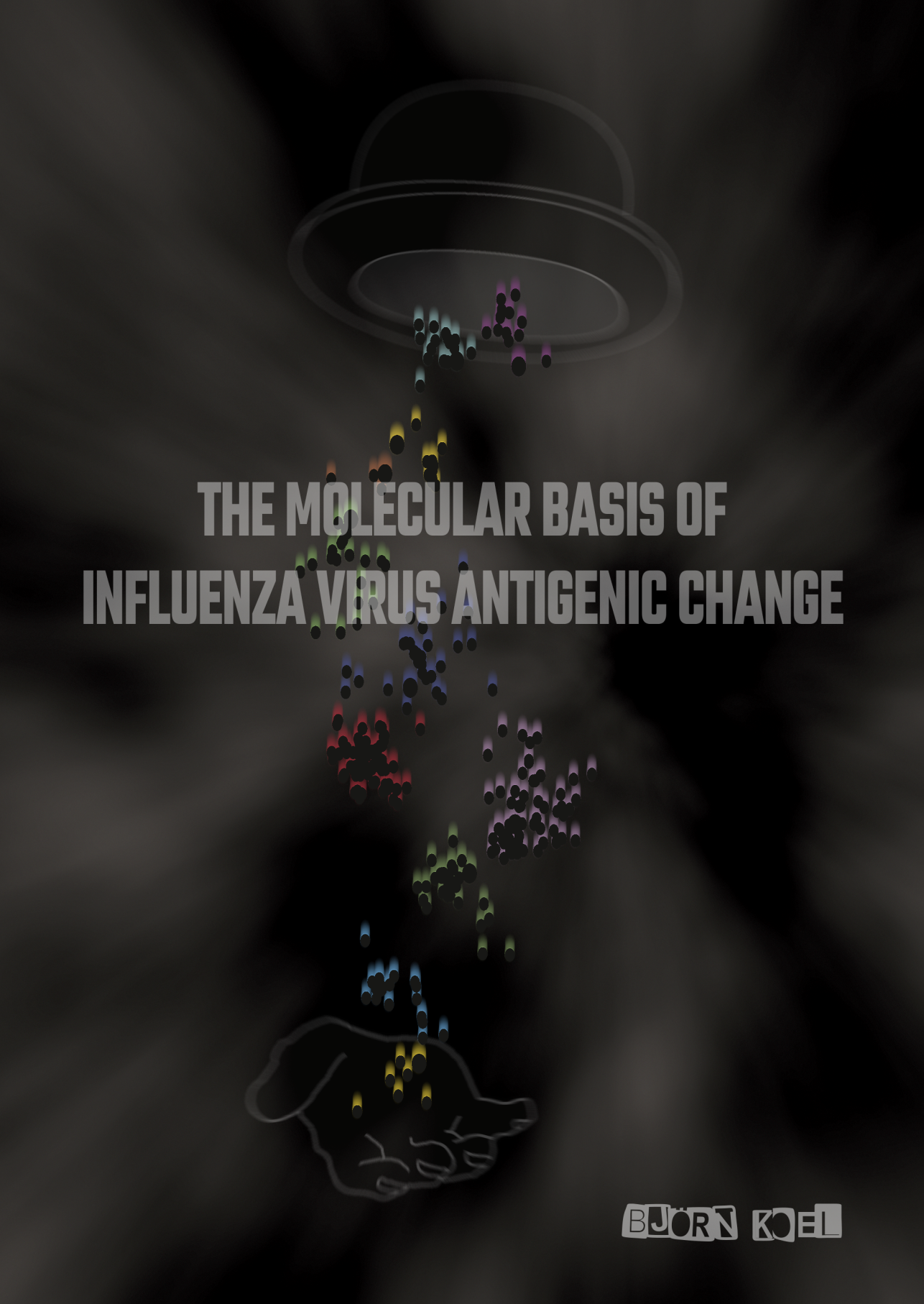


THE MOLECULAR BASIS OF INFLUENZA VIRUS ANTIGENIC CHANGE

BJÖRN KOEL

THE MOLECULAR BASIS OF INFLUENZA VIRUS ANTIGENIC CHANGE

Björn Ferry Koel

ISBN: 978-94-6182-672-5

Cover design: Björn Koel and Off Page, Amsterdam
Printing: Off Page, Amsterdam

The research described in this thesis was performed at the Department of Viroscience of the Erasmus Medical Centre, Rotterdam, The Netherlands, and financially supported by an NWO VICI grant, National Institute of Allergy and Infectious Diseases contract HHSN266200700010C, NIH Director's Pioneer Award DP1-OD000490-01, European Union FP7 program EMPERIE (223498), European Union FP7 program ANTIGONE (278976), and program grant P0050/2008 from the Human Frontier Science Program.

The research for this was performed within the framework of the Erasmus Postgraduate School Molecular Medicine.

THE MOLECULAR BASIS OF INFLUENZA VIRUS ANTIGENIC CHANGE

De moleculaire basis van de antigene evolutie van influenzavirus

Proefschrift

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

Prof.dr. H.A.P. Pols

en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op
donderdag 19 mei 2016 om 09:30 uur

door

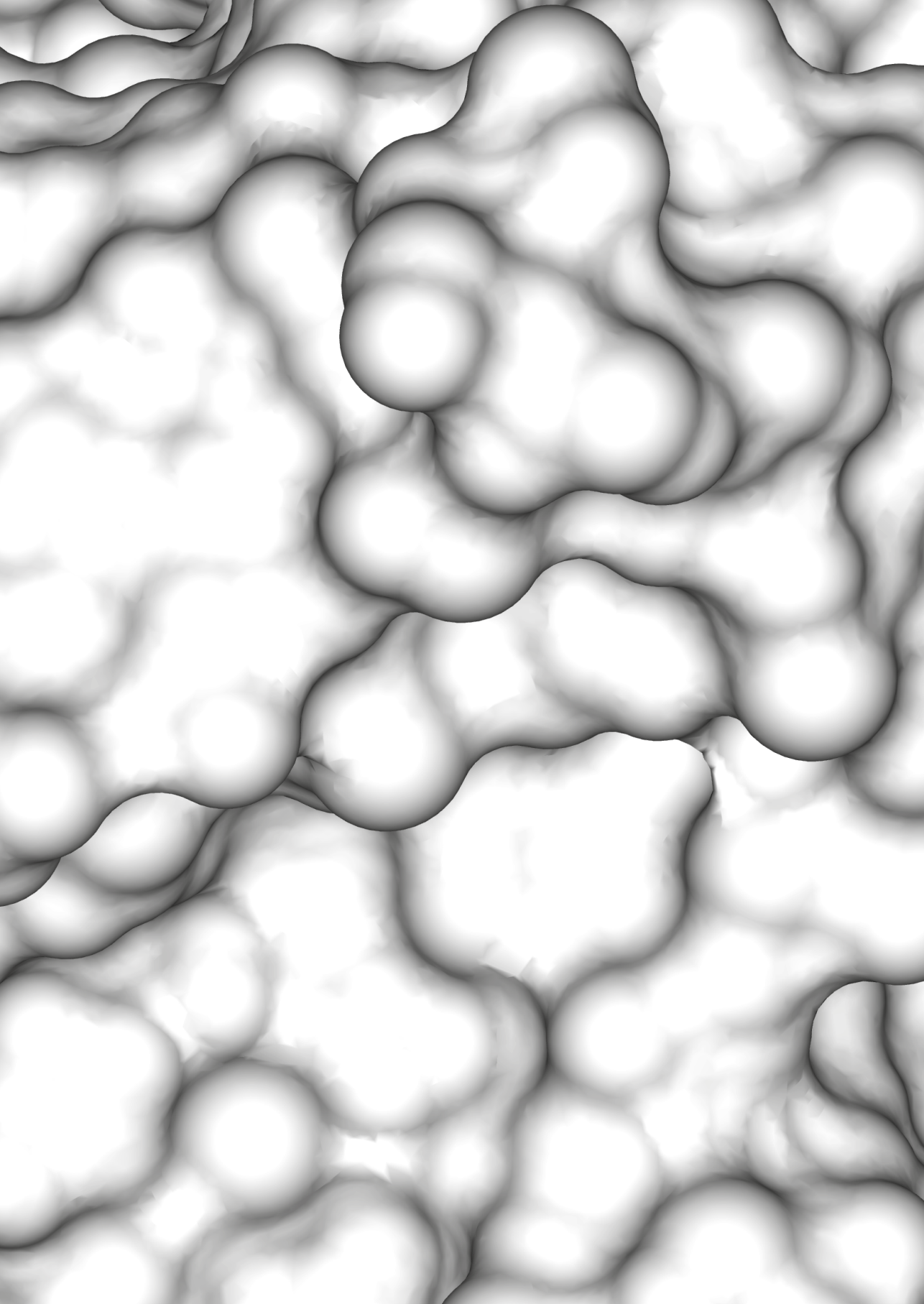
Björn Ferry Koel
geboren te 's Gravenhage

Erasmus University Rotterdam

PROMOTIECOMMISSIE

Promotoren:	Prof.dr. R.A.M. Fouchier
	Prof.dr. D.J. Smith
	Prof.dr. A.D.M.E. Osterhaus
Overige leden:	Prof.dr. G.F. Rimmelzwaan
	Prof.dr. M.D. de Jong
	Prof.dr. E.J. Snijder

TABLE OF CONTENTS



1

GENERAL INTRODUCTION

INFLUENZA VIRUSES

In humans, seasonal influenza is a highly contagious acute respiratory disease that is characterized by myalgia, headache, high fever, malaise, rhinitis and a non-productive cough. Highest morbidity rates are found in infants and particularly in the elderly over 65 years of age, and highest mortality rates in the latter, often in combination with underlying disease. Influenza A and B viruses cause the majority of influenza virus infections. Influenza C virus has a much lower prevalence and usually causes only minor symptoms. Recently, a novel influenza virus exclusively isolated from swine and cattle was proposed to represent a separate genus, the influenza D viruses (1).

Influenza viruses have a negative-sense, single-stranded, and segmented RNA genome and are classified in the family of *Orthomyxoviridae*. Influenza A viruses are further classified in subtypes based on their surface proteins hemagglutinin (HA) and neuraminidase (NA) (Fig. 1). To date, 18 HAs and 11 NAs have been identified (2-4). Nearly all combinations of H1-16 and N1-N9 are found in wild birds, which are considered the primary natural reservoir of influenza A viruses (2, 5). H17 and 18, and N10 and 11 are genetically and functionally distinct from all other known HAs and NAs and have exclusively been identified in bats (3, 4), suggesting that wild birds are not the only natural reservoir of influenza A viruses. Influenza A viruses infect a large variety of host species including humans, birds, pigs, horses, dogs, cats, bats, and sea mammals. Influenza B virus forms a single genus within the *Orthomyxoviridae* family and is only known to infect humans and seals (6).

INFLUENZA PANDEMICS AND EPIDEMICS

Influenza pandemics arise when antigenically novel influenza A virus strains enter the human population and subsequently spread globally. Zoonotic transmissions from wild birds into humans, either directly or via domestic animals such as pigs as an intermediate host, resulted in four influenza pandemics over the last century: A(H1N1) virus in 1918, A(H2N2) virus in 1957, A(H3N2) virus in 1968, and A(H1N1) in 2009 (Fig. 2). The latter is referred to as A(H1N1)pdm09 to differentiate the virus from the old seasonal A(H1N1) viruses that circulated during most of the 20th century. In 1957 and 1968, the newly introduced pandemic influenza viruses replaced the previously circulating subtypes. From 1977 onwards two influenza A virus subtypes co-circulated and caused annual epidemics. An A(H1N1) virus that was genetically and antigenically similar to A(H1N1) viruses that circulated in the mid 1940s was re-introduced in the human population in 1977 and circulated until it was replaced by A(H1N1)pdm09 virus in 2009. A(H3N2) viruses continued to circulate alongside both seasonal pre-2009 A(H1N1) and pandemic A(H1N1)pdm09 viruses. Influenza B viruses formed a single phylogenetic lineage, commonly referred to as the B/Lee/40-lineage, until the 1970s. This lineage diverged over time into two genetically and antigenically distinct lineages known as the B/Victoria/2/87 (B/Vic) and B/Yamagata/16/88 (B/Yam) lineages (7). Currently, two influenza A virus subtypes, A(H3N2) and A(H1N1)pdm09,

and viruses from both influenza B phylogenetic lineages cause frequent epidemics in humans (Fig. 2).

Influenza viruses are a major cause of morbidity and mortality. The 1918 Spanish influenza pandemic claimed an estimated 30 – 50 million lives (8). Subsequent influenza pandemics caused approximately 3.5 million deaths in total (9, 10). Seasonal human influenza viruses annually infect 5 -15% of the human population and are responsible for an estimated 250,000 - 500,000 deaths per year worldwide (11), resulting in a cumulative mortality that roughly equals that of pandemic influenza over the last century. The isolation of a highly pathogenic avian influenza (HPAI) A(H5N1) virus from a 3-year old boy in 1997 raised concerns about the possibility that influenza A virus subtypes not currently known to cause epidemics in humans may become pandemic (12, 13). To date, more than 700 laboratory-confirmed human cases of avian influenza A/H5N1 virus infection have been reported to the World Health Organization (WHO), 60% of which were fatal (14). Although it was recently shown that HPAI A(H5N1) viruses may evolve the ability to become aerosol transmissible (15, 16), no sustained human-to-human transmission of A(H5N1) viruses has been reported to date. However, since their identification in 1996, HPAI A(H5N1) viruses have spread to over 60 countries in the Eastern Hemisphere, diversified into multiple distinct genetic lineages, continued to cause large-scale outbreaks in domesticated poultry, and frequently crossed the species barrier into humans (17-20). HPAI A(H5N1) virus therefore remains a pathogen of significant concern from an economic, veterinary, and public health point of view. Although HPAI A(H5N1) virus is the most widespread zoonotic avian influenza virus, avian influenza viruses of HA subtypes H5, H6, H7, H9, and H10 have in various combinations with NA subtypes been isolated from and caused disease in humans (14, 21). Recent outbreaks in poultry populations include viruses of subtypes A(H5N8), A(H7N9), A(H9N2) and A(H10N8). The A(H7N9) virus is of particular concern because of persistent outbreaks in poultry populations, the absence of clinical disease in poultry, and the high number of human cases (22, 23). The dissemination of avian influenza viruses additional to A(H5N1) virus into poultry populations outside of Southeast Asia and potential for genetic diversification by mutation and reassortment that may lead to human adaptation are reason for continued surveillance for avian influenza viruses in poultry stocks worldwide.

HEMAGGLUTININ

Hemagglutinin (HA) is one of two major glycoproteins present on the surface of influenza viruses, the other one being neuraminidase (NA) (Fig 1). HA is synthesized as an HA_0 precursor protein. Post-translational modifications including glycosylation and cleavage by host cell proteases generate the disulfide-bond linked HA1 and HA2 subunits of the mature trimeric protein present on the viral surface (Fig. 3). HA is responsible for the initial stages of the replication cycle of influenza viruses. The receptor binding site (RBS) initiates binding of the viral particle to terminal sialic acid residues present on the host cell surface, resulting in receptor-mediated

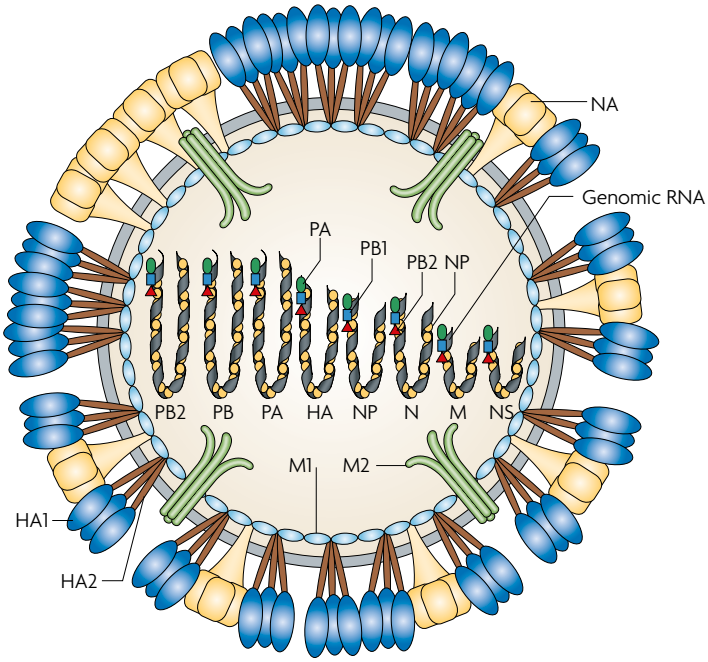


Fig. 1. Schematic representation of an influenza A virus. The surface proteins hemagglutinin (HA), neuraminidase (NA), and matrix 2 (M2) are inserted in the viral envelope, which is lined by the matrix 1 protein (M1). The eight RNA gene segments are present in the form of ribonucleoprotein-complexes that entail the nucleoprotein (NP) and the polymerase proteins (PB2, PB1, and PA). M1, matrix protein, NS; non-structural protein. The two HA subunits, HA1 (blue) and HA2 (brown), are indicated in the figure. Figure from Hedestam *et al.* (50).

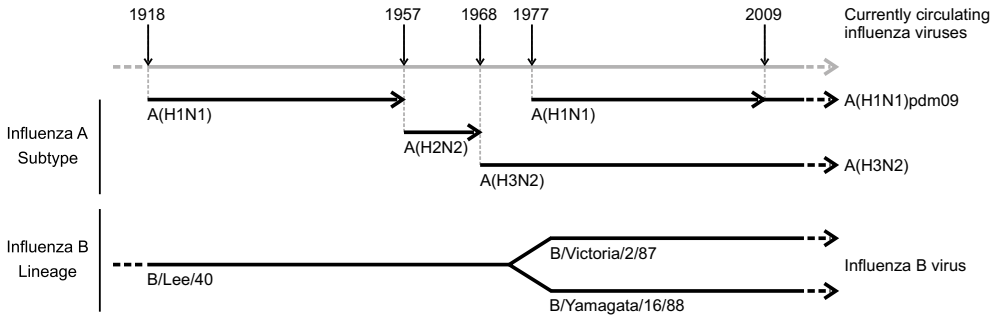


Fig. 2 Influenza A and B virus circulation in humans during the past century. Vertical arrows above the timeline (grey) indicate the influenza pandemics of the past century, horizontal arrows indicate the period of circulation after introduction into humans. The A(H1N1) viruses re-introduced in 1977 were descendants from the virus responsible for the 1918 influenza pandemic that circulated in the mid 1940s. Because these viruses were not antigenically novel the resulting outbreak is usually not considered a pandemic. Influenza B viruses formed a single phylogenetic lineage until the 1970s, after which they diverged over time into the two genetically and antigenically distinct lineages that continue to co-circulate until the present day.

endocytosis. The low pH inside the endosome triggers a conformational change of HA that exposes the fusion peptide at the N-terminus of the HA2 subunit. Fusion of the viral and endosomal membranes results in release of the gene segments into the cytoplasm. Subsequent trafficking of the viral RNA (vRNA) to the cell nucleus initiates the next stage of viral replication: the production of mRNA that serves as template for viral protein production and complementary RNA (cRNA) that serves as template for the production of vRNA, that will be incorporated in the newly formed viral particles.

The RBS is located on the membrane distal surface of HA and contains several structural elements that are fundamental to the HA of all subtypes (Fig. 3B). It forms a shallow concave pocket on the globular head of the HA molecule that is surrounded by the 130-loop, 220-loop, and 190 α-helix (24). Although sequence variation exists, the basic conformation of these elements is similar among influenza viruses. The amino acids on positions 98, 153, 183, and 195 are conserved throughout influenza A and B viruses and form the base of the RBS (24-26). Other amino acids within the RBS may also be conserved among subtypes with a similar host range and define the receptor specificity of HA. Human influenza viruses preferentially bind to α2,6-linked sialic acids (SAs), while avian influenza viruses have a preference for α2,3-linked SAs (27). Swine viruses bind either α2,6 or both α2,3 and α2,6 linked SAs.

HA is the main component of influenza vaccines because HA-specific antibodies can provide protective immunity against disease. Neutralizing antibodies prevent infection of the host cell by obstructing attachment of the RBS to SAs on the cell surface or preventing the membrane fusion reaction. Anti-HA antibodies are also involved in activation of other parts of the immune system that help viral clearance through opsonization and antibody-dependent cell-mediated cytotoxicity (28). The main antibody subtypes involved in virus neutralization and prevention from clinical disease are mucosal IgA in the upper respiratory tract and IgG in the lung (29). However, cross-reactivity between antibodies against the different influenza A virus subtypes and influenza B virus is very limited, necessitating the use of tri- or quadrivalent influenza vaccines.

INFLUENZA VIRUS ANTIGENIC CHANGE

The exceptionally high genetic diversity of influenza viruses results from the error prone influenza virus polymerase and allows rapid adaptation to changing environments, including evasion of neutralizing antibodies in the human population. The HA amino acid composition of seasonal human influenza viruses continually changes to avoid neutralization by antibodies generated in response to previous infections or vaccinations. The colonization of the European swine population by a human A(H3N2) virus in the early 1970s elucidates how antibody immunity shapes HA evolution. Whereas humans may experience several influenza virus infections during their lifetime, the short life span of pigs limits the buildup of immunity in the swine population. The human and swine A(H3N2) virus lineages had comparable rates of nucleotide and amino acid substitutions, but antigenic change of the swine virus

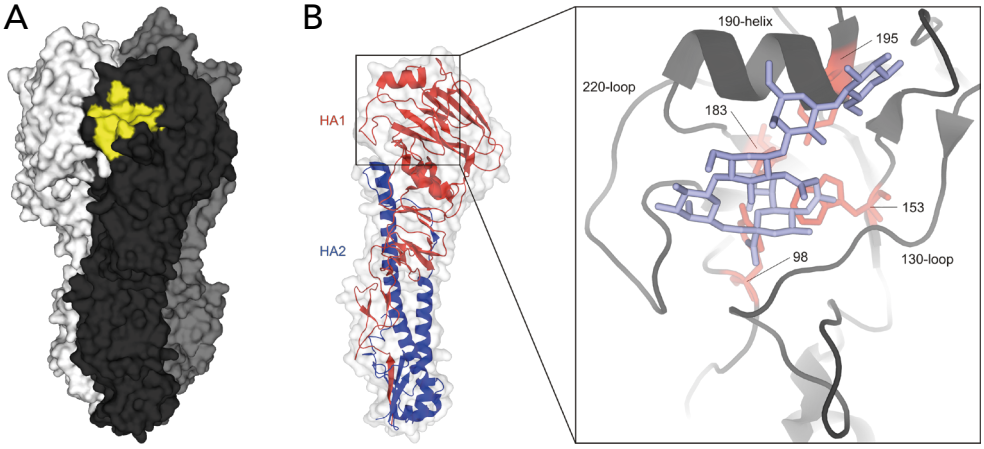


Fig. 3. Structure of an A/Aichi/2/1968 hemagglutinin. (A) The three monomers of HA (indicated in white, grey, and black) assemble during post-translational modification to form a mature trimeric HA protein. The receptor binding site (RBS, yellow) is located on the membrane distal globular head of HA. (B) Each monomer consists of an HA1 (red) and HA2 (blue) subunit. The amino acid composition of the RBS (inset) determines the receptor specificity of influenza virus. However, the overall shape of the RBS is similar for all influenza A and B viruses, and contains structural elements (the 130- and 220-loops and 190 α-helix) and amino acids (98Y, 153W, 183H, and 195Y, indicated in red) that are conserved among influenza A and B viruses. The human receptor analog LSTc is indicated in blue. PDB accession numbers; 5HMG (panel A), 2YPG (panel B).

lineage occurred at a rate approximately six times slower than the rate of antigenic drift in humans (30). Antigenic drift is the gradual antigenic change that results from the accumulation of point mutations in HA. Amino acid substitutions occur throughout the HA protein, but are most prominent in the HA1 domain. Nearly half of the HA1 amino acid positions of human A(H3N2) and A(H1N1) HAs, and the majority of surface exposed positions, have changed since 1968 and 1977, respectively, illustrating the highly variable nature of HA (31, 32). Owing to the rapid appearance of antigenic variants that escape population immunity, influenza vaccines must be frequently reformulated.

HA amino acid positions associated with antigenic drift of A(H3N2) and A(H1N1) viruses were identified in the early 1980s and the following decade by mapping genetic differences between strains from consecutive epidemics and mutations in escape variants after growing virus in the presence of antibodies (32-34). The 131 amino acid positions associated with antigenic change of A(H3N2) virus cover much of the HA globular head and were categorized in antigenic sites A – E (Fig. S6C of chapter 2) (32, 33). The antigenic sites identified for A(H1N1) and A(H5N1) viruses resemble the antigenic sites described for A(H3N2) virus (34-36). Four antigenic sites were identified for A(H1N1) virus (corresponding H3 sites in brackets): Ca – which is subdivided in Ca_1 and Ca_2 (D and A), Cb (C), Sa, and Sb (B). For A(H5N1) virus, three antigenic sites have been identified. Site 1 corresponds to site A of

H3, site 2 consists of two sub-sites; one that corresponds to site B of the H3 HA, the other to site Sa of H1 (neither of which have an equivalent H3 antigenic site). Based on the original publication of A(H3N2) antigenic sites it was suggested that at least one amino acid substitution in each of the four antigenic sites (A - D) was required for the production of new epidemic strains (33). However, after identification of a fifth site (E), that statement was adjusted to the widely used heuristic that each new drift variant of epidemiologic importance has generally had four or more amino acid substitutions located in two or more of the antigenic sites (32).

Antigenic differences among influenza viruses are measured routinely using the hemagglutination inhibition (HI) assay, which uses the property of influenza viruses to agglutinate erythrocytes and the ability of HA specific antibodies to block hemagglutination (37). In the HI assay, the highest dilution of antiserum that completely inhibits influenza viruses from agglutinating erythrocytes is determined. Comparing the reciprocal of endpoint dilutions, the HI titers, for a panel of antisera prepared against antigenically diverse viruses provides information about antigenic differences between the test viruses. An example HI table is shown in Figure 4. The use of antisera in serological assays is based on the finding that the level of serum antibodies reflects immunity to influenza virus (38-40). Ferret antisera obtained after inoculation with influenza viruses under laboratory-controlled conditions are the most sensitive reagents for detection of antigenic variation of influenza viruses, but antisera from other species are occasionally preferred. For example, chicken antisera, which have IgY instead of IgG, are often used in HI assays testing A(H5N1) viruses.

Antigenic cartography is a method to visualize and analyze HI data (31). The concept of antigenic cartography to generate "antigenic maps" is described in Figure 4. Because antisera are tested against multiple antigens, and antigens tested against multiple antisera, many measurements can be used to determine the position of the antigen and antiserum in an antigenic map, thus improving the resolution of HI data. Visualization of HI data in an antigenic map allows studying long-term virus evolution. Analyzing HI data of A(H3N2) viruses that circulated between 1968 and 2003 revealed that antigenic evolution was punctuated rather than gradual, and that antigenic evolution proceeded in a mostly linear fashion away from earlier circulating viruses, presumably related to escape from existing population immunity (31). Amino acid differences between the 11 antigenic clusters—the groups of antigenically similar viruses—were dispersed across the five A(H3N2) virus antigenic sites. Several studies used sequence information to infer amino acid positions associated with antigenic change (41-47). These studies also mostly identified amino acid positions scattered over the HA globular head located in multiple antigenic sites. However, only few studies pinpointed the exact substitutions responsible for antigenic change based on phenotypic analysis (31, 48, 49).

A

Virus Isolate:	BE/353/89	LY/1149/91	SP/22/90	HK/34/90	MA/G102/93	BE/32/92
A/Beijing/353/89	**1280**	320	640	160	10	40
A/Lyon/1149/91	640	**1280**	1280	20	<10	10
A/Madrid/G101/93	640	640	1280	N.D.	40	160
A/Madrid/G103/93	640	640	1280	N.D.	40	160
A/Singapore/22/90	40	40	**1280**	1280	40	160
A/Hong Kong/34/90	80	20	2560	**1280**	40	160
A/Lyon/459/91	80	40	N.D.	1280	40	80
A/Netherlands/938/92	80	40	1280	2560	160	640
A/Madrid/G102/93	40	10	1280	160	**1280**	1280
A/Beijing/32/92	80	20	320	2560	320	**1280**

B

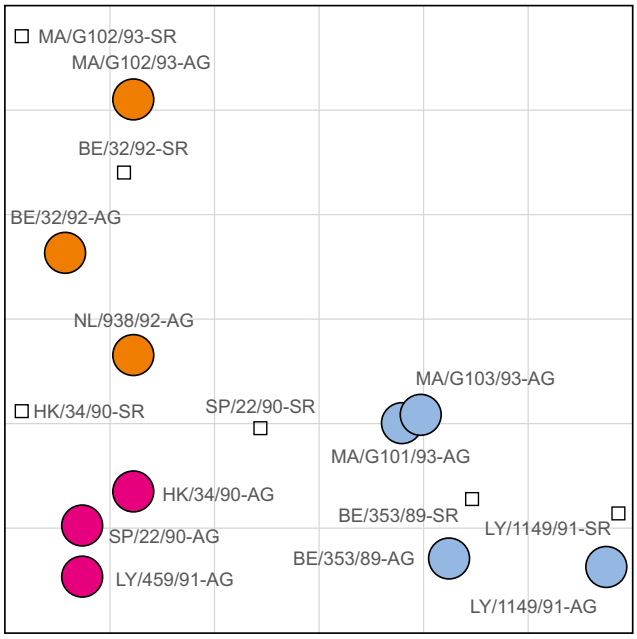


Fig. 4. Example of an antigenic map of influenza A(H3N2) virus. (A) HI assay results in table format. The antigens are indicated in the left most column, antisera in the top row. Each titer in the HI table can be thought of as specifying a target distance between an antigen and serum point in the antigenic map (N.D: not determined). (B) HI assay results represented in an antigenic map. The open squares represent the antisera; colored circles represent the antigens (viruses). The distance between antiserum point S and antigen point A corresponds to the difference between the log2 of the maximum titer observed for antiserum S against any antigen and the log2 of the titer for antiserum S against antigen A. Modified multidimensional scaling methods are then used to arrange the antigen and antiserum points in an antigenic map to best satisfy the target distances specified by the HI data. The result is a map in which the distance between the points represents antigenic distance as measured by the HI assay in which the distances between antigens and antisera are inversely related to the log2 HI titer. Each square in the grid of the antigenic map equals a two-fold difference in the HI assay. The colored rows in the table and the colors of the antigens in the map indicate antigenically similar viruses.

