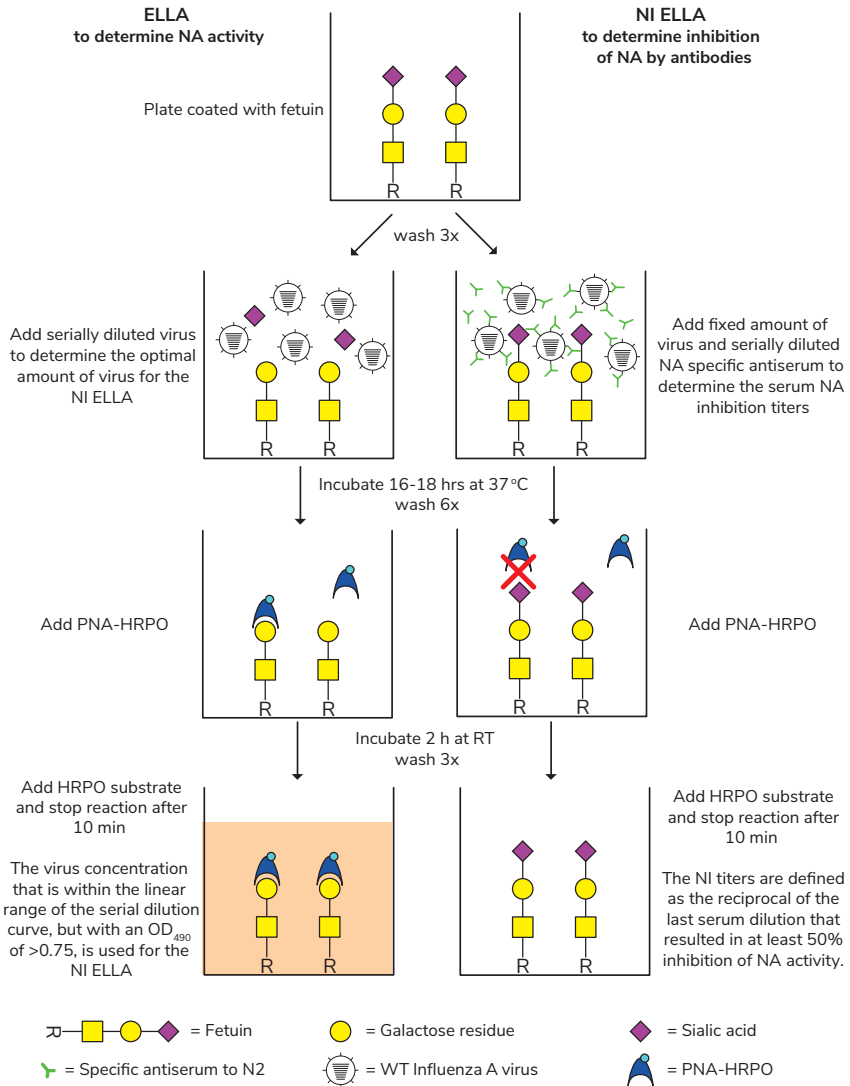


Figure 1 | Principles of the ELLA (left) and the NI ELLA (right). After addition of influenza virus to a fetuin-coated well, the NA cleaves sialic acid from the fetuin, thus unmasking a terminal galactose residue. This terminal galactose residue is specifically recognized by PNA-HRPO. Upon addition of the peroxidase substrate, a detectable colour change is produced which can be measured at  $OD_{490}$ . By using different dilutions of virus in the ELLA, the optimal virus dilution for the NI ELLA can be determined. For the NI ELLA, serially diluted serum is added to a fixed amount of virus ( $OD_{490} > 0.75$ ). If the NA activity is neutralized by NA-specific antibodies, sialic acid will not be cleaved from the terminal galactose residue and no colour change is detected.

antibodies present in the ferret sera will not interfere with wild type A(H3N2) viruses in the NI assay. The A(H7N7) virus was selected for its ability to replicate efficiently in ferrets, thereby producing high levels of antibodies, but in principle other influenza virus subtypes can be used.

### **Assay suitability with different influenza subtypes**

A panel of avian and human influenza viruses was tested in the ELLA and NI ELLA (Table 1) to evaluate the use of wild type viruses as antigens. First, the ELLA was conducted to determine the NA activity for each antigen at different concentrations. Next, the NI ELLAs were conducted using the antigen concentration that gave 50 % of the maximal NA activity. However, when these viruses were tested against naïve ferret sera there was already an inhibitory effect of the sera observed for some of these viruses. This inhibitory effect was highest for human A(H3N2) and avian A(H4N2) viruses (Table 1).

### **Treatment of serum to eliminate non-specific inhibition of NA activity**

Several options were investigated to eliminate the non-specific inhibitory effect of sera used in the NI ELLA: purification of (anti)serum, sucrose-gradient purification of viruses, commercial *Vibrio cholerae* neuraminidase (VCNA, Sigma–Aldrich) or RDE treatment of sera, and harvesting virus stocks at varying times after inoculation. VCNA and RDE treatment eliminated most of the non-specific inhibition, whereas the other approaches were unsuccessful (data not shown). Sera used in the HI assay are treated with RDE, followed by heat inactivation at 56 °C to inactivate non-specific inhibitors<sup>271</sup>. Similar inhibitors may explain the inhibitory effect of naïve ferret sera in the NI ELLA.

Similar to NA, RDE is a neuraminidase and should therefore be inactivated before use in the NI assay. RDE-treated sera were heat inactivated at 56 °C for 1, 2, 4 and 8 h and tested in the ELLA for residual neuraminidase activity of the RDE (data not shown). At least 4 h of heat inactivation at 56 °C was required to eliminate neuraminidase activity of RDE.

Next, untreated and RDE-treated sera were compared in the NI ELLA (Figure 2). Although the inhibitory effect of naïve ferret serum was eliminated upon treatment with RDE for 16 h and heat inactivation at 56 °C for 8 h, the specific inhibitory effect of antisera targeted to the N2 antigen was still present. To test the effect of the duration of the heat treatment on antibodies, a panel of anti-A(H7N2) sera were treated with RDE and subjected to various durations of heat inactivation at 56 °C. These sera were tested in NI ELLA and HI assay (Table 2). Each serum was tested against the homologous recombinant A(H7N2) virus in the HI assay. The same sera were tested against their N2-homologous

Table 1 | Non-specific inhibition of avian and human influenza viruses by naïve ferret sera as measured in the NI ELLA\*.

Type	Host	Subtype/ Lineage <sup>a</sup>	Strain	NI titre
A	Avian	H4N2	A/Mallard/Netherlands/8/2007	1,810
		H3N8	A/Mallard/Netherlands/37/2010	28
		H3N2	A/Mallard/Netherlands/51/2008	160
		H1N1	A/Eurasian Wigeon/Netherlands/6/2007	14
	Human	sH1N1	A/Netherlands/364/2006	20
			A/Netherlands/246/2008	20
		pH1N1	A/Netherlands/602/2009	160
			A/California/04/2009	113
			sH2N2 <sup>b</sup>	NL/57
		sH3N2	NL/68	160
			BI/68	640
			BI/76	905
			NL/82	640
			NL/92	2,560
B	Human	Victoria	A/Malaysia/2506/2004	< 10
		Yamagata	A/Florida/04/2006	< 10

\*All viruses were tested against two naïve ferret sera. These sera were not treated with RDE or heat inactivation. The geometric mean titre was calculated from these titres.

<sup>a</sup>Seasonal (s) and pandemic (p).

<sup>b</sup>Recombinant strains containing the internal genes of A/Puerto Rico/8/1934 (H1N1).

recombinant A(H2N2) or wild type A(H3N2) viruses in the NI ELLA. The duration of heat inactivation after the RDE treatment did not affect the HI and NI titres.

Thus, RDE treatment followed by heat inactivation at 56 °C for 8 h was sufficient to inactivate RDE activity without affecting the HI and NI titres of ferret antisera. Therefore, ferret sera treated with RDE and heat-inactivated at 56 °C for 8 h was subsequently used for the NI ELLAs.

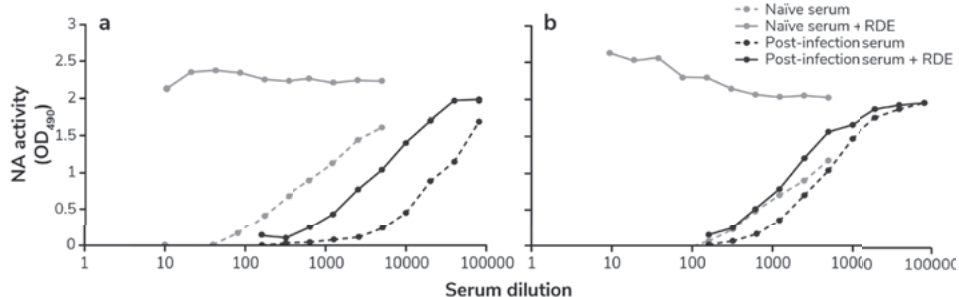


Figure 2 | **Neutralizing activity of naïve and post-infection sera after RDE treatment.** Viruses BI/76 (a) and NL/92 (b) were tested against serially diluted naïve ferret sera (grey lines) and their serially diluted N2-homologous A(H7N2) ferret sera (black lines). Sera were incubated with RDE (solid lines) and without RDE (dashed lines) at 37 °C for 16 h and RDE activity was heat inactivated by incubation at 56 °C for 8 h.

Table 2 | **HI and NI titres after different durations of RDE heat inactivation.**

A(H7N2) antisera	HI titres <sup>a</sup>				NI titre <sup>b</sup>			
	1 h	2 h	4 h	8 h	1 h	2 h	4 h	8 h
NL/57	1,280	761	1810	1,280	5,120	5,120	5,120	5,120
NL/68	640	538	640	640	2,560	2,560	2,560	2,560
BI/68	640	453	640	640	20,480	20,480	20,480	20,480
BI/76	269	320	320	320	5,120	3840	5,120	5,120
NL/82	640	538	640	640	160	320	226	226
NL/92	538	453	453	538	1,280	1,280	1,280	1,280
NL/09	160	160	160	160	2,560	2,560	2,560	2,560

A(H7N2) antisera were incubated with RDE for 16 h at 37 °C and heat inactivated at 56 °C for 1, 2, 4, or 8 h. The geometric mean titres were calculated from duplicate titres.

<sup>a</sup>Recombinant A(H7N2) viruses were tested against their homologous anti-A(H7N2) ferret sera in the HI assay.

<sup>b</sup>Recombinant A(H2N2) and wild type A(H3N2) viruses were tested in the NI ELLA against their N2-homologous A(H7N2) ferret sera.

### Assay sensitivity and reproducibility

The optimal amount of antigen to use in the NI ELLA was also determined. First A(H3N2) wild type viruses were titrated and measured in the ELLA. Most of these titration curves were sigmoidal, with a maximum signal of 3.0 and background of  $< 0.2$  at  $OD_{490}$ . Two viruses, NL/92 and NL/95, were tested in the NI ELLA at concentrations that spanned the antigen titration curve. Each 2-fold dilution of the antigen (virus) was tested with monospecific ferret antisera directed against the homologous N2 (Figure 3a and 3b) and NL/95 was also tested against two heterologous N2 (Figure 3c and 3d). The NI titre was dependent on the amount of N2 antigen, with low NI titres upon use of high amounts of N2 antigen, and high NI titres upon use of low amounts of N2 antigen. However, a minimal difference in NI titre was observed when the N2 antigen concentrations were within the linear trajectory of the titration curve, which represents approximately 21–87 % of total NA activity. This corresponded to an NA activity ( $OD_{490}$ ) between approximately 0.6–2.5. Therefore, to standardize the amount of N2 antigen in the assay, a dilution of virus that resulted in 50 % (minimum of 25 % and maximum of 75 %) of total NA activity was selected for the NI ELLA. The virus dilutions that were needed to reach 50 % of total NA activity ranged from 2- to 2,000-fold. It should be noted that HA titres do not necessarily correlate with NA titre, and can therefore not be used as an indicator for antigen dilution.

For viruses that gave low total NA activity, the amount of N2 antigen within the linear trajectory but with a minimum NA activity of  $> 0.75$  ( $OD_{490}$ ) was selected. If the total NA activity was lower than  $0.75$  ( $OD_{490}$ ), viruses were concentrated and retested to ensure a robust measurement of NA activity inhibition.

The assay variability was assessed for NI ELLAs conducted on different plates and different days (data not shown). The results obtained on different days showed high correlation ( $R^2 = 0.88$ ) and the largest difference observed for individual samples was fourfold.

The titration of the viruses in the ELLA as well as the NI ELLA take 2 days to perform. Once titrated in the ELLA, the aliquoted viruses can be used in multiple NI ELLAs without the need for re-titration. A maximum of 20 plates (10 viruses measured in duplicate against eight antisera) or 60 plates (30 viruses measured in duplicate against eight antisera) can be performed by one or by two people, respectively.

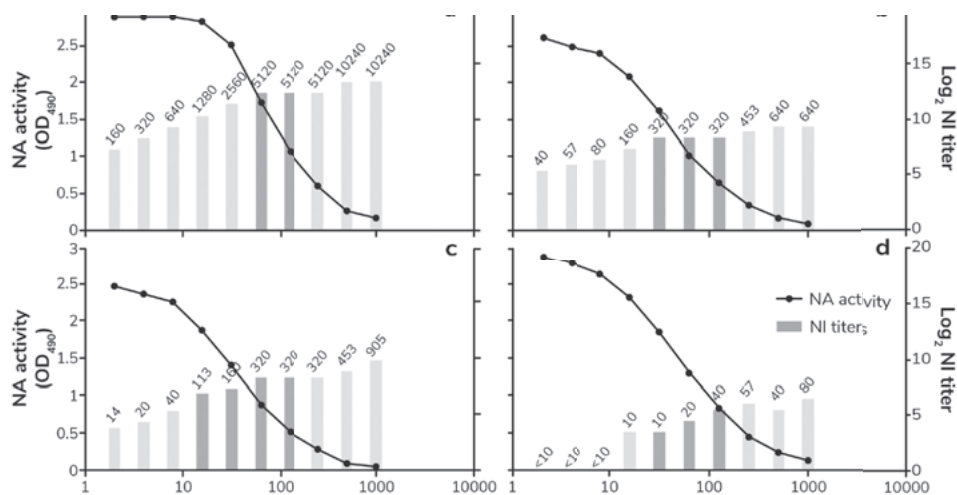


Figure 3 | **Determination of the optimal amount of antigen for routine antigenic characterization.** First, the NA activity (black circles) of serially diluted NL/95 (a, c and d) and NL/92 (b) virus was determined in the ELLA. Subsequently each of the dilutions of NL/95 and NL/92 antigens were tested against their N2-homologous (a and b, respectively) A(H7N2) ferret sera in the NI ELLA. In addition, NL/95 was tested against heterologous anti-NL/03 (c) and anti-NL/92 (d) A(H7N2) ferret sera. Absolute NI titres are shown as numbers above the log<sub>2</sub> NI titres (bars). The dark bars represent the dilutions that resulted in 25–75 % of the maximum NA activity. The assays were performed on different plates and on different days, resulting in some plate-to-plate or day-to-day variation.

### Specificity of the NI ELLA

To test the specificity of the NI ELLA<sup>272</sup> for antigenic characterization of wild type human A(H3N2) viruses, representative N2 antigens of the entire period of circulation of A(H2N2) and A(H3N2) viruses were selected. With these N2 antigens, recombinant A(H7N2) viruses with the internal genes of A/Netherlands/219/03 (H7N7) and the HA the A/Netherlands/33/03 (H7N7) were generated. The MBCS of the HA was removed to decrease pathogenicity. Monospecific ferret antisera elicited against these recombinant A(H7N2) viruses were obtained and tested against homologous and heterologous N2 antigens (Table 3). Homologous titres were highest for each antigen and ranged from 320 to > 10,240. Heterologous titres ranged from < 20 to > 10,240, and showed time dependence. For example, ferret anti-NL/57 did not recognize NL/09 virus (NI titre = < 20), but did recognize BI/68 virus (NI titre = 5,120); ferret anti-NL/09 did not recognize NL/57 virus (NI titre = < 20), but did recognize NL/03 virus (NI titre = 2,560). Overall, viruses with similar isolation dates were recognized by the same antisera. Viruses with



Table 3 | NI titres of A(H7N2) antisera against A(H3N2) viruses.

N2 anti-gens	Subtype	A(H7N2) antisera										
		NL/57	NL/68	BI/68	BI/72	BI/76	NL/82	NL/92	NL/95	NL/03	NL/09	
NL/57	H2N2	> <b>10,240</b>	2,560	> 10,240	40	80	< 20	20	< 20	< 20	< 20	
NL/68	H2N2	10,240	> <b>10,240</b>	> 10,240	40	80	< 20	160	320	< 20	< 20	
BI/68	H3N2	5,120	> 10,240	> <b>10,240</b>	160	160	20	80	320	40	< 20	
BI/72	H3N2	80	1,280	20	<b>5,120</b>	5,120	320	320	160	20	< 20	
BI/76	H3N2	40	640	20	2,560	<b>5,120</b>	160	160	80	< 20	< 20	
NL/82	H3N2	< 20	< 20	< 20	40	160	<b>320</b>	160	< 20	< 20	< 20	
NL/92	H3N2	< 20	< 20	< 20	40	80	320	<b>1,280</b>	320	< 20	< 20	
NL/95	H3N2	< 20	80	20	< 20	40	< 20	40	<b>2,560</b>	160	< 20	
NL/03	H3N2	< 20	40	20	20	< 20	< 20	20	160	<b>640</b>	< 20	
NL/09	H3N2	< 20	40	40	20	20	< 20	640	640	2,560	<b>5,120</b>	

The top row represents the various A(H7N2) antisera and the left column shows the N2 antigens, both ordered chronologically. N2-homologous NI titres are underscored and depicted in bold.

later isolation dates were not recognized by ferret sera specific for early viruses and vice versa.

## DISCUSSION

Monitoring functional antibody responses against NA, in tandem with HA-specific antibody analysis, could enhance the antigenic characterization of influenza viruses during routine surveillance for the purpose of influenza vaccine strain selection. The NI ELLA measures functional inhibition of NA activity by antibodies, and consequently, it has clear relevance to immunity *in vivo*<sup>130,143,144</sup>. In this study, the existing NI ELLA was optimized for rapid antigenic characterization of NA of wild type A(H3N2) viruses.

Similar to the TBA assay, the ELLA relies on the sialidase activity of NA. However, it has been demonstrated that the ELLA has a higher sensitivity compared to the TBA method<sup>257</sup>. The ELLA method, first described by Lambre et al., has been successfully applied in recent studies. Cate et al.<sup>246</sup> showed that NI antibody seroconversion rates are greater when vaccine dose is increased. Couch et al.<sup>255</sup> investigated the NA antibody response to current influenza vaccines and proved the importance of antibodies directed to the NA<sup>254</sup>. Couzens et al.<sup>273</sup> optimized the ELLA to measure influenza A virus NI titres in human sera.

NI assay results can be distorted by HA-specific antibodies interfering with substrate cleavage by NA<sup>137</sup>. To avoid this possible interference, various forms of NA have been used as antigens in NI assays: purified recombinant NA<sup>257</sup>, recombinant A(H6N1) and A(H6N2) viruses<sup>239,273</sup>, detergent split wild-type viruses<sup>246</sup>, and virus like particles<sup>254</sup>. However, generation of these antigens is time consuming and consequently unsuitable for large scale screening of viruses for vaccine strain selection. Therefore, the usage of wild type viruses was explored in the NI ELLA in combination with sera that were raised against HA-mismatched viruses.

Upon titration of naïve ferret sera against various influenza (sub)types, a non-specific inhibitory effect of the sera was occasionally observed. Non-specific inhibitors that interfere with the HI assay have been described previously, first by Hirst<sup>281</sup> and later by Francis<sup>282</sup>. Three classes of these non-specific inhibitors present in human and animal sera have been described:  $\alpha$ -,  $\beta$ -, and  $\gamma$ -class<sup>265,269,283</sup>. Inhibitors of the  $\alpha$ - and  $\gamma$ -classes express sialic acid residues that are specifically bound by the influenza virus HA<sup>265,284</sup> thereby blocking the receptor-binding sites of HA<sup>269</sup>. Inhibitors of the  $\beta$ -class<sup>265,285</sup> are, in contrast to inhibitors of the  $\alpha$ - and  $\gamma$ -classes<sup>286,287</sup>, thermolabile<sup>269</sup>. Sera used in the HI assay are treated beforehand with RDE to inactivate sialylated  $\alpha$ - and

$\gamma$ -inhibitors<sup>284,288</sup>, followed by incubation at 56 °C to inactivate the  $\beta$  inhibitors and the RDE<sup>265,289,290</sup>. The inhibitory effect of the naïve ferret sera seen in the NI ELLA may be explained by similar sialic acid expressing inhibitors that interact with the HA, the NA, or both, and thus preventing desialylation of fetuin. Since heat inactivation at 56 °C by itself did not affect the observed non-specific inhibitory effect, it indicates that the inhibitors were not  $\beta$ -class inhibitors<sup>289</sup> but rather  $\alpha$ - or, most likely,  $\gamma$ -inhibitors. RDE treatment of serum, followed by heat inactivation, eliminated most of the non-inhibitory effect without affecting specific virus-specific NI and HI titres.

Although the NI titres showed dependence on the amount of N2 antigen used, the NI titres did not differ when N2 antigen concentrations within the linear trajectory of the titration curve were chosen. These conditions resulted in limited variability between NI ELLAs performed on different days.

To be able to use wild type viruses as antigens, recombinant viruses with a mismatched HA must be used to generate antisera. The A(H7N7) virus was selected as a backbone to generate recombinant A(H7N2) viruses, as A(H7N7) replicates well in ferrets and antibodies directed to H7 do not recognize H3. However, in principle, any influenza A subtype (i.e. H9) can be selected as a backbone, providing that it replicates well in ferrets and antibodies directed to the HA do not recognize H3.

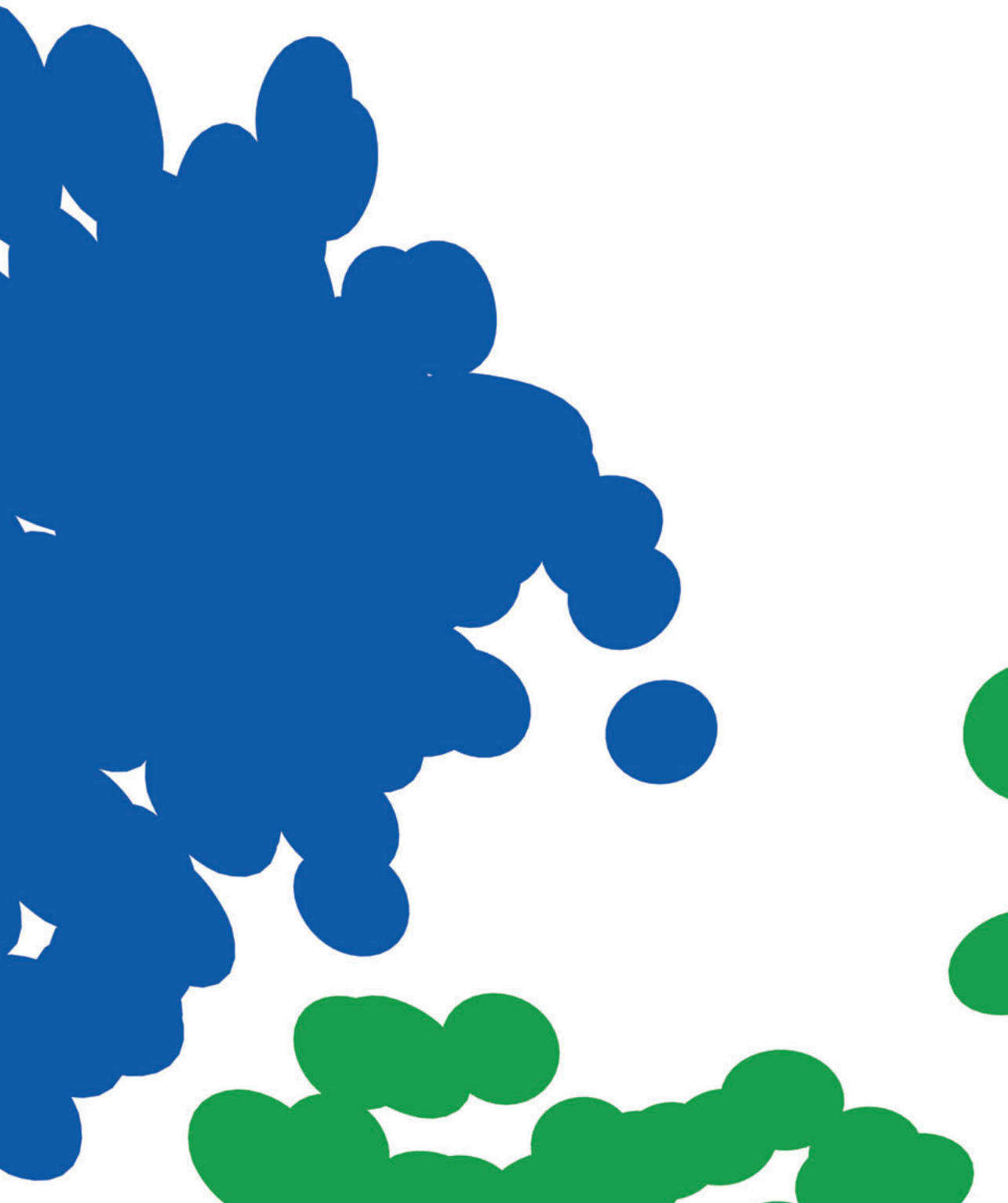
Using recombinant A(H7N2) viruses, ferret antisera yielding high NI titres without boosting or the usage of adjuvants could be generated. A panel of ferret anti-A(H7N2) sera were utilized in the NI ELLA to screen wild type A(H3N2) viruses. The observed intrasubtypic cross-reactivity of the ferret antisera permits screening of many viruses during influenza virus surveillance with a limited number of sera, while the specificity of the ferret antisera allows detection of drift variants. Although generation of the recombinant A(H7N2) viruses and anti-A(H7N2) ferret sera has to be done in a BSL-3 laboratory, sera can be used under lower biosafety conditions upon demonstration of the absence of virus and viral RNA. It is likely that this strategy will also be applicable for N1 antigens to screen circulating human A(H1N1) viruses. This approach would be more difficult for influenza B viruses given that there are no different subtypes. Perhaps an early influenza B virus (i.e. B/Lee/40) could be a suitable backbone if antibodies directed to its HA do not recognize the HA of currently circulating influenza B viruses. Alternatively one could use chimeric influenza A/B viruses, with the HA and internal genes of influenza A and a NA segment with the non-coding regions of influenza A and the ORF of influenza B<sup>291,292</sup>. But these viruses are possibly too attenuated to replicate efficiently in ferrets.