1

OUTLINE OF THIS THESIS

Understanding the genetic changes that allow influenza viruses to escape population immunity is a critical first step towards the prediction of influenza virus antigenic evolution. The current possibility to integrate genetics, classic serology, and bioinformatics provides an opportunity to study the molecular basis of influenza virus antigenic change in detail.

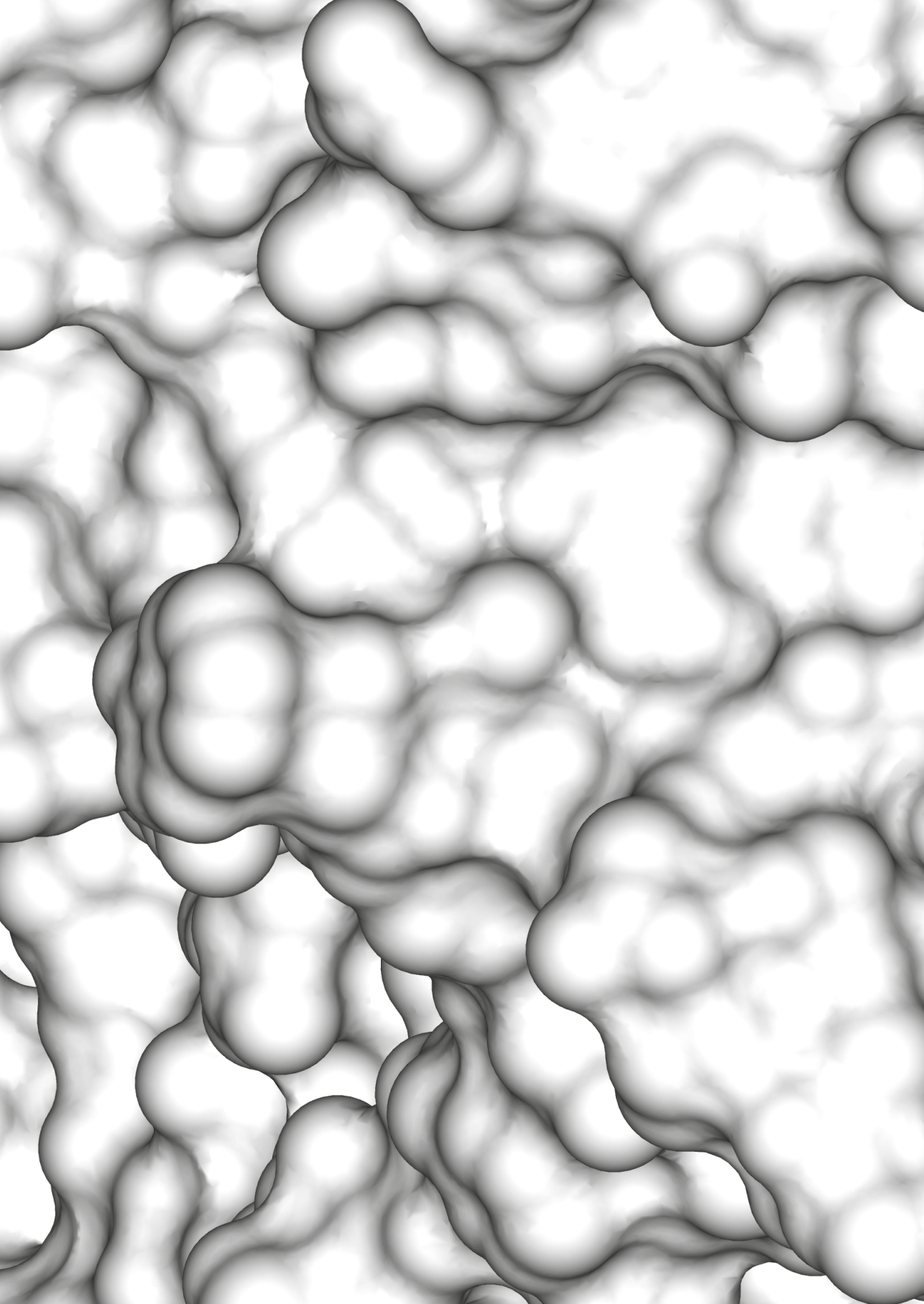
The knowledge that antigenic evolution of A(H3N2) virus is punctuated was used to identify the amino acid substitutions responsible for the antigenic differences between HAs of viruses that belong to consecutive antigenic clusters, and allowed us to delineate a small set of HA amino acid positions and substitutions that shaped antigenic evolution of human influenza viruses (chapter 2).

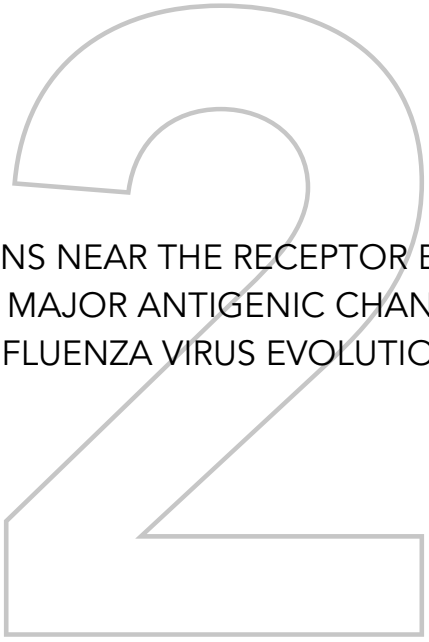
Genetic evolution of HPAI A(H5N1) viruses has been studied in detail, mostly focusing on the possibility that A(H5N1) viruses cross the species barrier and form a stable lineage in humans. Much less is known about the antigenic evolution of HPAI A(H5N1) viruses, hampering the design of intervention strategies that could help prevent outbreaks in poultry. We mapped antigenic variation of HA of A(H5N1) clade 2.1 viruses and identified the evolutionary patterns and molecular basis of the observed antigenic changes (chapter 3).

The work described in chapters 2 and 3 indicated that major antigenic changes of HA during influenza virus evolution were predominantly caused by single amino acid substitutions. Two such substitutions were introduced into the HA of temporally and genetically diverse A(H3N2) viruses to investigate the influence of the amino acid context in which a substitution occurs on antigenic change (chapter 4).

The majority of A(H1N1)pdm09 viruses isolated since the 2009 influenza pandemic remain antigenically similar to the pandemic virus. However, A(H1N1)pdm09 viruses are anticipated to change antigenically as a result of increasing population immunity. Building on the work from chapters 2 and 3, we introduced substitutions into a virus representing the antigenic phenotype of A(H1N1)pdm09 viruses to identify amino acid substitutions in HA that support escape from convalescent ferret antisera and human sera (chapter 5).

The main findings and implications of this thesis are discussed in chapter 6.





SUBSTITUTIONS NEAR THE RECEPTOR BINDING SITE DETERMINE MAJOR ANTIGENIC CHANGE DURING INFLUENZA VIRUS EVOLUTION

ABSTRACT

The molecular basis of antigenic drift was determined for the hemagglutinin (HA) of human influenza A/H3N2 virus. From 1968 to 2003 antigenic change was caused mainly by single amino acid substitutions, which occurred at only seven positions in HA immediately adjacent to the receptor binding site. Most of these substitutions were involved in antigenic change more than once. Equivalent positions were responsible for the recent antigenic changes of influenza B and A/H1N1 viruses. Substitution of a single amino acid at one of these positions substantially changed the virus-specific antibody response in infected ferrets. These findings have potentially far-reaching consequences for understanding the evolutionary mechanisms that govern influenza viruses.

Björn F. Koel,[1] David F. Burke,[2,3] Theo M. Bestebroer,[1] Stefan van der Vliet,[1] Gerben C. M. Zondag,[4,5] Gaby Vervaet,[1] Eugene Skepner,[2,3] Nicola S. Lewis,[2,3] Monique I. J. Spronken,[1] Colin A. Russell,[3,6] Mikhail Y. Eropkin,[7] Aeron C. Hurt,[8] Ian G. Barr,[8] Jan C. de Jong,[1] Guus F. Rimmelzwaan,[1] Albert D. M. E. Osterhaus,[1] Ron A. M. Fouchier,[1] Derek J. Smith[1,2,3,9]

[1]Department of Viroscience, Erasmus MC, 3015GE Rotterdam, Netherlands.
[2]Center for Pathogen Evolution, Department of Zoology, University of Cambridge, Cambridge CB2 3EJ, UK.
[3]WHO Collaborating Centre for Modeling Evolution and Control of Emerging Infectious Diseases, University of Cambridge, Cambridge CB2 3EJ, UK.
[4]BaseClear B.V., 2333CC Leiden, Netherlands.
[5]Luris, Leiden University, 2333AA Leiden, Netherlands.
[6]Department of Veterinary Medicine, University of Cambridge, Cambridge CB3 0ES, UK.
[7]Research Institute of Influenza, 197376 St. Petersburg, Russia.
[8]WHO Collaborating Centre for Reference and Research on Influenza, VIDRL, Melbourne, Victoria 3051, Australia.
[9]Fogarty International Center, National Institutes of Health, Bethesda, MD 20892, USA.

Science, 2013, 342(6161): 976-9

Influenza A/H3N2 virus is a major cause of morbidity and mortality in humans and poses a considerable economic burden (11, 51). Vaccination is the primary method to reduce this public health impact. The HA surface glycoprotein is the main component of influenza vaccines and antibodies against HA can prevent serious illnesses (40). However, influenza viruses can escape from antibody-mediated neutralization by accumulating mutations in HA in a process called antigenic drift, and as a consequence influenza vaccines require frequent updates. Several recent studies have focused on the identification of conserved domains of HA as targets of virus-neutralizing antibodies to circumvent this problem (52-55). Other recent work has focused on identifying the mechanisms of antigenic drift (56, 57), and on sequence-based prediction to identify positively selected codons (43, 44, 46, 58). This research has been restricted by our limited fundamental insight into the molecular basis of antigenic evolution.

Seminal work in the 1980s identified 131 amino acid positions in five antigenic sites (A-E) on the globular head of HA as main targets for specific antibodies and suggested that antigenic drift is caused by accumulation of amino acid substitutions in these sites (32, 33). This work led to the widely used heuristic that it takes at least four amino acid substitutions, spread between two or more different antigenic sites, to cause substantial antigenic change. Smith et al. (31) showed that eleven antigenic clusters of viruses have emerged since the introduction of the A/H3N2 virus in humans in 1968 until 2003, each of which was subsequently replaced by viruses with distinct antigenic properties. Between 1 and 13 amino acid substitutions were associated with each of the antigenic cluster transitions. Almost all of these cluster-difference substitutions were in the antigenic sites (31). Here, we investigate which of these substitutions actually caused the antigenic change.

We selected a representative virus from each antigenic cluster. The HA1 subunit amino acid sequence, which comprises the globular head domain of HA including the receptor binding site (RBS), of each representative virus was identical to the consensus sequence for all strains from the respective cluster (59). The consensus HA genes, representing natural circulating viruses, were used to make recombinant viruses in the context of the A/Puerto Rico/8/1934 reference virus (60). We also produced chimeric viruses with the full HA1 or with HA1 positions 109-301 of each antigenic cluster consensus strain in the context of HA of the Sichuan 1987 cluster consensus virus (Fig. S1). The antigenic properties of all viruses were analyzed in hemagglutination inhibition (HI) assays using a panel of 8-16 ferret antisera raised against A/H3N2 viruses from 1968-2006 (Table S1). The wild type, recombinant, and chimeric viruses all had similar antigenic properties in the HI assay (Figs. S2 and S3). Thus, HA1 positions 109 - 301 determined the antigenic phenotype of representative A/H3N2 strains circulating between 1968 and 2003.

Each of the 54 cluster-difference substitutions within the 109 - 301 HA1 region (Table S2) was introduced independently into the corresponding consensus HA and tested for antigenic effect. If no single substitution induced the antigenic change

to the subsequent cluster, combinations of substitutions were tested. We call the subset of cluster-difference substitutions that were responsible for the major antigenic change between clusters "cluster-transition" substitutions. The cluster-transition substitutions were also tested in reverse. For example, each of the 11 cluster-difference substitutions between the EN72 and VI75 clusters were introduced individually in the EN72 consensus virus. We found that substitution 189QK alone caused the change in antigenic phenotype from EN72-like to VI75-like. The reverse substitution 189KQ was then introduced into the VI75 consensus HA and resulted in reversion of the antigenic phenotype to EN72-like (Fig. S4B).

Surprisingly, seven of the ten cluster transitions were caused by only a single amino acid substitution (Fig. 1, Table S3). Two cluster transitions were caused by two substitutions, and one by three substitutions. In two of the three cluster transitions for which more than one substitution was required for the full antigenic change, one substitution was responsible for the majority of the antigenic change and the additional substitution(s) had a relatively modest effect on the antigenic phenotype (Fig. 1, Figs. S4C and E, Table S2). For the remaining transition the effect caused by the single substitution could not be determined (Fig. S4I). We further tested the results for all cluster transitions using virus neutralization assays and obtained similar results (Figs. S5 and S6, Table S2). In summary, we found that at least 9 of 10 A/H3N2 cluster transitions over the 35-year period were predominantly caused by single amino acid substitutions.

Strikingly, all the cluster-transition substitutions occurred at only seven positions, immediately adjacent to the RBS (Fig. 2). Six of seven positions align to form an antigenic ridge on the periphery of the RBS, while the other, at position 145, is located on a loop that is partly involved in receptor binding. Five of seven key positions were involved in a cluster transition at least twice (Fig. 1, Table S3).

Substitutions responsible for major antigenic change were located exclusively in antigenic sites A (position 145) and B (positions 155, 156, 158, 159, 189, and 193), with none in sites C, D, or E. Our results do not contradict the studies that found 131 positions that can cause antigenic change and that define sites A-E (32, 33). Rather we show that nature has, during 35 years of A/H3N2 virus circulation in humans, only selected for the substantial antigenic changes caused by substitutions at only seven of these 131 position; this is an important change in our understanding of the antigenic evolution of seasonal influenza viruses.

Even though antibodies have been found to bind many regions of the HA, the location of the key positions for antigenic change on the periphery of the RBS suggests that RBS region specific antibodies play the critical role in neutralizing influenza A/H3N2 viruses.

To test the surprising finding that the antigenic phenotype was determined by the amino acids at just seven positions we attempted to change the phenotype of the HK68 consensus virus to become FU02-like by introducing only the five cluster-transition substitutions that differ between the HK68 and FU02 antigenic clusters (Table S4). This test compares the effect of substitutions at just five positions (HK68+5)

to the 54 substitutions that occurred over 35 years between HK68 and FU02, 47 of which were in antigenic sites. Remarkably, the HK68+5 mutant is FU02-like antigenically: it had an HI titer only 3-fold lower than the FU02 consensus virus to a FU02 antiserum (Fig. S7, Table S4). Interestingly, although HK68+5 had a substantially (20-fold) reduced titer with the HK68 antiserum, it retained some residual low reactivity with the HK68 serum. In a similar experiment we incrementally introduced cluster-transition substitutions into the HK68 consensus virus. The antigenic properties of the mutants matched those of the consensus viruses from subsequent clusters other than in their HI titers to one of the two HK68 sera. After introduction of all cluster-transition substitutions between the first five clusters we were unable to rescue the mutant virus. In summary, even for highly divergent strains with as many as 54 substitutions and 35 years of evolution between them, only the amino acid substitutions at key positions determined the majority of the antigenic phenotype.

We next tested the effect of a cluster-transition substitution in HA on the antibody response in an infected host. Two groups of three ferrets each were inoculated with either the wildtype HK68 consensus virus or HK68 containing cluster-transition substitution 155TY, and antisera were collected two weeks later. Each antiserum was titrated in the HI assay to 31 viruses belonging to the HK68, EN72, or VI75 antigenic clusters (31). Antisera to the HK68 consensus virus had a 25-fold (4.6 $\log_2$) reduction in reactivity to EN72-like strains (compared to its reactivity with HK68-like viruses) and had no reactivity to the majority of the VI75-like viruses. In contrast, antisera to HK68 155TY had only 4-fold (2.1 $\log_2$) reduced titers to EN72-like strains, and had substantial titers to the viruses from the VI75 cluster (Fig. 3 and Fig. S9). Thus, a single cluster-transition substitution can also substantially increase reactivity of the antiserum raised to the mutant virus to strains from subsequent antigenic clusters while maintaining reactivity to the antigenic cluster of origin.

The identification of key influential positions in HA and single amino acid substitutions responsible for major antigenic changes provides a plausible explanation for the previously described punctuated nature of the antigenic evolution of influenza A/H3N2 virus (31): 9 out of 10 cluster transitions were predominantly caused by a single amino acid substitution.

Although we found that almost 90% of HA substitutions associated with cluster transitions had little or no measurable antigenic effect, these findings do not imply that substitutions away from the RBS are phenotypically neutral. Some may be neutral hitchhikers carried along by chance, but others may collectively add to the decreased recognition by antibodies raised to a strain in an earlier cluster (Figs. 3, S7 and S8), or may be compensatory mutations necessary to retain function.

Given the high mutation rate of influenza virus and the observation that single amino acid substitutions are sufficient to cause antigenic cluster transitions, it is surprising that new antigenic clusters appear as slowly as they do—on average every 3.3 years (31). One hypothesis is that antigenic change has an intrinsic fitness cost that slows down the antigenic evolution of the virus. We found that several mutant viruses with substitutions near the RBS could not be rescued (Figs. S4E and I, S8).

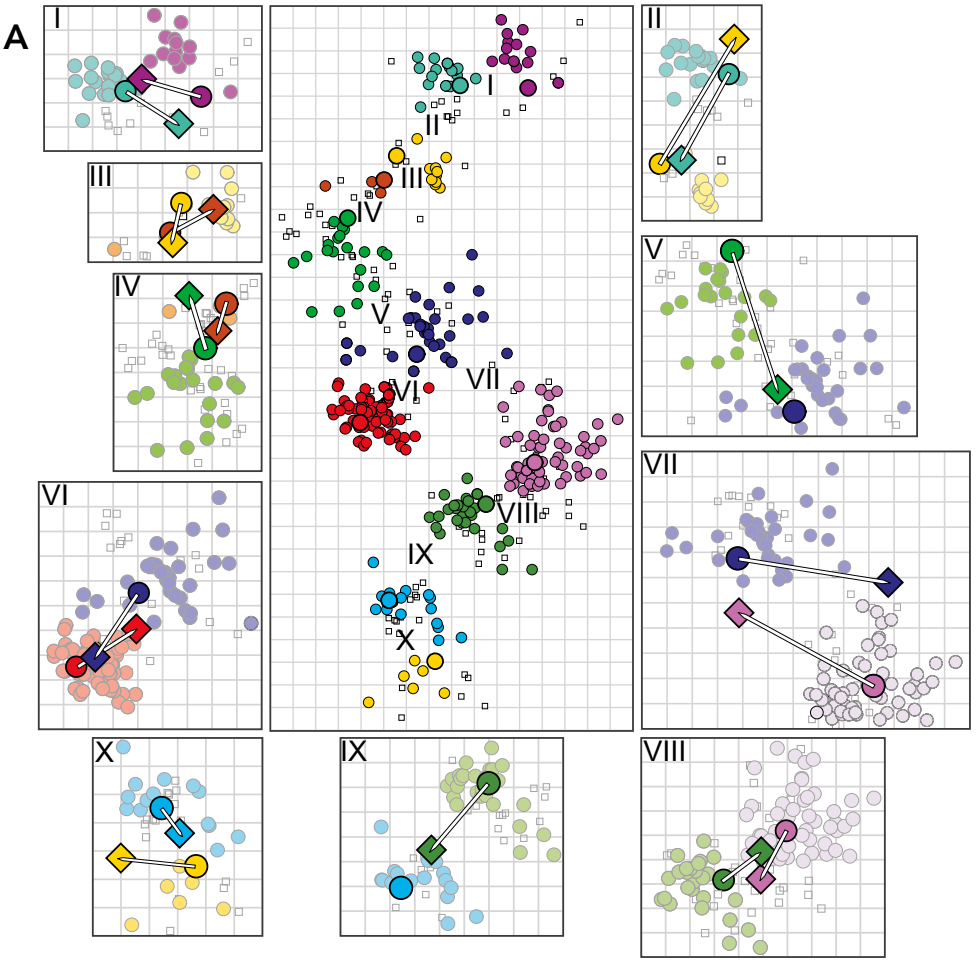
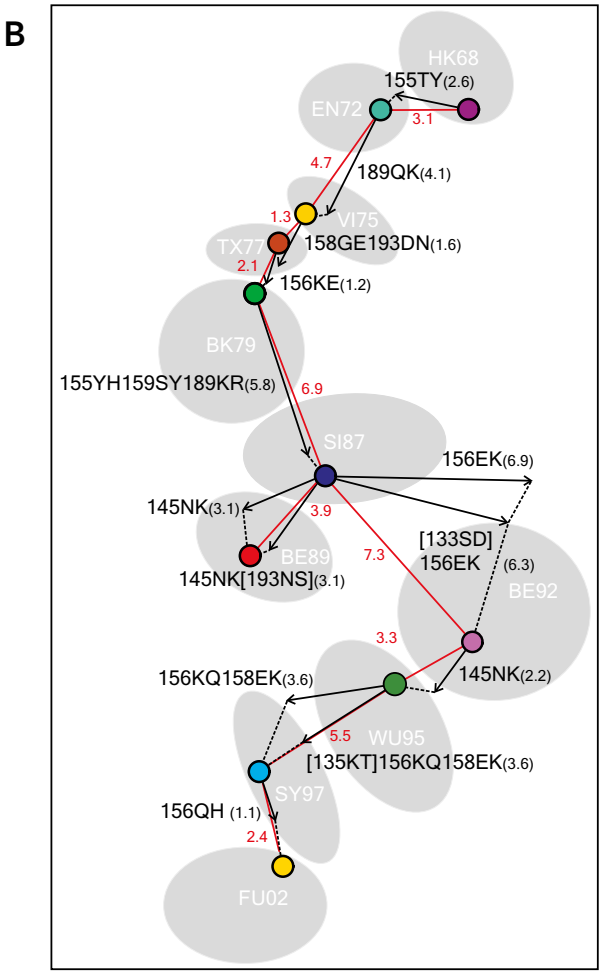


Fig. 1. Antigenic effect of cluster-transition substitutions. A) Antigenic maps of A/H3N2 virus evolution and cluster transition mutants. The central map provides an overview in which antisera and epidemic strains are indicated by open squares and colored circles respectively, and consensus viruses by larger colored circles. Both vertical and horizontal axes represent antigenic distance. The spacing between gridlines is one antigenic unit distance, corresponding to a two-fold difference in the HI assay. Cluster names are as indicated in B). Maps I-X show each cluster transition in more detail. Diamonds indicate the position of viruses with cluster-transition substitutions. B) Overview of cluster-transition substitutions. The clusters are named after the first vaccine strain in the cluster, with letters and digits referring to location and year of isolation (HK; Hong Kong, EN; England, VI; Victoria, TX; Texas, BK; Bangkok, SI; Sichuan, BE; Beijing, WU; Wuhan, SY; Sydney, FU; Fujian).



Colored circles indicate the consensus viruses. The gray background shapes indicate the cloud of strains that make up an antigenic cluster. The red line represents the evolutionary path from the HK68 to the FU02 consensus virus. Antigenic distances between consensus viruses are shown in red. Arrows indicate antigenic distance and direction of the cluster-transition amino acid substitutions. Antigenic distance between the mutant and consensus virus is indicated in parentheses. Substitutions between square brackets are accessory substitutions, which changed direction towards the subsequent cluster (Figs. 2, S6 and S10, Table S3), and did not significantly add to the antigenic distance.

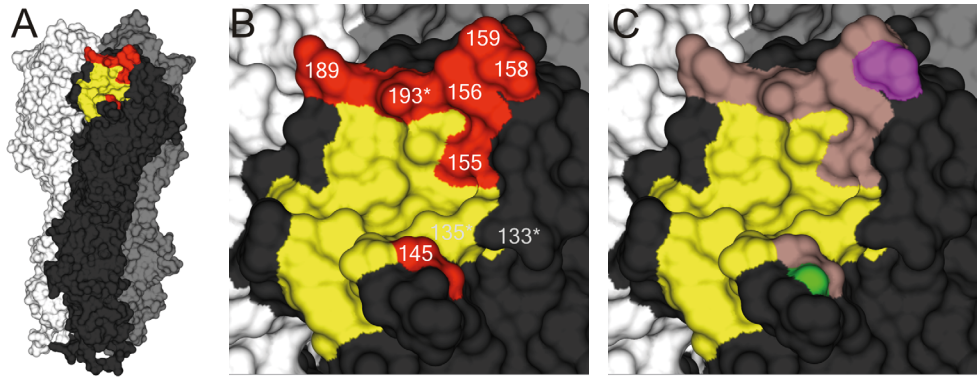


Fig. 2. Positions of the cluster-transition amino acid substitutions indicated on an A/Aichi/2/1968 HA trimer. The three monomers are shown in black, white and grey, the RBS in yellow. A) and B) The positions responsible for A/H3N2 cluster transitions are shown in red. An asterisk indicates accessory substitutions (Fig. S10). Position 193 is both a cluster-transition and accessory substitution (Fig. 1B). C) Positions of amino acid substitutions responsible for antigenic change of influenza A/H1N1 and B virus are shown in green and magenta, respectively. The positions responsible for cluster transitions of A/H3N2 virus are shown in light brown.

These results suggest that introduction of mutations on the periphery of the RBS can affect HA function, and that co-mutations may be crucial to retain viral fitness. The requirement for co-mutations that offset the fitness cost of antigenic evolution could provide an explanation for the paradox of high mutation rate and slow antigenic evolution.

Almost all substitutions responsible for cluster transitions resulted in substantial changes in the biophysical properties of the amino acids involved. For instance, in nine cluster transitions at least one charge change was involved, and substantial volume changes were involved in three transitions (Table S5). The amino acid composition of the key positions seems limited by their exposed nature. Hydrophobic amino acids are typically located on the interior of a protein where they can be shielded from solvent access. Tyrosine, which is only partially hydrophobic because of the hydroxyl group on its aromatic ring, was the only hydrophobic amino acid observed on the key positions. The necessity for maintaining a functional HA structure as well as escape from neutralizing antibodies may have restricted the range of possible amino acids.

The number of potential glycosylation sites on the A/H3N2 virus HA has steadily increased since 1968 (24), and the presence of carbohydrate side chains has been associated with shielding of antibody epitopes (61, 62). A total of 14 HA positions are associated with glycosylation of the A/H3N2 viruses that circulated from 1968 – 2003. Two positions, 133 and 144, are located adjacent to the RBS. However, changes in glycosylation did not coincide with cluster transitions and viruses belonging to the same antigenic cluster often had different glycosylation states (63). In agreement with these observations, the cluster-transition substitutions neither

introduced nor deleted any glycosylation sites, and glycosylation was therefore not directly involved in the antigenic changes between the clusters. The location of the majority of the potential glycosylation sites away from the RBS could be the reason why changes in glycosylation have not played a major role in the antigenic change of A/H3N2 viruses.

To further expand these results, we examined the most recent antigenic cluster transitions in the other human seasonal influenza viruses: the B/Yamagata and B/Victoria lineages of the B viruses, and the pre-2009 A/H1N1 viruses (the current A/H1N1pdm09 viruses are yet to undergo a cluster transition). We found the same results as for the A/H3N2 viruses: The major antigenic change was caused by a single amino acid substitution in the corresponding region close to the RBS (Fig. 2C, Figs. S11-13).

We find that although human A/H3N2 seasonal influenza viruses have fixed amino acid substitutions at 54 positions in antigenic sites, substitutions at only seven of these

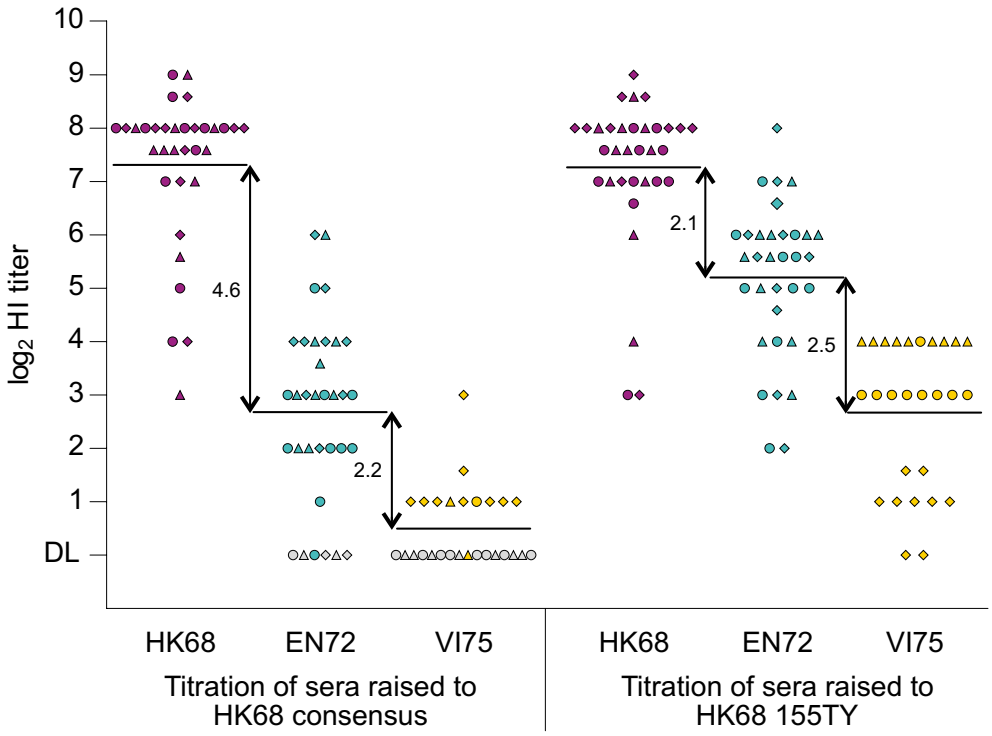


Fig. 3. Effect of a cluster-transition substitution on the antibody response of ferrets. Groups of three ferrets were inoculated with HK68 consensus virus or HK68+155TY, and antisera were tested to strains from the HK68 (purple), EN72 (cyan) or VI75 (yellow) clusters of Fig. 1. Circles, triangles and diamonds mark the three individual antisera. Black horizontal lines show the mean $\log_2$ HI titer. DL indicates the detection limit of the HI assay at the starting dilution (1/20) we used. Reactivity below detection was set to the value of the detection limit to calculate the means; these points are indicated in grey. The $\log_2$ reduction in mean reactivity to the different strains is indicated between the groups.

locations have been responsible for the major antigenic changes in these viruses to date. Moreover, these locations are all near the RBS of the HA, which suggests the mechanism for slowing the antigenic evolution of these viruses could be a reduction in receptor binding function. This small number of critical positions, and restricted amino acid usage involved in antigenic cluster transitions, suggests that possibilities for important antigenic change of seasonal influenza viruses may be more restricted than previously thought, with potentially far reaching consequences for understanding the underlying evolutionary mechanisms governing such viruses.