In conclusion, the NI ELLA described in this study provides a practical method to evaluate the antigenic properties of NAs from different viruses, and can be used for large-scale screening of wild type influenza A viruses. The antigenic drift of HA and NA was shown to be discordant<sup>159</sup>, and NA-specific antibodies—in addition to HA-specific antibodies—contribute to protection against disease. Therefore, information regarding antigenic drift of NA may facilitate selection of viruses that are antigenically-matched to circulating strains for influenza vaccine production.

### **ACKNOWLEDGEMENTS**

This work was supported by an NWO-VICI grant and NIH contract no. HHSN266200700010C. We gratefully thank Geert van Amerongen, Sander Herfst, Eefje Schrauwen, Oanh Vuong, Martin Linster and Ruud van Beek for excellent technical assistance.





## CHAPTER 8

# Summarizing discussion



## SUMMARIZING DISCUSSION

Influenza viruses are a significant cause of morbidity and mortality, with around three to five million cases of severe illness worldwide and around 290,000 to 650,000 deaths annually<sup>32</sup>. Influenza A H3N2 (A(H3N2)) and A H1N1pdm09 (A(H1N1)pdm09) viruses currently cocirculate with two lineages of influenza B virus (IBV) and cause seasonal epidemics. Of these, A(H3N2) virus causes the most severe disease<sup>156,345–347</sup> and have the highest rate of antigenic evolution<sup>348</sup>. New antigenic clusters of A(H3N2) viruses appear on average every 3.3 years<sup>107</sup>, enabling the virus to escape from herd immunity. To date, vaccination is still the most effective way to prevent influenza. Vaccine effectiveness (VE) is dependent on how well the vaccine strains match the circulating influenza strains, and estimates between 2004 and 2015 were lowest for A(H3N2) viruses; 33 % for matched viruses and 23 % for variant viruses<sup>349</sup>. A more accurate vaccine strain selection or prediction may increase the VE, but there is still much uncertainty about the underlying mechanisms that govern antigenic drift and drive the evolution of seasonal influenza viruses. However, significant progress has been made in recent years. Below, this progress will be discussed in the context of the work presented in this thesis.

### Genetic analysis of A(H3N2) viruses

*Genetic changes in the neuraminidase gene of influenza A(H3N2) viruses from 1968 to 2009 and its correspondence to haemagglutinin evolution*

The second major surface glycoprotein, neuraminidase (NA), uses its receptor-destroying activity to cleave terminal sialic acid from glycans on the host cell surface, thereby allowing the release of newly formed virus particles<sup>66–69</sup>. An association between the level of NA activity and the release of viral particles into the air has been reported<sup>350</sup> suggesting a link between NA activity and transmissibility<sup>350,351</sup>. This is likely related to the ability of NA to prevent aggregation of virus particles, since single particles are more easily transmitted via aerosols. It is well known that haemagglutinin (HA) is needed for viral entry, and NA for viral exit, but some evidence shows that NA might play a role in viral entry as well<sup>352</sup>. Recent studies demonstrated that NA also facilitates virus movement through mucus to allow the virus to reach their host cells<sup>353,354</sup>, and that influenza A virus (IAV) movement across the host cell surface was coordinated by both HA and NA<sup>355</sup>. These new findings highlight the importance of studying the evolution of NA.

In **chapter 2**, we studied the evolution of N2 between 1968 and 2009. Rates of nucleotide substitution were high and estimated to be  $3.15 \times 10^{-3}$  nucleotide substitutions



per site per year (nt subs/site/yr). Other studies found slightly higher rates which might be attributed to the smaller timespan of the study data sets compared to ours<sup>356,357</sup>. Rates of nucleotide substitution of N2 are generally higher compared to the NA of IBV<sup>357-359</sup>, which is expected since IBV exhibits lower rates of antigenic change<sup>299</sup>, possibly due to IBV having a less error-prone viral RNA-dependent RNA polymerase (RdRP)<sup>360</sup>, lower immunogenicity<sup>361</sup>, lower receptor binding avidity for sialic acid<sup>348,362</sup>, or fewer opportunities for the virus to evolve due to the smaller and less frequent epidemics<sup>363</sup>. In contrast to IBV, rates of N2 are often lower compared to the N1 of 2009 H1N1 pandemic (A(H1N1)pdm09) viruses<sup>357,364,365</sup>, which might be due to the relaxed selective constraints directly following zoonotic transmission at the time of the 2009 H1N1 pandemic emergence<sup>365</sup>.

The degree of natural selection acting on NA was estimated by looking at the mean number of amino acid-changing (non-synonymous or  $d_N$ ) and silent (synonymous or  $d_S$ ) substitutions per site ( $d_N/d_S$  ratio). The NA gene was under strong selection ( $d_N/d_S = 0.249$ ) indicating that there was an excess of nonsynonymous mutations that are eventually removed from the population by purifying selection. This is consistent with what others have recently found<sup>356,357,366,367</sup>. Similar  $d_N/d_S$  values have recently been reported for the NA of A(H1N1)pdm09 viruses and IBV<sup>357,358,365</sup>, suggesting that there was no difference in selection pressures on NA between contemporary seasonal viruses.

Early crystallographic studies of NA showed that antigenic regions surround the enzyme's highly conserved active site<sup>135,136</sup>. Antigenic sites A, B, and C<sup>134</sup> have been shown to be highly variable, most likely due to antigenic evolution<sup>122-124</sup>. To date, limited knowledge of the antigenic sites and epitopes that might be targeted by the immune system is available<sup>112,134,320,325,368-370</sup>. One way to investigate possible mutations that mediate escape from neutralizing antibodies is by looking at positive selection. We found a total of sixteen codons that were positively selected, of which most were located in the head region. Only six of these sites—199, 328, 334, 338, 367, and 370—are within known antigenic sites of NA<sup>134</sup> of which two—199 and 338—were recently found critical in an influenza A H9N2 (A(H9N2)) virus NA for the binding of mouse monoclonal antibodies (MAbs)<sup>316</sup>.

In 2004, Smith et al. studied the antigenic and genetic evolution of HA of human A(H3N2) virus from its introduction in humans in 1968 until 2003 revealing a punctuated pattern of eleven clusters each containing viruses that are antigenically similar<sup>107</sup>. The eleven antigenic clusters were named after the first vaccine strain in the cluster; Hong Kong 1968 (HK68), England 1972 (EN72), Victoria 1975 (VI75), Texas 1977 (TX77),

Bangkok 1979 (BK79), Sichuan 1987 (SI87), Beijing 1989 (BE89), Beijing 1992 (BE92), Wuhan 1995 (WU95), Sydney 1997 (SY97), and Fujian 2002 (FU02). In **chapter 2** we updated this data set with sequences until 2009 and compared the genetic changes to those of NA. The genetic distances between clades in the ML trees and in the genetic maps for NA and HA1 were often discordant, suggesting an asynchronous evolution of the two genes. Although variants on the trunk were the ancestors of all variants in future years, the HA1 phylogeny did have some variants that clustered away from the trunk in an evolutionary dead-end clade, most notably for viruses within the VI75 and the BE89 clusters. Compared to HA1, such dead-end clades were less obvious in the NA tree and in the genetic map of NA.

Reassortment of gene segments of IAVs facilitates genetic diversification, and has contributed to the emergence of novel seasonal and pandemic influenza strains<sup>39,169,371</sup>. Using phylogenetic methods such as tanglegrams and Graph-incompatibility-based Reassortment Finder (GiRaF), we were able to identify numerous reassortment events. Overall, our data indicated that, although mostly singular, reassortment events have occurred throughout the evolution of human influenza A(H3N2) viruses, especially during the periods of cocirculation of BE92- and WU95-like viruses. Most reassortment events were within antigenic clusters, rather than between antigenic clusters. Two reassortment events resulted in viruses that persisted in the human population for a longer period, but most reassortants did not persist, possibly suggesting that these reassortment events were neutral or detrimental.

The mean rate of nucleotide substitution for HA1 was higher compared to that of NA ( $5.15 \times 10^{-3}$  vs  $3.15 \times 10^{-3}$  nt subs/site/yr), but lower than what others have found for HA1<sup>356</sup>, most likely due to a smaller data set compared to ours.

HA1 was under strong selection ( $d_N/d_S = 0.362$ ) which is consistent with a study conducted in 2016<sup>356</sup>. For HA there was an abundance of negatively-selected sites and a limited number of positively-selected sites. *In silico* analyses showed a total of twenty-one positively-selected residues, mainly found in the antigenic sites of HA1<sup>53</sup>. Koel et al. showed that for HA of A(H3N2) viruses, all mutations that result in antigenic evolution occur at only seven positions, immediately adjacent to the receptor binding site (RBS); position 145, 155, 156, 158, 159, 189, and 193<sup>236</sup>. Koel et al. also detected two accessory substitutions: 133 and 135. Except for position 158, all positions were detected through these *in silico* analyses. It is possible that at least some of the other positively, or negatively, selected sites that we detected contributed to total viral fitness, as suggested by others<sup>372,373</sup>.

---

*Genomewide analysis of reassortment and evolution of human influenza A(H3N2) viruses circulating between 1968 and 2011*

In **chapter 3**, we extended the genetic analysis of **chapter 2** to the full genome and studied all eight gene segments and fifteen proteins. The mean rates of nucleotide substitution of the individual segments varied from  $2.07 \times 10^{-3}$  to  $3.99 \times 10^{-3}$  nt subs/site/yr, with the highest rates of nucleotide substitution for the major glycoproteins HA and NA, and the lowest rate for the matrix (M) segment. For A(H1N1)pdm09 viruses a similar trend was observed with respect to the M segment evolving at the lowest rate, and the HA and NA segments showing the highest rate, albeit with slightly higher mean rates of nucleotide substitution compared to A(H3N2)<sup>364,365</sup>. The HA1 subunit showed the highest rate of amino acid substitution, whereas the rate for the HA2 subunit was much lower ( $14.9 \times 10^{-3}$  compared to  $1.4 \times 10^{-3}$  amino acid substitutions per site per year. This was likely due to the fact that HA1 is the major immunogenic region of HA<sup>52,53</sup>. Basic polymerase 1-frame 2 (PB1-F2) and NA also displayed high rates of amino acid substitution. For all other proteins, rates of amino acid substitution were at least three times lower. Others who studied the A(H3N2) virus observed high rates of amino-acid substitutions for HA sites outside the epitope region<sup>374</sup>, and for genomic regions other than HA<sup>103</sup>, possibly signalling positive selective pressure by host immunity.

In order for reassortment to take place, more than one virus particle must enter a single host cell, followed by production of genome segments from each virus particle. It is known that influenza exhibits high levels of mixed infections in all major hosts<sup>201,375-377</sup>. Experimental systems showed that double infection of appropriate cells readily takes place<sup>378</sup>, that reassortment between two similar influenza viruses is efficient but also strongly dependent on dose and timing of the infections<sup>379</sup>, and also that NA can limit a second infection superimposed on an earlier one<sup>380</sup>. In any case, formation of viable infectious reassortants is dependent on the incorporation of one copy of each segment into a virus particle. As a consequence, suboptimal compatibility between vRNA packaging signals limit reassortment<sup>381</sup>. Thus, investigating reassortment with respect to all eight gene segments is a more natural approach compared to only looking at HA and NA. Our study on reassortment using tanglegrams showed abundant phylogenetic clustering differences suggestive of frequent reassortment over time. This was confirmed by GiRaF analyses, which revealed reassortment events particularly during the circulation of BE92-like and WU95-like viruses, which is in agreement with **chapter 2**. Several reassortment events persisted, mostly in more recent years, suggesting that these reassortment events aided in increasing genetic diversity of the virus to improve fitness or

to evade population immunity. However, most reassortment events were non-persistent, indicating that reassortment is rarely beneficial enough to reach fixation. It is important to note that IAV reassortment may be more prevalent in nature than one might expect based on the results of influenza surveillance studies<sup>379</sup>.

We further investigated positive and negative selection of all proteins using *in silico* tools. The membrane proteins HA, NA, M protein 2 (M2), and nonstructural protein 1 (NS1) showed relatively high overall  $d_N/d_S$  values compared to those of the other proteins. As expected, the HA1 subunit revealed a relatively higher overall  $d_N/d_S$  than that of HA2. Overall  $d_N/d_S$  values were higher for A(H1N1)pdm09 viruses, most likely due to adaptation to the new host after the introduction into humans<sup>365</sup>. Apart from the HA gene segment, overall  $d_N/d_S$  values were similar for IBV<sup>358</sup>, suggesting no difference between selection pressures on A(H3N2) virus and IBV. The higher overall  $d_N/d_S$  for H3 compared to the HA of IBV<sup>358,382</sup> is expected since IBV exhibits lower rates of antigenic evolution<sup>299</sup>.

All ORFs contained an abundance of negatively-selected codons. HA contained the largest number of positively-selected codons, most prominently located in the HA1 subunit, of which most are located in antigenic sites<sup>53</sup>. For NA, nineteen sites were positively selected, of which residue 199 and 338 were recently detected in the A(H9N2) virus NA of escape mutants<sup>316</sup>, and residues 267 and 401 were found to be part of antigenic epitopes in N1<sup>325,326</sup>. Both HA and NA are under antibody-mediated selection<sup>122-124,236,383,384</sup>, and at least some of the positively selected sites were likely due to immune evasion. It is likewise probable that some of the mutations, while not directly needed for evasion of the immune system, restore partial loss of viral fitness due to mutations causing antigenic change<sup>236</sup>. As in a study of A(H1N1)pdm09 viruses, we observed only few positively-selected sites in the polymerase genes compared to the abundance in HA<sup>365</sup>. Compared to A(H3N2), fewer positively-selected sites were detected in IBV, especially in the Yamagata lineage<sup>358</sup>, which is in line with the slower antigenic evolution<sup>299</sup> possibly due lower immune pressure.

In **chapter 3**, it became clear that viruses belonging to an antigenic cluster were not always seeded by the viruses from the antigenic cluster preceding it. Reassortment events during circulation of the HK68-like, VI75-like, BK79-like, BE92-like, WU95-like, SY97-like, FU02-like, CA04-like, and PE09-like viruses were identified by GiRaF analysis, although most did not persist over time. We investigated whether the timing of persistent reassortment events coincided with antigenic cluster transitions. From our GiRaF analysis, the only cluster transition that was linked to a persistent reassortment event was the SY97-to-FU02 antigenic cluster transition. A significant increase in transient amino

acid substitutions following reassortment primarily in the surface glycoprotein HA was observed by others, which was thought to affect virus fitness and directly influences antigenic variation<sup>385</sup>. For our data set, the direct contribution of reassortment to antigenic change seemed limited.

### **Antigenic analysis of NA**

Antibodies to NA contribute to protection against influenza virus infection and, similar to HA, antigenic drift has been observed for the NA. However, the NA is less studied, in part, due to a lack of suitable assays. Antigenic evolution of HA is mostly studied using haemagglutination inhibition (HI) titres<sup>236,386–390</sup>, and similarly, NA inhibition (NI) titres are measured to study the antigenic evolution of NA. Serology testing of NA using wild type virus is complicated, as antibodies that bind to the HA head domain exhibit strong NI activity due to steric hindrance<sup>239,324</sup>. To measure NA-specific inhibition, reassortant viruses with a non-matching HA are typically used to reduce the impact of anti-HA antibodies on the assays<sup>239</sup>.

#### *Discordant antigenic drift of neuraminidase and haemagglutinin H1N1 and H3N2 influenza viruses*

Traditionally, anti-NA assays have used the thiobarbituric acid (TBA) method. To reduce the use of highly hazardous chemicals, a mini-TBA assay has been established<sup>239</sup>. In **chapter 4** we have used this mini-TBA assay and the HI assay in combination with antigenic cartography to evaluate the antigenic evolution of HA and NA of the A(H1N1) and A(H3N2) viruses that have been used in vaccine formulations between 1995 and 2010.

For this study, we used panels of ferret antisera directed to wild type A(H1N1) and A(H3N2) virus and, as antigens, virus reassortants generated by reverse genetics to combine the targeted NA and a mismatched HA of the H6 subtype. There was minimal NI cross-reactivity between early NAs and antisera raised against recent seasonal strains of the A(H1N1) and A(H3N2) human lineages, demonstrating extensive antigenic evolution since introduction of these subtypes. As previously shown for HA, genetic changes in NA did not always lead to an antigenic change. The non-continuous pattern of NA evolution did not correspond closely with HA evolution in either subtype. This observation is in agreement with the reported independent evolution of the phenotype, with discordant patterns of antigenic evolution of HA and NA, as measured by serological assays<sup>122,137</sup>. The use of human sera showed that for the A(H1N1) and A(H3N2) components of

the 2006–2007 vaccine, human NA-specific antibodies reacted in a similar manner as monospecific ferret antisera, suggesting the antigenic evolution identified by analysis of ferret antisera is relevant to human serologic responses.

Using a small panel of sera from vaccinated individuals, it was noticed that current inactivated vaccines induced variable anti-NA responses, and did not reliably induce robust anti-NA immunity. This is in line with other studies<sup>255,391–394</sup>. It is well known that current inactivated influenza virus vaccines contain NA of variable quality and (non-standardized) quantity, potentially with lot-to-lot variability<sup>255,391–394</sup>, and this might influence the induction of anti-NA immunity. Several methods have been suggested to improve this, for example the addition of purified NA to, or alongside, regular inactivated vaccines as purified N2 was shown to be safe and immunogenic in humans<sup>321,395,396</sup>. The approach of adding recombinant NA<sup>397</sup> could also be used for trivalent inactivated vaccines (TIV) or to the recently approved recombinant HA vaccine 'Flublok'<sup>398,399</sup>. Clinical trials are needed to evaluate NA-only, NA/HA-only, NA-spiked inactivated vaccines, and newly developed vaccine formulations and strategies to assess which of these induce a more robust anti-NA immune response.

In this study we found that a single point mutation, E329K, in the NA of A/Brisbane/59/2007 was primarily responsible for the lack of inhibition by polyclonal antibodies specific for earlier strains. This residue has been put forward as an antigenic residue in earlier studies<sup>114,134–136,318,319</sup>. Our results demonstrate the impact of a single point mutation on NA antigenic evolution, and its consequence on recognition by human sera. This is also observed for the HA; seven of the ten cluster transitions of A(H3N2) viruses were caused by only a single amino acid substitution<sup>236</sup>.

#### *Enzyme-linked lectin assays to measure influenza A virus neuraminidase inhibition*

The miniaturized version of the TBA is still cumbersome and not suitable for routine serology. An enzyme-linked lectin assay (ELLA) described by Lambre et al.<sup>142</sup> is a practical alternative method, without hazardous chemicals, for measuring NI titres. First the antigen—whole virus or (purified) NA—is titrated with the ELLA. The ELLA method consists of coating plates with fetuin, a substrate for NA, and adding antigen. Active NA cleaves the terminal sialic acid residue from the fetuin, leaving an exposed galactose residue. Peanut agglutinin conjugated to horseradish peroxidase is then added, whereupon the horseradish peroxidase-conjugated peanut agglutinin lectin binds exposed terminal galactose. The intensity of the signal after addition of the substrate is dependent on the level of desialylation and thus NA activity. In the NI ELLA,

binding of NA by specific antibodies will inhibit the enzymatic function of NA and result in a reduction of desialylation and, hence, in a reduction of the signal. Results can be determined as 50 % endpoint titres. In **chapter 5** and **6**, we optimized the published ELLA method<sup>142</sup> for routine analysis of human sera, and rapid antigenic characterization of the NA using wild type viruses.

An optimized enzyme-linked lectin assay to measure influenza A virus neuraminidase inhibition antibody titres in human sera

Although the ELLA dates back to 1990, it was only more recently that laboratories began using the ELLA to measure NI antibody titres of clinical samples<sup>246,254,256,301,400</sup>. In **chapter 5**, we optimized the ELLA for routine analysis of human sera using reassortant viruses with a mismatched HA (H6) as antigen to avoid non-specific inhibition by H1 and H3-specific antibodies in human sera.

The optimized ELLA was subtype-specific, reproducible, had minimal plate-to-plate variability, and had operator-to-operator repeatability. A subsequent inter-laboratory study of ELLA variability showed that the assay had good reproducibility when performed in different laboratories, and that inclusion of a standard can further reduce variability in results<sup>339</sup>. While the titres measured by ELLA were somewhat greater than those measured by a mini-TBA method, seroconversion rates were the same, suggesting similarity in assay sensitivity under these optimized conditions. The ELLA described in **chapter 5** provides a practical format for routine evaluation of human antibody responses to NA. In addition, a protocol has now been published for routine evaluation of human antibody responses following influenza infection or vaccination<sup>401</sup>.

Optimization of an enzyme linked lectin assay suitable for rapid antigenic characterization of the neuraminidase of human influenza A(H3N2) viruses.

In **chapter 6** we provided details of the critical steps that are needed to modify and optimize the NI ELLA in a format that is reproducible, highly sensitive, and useful for large-scale influenza surveillance to monitor antigenic evolution of NA.

To prevent interference of HA-specific antibodies, most NI assays are performed with recombinant viruses containing a mismatched HA. However, generation of these viruses is time consuming and unsuitable for large-scale influenza surveillance. The feasibility of using the ELLA to evaluate the antigenic relatedness of NA of wild type A(H3N2) viruses was assessed. Rather than using recombinant viruses, wild type A(H3N2) viruses were used as antigen with ferret sera elicited against recombinant

viruses with a mismatched HA. These recombinant influenza A H7N2 (A(H7N2)) viruses contain the NA of A(H3N2) viruses and the HA and internal gene segments of an influenza A H7N7 (A(H7N7)) virus. Therefore, the HA-specific antibodies present in the ferret sera will not interfere with wild type A(H3N2) viruses in the NI assay. The A(H7N7) virus was selected for its ability to replicate efficiently in ferrets, thereby producing high levels of antibodies, but in principle other influenza virus subtypes can be used.

Some non-specific inhibition was observed when sera were tested in the ELLA even after heat inactivation. This inhibition was removed by treatment of the serum samples with a small amount of sialidase, such as Burnet's receptor-destroying enzyme (RDE), followed by heat inactivation to inactivate RDE activity without affecting the HI and NI titres of ferret antisera.