ACKNOWLEDGEMENTS

We thank M. Aban, G. van Amerongen, C. Baas, R. van Beek, M. de Graaf, S. Herfst, S. James, M. Linster, K. Sutherland-Cash, and C. Whittleston for excellent technical assistance and discussions. D.F.B and D.J.S acknowledge the use of the CamGrid distributed computing resource. This work was financed through an NWO VICI grant, National Institute of Allergy and Infectious Diseases-NIH contract HHSN266200700010C, NIH Director's Pioneer Award DP1-OD000490-01, European Union FP7 program EMPERIE (223498), European Union FP7 program ANTIGONE (278976), and program grant P0050/2008 from the Human Frontier Science Program. The Melbourne WHO Collaborating Centre for Reference and Research on Influenza is supported by the Australian Government Department of Health and Ageing. A.D.M.E.O. (on behalf of Viroclinics Biosciences BV) has advisory affiliations with GSK, Novartis, and Roche. A.D.M.E.O. and G.F.R are consultants for Viroclinics Biosciences BV. A/H3N2 virus sequences were previously published by Smith et al. (16). The sequences of the influenza A/H1N1 and B viruses used in this study are available from the GISAID EpiFluTM Database (www.gisaid.org) and are listed in Table S6.

SUPPLEMENTARY MATERIALS

Substitutions near the receptor binding site determine major antigenic change during influenza virus evolution

Björn F. Koel, David F. Burke, Theo M. Bestebroer, Stefan van der Vliet, Gerben C.M. Zondag, Gaby Vervaet, Eugene Skepner, Nicola S. Lewis, Monique I.J. Spronken, Colin A. Russell, Mikhail Y. Eropkin, Aeron C. Hurt, Ian G. Barr, Jan C. de Jong, Guus F. Rimmelzwaan, Albert D.M.E Osterhaus, Ron A.M. Fouchier, Derek J. Smith.

Materials and Methods
Figs. S1 to S13
Tables S1 to S6

A zoomable pdf of the Supplementary Materials is available from:
www.sciencemag.org/content/342/6161/976/suppl/DC1

MATERIALS METHODS

Cells

293T cells were cultured in DMEM (Cambrex, Heerhugowaard, The Netherlands) supplemented with 10% FCS, 100 IU/ml penicillin, 100 µg/ml streptomycin, 2mM glutamine, 1mM sodiumpyruvate, and non-essential amino acids (MP Biomedicals, Europe, Illkirch, France). Madin-Darby Canine Kidney (MDCK) cells were cultured in EMEM (Cambrex) supplemented with 10% FCS, 100 IU/ml penicillin, 100 µg/ml streptomycin, 2mM glutamine, 1.5 mg/ml sodiumbicarbonate (Cambrex), 10mM Hepes (Cambrex), and non-essential amino acids (MP Biomedicals).

Recombinant, chimeric and mutant virus construction

Full hemagglutinin (HA) genes of the consensus virus of each antigenic cluster were cloned in the modified pHW2000 expression plasmid (64, 65), where consensus viruses are those viruses with the HA amino acid consensus sequence for that cluster (HK68; A/Bilthoven/16190/68, EN72; A/Bilthoven/21793/72, VI75; A/Bilthoven/1761/76, TX77; A/Bilthoven/2271/76, BK79; A/Netherlands/233/82 193N, SI87; A/Netherlands/620/89, BE89; A/Netherlands/823/92, BE92; A/Netherlands/179/93, WU95; A/Netherlands/178/95 145K, SY97; A/Netherlands/427/98 and FU02; A/Netherlands/213/03. Cluster names as defined in Fig. 1). For nine of the eleven clusters the HA amino acid consensus sequence was identical to that of a naturally occurring virus from that cluster. For two clusters the HA consensus sequences were constructed by genetic modification of the HA of a virus from that cluster. The BK79 cluster consensus virus was genetically engineered to represent the amino acid consensus sequence of the BK79 cluster by introducing the 193KN substitution in

the A/Netherlands/233/82 HA. The virus representing the WU95 cluster was genetically engineered to be antigenically WU95-like by introduction of the 145NK substitution in the NL/178/95 HA (31), and differs from the true WU95 consensus sequence by two amino acids (135TK and 262SN). A/Netherlands/301/99 was initially selected to represent the SY97 cluster. The virus grows to normal $TCID_{50}$ titers but wildtype, recombinant, and mutant viruses had very low HA titers, which potentially affected HI assay results. We therefore selected A/Netherlands/427/98 to represent the SY97 cluster. This virus differs from NL/301/99 by five amino acid substitutions that occur only in the SY97-like viruses from 1999 onwards (and that may have been responsible for the poor hemagglutination). A/Netherlands/427/98 and A/Netherlands/301/99 are antigenically very similar and differ by less than one antigenic unit in the antigenic map.

HA chimeric constructs were made by digestion of consensus HAs with SalI – NdeI (HA1, amino acids 1-301), SalI – Bsu36I (5'-HA1, amino acids 1-108), or Bsu36I – NdeI (3'-HA1, amino acids 109 – 301), and ligation in the pHW2000 expression plasmid containing the SI87 HA gene (see Fig. S1). SI87 was chosen because of its central position in the antigenic map. The NdeI site was not present in the A/Netherlands/178/95 HA sequence. Therefore NsiI was used for digestion, yielding fragments of slightly different sizes; SalI-NsiI (HA1): amino acids 1-282, and Bsu36I-NsiI (3'-HA1): amino acids 109-282. The HA of A/Netherlands/233/82 which, except for the Lysine (K) at position 193, has the amino acid consensus sequence for the BK79 cluster and was used to construct BK79 chimeric viruses. We were unable to rescue the SY97-HA1 chimeric virus.

Introduction of 50 initial point mutations was outsourced to Baseclear (Leiden, The Netherlands). Introduction of mutations for the remaining 70 mutants (additional single and combined mutations) was performed using the QuickChange multi-site directed mutagenesis kit (Stratagene, Leusden, The Netherlands) according to the manufacturers instructions. Modified plasmids were subsequently used to generate recombinant viruses consisting of the modified HA gene and remaining gene segments of the A/Puerto Rico/8/34 strain by reverse genetics as described elsewhere (64). The presence of correct regions/mutations and absence of undesired mutations was confirmed by sequence analysis of the HA gene.

The HA of the A/Singapore/92/05 A/H1N1 virus was genetically modified to match the HA1 amino acid sequence of A/Cheju/960/05 (NC99; A/New Caledonia/20/99-like) and A/Daejeion/1000/05 (SS06; A/Solomon Islands/3/06-like). The recombinant viruses were rescued in the context of A/Puerto Rico/8/34 gene segments. All influenza B viruses were wildtype strains belonging to the B/Florida/4/06 (FL06) - B/Wisconsin/1/10 (WN10) antigenic phenotypes (influenza B virus Yamagata-lineage), or B/Malaysia/2506/04 (ML04) - B/Brisbane/60/08 (BR08) antigenic phenotypes (influenza B virus Victoria-lineage).

Accession numbers

In addition to the previously published A/H3N2 sequences by Smith *et al.* (31) the sequences of the influenza A/H1N1 and B viruses used in this study are available from the GISAID EpiFlu™ Database (www.gisaid.org) and selected sequences are listed in Table S6.

Virus production

Virus stocks were generated by inoculation of Madin Darby Canine Kidney (MDCK) cells with 2mL 293T transfection supernatant. Inoculum was removed after 2 hours and replaced by MDCK infection medium, consisting of EMEM (Cambrex), 100 IU/ml penicillin, 100 µg/ml streptomycin, 2mM glutamine, 1.5 mg/ml sodium bicarbonate, 10mM Hepes, non-essential amino acids, and 25 µg/ml trypsin. Subsequently, plates were incubated at 37°C / 5% CO_2. If hemagglutination titers were below 12 hemagglutinating units, culture supernatants were subjected to ultracentrifugation to concentrate viral particles.

Hemagglutination inhibition assay

Hemagglutination inhibition (HI) assays using a panel of post-infection ferret antisera were performed using standard procedures (37). Briefly, ferret antisera were prepared upon intranasal inoculation and collecting blood 14 days (A/H3N2 and A/H1N1 viruses) or 21 days (influenza B viruses) later. An overview of antisera used in this study is provided in Table S1. Antisera were pre-treated overnight at 37 °C with receptor destroying enzyme (*Vibrio cholerae* neuraminidase) and incubated at 56°C for 1h. Two-fold serial dilutions of the antisera, starting at a 1:20 dilution, were mixed with 25 µL of a virus stock containing 4 hemagglutinating units and were incubated at 37°C for 30 minutes. Subsequently, 25 µL 1% turkey erythrocytes was added and the mixture was incubated at 4°C for 1h. Hemagglutination inhibition patterns were read and the HI titer was expressed as the reciprocal value of the highest dilution of the serum that completely inhibited agglutination of turkey erythrocytes.

Virus neutralization assay

Virus neutralization assays were performed as described previously (66). Briefly, antisera were heated for 30 minutes at 56°C and twofold serial dilutions of the antisera starting at a 1:10 dilution, were mixed 1:1 with 100 tissue culture infectious $dose_{50}$ ($TCID_{50}$) of the virus stocks. After incubation at 37°C for 2h the antiserum – virus mixture was transferred to 96-wells plates containing MDCK cells, which were washed twice with PBS prior to inoculation. Plates were incubated for 2h at 37°C, and inoculum was replaced by 200 µL infection medium. After seven days end-point dilutions were read by hemagglutination assay.

Plaque reduction assay

Plaque reduction assays were performed essentially similar to the virus neutralization assays. Inoculum consisted of antiserum dilutions mixed with 100 plaque-forming units of virus, and was replaced after two hours of inoculation by a 1:1 mixture of 2.4% Avicell (67) and double strength infection medium. Presence of plaques was read by immunostaining of influenza virus nucleoprotein expressed in the MDCK cells two days after inoculation. The end-point dilution was defined as the dilution in which there was complete inhibition of plaque formation.

Antigenic cartography

Antigenic map construction was performed as described previously (31). Briefly, antigenic cartography is a method to increase the resolution of, resolve paradoxes in, and visualize HI data. In an antigenic map, the distance between antiserum point S and antigen point A corresponds to the difference between the $\log_2$ of the maximum titer observed for antiserum S against any antigen and the $\log_2$ of the titer for antiserum S against antigen A. Thus, each titer in an HI table can be thought of as specifying a target distance for the points in an antigenic map. Modified multidimensional scaling methods are then used to arrange the antigen and antiserum points in an antigenic map to best satisfy the target distances specified by the HI data. The result is a map in which the distance between the points represents antigenic distance as measured by the HI assay in which the distances between antigens and antisera are inversely related to the $\log_2$ HI titer. Because antisera are tested against multiple antigens, and antigens tested against multiple antisera, many measurements can be used to determine the position of the antigen and antiserum in an antigenic map, thus improving the resolution of HI data.

Positions of the cluster-transition amino acid substitutions in the HA structure

The positions identified as responsible for significant antigenic change were plotted on the crystal structure of the A/Aichi/2/1968 (A/H3N2) virus HA [PDB accession code 5HMG] using MacPyMOL (The PyMOL Molecular Graphics System, Version 1.3, Schrödinger, LLC) for visualization of the trimer.

Rationale of substitution choices

Smith *et al.* (31) previously identified 67 cluster-difference substitutions associated with the antigenic evolution of A/H3N2 viruses. An amino acid substitution X to Y at location L was considered a cluster-difference substitution between clusters A and B if all (or all but one) strains in cluster A have amino acid X at location L, and all (or all but one) strains in cluster B have amino acid Y at location L. The analysis shown in Figs. S1, S2 and S3 indicated that only the amino acid positions 109 – 301 in HA1 were relevant for the antigenic phenotype. Therefore, the 54 substitutions that were within region 109 – 301 from the set of 67 were introduced in the appropriate consensus HA.

Initially all substitutions were tested individually. If single substitutions did not induce the antigenic change necessary for the cluster transition, combinations of substitutions were tested. Where possible the substitution(s) responsible for a cluster transition were tested both in chronological time order (forward mutant, e.g. HK68 – EN72 transition) and in reverse time order (reverse mutant, e.g. EN72 – HK68 transition). Fig. S4 describes each cluster transition in detail, motivates the choice of mutations introduced, and explains why particular mutants could not be tested.

Substitutions responsible for major antigenic change were located exclusively in antigenic sites A and B, and generally involved gross changes in the physical properties of the amino acids. We tested mutants with cluster-transition substitutions, and mutants with cluster-difference substitutions in sites C, D and E (Fig. S6) that had large changes in amino acid properties (Table S2) in virus neutralization assays to test if results matched those obtained by HI assay (Fig. S5).

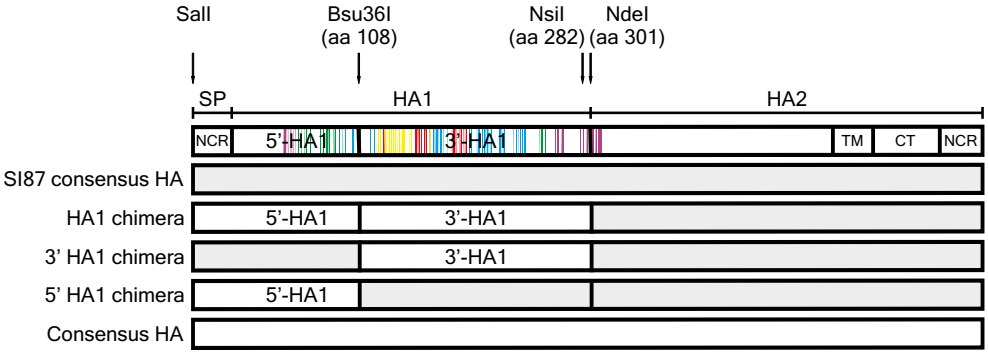


Fig. S1. Construction of HA chimeric viruses. The top bar shows the different regions of the HA gene, SP: signal peptide, NCR: non-coding region, TM: transmembrane region, and CT: cytoplasmic tail. Colors indicate the amino acid positions located in the antigenic sites, yellow: site A, red: site B, purple: site C, blue: site D, green: site E. White portions in the lower bars indicate regions derived from consensus strains, grey regions are derived from the SI87 consensus virus. Vertical arrows indicate the location of restriction sites. Amino acid positions of restriction sites are indicated between brackets.

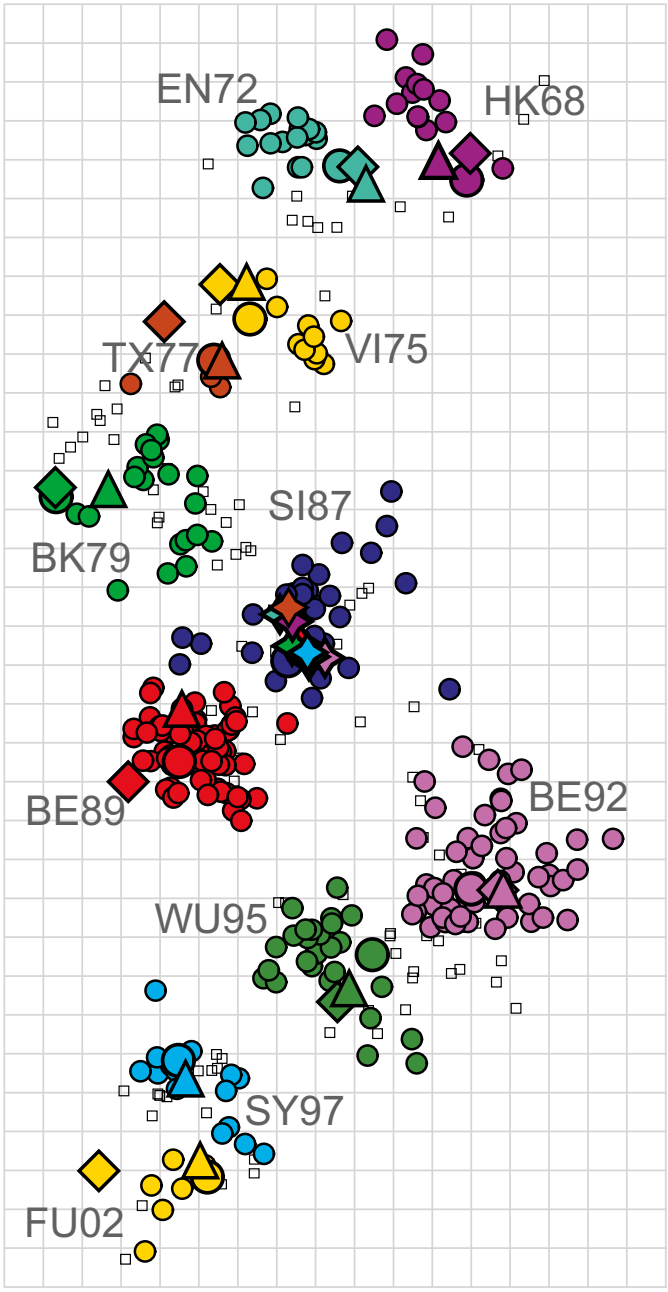


Fig. S2. Antigenic properties of consensus and chimeric viruses. The colored symbols represent strains, open squares represent antisera. Small circles represent wild type viruses, large circles recombinant consensus viruses, triangles HA1 chimeric viruses, and diamonds 3'-HA1 chimeric viruses. The viruses are color coded by antigenic similarity. Star shapes indicate 5'-HA1 chimeric viruses and are color coded by their 5'-HA1 genetic composition (see Fig. S1). Both vertical and horizontal axes represent antigenic distance. The spacing between gridlines is one antigenic unit distance, corresponding to a two-fold difference in the HI assay.

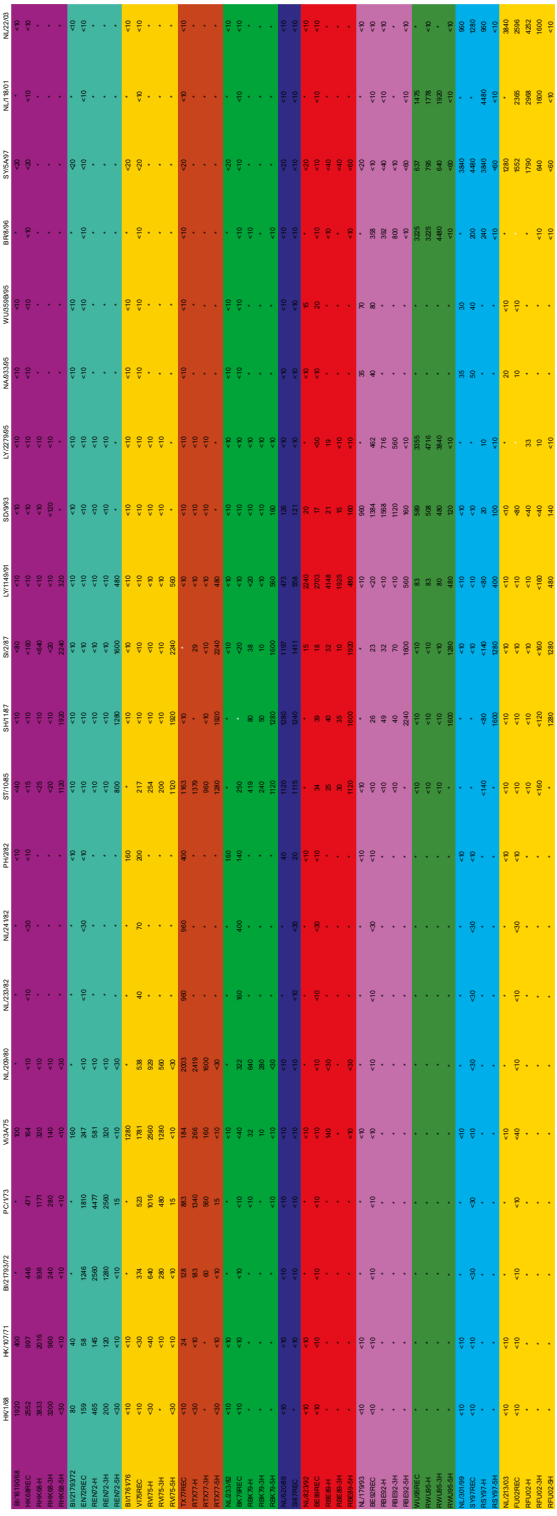


Fig. S3. HI data for chimeric viruses. Titers reflect the geometric average of multiple HI tests. The chimeras are grouped and color-coded by HA1, 3′-HA1 and 5′-HA1 genotype. The prefix R indicates that the chimera is a 7:1 recombinant virus, suffixes -H, -3H and -5H specify HA1, 3′-HA1 and 5′-HA1 chimeras respectively. A dash in the main body of the table marks the absence of (reliable) HI data for this antigen – antiserum combination, which may occur when the standard deviation of repeat measurements is more than one, a black asterisk marks the lack of cross-titration for that antigen – antiserum set. When the highest concentration of antiserum did not inhibit hemagglutination, the titer is indicated as a threshold titer (e.g. <10).

Fig. S4

Antigenic maps of cluster transitions. The figure is similar to the small maps in Fig. 1, but in addition the positions of all tested mutants is indicated. The larger circles represent the consensus viruses for the two clusters shown in the map, small circles with arrows indicate mutants, small faint colored circles are wildtype strains, and diamonds indicate the mutant with the minimal set of substitutions necessary for the cluster transition. Below each map we show the HI titers for the mutants and consensus viruses in the map. In the HI tables consensus viruses are indicated in bold and with the suffix REC. The tables are subdivided in forward and reverse mutants, and occasionally in single and double/triple mutants if more than one substitution was necessary for a cluster transition. Only those antiserum columns with at least one titer above the detection limit are shown in the tables. A dash in the main body of the table marks the absence of (reliable) HI data for this antigen – antiserum combination, which may occur when the standard deviation of repeat measurements is more than one, a black asterisk indicates that that antigen – antiserum combination was not titrated. When the highest concentration of antiserum did not inhibit hemagglutination, the titer is indicated as a threshold titer (e.g. <10). The tables can be zoomed to show details. All mutants were initially analyzed in small antigenic maps constructed using only the HI data obtained from the assays in which the mutants were tested. When the mutations responsible for the cluster transitions were identified they were integrated into, and re-analyzed in the context of, the larger 35-year map (see Smith *et al*. (31)).

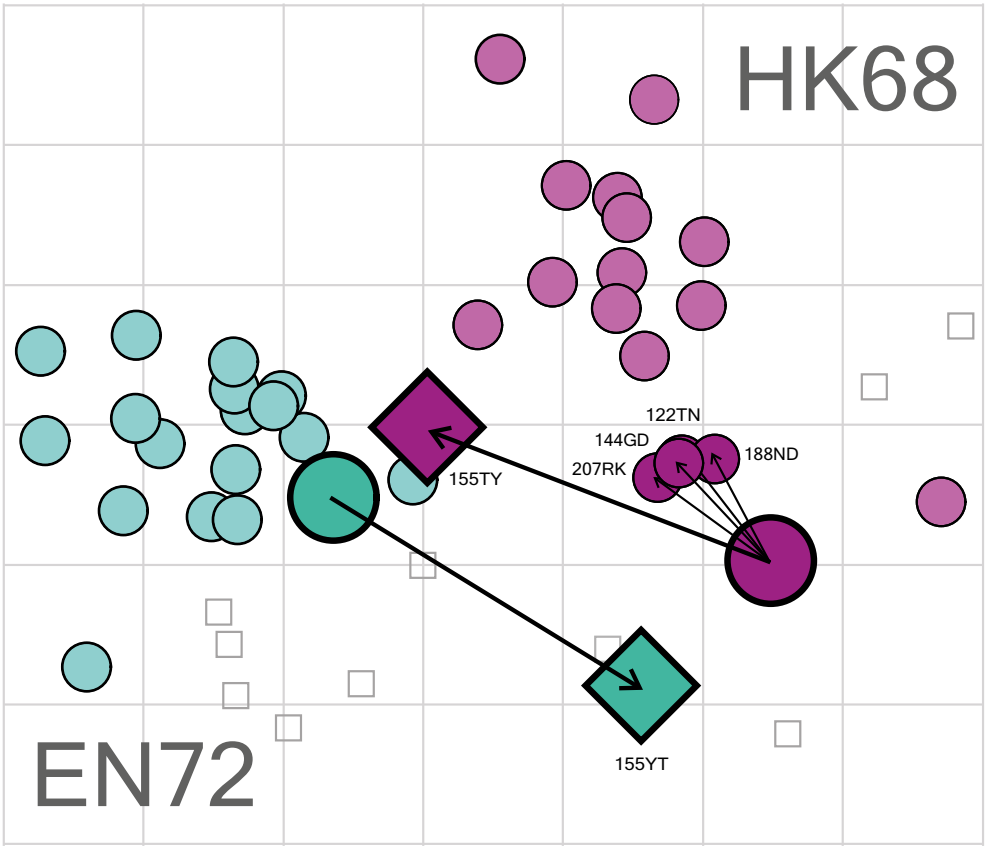


Fig. S4A. HK68 – EN72 cluster transition. The five cluster-difference substitutions were initially tested as forward mutants. Only 155TY induced a significant antigenic change. No antigenic changes were observed for the remaining four mutants. The symbols for 144GD and 122TN largely overlap and therefore only four forward mutants are visible in the map. Because new batches of antisera were raised, the 155TY forward and 155YT reverse mutants as well as the consensus strains were retested, as shown in the top and bottom right panels of the table. Hence, we conclude that the single cluster-transition substitution for the HK68 to EN72 cluster transition was 155TY.

	HK/1/68	HK/107/71	BI/21793/72	PC/1/73	VI/3A/75	HK/1A/68	HK/107A/71	BI/21793A/72	PC/1B/73
HK68REC	2606	907	277	500	155	2715	1280	320	86
HK68122TN	2560	1120	640	960	200	*	*	*	*
HK68144GD	2560	3840	1120	1280	280	*	*	*	*
HK68155TY	640	40	640	1280	240	1120	80	800	200
HK68188ND	3200	1280	800	320	240	*	*	*	*
HK68207RK	2560	1280	640	1120	280	*	*	*	*
EN72REC	218	59	1301	1920	240	226	80	2048	363

	HK/1A/68	HK/107A/71	BI/21793A/72	PC/1B/73
EN72REC	226	80	2048	363
EN72155YT	800	1920	960	160
HK68REC	2715	1280	320	86

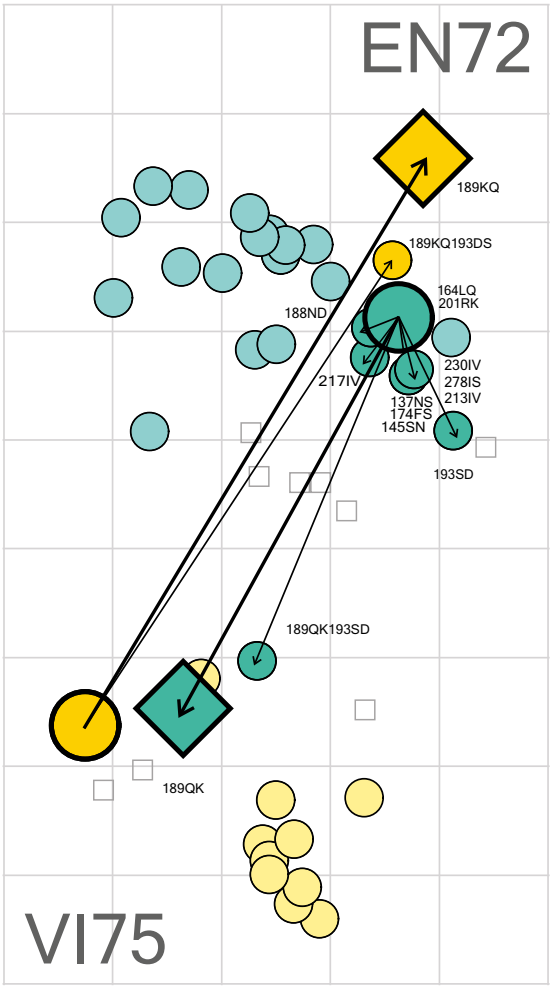


Fig. S4B. EN72 – VI75 cluster transition. All cluster-difference mutants were initially tested to the antisera as shown in the left panel of the table. Out of 11 forward pointmutants only 189QK showed substantial antigenic difference compared to the EN72 consensus virus. Substitution 193SD showed more than 1 unit antigenic divergence from the consensus virus, and was therefore combined with 189QK. Both forward and reverse mutants with only the 189 substitution elicited a larger antigenic change than the double mutants. Because new batches of antisera were raised, the mutants that were antigenically different from the EN72REC strain were re-analyzed (top and bottom right panels of the table). Therefore we conclude that the only cluster-transition substitution for the EN72 to VI75 cluster transition was 189QK.

	HK/1/68	HK/107/71	BI/21793/72	PC/1/73	VI/3A/75	NL/209/80	HK/1A/68	HK/107A/71	BI/21793A/72	SL/840/74	VI/3D/75	NL/209B/80
EN72REC	218	59	1301	1920	240	<10	226	80	2048	240	104	20
EN72137NS	240	50	1600	1920	400	<10	*	*	*	*	*	*
EN72145SN	160	40	1280	1280	320	<10	*	*	*	*	*	*
EN72164LQ	160	40	1600	2560	280	<10	*	*	*	*	*	*
EN72174FS	240	60	1920	2560	320	<10	*	*	*	*	*	*
EN72193SD	240	120	3840	4480	640	<10	*	*	*	*	*	*
EN72201RK	240	60	1280	2240	320	<10	*	*	*	*	*	*
EN72213IV	240	70	1600	3200	320	<10	*	*	*	*	*	*
EN72217IV	240	60	1280	2560	280	<10	*	*	*	*	*	*
EN72230IV	240	50	1600	2240	280	<10	*	*	*	*	*	*
EN72278IS	240	80	1600	3200	320	<10	*	*	*	*	*	*
EN72189QK	80	60	1280	2560	1280	2240	20	60	1920	480	2240	1920
EN72189QK193SD	120	40	640	*	1280	320	20	70	1120	240	960	200
VI75REC	<10	<30	397	570	1070	452	<10	32	378	40	640	358

	HK/1A/68	HK/107A/71	BI/21793A/72	SL/840/74	VI/3D/75	NL/209B/80
VI75REC	<10	32	378	40	640	358
VI75189KQ	70	40	320	30	40	20
VI75189KQ193DS	100	20	640	40	40	25
EN72REC	226	80	2048	240	104	20

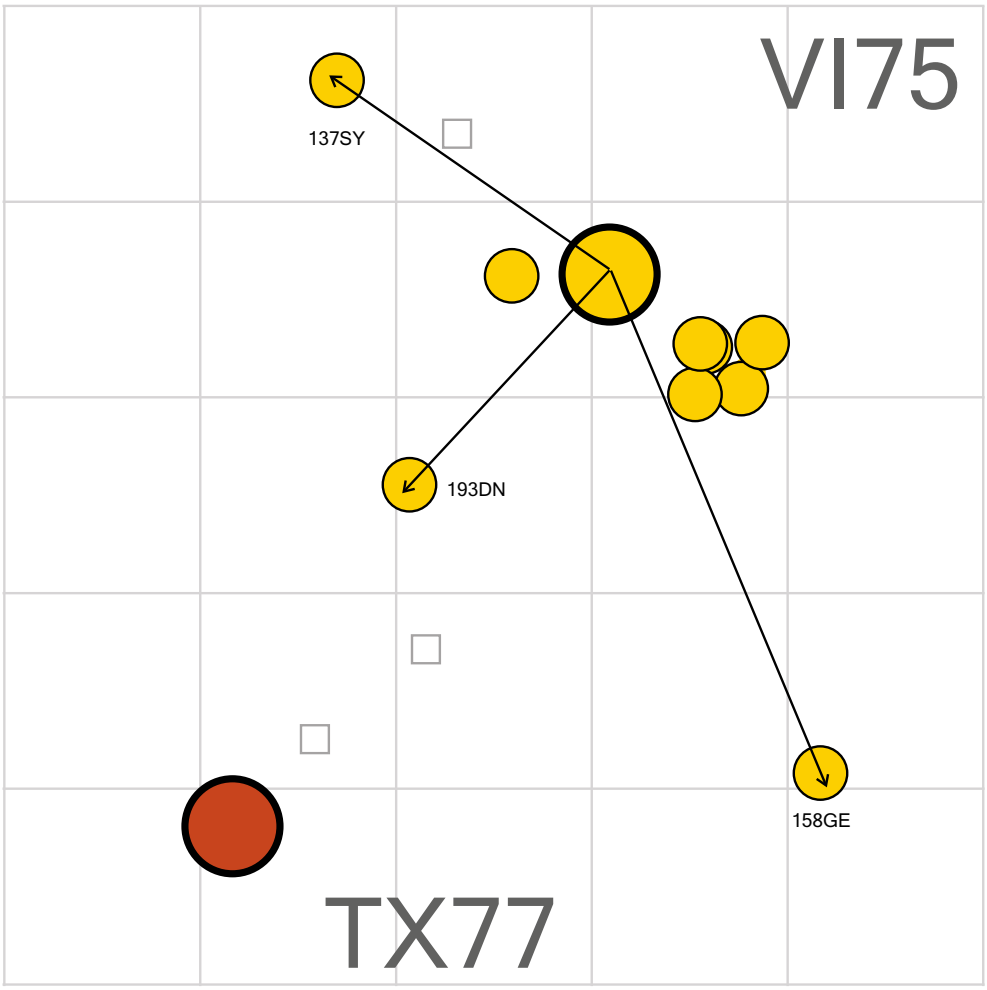


Fig. S4C. VI75 – TX77 cluster transition. Nine cluster-difference substitutions were tested as single mutants initially, none of which could induce the cluster transition. Three mutants (137SY, 193DN and 158GE were combined to make double and triple mutants. Both 158GE193DN and 137SY158GE193DN have a similar phenotype as the TX77 strains. The directionality of the reverse mutant 137YS158EG193ND is incorrect and does not revert the phenotype to VI75-like, while the double reverse mutant 158EG193ND displays similar antigenic properties as the VI75-like strains. Hence, we conclude that 158GE and 193ND together were the cluster-transition substitutions for the VI75 – TX77 cluster transition.

	HK/107/71	BI/21793/72	PC/1/73	VI/3A/75	NL/209/80	ST/10/85	HK/107A/71	BI/21793A/72	VI/3D/75	NL/209B/80	PH/2/82	ST/10A/85
VI75REC	<30	397	570	1070	452	232	32	378	640	358	211	358
VI75158GE	<10	80	400	320	640	160	*	*	*	*	*	*
VI75164QL	<10	320	640	1920	640	240	*	*	*	*	*	*
VI75174SF	<10	480	640	1600	640	160	*	*	*	*	*	*
VI75193DN	25	560	1920	1920	2240	960	*	*	*	*	*	*
VI75201KR	20	560	800	1920	640	160	*	*	*	*	*	*
VI75213VI	10	640	640	2240	640	200	*	*	*	*	*	*
VI75230VI	<10	320	640	1280	400	140	*	*	*	*	*	*
VI75137SY	40	640	1120	1280	160	70	*	*	*	*	*	*
VI75260MI	10	640	800	1920	640	200	*	*	*	*	*	*
VI75137SY158GE	*	100	560	240	280	240	*	*	*	*	*	*
VI75137SY193DN	*	280	960	960	1280	960	*	*	*	*	*	*
VI75158GE193DN	*	<80	1120	240	3200	2240	40	320	1120	2240	960	1600
VI75137SY158GE193DN	*	<80	640	160	1920	960	*	*	*	*	*	*
TX77REC	14	98	854	120	1523	781	26	302	557	2030	301	943

	HK/107A/71	BI/21793A/72	VI/3D/75	NL/209B/80	PH/2/82	ST/10A/85
TX77REC	26	302	557	2030	301	943
TX77158EG193ND	80	960	640	320	80	320
TX77137YS158EG193ND	10	–	358	400	253	453
VI75REC	32	378	640	358	211	358

Fig. S4D-I. TX77 – BK79 cluster transition. Because of the limited number of strains in the TX77 cluster, which consists of only 3 strains (Smith *et al.* (31)), there are many cluster-difference substitutions between the TX77 and BK79 clusters. All 13 cluster-difference substitutions were initially tested as single mutants. Only TX77 156KE was substantially different from the TX77 consensus virus, and resembled the BK79 consensus virus antigenically (Fig. S4 D-I and Fig. S 4 D-II). Addition of the cluster-difference substitutions with large changes in amino acid properties—e.g. large changes in amino acid volume and/or charge changes—or substitutions near the receptor binding site to the TX77 156KE mutant did not induce additional antigenic change. Similarly, BK79 156EK reverted the antigenic phenotype to TX77-like and additional substitutions did not change the antigenic the antigenic properties as compared to 156EK alone. Thus, we conclude that 156KE alone was responsible for the antigenic difference between the TX77 and BK79 clusters.

	BI/21793A/72	PC/1B/73	VI/3D/75	NL/241/82	PH/2/82	ST/10A/85	SI/2/87	BI/21793/72	PC/1/73	VI/3A/75	TX/1C/77	BK/1/79	NL/209B/80	NL/209/80	ST/10/85
TX77REC	390	164	761	921	316	890	28	101	900	124	2560	160	2455	1617	1091
TX77133NS	640	240	1120	800	400	1120	20	80	480	100	*	*	*	1693	480
TX77143PS	145	77	994	689	320	1253	25	80	1280	160	*	*	2933	2074	1280
TX77146GS	240	240	640	800	320	960	50	240	1920	200	*	*	*	2263	1920
TX77156KE	1112	262	1016	4256	2395	2189	160	160	640	320	1920	2560	5431	5724	3840
TX77160TK	320	160	640	480	160	800	35	200	1920	160	*	*	*	1810	960
TX77172GD	320	160	640	400	200	640	70	*	*	240	*	*	*	1280	1920
TX77173NK	640	196	800	640	320	905	*	*	*	100	*	*	*	1893	*
TX77197QR	320	160	640	480	240	640	70	400	1600	160	*	*	*	1568	800
TX77213IV	320	150	640	800	299	640	*	*	*	80	*	*	*	1466	*
TX77217VI	640	320	800	640	320	1120	20	120	1120	160	*	*	*	1920	800
TX77244VL	320	160	640	560	240	640	25	100	1280	160	*	*	*	1197	640
TX77248NT	640	169	640	640	320	1076	10	*	*	80	*	*	2560	1280	*
TX77307KR	*	120	*	*	400	560	*	*	*	100	*	*	*	1280	*
TX77143PS156KE	160	57	905	4148	2560	3135	226	*	*	*	2560	3840	6271	*	*
TX77146GS156KE	960	640	1600	8960	5120	1920	280	*	*	*	3840	2560	3840	*	*
TX77156KE172DG	1280	640	1280	5120	3200	1920	160	*	*	*	2560	1600	4480	*	*
TX77156KE197QR	1120	640	1280	5120	2560	1280	120	*	*	*	2560	1280	3200	*	*
TX77156KE244VL	1280	640	1280	8960	5120	1600	160	*	*	*	2560	1600	6400	*	*
TX77156KE248TN	1280	320	960	5120	*	2560	80	*	*	*	*	*	10240	*	*
BK79REC	34	<10	320	2871	1087	1613	48	*	*	*	1197	1012	3401	2074	1431

	BI/21793A/72	PC/1B/73	VI/3D/75	NL/209B/80	NL/241/82	PH/2/82	ST/10A/85	NL/21793/72	VI/3/75	TX/1C/77	BK/1/79	NL/209/80	PH/2A/82	PH/2B/82	ST/10/85
BK79REC	34	<10	320	3401	2871	1087	1613	40	160	1197	1012	2074	640	2560	1431
BK79143SP	253	62	320	2715	2873	1417	1810	226	226	1280	1280	1810	640	2560	1197
BK79156EK	35	<80	320	2024	359	320	640	40	226	1280	160	905	160	960	453
BK79248TN	80	45	640	3200	*	2560	3200	60	640	*	*	2560	*	*	2240
BK79143SP156EK	113	40	320	1810	320	320	506	80	320	1280	80	640	120	960	640
BK79143SP248TN	*	80	*	*	<20	2560	*	160	640	*	*	3840	*	*	2560
BK79156EK172GD	90	80	320	2560	640	320	640	*	*	1600	80	*	*	*	*
BK79156EK244VL	80	160	640	3840	1280	480	1280	*	*	1920	140	*	*	*	*
BK79156EK248TN	80	40	640	2560	320	*	1280	*	*	*	*	*	*	*	*
BK79143SP156EK248TN	160	<60	320	1280	320	*	640	*	*	*	*	*	*	*	*
TX77REC	390	164	761	2455	921	316	890	274	640	2560	160	1617	160	1280	1091

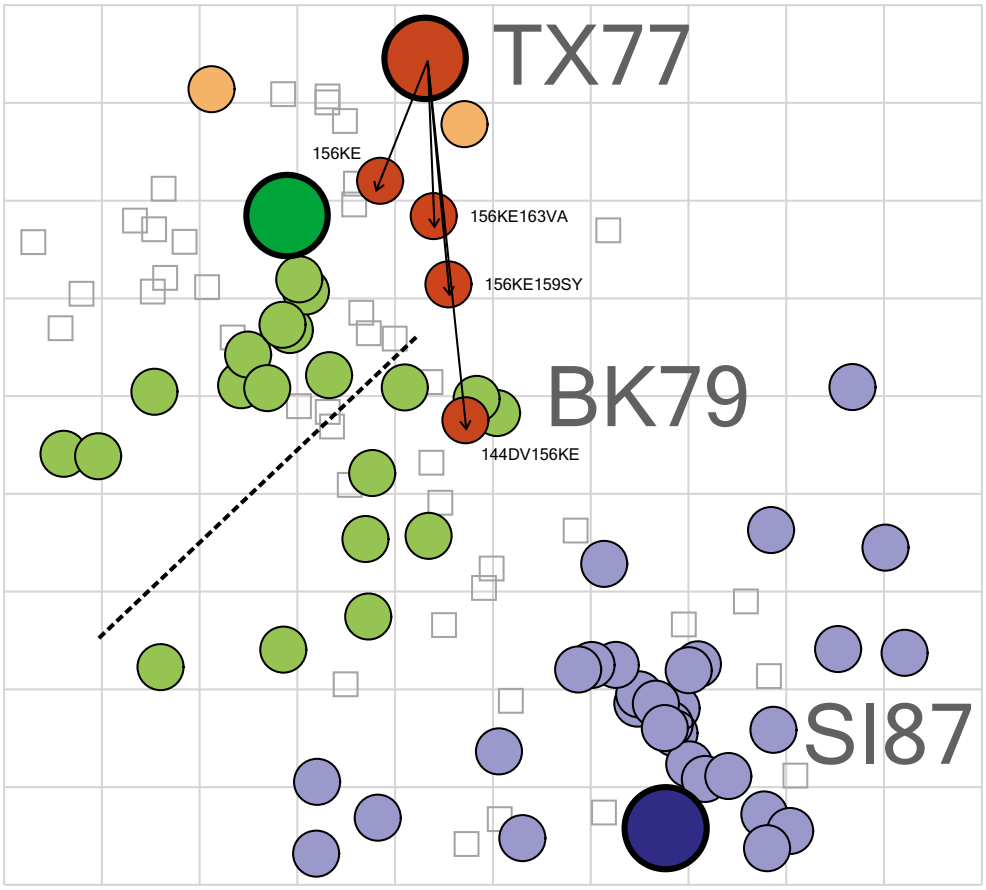


Fig. S4D-II. Antigenic change within the BK79 antigenic cluster. The viruses representing the antigenic clusters have an HA1 amino acid sequence that is the consensus sequence of all strains from the respective cluster. Such a virus is usually an adequate representative of the antigenic consensus of the cluster, however for the BK79 cluster this is not the case. We added to the TX77 156KE cluster-transition mutant three substitutions that occurred within the BK79 cluster and which were different for BK79 strains located above or below the dashed line: 144DV, 163VA and 159SY. All three TX77 mutants are BK79-like, and 159SY and 144DV changed antigenically as compared to 156KE alone and are located near the main body of the BK79 cluster. The addition of substitutions 144DV or 159SY to the TX77 156KE cluster transition mutant thus resulted in viruses with antigenic properties that represented the BK79 cluster better than the amino acid consensus virus of the BK79 cluster that does not have these substitutions.

	BI/21793/72	PC/1B/73	VI/3D/75	TX/1/77	BK/1/79	NL/209B/80	PH/2/82	NL/241/82	ST/10A/85	SH/11A/87	SI/2/87	LY/1149A/91	BE/32/92
TX77REC	480	400	1120	2560	160	3840	320	1600	800	<30	40	<10	<10
TX77144DV156KE	80	40	480	1280	640	2560	2240	1920	2560	400	1280	160	320
TX77156KE159SY	2560	1120	1280	3200	1920	4480	4480	8960	6400	1280	160	320	70
TX77156KE163VA	2240	960	1280	2560	1600	5120	4480	7680	2560	<30	240	120	40
BK79REC	40	<10	320	1120	800	3200	1280	6400	1280	<30	80	<10	<10
SI87REC	<10	<10	<10	<20	<40	<10	<10	<40	640	1280	5120	320	320

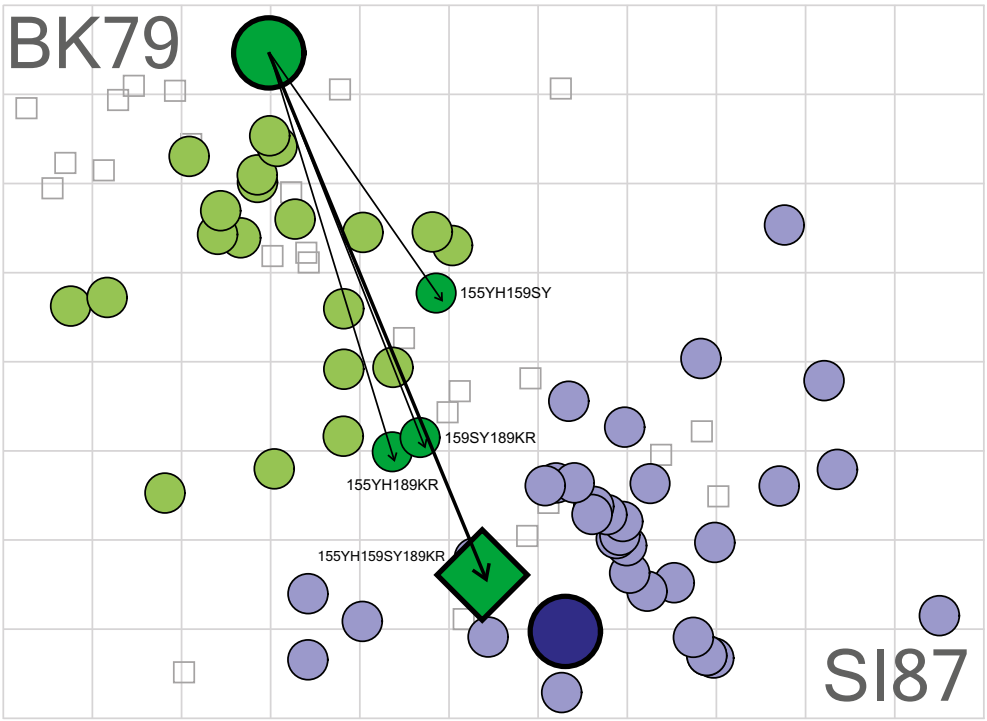


Fig. S4E. BK79 – SI87 cluster transition. Three cluster-difference substitutions were associated with the BK79 to SI87 cluster transition: 124GD, 155YH and 189KR. These mutations were initially introduced in the wildtype HA of NL/233/82, which has the amino acid consensus sequence of the BK79 cluster except for 193NK. Both 155YH and 189KR changed antigenically but did not induce the full transition to SI87-like, and also combination of 155YH and 189KR did not entirely change the virus to SI87-like. Because BK79 193K viruses reacted with low titers to available antisera in HI tests, we introduced 193KN in the HA of A/Netherlands/233/82 to obtain the exact BK79 cluster genetic consensus. HA position 159 changed from Serine (S) to Tyrosine (Y) during the period that the BK79-like strains circulated, and structural analysis of HA suggested that this substitution could have a major impact on the antigenic properties of BK79-like viruses. The combination 155YH159SY189KR induced a 5.8 antigenic unit change, and had the same phenotype as the SI87 consensus virus. Interestingly, the 155HY159YS189RK reverse mutant was constructed, but we were unable to rescue this virus. Thus, we conclude that three substitutions were responsible for the antigenic difference between the BK79 and SI87 clusters: 155YH, 159SY and 189KR.

	BI/21793A/72	PC/1B/73	VI/3D/75	NL/209B/80	NL/241/82	PH/2/82	ST/10A/85	SH/11A/87	SI/2/87	HK/34A/90	LY/1149A/91
BK79REC	40	<40	320	3505	3474	1140	1280	<10	48	<20	<10
BK79155YH159SY	70	40	320	6400	3200	1280	3840	2240	480	40	160
BK79155YH189KR	80	<10	<10	320	320	160	140	560	640	<30	40
BK79159SY189KR	40	30	<10	320	320	320	640	960	320	<30	160
BK79155YH159SY189KR	40	<10	<10	160	80	80	200	1280	140	80	200
SI87REC	<10	<10	<10	<40	<10	<10	453	1141	2065	361	312

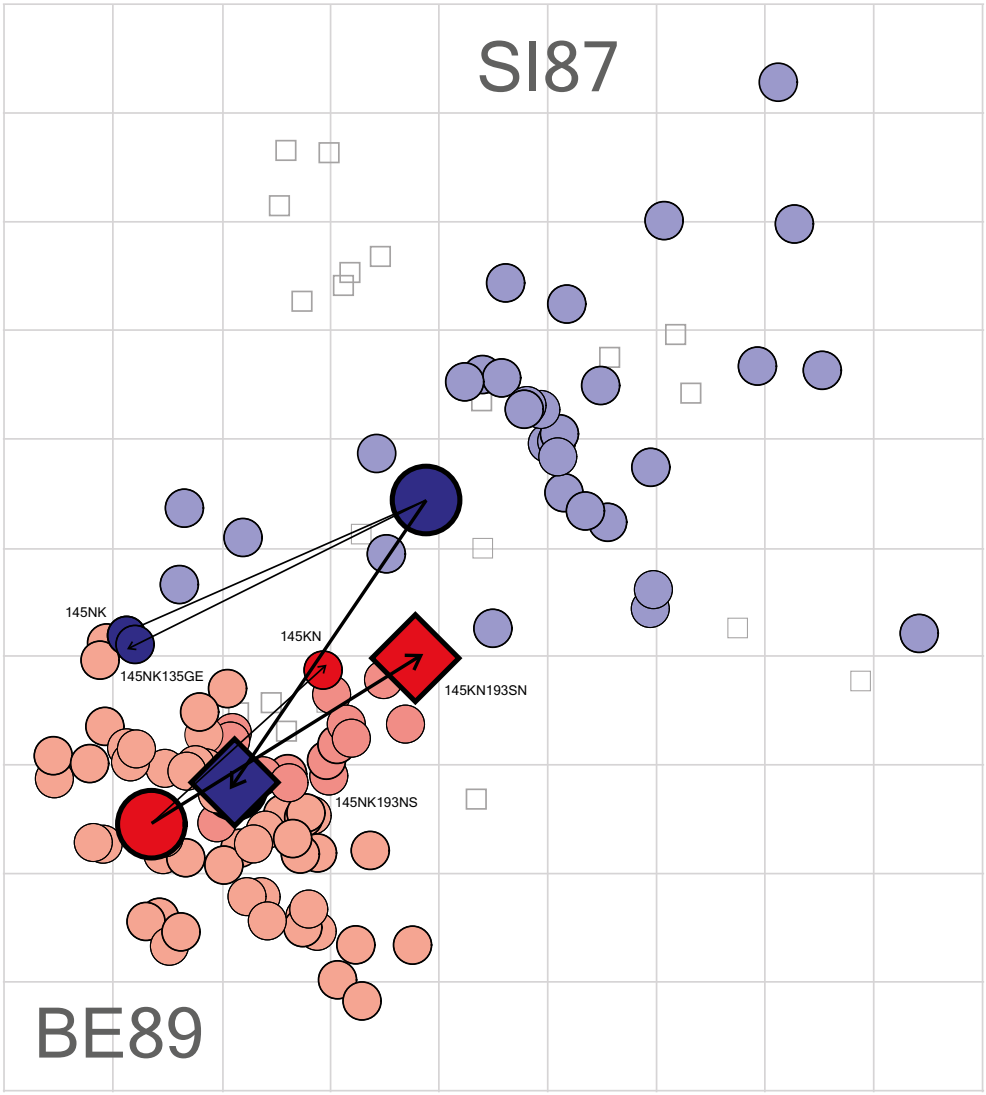


Fig. S4F. SI87 – BE89 cluster transition. Only one cluster-difference substitution occurs between the SI87 and BE89 clusters. This substitution, 145NK, has been described previously by Smith *et al.* (31). The 135GE and 193NS substitutions, which have few exceptions to the cluster-difference substitution criterion, were added to 145NK to test if these have additional significance to 145NK alone. The 145NK135GE mutant did not change antigenically compared to 145NK. Addition of 193NS did not affect the distance from the SI87 consensus virus, but better resembles the antigenic properties of the BE89 consensus virus than 145NK alone (see Fig. S10 for details). The reverse 145KN and 145KN193SN mutants, although located on the interface of the SI87 and BE89 clusters, are 1.7 and 1.5 units away from the SI87 consensus virus, respectively. Hence we conclude that the cluster-transition substitution for the SI87 to BE89 cluster transition was 145NK.

	ST/10/85	SI/2/87	SH/11/87	BE/353/89	HK/34/90	VI/2/90	LY/1149/91	BE/32/92	NL/823/92	PA/548/92	NL/209/80
SI87REC	525	1277	1280	159	264	660	339	59	403	186	<40
SI87145NK	160	109	291	113	57	905	640	<10	716	506	35
SI87145NK135GE	320	80	800	160	40	1280	1280	<10	1280	1280	*
SI87145NK193NS	139	144	396	104	160	1485	1253	<10	1677	883	20
BE89REC	<30	25	35	480	25	700	1041	<40	1459	1329	<10

	ST/10/85	ST/10A/85	SH/11/87	SH/11A/87	SI/2/87	BE/353/89	HK/34/90	BE/32/92	SD/9/93	JO/33/94
BE89REC	<30	<10	35	<160	25	480	25	<40	<36	<10
BE89145KN	80	160	392	560	689	212	320	80	109	98
BE89145KN193SN	160	160	480	640	1016	196	784	145	113	120
SI87REC	525	559	1280	1140	1277	159	264	59	120	<10

ST/10A/85	ST/10B/85	SH/11A/87	SH/11B/87	LY/2279A/95	BE/353C/89	HK/34A/90	VI/2B/90	LY/1149A/91	NL/823B/92	PA/548B/92	LY/2279B/95	WU/359B/95	BR/8B/96
559	1465	1140	635	<30	70	237	1252	320	640	320	<10	<10	<10
100	400	240	160	25	*	*	*	*	*	*	*	*	*
*	*	*	*	*	*	*	*	*	*	*	*	*	*
80	*	506	*	80	60	160	1920	1920	1920	1280	160	40	120
<10	80	53	60	<40	35	<40	664	1970	1970	1970	<10	<20	20

ST/10B/85	SH/11B/87	LY/1149/91	PA/548/92	NL/823/92	BE/353C/89	HK/34A/90	VI/2B/90	LY/1149A/91	NL/823B/92	PA/548B/92	JO/33A/94	WU/359B/95	BR/8B/96
80	60	1041	1329	1459	35	<40	664	1970	1970	1970	<10	<20	20
640	560	1737	1810	2560	*	*	*	*	*	*	*	*	*
*	*	*	*	*	80	160	1280	1280	3200	2560	35	10	20
1465	635	339	186	403	70	237	1252	320	640	320	<30	<10	<10

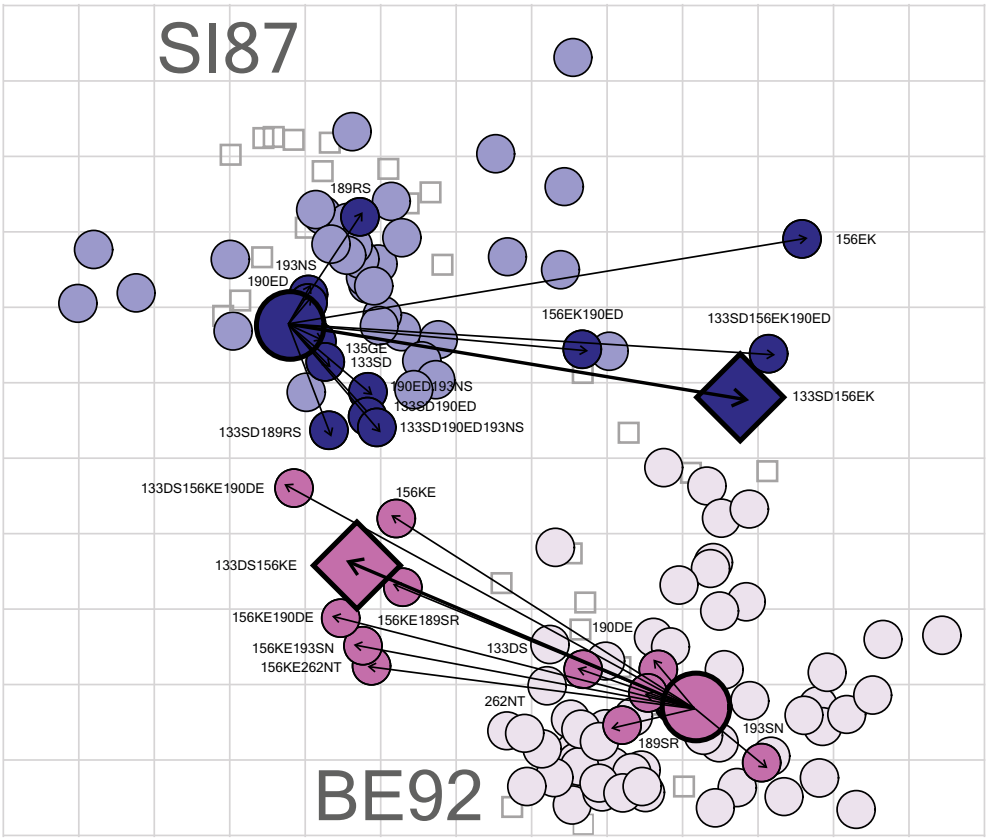


Fig. S4G. SI87 – BE92 cluster transition. Given that from phylogenetic analysis the most likely ancestor of the BE92 cluster was a SI87-like strain rather than a BE89-like strain (31), cluster- difference substitutions as defined by Smith et al. (31) were redefined comparing SI87 and BE92-clusters (Table S2). None of the cluster-difference substitutions 133SD, 190ED or 193NS, alone or as double or triple mutants, caused the antigenic change from the SI87 to the BE92 cluster. The reverse mutants were also tested, but similarly none could revert the antigenic phenotype to SI87-like. We therefore tested three additional substitutions that had few exceptions to the cluster-difference substitution criterion. Substitution 156EK resulted in a substantial antigenic difference from the SI87-like strains, but did not match the strains of the BE92 cluster antigenically. It was therefore combined with substitutions that induced more than 1 unit antigenic change from either SI87 or BE92. The combinations with the 156 substitution resulted in an antigenic change that was at least 2 units larger than any other combination (between 3.8 – 6.3 antigenic units for combinations with the 156 substitution, between 1.4 – 1.8 for any other combination). For the reverse mutants no significant changes were observed by addition of other substitutions to 156KE. The addition of 133SD to 156EK in the forward mutant, however, skewed the phenotype to BE92-like, without affecting antigenic distance from the SI87 consensus virus. Since 133SD alone, or combined with all other substitutions except 156KE, did not result in evident antigenic change, we conclude that 133SD is considered to be accessory to the 156EK for this cluster transition (see Fig. S10 for details). We therefore conclude that 156EK was the cluster-transition substitution for the SI87 to BE92 cluster transition.