#### *Mapping the Antigenic Evolution of the N2 Neuraminidase of Human Influenza A Viruses from 1957 to 2012*

In **chapter 7** we have mapped the antigenic evolution of N2 from its introduction in humans in 1957 until 2012 using the NI ELLA described in **chapter 6**. We tested 285 A(H3N2) viruses and two A(H2N2) viruses in NI ELLA using 19 post-infection ferret antisera directed to A(H7N2) recombinants. The results showed clear antigenic evolution: viruses were recognised by antisera mostly in a time-dependent manner with high titres between homologous antigens and antisera, and low titres between early antisera and late antigens, and vice versa. Next, antigenic cartography was used to visualise the NI titres in an antigenic map, similar to what was done for HA. Using extensive testing it was determined that the NI data was best represented in a 2D map, and that this map had high resolution. The NA antigenic map showed that the antigens tend to group in clusters rather than a continuous antigenic lineage, although much less than for HA. Six antigenic clusters, determined by a k-means clustering algorithm, were observed for NA, and were named after the first vaccine strain of the cluster that had not already been used to name an HA antigenic map cluster; Netherlands 1957 (H2N2 virus), Port Chalmers 1973, Philippines 1982, Guizhou 1989, Nanchang 1995, and New York 2004. The NA antigenic map showed that a simple (non-intersecting) curve could be drawn through the cluster centroids in temporal order. On average, antigenic NA clusters continued to cocirculate longer than the average of 2 years of cocirculation of strains from different clusters described previously for HA<sup>107</sup>. This could be due to the fact that NA is immuno-subdominant to the HA<sup>263,297,402</sup> and thus encounters lower herd immunity.



By comparing sequences from the six antigenic clusters several cluster-difference substitutions were detected, but it is not yet known which of these mutations are responsible for antigenic change or cluster-transitions. For HA of A(H3N2) viruses, all cluster-transition substitutions occurred at only seven positions, immediately adjacent to the RBS<sup>236</sup>, although there were many more cluster-difference substitutions. Mutations at these seven positions also resulted in antigenic change for influenza A H5N1 (A(H5N1))<sup>386</sup> and A(H1N1)pdm09 viruses<sup>403</sup>. For NA, six of the 46 cluster-difference substitutions—153, 155, 197, 199, 221, and 370—were within 3 Å of any atom in the active site and were all surface exposed. Two of these six substitutions—199 and 370—were involved in multiple cluster transitions and were found to be positively selected (**chapter 2 and 3**). These findings indicate an important role for the described amino acid substitutions in the antigenic evolution of N2, however future studies using mutant viruses are needed to truly assess the effect of these mutations on the antigenic phenotype. The amino acid substitutions that were involved in the major antigenic changes for HA of A(H3N2)<sup>236</sup> and A(H5N1) viruses<sup>386</sup> all resulted in large changes of biophysical properties. It would be interesting to assess whether this also holds true for NA.

Previously, the antigenic evolution of A(H3N2) viruses from their introduction into humans in 1968 to 2003 was characterized using HI assays and antigenic cartography<sup>107</sup>. To be able to compare the mapped antigenic evolution of NA from 1968 to 2011, we updated the HA antigenic map from 2003 to 2011. This updated antigenic map revealed two additional new antigenic HA clusters, California 2004 and Perth 2009, adding to the eleven previously described clusters. For HA, the antigenic clusters remain dominant for on average 3.3 years<sup>107</sup>. This is remarkable if one considers the high mutation rate of influenza virus<sup>177,178</sup>, especially for HA<sup>95,97,179</sup>, and that single amino acid substitutions are often sufficient to cause antigenic cluster transitions in HA<sup>236</sup>. There are several hypotheses proposed to explain this paradox. Computational models demonstrate that antigenic evolution may come with a substantial fitness cost<sup>372,373,404</sup>, and this is supported by the observation that some mutant viruses with substitutions near the RBS of HA could not be rescued<sup>236</sup>. HA receptor binding avidity could limit the rate of antigenic evolution since many of the amino acid substitutions needed for adjusting receptor binding avidity for sialic acid also alter antigenicity<sup>362</sup>.

As described in previous studies using much smaller data sets<sup>122,137,159</sup>, our study showed that the timing of antigenic evolution of HA and NA is discordant. We show that when NA is evolving antigenically, HA does not and vice versa. This alternation was seen repeatedly during the circulation of A(H3N2) viruses from 1968 until 2011. The

discordance was also observed upon comparison of the time span of circulating viruses within the antigenic clusters of NA with those of HA. However, this pattern was almost absent when comparing the genetic changes of HA and NA. The opposite roles of HA and NA during influenza virus replication require functional balance between HA and NA<sup>171,172</sup>. The discordance in antigenic evolution between NA and HA, seen in **chapter 4** and **7**, might also affect the HA-NA balance<sup>405</sup>. Functional mismatches between HA and the NA are known to decrease viral fitness<sup>172,406</sup>. The data presented here might aid in the characterization and understanding of the functional balance of these two glycoproteins.

### **Future perspective**

*The use of antigenic cartography for visualizing bindings data of other pathogens*

Since its initial design in 2004<sup>107</sup>, antigenic cartography has become a core component of the World Health Organization (WHO) Global Influenza Surveillance and Response System (GISRS)<sup>303</sup>. Several studies of influenza viruses using the HI assay and antigenic cartography have been published since: swine A(H3N2) virus<sup>407</sup>, swine H1 virus<sup>390</sup>, avian H7 virus<sup>389</sup>, equine A(H3N8) virus<sup>388</sup>, A(H3N2) virus<sup>387</sup>, and avian H9 virus<sup>408</sup>. In **chapter 4** and **7**, we used antigenic cartography and NI titres to map the antigenic properties of the NA of seasonal influenza viruses. Antigenic cartography played a huge role in revealing the key positions involved in the antigenic evolution of A(H3N2)<sup>236</sup> and A(H5N1)<sup>386</sup> viruses. A combined approach of an antigenic map and a phylogenetic tree was used to investigate the dynamics of seasonal influenza viruses<sup>299</sup>. Antigenic cartography is not only restricted to influenza and HI or NI assays. Various other viruses and a bacterium have been antigenically characterized using neutralization assays and antigenic cartography<sup>409–414</sup>. The methods behind antigenic cartography functioned as a basis for a similar computational framework<sup>415–417</sup>, which has led to an array of recent publications<sup>358,365,418–426</sup>, not only limited to the influenza virus<sup>427–429</sup>. This shows the potential of antigenic cartography through the visualisation of complex data, thereby providing new insights in pathogen evolution.

### *Antigenic sites and epitopes of neuraminidase*

In 1972, Walter Dowdle advocated for a focus on both the HA and the NA during antigenic influenza surveillance and vaccine formulation<sup>430</sup>. During the last decade, more researches acknowledged this idea<sup>342,431–439</sup>, and an NA focus group, NAction!, was formed to promote research that helps to understand NA-based immunity and how it can contribute to the design of better influenza virus vaccines<sup>440</sup>. To date, there

is still limited data on the antigenic sites and epitopes of NA that might be targeted by the immune system<sup>112,134,320,325,368–370</sup>. We have now identified six antigenic clusters and revealed several cluster-difference mutations that can play a role in antigenic change of the NA. Based on these results, a similar approach as Koel et al. used to map the cluster defining mutations for HA<sup>236</sup> could be used for NA. To do this, a selection of representative viruses from each antigenic cluster should be made in which cluster-difference substitutions are introduced using reverse genetics. These viruses can then be tested in the NI ELLA to determine whether the substitution is responsible for the cluster transition. There are 46 cluster-difference substitutions (**chapter 7**), but of high interest are the residues near the active site of NA, as all key positions in HA were immediately adjacent to the RBS<sup>236</sup>. All key positions for antigenic change of HA (with the exception of position 158) found by Koel et al.<sup>236</sup> were detected through our *in silico* analyses used in **chapter 2** and **3**. In parallel, for NA it would be preferable to start with residues 199 and 370. These positions are within 3 Å of any atom in the active site, are surface exposed, and were found to be positively selected (**chapter 2** and **3**).

#### *Characterization of neuraminidase inhibition for other influenza subtypes*

For the HA of seasonal influenza viruses, H1 of pre-2009 H1N1 pandemic viruses and the HA of IBV displayed slower rates of antigenic evolution compared to the HA of A(H3N2) viruses<sup>348</sup>. It would be interesting to see whether N1 evolves more slowly antigenically compared to N2. Recently, H6NB viruses have been rescued<sup>441</sup>, creating the possibility of studying the antigenic evolution of NA of IBV. Once key positions responsible for antigenic change in N2 are determined, it will be interesting to see whether these also play a role in the antigenic evolution of other NA subtypes.

Usage of recombinant influenza viruses with an antigenically-mismatched avian HA is still considered the gold standard assay for measuring NI activity<sup>440</sup>. However, four things need to be kept in mind: 1) HA stalk-reactive antibodies might interfere with the assay<sup>324,441,442</sup>, 2) NI activities differed slightly between wild type viruses and H6Nx viruses using MAbs<sup>441</sup>, 3) the requirement of a permit to work with low pathogenic avian viruses, 4) emergence of H6Nx<sup>443</sup> or H7Nx viruses<sup>444,445</sup> in humans. The first could be overcome by including a negative control with a mismatched NA to determine the NI activity background caused by the HA stalk-reactive antibodies. The second should be investigated further to determine whether this is restricted to MAbs or that it also has impact on polyclonal sera. The third could be solved by sharing inactivated reassortant viruses between laboratories, and several sources of H6Nx viruses are now available

for distribution<sup>440</sup>. The final obstacle can be overcome by using purified recombinant NA<sup>257,446,447</sup> as antigen, which should be expressed as a tetramer with fully functional enzymatic activity to preserve the antigenic structure<sup>391,446,448,449</sup>. The downside of using soluble NA is that the NA might behave differently from a membrane-bound NA. In addition, HA is needed since it plays an important role in bringing NA and the substrate into close proximity<sup>450</sup>. Other options are NA-expressing vectors<sup>451,452</sup>, NA-only virus like particles<sup>327</sup>, and pseudotyped viruses devoid of any HA<sup>453</sup>. In the end, the ELLA using recombinant viruses with a mismatched HA is still recommended as a standard assay for measuring NI activity. Several HA subtypes have been reported to infect humans; H1-H3, H5-H7, H9, and H10<sup>20,37,443,454-458</sup> and this should be considered when selecting a “mismatched” HA for studying human antisera after infection or vaccination. Although previous studies using human sera showed similar results for NA antigenic evolution as observed for ferret antisera, it would be interesting to test human sera of young children after a primary influenza infection in the NI ELLA. Fonville et al. showed that human sera of young children after a primary influenza infection could recapitulate the antigenic evolution seen for HA while using ferret post-infection antisera<sup>459</sup>, while human post-vaccination antisera of individuals that encountered multiple influenza virus infections responded differently than ferret post-infection antisera to H3 viruses<sup>426</sup>.

### *Influenza vaccines*

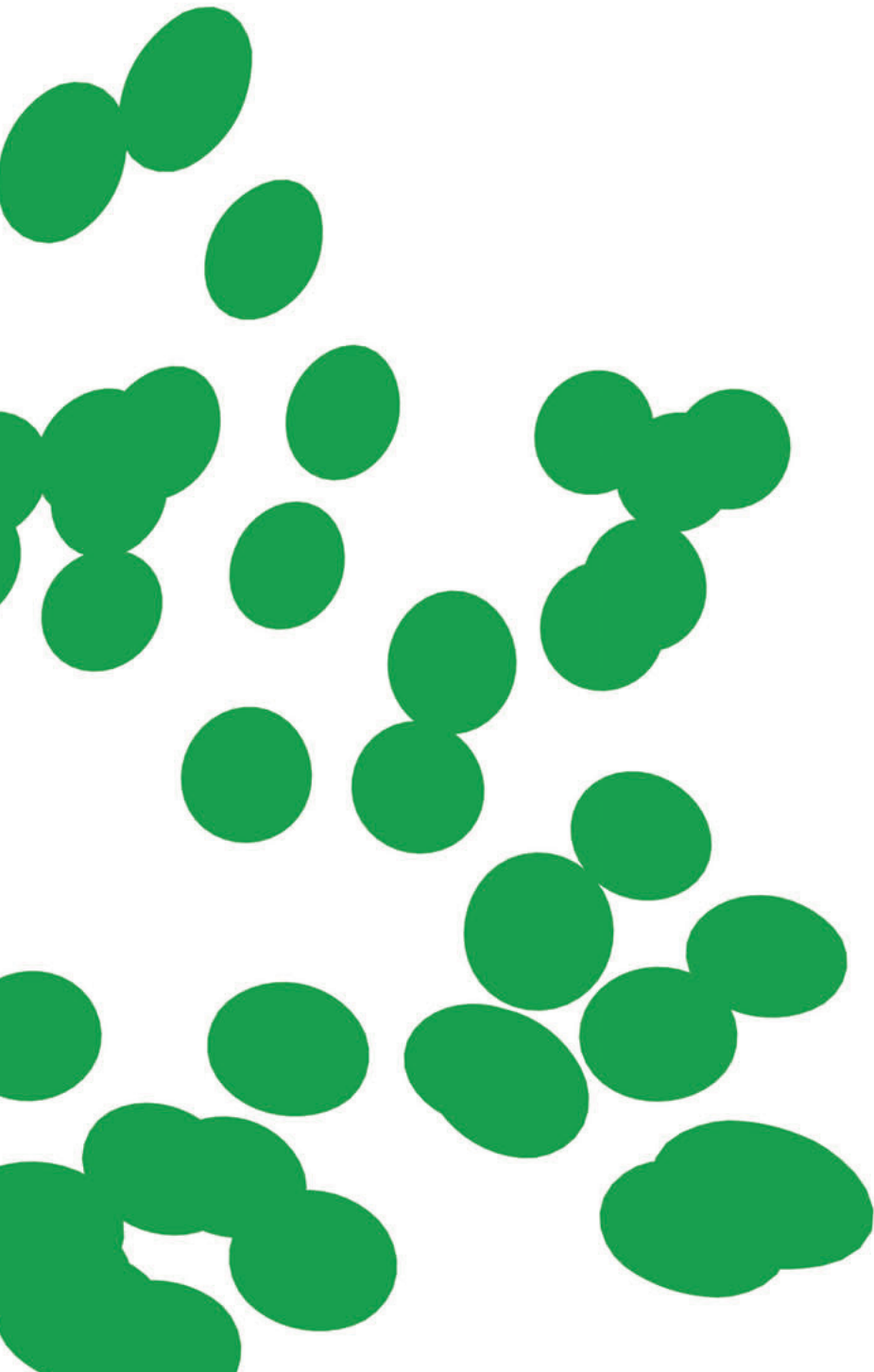
Influenza vaccination is still the most effective method of preventing influenza virus infection and its potentially severe complications. However, even in years when influenza vaccines are well matched to circulating viruses, estimates of VE range from 40 to 60 %<sup>460</sup>, which is lower than with other vaccines<sup>461-464</sup>. Antibodies are produced against both HA and NA in response to infection or vaccination. Although HA antibody levels have been shown to correlate with protection, it has been shown that NA is immunogenic and induces antibody responses in mice<sup>259,323-329,332,333,335,465</sup>, guinea pigs<sup>466</sup>, ferrets<sup>337,338,467</sup>, chickens<sup>336</sup>, and humans<sup>246,256,257,301,322</sup>, and may also correlate with protection as well as reduced severity of illness<sup>127,128,254,301,438,468</sup>. Influenza virus vaccines that are currently licensed include immunogenic quantities of NA, but the contribution of NA to vaccine efficacy is not known<sup>431</sup>, the type of NA is not matched with circulating viruses, and the amount of NA is not standardized. An accurate mass spectrometry-based method has been developed to quantify both NA and HA in influenza vaccines<sup>469</sup>. While the assay is not necessarily reflective of NA immunogenicity, it provides a way to measure the absolute concentration of each

---

NA type/subtype in seasonal and pandemic vaccines<sup>470</sup>. An assay to measure NA immunogenicity—a potency assay—that can differentiate between NA subtypes is a capture enzyme-linked immunosorbent assay to quantify the native form of NA in monovalent or multivalent vaccines<sup>470</sup>. Alternatively, the potency as a proxy for immunogenicity can be evaluated with the VaxArray NA assay<sup>471</sup>. Employing these methods to evaluate novel vaccine formulations and strategies are likely to improve the effectiveness of these vaccines.

### *Concluding remarks*

Our work clearly shows that there is antigenic evolution for NA, thus warranting the inclusion of NAs representing emerging influenza A strains in vaccines. However, to select these strains, current influenza surveillance strategies should be adapted. Increasing knowledge on which mutations cause changes in the phenotype can help to perform more targeted influenza surveillance. This has been done for mutations that render A(H5N1) virus transmissible<sup>472</sup>, and for studies on antiviral resistance<sup>473–475</sup>. Influenza surveillance where sequence data is integrated with epidemiological, geographical, and/or antigenic data<sup>373,476,477</sup> can help to gain insights into the evolution and epidemiology of influenza viruses, and ultimately to help prediction of influenza virus evolution. For example, real-time tracking of influenza virus evolution with NextFlu<sup>478</sup>, also as a web-application <https://nextstrain.org>, was used to analyse seasonal influenza circulation patterns and provided projections for the 2017–2018 flu season<sup>479</sup>. Once the key residues responsible for antigenic change in NA have been revealed, it would be advisable to integrate genetic and antigenic NA data in these influenza surveillance applications. Including NA sequencing and measurement of NI antibody titres into routine influenza surveillance will facilitate consideration of NA content, and improve next generation influenza vaccines.



## CHAPTER 9

# References

## REFERENCES

1. Shope, R. E. Swine influenza: III. Filtration experiments and etiology. *J. Exp. Med.* **54**, 373–385 (1931).
2. Smith, W., Andrewes, C. H., Laidlaw, P. P. & Others. A virus obtained from influenza patients. *Lancet* 66–68 (1933).
3. Lupiani, B. & Reddy, S. M. The history of avian influenza. *Comp. Immunol. Microbiol. Infect. Dis.* **32**, 311–323 (2009).
4. Wright, P. F., Neumann, G. & Kawaoka, Y. Orthomyxoviruses. *Fields Virology*, eds Knipe DM, et al. (2013).
5. King, A. M. Q. et al. Changes to taxonomy and the International Code of Virus Classification and Nomenclature ratified by the International Committee on Taxonomy of Viruses (2018). *Arch. Virol.* (2018). doi:10.1007/s00705-018-3847-1
6. Shaw, M. L. & Palese, P. Orthomyxoviridae. *Fields virology* **1**, 1151–1185 (2013).
7. Su, S., Fu, X., Li, G., Kerlin, F. & Veit, M. Novel Influenza D virus: Epidemiology, pathology, evolution and biological characteristics. *Virulence* **8**, 1580–1591 (2017).
8. Labuda, M. & Nuttall, P. A. Tick-borne viruses. *Parasitology* **129 Suppl**, S221–45 (2004).
9. Hubálek, Z. & Rudolf, I. Tick-borne viruses in Europe. *Parasitol. Res.* **111**, 9–36 (2012).
10. Presti, R. M. et al. Quarantfil, Johnston Atoll, and Lake Chad viruses are novel members of the family Orthomyxoviridae. *J. Virol.* **83**, 11599–11606 (2009).
11. Shi, M. et al. The evolutionary history of vertebrate RNA viruses. *Nature* **556**, 197–202 (2018).
12. Chanock, R. H. et al. A revised system of influenza virus nomenclature. *Virology* **47**, 854–856 (1972).
13. Tong, S. et al. New world bats harbor diverse influenza A viruses. *PLoS Pathog.* **9**, e1003657 (2013).
14. Tong, S. et al. A distinct lineage of influenza A virus from bats. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 4269–4274 (2012).



15. Fouchier, R. A. M. et al. Characterization of a novel influenza A virus haemagglutinin subtype (H16) obtained from black-headed gulls. *J. Virol.* **79**, 2814–2822 (2005).
16. Li, Q. et al. Structural and functional characterization of neuraminidase-like molecule N10 derived from bat influenza A virus. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 18897–18902 (2012).
17. Schild, G. C., Newman, R. W., Webster, R. G., Major, D. & Hinshaw, V. S. Antigenic analysis of influenza A virus surface antigens: considerations for the nomenclature of influenza virus. Brief review. *Arch. Virol.* **63**, 171–184 (1980).
18. Ciminski, K., Thamamongood, T., Zimmer, G. & Schwemmler, M. Novel insights into bat influenza A viruses. *J. Gen. Virol.* **98**, 2393–2400 (2017).
19. World Health Organization. A revision of the system of nomenclature for influenza viruses: a WHO memorandum. *Bull. World Health Organ.* **58**, 585–591 (1980).
20. World Health Organization. Cumulative number of confirmed human cases for avian influenza A(H5N1) reported to WHO, 2003-2018. *World Health Organization (WHO)* (2018). Available at: [http://www.who.int/influenza/human\\_animal\\_interface/H5N1\\_cumulative\\_table\\_archives/en/](http://www.who.int/influenza/human_animal_interface/H5N1_cumulative_table_archives/en/). (Accessed: 9th August 2018)
21. de Jong, J. C., Claas, E. C., Osterhaus, A. D., Webster, R. G. & Lim, W. L. A pandemic warning? *Nature* **389**, 554 (1997).
22. Ogata, T. et al. Human H5N2 avian influenza infection in Japan and the factors associated with high H5N2-neutralizing antibody titer. *J. Epidemiol.* **18**, 160–166 (2008).
23. Ostrowsky, B. et al. Low pathogenic avian influenza A (H7N2) virus infection in immunocompromised adult, New York, USA, 2003. *Emerg. Infect. Dis.* **18**, 1128–1131 (2012).
24. Editorial team Collective. Avian influenza A/(H7N2) outbreak in the United Kingdom. *Euro Surveill.* **12**, E070531.2 (2007).
25. Tweed, S. A. et al. Human illness from avian influenza H7N3, British Columbia. *Emerg. Infect. Dis.* **10**, 2196–2199 (2004).

26. Kurtz, J., Manvell, R. J. & Banks, J. Avian influenza virus isolated from a woman with conjunctivitis. *Lancet* **348**, 901–902 (1996).
27. Fouchier, R. A. M. et al. Avian influenza A virus (H7N7) associated with human conjunctivitis and a fatal case of acute respiratory distress syndrome. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 1356–1361 (2004).
28. Peiris, M. et al. Human infection with influenza H9N2. *Lancet* **354**, 916–917 (1999).
29. Butt, K. M. et al. Human infection with an avian H9N2 influenza A virus in Hong Kong in 2003. *J. Clin. Microbiol.* **43**, 5760–5767 (2005).
30. Cheng, V. C. C. et al. Infection of immunocompromised patients by avian H9N2 influenza A virus. *J. Infect.* **62**, 394–399 (2011).
31. Chen, M.-J. et al. Genetic and phylogenetic analysis of multi-continent human influenza A(H1N2) reassortant viruses isolated in 2001 through 2003. *Virus Res.* **122**, 200–205 (2006).
32. World Health Organization. Fact Sheet - Influenza (Seasonal). World Health Organization (2018). Available at: [http://www.who.int/news-room/fact-sheets/detail/influenza-\(seasonal\)](http://www.who.int/news-room/fact-sheets/detail/influenza-(seasonal)). (Accessed: 26th May 2018)
33. van Genugten, M. L., Heijnen, M. L. & Jager, J. C. Scenario analysis of the expected number of hospitalisations and deaths due to pandemic influenza in the Netherlands, RIVM report 282701002. (2002).
34. Sprenger, M. J., Mulder, P. G., Beyer, W. E., Van Strik, R. & Masurel, N. Impact of influenza on mortality in relation to age and underlying disease, 1967-1989. *Int. J. Epidemiol.* **22**, 334–340 (1993).
35. World Health Organization. WHO global influenza preparedness plan: the role of WHO and recommendations for national measures before and during pandemics. World Health Organization (2005). Available at: [http://www.who.int/csr/resources/publications/influenza/WHO\\_CDS\\_CSR\\_GIP\\_2005\\_5.pdf](http://www.who.int/csr/resources/publications/influenza/WHO_CDS_CSR_GIP_2005_5.pdf). (Accessed: 9th August 2018)
36. Fauci, A. S. Pandemic influenza threat and preparedness. *Emerg. Infect. Dis.* **12**, 73–77 (2006).
37. Centers for Disease Control and Prevention. Past Pandemics. Centers for Disease Control and Prevention (2017). Available at: <https://www.cdc.gov/flu/pandemic-resources/basics/past-pandemics.html>. (Accessed: 23rd March 2018)

38. Lindstrom, S. E., Cox, N. J. & Klimov, A. Genetic analysis of human H2N2 and early H3N2 influenza viruses, 1957-1972: evidence for genetic divergence and multiple reassortment events. *Virology* **328**, 101–119 (2004).
39. Smith, G. J. D. et al. Dating the emergence of pandemic influenza viruses. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 11709–11712 (2009).
40. Stöhr, K. Influenza—WHO cares. *Lancet Infect. Dis.* (2002).
41. Nakajima, K., Desselberger, U. & Palese, P. Recent human influenza A (H1N1) viruses are closely related genetically to strains isolated in 1950. *Nature* **274**, 334–339 (1978).
42. Wise, H. M. et al. A complicated message: Identification of a novel PB1-related protein translated from influenza A virus segment 2 mRNA. *J. Virol.* **83**, 8021–8031 (2009).
43. Chen, W. et al. A novel influenza A virus mitochondrial protein that induces cell death. *Nat. Med.* **7**, 1306–1312 (2001).
44. Jagger, B. W. et al. An overlapping protein-coding region in influenza A virus segment 3 modulates the host response. *Science* **337**, 199–204 (2012).
45. Muramoto, Y., Noda, T., Kawakami, E., Akkina, R. & Kawaoka, Y. Identification of novel influenza A virus proteins translated from PA mRNA. *J. Virol.* **87**, 2455–2462 (2013).
46. O'Neill, R. E., Talon, J. & Palese, P. The influenza virus NEP (NS2 protein) mediates the nuclear export of viral ribonucleoproteins. *EMBO J.* **17**, 288–296 (1998).
47. Karlsson Hedestam, G. B. et al. The challenges of eliciting neutralizing antibodies to HIV-1 and to influenza virus. *Nat. Rev. Microbiol.* **6**, 143 (2008).
48. Vasin, A. V. et al. Molecular mechanisms enhancing the proteome of influenza A viruses: an overview of recently discovered proteins. *Virus Res.* **185**, 53–63 (2014).
49. Skehel, J. J. & Waterfield, M. D. Studies on the primary structure of the influenza virus haemagglutinin. *Proc. Natl. Acad. Sci. U. S. A.* **72**, 93–97 (1975).
50. Lazarowitz, S. G. & Choppin, P. W. Enhancement of the infectivity of influenza A and B viruses by proteolytic cleavage of the haemagglutinin polypeptide. *Virology* **68**, 440–454 (1975).
51. Klenk, H. D., Rott, R., Orlich, M. & Blödorn, J. Activation of influenza A viruses by trypsin treatment. *Virology* **68**, 426–439 (1975).

52. Wilson, I. A. & Cox, N. J. Structural basis of immune recognition of influenza virus haemagglutinin. *Annu. Rev. Immunol.* **8**, 737–771 (1990).
53. Wiley, D. C., Wilson, I. A. & Skehel, J. J. Structural identification of the antibody-binding sites of Hong Kong influenza haemagglutinin and their involvement in antigenic variation. *Nature* **289**, 373–378 (1981).
54. Cross, K. J., Langley, W. A., Russell, R. J., Skehel, J. J. & Steinhauer, D. A. Composition and functions of the influenza fusion peptide. *Protein Pept. Lett.* **16**, 766–778 (2009).
55. Nicholson, K. G., Wood, J. M. & Zambon, M. Influenza. *Lancet* **362**, 1733–1745 (2003).
56. Hampson, A. W. & Mackenzie, J. S. The influenza viruses. *Med. J. Aust.* **185**, S39–43 (2006).
57. Hilleman, M. R. Realities and enigmas of human viral influenza: pathogenesis, epidemiology and control. *Vaccine* **20**, 3068–3087 (2002).
58. Sauter, N. K. et al. Hemagglutinins from two influenza virus variants bind to sialic acid derivatives with millimolar dissociation constants: a 500-MHz proton nuclear magnetic resonance study. *Biochemistry* **28**, 8388–8396 (1989).
59. Pinto, L. H. & Lamb, R. A. The M2 proton channels of influenza A and B viruses. *J. Biol. Chem.* **281**, 8997–9000 (2006).
60. Bullough, P. A., Hughson, F. M., Skehel, J. J. & Wiley, D. C. Structure of influenza haemagglutinin at the pH of membrane fusion. *Nature* **371**, 37–43 (1994).
61. Wu, W. W. H., Sun, Y.-H. B. & Panté, N. Nuclear import of influenza A viral ribonucleoprotein complexes is mediated by two nuclear localization sequences on viral nucleoprotein. *Virology* **4**, 49 (2007).
62. Fechter, P. & Brownlee, G. G. Recognition of mRNA cap structures by viral and cellular proteins. *J. Gen. Virol.* **86**, 1239–1249 (2005).
63. Akarsu, H. et al. Crystal structure of the M1 protein-binding domain of the influenza A virus nuclear export protein (NEP/NS2). *EMBO J.* **22**, 4646–4655 (2003).
64. Neumann, G., Hughes, M. T. & Kawaoka, Y. Influenza A virus NS2 protein mediates vRNP nuclear export through NES-independent interaction with hCRM1. *EMBO J.* **19**, 6751–6758 (2000).

65. Rossman, J. S. & Lamb, R. A. Influenza virus assembly and budding. *Virology* **411**, 229–236 (2011).
66. Liu, C., Eichelberger, M. C., Compans, R. W. & Air, G. M. Influenza type A virus neuraminidase does not play a role in viral entry, replication, assembly, or budding. *J. Virol.* **69**, 1099–1106 (1995).
67. Palese, P. & Compans, R. W. Inhibition of influenza virus replication in tissue culture by 2-deoxy-2,3-dehydro-N-trifluoroacetylneuraminic acid (FANA): mechanism of action. *J. Gen. Virol.* **33**, 159–163 (1976).
68. Palese, P., Tobita, K., Ueda, M. & Compans, R. W. Characterization of temperature sensitive influenza virus mutants defective in neuraminidase. *Virology* **61**, 397–410 (1974).
69. Gottschalk, A. Neuraminidase: the specific enzyme of influenza virus and *Vibrio cholerae*. *Biochim. Biophys. Acta* **23**, 645–646 (1957).
70. von Itzstein, M. The war against influenza: discovery and development of sialidase inhibitors. *Nat. Rev. Drug Discov.* **6**, 967–974 (2007).
71. Fujii, Y., Goto, H., Watanabe, T., Yoshida, T. & Kawaoka, Y. Selective incorporation of influenza virus RNA segments into virions. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 2002–2007 (2003).
72. Marsh, G. A., Hatami, R. & Palese, P. Specific residues of the influenza A virus haemagglutinin viral RNA are important for efficient packaging into budding virions. *J. Virol.* **81**, 9727–9736 (2007).
73. Harris, A. et al. Influenza virus pleiomorphy characterized by cryoelectron tomography. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 19123–19127 (2006).
74. Noda, T. et al. Architecture of ribonucleoprotein complexes in influenza A virus particles. *Nature* **439**, 490–492 (2006).
75. Yamaguchi, M., Danev, R., Nishiyama, K., Sugawara, K. & Nagayama, K. Zernike phase contrast electron microscopy of ice-embedded influenza A virus. *J. Struct. Biol.* **162**, 271–276 (2008).
76. Watanabe, T., Watanabe, S., Noda, T., Fujii, Y. & Kawaoka, Y. Exploitation of nucleic acid packaging signals to generate a novel influenza virus-based vector stably expressing two foreign genes. *J. Virol.* **77**, 10575–10583 (2003).
77. Ozawa, M. et al. Contributions of two nuclear localization signals of influenza A virus nucleoprotein to viral replication. *J. Virol.* **81**, 30–41 (2007).

78. Liang, Y., Hong, Y. & Parslow, T. G. cis-Acting packaging signals in the influenza virus PB1, PB2, and PA genomic RNA segments. *J. Virol.* **79**, 10348–10355 (2005).
79. Liang, Y., Huang, T., Ly, H., Parslow, T. G. & Liang, Y. Mutational analyses of packaging signals in influenza virus PA, PB1, and PB2 genomic RNA segments. *J. Virol.* **82**, 229–236 (2008).
80. Hutchinson, E. C., Curran, M. D., Read, E. K., Gog, J. R. & Digard, P. Mutational analysis of cis-acting RNA signals in segment 7 of influenza A virus. *J. Virol.* **82**, 11869–11879 (2008).
81. Gog, J. R. et al. Codon conservation in the influenza A virus genome defines RNA packaging signals. *Nucleic Acids Res.* **35**, 1897–1907 (2007).
82. Muramoto, Y. et al. Hierarchy among viral RNA (vRNA) segments in their role in vRNA incorporation into influenza A virions. *J. Virol.* **80**, 2318–2325 (2006).
83. de Wit, E., Spronken, M. I. J., Rimmelzwaan, G. F., Osterhaus, A. D. M. E. & Fouchier, R. A. M. Evidence for specific packaging of the influenza A virus genome from conditionally defective virus particles lacking a polymerase gene. *Vaccine* **24**, 6647–6650 (2006).
84. Scholtissek, C., Rohde, W., Von Hoyningen, V. & Rott, R. On the origin of the human influenza virus subtypes H2N2 and H3N2. *Virology* **87**, 13–20 (1978).
85. Guo, Y. J., Xu, X. Y. & Cox, N. J. Human influenza A (H1N2) viruses isolated from China. *J. Gen. Virol.* **73** ( Pt 2), 383–387 (1992).
86. Gregory, V. et al. Emergence of influenza A H1N2 reassortant viruses in the human population during 2001. *Virology* **300**, 1–7 (2002).
87. Nishikawa, F. & Sugiyama, T. Direct isolation of H1N2 recombinant virus from a throat swab of a patient simultaneously infected with H1N1 and H3N2 influenza A viruses. *J. Clin. Microbiol.* **18**, 425–427 (1983).
88. Kendal, A. P. et al. Laboratory-based surveillance of influenza virus in the united states during the winter of 1977--1978: II. I isolation of a mixture of a/victoria-and a/ussr-like viruses from a single person during an epidemic in Wyoming, USA, January 1978. *Am. J. Epidemiol.* **110**, 462–468 (1979).
89. Ellis, J. S. et al. Influenza AH1N2 viruses, United Kingdom, 2001-02 influenza season. *Emerg. Infect. Dis.* **9**, 304–310 (2003).

90. Barr, I. G. et al. Reassortants in recent human influenza A and B isolates from South East Asia and Oceania. *Virus Res.* **98**, 35–44 (2003).
91. Bean, W. J., Jr, Cox, N. J. & Kendal, A. P. Recombination of human influenza A viruses in nature. *Nature* **284**, 638–640 (1980).
92. Yamane, N., Arikawa, J., Odagiri, T., Sukeno, N. & Ishida, N. Isolation of three different influenza A viruses from an individual after probable double infection with H3N2 and H1N1 viruses. *Jpn. J. Med. Sci. Biol.* **31**, 431–434 (1978).
93. Xu, X. et al. Intercontinental circulation of human influenza A(H1N2) reassortant viruses during the 2001-2002 influenza season. *J. Infect. Dis.* **186**, 1490–1493 (2002).
94. Garten, R. J. et al. Antigenic and genetic characteristics of swine-origin 2009 A(H1N1) influenza viruses circulating in humans. *Science* **325**, 197–201 (2009).
95. Rambaut, A. et al. The genomic and epidemiological dynamics of human influenza A virus. *Nature* **453**, 615–619 (2008).
96. Holmes, E. C. et al. Whole-genome analysis of human influenza A virus reveals multiple persistent lineages and reassortment among recent H3N2 viruses. *PLoS Biol.* **3**, e300 (2005).
97. Nelson, M. I. et al. Stochastic processes are key determinants of short-term evolution in influenza a virus. *PLoS Pathog.* **2**, e125 (2006).
98. Schweiger, B., Bruns, L. & Meixenberger, K. Reassortment between human A(H3N2) viruses is an important evolutionary mechanism. *Vaccine* **24**, 6683–6690 (2006).
99. Wolf, Y. I., Viboud, C., Holmes, E. C., Koonin, E. V. & Lipman, D. J. Long intervals of stasis punctuated by bursts of positive selection in the seasonal evolution of influenza A virus. *Biol. Direct* **1**, 34 (2006).
100. Nelson, M. I., Simonsen, L., Viboud, C., Miller, M. A. & Holmes, E. C. Phylogenetic analysis reveals the global migration of seasonal influenza A viruses. *PLoS Pathog.* **3**, 1220–1228 (2007).
101. Nelson, M. I. et al. Multiple reassortment events in the evolutionary history of H1N1 influenza A virus since 1918. *PLoS Pathog.* **4**, e1000012 (2008).
102. Simonsen, L. et al. The genesis and spread of reassortment human influenza A/H3N2 viruses conferring adamantane resistance. *Mol. Biol. Evol.* **24**, 1811–1820 (2007).

103. Suzuki, Y. Natural selection on the influenza virus genome. *Mol. Biol. Evol.* **23**, 1902–1911 (2006).
104. Rimmelzwaan, G. F. et al. Sequence variation in the influenza A virus nucleoprotein associated with escape from cytotoxic T lymphocytes. *Virus Res.* **103**, 97–100 (2004).
105. Schild, G. C., Oxford, J. S. & Dowdle, W. R. Antigenic variation in current influenza A viruses: evidence for a high frequency of antigenic 'drift' for the Hong Kong virus. *Bull. World Health Organ.* (1974).
106. Hirst, G. K. The agglutination of red cells by allantoic fluid of chick embryos infected with influenza virus.. *Science* **94**, 22–23 (1941).
107. Smith, D. J. et al. Mapping the antigenic and genetic evolution of influenza virus. *Science* **305**, 371–376 (2004).
108. 108.Gottschalk, A. Neuraminic acid; the functional group of some biologically active mucoproteins. *Yale J. Biol. Med.* **28**, 525–537 (1956).
109. Laver, W. G. STRUCTURAL STUDIES ON THE PROTEIN SUBUNITS FROM THREE STRAINS OF INFLUENZA VIRUS. *J. Mol. Biol.* **9**, 109–124 (1964).
110. Bossart, P. J., Babu, Y. S., Cook, W. J., Air, G. M. & Laver, W. G. Crystallization and preliminary X-ray analyses of two neuraminidases from influenza B virus strains B/Hong Kong/8/73 and B/Lee/40. *J. Biol. Chem.* **263**, 6421–6423 (1988).
111. Laver, W. G. Crystallization and peptide maps of neuraminidase 'heads' from H2N2 and H3N2 influenza virus strains. *Virology* **86**, 78–87 (1978).
112. Colman, P. M. et al. Three-dimensional structures of influenza virus neuraminidase-antibody complexes. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **323**, 511–518 (1989).
113. Laver, W. G. et al. Crystallization and preliminary X-ray analysis of type B influenza virus neuraminidase complexed with antibody Fab fragments. *Virology* **167**, 621–624 (1988).
114. Laver, W. G., Webster, R. G. & Colman, P. M. Crystals of antibodies complexed with influenza virus neuraminidase show isosteric binding of antibody to wild-type and variant antigens. *Virology* **156**, 181–184 (1987).
115. Varghese, J. N., McKimm-Breschkin, J. L., Caldwell, J. B., Kortt, A. A. & Colman, P. M. The structure of the complex between influenza virus neuraminidase and sialic acid, the viral receptor. *Proteins* **14**, 327–332 (1992).



116. Tulip, W. R. et al. Refined atomic structures of N9 subtype influenza virus neuraminidase and escape mutants. *J. Mol. Biol.* **221**, 487–497 (1991).
117. Bossart-Whitaker, P. et al. Three-dimensional structure of influenza A N9 neuraminidase and its complex with the inhibitor 2-deoxy 2,3-dehydro-N-acetyl neuraminic acid. *J. Mol. Biol.* **232**, 1069–1083 (1993).
118. Burmeister, W. P., Ruigrok, R. W. & Cusack, S. The 2.2 Å resolution crystal structure of influenza B neuraminidase and its complex with sialic acid. *EMBO J.* **11**, 49–56 (1992).
119. Colman, P. M. Influenza virus neuraminidase: structure, antibodies, and inhibitors. *Protein Sci.* **3**, 1687–1696 (1994).
120. Bucher, D. J. & Kilbourne, E. D. A 2 (N2) neuraminidase of the X-7 influenza virus recombinant: determination of molecular size and subunit composition of the active unit. *J. Virol.* **10**, 60–66 (1972).
121. Paterson, R. G. & Lamb, R. A. Conversion of a class II integral membrane protein into a soluble and efficiently secreted protein: multiple intracellular and extracellular oligomeric and conformational forms. *J. Cell Biol.* **110**, 999–1011 (1990).
122. Kilbourne, E. D., Johansson, B. E. & Grajower, B. Independent and disparate evolution in nature of influenza A virus haemagglutinin and neuraminidase glycoproteins. *Proc. Natl. Acad. Sci. U. S. A.* **87**, 786–790 (1990).
123. Laver, W. G., Air, G. M., Webster, R. G. & Markoff, L. J. Amino acid sequence changes in antigenic variants of type A influenza virus N2 neuraminidase. *Virology* **122**, 450–460 (1982).
124. Luther, P., Bergmann, K. C. & Oxford, J. S. An investigation of antigenic drift of neuraminidases of influenza A (H1N1) viruses. *J. Hyg.* **92**, 223–229 (1984).
125. Schulman, J. L., Khakpour, M. & Kilbourne, E. D. Protective effects of specific immunity to viral neuraminidase on influenza virus infection of mice. *J. Virol.* **2**, 778–786 (1968).
126. Webster, R. G., Reay, P. A. & Laver, W. G. Protection against lethal influenza with neuraminidase. *Virology* **164**, 230–237 (1988).
127. Clements, M. L., Betts, R. F., Tierney, E. L. & Murphy, B. R. Serum and nasal wash antibodies associated with resistance to experimental challenge with influenza A wild-type virus. *J. Clin. Microbiol.* **24**, 157–160 (1986).

128. Murphy, B. R., Kasel, J. A. & Chanock, R. M. Association of serum anti-neuraminidase antibody with resistance to influenza in man. *N. Engl. J. Med.* **286**, 1329–1332 (1972).
129. Brett, I. C. & Johansson, B. E. Immunization against influenza A virus: comparison of conventional inactivated, live-attenuated and recombinant baculovirus produced purified haemagglutinin and neuraminidase vaccines in a murine model system. *Virology* **339**, 273–280 (2005).
130. Couch, R. B., Kasel, J. A., Gerin, J. L., Schulman, J. L. & Kilbourne, E. D. Induction of Partial Immunity to Influenza by a Neuraminidase-specific Vaccine. *J. Infect. Dis.* **129**, 411–420 (1974).
131. Johansson, B. E., Grajower, B. & Kilbourne, E. D. Infection-permissive immunization with influenza virus neuraminidase prevents weight loss in infected mice. *Vaccine* **11**, 1037–1039 (1993).
132. Kilbourne, E. D. Comparative efficacy of neuraminidase-specific and conventional influenza virus vaccines in induction of antibody to neuraminidase in humans. *J. Infect. Dis.* **134**, 384–394 (1976).
133. Compans, R. W., Dimmock, N. J. & Meier-Ewert, H. Effect of antibody to neuraminidase on the maturation and hemagglutinating activity of an influenza A2 virus. *J. Virol.* **4**, 528–534 (1969).
134. Air, G. M., Els, M. C., Brown, L. E., Laver, W. G. & Webster, R. G. Location of antigenic sites on the three-dimensional structure of the influenza N2 virus neuraminidase. *Virology* **145**, 237–248 (1985).
135. Colman, P. M. et al. Three-dimensional structure of a complex of antibody with influenza virus neuraminidase. *Nature* **326**, 358–363 (1987).
136. Colman, P. M., Varghese, J. N. & Laver, W. G. Structure of the catalytic and antigenic sites in influenza virus neuraminidase. *Nature* **303**, 41–44 (1983).
137. Schulman, J. L. & Kilbourne, E. D. Independent variation in nature of haemagglutinin and neuraminidase antigens of influenza virus: distinctiveness of haemagglutinin antigen of Hong Kong-68 virus. *Proc. Natl. Acad. Sci. U. S. A.* **63**, 326–333 (1969).
138. Kilbourne, E. D., Laver, W. G., Schulman, J. L. & Webster, R. G. Antiviral activity of antiserum specific for an influenza virus neuraminidase. *J. Virol.* **2**, 281–288 (1968).

139. Warren, L. The thiobarbituric acid assay of sialic acids. *J. Biol. Chem.* **234**, 1971–1975 (1959).
140. Webster, R. G. & Laver, W. G. Preparation and properties of antibody directed specifically against the neuraminidase of influenza virus. *J. Immunol.* **99**, 49–55 (1967).
141. World Health Organization. WHO manual on animal influenza diagnosis and surveillance. *World Health Organization - Western Pacific Region* (2002). Available at: [http://www.wpro.who.int/entity/emerging\\_diseases/documents/docs/manualonanimalaidiagnosisandsurveillance.pdf](http://www.wpro.who.int/entity/emerging_diseases/documents/docs/manualonanimalaidiagnosisandsurveillance.pdf). (Accessed: 9th August 2018)
142. Lambré, C. R., Terzidis, H., Greffard, A. & Webster, R. G. Measurement of anti-influenza neuraminidase antibody using a peroxidase-linked lectin and microtitre plates coated with natural substrates. *J. Immunol. Methods* **135**, 49–57 (1990).
143. Beutner, K. R. et al. Evaluation of a neuraminidase-specific influenza A virus vaccine in children: antibody responses and effects on two successive outbreaks of natural infection. *J. Infect. Dis.* **140**, 844–850 (1979).
144. Ogra, P. L. et al. Clinical and immunologic evaluation of neuraminidase-specific influenza A virus vaccine in humans. *J. Infect. Dis.* **135**, 499–506 (1977).
145. De Clercq, E. Antiviral agents active against influenza A viruses. *Nat. Rev. Drug Discov.* **5**, 1015–1025 (2006).
146. Salk, J. E. & Suriano, P. C. Importance of antigenic composition of influenza virus vaccine in protecting against the natural disease; observations during the winter of 1947-1948. *Am. J. Public Health Nations. Health* **39**, 345–355 (1949).
147. Payne, A. M. The influenza programme of WHO. *Bull. World Health Organ.* **8**, 755–774 (1953).
148. World Health Organization. Constitution of World Health Organization. *World Health Organization (WHO)* (2005). Available at: <http://apps.who.int/gb/bd/PDF/bd47/EN/constitution-en.pdf?ua=1>. (Accessed: 9th August 2018)
149. Zhang, W. & Wood, J. The Global Influenza Surveillance and Response System (GISRS)-65 years of building trust and sharing and a role model for global health security. *Influenza Other Respi. Viruses* (2018).

150. Allison, J. E., Glezen, W. P., Taber, L. H., Paredes, A. & Webster, R. G. Reactogenicity and immunogenicity of bivalent influenza A and monovalent influenza B virus vaccines in high-risk children. *J. Infect. Dis.* **136 Suppl**, S672–6 (1977).
151. Russell, C. A. et al. Influenza vaccine strain selection and recent studies on the global migration of seasonal influenza viruses. *Vaccine* **26 Suppl 4**, D31–4 (2008).
152. Fiore, A. E. et al. Prevention and control of seasonal influenza with vaccines: recommendations of the Advisory Committee on Immunization Practices (ACIP), 2009. *MMWR Recomm. Rep.* **58**, 1–52 (2009).
153. Belongia, E. A. et al. Effectiveness of inactivated influenza vaccines varied substantially with antigenic match from the 2004-2005 season to the 2006-2007 season. *J. Infect. Dis.* **199**, 159–167 (2009).
154. World Health Organization. Fact Sheet - Influenza (Seasonal). World Health Organization (WHO) (2003). Available at: <http://www.who.int/mediacentre/factsheets/fs211>. (Accessed: 2nd December 2013)
155. Barr, I. G. et al. Epidemiological, antigenic and genetic characteristics of seasonal influenza A(H1N1), A(H3N2) and B influenza viruses: basis for the WHO recommendation on the composition of influenza vaccines for use in the 2009-2010 Northern Hemisphere season. *Vaccine* **28**, 1156–1167 (2010).
156. Taubenberger, J. K. & Morens, D. M. 1918 Influenza: the mother of all pandemics. *Emerg. Infect. Dis.* **12**, 15–22 (2006).
157. Schild, G. C. et al. Antigenic variation in current influenza A viruses: evidence for a high frequency of antigenic 'drift' for the Hong Kong virus. *Bull. World Health Organ.* **51**, 1–11 (1974).
158. Fiore, A. E. et al. Prevention and control of influenza with vaccines: recommendations of the Advisory Committee on Immunization Practices (ACIP), 2010. *MMWR Recomm. Rep.* **59**, 1–62 (2010).
159. Sandbulte, M. R. et al. Discordant antigenic drift of neuraminidase and haemagglutinin in H1N1 and H3N2 influenza viruses. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 20748–20753 (2011).
160. De Jong, J. C. et al. Het influenzaseizoen 2010/2011 in Nederland: het nieuwe A (H1N1)-virus van 2009 blijft actief. *Ned Tijdschr Med Microbiol* **19**, 21–27 (2011).