	ST/10A/85	ST/10B/85	SH/11A/87	SH/11B/87	SI/2/87	LY/1149/91	SD/9/93	LY/2279A/95	ST/10/85	HK/34/90	BE/32/92	PA/548/92	JO/33/94	LY/2279/95
SI87REC	640	1280	1353	800	1378	282	120	<30	599	320	51	160	<20	<10
SI87133SD	320	1120	800	400	320	480	160	15	*	*	*	*	*	*
SI87135GE	1120	1120	1280	640	480	640	90	60	*	*	*	*	*	*
SI87189RS	640	1280	1600	640	1600	80	80	<10	*	*	*	*	*	*
SI87190ED	640	1920	2560	960	1280	560	160	15	*	*	*	*	*	*
SI87193NS	800	1600	1280	1120	960	480	70	20	*	*	*	*	*	*
SI87156EK	*	*	*	*	1280	<10	20	*	100	640	160	40	40	<30
BE92REC	<20	35	<10	<10	33	<10	651	367	<30	1280	95	<10	1600	640

	SH/11A/87	SI/2/87	BE/353C/89	HK/34A/90	BE/32/92	SD/9B/93	JO/33A/94	VI/2B/90	LY/1149A/91	NL/823B/92	PA/548B/92	LY/2279B/95	NA/933/95	WU/359B/95	BR/8B/96	NL/209/80	PH/2/82	ST/10A/85	SD/9A/93	LY/2279A/95
SI87REC	1353	1378	92	193	51	<20	<30	1197	320	640	320	<10	<10	<20	<10	<40	20	640	20	<30
SI87133SD156EK	<20	320	20	640	160	40	80	<10	<10	<10	<10	<10	<10	<20	40	*	*	*	*	*
SI87156EK190ED	80	960	40	1280	160	80	140	20	40	20	40	<10	<10	<20	35	*	*	*	*	*
SI87133SD156EK190ED	<20	140	20	1280	160	80	80	<10	20	<10	20	<10	<10	10	40	*	*	*	*	*
SI87133SD189RS	1280	280	80	320	80	160	20	*	*	*	*	*	*	*	*	<10	20	800	40	40
SI87133SD190ED	1280	160	160	640	160	160	20	*	*	*	*	*	*	*	*	10	25	640	80	40
SI87190ED193NS	1280	640	100	560	200	160	30	*	*	*	*	*	*	*	*	20	40	640	40	40
SI87133SD190ED193NS	960	100	80	640	320	160	20	*	*	*	*	*	*	*	*	20	20	560	80	40
BE92REC	<10	33	<40	350	95	547	1213	<10	<40	17	<30	599	20	80	320	<10	<10	<20	80	367

	NL/209B/80	ST/10A/85	SH/11A/87	SI/2/87	BE/353C/89	HK/34A/90	VI/2B/90	LY/1149A/91	BE/32/92	NL/823B/92	PA/548B/92
BE92REC	<10	<20	<10	33	<40	350	<10	<40	95	17	<30
BE92133DS	<10	20	20	80	20	160	20	40	80	40	40
BE92189SR	<10	<10	<10	40	<10	320	<10	80	160	40	15
BE92190DE	<10	30	<10	240	<10	240	<10	40	80	25	<10
BE92193SN	<10	<10	<10	40	<10	100	<10	<10	20	<10	<10
BE92262NT	<10	<10	<10	30	20	240	<10	15	80	20	<10
BE92156KE	40	640	1280	89	80	200	1120	280	57	560	400
SI87REC	<10	640	1353	1378	92	193	1197	320	51	640	320

	NL/209/80	ST/10/85	SH/11/87	SI/2/87	BE/353/89	HK/34/90	VI/2/90	LY/1149/91	BE/32/92	NL/823/92	PA/548/92
BE92REC	<10	<30	<60	33	10	1280	20	<10	95	<10	<10
BE92156KE133DS	20	240	800	160	80	320	960	320	40	320	160
BE92156KE189SR	35	160	640	80	<10	640	1280	480	160	320	320
BE92156KE190DE	<10	320	640	320	40	160	800	480	40	320	320
BE92156KE193SN	<10	320	640	80	40	320	640	160	40	320	240
BE92156KE262NT	<10	160	640	80	40	320	640	160	40	160	160
BE92133DS156KE190DE	*	*	*	480	*	*	*	*	20	*	*
SI87REC	<40	599	1012	1378	80	320	1280	282	51	320	160

SD/9B/93	JO/33A/94	LY/2279B/95	NL/209/80	ST/10/85	SH/11/87	HK/34/90	VI/2/90	LY/1149/91	NL/823/92	PA/548/92	SD/9/93	JO/33/94	LY/2279/95
547	1213	599	<10	<30	<60	1280	20	<10	<10	<10	651	1600	640
800	3200	320	*	*	*	*	*	*	*	*	*	*	*
160	3840	320	*	*	*	*	*	*	*	*	*	*	*
320	3840	640	*	*	*	*	*	*	*	*	*	*	*
80	1600	280	*	*	*	*	*	*	*	*	*	*	*
640	3200	480	*	*	*	*	*	*	*	*	*	*	*
80	2240	320	20	320	640	560	960	160	320	240	80	1120	320
<20	<30	<10	<40	599	1012	320	1280	282	320	160	120	<20	<10

SD/9/93	JO/33/94	LY/2279/95	SH/11A/87	BE/353C/89	HK/34A/90	VI/2B/90	LY/1149A/91	NL/823B/92	PA/548B/92	SD/9B/93	LY/2279B/95	WU/359B/95	BR/8B/96
651	1600	640	<10	<40	350	<10	<40	17	<30	547	599	80	320
40	320	140	*	*	*	*	*	*	*	*	*	*	*
40	1120	320	*	*	*	*	*	*	*	*	*	*	*
40	640	320	*	*	*	*	*	*	*	*	*	*	*
40	960	320	*	*	*	*	*	*	*	*	*	*	*
40	960	320	*	*	*	*	*	*	*	*	*	*	*
*	640	*	1280	20	160	960	480	400	480	20	160	20	60
120	<20	<10	1353	92	193	1197	320	640	320	<20	<10	<20	<10

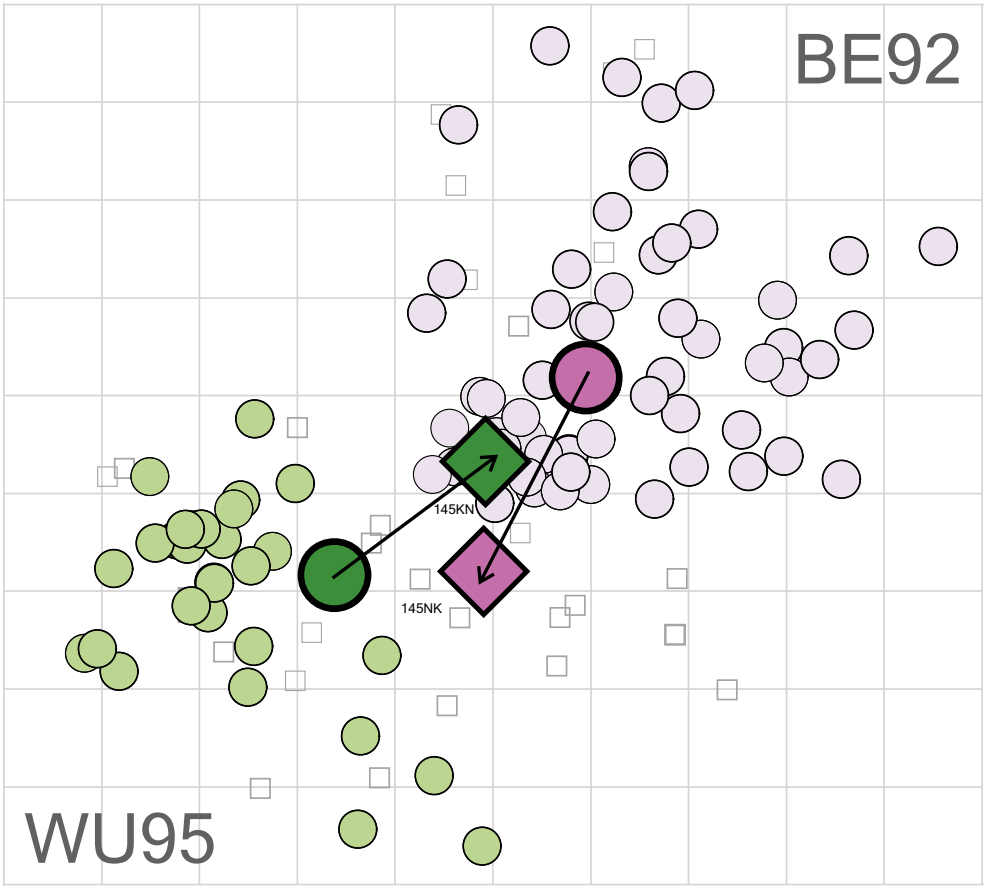


Fig. S4H. BE92 – WU95 cluster transition. Smith *et al.* (31) previously showed that the 145NK substitution is responsible for the BE92 – WU95 cluster transition. For this study we additionally tested the reverse mutant WU95145KN, which reverts the phenotype to BE92-like. Thus, we conclude that the single cluster-transition substitution for the BE92 to WU95 cluster transition was 145NK.

	HK/34A/90	VI/2B/90	LY/1149A/91	LY/1149/91	BE/32/92	NL/823B/92	PA/548B/92	SD/9/93	SD/9B/93	JO/33A/94	LY/2279B/95	LY/2279/95	NA/933/95	WU/359B/95	BR/8/96	BR/8B/96	SY/5A/97
BE92REC	402	<10	<10	<10	83	12	<10	736	460	1213	423	610	20	76	360	320	<30
BE92N145K	160	<10	40	25	20	80	10	320	40	320	1600	1920	80	320	640	800	<30
WU95K145N	640	20	20	60	240	20	<10	2080	320	2560	1120	1920	106	226	1600	1280	<30
WU95REC	104	<40	88	109	21	83	26	560	102	190	1942	2581	102	695	2328	2560	479

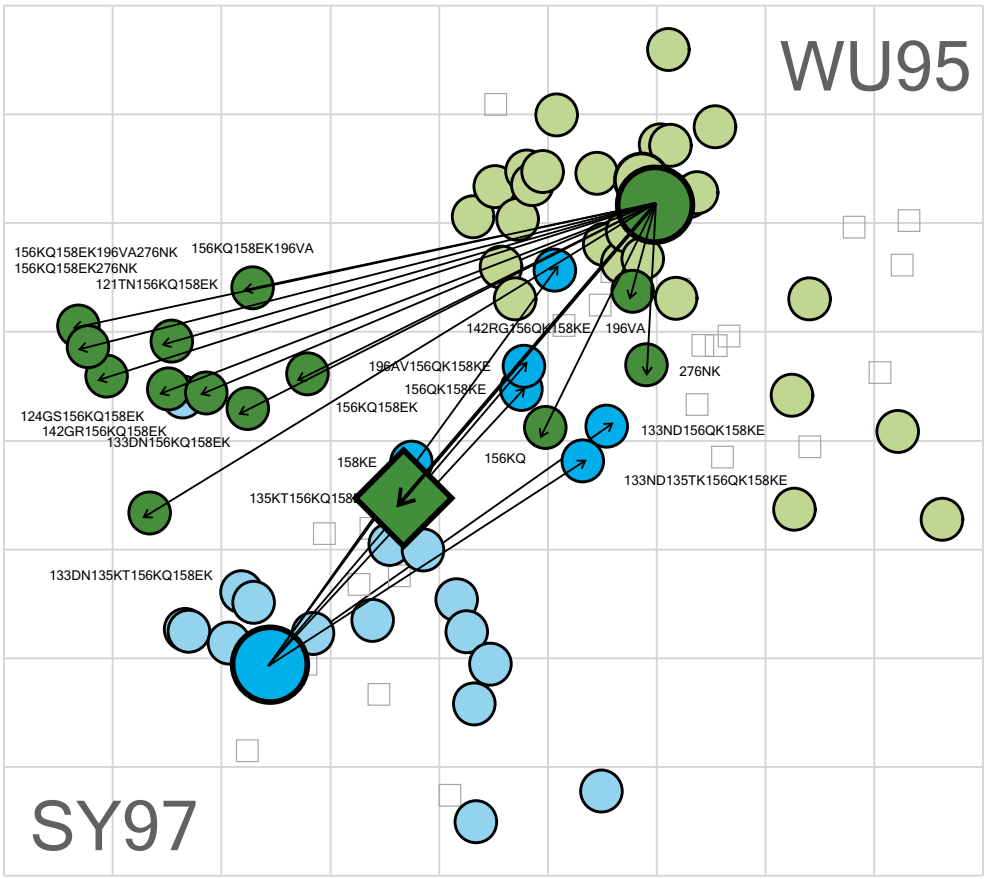


Fig. S4I. WU95 – SY97 cluster transition. Four cluster-difference substitutions occur between the WU95 and SY97 clusters. We were unable to rescue the 158EK mutant, and therefore could not test it as single mutant. The combination of 156KQ and 158EK induced an antigenic change of 3.6 antigenic units from WU95REC, which is similar to specific strains of the SY97 antigenic phenotype. Its position, however, is away from the main body of the SY97 cluster. Addition of the remaining cluster-difference mutations to this double mutant did not result in a position more similar to the majority of the SY97-like strains. To test if substitutions that do not fit the cluster-difference criterion were involved in this cluster transition we combined the 156KQ158EK mutant with five additional substitutions. Four (121TN, 124GS, 133DN, and 142GR) had few exceptions to the criterion for a cluster-difference substitution and may therefore be involved in the antigenic change, and one (135KT) is typical for late WU95-like strains and is present in the majority of strains from the WU95 cluster. Furthermore, the 133DN135KT combination was added because it introduces a glycosylation pattern. Addition of 121TN, 124GS, 133DN, or 142GR did not result in a position closer to the main body of the SY97 cluster. Addition of 135KT skewed the phenotype towards that typical of the majority of SY97-like strains without increasing antigenic distance from WU95REC (see Fig. S10 for details). Addition of both 133DN and 135KT further increased the distance from WU95REC but within the SY97 cluster. The reverse mutant SY97 156QK158KE reverted the phenotype to WU95-like. We therefore conclude that the cluster-transition substitutions for the WU95 to SY97 cluster transition were 156KQ and 158EK.

	LY/1149A/91	SD/9B/93	JO/33A/94	LY/2279A/95	NA/933/95	WU/359B/95	BR/8B/96	SY/5A/97	NL/118/01	FU/411/02	NL/22/03	NL/88/03	WE/001/04	WN/67/05	NL/42/06
SY97REC	44	<10	<10	65	43	48	95	3200	3200	747	842	1016	105	38	40
SY97158KE	280	40	80	320	160	160	640	1280	2560	160	240	960	40	20	20
WU95REC	131	141	232	1406	190	563	1383	369	966	<10	<10	49	<10	<10	<10

	LY/1149/91	SD/9/93	LY/2279/95	BR/8/96	SY/5A/97	NL/118/01
WU95REC	71	453	3200	2059	369	966
WU95156KQ	40	60	1120	560	640	1280
WU95196VA	70	240	1920	960	240	640
WU95276NK	40	320	2560	1280	480	960
SY97REC	*	*	*	*	3200	3200

	LY/1149A/91	SD/9B/93	JO/33A/94	LY/2279A/95	LY/2279B/95	NA/933/95	WU/359B/95	BR/8B/96	SY/5A/97	NL/118/01	FU/411/02	NL/22/03	NL/88/03	WE/001/04	WN/67/05	NL/42/06
WU95REC	131	141	232	1406	1633	190	563	1383	369	966	<10	<10	49	<10	<10	<10
WU95156KQ158EK	20	80	98	434	80	113	104	640	1024	1012	165	89	226	<10	<10	<10
WU95121TN156KQ158EK	*	*	20	*	240	<10	80	*	1280	1280	40	40	40	20	<10	<10
WU95124GS156KQ158EK	*	*	<10	*	160	<10	80	*	1280	1120	50	80	80	<10	<10	<10
WU95135KT156KQ158EK	80	80	<10	160	*	480	640	1280	1600	2560	320	320	320	15	30	
WU95142GR156KQ158EK	*	*	<10	160	80	<10	160	*	2240	1600	80	80	160	<10	<10	<10
WU95156KQ158EK196VA	*	*	*	640	150	<40	139	*	1094	784	57	80	32	<10	<10	<10
WU95156KQ158EK276NK	*	*	*	*	139	<10	40	*	784	876	69	89	40	<10	<10	<10
WU95133DN156KQ158EK	160	<10	<40	80	120	*	80	320	1171	2024	183	160	202	*	<40	
WU95133DN135KT156KQ158EK	<10	<10	<10	57	*	80	160	240	1280	2560	237	240	226	<10	<10	<10
WU95156KQ158EK196VA276NK	*	*	*	*	126	<10	37	*	716	716	57	69	40	<10	<10	<10
SY97REC	44	<10	<10	65	*	43	48	95	3200	3200	747	842	1016	105	38	40

	LY/1149A/91	BE/32/92	SD/9B/93	JO/33A/94	LY/2279A/95	NA/933/95	WU/359B/95	BR/8B/96	SY/5A/97	NL/118/01	FU/411/02	NL/22/03	NL/88/03	WE/001/04	WN/67/05	NL/42/06
SY97REC	44	<10	<10	<10	65	43	48	95	3200	3200	747	842	1016	105	38	40
SY97156QK158KE	160	280	320	160	800	400	640	1280	1280	2560	400	320	320	80	<10	<10
SY97133ND156QK158KE	160	320	560	320	1280	640	1280	1280	640	1600	120	240	640	80	40	60
SY97142RG156QK158KE	160	280	320	160	640	320	480	640	320	640	30	160	160	40	<10	<10
SY97196AV156QK158KE	80	160	320	160	640	320	640	640	640	1280	160	160	320	30	<10	<10
WU95REC	131	20	141	232	1406	190	563	1383	369	966	<10	<10	49	<10	<10	<10

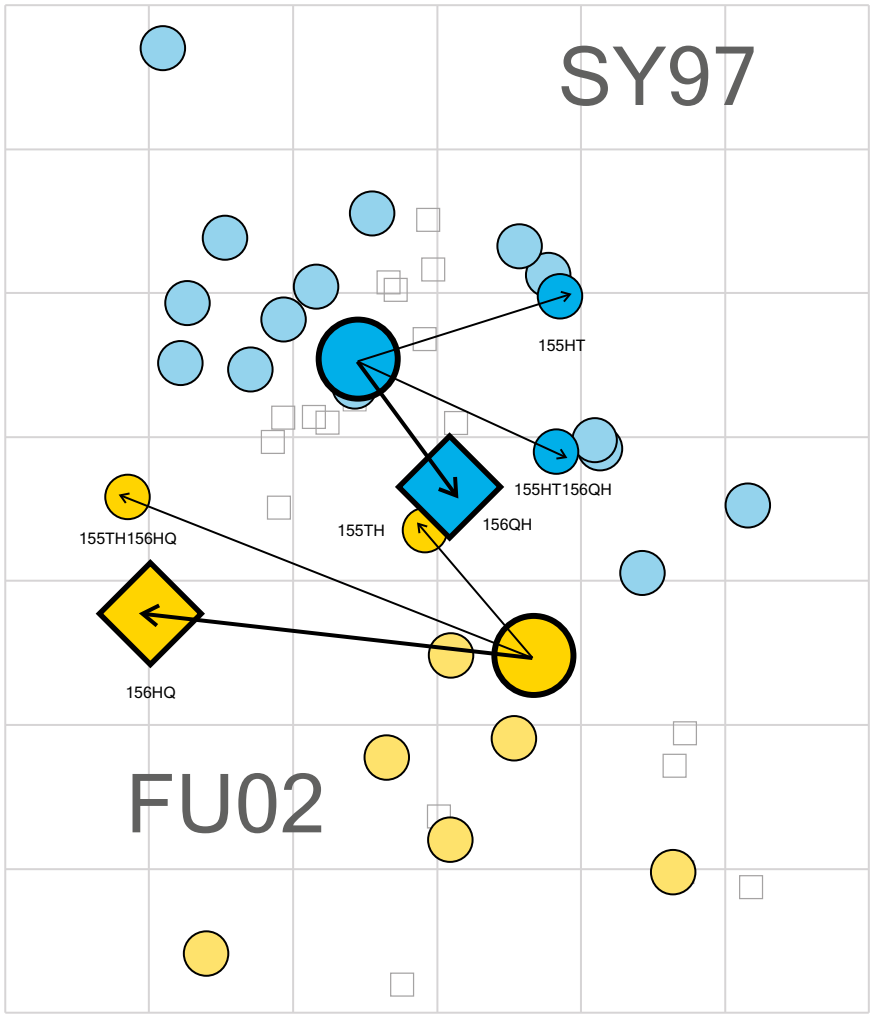


Fig. S4J. SY97 – FU02 cluster transition. Four cluster-difference substitutions occur between the SY97 and FU02 clusters, each of which was constructed and tested as a single mutation. The panel of 16 antisera used in the analysis did not recognize the 155HT, 156QH or the double 155HT156QH mutants. Each mutant was extensively tested to exclude contamination and sequence anomalies. The remaining two mutants, 131AT and 225GD, reacted normally but were SY97-like antigenically (data not shown). Jin *et al*. (48) reported that a combination of 155HT and 156QH, but not the single substitutions, reverted the phenotype from FU02 to SY97-like and concluded that both 155HT and 156QH were required for the SY97 to FU02 transition, but also could not test the single forward mutants. The consensus virus that represents SY97 was replaced because both the consensus virus as well as mutants made from the consensus HA grew to very low HA titers, which may have caused the unexpected results (see materials and methods for details). We rescued and tested the three forward and reverse mutants in the new consensus virus. The phenotypes of the 155HT156QH double and 156QH single mutants were very similar, also for the reverse mutants, and both the induced the cluster transition. The directionality of the 155HT alone was more towards earlier clusters. Thus, 155HT did not contribute to the antigenic effect of 156QH. Hence, we conclude that the cluster-transition substitution for the SY97 to FU02 cluster transition was 156QH.

	LY/1149A/91	BE/32/92	SD/98/93	LY/2279A/95	NA/933/95	WU/359B/95	BR/8B/96	SY/5A/97	NL/118/01	FU/411/02	NL/22/03	NL/88/03	WE/001/04	WN/67/05	NL/42/06
SY97REC	49	<10	<10	80	53	57	113	4280	3200	1140	1109	1522	139	49	40
SY97155HT	299	57	98	150	277	196	367	2677	3387	905	1197	2074	480	277	506
SY97156QH	<10	<10	<10	98	37	57	49	2024	3387	1584	1339	453	80	53	277
SY97155HT156QH	130	10	17	179	150	98	98	2479	3200	2715	3135	669	847	170	1197
FU02REC	<40	<10	<20	28	57	<10	40	1351	1949	2191	2090	421	871	452	916

	LY/1149A/91	LY/2279A/95	NA/933/95	WU/359B/95	BR/8B/96	SY/5A/97	NL/118/01	FU/411/02	NL/22/03	NL/88/03	WE/001/04	WE/1-IVR-139/04	WN/67/05	NL/42/06
FU02REC	<40	28	57	<10	40	1351	1949	2191	2090	421	871	471	452	916
FU02155TH	<10	15	*	<40	60	1893	2933	847	1431	847	280	160	15	70
FU02156HQ	<10	<10	<10	<10	30	1466	2217	599	733	506	140	80	80	*
FU02155TH156HQ	<10	10	<20	<10	50	1920	2862	320	518	1197	70	80	42	49
SY97REC	49	80	53	57	113	4280	3200	1140	1109	1522	139	*	49	40

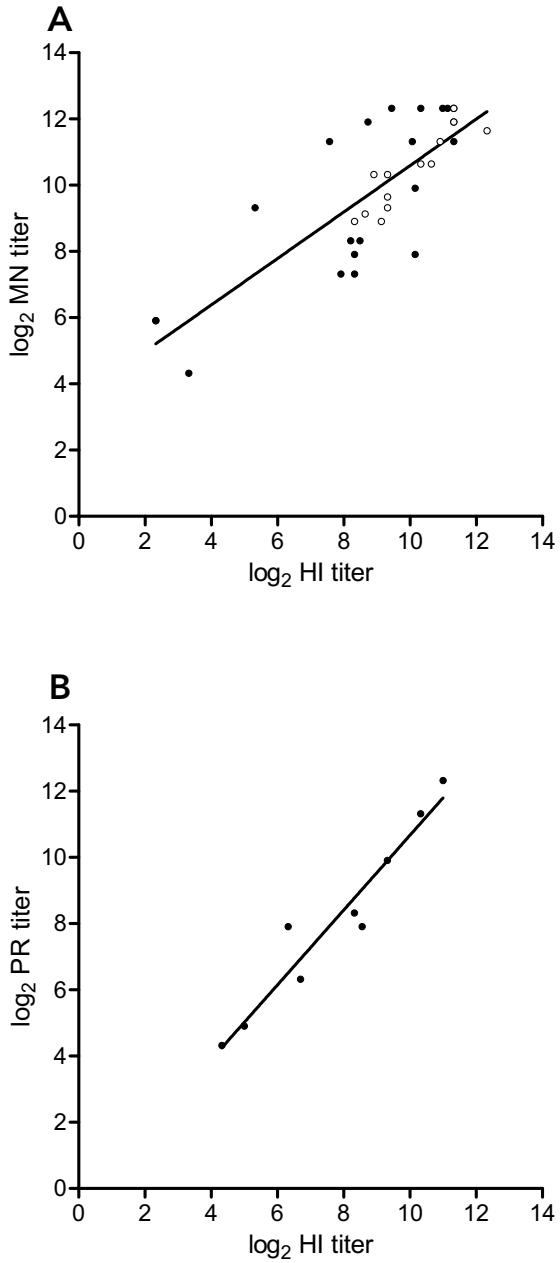


Fig. S5. Correlation between hemagglutination inhibition and virus neutralization assays. A) Consensus viruses and their associated forward and reverse cluster-transition mutants (black symbols), and non-cluster-transition mutants in sites C, D and E (open symbols), were tested in both the HI and micro neutralization (MN) assays. Titers are expressed as the $\log_2$ of the resulting value found in each assay. $R^2 = 0.65$, 95% CI = 0.64 – 0.90, P-value (two-tailed) < 0.0001. B) Consensus viruses used in the analysis of cluster-defining substitutions were tested in both the HI and plaque reduction (PR) assays. Titers are expressed as the $\log_2$ of the resulting value found in each assay. $R^2 = 0.96$, 95% CI = 0.83 – 0.99, P value (two-tailed) < 0.0001.

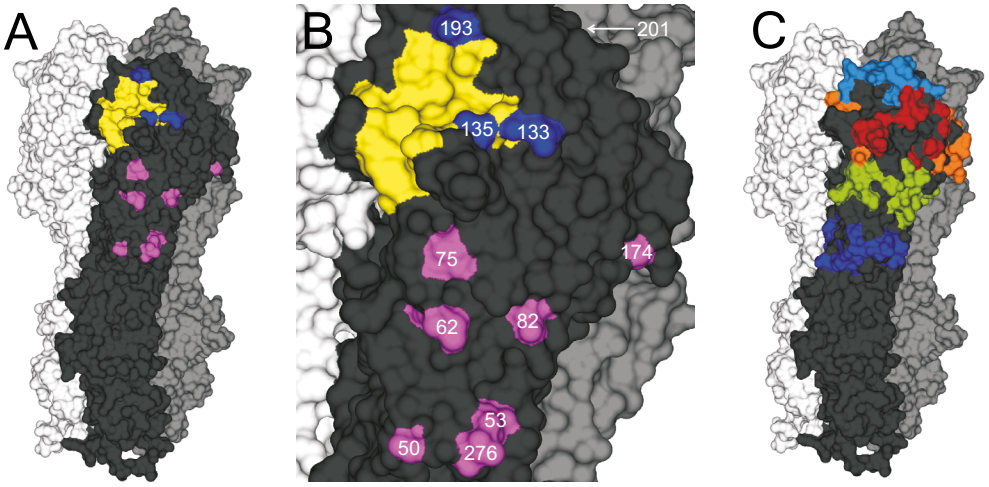
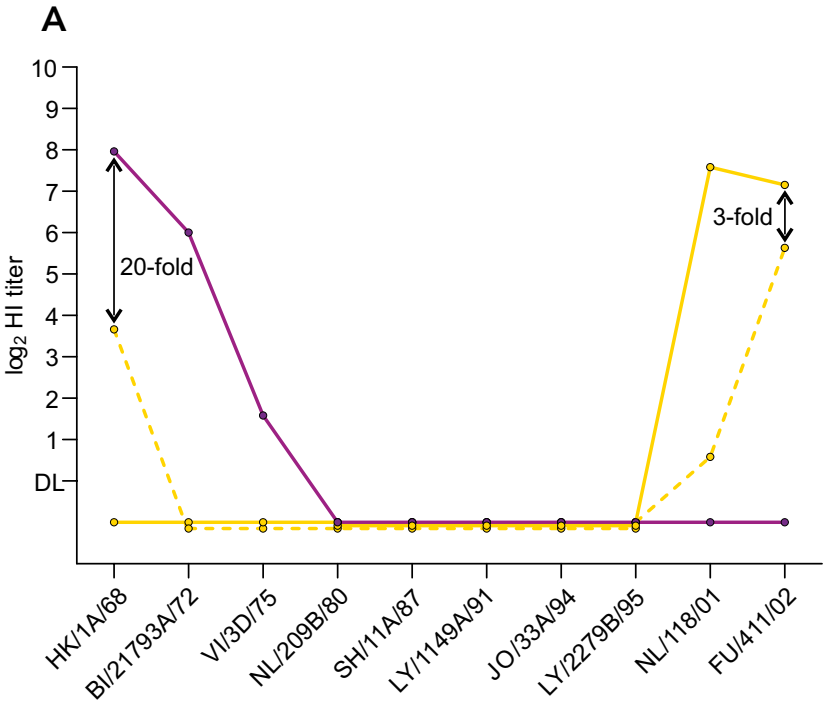


Fig. S6. Structure of an A/Aichi/2/1968 HA trimer. Accessory substitutions and cluster-difference amino acid substitutions tested in both hemagglutination inhibition and virus neutralization assays indicated on an A/Aichi/2/1968 HA trimer. The three monomers are shown in black, white and grey, the RBS in yellow. For reference, panel C shows the positions of antigenic sites A-E. A) Positions of accessory substitutions (blue) and cluster-difference amino acid substitutions tested in virus neutralization assays (magenta). Mutants tested were: EN72 53ND; site C, EN72 201RK; site D, VI75 50KR; site C, VI75 82EK; site E, VI75 174SF; site D, WU95 62KE; site E, WU95 276NK; site C, SY97 50RG; site C, and SY97 75HQ; site E. B) Zoomed in on the area around the RBS. Position 193 is a cluster-transition substitution in the VI75 to TX77 antigenic cluster transition, but is an accessory substitution in the SI87 to BE89 cluster transition (Fig. S10). Position 201 is located on the trimer interface behind the dark monomer in this figure and is therefore not visible. C) Position of antigenic sites A-E (32, 33). Site A; red, site B; light blue, site C; blue, site D; orange, and site E; green.