161. Fitch, W. M., Bush, R. M., Bender, C. A. & Cox, N. J. Long term trends in the evolution of H(3) HA1 human influenza type A. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 7712–7718 (1997).
162. Nagarajan, N. & Kingsford, C. GiRaF: robust, computational identification of influenza reassortments via graph mining. *Nucleic Acids Res.* **39**, e34 (2011).
163. Pybus, O. G. et al. Phylogenetic evidence for deleterious mutation load in RNA viruses and its contribution to viral evolution. *Mol. Biol. Evol.* **24**, 845–852 (2007).
164. Pond, S. L. K. et al. Adaptation to different human populations by HIV-1 revealed by codon-based analyses. *PLoS Comput. Biol.* **2**, e62 (2006).
165. Kosakovsky Pond, S. L. et al. A random effects branch-site model for detecting episodic diversifying selection. *Mol. Biol. Evol.* **28**, 3033–3043 (2011).
166. Pond, S. L. K., Frost, S. D. W. & Muse, S. V. HyPhy: hypothesis testing using phylogenies. *Bioinformatics* **21**, 676–679 (2005).
167. Kosakovsky Pond, S. L., Poon, A. F. Y., Leigh Brown, A. J. & Frost, S. D. W. A maximum likelihood method for detecting directional evolution in protein sequences and its application to influenza A virus. *Mol. Biol. Evol.* **25**, 1809–1824 (2008).
168. Webster, R. G., Bean, W. J., Gorman, O. T., Chambers, T. M. & Kawaoka, Y. Evolution and ecology of influenza A viruses. *Microbiol. Rev.* **56**, 152–179 (1992).
169. Smith, G. J. D. et al. Origins and evolutionary genomics of the 2009 swine-origin H1N1 influenza A epidemic. *Nature* **459**, 1122–1125 (2009).
170. Kaverin, N. V. et al. Intergenic HA–NA interactions in influenza A virus: postreassortment substitutions of charged amino acid in the haemagglutinin of different subtypes. *Virus Res.* **66**, 123–129 (2000).
171. Mitnaul, L. J. et al. Balanced haemagglutinin and neuraminidase activities are critical for efficient replication of influenza A virus. *J. Virol.* **74**, 6015–6020 (2000).
172. Wagner, R., Matrosovich, M. & Klenk, H.-D. Functional balance between haemagglutinin and neuraminidase in influenza virus infections. *Rev. Med. Virol.* **12**, 159–166 (2002).

173. Kendal, A. P. et al. Co-circulation of two influenza A (H3N2) antigenic variants detected by virus surveillance in individual communities. *Am. J. Epidemiol.* **108**, 308–311 (1978).
174. Pereira, M. S. & Chakraverty, P. The laboratory surveillance of influenza epidemics in the United Kingdom 1968–1976. *Epidemiology & Infection* **79**, 77–87 (1977).
175. Schild, G. C. et al. Antigenic variation in current human type A influenza viruses: antigenic characteristics of the variants and their geographic distribution. *Bull. World Health Organ.* **48**, 269–278 (1973).
176. Xu, X., Cox, N. J., Bender, C. A., Regnery, H. L. & Shaw, M. W. Genetic variation in neuraminidase genes of influenza A (H3N2) viruses. *Virology* **224**, 175–183 (1996).
177. Hanada, K., Suzuki, Y. & Gojobori, T. A large variation in the rates of synonymous substitution for RNA viruses and its relationship to a diversity of viral infection and transmission modes. *Mol. Biol. Evol.* **21**, 1074–1080 (2004).
178. Jenkins, G. M., Rambaut, A., Pybus, O. G. & Holmes, E. C. Rates of molecular evolution in RNA viruses: a quantitative phylogenetic analysis. *J. Mol. Evol.* **54**, 156–165 (2002).
179. Bhatt, S., Holmes, E. C. & Pybus, O. G. The genomic rate of molecular adaptation of the human influenza A virus. *Mol. Biol. Evol.* **28**, 2443–2451 (2011).
180. Chen, R. & Holmes, E. C. The evolutionary dynamics of human influenza B virus. *J. Mol. Evol.* **66**, 655–663 (2008).
181. Blok, J. & Air, G. M. Variation in the membrane-insertion and 'stalk' sequences in eight subtypes of influenza type A virus neuraminidase. *Biochemistry* **21**, 4001–4007 (1982).
182. Gulati, U. et al. Antibody epitopes on the neuraminidase of a recent H3N2 influenza virus (A/Memphis/31/98). *J. Virol.* **76**, 12274–12280 (2002).
183. Lentz, M. R., Air, G. M., Laver, W. G. & Webster, R. G. Sequence of the neuraminidase gene of influenza virus A/Tokyo/3/67 and previously uncharacterized monoclonal variants. *Virology* **135**, 257–265 (1984).
184. Hall, T. A. & Others. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. in *Nucleic acids symposium series* **41**, 95–98 ([London]: Information Retrieval Ltd., c1979-c2000., 1999).

- 
185. Posada, D. jModelTest: phylogenetic model averaging. *Mol. Biol. Evol.* **25**, 1253–1256 (2008).
  186. Darriba, D., Taboada, G. L., Doallo, R. & Posada, D. ProtTest 3: fast selection of best-fit models of protein evolution. *Bioinformatics* **27**, 1164–1165 (2011).
  187. Guindon, S. & Gascuel, O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* **52**, 696–704 (2003).
  188. Zwickl, D. J. Genetic algorithm approaches for the phylogenetic analysis of large biological sequence data sets under the maximum likelihood criterion. (2006).
  189. Swofford, D. L. PAUP\*: phylogenetic analysis using parsimony, version 4.0 b10. (2003).
  190. Ronquist, F. & Huelsenbeck, J. P. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**, 1572–1574 (2003).
  191. Huelsenbeck, J. P. & Ronquist, F. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**, 754–755 (2001).
  192. Drummond, A., Pybus, O. G. & Rambaut, A. Inference of viral evolutionary rates from molecular sequences. *Adv. Parasitol.* **54**, 331–358 (2003).
  193. Drummond, A. J. & Rambaut, A. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* **7**, 214 (2007).
  194. Shapiro, B., Rambaut, A. & Drummond, A. J. Choosing appropriate substitution models for the phylogenetic analysis of protein-coding sequences. *Mol. Biol. Evol.* **23**, 7–9 (2006).
  195. Minin, V. N., Bloomquist, E. W. & Suchard, M. A. Smooth skyride through a rough skyline: Bayesian coalescent-based inference of population dynamics. *Mol. Biol. Evol.* **25**, 1459–1471 (2008).
  196. Yang, Z. PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**, 1586–1591 (2007).
  197. Goldman, N. & Yang, Z. A codon-based model of nucleotide substitution for protein-coding DNA sequences. *Mol. Biol. Evol.* **11**, 725–736 (1994).
  198. Westgeest, K. B. et al. Genetic evolution of the neuraminidase of influenza A (H3N2) viruses from 1968 to 2009 and its correspondence to haemagglutinin evolution. *J. Gen. Virol.* **93**, 1996–2007 (2012).

199. Zhou, B. et al. Single-reaction genomic amplification accelerates sequencing and vaccine production for classical and Swine origin human influenza A viruses. *J. Virol.* **83**, 10309–10313 (2009).
200. Zhou, B. & Wentworth, D. E. Influenza A virus molecular virology techniques. *Methods Mol. Biol.* **865**, 175–192 (2012).
201. Dugan, V. G. et al. The evolutionary genetics and emergence of avian influenza viruses in wild birds. *PLoS Pathog.* **4**, e1000076 (2008).
202. Djikeng, A. et al. Viral genome sequencing by random priming methods. *BMC Genomics* **9**, 5 (2008).
203. Bahl, J. et al. Temporally structured metapopulation dynamics and persistence of influenza A H3N2 virus in humans. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 19359–19364 (2011).
204. Kass, R. E. & Raftery, A. E. Bayes Factors. *J. Am. Stat. Assoc.* **90**, 773–795 (1995).
205. Kosakovsky Pond, S. L. & Frost, S. D. W. Not so different after all: a comparison of methods for detecting amino acid sites under selection. *Mol. Biol. Evol.* **22**, 1208–1222 (2005).
206. Murrell, B. et al. FUBAR: a fast, unconstrained bayesian approximation for inferring selection. *Mol. Biol. Evol.* **30**, 1196–1205 (2013).
207. Delport, W., Poon, A. F. Y., Frost, S. D. W. & Kosakovsky Pond, S. L. Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology. *Bioinformatics* **26**, 2455–2457 (2010).
208. Pond, S. L. K. & Frost, S. D. W. Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. *Bioinformatics* **21**, 2531–2533 (2005).
209. Fang, R., Min Jou, W., Huylebroeck, D., Devos, R. & Fiers, W. Complete structure of A/duck/Ukraine/63 influenza haemagglutinin gene: animal virus as progenitor of human H3 Hong Kong 1968 influenza haemagglutinin. *Cell* **25**, 315–323 (1981).
210. Kawaoka, Y., Krauss, S. & Webster, R. G. Avian-to-human transmission of the PB1 gene of influenza A viruses in the 1957 and 1968 pandemics. *J. Virol.* **63**, 4603–4608 (1989).
211. Masurel, N. & Marine, W. M. Recycling of Asian and Hong Kong influenza A virus haemagglutinins in man. *Am. J. Epidemiol.* **97**, 44–49 (1973).



- 
212. Abed, Y., Hardy, I., Li, Y. & Boivin, G. Divergent evolution of haemagglutinin and neuraminidase genes in recent influenza A:H3N2 viruses isolated in Canada. *J. Med. Virol.* **67**, 589–595 (2002).
  213. Bragstad, K., Nielsen, L. P. & Fomsgaard, A. The evolution of human influenza A viruses from 1999 to 2006: a complete genome study. *Virology* **5**, 40 (2008).
  214. Li, Y., Chen, Z. Y., Wang, W., Baker, C. C. & Krug, R. M. The 3'-end-processing factor CPSF is required for the splicing of single-intron pre-mRNAs in vivo. *RNA* **7**, 920–931 (2001).
  215. Melén, K. et al. Nuclear and nucleolar targeting of influenza A virus NS1 protein: striking differences between different virus subtypes. *J. Virol.* **81**, 5995–6006 (2007).
  216. Du, J., Cross, T. A. & Zhou, H.-X. Recent progress in structure-based anti-influenza drug design. *Drug Discov. Today* **17**, 1111–1120 (2012).
  217. Naffakh, N., Tomoiu, A., Rameix-Welti, M.-A. & van der Werf, S. Host restriction of avian influenza viruses at the level of the ribonucleoproteins. *Annu. Rev. Microbiol.* **62**, 403–424 (2008).
  218. Boulo, S., Akarsu, H., Ruigrok, R. W. H. & Baudin, F. Nuclear traffic of influenza virus proteins and ribonucleoprotein complexes. *Virus Res.* **124**, 12–21 (2007).
  219. Hiromoto, Y. et al. Phylogenetic analysis of the three polymerase genes (PB1, PB2 and PA) of influenza B virus. *J. Gen. Virol.* **81**, 929–937 (2000).
  220. Bean, W. J. et al. Evolution of the H3 influenza virus haemagglutinin from human and nonhuman hosts. *J. Virol.* **66**, 1129–1138 (1992).
  221. Trifonov, V., Racaniello, V. & Rabadan, R. The Contribution of the PB1-F2 Protein to the Fitness of Influenza A Viruses and its Recent Evolution in the 2009 Influenza A (H1N1) Pandemic Virus. *PLoS Curr.* **1**, RRN1006 (2009).
  222. Chen, C.-J. et al. Differential localization and function of PB1-F2 derived from different strains of influenza A virus. *J. Virol.* **84**, 10051–10062 (2010).
  223. Shi, M. et al. Evolutionary conservation of the PA-X open reading frame in segment 3 of influenza A virus. *J. Virol.* **86**, 12411–12413 (2012).
  224. Dundon, W. G. & Capua, I. A Closer Look at the NS1 of Influenza Virus. *Viruses* **1**, 1057–1072 (2009).
  225. Feng, J. et al. Influenza A virus infection engenders a poor antibody response against the ectodomain of matrix protein 2. *Virology* **3**, 102 (2006).

226. Bright, R. A., Shay, D. K., Shu, B., Cox, N. J. & Klimov, A. I. Adamantane resistance among influenza A viruses isolated early during the 2005-2006 influenza season in the United States. *JAMA* **295**, 891–894 (2006).
227. Li, S., Min, J.-Y., Krug, R. M. & Sen, G. C. Binding of the influenza A virus NS1 protein to PKR mediates the inhibition of its activation by either PACT or double-stranded RNA. *Virology* **349**, 13–21 (2006).
228. Fernandez-Sesma, A. et al. Influenza virus evades innate and adaptive immunity via the NS1 protein. *J. Virol.* **80**, 6295–6304 (2006).
229. Baigent, S. J. & McCauley, J. W. Influenza type A in humans, mammals and birds: determinants of virus virulence, host-range and interspecies transmission. *Bioessays* **25**, 657–671 (2003).
230. Chen, R. & Holmes, E. C. Avian influenza virus exhibits rapid evolutionary dynamics. *Mol. Biol. Evol.* **23**, 2336–2341 (2006).
231. Both, G. W., Sleight, M. J., Cox, N. J. & Kendal, A. P. Antigenic drift in influenza virus H3 haemagglutinin from 1968 to 1980: multiple evolutionary pathways and sequential amino acid changes at key antigenic sites. *J. Virol.* **48**, 52–60 (1983).
232. Boon, A. C. M. et al. Sequence variation in a newly identified HLA-B35-restricted epitope in the influenza A virus nucleoprotein associated with escape from cytotoxic T lymphocytes. *J. Virol.* **76**, 2567–2572 (2002).
233. Voeten, J. T. et al. Antigenic drift in the influenza A virus (H3N2) nucleoprotein and escape from recognition by cytotoxic T lymphocytes. *J. Virol.* **74**, 6800–6807 (2000).
234. DiBrino, M. et al. The HLA-B14 peptide binding site can accommodate peptides with different combinations of anchor residues. *J. Biol. Chem.* **269**, 32426–32434 (1994).
235. Koelle, K., Cobey, S., Grenfell, B. & Pascual, M. Epochal evolution shapes the phylodynamics of interpandemic influenza A (H3N2) in humans. *Science* **314**, 1898–1903 (2006).
236. Koel, B. F. et al. Substitutions near the receptor binding site determine major antigenic change during influenza virus evolution. *Science* **342**, 976–979 (2013).
237. Webster, R. G., Laver, W. G., Air, G. M. & Schild, G. C. Molecular mechanisms of variation in influenza viruses. *Nature* **296**, 115–121 (1982).

- 
238. Hensley, S. E. et al. Influenza A virus haemagglutinin antibody escape promotes neuraminidase antigenic variation and drug resistance. *PLoS One* **6**, e15190 (2011).
239. Sandbulte, M. R., Gao, J., Straight, T. M. & Eichelberger, M. C. A miniaturized assay for influenza neuraminidase-inhibiting antibodies utilizing reverse genetics-derived antigens. *Influenza Other Respi. Viruses* **3**, 233–240 (2009).
240. Centers for Disease Control and Prevention (CDC). Influenza activity--United States and worldwide, 2007-08 season. *MMWR Morb. Mortal. Wkly. Rep.* **57**, 692–697 (2008).
241. de Jong, J. C., Beyer, W. E., Palache, A. M., Rimmelzwaan, G. F. & Osterhaus, A. D. Mismatch between the 1997/1998 influenza vaccine and the major epidemic A(H3N2) virus strain as the cause of an inadequate vaccine-induced antibody response to this strain in the elderly. *J. Med. Virol.* **61**, 94–99 (2000).
242. Bridges, C. B. et al. Effectiveness and cost-benefit of influenza vaccination of healthy working adults: A randomized controlled trial. *JAMA* **284**, 1655–1663 (2000).
243. Xu, X., Shaw, J., Smith, C. B., Cox, N. J. & Klimov, A. I. Multiple lineages co-circulation and genetic reassortment of the neuraminidase and hemagglutinin genes within influenza viruses of the same type/subtype. *Int. Congr. Ser.* **1219**, 383–387 (2001).
244. Air, G. M., Laver, W. G., Webster, R. G., Els, M. C. & Luo, M. Antibody recognition of the influenza virus neuraminidase. *Cold Spring Harb. Symp. Quant. Biol.* **54 Pt 1**, 247–255 (1989).
245. Pushko, P. et al. Recombinant H1N1 virus-like particle vaccine elicits protective immunity in ferrets against the 2009 pandemic H1N1 influenza virus. *Vaccine* **28**, 4771–4776 (2010).
246. Cate, T. R. et al. A high dosage influenza vaccine induced significantly more neuraminidase antibody than standard vaccine among elderly subjects. *Vaccine* **28**, 2076–2079 (2010).
247. Monto, A. S. & Kendal, A. P. Effect of neuraminidase antibody on Hong Kong influenza. *Lancet* **1**, 623–625 (1973).
248. Viboud, C. et al. Multinational impact of the 1968 Hong Kong influenza pandemic: evidence for a smoldering pandemic. *J. Infect. Dis.* **192**, 233–248 (2005).

249. Schild, G. C. & Newman, R. W. Immunological relationships between the neuraminidases of human and animal influenza viruses. *Bull. World Health Organ.* **41**, 437–445 (1969).
250. Schulman, J. L. The role of antineuraminidase antibody in immunity to influenza virus infection. *Bull. World Health Organ.* **41**, 647–650 (1969).
251. Sandbulte, M. R. et al. Cross-reactive neuraminidase antibodies afford partial protection against H5N1 in mice and are present in unexposed humans. *PLoS Med.* **4**, e59 (2007).
252. Palmer, D., Dowdle, W., Coleman, M. & Schild, G. Advanced laboratory techniques for influenza diagnosis: Immunology Series No 6. Part. 2: Procedural Guide. (1975).
253. Sali, A. & Blundell, T. L. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234**, 779–815 (1993).
254. Couch, R. B. et al. Antibody correlates and predictors of immunity to naturally occurring influenza in humans and the importance of antibody to the neuraminidase. *J. Infect. Dis.* **207**, 974–981 (2013).
255. Couch, R. B. et al. Randomized comparative study of the serum antihaemagglutinin and antineuraminidase antibody responses to six licensed trivalent influenza vaccines. *Vaccine* **31**, 190–195 (2012).
256. Fries, L. F., Smith, G. E. & Glenn, G. M. A recombinant viruslike particle influenza A (H7N9) vaccine. *N. Engl. J. Med.* **369**, 2564–2566 (2013).
257. Fritz, R. et al. A Vero Cell–Derived Whole-Virus H5N1 Vaccine Effectively Induces Neuraminidase-Inhibiting Antibodies. *J. Infect. Dis.* **205**, 28–34 (2012).
258. Hoffmann, E., Neumann, G., Kawaoka, Y., Hobom, G. & Webster, R. G. A DNA transfection system for generation of influenza A virus from eight plasmids. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 6108–6113 (2000).
259. Sultana, I., Gao, J., Markoff, L. & Eichelberger, M. C. Influenza neuraminidase-inhibiting antibodies are induced in the presence of zanamivir. *Vaccine* **29**, 2601–2606 (2011).
260. Eichelberger, M. C. et al. Qualitative differences in T cell responses to live, attenuated and inactivated influenza vaccines. *J. Clin. Cell. Immunol.* **4**, 002 (2011).

- 
261. Bland, J. M. & Altman, D. G. Measuring agreement in method comparison studies. *Stat. Methods Med. Res.* **8**, 135–160 (1999).
  262. Hassantoufighi, A. et al. A practical influenza neutralization assay to simultaneously quantify haemagglutinin and neuraminidase-inhibiting antibody responses. *Vaccine* **28**, 790–797 (2010).
  263. Johansson, B. E. & Kilbourne, E. D. Dissociation of influenza virus haemagglutinin and neuraminidase eliminates their intravirionic antigenic competition. *J. Virol.* **67**, 5721–5723 (1993).
  264. Johansson, B. E. & Kilbourne, E. D. Immunization with dissociated neuraminidase, matrix, and nucleoproteins from influenza A virus eliminates cognate help and antigenic competition. *Virology* **225**, 136–144 (1996).
  265. Anders, E. M., Hartley, C. A. & Jackson, D. C. Bovine and mouse serum beta inhibitors of influenza A viruses are mannose-binding lectins. *Proc. Natl. Acad. Sci. U. S. A.* **87**, 4485–4489 (1990).
  266. White, M. R. et al. Cooperative anti-influenza activities of respiratory innate immune proteins and neuraminidase inhibitor. *Am. J. Physiol. Lung Cell. Mol. Physiol.* **288**, L831–40 (2005).
  267. Hartshorn, K. L. et al. Conglutinin acts as an opsonin for influenza A viruses. *J. Immunol.* **151**, 6265–6273 (1993).
  268. Rogers, G. N., Pritchett, T. J., Lane, J. L. & Paulson, J. C. Differential sensitivity of human, avian, and equine influenza A viruses to a glycoprotein inhibitor of infection: selection of receptor specific variants. *Virology* **131**, 394–408 (1983).
  269. Matrosovich, M., Gao, P. & Kawaoka, Y. Molecular mechanisms of serum resistance of human influenza H3N2 virus and their involvement in virus adaptation in a new host. *J. Virol.* **72**, 6373–6380 (1998).
  270. Job, E. R. et al. Serum amyloid P is a sialylated glycoprotein inhibitor of influenza A viruses. *PLoS One* **8**, e59623 (2013).
  271. Cox, N., Webster, R. & Stohr, K. WHO Manual on Animal Influenza Diagnosis and Surveillance. *World Health Organization Department Communicable Disease Surveillance and Response.* (2002).
  272. Lambre, C. R., Terzidis, H., Greffard, A. & Webster, R. G. An enzyme-linked lectin assay for sialidase. *Clin. Chim. Acta* **198**, 183–193 (1991).

273. Couzens, L. et al. An optimized enzyme-linked lectin assay to measure influenza A virus neuraminidase inhibition antibody titers in human sera. *J. Virol. Methods* **210C**, 7–14 (2014).
274. de Wit, E. et al. Efficient generation and growth of influenza virus A/PR/8/34 from eight cDNA fragments. *Virus Res.* **103**, 155–161 (2004).
275. de Wit, E. et al. Molecular determinants of adaptation of highly pathogenic avian influenza H7N7 viruses to efficient replication in the human host. *J. Virol.* **84**, 1597–1606 (2010).
276. Schrauwen, E. J. et al. Insertion of a multibasic cleavage site in the haemagglutinin of human influenza H3N2 virus does not increase pathogenicity in ferrets. *J. Gen. Virol.* **92**, 1410–1415 (2011).
277. Chutinimitkul, S. et al. In vitro assessment of attachment pattern and replication efficiency of H5N1 influenza A viruses with altered receptor specificity. *J. Virol.* **84**, 6825–6833 (2010).
278. Fouchier, R. A. et al. Detection of influenza A viruses from different species by PCR amplification of conserved sequences in the matrix gene. *J. Clin. Microbiol.* **38**, 4096–4101 (2000).
279. Burnet, F. M. & Stone, J. D. The receptor-destroying enzyme of *V. cholerae*. *Aust. J. Exp. Biol. Med. Sci.* **25**, 227–233 (1947).
280. Hirst, G. K. Studies of Antigenic Differences among Strains of Influenza a by Means of Red Cell Agglutination. *J. Exp. Med.* **78**, 407–423 (1943).
281. Hirst, G. K. The Quantitative Determination of Influenza Virus and Antibodies by Means of Red Cell Agglutination. *J. Exp. Med.* **75**, 49–64 (1942).
282. Francis, T. Dissociation of Hemagglutinating and Antibody-Measuring Capacities of Influenza Virus. *J. Exp. Med.* **85**, 1–7 (1947).
283. Ryan-Poirier, K. A. & Kawaoka, Y. Distinct glycoprotein inhibitors of influenza A virus in different animal sera. *J. Virol.* **65**, 389–395 (1991).
284. Pritchett, T. J. & Paulson, J. C. Basis for the potent inhibition of influenza virus infection by equine and guinea pig alpha 2-macroglobulin. *J. Biol. Chem.* **264**, 9850–9858 (1989).
285. Hartley, C. A., Jackson, D. C. & Anders, E. M. Two distinct serum mannose-binding lectins function as beta inhibitors of influenza virus: identification of bovine serum beta inhibitor as conglutinin. *J. Virol.* **66**, 4358–4363 (1992).

- 
286. Krizanova, O. & Rathova, V. Serum inhibitors of myxoviruses. *Curr. Top. Microbiol. Immunol.* **47**, 125–151 (1969).
287. Subbarao, E. K., Kawaoka, Y., Ryan-Poirier, K., Clements, M. L. & Murphy, B. R. Comparison of different approaches to measuring influenza A virus-specific haemagglutination inhibition antibodies in the presence of serum inhibitors. *J. Clin. Microbiol.* **30**, 996–999 (1992).
288. Ryan-Poirier, K. A. & Kawaoka, Y. Alpha 2-macroglobulin is the major neutralizing inhibitor of influenza A virus in pig serum. *Virology* **193**, 974–976 (1993).
289. Konno, J. Studies on several inhibitors against influenza viruses. 2. beta-Inhibitor, its biological and physicochemical properties with particular emphasis on the differences from alpha-inhibitor, immune serum and properdin. *Tohoku J. Exp. Med.* **67**, 391–405 (1958).
290. Cohen, A. & Belyavin, G. Haemagglutination inhibition of Asian influenza viruses: a new pattern of response. *Virology* **7**, 59–74 (1959).
291. Flandorfer, A., Garcia-Sastre, A., Basler, C. F. & Palese, P. Chimeric influenza A viruses with a functional influenza B virus neuraminidase or haemagglutinin. *J. Virol.* **77**, 9116–9123 (2003).
292. Baker, S. F. et al. Influenza A and B virus intertypic reassortment through compatible viral packaging signals. *J. Virol.* **88**, 10778–10791 (2014).
293. Nichol, K. L. Efficacy and effectiveness of influenza vaccination. *Vaccine* **26 Suppl 4**, D17–22 (2008).
294. Nichol, K. L. The efficacy, effectiveness and cost-effectiveness of inactivated influenza virus vaccines. *Vaccine* **21**, 1769–1775 (2003).
295. Pica, N. & Palese, P. Toward a universal influenza virus vaccine: prospects and challenges. *Annu. Rev. Med.* **64**, 189–202 (2013).
296. Cox, R. J. Correlates of protection to influenza virus, where do we go from here? *Hum. Vaccin. Immunother.* **9**, 405–408 (2013).
297. Johansson, B. E., Moran, T. M. & Kilbourne, E. D. Antigen-presenting B cells and helper T cells cooperatively mediate intravirionic antigenic competition between influenza A virus surface glycoproteins. *Proc. Natl. Acad. Sci. U. S. A.* **84**, 6869–6873 (1987).

298. Wilson, I. A., Skehel, J. J. & Wiley, D. C. Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 Å resolution. *Nature* **289**, 366–373 (1981).
299. Bedford, T. et al. Integrating influenza antigenic dynamics with molecular evolution. *Elife* **3**, e01914 (2014).
300. Kilbourne, E. D., Cerini, C. P., Khan, M. W., Mitchell, J. W., Jr & Ogra, P. L. Immunologic response to the influenza virus neuraminidase is influenced by prior experience with the associated viral haemagglutinin. I. Studies in human vaccinees. *J. Immunol.* **138**, 3010–3013 (1987).
301. Monto, A. S. et al. Antibody to Influenza Virus Neuraminidase: An Independent Correlate of Protection. *J. Infect. Dis.* **212**, 1191–1199 (2015).
302. Westgeest, K. B. et al. Optimization of an enzyme-linked lectin assay suitable for rapid antigenic characterization of the neuraminidase of human influenza A(H3N2) viruses. *J. Virol. Methods* **217**, 55–63 (2015).
303. Fouchier, R. A. M. & Smith, D. J. Use of antigenic cartography in vaccine seed strain selection. *Avian Dis.* **54**, 220–223 (2010).
304. Hirst - Journal of Experimental Medicine, G. K. & 1943. Studies of antigenic differences among strains of influenza A by means of red cell agglutination. *jem.rupress.org* (1943).
305. Westgeest, K. B. et al. Genomewide analysis of reassortment and evolution of human influenza A(H3N2) viruses circulating between 1968 and 2011. *J. Virol.* **88**, 2844–2857 (2014).
306. Garreta, R. & Moncecchi, G. *Learning scikit-learn: Machine Learning in Python.* (Packt Publishing Ltd, 2013).
307. Lloyd, S. Least squares quantization in PCM. *IEEE Trans. Inf. Theory* **28**, 129–137 (1982).
308. Arthur, D. & Vassilvitskii, S. k-means++: The advantages of careful seeding. *Proceedings of the eighteenth annual ACM* (2007).
309. Ester, M., Kriegel, H. P., Sander, J. & Xu, X. A density-based algorithm for discovering clusters in large spatial databases with noise. in *Second International Conference on Knowledge Discovery and Data Mining* (eds. Simoudis, E., Han, J. & Fayyad, U. M.) 226–231 (1996).



- 
310. Frey, B. J. & Dueck, D. Clustering by passing messages between data points. *Science* **315**, 972–976 (2007).
  311. Rosenberg, A., Hirschberg - EMNLP-CoNLL, J. & 2007. V-Measure: A Conditional Entropy-Based External Cluster Evaluation Measure. *aclweb.org* (2007).
  312. Hubert, L. & Arabie, P. Comparing partitions. *J. Classification* **2**, 193–218 (1985).
  313. Vinh, N. X., Epps, J. & Bailey, J. Information Theoretic Measures for Clusterings Comparison: Variants, Properties, Normalization and Correction for Chance. *J. Mach. Learn. Res.* **11**, 2837–2854 (2010).
  314. Rousseeuw, P. J. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* **20**, 53–65 (1987).
  315. Mögling, R. et al. Neuraminidase-mediated haemagglutination of recent human influenza A(H3N2) viruses is determined by arginine 150 flanking the neuraminidase catalytic site. *J. Gen. Virol.* **98**, 1274–1281 (2017).
  316. Wan, Z. et al. Identification of amino acids in H9N2 influenza virus neuraminidase that are critical for the binding of two mouse monoclonal antibodies. *Vet. Microbiol.* **187**, 58–63 (2016).
  317. Saito, T., Taylor, G., Laver, W. G., Kawaoka, Y. & Webster, R. G. Antigenicity of the N8 influenza A virus neuraminidase: existence of an epitope at the subunit interface of the neuraminidase. *J. Virol.* **68**, 1790–1796 (1994).
  318. Webster, R. G. et al. Antigenic structure and variation in an influenza virus N9 neuraminidase. *J. Virol.* **61**, 2910–2916 (1987).
  319. Tulloch, P. A. et al. Electron and X-ray diffraction studies of influenza neuraminidase complexed with monoclonal antibodies. *J. Mol. Biol.* **190**, 215–225 (1986).
  320. Venkatramani, L. et al. An Epidemiologically Significant Epitope of a 1998 Human Influenza Virus Neuraminidase Forms a Highly Hydrated Interface in the NA–Antibody Complex. *J. Mol. Biol.* **356**, 651–663 (2006).
  321. Kilbourne, E. D. et al. Purified influenza A virus N2 neuraminidase vaccine is immunogenic and non-toxic in humans. *Vaccine* **13**, 1799–1803 (1995).
  322. van der Velden, M. V. W. et al. Safety and immunogenicity of a vero cell culture-derived whole-virus influenza A(H5N1) vaccine in a pediatric population. *J. Infect. Dis.* **209**, 12–23 (2014).

323. Wohlbold, T. J., Hirsh, A. & Krammer, F. An H10N8 influenza virus vaccine strain and mouse challenge model based on the human isolate A/Jiangxi-Donghu/346/13. *Vaccine* **33**, 1102–1106 (2015).
324. Wohlbold, T. J. et al. Hemagglutinin Stalk- and Neuraminidase-Specific Monoclonal Antibodies Protect against Lethal H10N8 Influenza Virus Infection in Mice. *J. Virol.* **90**, 851–861 (2015).
325. Wan, H. et al. Structural characterization of a protective epitope spanning A(H1N1)pdm09 influenza virus neuraminidase monomers. *Nat. Commun.* **6**, 6114 (2015).
326. Wan, H. et al. Molecular basis for broad neuraminidase immunity: conserved epitopes in seasonal and pandemic H1N1 as well as H5N1 influenza viruses. *J. Virol.* **87**, 9290–9300 (2013).
327. Easterbrook, J. D. et al. Protection against a lethal H5N1 influenza challenge by intranasal immunization with virus-like particles containing 2009 pandemic H1N1 neuraminidase in mice. *Virology* **432**, 39–44 (2012).
328. Liu, W.-C., Lin, C.-Y., Tsou, Y.-T., Jan, J.-T. & Wu, S.-C. Cross-Reactive Neuraminidase-Inhibiting Antibodies Elicited by Immunization with Recombinant Neuraminidase Proteins of H5N1 and Pandemic H1N1 Influenza A Viruses. *J. Virol.* **89**, 7224–7234 (2015).
329. Jiang, L. et al. Comparative Efficacy of Monoclonal Antibodies That Bind to Different Epitopes of the 2009 Pandemic H1N1 Influenza Virus Neuraminidase. *J. Virol.* **90**, 117–128 (2015).
330. Martinet, W., Saelens, X., Deroo, T. & Neiryneck, S. Protection of mice against a lethal influenza challenge by immunization with yeast-derived recombinant influenza neuraminidase. *FEBS J.* (1997).
331. Chen, Z. et al. Cross-protection against a lethal influenza virus infection by DNA vaccine to neuraminidase. *Vaccine* **18**, 3214–3222 (2000).
332. Wu, C.-Y. et al. A VLP vaccine induces broad-spectrum cross-protective antibody immunity against H5N1 and H1N1 subtypes of influenza A virus. *PLoS One* **7**, e42363 (2012).
333. Quan, F.-S. et al. Influenza M1 VLPs containing neuraminidase induce heterosubtypic cross-protection. *Virology* **430**, 127–135 (2012).

- 
334. Johansson, B. E., Pokorny, B. A. & Tiso, V. A. Supplementation of conventional trivalent influenza vaccine with purified viral N1 and N2 neuraminidases induces a balanced immune response without antigenic competition. *Vaccine* **20**, 1670–1674 (2002).
335. Yang, Y.-L., Chang, S.-H., Gong, X., Wu, J. & Liu, B. Expression, purification and characterization of low-glycosylation influenza neuraminidase in  $\alpha$ -1,6-mannosyltransferase defective *Pichia pastoris*. *Mol. Biol. Rep.* **39**, 857–864 (2012).
336. Halbherr, S. J. et al. Biological and protective properties of immune sera directed to the influenza virus neuraminidase. *J. Virol.* **89**, 1550–1563 (2015).
337. Bosch, B. J. et al. Recombinant soluble, multimeric HA and NA exhibit distinctive types of protection against pandemic swine-origin 2009 A(H1N1) influenza virus infection in ferrets. *J. Virol.* **84**, 10366–10374 (2010).
338. Rockman, S. et al. Neuraminidase-inhibiting antibody is a correlate of cross-protection against lethal H5N1 influenza virus in ferrets immunized with seasonal influenza vaccine. *J. Virol.* **87**, 3053–3061 (2013).
339. Eichelberger, M. C. et al. Comparability of neuraminidase inhibition antibody titers measured by enzyme-linked lectin assay (ELLA) for the analysis of influenza vaccine immunogenicity. *Vaccine* **34**, 458–465 (2016).
340. Memoli, M. J. et al. Evaluation of Antihaemagglutinin and Antineuraminidase Antibodies as Correlates of Protection in an Influenza A/H1N1 Virus Healthy Human Challenge Model. *MBio* **7**, e00417–16 (2016).
341. Dunning, A. J. et al. Correlates of Protection against Influenza in the Elderly: Results from an Influenza Vaccine Efficacy Trial. *Clin. Vaccine Immunol.* **23**, 228–235 (2016).
342. Air, G. M. Influenza neuraminidase. *Influenza Other Respi. Viruses* **6**, 245–256 (2012).
343. Air, G. M., Laver, W. G. & Webster, R. G. Antigenic variation in influenza viruses. *Contrib. Microbiol. Immunol.* **8**, 20–59 (1987).
344. Malby, R. L. et al. The structure of a complex between the NC10 antibody and influenza virus neuraminidase and comparison with the overlapping binding site of the NC41 antibody. *Structure* **2**, 733–746 (1994).

345. Greene, S. K., Ionides, E. L. & Wilson, M. L. Patterns of Influenza-associated Mortality among US Elderly by Geographic Region and Virus Subtype, 1968–1998. *Am. J. Epidemiol.* **163**, 316–326 (2006).
346. Vestergaard, L. S. et al. Excess all-cause and influenza-attributable mortality in Europe, December 2016 to February 2017. *Euro Surveill.* **22**, (2017).
347. Budd, A. P. et al. Update: Influenza Activity - United States, October 1, 2017-February 3, 2018. *MMWR Morb. Mortal. Wkly. Rep.* **67**, 169–179 (2018).
348. Bedford, T. et al. Global circulation patterns of seasonal influenza viruses vary with antigenic drift. *Nature* **523**, 217–220 (2015).
349. Belongia, E. A. et al. Variable influenza vaccine effectiveness by subtype: a systematic review and meta-analysis of test-negative design studies. *Lancet Infect. Dis.* **16**, 942–951 (2016).
350. Lakdawala, S. S. et al. Eurasian-origin gene segments contribute to the transmissibility, aerosol release, and morphology of the 2009 pandemic H1N1 influenza virus. *PLoS Pathog.* **7**, e1002443 (2011).
351. Yen, H.-L. et al. Hemagglutinin-neuraminidase balance confers respiratory-droplet transmissibility of the pandemic H1N1 influenza virus in ferrets. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 14264–14269 (2011).
352. Su, B. et al. Enhancement of the influenza A haemagglutinin (HA)-mediated cell-cell fusion and virus entry by the viral neuraminidase (NA). *PLoS One* **4**, e8495 (2009).
353. Cohen, M. et al. Influenza A penetrates host mucus by cleaving sialic acids with neuraminidase. *Viol. J.* **10**, 321 (2013).
354. Yang, X. et al. A beneficiary role for neuraminidase in influenza virus penetration through the respiratory mucus. *PLoS One* **9**, e110026 (2014).
355. Sakai, T., Nishimura, S. I., Naito, T. & Saito, M. Influenza A virus haemagglutinin and neuraminidase act as novel motile machinery. *Sci. Rep.* **7**, 45043 (2017).
356. Mostafa, A. et al. Phylogenetic analysis of human influenza A/H3N2 viruses isolated in 2015 in Germany indicates significant genetic divergence from vaccine strains. *Arch. Virol.* **161**, 1505–1515 (2016).
357. Tewawong, N. et al. Evolution of the neuraminidase gene of seasonal influenza A and B viruses in Thailand between 2010 and 2015. *PLoS One* **12**, e0175655 (2017).

- 
358. Vijaykrishna, D. et al. The contrasting phylodynamics of human influenza B viruses. *Elife* **4**, e05055 (2015).
  359. Oong, X. Y. et al. Epidemiological and Evolutionary Dynamics of Influenza B Viruses in Malaysia, 2012–2014. *PLoS One* **10**, e0136254 (2015).
  360. Nobusawa, E. & Sato, K. Comparison of the mutation rates of human influenza A and B viruses. *J. Virol.* **80**, 3675–3678 (2006).
  361. Monto, A. S. & Maassab, H. F. Ether treatment of type B influenza virus antigen for the haemagglutination inhibition test. *J. Clin. Microbiol.* **13**, 54–57 (1981).
  362. Hensley, S. E. et al. Hemagglutinin receptor binding avidity drives influenza A virus antigenic drift. *Science* **326**, 734–736 (2009).
  363. Petrova, V. N. & Russell, C. A. The evolution of seasonal influenza viruses. *Nat. Rev. Microbiol.* **16**, 60 (2018).
  364. Makkoch, J. et al. Whole Genome Characterization, Phylogenetic and Genome Signature Analysis of Human Pandemic H1N1 Virus in Thailand, 2009–2012. *PLoS One* **7**, e51275 (2012).
  365. Su, Y. C. F. et al. Phylodynamics of H1N1/2009 influenza reveals the transition from host adaptation to immune-driven selection. *Nat. Commun.* **6**, 7952 (2015).
  366. Lin, J.-H. et al. Molecular epidemiology and antigenic analyses of influenza A viruses H3N2 in Taiwan. *Clin. Microbiol. Infect.* **17**, 214–222 (2011).
  367. Kim, J. I. et al. Phylogenetic relationships of the HA and NA genes between vaccine and seasonal influenza A(H3N2) strains in Korea. *PLoS One* **12**, e0172059 (2017).
  368. Lee, J. T. & Air, G. M. Interaction between a 1998 human influenza virus N2 neuraminidase and monoclonal antibody Mem5. *Virology* **345**, 424–433 (2006).
  369. Lee, J. T. & Air, G. M. Contacts between influenza virus N9 neuraminidase and monoclonal antibody NC10. *Virology* **300**, 255–268 (2002).
  370. Nuss, J. M., Whitaker, P. B. & Air, G. M. Identification of critical contact residues in the NC41 epitope of a subtype N9 influenza virus neuraminidase. *Proteins* **15**, 121–132 (1993).
  371. Neumann, G., Noda, T. & Kawaoka, Y. Emergence and pandemic potential of swine-origin H1N1 influenza virus. *Nature* **459**, 931 (2009).

372. Strelkowa, N. & Lässig, M. Clonal interference in the evolution of influenza. *Genetics* **192**, 671–682 (2012).
373. Luksza, M. & Lässig, M. A predictive fitness model for influenza. *Nature* **507**, 57–61 (2014).
374. Meyer, A. G. & Wilke, C. O. Geometric Constraints Dominate the Antigenic Evolution of Influenza H3N2 Hemagglutinin. *PLoS Pathog.* **11**, e1004940 (2015).
375. Ghedin, E. et al. Mixed infection and the genesis of influenza virus diversity. *J. Virol.* **83**, 8832–8841 (2009).
376. Murcia, P. R. et al. Evolution of an Eurasian avian-like influenza virus in naive and vaccinated pigs. *PLoS Pathog.* **8**, e1002730 (2012).
377. Hughes, J. et al. Transmission of equine influenza virus during an outbreak is characterized by frequent mixed infections and loose transmission bottlenecks. *PLoS Pathog.* **8**, e1003081 (2012).
378. Bodewes, R. et al. Use of influenza A viruses expressing reporter genes to assess the frequency of double infections in vitro. *J. Gen. Virol.* **93**, 1645–1648 (2012).
379. Marshall, N., Priyamvada, L., Ende, Z., Steel, J. & Lowen, A. C. Influenza virus reassortment occurs with high frequency in the absence of segment mismatch. *PLoS Pathog.* **9**, e1003421 (2013).
380. Huang, I.-C. et al. Influenza A virus neuraminidase limits viral superinfection. *J. Virol.* **82**, 4834–4843 (2008).
381. Gerber, M., Isel, C., Moules, V. & Marquet, R. Selective packaging of the influenza A genome and consequences for genetic reassortment. *Trends Microbiol.* **22**, 446–455 (2014).
382. Piralla, A. et al. Molecular epidemiology of influenza B virus among hospitalized pediatric patients in Northern Italy during the 2015-16 season. *PLoS One* **12**, e0185893 (2017).
383. Harvey, W. T. et al. Identification of low- and high-impact haemagglutinin amino acid substitutions that drive antigenic drift of influenza A (H1N1) viruses. *PLoS Pathog.* **12**, e1005526 (2016).
384. Chambers, B. S., Parkhouse, K., Ross, T. M., Alby, K. & Hensley, S. E. Identification of Hemagglutinin Residues Responsible for H3N2 Antigenic Drift during the 2014-2015 Influenza Season. *Cell Rep.* **12**, 1–6 (2015).

- 
385. Neverov, A. D., Lezhnina, K. V., Kondrashov, A. S. & Bazykin, G. A. Intrasubtype reassortments cause adaptive amino acid replacements in H3N2 influenza genes. *PLoS Genet.* **10**, e1004037 (2014).
386. Koel, B. F. et al. Antigenic variation of clade 2.1 H5N1 virus is determined by a few amino acid substitutions immediately adjacent to the receptor binding site. *MBio* **5**, e01070–14 (2014).
387. Fonville, J. M. et al. Antigenic Maps of Influenza A(H3N2) Produced With Human Antisera Obtained After Primary Infection. *J. Infect. Dis.* **213**, 31–38 (2016).
388. Lewis, N. S. et al. Antigenic and genetic evolution of equine influenza A (H3N8) virus from 1968 to 2007. *J. Virol.* **85**, 12742–12749 (2011).
389. Abbas, M. A. et al. H7 avian influenza virus vaccines protect chickens against challenge with antigenically diverse isolates. *Vaccine* **29**, 7424–7429 (2011).
390. Lorusso, A. et al. Genetic and antigenic characterization of H1 influenza viruses from United States swine from 2008. *J. Gen. Virol.* **92**, 919–930 (2011).
391. Wohlbold, T. J. et al. Vaccination with adjuvanted recombinant neuraminidase induces broad heterologous, but not heterosubtypic, cross-protection against influenza virus infection in mice. *MBio* **6**, e02556 (2015).
392. Nachbagauer, R. et al. Induction of broadly reactive anti-haemagglutinin stalk antibodies by an H5N1 vaccine in humans. *J. Virol.* **88**, 13260–13268 (2014).
393. Powers, D. C., Kilbourne, E. D. & Johansson, B. E. Neuraminidase-specific antibody responses to inactivated influenza virus vaccine in young and elderly adults. *Clin. Diagn. Lab. Immunol.* **3**, 511–516 (1996).
394. Laguio-Vila, M. R. et al. Comparison of serum haemagglutinin and neuraminidase inhibition antibodies after 2010–2011 trivalent inactivated influenza vaccination in healthcare personnel. in *Open forum infectious diseases* **2**, (Oxford University Press, 2015).
395. Johansson, B. E., Matthews, J. T. & Kilbourne, E. D. Supplementation of conventional influenza A vaccine with purified viral neuraminidase results in a balanced and broadened immune response. *Vaccine* **16**, 1009–1015 (1998).
396. Krammer, F. & Palese, P. Advances in the development of influenza virus vaccines. *Nat. Rev. Drug Discov.* **14**, 167–182 (2015).

397. Dalakouras, T., Smith, B. J., Platis, D., Cox, M. M. J. & Labrou, N. E. Development of recombinant protein-based influenza vaccine. Expression and affinity purification of H1N1 influenza virus neuraminidase. *J. Chromatogr. A* **1136**, 48–56 (2006).
398. Food, Administration, D. & Others. Summary basis for regulatory action—Flublok Quadrivalent. Silver Spring, MD: US Department of Health and Human Services. *FDA Med. Bull.* (2016).
399. Traynor, K. First recombinant flu vaccine approved. *Am. J. Health. Syst. Pharm.* **70**, 382 (2013).
400. Fritz, R. et al. Neuraminidase-Inhibiting Antibody Response to H5N1 Virus Vaccination in Chronically Ill and Immunocompromised Patients. *Open Forum Infect Dis* **1**, ofu072 (2014).
401. Gao, J., Couzens, L. & Eichelberger, M. C. Measuring Influenza Neuraminidase Inhibition Antibody Titers by Enzyme-linked Lectin Assay. *J. Vis. Exp.* (2016). doi:10.3791/54573
402. Johansson, B. E., Moran, T. M., Bona, C. A., Popple, S. W. & Kilbourne, E. D. Immunologic response to influenza virus neuraminidase is influenced by prior experience with the associated viral haemagglutinin. II. Sequential infection of mice simulates human experience. *of Immunology* (1987).
403. Koel, B. F. et al. Identification of amino acid substitutions supporting antigenic change of influenza A(H1N1)pdm09 viruses. *J. Virol.* **89**, 3763–3775 (2015).
404. Koelle, K. & Rasmussen, D. A. The effects of a deleterious mutation load on patterns of influenza A/H3N2's antigenic evolution in humans. *Elife* (2015).
405. Gaymard, A., Le Briand, N., Frobert, E., Lina, B. & Escuret, V. Functional balance between neuraminidase and haemagglutinin in influenza viruses. *Clin. Microbiol. Infect.* **22**, 975–983 (2016).
406. Neverov, A. D., Kryazhimskiy, S. & Plotkin, J. B. Coordinated evolution of influenza A surface proteins. *PLoS Genet.* (2015).
407. De Jong, J. C. et al. Antigenic and genetic evolution of swine influenza A (H3N2) viruses in Europe. *J. Virol.* **81**, 4315–4322 (2007).
408. Wang, Y., Davidson, I., Fouchier, R. & Spackman, E. Antigenic Cartography of H9 Avian Influenza Virus and Its Application to Vaccine Selection. *Avian Dis.* **60**, 218–225 (2016).



- 
409. Huang, S.-W. et al. Mapping Enterovirus A71 Antigenic Determinants from Viral Evolution. *J. Virol.* **89**, 11500–11506 (2015).
410. Huang, S.-W. et al. Reemergence of enterovirus 71 in taiwan: dynamics of genetic and antigenic evolution from 1998 to 2008. *J. Clin. Microbiol.* **47**, 3653–3662 (2009).
411. Mansfield, K. L. et al. Flavivirus-induced antibody cross-reactivity. *J. Gen. Virol.* **92**, 2821–2829 (2011).
412. Ludi, A. B. et al. Antigenic variation of foot-and-mouth disease virus serotype A. *J. Gen. Virol.* **95**, 384–392 (2014).
413. Nolden, T. et al. Comparative studies on the genetic, antigenic and pathogenic characteristics of Bokeloh bat lyssavirus. *J. Gen. Virol.* **95**, 1647–1653 (2014).
414. James, S. L. et al. Antigenic Relationships among Human Pathogenic Orientia tsutsugamushi Isolates from Thailand. *PLoS Negl. Trop. Dis.* **10**, e0004723 (2016).
415. Cai, Z., Zhang, T. & Wan, X.-F. A computational framework for influenza antigenic cartography. *PLoS Comput. Biol.* **6**, e1000949 (2010).
416. Cai, Z., Zhang, T. & Wan, X.-F. Concepts and applications for influenza antigenic cartography. *Influenza Other Respi. Viruses* **5 Suppl 1**, 204–207 (2011).
417. Barnett, J. L., Yang, J., Cai, Z., Zhang, T. & Wan, X.-F. AntigenMap 3D: an online antigenic cartography resource. *Bioinformatics* **28**, 1292–1293 (2012).
418. Xia, J. et al. Genetic and antigenic evolution of H9N2 subtype avian influenza virus in domestic chickens in southwestern China, 2013–2016. *PLoS One* **12**, e0171564 (2017).
419. Martin, B. E. et al. Detection of Antigenic Variants of Subtype H3 Swine Influenza A Viruses from Clinical Samples. *J. Clin. Microbiol.* **55**, 1037–1045 (2017).
420. Bailey, E. et al. Antigenic Characterization of H3 Subtypes of Avian Influenza A Viruses from North America. *Avian Dis.* **60**, 346–353 (2016).
421. Wei, Y. et al. Antigenic evolution of H9N2 chicken influenza viruses isolated in China during 2009–2013 and selection of a candidate vaccine strain with broad cross-reactivity. *Vet. Microbiol.* **182**, 1–7 (2016).