Fig. S7. Effect of exchanging amino acids on key positions in the HK68 and FU02 consensus viruses. Solid purple and yellow lines show the $\log_2$ HI titers for the HK68 and FU02 consensus viruses to a set of antisera that are representative of the antigenic clusters. DL indicates the detection limit of the HI assay used. Vertical arrows denote the fold difference from homologous titers with HK68 and FU02 consensus viruses. A) The dashed yellow line shows the $\log_2$ HI titers for HK68+5, which is genetically fully HK68 consensus except for the five substitutions on the seven key positions that differ between the HK68 and FU02 viruses. B) The dashed purple line indicates the $\log_2$ HI titers for FU02-5 (FU02 minus 5) that has the five amino acids typical of the HK68 consensus virus on the seven key positions, but is otherwise genetically fully FU02 consensus. The amino acid substitutions in HK68+5 and FU02-5 compared to the HK68 and FU02 consensus viruses, as well as the HI titers, are listed in Table S4. The antisera used in this experiment were raised to viruses from the following antigenic clusters: HK/1A/68; HK68, BI/21793/72; EN72, VI/3D/75; VI75, NL/209B/80; BK79, SH/11A/87; SI87, LY/1149/91; BE89, JO/33A/94; BE92, LY/2279/95; WU95; NL/118/01; SY97, FU/411/02; FU02. The data in this figure are an extreme test of the findings for 'natural' cluster transitions (cluster transitions between neighboring clusters) by incorporating only the cluster-transition substitutions for nine consecutive cluster transitions from HK68 to FU02 into a single virus. Interestingly, the forward mutant virus in panel A retains some binding to the HK68 antiserum (HK/1A/68), gains titers to the latest antisera (NL/118/01 and FU/441/02) and has undetectable titers to the other antisera in the panel—other than residual titers to the HK/1A/68 antiserum, this highly artificial virus is essentially FU02-like.


A
10
9
8
7
6
5
4
3
2
1
DL
log2 HI titer
20-fold
3-fold
HK/1A/68
BI/21793A/72
VI/3D/75
NL/209B/80
SH/11A/87
LY/1149A/91
JO/33A/94
LY/2279B/95
NL/118/01
FU/411/02

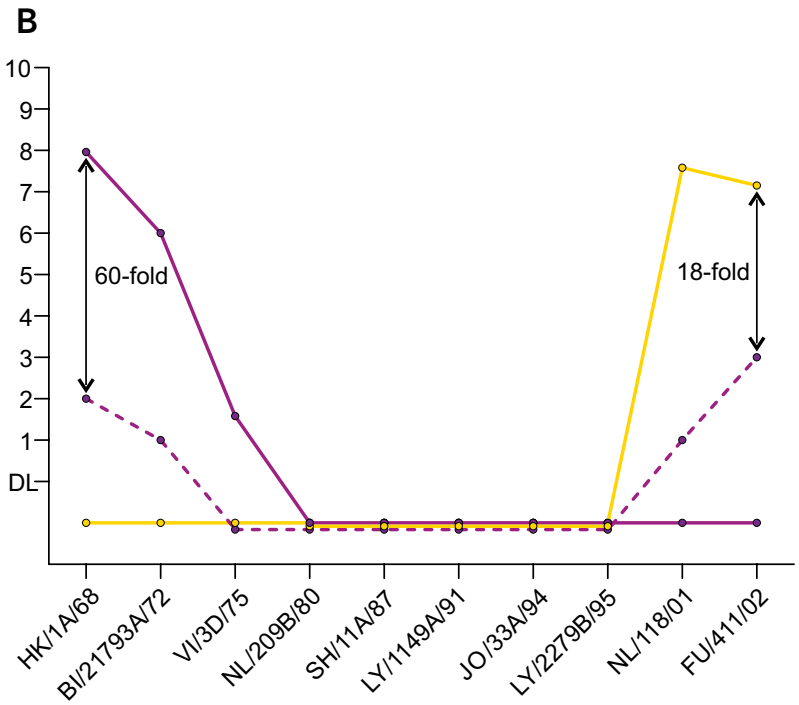
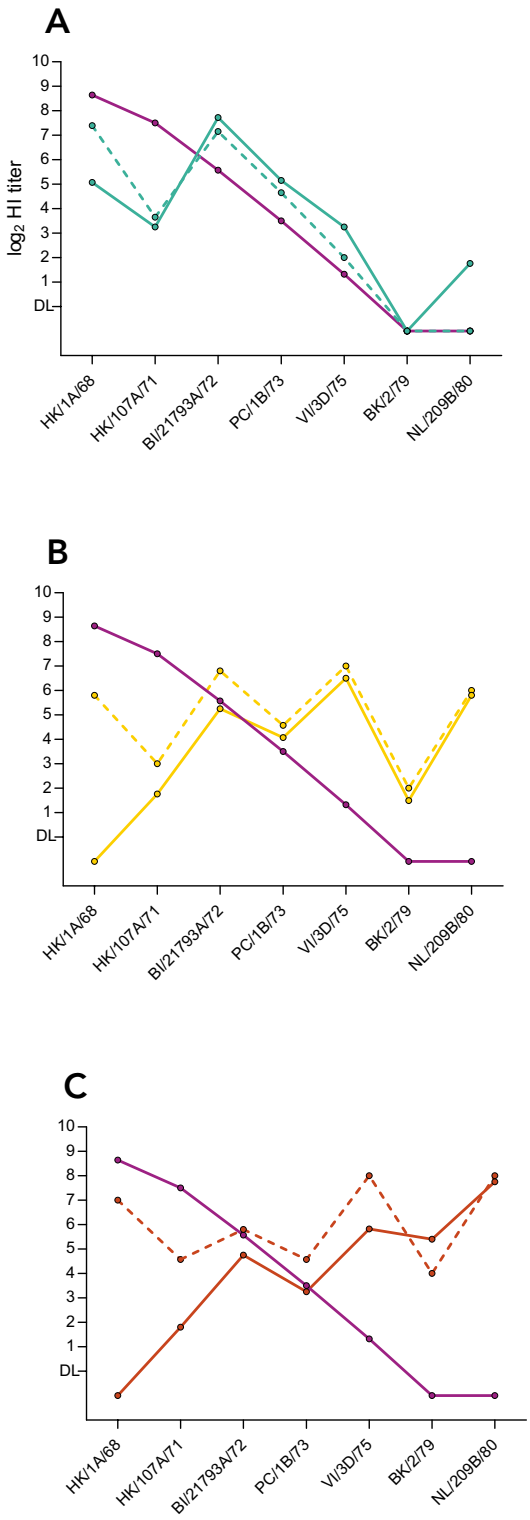
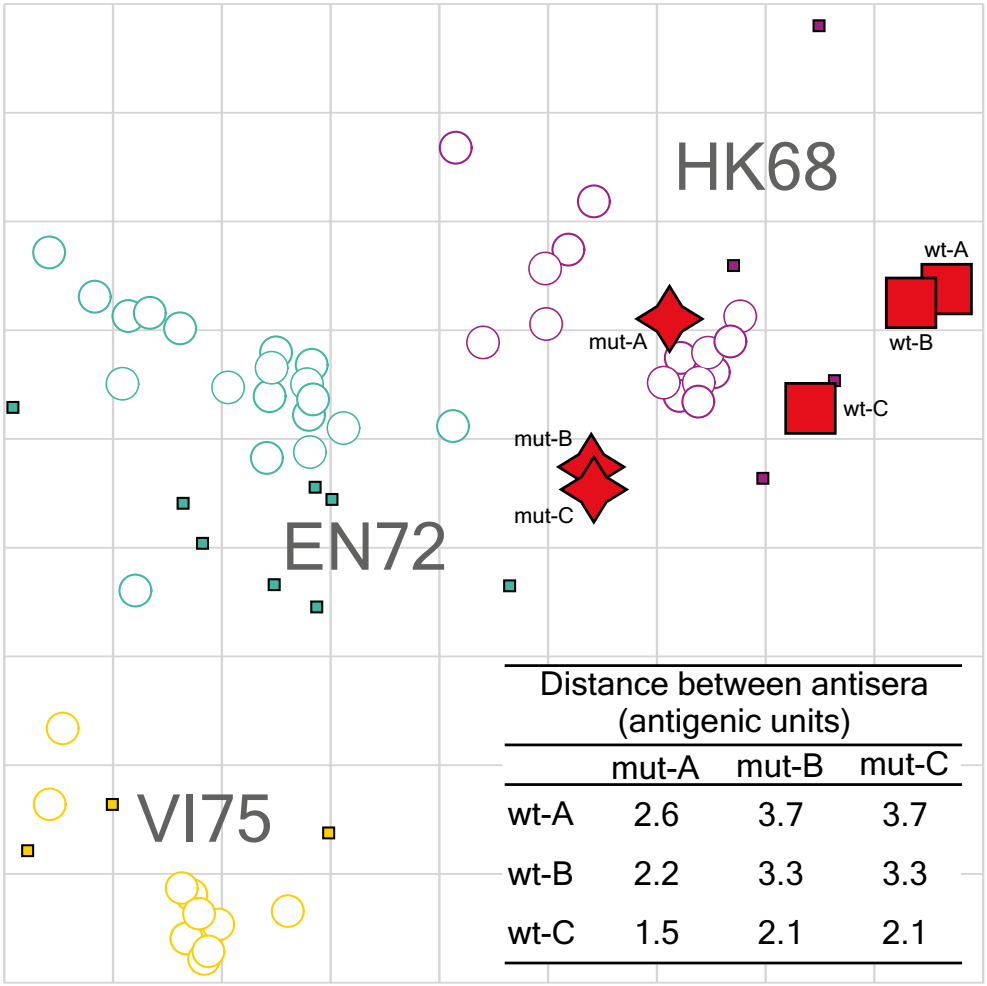

B
10
9
8
7
6
5
4
3
2
1
DL
60-fold
18-fold
HK/1A/68
BI/21793A/72
VI/3D/75
NL/209B/80
SH/11A/87
LY/1149A/91
JO/33A/94
LY/2279B/95
NL/118/01
FU/411/02

Fig. S8. Incremental addition of cluster-transition substitutions. A) Solid purple and cyan lines indicate $\log_2$ HI titers for HK68 and EN72 consensus viruses, respectively. The dashed cyan line shows $\log_2$ HI titers for HK68 155TY, which has the substitution responsible for the major antigenic change between the HK68 to EN72 cluster but not the remaining four cluster-difference substitutions. In all panels, DL indicates the detection limit of the HI assay used. B) Solid purple and yellow lines indicate $\log_2$ HI titers for the HK68 and VI75 consensus viruses, respectively. The dashed yellow line shows the $\log_2$ HI titers for HK68 155TY189QK—the cluster-transition substitutions responsible for the HK68 to EN72 and EN72 to VI75 cluster transitions, but not the remaining 14 cluster-difference substitutions. C) Solid purple and brown lines indicate $\log_2$ HI titers for HK68 and TX77 consensus viruses, respectively. The dashed brown line shows $\log_2$ HI titers for HK68 155TY189QK158GE193DN that has the cluster-transition substitutions for the HK68 to EN72, EN72 to VI75 and VI75 to TX77 cluster transitions, but not the remaining 21 cluster-difference substitutions. We constructed the HK68 consensus HA with all cluster-transition substitutions between HK68 and BK79, but could not rescue this mutant. The antisera used in this experiment were raised to viruses from the following antigenic clusters: HK/1A/68; HK68, HK/107A/71; HK68, BI/21793/72; EN72, PC/1B/73; EN72, VI/3D/75; VI75, BK/2/79; BK79, NL/209B/80; BK79. Panels B and C test the extremities of the findings for 'natural' cluster transitions (cluster transitions between neighboring clusters) by incorporating the cluster-transition substitutions for multiple (two and three respectively) consecutive cluster transitions into a single virus. Interestingly, the mutant viruses retain some binding to the earliest HK68 antiserum (HK/1A/68), but have mostly similar titers as the consensus viruses to the other HK68 antiserum (HK/107A/71) and to the other antisera in the panel.


A
log2 HI titer
10
9
8
7
6
5
4
3
2
1
DL
HK/1A/68
HK/107A/71
BI/21793A/72
PC/1B/73
VI/3D/75
BK/2/79
NL/209B/80
B
C



Distance between antisera (antigenic units)

	mut-A	mut-B	mut-C
wt-A	2.6	3.7	3.7
wt-B	2.2	3.3	3.3
wt-C	1.5	2.1	2.1

Fig. S9. Antigenic map of antisera raised to the wildtype or a mutant HK68-like virus. The map shows the positions of antisera raised to the HK68 consensus virus (red squares) and to HK68 155TY (red stars). The position of the HK68, EN72 and VI75-like strains are indicated by purple, cyan and yellow outlines, respectively. The inset table shows the distances between the antiserum positions. The average distance between the wildtype and mutant antisera is 2.7 antigenic units (a 6.5-fold dilution difference). This distance is similar to the distance between the HK68 consensus virus and the HK68 155TY mutant, and the distance between the HK68 and EN72 consensus viruses (2.6 and 3.1 antigenic units. Fig. 1 and Table S3).

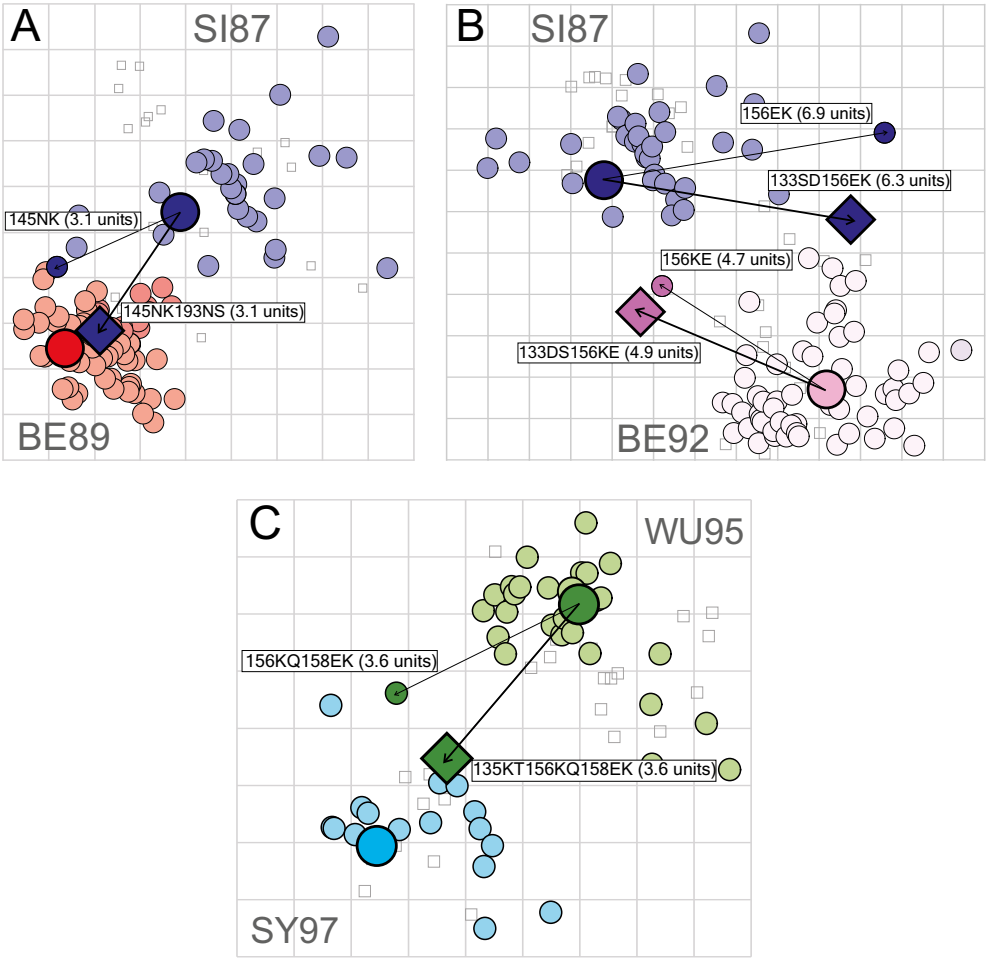
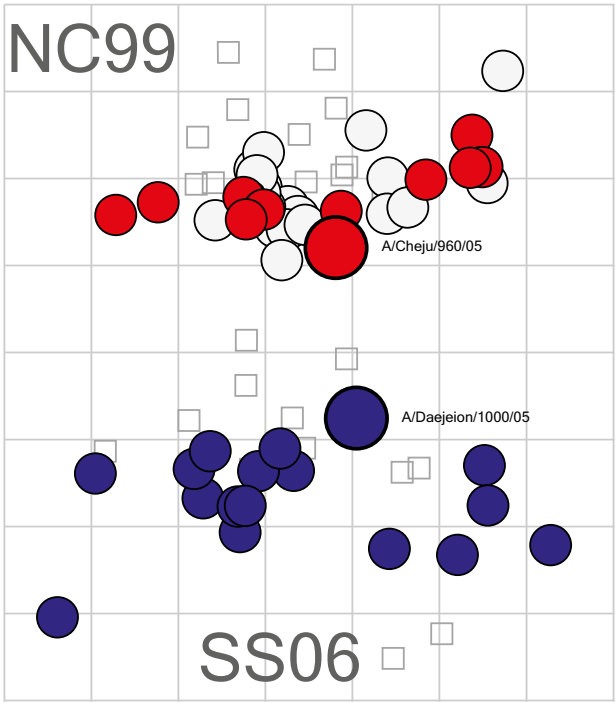


Fig. S10. Antigenic effect of accessory substitutions. For three cluster transitions an accessory amino acid substitution was required to change the antigenic phenotype to that characteristic of the subsequent cluster. The larger circles represent the consensus viruses for the two clusters shown in the map, small faint colored circles are wildtype strains, small filled circles indicate the cluster-transition mutant without the accessory substitution, diamonds indicate the cluster-transition mutant with the accessory substitution. The distances in antigenic units between the consensus virus and mutants are indicated in the figures. A) The accessory substitution for the SI87 – BE89 transition was 193NS. B) The accessory substitution for the SI87 – BE92 transition was 133SD. C) The accessory substitution for the WU95 – SY97 transition was 135KT. Addition of the accessory substitutions did not result in increased antigenic distance from the consensus virus in which the substitution was introduced, but corrected for directionality towards the subsequent antigenic phenotype in the antigenic maps. Because no additional antigenic distance is acquired by addition of an accessory substitution, we did not consider these to be essential molecular determinants of the major antigenic change between clusters.



	A/New Caledonia/020/99	A/Netherlands/300/00	A/Netherlands/352/02	A/Solomon Islands/003/06	A/Netherlands/361/06	A/Netherlands/345/07	A/Netherlands/165/08
A/New Caledonia/020/99	640	960	640	80	800	80	80
A/Cheju/960/2005	480	960	640	360	1120	640	960
A/Daejeion/1000/2005	160	160	320	960	1280	800	1280
A/Solomon Islands/003/06	80	100	60	400	1280	320	640

Fig. S11. Antigenic change of epidemic A/H1N1 viruses. NC99; A/New Caledonia/20/99, SS06; A/Solomon Islands/3/06. The strains are colored according to the amino acid at position 140 (H1 numbering); red circles have a Lysine (K) at position 140, whereas blue circles have a Glutamic acid (E) at this position. There are no other substitutions that are consistently different between viruses of the NC99 and SS06 phenotypes. No sequence information was available for the grey filled circles. The larger circles represent recombinant viruses that have the HA1 amino acid sequence from A/Cheju/960/05 and A/Daejeion/1000/05. The HA1 amino acid sequence of these viruses differs only by 140KE. Therefore, 140KE alone is sufficient to induce the antigenic transition from NC99 to SS06-like.

	B/Malaysia/2506/04	B/Darwin/40/06	B/Singapore/616/08	B/Philippines/6363/09	B/Cambodia/30/11	B/Brisbane/33/08	B/Brisbane/60A/08	B/Brisbane/60B/08	B/Hong Kong/90/08	B/Sydney/508/10	B/South Australia/11/12
B/Malaysia/20/04	40	<20	40	640	160	20	<20	<20	40	<20	<20
B/Malaysia/283/05	20	<20	40	640	80	<20	<20	<20	20	<20	<20
B/Macau/388/05	20	<20	40	640	160	<20	<20	<20	40	<20	<20
B/Perth/1/06	20	20	40	320	80	<20	<20	<20	40	<20	<20
B/Darwin/2/06	20	20	80	640	160	20	<20	20	40	<20	20
B/Victoria/303/06	<20	320	40	320	40	160	80	320	160	80	160
B/Brisbane/6/07	<20	640	40	160	40	160	160	320	320	160	160

Fig. S12. Antigenic change of influenza B/Victoria-lineage viruses (7). ML04; B/Malaysia/2506/04, BR08; B/Brisbane/60/08. The strains are colored according to the amino acid at position 165 (B/Victoria numbering): Blue circles have 165 Asparagine (N), orange circles have a Lysine (K) at this position. The HA1 amino acid sequences of the large blue and orange circles differ only by the 165NK substitution. Thus, change of an N to a K at position 165 was responsible for the antigenic change between the ML04 and BR08 antigenic phenotypes.

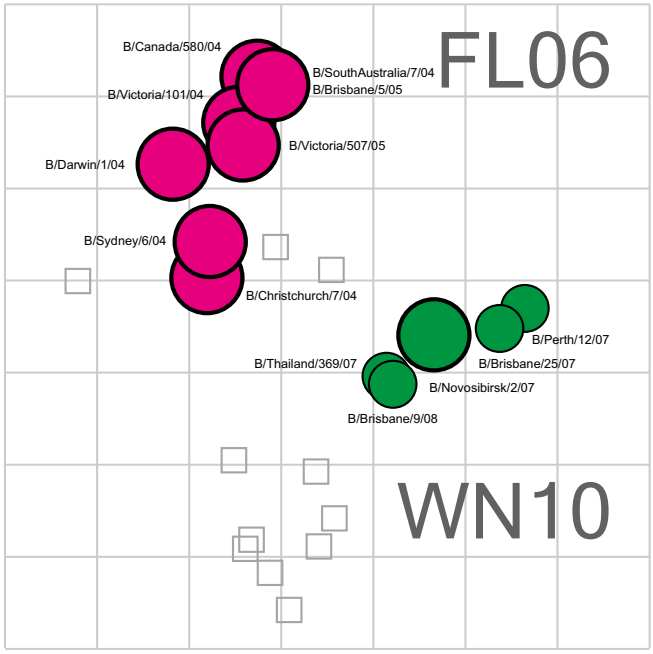


Fig. S13. Antigenic change of influenza B/Yamagata-lineage viruses (7). FL06; B/Florida/4/06, WN10; B/Wisconsin/1/10. The strains are colored according to the amino acid at position 166 (B/Victoria numbering): strains that have an Asparagine (N) at position 166 are shown in magenta, those with 166 Tyrosine (Y) in green. The HA1 amino acid sequences of the large magenta and green circles differ by only 166NY. We conclude that the 166NY substitution was solely responsible for the antigenic difference between the FL06 and WN10 antigenic phenotypes.

	B/Florida/4/06	B/Brisbane/3A/07	B/Bangladesh/3333/07	B/Indiana/1/08	B/South Australia/5/08	B/Brisbane/9/08	B/Hubei/Wujiagang/158A/09	B/Wisconsin/1/10	B/Brisbane/2/12	B/South Australia/6/12	B/Malaysia/412/12
B/Canada/580/04	320	320	40	80	40	40	40	20	40	40	160
B/Darwin/1/04	320	320	80	160	80	80	80	40	40	80	320
B/South Australia/7/04	320	320	40	80	40	40	80	20	40	40	160
B/Sydney/6/04	640	640	160	160	120	160	160	80	80	160	480
B/Victoria/101/04	640	320	40	160	40	80	80	30	40	40	240
B/Victoria/507/05	640	640	80	160	40	40	160	40	40	40	320
B/Brisbane/5/05	320	320	40	80	40	40	80	20	40	40	160
B/Christchurch/7/04	640	640	160	160	160	160	320	120	160	160	640
B/Novosibirsk/2/07	320	640	160	160	160	320	320	160	160	160	80
B/Thailand/369/07	320	640	160	240	320	640	320	320	240	320	160
B/Perth/12/07	160	320	80	160	160	120	160	60	80	80	40
B/Brisbane/9/08	320	640	320	320	320	480	640	160	240	240	80
B/Brisbane/25/07	160	480	160	160	160	160	320	80	80	80	40

Table S1. Ferret antisera used in this study. Each antiserum was raised in ferrets as described in Supplementary Methods. When antisera were raised to the same strain multiple times the repeat antisera were labeled A, B, C, etc., additional to the virus strain name to which the serum was raised (e.g. A/Hong Kong/1/68, A/Hong Kong/1A/68, A/Hong Kong/1B/68). Data obtained using different antisera raised to the same strain were treated as individual antisera, i.e. data from A/Hong Kong/1/68 and A/Hong Kong/1A/68 were not averaged. [a]Influenza B virus Yamagata-lineage (7). [b]Influenza B virus Victoria-lineage (7).

Antigenic cluster	Subtype or lineage	Antisera		
HK68	A/H3N2	A/Hong Kong/1/68	A/Hong Kong/107/71	
EN72	A/H3N2	A/Bilthoven/21793/72	A/Port Chalmers/1/73	A/Scotland/840/74
VI75	A/H3N2	A/Victoria/3/75		
TX77	A/H3N2	-		
BK79	A/H3N2	A/Netherlands/209/80	A/Philippines/2/82	A/Netherlands/233/82
		A/Netherlands/241/82	A/Stockholm/10/85	
SI87	A/H3N2	A/Sichuan/2/87	A/Shanghai/11/87	A/Hong Kong/34/90
BE89	A/H3N2	A/Beijing/353/89	A/Victoria/2/90	A/Lyon/1149/91
		A/Paris/548/92	A/Netherlands/823/92	
BE92	A/H3N2	A/Beijing/32/92	A/Shangdong/9/93	A/Johannesburg/33/94
WU95	A/H3N2	A/Wuhan/359/95	A/Nanchang/933/95	A/Lyon/2279/95
		A/Brisbane/8/96		
SY97	A/H3N2	A/Sydney/5/97	A/Netherlands/118/01	A/Netherlands/88/03
FU02	A/H3N2	A/Fujian/411/02	A/Netherlands/22/03	A/Wellington/1-IVR-139/04
		A/Wisconsin/67/05	A/Netherlands/42/06	
NC99	A/H1N1	A/New Caledonia/20/99	A/Netherlands/271/99	A/Netherlands/300/00
		A/Netherlands/352/02	A/Netherlands/128/04	A/Netherlands/239/06
		A/Netherlands/364/06	A/St.Petersburg/10/07	
SS06	A/H1N1	A/Solomon Islands/3/06	A/Fukushima/97/06	A/Netherlands/361/06
		A/Brisbane/59/07	A/Netherlands/345/07	A/Netherlands/165/08
FL06	B Yam[a]	B/Florida/4/06	B/Brisbane/3/07	B/South Australia/6/12
		B/Malaysia/412/12		
WN10	B Yam[a]	B/Bangladesh/3333/07	B/Indiana/1/08	B/South Australia/5/08
		B/Brisbane/9/08	B/Hubei/Wujiagang/158/09	B/Wisconsin/1/10
		B/Brisbane/2/12		
ML04	B Vic[b]	B/Malaysia/2506/04	B/Darwin/40/06	B/Singapore/616/08
		B/Philippines/6363/09	B/Cambodia/30/11	
BR08	B Vic[b]	B/Brisbane/33/08	B/Brisbane/60/08	B/Hong Kong/90/08
		B/Sydney/508/10	B/South Australia/11/12	

Table S2. Antigenic distances by cluster-difference substitutions. Smith *et al.* previously identified 67 amino acid substitutions associated with the antigenic evolution of A/H3N2 viruses (31), 54 of which were in the 109 – 301 amino acid region of HA. These 54 substitutions were introduced in the appropriate consensus HA by site directed mutagenesis and the antigenic properties of recombinant viruses were tested in HI assays. Note that from phylogenetic analysis the most likely ancestor of the BE92 cluster was a SI87-like strain rather than a BE89-like strain (31), and cluster-difference substitutions as defined by Smith *et al.* (31) were redefined comparing the SI87 and BE92-clusters. Substitutions in bold indicate cluster-transition substitutions. Antigenic change of substitutions in italic was tested in both hemagglutination inhibition and virus neutralization assays (Figs. S5 and S6). Antigenic distance is expressed as antigenic units, and shows the distance between the consensus virus in which the mutations were introduced and the mutant. "N.A" indicates substitutions that were assigned as cluster-difference substitutions that are outside the 109 - 301 amino acid region of HA, and were therefore not tested. "N.D" indicates the mutants we could not grow.

	Cluster-difference substitutions											
Cluster transition	Antigenic Site A	Antigenic Distance	Antigenic Site B	Antigenic Distance	Antigenic Site C	Antigenic Distance	Antigenic Site D	Antigenic Distance	Antigenic Site E	Antigenic Distance	Other	Antigenic Distance
HK68-EN72	122TN	0.9	155TY	2.6			207RK	0.9				
	144GD	0.9	188ND	0.6								
EN72-VI75	137NS	0.5	164LQ	0.5	53ND	N.A.	174FS	0.5				
	145SN	0.3	189QK	4.1	278IS	0.5	201RK	0.3				
			193SD	1.2			213IV	0.5				
							217IV	0.3				
							230IV	0.3				
VI75-TX77	137SY	2.2	158GE	0.5	50KR	N.A.	174SF	0.9	82EK	N.A.		
			164LQ	1.0	53DN	N.A.	201KR	1.4	260MI	1.2		
			193DN	0.9			213VI	1.3				
							230VI	0.9				
TX77-BK79	133NS	0.2	156KE	1.2	53ND	N.A.	172DG	0.3	62IK	N.A.		
	143PS	0.3	160TK	0.7	54NS	N.A.	217VI	0.4	82EK	N.A.		
	146GS	0.5	197QR	0.7			244VL	0.5				
BK79-SI87	124GD	0.9	155YH	1.4								
			159SY	4.1								
			189KR	3.3								

Table S2. (continued)

	Cluster-difference substitutions											
Cluster transition	Antigenic Site A	Antigenic Distance	Antigenic Site B	Antigenic Distance	Antigenic Site C	Antigenic Distance	Antigenic Site D	Antigenic Distance	Antigenic Site E	Antigenic Distance	Other	Antigenic Distance
SI87-BE89	145NK	3.1	193NS	N.D.								
SI87-BE92	133SD	0.7	156EK	6.9								
			190ED	0.4								
			193NS	0.5								
BE92-WU95	145NK	2.2										
WU95-SY97			156KQ	2.3	276NK	1.5			62KE	N.A.		
			158EK	N.D.								
			196VA	0.8								
SY97-FU02	131AT	N.D.	155TH	1.5	50RG	N.A.			75HQ	N.A.	25LI	N.A.
			156HQ	1.1					83EK	N.A.	202VI	N.D.
											222WR	N.D.
											225GD	N.D.

Table S3. Cluster-transition substitutions and distance between mutant and consensus virus into which the substitutions were introduced. Cluster-transition amino acid substitutions as defined in the main text. Substitutions between brackets are accessory substitutions. The addition of these substitutions corrected for directionality towards the consensus virus of the subsequent cluster only and did not significantly add to the antigenic distance from the consensus virus in which the substitutions were introduced (Fig. S10).

Cluster transition	Cluster-transition substitutions				Distance between consensus viruses (antigenic units)	Distance mutant from consensus virus (antigenic units)
	Antigenic Site A	Antigenic Site B				
HK68-EN72		155TY			3.1	2.6
EN72-VI75		189QK			4.7	4.1
VI75-TX77		158GE	193DN		1.3	1.6
TX77-BK79		156KE			2.1	1.2
BK79-SI87		155YH	159SY	189KR	6.9	5.8
SI87-BE89	145NK	(193NS)			3.9	3.1
SI87-BE92	(133SD)	156EK			7.3	6.3
BE92-WU95	145NK				3.3	2.2
WU95-SY97	(135KT)	156KQ	158EK		5.5	3.6
SY97-FU02			156QH		2.4	1.1

Table S4. Effect of exchanging amino acids on key positions in the HK68 and FU02 consensus viruses. Table S4A shows the positions responsible for the major antigenic change between the HK68 and FU02 clusters. The HK68+5 virus is genetically fully HK68-like and, vice versa, the FU02-5 (FU02 minus 5) is genetically fully FU02 like, except for the five introduced cluster-transition substitutions. The HI titers in Table S4B are a measure of cross-reactivity of serum antibodies with the tested viruses, where a high value indicates good reactivity. When the highest concentration of antiserum did not inhibit hemagglutination, the titer is indicated as a threshold titer (e.g. <10). The viruses were tested to a set of antisera that are representative of the clusters shown in the antigenic map in Fig. 1. Titers shown are the geometric mean value of multiple independent experiments.

Table S4A.

	HA position						
	145	155	156	158	159	189	193
HK68 consensus	S	T	K	G	S	Q	S
HK68+5	K	.	H	K	Y	S	.
FU02-5	S	.	K	G	S	Q	.
FU02 consensus	K	.	H	K	Y	S	.

Table S4B.

	HI titer from antisera to									
	HK68	EN72	VI75	BK79	SI87	BE89	BE92	WU95	SY97	FU02
	HK/1A/68	BI/21793A/72	VI/3D/75	NL/209B/80	SH/11A/87	LY/1149A/91	JO/33A/94	LY/2279B/95	NL/118/01	FU/411/02
HK68 consensus	2506	640	30	<10	<20	<10	<10	<10	<10	<10
HK68+5	127	<10	<10	<10	<20	<10	<10	<10	15	497
FU02+5	40	20	<10	<10	<20	<10	<10	<10	20	80
FU02 consensus	<10	<10	<10	<10	<20	<10	<10	<10	1920	1421

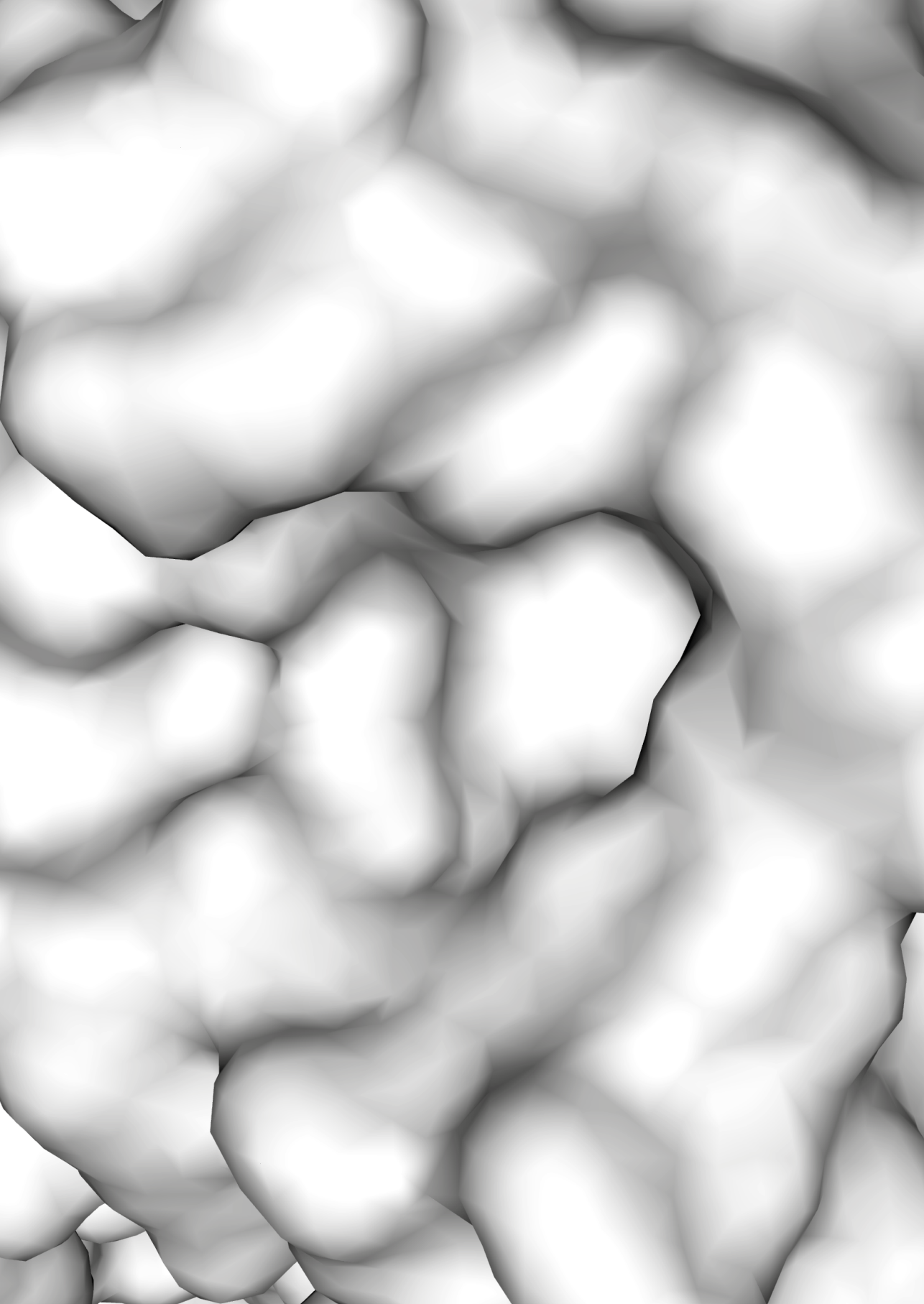
Table S5. Charge changes induced by cluster-transition and cluster-difference substitutions. Cluster-transition and cluster-difference amino acid substitutions as defined in the main text and Smith *et al.* (31). Substitutions between brackets are accessory substitutions (Fig. S10). Neutr, pos and neg: uncharged, positively or negatively charged at physiological pH, respectively. [v]Substantial change in amino acid volume after substitution. [a]Influenza A/H1N1 virus. [b]Influenza B virus Yamagata-lineage (7). [c]Influenza B virus Victoria-lineage (7). N.D. not determined.

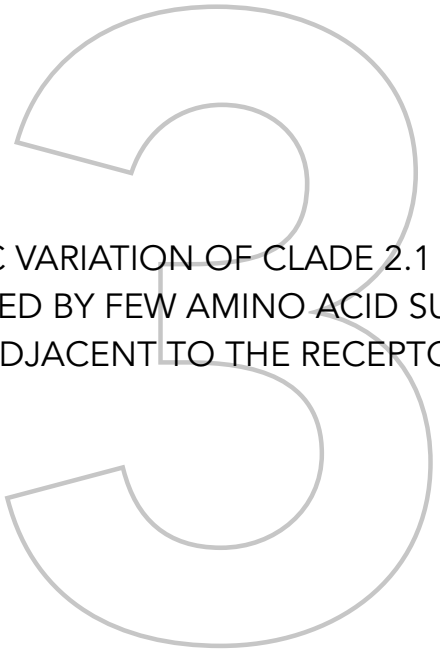
Cluster transition	Cluster transition-substitution	Change by cluster-transition substitution	Total charge change cluster-transition substitution(s)	Total charge change cluster-difference substitution(s)
HK68-EN72	155TY	-[v]	0	-2
EN72-VI75	189QK	neutr – pos	+1	-1
VI75-TX77	158GE	neutr – neg[v]	0	+3
	193DN	neg – neutr		
TX77-BK79	156KE	pos – neg	-2	+3
BK79-SI87	155YH	neutr – pos	+1	0
	159SY	-[v]		
	189KR	-		
SI87-BE89	145NK	neutr - pos	+1	+1
	(193NS)	-		
SI87-BE92	156EK	neg - pos	+2	+1
	(133SD)	neutr – neg		
BE92-WU95	145NK	neutr – pos	+1	+1
WU95-SY97	156KQ	pos – neutr	+1	0
	158EK	neg – pos		
	(135KT)	pos – neutr		
SY97-FU02	156QH	neutr – pos	+1	0
NC99-SS06[a]	140KE	pos – neg	-2	N.D.
FL06-WN10[b]	166NY	-[v]	0	N.D.
ML04-BR08[c]	165NK	neutr – pos	+1	N.D.

Table S6. Influenza A/H1N1 and B virus sequence information. We gratefully acknowledge the authors, originating and submitting laboratories of the sequences from GISAID's EpiFlu™ Database on which this research is based. The list is detailed below. [1]Centers for Disease Control and Prevention, Atlanta, United States of America. [2]Queensland Health Forensic and Scientific Services, Brisbane, Australia. [3]WHO Collaborating Centre for Reference and Research on Influenza, Melbourne, Australia. [4]Pathwest QE II Medical Centre, Perth, Australia. [5]Institute of Environmental Science & Research, Auckland, New Zealand. [6]WHO National Influenza Centre National Institute of Medical Research, London, United Kingdom.

Segment ID	Segment	Country	Collection date	Isolate name	Originating laboratory	Submitting laboratory	Authors
EPI158135	HA	Singapore	2005-09-20	A/Singapore/92/2005		CDC[1]	Smith, Catherine
EPI157517	HA	Republic of Korea	2005-12-23	A/Cheju/960/2005		CDC	Smith, Catherine
EPI157555	HA	Republic of Korea	2005-12-21	A/Daejeion/1000/2005		CDC	Smith, Catherine
EPI117575	HA	Malaysia	2004	B/Malaysia/20/2004			
EPI117523	HA	Malaysia	2005-03-07	B/Malaysia/283/2005			
EPI123913	HA	Australia	2006	B/Perth/1/2006			
EPI117567	HA	China	2005-04-13	B/Macau/388/2005			
EPI123923	HA	Australia	2006-06-07	B/Darwin/2/2006			
EPI123907	HA	Australia	2006-06-06	B/Victoria/303/2006			
EPI154535	HA	Australia	2007-06-26	B/Brisbane/6/2007	QHFSS[2]	WHO CC[3]	Komadina, Naomi
EPI11756	HA	Australia	2005-06-09	B/Brisbane/5/2005			
EPI156897	HA	Canada	2004-06-09	B/Canada/580/2004		CDC	Smith, Catherine
EPI117622	HA	New Zealand	2004-07-20	B/Christchurch/7/2004			
EPI117491	HA	Australia	2004-09-10	B/South Australia/7/2004			
EPI117616	HA	Australia	2004	B/Victoria/101/2004			
EPI117604	HA	Australia	2004	B/Darwin/1/2004			
EPI117539	HA	Australia	2005-06-03	B/Victoria/507/2005			
EPI156071	HA	Thailand	2007-03-06	B/Thailand/369/2007	NIMR[6]	WHO CC	Komadina, Naomi
EPI16171	HA	Australia	2007-06-18	B/Perth/12/2007	Pathwest	WHO CC	Komadina, Naomi
EPI163088	HA	Australia	2007-10-02	B/Brisbane/25/2007	QHFSS	WHO CC	Komadina, Naomi
EPI155552	HA	Russian Federation	2007-02-19	B/Novosibirsk/02/2007		CDC	Smith, Catherine

All submitters of data may be contacted directly via the GISAID website www.gisaid.org





ANTIGENIC VARIATION OF CLADE 2.1 H5N1 VIRUS IS DETERMINED BY FEW AMINO ACID SUBSTITUTIONS IMMEDIATELY ADJACENT TO THE RECEPTOR BINDING SITE

ABSTRACT

Highly pathogenic avian influenza (HPAI) viruses of the H5N1 subtype are genetically highly variable and have diversified into multiple phylogenetic clades over the past decade. Antigenic drift is a well-studied phenomenon for seasonal human influenza viruses, but much less is known about the antigenic evolution of HPAI H5N1 viruses that circulate in poultry. We hypothesized that the molecular basis for antigenic change of avian H5N1 viruses circulating in poultry is similar to that of seasonal influenza viruses circulating in humans. In this study we focused on HPAI H5N1 clade 2.1 viruses, which are enzootic to Indonesia. We selected a set of representative viruses from genetically distinct lineages that are currently circulating and determined their antigenic properties by hemagglutination inhibition assays. At least six antigenic variants have circulated between 2003, when H5N1 clade 2.1 viruses were first detected in Indonesia, and 2011. During this period, multiple antigenic variants co-circulated in the same geographic regions. Tests with mutant viruses constructed by site-directed mutagenesis revealed that antigenic differences between clade 2.1 viruses were due to single, double, or quadruple amino acid substitutions immediately adjacent to the receptor binding site. Antigenic variants of H5N1 virus evaded recognition by both ferret and chicken antibodies. The molecular basis for antigenic change of clade 2.1 viruses closely resembled that of seasonal human influenza viruses, indicating that the hemagglutinin of influenza viruses from different hosts and subtypes may be similarly restricted to evade antibody recognition.

Björn F. Koel,[1] Stefan van der Vliet,[1] David F. Burke,[2,3] Theo M. Bestebroer,[1] Eny E. Bharoto,[4] I. Wayan W. Yasa,[4] Inna Herliana,[4] Brigitta M. Laksono,[1,4] Kemin Xu,[6] Eugene Skepner,[2,3] Colin A. Russell,[3,7] Guus F. Rimmelzwaan,[1] Daniel R. Perez,[6] Albert D. M. E. Osterhaus,[1] Derek J. Smith,[1,2,3] Teguh Y. Prajitno,[4] Ron A. M. Fouchier[1]

[1]Department of Viroscience, Erasmus MC, Rotterdam, The Netherlands.
[2]Center for Pathogen Evolution, Department of Zoology, University of Cambridge, Cambridge, United Kingdom.
[3]WHO Collaborating Centre for Modeling Evolution and Control of Emerging Infectious Diseases, University of Cambridge, Cambridge, United Kingdom.
[4]JAPFA Comfeed, Indonesia.
[5]Department of Veterinary Medicine, University of Maryland, College Park, and Virginia-Maryland.
[6]Regional College of Veterinary Medicine, College Park, Maryland, USA.
[7]Department of Veterinary Medicine, University of Cambridge, Cambridge, United Kingdom.

MBio. 2014, 10; 5(3): e01070-14

IMPORTANCE

HPAI H5N1 viruses are responsible for severe outbreaks in both commercial and backyard poultry, causing considerable economic losses and regular zoonotic transmissions to humans. Vaccination is a primary method to reduce the burden of HPAI H5N1 virus in poultry. Influenza viruses can escape from recognition by antibodies induced upon vaccination or infection through genetic changes in the hemagglutinin protein. The evolutionary patterns and molecular basis of antigenic change of HPAI H5N1 viruses are poorly understood, hampering formulation of optimal vaccination strategies. We show that HPAI H5N1 viruses in Indonesia diversified into multiple antigenic variants, that antigenic differences were due to one or very few substitutions near the receptor binding site and that the molecular basis for antigenic change was remarkably similar to that of seasonal human influenza viruses. These findings have consequences for future vaccination and surveillance considerations, and contribute to the understanding of the antigenic evolution of influenza viruses.

INTRODUCTION

Since their first detection in Southeast Asia in 1997, highly pathogenic avian influenza (HPAI) viruses of subtype H5N1 have spread to numerous countries of the Eastern hemisphere where they caused outbreaks in poultry and diversified into multiple distinct genetic lineages or clades (17-19). In Indonesia, HPAI H5N1 viruses were first detected in Central Java in August 2003 (17). The viruses were classified to belong to clade 2.1, and from early 2004 onwards, multiple persistent genetic lineages within this clade have emerged in Indonesia (68). Within the first two years after initial detection, the clade 2.1 viruses spread over large parts of the country (69) and frequently caused severe outbreaks in both backyard flocks and commercial poultry farms. Economic losses resulting from HPAI H5N1 virus outbreaks in poultry were estimated at US$ 470 million by 2008 (70). Indonesia's first human case of H5N1 virus infection was identified in 2005 (71). Further zoonotic transmissions resulted in 193 reported human cases of infection, 161 of which were fatal (72).

The Indonesian government launched a mass vaccination program in poultry in February 2004 using A/Chicken/Legok/2003 as the recommended vaccine strain, though imported vaccines — often of different clades or subtypes (e.g. clade 1 or subtype H5N2 and H5N9) — were used by the poultry industry initially. Emergence of antigenic variants to which the vaccines did not provide protection was detected in poultry already in 2005. After a series of HPAI H5N1 virus outbreaks from 2006 until 2008, and as a result of monitoring avian influenza virus variants by the Indonesian OFFLU project (73), the poultry vaccine recommendation was updated to contain an A/Chicken/West Java/30/07 (A/Chicken/Nagrak/30/07) -like strain. With time, vaccination intensified in commercial poultry farms with nearly 100% vaccine coverage in breeders and layers, while broilers were vaccinated only during the rainy season. However, implementing a vaccination campaign for backyard poultry appeared to

be more challenging and vaccination coverage by 2008 was estimated to be less than 40% of the standing backyard and commercial poultry population in Indonesia (74). Despite vaccination campaigns, improved hygiene measures, movement controls, and pre-emptive culling of infected flocks, the virus is still enzootic in large parts of Indonesia.

In 2007, an international consultative group of scientists convened by the World Health Organization (WHO), the World Organization for Animal Health (OIE), and the Food and Agriculture Organization of the United Nations (FAO), reported on the diversification of clade 2.1 into clades 2.1.1, 2.1.2, and 2.1.3. Since 2007, clades 2.1.1 and 2.1.2 seemed to have gone out of circulation, while clade 2.1.3 viruses were further reported to have diversified into fourth order clades; 2.1.3.1, 2.1.3.2, and 2.1.3.3 (17-19, 75). Given the large genetic diversity of clade 2.1 viruses and continuing outbreaks in vaccinated poultry, we anticipated that clade 2.1 viruses also changed — and possibly diversified — antigenically over time.

Antibodies targeting the hemagglutinin (HA) surface glycoprotein are a critical component of the immune defense against influenza virus infections in both humans and animals (32). In humans, amino acid substitutions in HA have been shown to lead to escape from antibody-mediated neutralization. Studies performed in the 1980s defined five antigenic sites in the globular head of HA that contain the amino acid positions associated with antigenic change of seasonal human influenza viruses (32-34). Similar sites have been identified for HPAI H5N1 viruses (35, 36) and a recent study suggested that antigenic change of HPAI H5N1 viruses in poultry could be attributed to amino acid substitutions in antigenic sites similar to those defined for human influenza viruses (76). However, it was recently shown that antigenic change of seasonal human influenza viruses is primarily caused by amino acid substitutions in a small dominant antigenic domain immediately adjacent to the HA receptor binding site (RBS), rather than across nearly the entire globular head of HA (77). Here, we mapped the antigenic diversity of clade 2.1 viruses and identified the substitutions responsible for antigenic differences, to test the hypothesis that the molecular basis for antigenic change of avian clade 2.1 viruses circulating in poultry is similar to that of seasonal influenza viruses circulating in humans. Analysis of a subset of viruses that represented genetically distinct lineages within clade 2.1 revealed that at least six antigenically diverse variants circulated in Indonesia from 2003 – 2011. Naturally occurring single, double, or quadruple amino acid substitutions were found to determine all major antigenic differences between antigenically distinct viruses. These substitutions occurred at positions immediately adjacent to the RBS, indicating that clade 2.1 viruses evolved to evade neutralization by antibodies directed to the RBS region. These findings imply that avian clade 2.1 H5N1 viruses and seasonal human influenza viruses can escape antibody neutralization in a similar way.

RESULTS

Phylogenetic analysis and selection of representative viruses

The HA sequences of 96 avian clade 2.1 viruses isolated between 2003 and 2011 during HPAI H5N1 virus outbreaks in commercial poultry farms and backyard flocks in Indonesia were determined as part of ongoing surveillance studies (Table S1). Of these 96 HA sequences, 79 were from viruses isolated during the currently underreported period since 2008. To evaluate recent genetic variation of clade 2.1 viruses we constructed a maximum-likelihood phylogenetic tree using the sequences described in the present study and publically available clade 2.1 sequences (17-19, 78) (Figs. 1 and S1).

In general, viruses isolated between 2008 and 2011 were genetically more diverse than those isolated prior to 2008. Almost 85% of the viruses isolated since 2008 belonged to clade 2.1.3.2. Of the 92 clade 2.1.3.2 viruses isolated since 2008, 77 grouped in a single branch that forms the lower part of clade 2.1.3.2 and 94% of the 2010 and 2011 viruses grouped in this part of the tree. Viruses belonging to the other presently circulating clades were much less prevalent in recent years; only seven clade 2.1.3.1 viruses and five clade 2.1.3.3 viruses were isolated since 2008. Two viruses, A/Chicken/West Java/34/2008 and A/Chicken/West Java/6-1/2008, belong to a group of viruses that did not fit in any of the defined clades and are most closely related to clade 2.1.3.

Analysis of antigenic diversity

We next investigated the antigenic diversity of clade 2.1 viruses. We selected viruses from each of the genetically distinct lineages that circulated since 2008, including multiple viruses from the most prominent lineage, to represent the recent genetic diversity of clade 2.1 viruses (Fig. 1, Table 1). Two additional representative viruses were included; A/Chicken/West Java/30/07, which has been used as poultry vaccine strain since 2010, and A/Indonesia/5/05, which is the current clade 2.1.3.2 candidate vaccine virus for human use (79). The HA genes of the selected viruses were cloned without a multi-basic cleavage site and used to make 6:2 recombinant viruses in the context of the A/Puerto Rico/8/1934 reference virus with the neuraminidase gene segment of an H5N1 virus (60). We were unable to construct A/Chicken/West Java/6-1/08.

To select a panel of antisera that covered the antigenic variation between the representative viruses, we initially prepared antisera against a subset of the representative viruses that were genetically most divergent. The antisera were subsequently tested in hemagglutination inhibition (HI) assays using all representative viruses, and additional antisera were generated against the viruses that showed a divergent HI pattern from the viruses used for generation of the initial antiserum

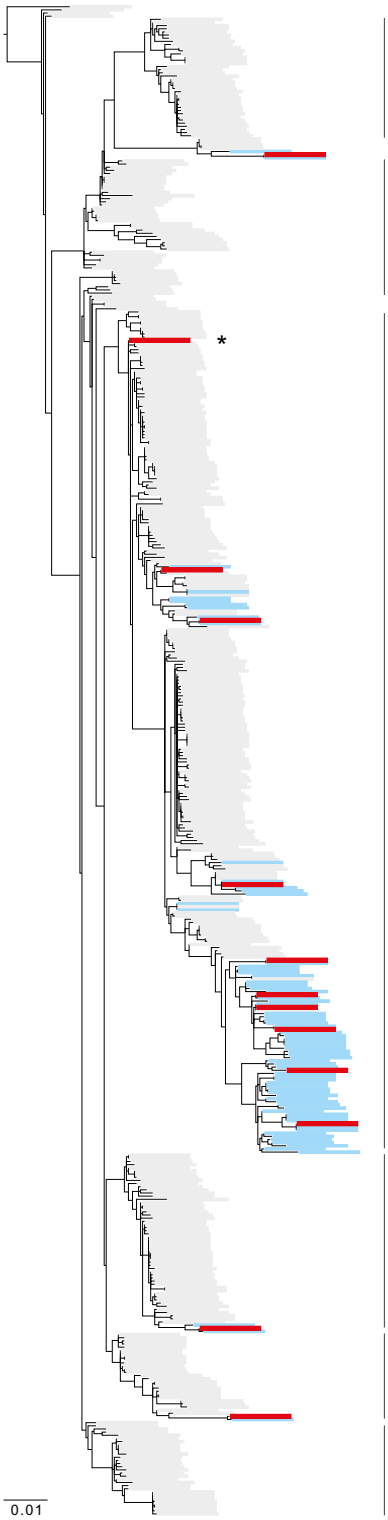


Fig. 1. Maximum likelihood phylogenetic tree for the HA gene of clade 2.1 viruses. HA sequences of 1600 nucleotides in length were used, with A/Goose/Guangdong/1/1996 as outgroup. The clades as defined previously are specified (17-19). The tree contains sequences from the present study from 2008 – 2011 (in blue) and publically available sequences (17-19, 78). Red bars indicate viruses selected for antigenic characterization (Table 1), including the recommended poultry vaccine strain (A/Chicken/West Java/30/07), and the candidate vaccine strain for human use (A/Indonesia/5/05). An asterisk indicates the position of A/Indonesia/5/05. The full tree is available as Fig. S1.

panel. The antigenic properties of all representative viruses were analyzed in HI assays using a panel of 8 – 16 ferret antisera (Table S2).

The representative viruses differentiated into six distinct antigenic variants (Fig. 2). A/Indonesia/5/05, A/Chicken/Central Java/51/09 and A/Chicken/North Sumatra/27/09 that belong to clade 2.1.3.2 were antigenically similar. Three other antigenic variants were distinguished within clade 2.1.3.2; A/Chicken/West Java/30/07, A/Chicken/West Java/119/10, and A/Chicken/East Java/121/10. The latter was antigenically similar to four other representative viruses of clade 2.1.3.2: A/Chicken/West Java/47/09, A/Chicken/Central Java/52/09, A/Chicken/West Java/59/09, and A/Chicken/West Java/90/10. The representative viruses for clades 2.1.3.1 and 2.1.3.3, A/Chicken/South Sulawesi/157/11 and A/Chicken/North Sumatra/72/10, respectively, were antigenically distinct from each other and all other representative viruses.

Surprisingly, A/Chicken/South Sulawesi/157/11 (clade 2.1.3.1) and A/Chicken/North Sumatra/72/10 (clade 2.1.3.3) were antigenically more closely related to A/Indonesia/5/05 (clade 2.1.3.2) than other antigenic variants of clade 2.1.3.2 (Fig. 2). *Vice versa*, A/Chicken/West Java/30/07 was genetically very similar to A/Indonesia/5/05 (Fig. S1), yet it was antigenically substantially more distinct from A/Indonesia/5/05 than A/Chicken/Central Java/51/09 and A/Chicken/North Sumatra/27/09, A/Chicken/South Sulawesi/157/11, and A/Chicken/North Sumatra/72/10 (Fig. 2). The correlation between the antigenic distance of the representative viruses from A/Indonesia/5/05 and the number of HA1 amino acid substitutions from A/Indonesia/5/05 was 0.4828, but was not statistically significant ($p = 0.1325$) (Table S3, Pearson correlation).

Mapping the molecular basis of antigenic change of clade 2.1 viruses

To gain insight in the molecular basis of antigenic change of clade 2.1 viruses we next performed site-directed mutagenesis of the HA of A/Indonesia/5/05 and generated

Table 1. Representative viruses selected for mapping the antigenic diversity of clade 2.1 viruses.

Strain	Clade
A/Indonesia/5/05	2.1.3.2
A/Chicken/West Java/30/07	2.1.3.2
A/Chicken/West Java/6-1/08	2.1.3.2
A/Chicken/North Sumatra/27/09	2.1.3.2
A/Chicken/West Java/47/09	2.1.3.2
A/Chicken/Central Java/51/09	2.1.3.2
A/Chicken/Central Java/52/09	2.1.3.2
A/Chicken/West Java/59/09	2.1.3.2
A/Chicken/North Sumatra/72/10	2.1.3.3
A/Chicken/West Java/90/10	2.1.3.2
A/Chicken/West Java/119/10	2.1.3.2
A/Chicken/East Java/121/10	2.1.3.2
A/Chicken/South Sulawesi/157/11	2.1.3.1

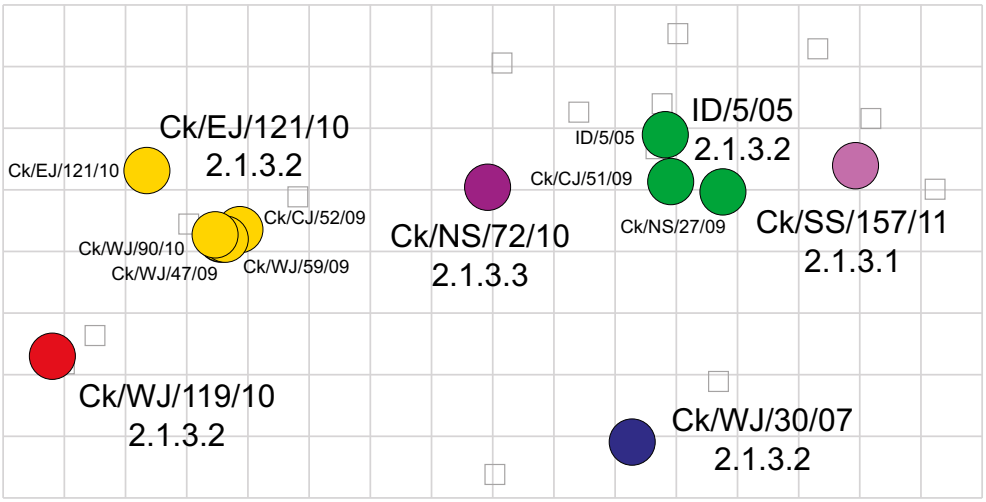


Fig 2. Antigenic map of selected clade 2.1 viruses. Filled circles and open squares indicate the positions of viruses and antisera, respectively. Both axes represent antigenic distance. The spacing between gridlines is one antigenic unit distance, which equals a two-fold difference in the HI assay. Positions of viruses and antisera in the map were generated using antigenic cartography methods as described previously (31). Briefly, the distance between a virus—antiserum pair is inversely related to the HI titer of the virus to that antiserum. Modified multidimensional scaling methods were used to arrange the relative positions of virus and antiserum points in the map such that they reflect the HI titers with minimal error. The viruses are color-coded to indicate the different antigenic variants: A/Indonesia/5/05; green, A/Chicken/West Java/30/07; blue, A/Chicken/North Sumatra/72/10; purple, A/Chicken/East Java/121/10; yellow, A/Chicken/West Java/119/10; red, and A/Chicken/South Sulawesi/157/11; pink. We considered viruses to be substantially antigenically different if the distance between them is at least two antigenic units. Abbreviations in the virus names are as follows: Ck, Chicken; CJ, Central Java; EJ, East Java; ID, Indonesia; NS, North Sumatra; SS, South Sulawesi; WJ, West Java.

recombinant viruses that were tested by HI assay. It was shown recently that all major antigenic changes of seasonal human influenza viruses were due to amino acid substitutions immediately adjacent to the RBS (77), and we hypothesized that the same would apply to H5N1 viruses. To test this hypothesis, the amino acid differences between HA of A/Indonesia/5/05 and the representative viruses from the remaining five antigenic variants were plotted on an H5N1 HA crystal structure to identify the amino acid substitutions that occurred near the RBS (Figs. S2 and S3). Subsequently, each of these substitutions was introduced independently into the A/Indonesia/5/05 HA gene and HI assays were performed with the mutant viruses to test if the substitution could fully explain the antigenic difference between A/Indonesia/5/05 and the strain in which the substitution was observed. Combinations of substitutions were tested when single substitutions did not fully explain the antigenic difference.