422. Spackman, E. et al. Variation in protection of four divergent avian influenza virus vaccine seed strains against eight clade 2.2.1 and 2.2.1.1. Egyptian H5N1 high pathogenicity variants in poultry. *Influenza Other Respi. Viruses* **8**, 654–662 (2014).
423. Yang, J., Zhang, T. & Wan, X.-F. Sequence-based antigenic change prediction by a sparse learning method incorporating co-evolutionary information. *PLoS One* **9**, e106660 (2014).
424. Kandeil, A. et al. Genetic and antigenic evolution of H9N2 avian influenza viruses circulating in Egypt between 2011 and 2013. *Arch. Virol.* **159**, 2861–2876 (2014).
425. Kayali, G. et al. Active surveillance for avian influenza virus, Egypt, 2010–2012. *Emerg. Infect. Dis.* **20**, 542–551 (2014).
426. Xie, H. et al. H3N2 Mismatch of 2014–15 Northern Hemisphere Influenza Vaccines and Head-to-head Comparison between Human and Ferret Antisera derived Antigenic Maps. *Sci. Rep.* **5**, 15279 (2015).
427. Swanstrom, J. A. et al. Dengue Virus Envelope Dimer Epitope Monoclonal Antibodies Isolated from Dengue Patients Are Protective against Zika Virus. *MBio* **7**, (2016).
428. Debbink, K. et al. Within-host evolution results in antigenically distinct GII.4 noroviruses. *J. Virol.* **88**, 7244–7255 (2014).
429. Lindesmith, L. C. et al. Broad blockade antibody responses in human volunteers after immunization with a multivalent norovirus VLP candidate vaccine: immunological analyses from a phase I clinical trial. *PLoS Med.* **12**, e1001807 (2015).
430. Dowdle, W. R. Influenza anti-neuraminidase: the second best antibody. *N. Engl. J. Med.* **286**, 1360–1361 (1972).
431. Eichelberger, M. C. & Wan, H. Influenza neuraminidase as a vaccine antigen. *Curr. Top. Microbiol. Immunol.* **386**, 275–299 (2015).
432. Wohlbold, T. J. & Krammer, F. In the shadow of haemagglutinin: a growing interest in influenza viral neuraminidase and its role as a vaccine antigen. *Viruses* **6**, 2465–2494 (2014).
433. Petrie, J. G. & Gordon, A. Epidemiological Studies to Support the Development of Next Generation Influenza Vaccines. *Vaccines (Basel)* **6**, (2018).

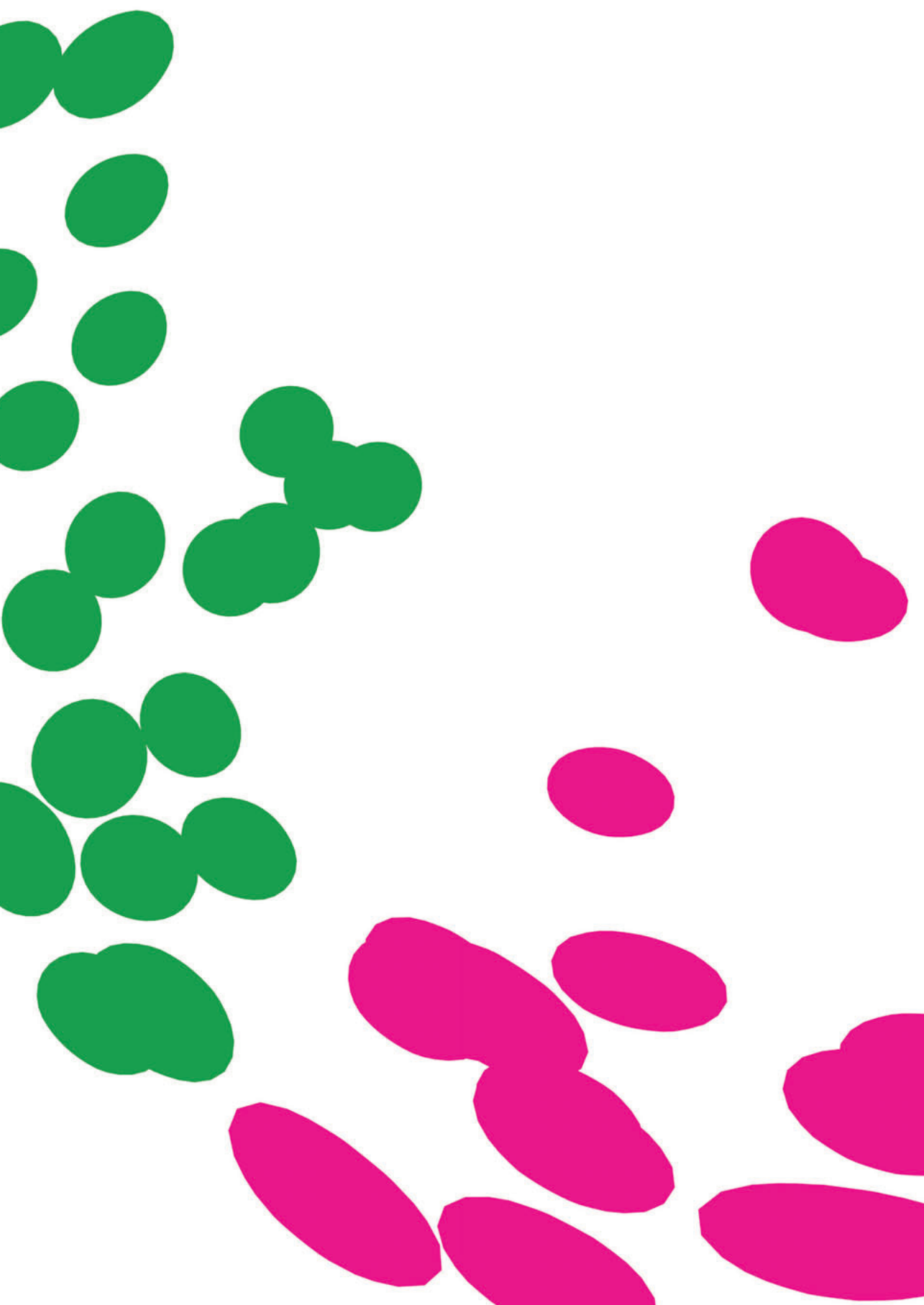
- 
434. Sautto, G. A., Kirchenbaum, G. A. & Ross, T. M. Towards a universal influenza vaccine: different approaches for one goal. *Virol. J.* **15**, 17 (2018).
  435. Hampson, A. et al. Improving the selection and development of influenza vaccine viruses - Report of a WHO informal consultation on improving influenza vaccine virus selection, Hong Kong SAR, China, 18-20 November 2015. *Vaccine* **35**, 1104–1109 (2017).
  436. Jagadesh, A., Salam, A. A. A., Mudgal, P. P. & Arunkumar, G. Influenza virus neuraminidase (NA): a target for antivirals and vaccines. *Arch. Virool.* **161**, 2087–2094 (2016).
  437. Treanor, J. J. Prospects for Broadly Protective Influenza Vaccines. *Am. J. Prev. Med.* **49**, S355–63 (2015).
  438. Johansson, B. E. & Cox, M. M. J. Influenza viral neuraminidase: the forgotten antigen. *Expert Rev. Vaccines* **10**, 1683–1695 (2011).
  439. Eichelberger, M. C., Morens, D. M. & Taubenberger, J. K. Neuraminidase as an influenza vaccine antigen: a low hanging fruit, ready for picking to improve vaccine effectiveness. *Curr. Opin. Immunol.* **53**, 38–44 (2018).
  440. Krammer, F. et al. NAction! How Can Neuraminidase-Based Immunity Contribute to Better Influenza Virus Vaccines? *MBio* **9**, (2018).
  441. Rajendran, M. et al. Analysis of Anti-Influenza Virus Neuraminidase Antibodies in Children, Adults, and the Elderly by ELISA and Enzyme Inhibition: Evidence for Original Antigenic Sin. *MBio* **8**, (2017).
  442. Tan, G. S. et al. Characterization of a broadly neutralizing monoclonal antibody that targets the fusion domain of group 2 influenza A virus haemagglutinin. *J. Virool.* **88**, 13580–13592 (2014).
  443. Yuan, J. et al. Origin and molecular characteristics of a novel 2013 avian influenza A(H6N1) virus causing human infection in Taiwan. *Clin. Infect. Dis.* **57**, 1367–1368 (2013).
  444. Gao, R. et al. Human infection with a novel avian-origin influenza A (H7N9) virus. *N. Engl. J. Med.* **368**, 1888–1897 (2013).
  445. Wang, X. et al. Epidemiology of avian influenza A H7N9 virus in human beings across five epidemics in mainland China, 2013–17: an epidemiological study of laboratory-confirmed case series. *Lancet Infect. Dis.* **17**, 822–832 (2017).

446. Prevato, M. et al. Expression and Characterization of Recombinant, Tetrameric and Enzymatically Active Influenza Neuraminidase for the Setup of an Enzyme-Linked Lectin-Based Assay. *PLoS One* **10**, e0135474 (2015).
447. Schotsaert, M. et al. Long-Lasting Cross-Protection Against Influenza A by Neuraminidase and M2e-based immunization strategies. *Sci. Rep.* **6**, 24402 (2016).
448. Margine, I., Palese, P. & Krammer, F. Expression of functional recombinant haemagglutinin and neuraminidase proteins from the novel H7N9 influenza virus using the baculovirus expression system. *J. Vis. Exp.* e51112 (2013). doi:10.3791/51112
449. Schmidt, P. M., Attwood, R. M., Mohr, P. G., Barrett, S. A. & McKimm-Breschkin, J. L. A generic system for the expression and purification of soluble and stable influenza neuraminidase. *PLoS One* **6**, e16284 (2011).
450. Kosik, I. & Yewdell, J. W. Influenza A virus haemagglutinin specific antibodies interfere with virion neuraminidase activity via two distinct mechanisms. *Virology* **500**, 178–183 (2017).
451. Kingstad-Bakke, B., Kamlangdee, A. & Osorio, J. E. Mucosal administration of raccoonpox virus expressing highly pathogenic avian H5N1 influenza neuraminidase is highly protective against H5N1 and seasonal influenza virus challenge. *Vaccine* **33**, 5155–5162 (2015).
452. Lei, H. et al. *Lactococcus lactis* displayed neuraminidase confers cross protective immunity against influenza A viruses in mice. *Virology* **476**, 189–195 (2015).
453. Prevato, M. et al. An Innovative Pseudotypes-Based Enzyme-Linked Lectin Assay for the Measurement of Functional Anti-Neuraminidase Antibodies. *PLoS One* **10**, e0135383 (2015).
454. Richard, M., de Graaf, M. & Herfst, S. Avian influenza A viruses: from zoonosis to pandemic. *Future Virol.* **9**, 513–524 (2014).
455. Meijer, A. et al. Case of seasonal reassortant A(H1N2) influenza virus infection, the Netherlands, March 2018. *Euro Surveill.* **23**, (2018).
456. Chen, H. et al. Clinical and epidemiological characteristics of a fatal case of avian influenza A H10N8 virus infection: a descriptive study. *Lancet* **383**, 714–721 (2014).
457. Lopez-Martinez, I. et al. Highly pathogenic avian influenza A(H7N3) virus in poultry workers, Mexico, 2012. *Emerg. Infect. Dis.* **19**, 1531–1534 (2013).

- 
458. Koopmans, M. et al. Transmission of H7N7 avian influenza A virus to human beings during a large outbreak in commercial poultry farms in the Netherlands. *Lancet* **363**, 587–593 (2004).
459. Fonville, J. M. et al. Antibody landscapes after influenza virus infection or vaccination. *Science* **346**, 996–1000 (2014).
460. Centers for Disease Control and Prevention. Vaccine Effectiveness - How Well Does the Flu Vaccine Work? Centers for Disease Control and Prevention (CDC) (2017). Available at: <https://www.cdc.gov/flu/about/qa/vaccineeffect.htm>. (Accessed: 26th May 2018)
461. Chang, M.-H. & Chen, D.-S. Prevention of hepatitis B. *Cold Spring Harb. Perspect. Med.* **5**, a021493 (2015).
462. Kjaer, S. K. et al. A 12-Year Follow-up on the Long-Term Effectiveness of the Quadrivalent Human Papillomavirus Vaccine in 4 Nordic Countries. *Clin. Infect. Dis.* **66**, 339–345 (2018).
463. Wutzler, P. et al. Varicella vaccination - the global experience. *Expert Rev. Vaccines* **16**, 833–843 (2017).
464. McLean, H. Q., Fiebelkorn, A. P., Temte, J. L., Wallace, G. S. & Centers for Disease Control and Prevention. Prevention of measles, rubella, congenital rubella syndrome, and mumps, 2013: summary recommendations of the Advisory Committee on Immunization Practices (ACIP). *MMWR Recomm. Rep.* **62**, 1–34 (2013).
465. Easterbrook, J. D. et al. Immunization with 1976 swine H1N1- or 2009 pandemic H1N1-inactivated vaccines protects mice from a lethal 1918 influenza infection. *Influenza Other Respi. Viruses* **5**, 198–205 (2011).
466. Wodal, W. et al. A cell culture-derived whole-virus H9N2 vaccine induces high titer antibodies against haemagglutinin and neuraminidase and protects mice from severe lung pathology and weight loss after challenge with a highly virulent H9N2 isolate. *Vaccine* **30**, 4625–4631 (2012).
467. Chen, Z., Kim, L., Subbarao, K. & Jin, H. The 2009 pandemic H1N1 virus induces anti-neuraminidase (NA) antibodies that cross-react with the NA of H5N1 viruses in ferrets. *Vaccine* **30**, 2516–2522 (2012).
468. Ohmit, S. E., Petrie, J. G., Cross, R. T., Johnson, E. & Monto, A. S. Influenza haemagglutination-inhibition antibody titer as a correlate of vaccine-induced protection. *J. Infect. Dis.* **204**, 1879–1885 (2011).

469. Williams, T. L., Pirkle, J. L. & Barr, J. R. Simultaneous quantification of haemagglutinin and neuraminidase of influenza virus using isotope dilution mass spectrometry. *Vaccine* **30**, 2475–2482 (2012).
470. Wan, H., Sultana, I., Couzens, L. K., Mindaye, S. & Eichelberger, M. C. Assessment of influenza A neuraminidase (subtype N1) potency by ELISA. *J. Virol. Methods* **244**, 23–28 (2017).
471. Kuck, L. R. et al. VaxArray for haemagglutinin and neuraminidase potency testing of influenza vaccines. *Vaccine* **36**, 2937–2945 (2018).
472. Russell, C. A. et al. The potential for respiratory droplet-transmissible A/H5N1 influenza virus to evolve in a mammalian host. *Science* **336**, 1541–1547 (2012).
473. Chao, D. L. Modeling the global transmission of antiviral-resistant influenza viruses. *Influenza Other Respi. Viruses* **7**, 58–62 (2013).
474. Dobrovolny, H. M. & Beauchemin, C. A. A. Modelling the emergence of influenza drug resistance: The roles of surface proteins, the immune response and antiviral mechanisms. *PLoS One* **12**, e0180582 (2017).
475. Leung, K., Lipsitch, M., Yuen, K. Y. & Wu, J. T. Monitoring the fitness of antiviral-resistant influenza strains during an epidemic: a mathematical modelling study. *Lancet Infect. Dis.* **17**, 339–347 (2017).
476. Neher, R. A., Bedford, T., Daniels, R. S., Russell, C. A. & Shraiman, B. I. Prediction, dynamics, and visualization of antigenic phenotypes of seasonal influenza viruses. *Proc. Natl. Acad. Sci. U. S. A.* **113**, E1701–9 (2016).
477. Nolte, N., Kurzawa, N., Eils, R. & Herrmann, C. MapMyFlu: visualizing spatio-temporal relationships between related influenza sequences. *Nucleic Acids Res.* **43**, W547–51 (2015).
478. Neher, R. A. & Bedford, T. nextflu: real-time tracking of seasonal influenza virus evolution in humans. *Bioinformatics* **31**, 3546–3548 (2015).
479. Bedford, T. & Neher, R. A. Seasonal influenza circulation patterns and projections for 2017–2018. *bioRxiv* (2017).







## CHAPTER 10

# Nederlandse samenvatting



## NEDERLANDSE SAMENVATTING

Elke winter krijgt ongeveer vijf tot vijftien procent van de wereldbevolking te maken met de zeer besmettelijke acute luchtweginfectie genaamd 'influenza'. Dit jaarlijkse verschijnsel wordt 'seizoensgriep', 'wintergriep' of 'epidemische griep' genoemd en resulteert wereldwijd in circa drie tot vijf miljoen ziekenhuisopnames en 290.000 tot 650.000 doden per jaar. Epidemische griep wordt veroorzaakt door de influenza A- en B-virussen.

Influenza A- en B-virussen behoren tot de Orthomyxoviridae-familie en hebben een gesegmenteerd enkelstrengs negatief RNA-genoom. Dit genoom, dat de genetische samenstelling het virus bevat, bestaat uit acht gensegmenten. Deze acht gensegmenten coderen voor basisch polymerase 2 (PB2), basisch polymerase 1 (PB1), zuur polymerase (PA), haemagglutinine (HA), nucleoproteïne (NP), neuraminidase (NA), matrix (M) en niet-structureel eiwit (NS). Door de opdeling van het genoom in gensegmenten is het mogelijk dat deze gensegmenten, tijdens een gelijktijdige infectie van een gastheer tussen verschillende influenzavirussen, kunnen worden uitgewisseld. Dit proces heet reassortering. Reassortering tussen varkens-, vogel- en/of humane influenza A-virussen hebben geleid tot verschillende pandemieën. Deze pandemische virussen veroorzaken vervolgens de jaarlijkse epidemische griep bij mensen.

Influenza A-virussen worden onderverdeeld in subtypen op basis van de antigene eigenschappen van de oppervlakte-eiwitten HA en NA. Tot op heden zijn er achttien HA en elf NA subtypen in de natuur gevonden. Subtypen H1N1 (A(H1N1)) en H3N2 (A(H3N2)) zijn, tezamen met influenza B-virussen, momenteel de veroorzakers van epidemische griep. De A(H3N2) influenzavirussen zijn de meest voorkomende en virulente van de twee influenza A-subtypen en vertonen de sterkste antigene evolutie.

Antigene drift ontstaat doordat het virus ontsnapt aan de neutraliserende werking van antistoffen die door eerdere infecties of vaccinaties in de bevolking zijn opgewekt. Deze antistoffen herkennen en binden antigene regio's op het HA of NA. Aminozuurveranderingen in deze antigene regio's kunnen bijdragen aan antigene variatie en zorgen uiteindelijk voor antigene drift. Griepvaccins zijn effectief tegen epidemische griep, maar moeten vaak worden aangepast vanwege deze antigene drift. Door middel van influenzasurveillance, waarbij de verspreiding van influenzavirussen continu wordt gevolgd zodat virussen die antigene drift hebben ondergaan vroegtijdig worden gedetecteerd, kunnen de meest geschikte virusstammen worden geïdentificeerd om in griepvaccins te gebruiken.

De antigene evolutie van de HA-eiwitten van A(H3N2) influenzavirussen is in kaart gebracht vanaf de introductie van het virus in de mens in 1968, tot en met 2003. In deze antigene kaart staat de afstand tussen twee virussen gelijk aan in welke mate de virussen antigeen overeenkomen: hoe groter de afstand, des te meer twee virussen verschillen en hoe kleiner de afstand, des te meer twee virussen op elkaar lijken. De antigene kaart toont aan dat de evolutie voor HA niet geleidelijk, maar sprongsgewijs optreedt. Tevens vormt de antigene kaart zich in een patroon van elf clusters, genoemd naar de eerste griepvaccinstam in het cluster: Hong Kong 1968 (HK68), Engeland 1972 (EN72), Victoria 1975 (VI75), Texas 1977 (TX77), Bangkok 1979 (BK79), Sichuan 1987 (SI87), Beijing 1989 (BE89), Beijing 1992 (BE92), Wuhan 1995 (WU95), Sydney 1997 (SY97) en Fujian 2002 (FU02). Elk van deze antigene clusters bevat virussen die enige tijd antigeen vergelijkbaar blijven, waarna virussen antigeen veranderen en een nieuw cluster vormen. Voor iedere 'clusterovergang' moet ten minste één griepvaccinaanpassing gemaakt worden. De antigene clusters bleven gemiddeld 3,3 jaar dominant en virussen werden geobserveerd tot 2 jaar voor en tot 2 jaar na de periode waarin dat antigene cluster dominant was.

Er bestaat nog steeds veel onzekerheid over de onderliggende mechanismen die de antigene drift van influenzavirussen reguleren en er zijn verschillende theorieën beschreven over deze onderliggende mechanismen. Echter beschouwen deze theorieën het primaire oppervlakte-eiwit 'HA' als de belangrijkste drijvende kracht, terwijl het secundaire oppervlakte-eiwit 'HA' en de andere virale eiwitten, grotendeels zijn genegeerd.

HA verkrijgt virale toegang tot de gastheercel door zich te binden aan de sialzuren op het oppervlak van de gastheercel. Vervolgens wordt het HA-molecuul 'HA0' proteolytisch geknipt in de subeenheden 'HA1' en 'HA2', waardoor het virale membraan kan fuseren met het membraan van de gastheercel. NA is een enzym dat sialidase-activiteit, ofwel receptorvernietigende-activiteit, bezit. NA gebruikt deze activiteit om de sialzuren op het oppervlak van de gastheercel te knippen, waardoor nieuw gevormde virusdeeltjes loskomen van de cel en een volgende cel kunnen infecteren. Een virusdeeltje heeft gemiddeld dertig tot veertig NA-tetrameren en 290-300 HA-trimeren. Elke identieke subeenheid van de NA-tetrameer bevat het actief centrum van het enzym. Antigene drift is niet alleen een kenmerk van HA, maar ook van NA. Hoewel antistoffen tegen NA de infectie niet kunnen voorkomen, tonen preklinische en klinische onderzoeken aan dat NA-specifieke immuniteit de ernst van de ziekte kan

verminderen. Antistoffen gericht tegen NA remmen het loskomen en de verspreiding van nieuw gevormde virusdeeltjes uit geïnfecteerde cellen.

In **hoofdstuk 2** hebben we de evolutie van NA-subtype 2 (N2) bestudeerd op basis van 291 virussen geïsoleerd tussen 1968 en 2009. De mutatiesnelheid van NA was hoog; naar schatting  $3,15 \times 10^{-3}$  nucleotideveranderingen per nucleotidepositie per jaar. De mate van natuurlijke selectie op het volledige NA-gensegment is geschat door te kijken naar verhouding van nucleotideveranderingen die in een aminozuurverandering resulteren ( $d_N$ ) en nucleotideveranderingen die niet in een aminozuurverandering resulteren ( $d_S$ ) ( $d_N/d_S$ -ratio). Bij een  $d_N/d_S$ -ratio van 1 heeft er geen selectie plaatsgevonden. Als de  $d_N/d_S$ -ratio lager is dan 1, kan er worden gesproken van negatieve selectie; als de  $d_N/d_S$ -ratio hoger is dan 1, wordt er gesproken van positieve selectie. Het NA-gen had een sterke selectie ( $d_N/d_S = 0,249$ ), wat aantoont dat er een overmaat aan nucleotideveranderingen waren die in een aminozuurverandering resulteerde en uiteindelijk zijn verwijderd door negatieve selectie. Kristallografische studies van NA toonden aan dat het sterk geconserveerde actief centrum van het enzym wordt omringd door antigene regio's. Tot op heden is er van de antigene regio's van NA beperkte kennis. Een methode om mogelijke veranderingen die de ontsnapping aan de neutraliserende werking van antistoffen veroorzaken te onderzoeken, is om positieve selectie per aminozuurpositie te bestuderen. Op basis hiervan zijn een totaal van zestien aminozuurposities, waarvoor positief werd geselecteerd, gevonden. Slechts zes van deze aminozuurposities — 199, 328, 334, 338, 367 en 370 — bevinden zich binnen bekende antigene regio's.

Vervolgens hebben we de evolutie van NA met die van HA vergeleken. Daarbij werd duidelijk dat er asynchrone evolutie tussen de twee genen is. Reassortering van gensegmenten van humane influenza A-virussen maakt genetische diversificatie mogelijk en heeft daarnaast bijgedragen aan nieuwe varianten van de epidemische influenzavirussen. Met behulp van fylogenetische analyses waren we in staat om verscheidene momenten, waarbij reassortering heeft plaatsgevonden, te identificeren. Deze werden met name tijdens de co-circulatie van BE92- en WU95-achtige virussen waargenomen. De meeste reassorteringen vonden plaats *binnen* antigene clusters, in plaats van tussen antigene clusters. Twee reassortering-incidenten hadden virussen tot resultaat die gedurende een langere periode in de menselijke populatie circuleerden, maar voor de meeste reassortanten was dit niet het geval. Dit suggereert dat reassortering vaak geen invloed heeft op, of zelfs nadelig is voor, het virus.

De gemiddelde mutatiesnelheid voor HA1, het immunogene gedeelte, was hoger in vergelijking met die van NA ( $5,15 \times 10^{-3}$  versus  $3,15 \times 10^{-3}$  nucleotideveranderingen per

nucleotidepositie per jaar). HA1 had een sterke selectie ( $d_N/d_S = 0,362$ ). Een totaal van 21 positief-geselecteerde aminozuurposities voor HA1 werden gevonden, voornamelijk in de antigene regio's van HA. Een studie uit 2013 toont aan dat voor HA van A(H3N2) influenzavirussen alle veranderingen met antigene variatie als resultaat, plaatsvinden op zeven aminozuurposities, onmiddellijk grenzend aan de receptor bindingsplaats: aminozuurpositie 145, 155, 156, 158, 159, 189 en 193. Met uitzondering van aminozuurpositie 158 werden alle aminozuurposities gedetecteerd door onze analyses. Het is aannemelijk dat ten minste enkele van de andere positief- of negatief-geselecteerde aminozuurposities die we hebben gedetecteerd, bijgedragen hebben aan de efficiëntie waarmee dit virus zich door de populatie verspreid.

De dataset gebruikt in **hoofdstuk 2** werd uitgebreid tot en met het griepseizoen 2010/2011 en vervolgens gebruikt om de volledige genomen van 286 A(H3N2) influenzavirussen te analyseren in **hoofdstuk 3**. Dit gaf de mogelijkheid om acht gensegmenten en vijftien eiwitten te bestuderen. De gemiddelde mutatiesnelheden op nucleotideniveau van de afzonderlijke gensegmenten varieerden van  $2,07 \times 10^{-3}$  tot en met  $3,99 \times 10^{-3}$  nucleotideveranderingen per nucleotidepositie per jaar. Hiervan was de hoogste mutatiesnelheid voor HA en NA en de laagste mutatiesnelheid voor het M-segment. De HA1-subeenheid vertoonde de hoogste mutatiesnelheid op aminozuurniveau in aminozuurveranderingen per aminozuurpositie per jaar, terwijl de mutatiesnelheid voor de HA2-subeenheid veel lager was ( $14,9 \times 10^{-3}$  in vergelijking met  $1,4 \times 10^{-3}$  aminozuurveranderingen per aminozuurpositie per jaar). Dit is waarschijnlijk te wijten aan het feit dat HA1 het belangrijkste immunogene deel van HA is. PB1-frame 2 (PB1-F2) en NA vertoonden ook hoge mutatiesnelheden op aminozuurniveau, voor alle andere eiwitten waren deze ten minste driemaal lager.

Ons onderzoek toont aan dat reassortering met name tijdens de circulatie van BE92- en WU95-achtige virussen plaatsvond. Dit is in overeenstemming met **hoofdstuk 2**. Verschillende van deze gereassorteerde virussen waren blijvend in de populatie (10/59), meestal in recentere jaren. Dit suggereert dat reassortering geholpen heeft bij het vergroten van de genetische diversiteit van het virus, om de conditie te verbeteren of om aan de immuniteit van de populatie te ontsnappen. Echter, in de meeste gevallen resulteerde reassortering niet in virussen die gedurende langere tijd in de humane populatie circuleerden.

We hebben positieve en negatieve selectie van alle eiwitten onderzocht met behulp van data-analyse. HA, NA, M-proteïne 2 (M2) en NS-eiwit 1 (NS1) vertoonden relatief hoge  $d_N/d_S$ -ratio's in vergelijking met die van de resterende eiwitten. HA bevatte

het grootste aantal positief-geselecteerde aminozuurposities, het meest prominent aanwezig in de antigene regio's. Voor NA werden negentien aminozuurposities positief-geselecteerd. Gezien het feit dat zowel HA als NA een belangrijk doelwit zijn van antistoffen, waren ten minste enkele van de positief-geselecteerde aminozuurposities waarschijnlijk gerelateerd aan het ontsnappen aan de neutraliserende werking van antistoffen.

Mede door het gebrek aan geschikte serologische NA testen is er weinig onderzoek gedaan naar antigene variatie van NA. Antigene variatie van HA wordt frequent bestudeerd met behulp van haemagglutineringsremmingstest (HAR)-titers. Met de HAR wordt de virus-neutraliserende werking van antistoffen in antisera op HA gemeten. Er kan worden bepaald in welke mate de antigene eigenschappen van de huidige virussen overeenkomen met die van eerdere virussen. Doorgaans wordt fretten-antiserum gebruikt omdat de infectie een vergelijkbaar verloop heeft in fretten als in de mens. Op vergelijkbare wijze worden NA-remming (NAR)-titers gemeten om de antigene evolutie van NA te bestuderen. Het testen van wild-type virussen met antisera in serologisch NA-testen is gecompliceerd omdat in het antiserum zich zowel antistoffen bevinden die binden aan HA als antistoffen die binden aan NA. De antistoffen die aan HA binden, vertonen sterische hindering in de NAR waardoor het lijkt alsof er sterke NA-remming is. Om NA-specifieke remming te meten worden doorgaans reassortante virussen met een niet-overeenkomend HA gebruikt die de invloed van anti-HA-antistoffen op de serologische NA-testen verminderen. Traditiegetrouw werd de thiobarbituurzuur (TBA)-test gebruikt om de NAR te meten. Echter, om deze test uit te voeren, worden schadelijke chemicaliën gebruikt. Om het gebruik hiervan te verminderen is een miniTBA-test opgezet. In **hoofdstuk 4** hebben we de antigene evolutie van HA en NA van A(H1N1) en A(H3N2) influenzavirussen, die in de griepvaccinsamenstelling tussen 1995 en 2010 werden gebruikt, geëvalueerd door analyse van HAR- en NAR-titers en antigene cartografie. De NAR-titers zijn verkregen door middel van de miniTBA-test waarbij gebruik is gemaakt van fretten-antiserum in combinatie met reassortante virussen met een niet-overeenkomend HA (H6) als antigeen. Zoals eerder aangetoond voor HA leidden aminozuurveranderingen in NA niet altijd tot een antigene verandering. Net zoals voor HA observeerden we voor NA verschillende antigene clusters. Voor zowel het H1N1- als het H3N2-subtype liepen de clusterovergangen van NA niet synchroon aan die van HA. Daarnaast tonen we aan dat antistoffen in humane antisera de virussen op een vergelijkbare wijze herkennen als de antistoffen in fretten-antiserum. Opmerkelijk is dat een enkele verandering in NA van het A(H1N1) influenzavirus, gebruikt in de griepvaccinsamenstelling tussen 2008 en

2010, genoeg is om het virus antigeen te veranderen, waardoor antistoffen gericht tegen virussen van voor 2007 dit virus niet meer herkenden. Deze gegevens onderstrepen het belang van de NAR om antigene variatie te definiëren wanneer er sequentieveranderingen zijn in NA.

Omdat er ook bij de miniTBA-test nog schadelijke chemicaliën worden gebruikt, is deze test niet geschikt voor routinematige serologie. Een enzym-gebonden lectine test (ELLA) is een praktisch alternatief (zonder schadelijke chemicaliën) voor het meten van NAR-titers. Eerst wordt het antigeen, heel virus of (gezuiverd) NA, getitreerd met de ELLA. Dit wordt bewerkstelligd door een microtiterplaat te coaten met 'fetuine' (een substraat voor NA) gevolgd door de toevoeging van het antigeen in een vooraf bepaalde verdunningsreeks. Actief NA knipt het terminale siaalzuur-residu van de fetuine, resulterend in een blootgesteld galactose-residu. Vervolgens wordt pinda-agglutinine, gelabeld aan mierikswortelperoxidase (PNA-HRPO), toegevoegd, waarna PNA-HRPO bindt aan de terminale galactose. De intensiteit van het signaal na toevoeging van het substraat is afhankelijk van de hoeveelheid terminale siaalzuur-residuen die zijn geknipt door NA en dus NA-activiteit. In de NAR-ELLA zal binding van NA door specifieke antistoffen de enzymatische functie van NA remmen, resulterend in een vermindering van de hoeveelheid terminale siaalzuur-residuen die zijn geknipt door NA en derhalve in een vermindering van het signaal. In **hoofdstuk 5** hebben we de ELLA geoptimaliseerd voor routinematige analyse van humane antisera met behulp van reassortante virussen met een niet-overeenkomend HA (H6) als antigeen, om niet-specifieke remming door H1- en H3-specifieke antistoffen in humane antisera te voorkomen. De geoptimaliseerde ELLA is subtype-specifiek en reproduceerbaar. Terwijl de waardes gemeten met ELLA enigszins groter zijn dan die gemeten door de miniTBA-test, zijn de seroconversiesnelheden gelijk, wat suggereert dat de gevoeligheid van deze twee testen overeenkomt onder deze geoptimaliseerde omstandigheden.

Om te voorkomen dat HA-specifieke antistoffen storing veroorzaken wordt worden de meeste NAR-testen uitgevoerd met reassortante virussen die een niet-overeenkomend HA bevatten. Het genereren van deze virussen is echter tijdrovend en dit maakt het ongeschikt voor grootschalige influenzasurveillance. In **hoofdstuk 6** werd de ELLA geoptimaliseerd voor snelle antigene karakterisering van de NA met behulp van wild-type virussen. In plaats van reassortante virussen te gebruiken, zijn wild-type A(H3N2) influenzavirussen gebruikt als antigeen in combinatie met fretten-antisera. Deze fretten-antisera werden opgewekt tegen reassortante virussen met niet-overeenkomend HA. Deze reassortante A(H7N2) influenzavirussen bevatten de NA van A(H3N2)

influenzavirussen en de overige zeven gensegmenten van een A(H7N7) influenzavirus. Hierdoor zullen de HA-specifieke antistoffen die aanwezig zijn in de fretten-antiseren niet kruisreageren met wild-type A(H3N2) influenzavirussen in de NAR-test. Door de behandeling met een kleine hoeveelheid sialidase kon, zonder HAR- en NAR-titers van fretten-antiseren te beïnvloeden, de niet-specifieke remming van antiseren worden verwijderd. De geoptimaliseerde NAR-ELLA is reproduceerbaar, zeer gevoelig en bruikbaar voor influenzasurveillance om antigene veranderingen van NA te monitoren.

In **hoofdstuk 7** hebben we de antigene evolutie van N2 vanaf de introductie in mensen in 1957 tot en met 2012 in kaart gebracht. Dit is bewerkstelligd met behulp van de NAR-ELLA in combinatie met (H2N2) en A(H3N2) influenzavirussen en fretten-antiseren gericht tegen reassortante A(H7N2) influenzavirussen zoals beschreven in **hoofdstuk 6**. De resultaten tonen duidelijke antigene evolutie aan, dat wil zeggen: veel herkenning van antigenen door antiseren uit hetzelfde isolatiejaar en weinig herkenning van late antigenen door vroege antiseren en vice versa. Vervolgens is antigene cartografie gebruikt om de NAR-titers in een antigene kaart te visualiseren. De NA-antigene kaart toont aan dat de antigenen in zes antigenen clusters zijn waargenomen, genoemd naar de eerste griepvaccinstam van het cluster dat nog niet was gebruikt om voor een HA-antigene cluster; Netherlands 1957 (NL57, H2N2-virus), Port Chalmers 1973 (PC73), Philippines 1982 (PH82), Guizhou 1989 (GU89), Nanchang 1995 (NA95) en New York 2004 (NY04). Gemiddeld circuleerden antigenen NA-clusters langer dan HA-clusters.

Op basis van de NA-clusters en sequentieanalyse zijn de verschillende aminozuurveranderingen geïdentificeerd die mogelijk de NA-antigeniciteit beïnvloeden. Voor HA van A(H3N2) influenzavirussen kwamen de veranderingen die de clusterovergangen veroorzaakten voor op zeven aminozuurposities, direct grenzend aan de receptor bindingsplaats. Voor NA zijn zes van de 46 aminozuurveranderingen die variëren per cluster — 153, 155, 197, 199, 221 en 370 — direct grenzend aan het actief centrum. Deze aminozuurposities bevinden zich aan de oppervlakte van het eiwit, en zijn dus potentieel toegankelijk voor antilichamen. Tevens zijn twee van deze aminozuurveranderingen die variëren per cluster — 199 en 370 — betrokken bij twee van de clusterovergangen in de antigene kaart van NA. Ook werden aminozuurposities 199 en 370 positief geselecteerd in de data-analyses van **hoofdstuk 2** en **3**. Deze bevindingen wijzen op een belangrijke rol voor deze aminozuurposities in de antigene evolutie van N2.

Om de in kaart gebrachte antigene evolutie van NA te kunnen vergelijken met die van HA, hebben we naast de antigene evolutie van 1968 tot 2003, die al bekend was, ook naar de antigene evolutie van HA vanaf 2003 tot en met 2011 gekeken. Gedurende



deze periode circuleerde twee nieuwe antigene HA-clusters, California 2004 (CA04) en Perth 2009 (PE09), wat het totaal van antigene clusters tussen 1968 en 2011 op dertien brengt. Voor HA blijven de antigene clusters gemiddeld 3,3 jaar dominant. Dit lijkt een lange periode als men de hoge mutatiesnelheid van influenzavirus in beschouwing neemt, in het bijzonder voor HA. Daarnaast zijn enkele aminozuurveranderingen vaak voldoende om antigene clusterovergangen in HA te veroorzaken. Zoals beschreven in eerdere studies met veel kleinere datasets, toont ons onderzoek aan dat de timing van antigene evolutie van HA en NA ongelijk is. We laten zien dat de clusterovergangen van NA en HA niet synchroon verlopen en elkaar zelfs afwisselen. Dit patroon is echter bijna afwezig bij het vergelijken van de genetische veranderingen van de aminozuursequenties van HA en NA. De tegenovergestelde rollen van HA en NA tijdens influenzavirusreplicatie vereisen een functioneel evenwicht tussen HA en NA. Het is nog onbekend hoe dit evenwicht wordt beïnvloed door de asynchrone antigene evolutie.

Ons werk toont duidelijk aan dat er antigene variatie is van NA. Het is dus belangrijk om te zorgen dat er, net zoals voor HA, gekeken wordt welk NA opgenomen zou moeten worden voor het ontwikkelen van het griepvaccin. Hiernaast is het ook van belang dat de hoeveelheid NA in het griepvaccin gestandaardiseerd wordt. Om deze NA's te selecteren, moet de huidige influenzasurveillance strategie worden gewijzigd. Momenteel wordt er alleen een genetische karakterisering van het NA uitgevoerd, maar het is belangrijk om dit te combineren met serologische testen die het antigene fenotype van het NA van deze virussen kunnen bepalen. Daarnaast is het cruciaal om te onderzoeken welke aminozuurposities veranderingen in het antigene fenotype van NA veroorzaken. Het ontrafelen van deze aminozuurposities kan helpen om meer gerichte influenzasurveillance uit te voeren. Influenzasurveillance waarbij sequentiegegevens worden geïntegreerd met epidemiologische, geografische en antigene gegevens kan helpen inzicht te krijgen in de evolutie en epidemiologie van influenzavirussen, om uiteindelijk de evolutie van griep te kunnen helpen voorspellen. Het opnemen van gerichte NA-sequentiebepaling, het meten van NAR-titers in routinematige influenzasurveillance en de standaardisatie van NA in bestaande griepvaccins zullen de volgende generatie griepvaccins verbeteren.



## CHAPTER 11

# Dankwoord + About the Author



## ABOUT THE AUTHOR

### Curriculum Vitae

Kim Birgitta Westgeest was born on May 28, 1982 in The Hague and grew up in Wassenaar, the Netherlands. In 2000, she obtained her HAVO diploma from Adelbert College (Wassenaar) and she studied Laboratory Science at Hogeschool Larenstein (Velp) from 2002 up and until 2006. Her first internship was at the department of Biomolecular Detection (RIKILT, Wageningen) where she developed an ELISA for the detection of soy proteins in food products under the supervision of Dr. M. Bremer. She stayed at the department for a holiday job as a research technician. Kim performed her second internship in the group of Prof. Dr. P. Quax at the department of Vascular and Metabolic diseases (TNO, Leiden). Here she studied silencing of NF- $\kappa$ B activity via RNAi, aiming at diminishing the re-narrowing of the arterial wall after therapeutic interventions under supervision of S. Droog and J. Grimbergen. She received her Bachelor of Applied Science degree in 2006. Hereafter, Kim worked as a research technician in the team of Dr. F. Schuren at the department of Microbiology – Genomics (TNO, Zeist). In 2007 she started the Master of Science programme Medical Biotechnology at Wageningen University. Her MSc research internship was in the group of Prof. Dr. Ron Fouchier at the department of Viroscience (Erasmus MC, Rotterdam). Here she focussed on generating a plasmid-driven influenza A/PR/8/34 virus-like particle system to investigate the process of genome packaging under supervision of M.I. Spronken and Prof. Dr. R.A.M. Fouchier. She obtained her master's degree in 2009 and continued as a PhD student under supervision of Prof. Dr. Ron Fouchier and Dr. Miranda de Graaf in the department of Viroscience. Her PhD research focussed on the evolution of human influenza A (H3N2) viruses circulating between 1968 and 2011, with special emphasis on neuraminidase. The results of these studies are presented in this thesis.

**PhD portfolio**

Name: Kim Birgitta Westgeest  
 Department: Viroscience, Erasmus MC  
 Graduate school: The Erasmus Postgraduate school Molecular Medicine (MolMed)  
 PhD period: 2009-2019  
 Promotor: Prof. dr. Ron A. M. Fouchier  
 Co-promotor: Dr. Miranda de Graaf

*Education*

2009-2019	PhD program	Erasmus MC
2007-2009	Master of Science (Medical Biotechnology)	Wageningen University
2002-2006	Bachelor of Laboratory Science (Toxicology/Biotechnology)	Hogeschool Larenstein

*In-depth courses, certificates*

2016	Basiskwalificatie Didactische Bekwaamheid	Hogeschool Leiden
2016	MGC Next Generation Sequencing data analysis	LUMC, Leiden
2016	Basiskwalificatie Examinering	ICLON, Leiden
2012	Biomedical English Writing and Communication	Erasmus MC
2010	16th International Bioinformatics Workshop.	Hopkins, US
2010	Basic Course on 'R'	MolMed
2010	Workshop on Photoshop and Illustrator CS4	MolMed
2009	15th International Bioinformatics Workshop	Erasmus MC
2009	Course Molecular Diagnostics IV	MolMed
2009	Get Out of Your Lab Days 2009!	MolMed
2008	Course in phylogeny	Dep. Viroscience
2008	Course in Virology	MolMed & Viroscience
2008	Advanced Immunology	MolMed

*Poster presentations*

2013	Options for the Control of Influenza VIII	Cape Town, ZA
2012	6th Orthomyxovirus Research Conference	Quebec, CA
2013	17th MolMed Day	Rotterdam
2012	16th MolMed Day	Rotterdam
2011	15th MolMed Day	Rotterdam
2010	16th International Bioinformatics Workshop	Rockville, US
2010	14th MolMed Day	Rotterdam
2009	15th International Bioinformatics Workshop	Rotterdam

*Oral presentations*

2013	Invited speaker group Dr. M. van Boven	RIVM, Bilthoven
2011	Fourth ESWI Influenza Conference.	Malta
2009-2014	Viroscience department seminars (twice a year)	Erasmus MC

*Attended congresses, symposia and meetings*

2010	Dutch annual virology symposium	Amsterdam
2009	Virology meeting	Dakar, SN
2009	Dutch annual virology symposium	Amsterdam
2009	13th MolMed Day	Rotterdam
2009-2014	Department of Zoology research meetings	Cambridge, UK
2008	12th MolMed Day	Rotterdam
2008-2014	Viroscience department seminars (twice, weekly)	Erasmus MC
2008-2014	Viroscience labmeetings (weekly)	Erasmus MC

**Teaching**

2019-present	Tutor Life Sciences	EUR
2015-2017	Teacher bio-informatics	HSLeiden
2013	Assisted in the practical course on phylogenetics for the Research Master programme 'Infection & Immunity'	MolMed
2009-2012	Organizing lab rotations on phylogenetics for the Research Master programme 'Infection & Immunity'	Molmed

**Awards & Grants**

2011	Young Scientist travel grant	Malta
2010	Award for best poster, 16th International Bioinformatics Workshop	Rockville, US

**Miscellaneous**

2009	Academische jaarprijs - FluFighters "Fighting the great flu", Team Osterhaus	
------	--	--

## Publications

**Westgeest KB**, Bestebroer TM, van der Vliet S, Spronken MI, van Beek R, van Amerongen G, de Jong JC, Rimmelzwaan GF, Osterhaus AD, de Graaf M, Smith DJ, Fouchier RA. Mapping the Antigenic Evolution of the N2 Neuraminidase of Human Influenza A Viruses from 1957 to 2012. *In preparation*.

Lewis NS, Verhagen JH, Javakhishvili Z, Russell CA, Lexmond P, **Westgeest KB**, Bestebroer TM, Smith GJ, Bahl J, Wentworth DE, Waldenström J, Fouchier RAM, de Graaf M (2015). Influenza A virus evolution and spatio-temporal dynamics in Eurasian Wild Birds: A phylogenetic and phylogeographic study of whole-genome sequence data. *J Gen Virol. Aug;96(8):2050-60*

van de Sandt CE, Dou Y, Vogelzang-van Trierum SE, **Westgeest KB**, Pronk MR, Osterhaus AD, Fouchier RA, Rimmelzwaan GF, Hillaire ML (2015). Influenza B virus-specific CD8+ T-lymphocytes strongly cross-react with viruses of the opposing influenza B lineage. *J Gen Virol. Aug;96(8):2061-73*.

**Westgeest KB**, Bestebroer TM, Spronken MI, Gao J, Couzens L, Osterhaus AD, Eichelberger M, Fouchier RA, de Graaf M. (2015). Optimization of an Enzyme-Linked Lectin Assay suitable for High-Throughput Antigenic Characterization of the Neuraminidase of Human Influenza A(H3N2) Viruses. *J Virol Methods. Jun 1;217:55-63*.

Couzens L, Gao J, **Westgeest K**, Sandbulte M, Lugovtsev V, Fouchier R, Eichelberger M (2014). An optimized enzyme-linked lectin assay to measure influenza A virus neuraminidase inhibition antibody titres in human sera. *J Virol Methods Sep 16. pii: S0166-0934(14)00345-0*.

**Westgeest KB**, Russell CA, Lin X, Spronken MI, Bestebroer TM, Bahl J, van Beek R, Skepner E, Halpin RA, de Jong JC, Rimmelzwaan GF, Osterhaus AD, Smith DJ, Wentworth DE, Fouchier RA, de Graaf M (2014). Genome-wide Analysis of Reassortment and Genetic Drift of Human Influenza A(H3N2) Viruses Circulating between 1968 and 2011. *J Virol. Mar;88(5):2844-57*



**Westgeest KB**, de Graaf M, Fourment M, Bestebroer TM, Beek R, M. Spronken MI, de Jong JC, Rimmelzwaan GF, C. Russell CA, Osterhaus AD, Smith GJ, Smith DJ, Fouchier RA (2012). The Genetic Evolution of Neuraminidase of Influenza A (H3N2) Viruses from 1968 to 2009 and its Correspondence to Haemagglutinin. *J Gen Virol.* 93(Pt 9):1996-2007

Sandbulte MR, **Westgeest KB**, Gao J, Xu X, Klimov AI, Russell CA, Burke DF, Smith DJ, Fouchier RA, Eichelberger MC (2011). Discordant antigenic drift of neuraminidase and haemagglutinin in H1N1 and H3N2 influenza viruses. *Proc Natl Acad Sci U S A.* 20;108(51):20748-53