A/Indonesia/5/05 and A/Chicken/West Java/30/07 differed by five amino acids in the HA1 region (Fig. S2), of which substitutions 133SA and 185AE were located adjacent to the RBS. Single mutants of A/Indonesia/5/05 with either 133SA or 185AE in HA did not have the same antigenic properties as A/Chicken/West Java/30/07 (Fig. S4A). However, introduction of both 133SA and 185AE in A/Indonesia/5/05 resulted in a virus with antigenic properties that were similar to A/Chicken/West Java/30/07 (Figs. 3A and S4A). Thus, the antigenic difference between A/Indonesia/5/05 and A/Chicken/West Java/30/07 was due to substitutions 133SA and 185AE.

The five viruses that were antigenically A/Chicken/East Java/121/10-like (Fig. 2) shared six amino acids in the HA1 region that were different from A/Indonesia/5/05 (Fig. S2), four of which were located adjacent to the RBS; 155SN, 183DN, 184AE and 189RM. When introduced individually, none of these substitutions changed the antigenic properties of A/Indonesia/5/05 to become A/Chicken/East Java/121/10-like. However, substitutions 183DN and 189RM each had a substantial impact on the antigenic properties of A/Indonesia/5/05 (Fig. S4B), and we therefore constructed a double mutant virus. A/Indonesia/5/05 with substitutions 183DN and 189RM in HA was antigenically similar to the A/Chicken/East Java/121/10-like viruses,

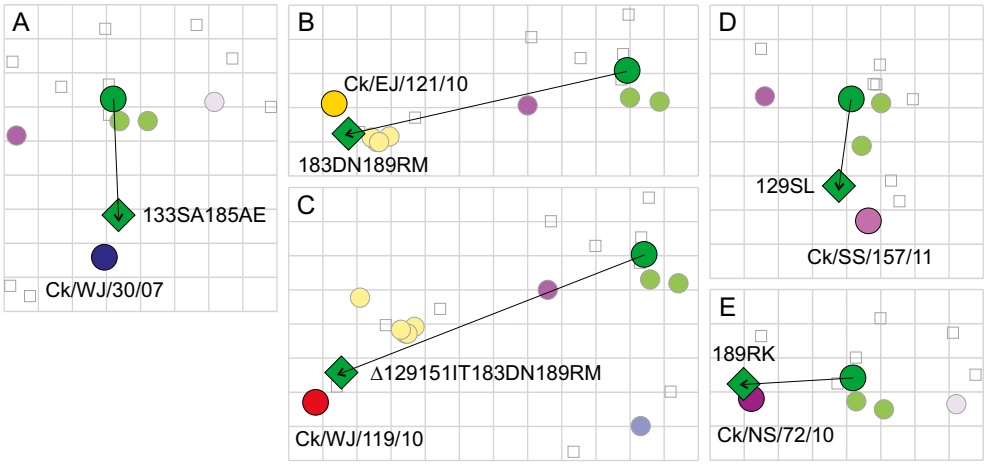


Fig 3. Summary of substitutions responsible for antigenic differences between representative viruses and A/Indonesia/5/05. The mutants were analyzed in the context of the antigenic map shown in Fig. 2. Panels A – E display a detail of the antigenic map and show the mutants with the substitutions responsible for the antigenic change from A/Indonesia/5/05 to: (A) A/Chicken/West Java/30/07, (B) A/Chicken East Java/121/10, (C) A/Chicken/West Java/119/10, (D) A/Chicken/South Sulawesi/157/11, and (E) A/Chicken/North Sumatra/72/10. A representative virus for each antigenic variant is shown as a large colored circle, the remaining representative viruses are shown as smaller faint colored circles. Mutants are indicated by green diamond shapes, A/Indonesia/5/05 as large green circles. Gridlines and antigenic map construction as in Fig. 1. Antigenic maps with all mutants made to investigate the molecular basis of antigenic change from A/Indonesia/5/05 are shown in Fig. S4. The antigenic distances between A/Indonesia/5/05, representative viruses and mutants are listed in Table S4.

indicating that the combination of substitutions 183DN and 189RM was responsible for the antigenic difference between A/Indonesia/5/05 and A/Chicken/East Java/121/10-like viruses (Figs. 3B and S4B).

A/Chicken/West Java/119/10 displayed 17 amino acid differences in HA1 as compared to A/Indonesia/5/05 (Fig. S2), of which 5 occurred at positions adjacent to the RBS: 136, 151, 159, 183, and 189. In addition, A/Chicken/West Java/119/10 HA had a deletion of the amino acid at position 129 as compared to the HA of A/Indonesia/5/05. These six amino acid changes were introduced individually in the HA of A/Indonesia/5/05, but none resulted in a change in antigenic properties to become A/Chicken/West Java/119/10-like (Fig. S4C). The Δ129151IT combination was previously shown to cause conformational changes around the RBS (22), which could affect antibody recognition. Mutations Δ129151IT in HA of A/Indonesia/5/05 indeed had a substantial effect on the antigenic properties of the virus (Fig. S4C). Since substitutions 183DN and 189RM were collectively responsible for the antigenic difference between A/Indonesia/5/05 and A/Chicken/East Java/121/10 (Fig. 3B), and both substitutions were also present in A/Chicken/West Java/119/10, we constructed a virus to combine Δ129151IT with 183DN and 189RM. This quadruple mutant virus was antigenically similar to A/Chicken/West Java/119/10 (Figs. 3B and S4C). Thus, we concluded that the deletion at position 129 and substitutions 151IT, 183DN, and 189RM were responsible for the antigenic difference between A/Indonesia/5/05 and A/Chicken/West Java/119/10.

A/Chicken/South Sulawesi/157/11 displayed 15 amino acid differences to the HA1 of A/Indonesia/5/05 (Fig. S2). Positions 129, 155 and 184 were located close to the RBS, and the corresponding substitutions were introduced individually into the HA of A/Indonesia/5/05. Substitutions 155SN and 184AE had no antigenic effect (Fig. S4D). Introduction of substitution 129SL into A/Indonesia/5/05 changed it antigenically to A/Chicken/South Sulawesi/157/11 (Figs. 3D and S4D).

There were 10 amino acid differences between HA1 of A/Indonesia/5/05 and A/Chicken/North Sumatra/72/10 (Fig. S2). We constructed pointmutants of the two substitutions that were located close to the RBS; 155SN and 189RK. Substitution 155SN did not have an antigenic effect (Fig. S4E). Introduction of 189RK was sufficient to change the antigenic properties of A/Indonesia/5/05 to become A/Chicken/North Sumatra/72/10-like (Fig. 3E and S4E).

To summarize these results, the amino acid changes responsible for antigenic change among all selected representative clade 2.1 viruses occurred at six key positions; 129, 133, 151, 183, 185 and 189. Five positions form a nearly continuous ridge located on the periphery of the RBS. The sixth, position 133, is located in the 130-loop that includes positions that are part of the RBS (Fig. 4).

Seven of the eight amino acid changes responsible for the antigenic differences from A/Indonesia/5/05 involved a large change in the biophysical properties of the amino acids (Table S5). Seven changes resulted in a modification of the hydrophilicity of the amino acids, five in a charge change, and a single

substitution involved a substantial increase in volume of the amino acid side chain. The A/Indonesia/5/05 HA amino acid sequence contains six predicted N-linked glycosylation sites (N-X-S/T-X) but these sites were not affected by the substitutions responsible for the antigenic changes, nor were any new glycosylation sites introduced by these substitutions.

We next marked the amino acid substitutions that were responsible for antigenic change of clade 2.1 viruses in the phylogenetic tree of Fig. 1 and analyzed in which years and geographic regions the antigenic variants identified in this study may have circulated (Fig. 5). The single substitutions responsible for the antigenic change to A/Chicken/North Sumatra/72/10 or A/Chicken/South Sulawesi/157/11 were detected in viruses of all third and fourth order clades of clade 2.1. These viruses were isolated intermittently between 2004 and 2011. Viruses that contained the 2 - 4 substitutions responsible for the antigenic change to A/Chicken/West Java/30/07, A/Chicken/East Java/121/10, and A/Chicken/West Java/119/10 were more confined to specific regions of the tree. The quadruple change that was responsible for the antigenic change to A/Chicken/West Java/119/10 was found in two unique sequences only, both of which were isolated in 2010.

Viruses with the substitutions typical for the A/Chicken/South Sulawesi/157/11, A/Chicken/East Java/121/10, A/Chicken/North Sumatra/72/10, and A/Indonesia/5/05 antigenic variants were isolated from multiple geographic regions of Indonesia (Fig. 5). The A/Chicken/West Java/119/10-like and A/Chicken/West Java/30/07-like viruses from our dataset were exclusively isolated on the island of Java. Co-circulation

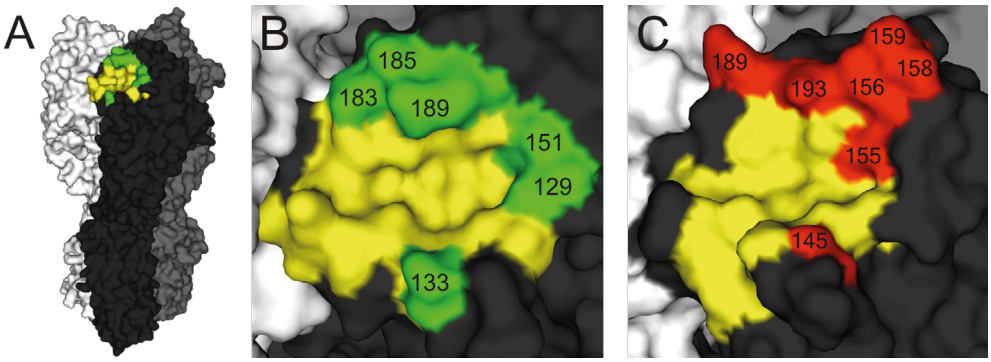


Fig 4. Amino acid positions responsible for antigenic change of clade 2.1 viruses plotted on an A/Indonesia/5/05 HA crystal structure. (A) Monomers are shown in black, grey and white, the RBS in yellow, and amino acid positions responsible for antigenic change in green. (B) Zoomed in on the area around the RBS. The amino acid positions are based on H5 numbering. (C) Amino acid positions responsible for antigenic change of seasonal human H3N2 viruses shown for comparison (77). Key positions are plotted on an A/Aichi/2/68 HA trimer and are indicated in red. The amino acid positions are based on H3 numbering. H5 amino acid positions 133, 151, 183, 185, and 189 correspond to H3 positions 137, 155, 187, 189, and 193, respectively. There is no equivalent of H5 amino acid position 129 in H3N2 viruses.

in the same geographic region of two or more viruses that had the substitutions typical for the antigenic variants identified here was noticed from 2005 onwards.

3

Antigenic effect of substitutions tested using chicken sera

Ferret antisera obtained after inoculation with influenza viruses under laboratory-controlled conditions are the most sensitive reagents for detection of antigenic variation of influenza viruses. We next examined if substitutions leading to antigenic variation as tested by ferret antisera also lead to evasion of recognition by antibodies in antisera of vaccinated chickens. Ten chickens were vaccinated with a vaccine prepared from A/Indonesia/5/05. Three weeks post-vaccination the antisera were harvested and tested in HI assays using the same protocol as for HI tests using ferret antisera.

The HI antibody titers against A/Chicken/West Java/30/07 were at least 4-fold lower than the HI antibody titers against A/Indonesia/5/05 with all ten A/Indonesia/5/05 antisera (Fig. 6A). Of the ten A/Indonesia/5/05 antisera, four had also at least four-fold lower HI titers against the 133SA185AE mutant as compared to A/Indonesia/5/05, five had two-fold or less lower HI titers, and the remaining antiserum had the same HI titers against the wildtype and mutant viruses (Fig. 6B). The HI antibody titers against A/Chicken/East Java/121/10 were also at least four-fold lower than the HI antibody titers to A/Indonesia/5/05 for eight of the ten A/Indonesia/5/05 antisera. The remaining two antisera had either a 2.7-fold lower HI titer against A/Chicken/East Java/121/10 as compared to A/Indonesia/5/05 or had the same HI titers (Fig. 6C). The HI antibody titers against the 183DN189RM mutant virus were at least 4-fold lower than against A/Indonesia/5/05 with all ten chicken antisera (Fig. 6D). These data thus show that two amino acid substitutions near the RBS can also be sufficient to substantially decrease recognition by chicken antibodies.

We next tested the effect of the same amino acid substitutions - 133SA185AE and 183DN189RM – on the antibody response of chickens upon vaccination. To this end, we vaccinated chickens with A/Indonesia/5/05 with substitutions 133SA185AE or A/Indonesia/5/05 with substitutions 183DN189RM. These post-vaccination antisera were compared with antisera raised against wildtype A/Indonesia/5/05 vaccine upon titration in HI assays against A/Indonesia/5/05 and A/Chicken/West Java/30/07 or A/Chicken/East Java/121/10. Antisera prepared by vaccination with the 133SA185AE mutant had on average 2.5-fold (1.2 $\log_2$) higher HI antibody titers against A/Chicken/West Java/30/07 than antisera prepared with A/Indonesia/5/05 vaccine (Fig. 7A). Antisera prepared by vaccination with the 183DN189RM mutant had on average 8.3-fold (3.2 $\log_2$) higher titers to A/Chicken/East Java/121/10 than antisera prepared with A/Indonesia/5/05 vaccine (Fig. 7B). We then tested if antisera prepared against the 133SA185AE vaccine inhibited the binding of A/Chicken/West Java/30/07 in HI assays, and did the same for 183DN189RM antisera and A/Chicken/East Java/121/10. Using the antisera raised against 133SA185AE vaccine, HI antibody titers against A/Chicken/West Java/30/07 were two to four-fold lower than to the homologous

virus for five of ten antisera, four-fold lower for four antisera, and six-fold lower for one antiserum (Fig. 7C). Using the antisera raised against the 183DN189RM vaccine, the HI antibody titers against A/Chicken/East Java/121/10 and homologous virus were the same for three of the ten antisera, less than two-fold lower than to the homologous virus for six of the ten antisera, and three-fold lower for one antiserum (Fig. 7D). Thus, amino acid substitutions responsible for antigenic change as defined with ferret antisera substantially affected the reactivity of chicken antisera upon vaccination.

DISCUSSION

Following the initial detection of H5N1 clade 2.1 viruses in 2003 in Central Java, the viruses diversified into three distinct genetic lineages (clades 2.1.1, 2.1.2, 2.1.3) from 2004 onwards and became enzootic in most parts of Indonesia (17-19, 69, 78). Since 2004, the ongoing genetic diversification of clade 2.1.3 viruses necessitated further classification into fourth order clades (2.1.3.1, 2.1.3.2, 2.1.3.3), while viruses

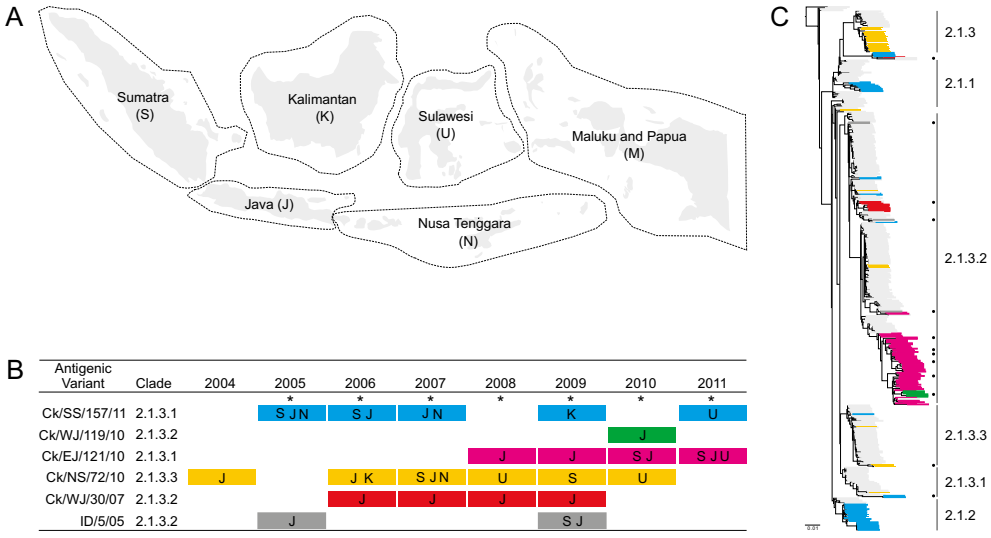


Fig 5. Identification of years and geographic region of isolation of clade 2.1 virus antigenic variants. (A) Geographical regions of Indonesia defined for this study. (B) Colored bars specify the years in which viruses with the substitutions that define the antigenic variants were isolated. The letters S, J, K, U, and N correspond to the geographical region from which isolates were obtained in a given year, and are as defined in panel A. An asterisk indicates the years in which multiple antigenic variants circulated in the same geographical region. Abbreviations in the virus names are provided in the legend to Fig. 2. (C) The phylogenetic tree of Fig. 1 color-coded for the substitutions that define the antigenic variants (Fig. 3). Color-coding corresponds to the antigenic variants specified in panel B. The positions of the representative viruses in the tree are indicated by a black dot. Sequences that did not contain the identified (combinations of) substitutions responsible for antigenic variation are shown in light grey.

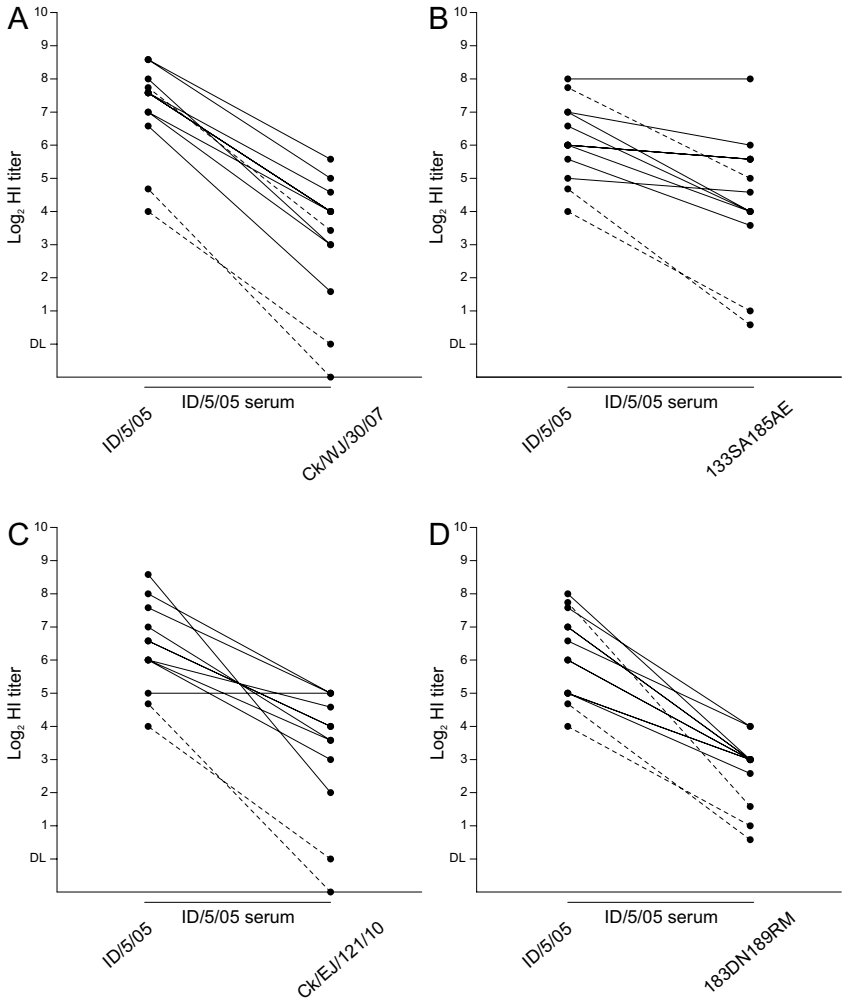


Fig 6. Reactivity of viruses with wildtype or mutant HAs to A/Indonesia/5/05 chicken antisera. Antisera obtained from chickens vaccinated with A/ Indonesia/5/05 were tested in HI assays to A/Indonesia/5/05 and Ck/West Java/30/07 (A), A/Indonesia/5/05 and A/Indonesia/5/05 + 133SA185AE (B), A/Indonesia/5/05 and Ck/East Java/121/10 (C), or A/Indonesia/5/05 and A/Indonesia/5/05 + 183DN189RM (D). Solid lines connect the $\log_2$ HI titers obtained for viruses tested to the same chicken antiserum. Dashed lines connect $\log_2$ HI titers obtained from HI assays using ferret antisera. DL indicates the detection limit of the HI assay at the 1:20 starting dilution that was used.

of clades 2.1.1 and 2.1.2 have apparently disappeared. However, data has been extremely sparse since 2008, which has hampered further analysis of clade 2.1 virus evolution. The present study shows that viruses from clades 2.1.3.1, 2.1.3.2, and 2.1.3.3 have circulated between 2008 and 2011, while viruses that belong to clades 2.1.1 or 2.1.2 were not detected. This finding further substantiates the suggestion that

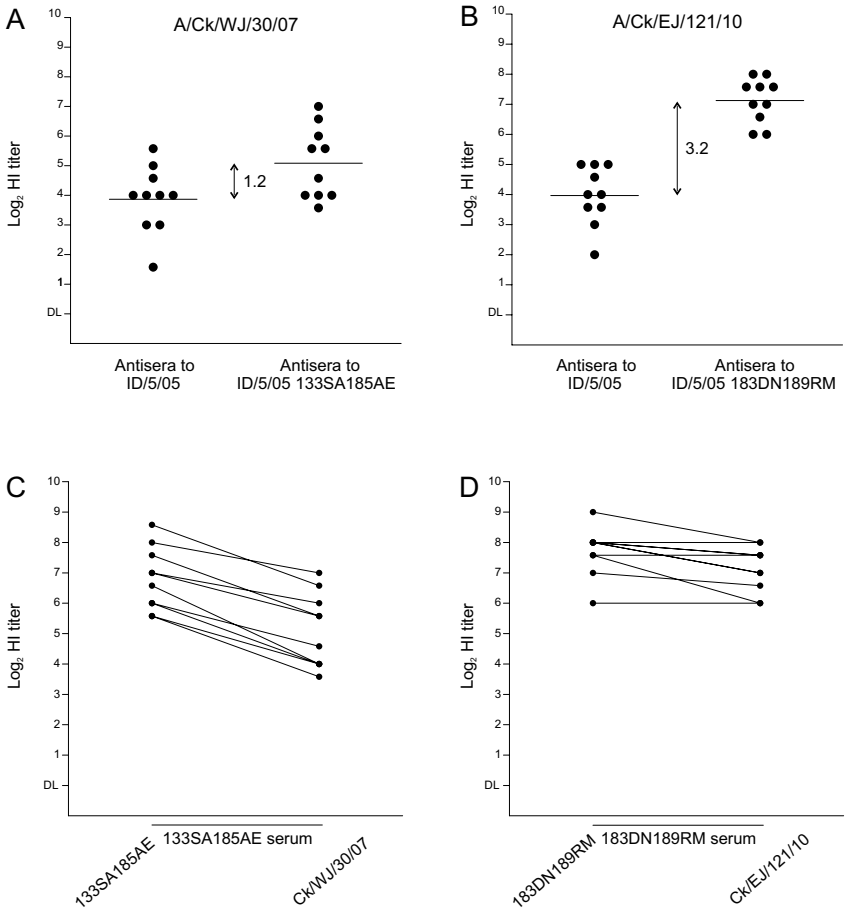


Fig 7. Effect of substitutions responsible for antigenic change on the antibody response of chickens. Antisera obtained from chickens vaccinated with A/Indonesia/5/05, A/Indonesia/5/05 with substitutions 133SA185AE, or A/Indonesia/5/05 with substitutions 183DN189RM were titrated in HI assays to A/Indonesia/5/05 and A/Chicken/West Java/30/07 (A), or A/Indonesia/5/05 and A/Chicken/East Java/121/10 (B). Each point represents the $\log_2$ HI titer for an individual antiserum. Horizontal lines show the mean $\log_2$ HI titer. The mean $\log_2$ HI titer difference is indicated between the groups. Antisera obtained from chickens vaccinated with A/Indonesia/5/05 with substitutions 133SA185AE or A/Indonesia/5/05 with substitutions 183DN189RM were titrated in HI assays to homologous virus and A/Chicken/West Java/30/07 (C), or homologous virus and A/Chicken/East Java/121/10 (D). Solid lines connect the $\log_2$ HI titers obtained for viruses tested to the same chicken antiserum. DL indicates the detection limit of the HI assay at the 1:20 starting dilution that was used.

viruses of these latter clades are no longer circulating in poultry. Although viruses from clades 2.1.3.1 and 2.1.3.3 were isolated, the majority of viruses belonged to clade 2.1.3.2 indicating that this was the dominant genotype in recent years. The dataset used in this study contains sequences from viruses isolated throughout Indonesia, but the majority of samples originate from outbreaks on the island of Java. This

geographical bias is (at least in part) explained by the poultry population densities within Indonesia; 70% of the commercial poultry population and poultry production is on the island of Java.

We selected viruses to represent the most prominent recently circulating lineages and used these viruses to map the antigenic diversity of clade 2.1 viruses. Using antigenic cartography methods, we show that at least six antigenically distinct variants circulated in Indonesia since 2003. Representative viruses that belonged to different phylogenetic clades, or to distinct lineages within the same clade, were generally also antigenically distinct. However, antigenic distances between the variants correlated poorly with distances in the phylogenetic tree or with the number of amino acid differences in HA1. As was suggested previously (35), antigenic differences between clade 2.1 viruses are therefore not necessarily predictable from (phylo)genetic information alone. Antigenic analysis of the representative viruses also indicated that new antigenic variants emerged with time, but that these new variants did not replace previously circulating viruses *per se*, and existing antigenic variants continued to circulate alongside newly emerged, antigenically distinct, viruses. However, the predominance of isolates with the 183DN189RM substitutions that define the A/Chicken/East Java/121/10 antigenic variant in recent years may also signal ongoing directional selection towards viruses of this antigenic phenotype or genotype.

Previously, we reported that major antigenic changes during the evolution of seasonal human H3N2, H1N1, and influenza B viruses were caused by very few amino acid substitutions that occurred exclusively at positions immediately adjacent to the RBS (77). Although amino acid substitutions that affect the antigenic properties of H3N2 and H1N1 influenza viruses can occur on positions distributed over large parts of the HA globular head (32-34) nature has selected for major antigenic change of these subtypes caused by substitutions at only few positions located adjacent to the RBS. In the present study we tested the hypothesis that the molecular basis for antigenic change of clade 2.1 viruses resembles that of seasonal human influenza viruses. Indeed, antigenic differences from A/Indonesia/5/05 were due to single, double, or, in a single case, quadruple amino acid substitutions only. Each of these substitutions occurred at a small set of amino acid positions located immediately adjacent to the RBS. Five key positions are located on the membrane distal periphery of the RBS and form a nearly continuous antigenic ridge, while the sixth position is located on the membrane proximal part of the RBS and is located in the 130-loop that is involved in receptor binding.

The location of substitutions responsible for antigenic change of clade 2.1 viruses is thus remarkably similar to that of seasonal human influenza viruses. The key substitutions structurally map to positions comparable to those responsible for antigenic change of human H3N2 virus (Fig. 4) (77), and a single or very few substitutions near the RBS were responsible for the antigenic differences. Interestingly, antigenic change of a clade 2.2.1 virus from Egypt was also due to substitutions

near the RBS (76). The combination of Δ129 and 151IT involved in the antigenic change from A/Indonesia/5/05 to A/Chicken/West Java/119/10 was observed in clade 2.2.1 viruses isolated between 2007 and 2009 (80). Consistent with the finding that Δ129151IT results in a conformational change of the area around the RBS, we found that combination of these amino acid changes resulted in reduced antibody recognition of clade 2.1 viruses.

We previously found that large changes in the biophysical properties were involved in all major antigenic changes of H3N2 viruses (77). The antigenic changes of clade 2.1 viruses similarly involved substitutions with large changes in the biophysical properties of the amino acids responsible for the antigenic difference from A/Indonesia/5/05 to four of the five antigenic variants. Charge changes of the amino acids seemed to play an important role in antigenic change of clade 2.1 viruses as well as human H3N2 viruses. Interestingly, substitutions that involved hydrophilicity changes were common in clade 2.1 virus, but were almost completely absent among substitutions that led to antigenic change of H3N2 viruses. Variation in the number of carbohydrate side chains on HA has been associated with changes in the antigenic properties of influenza A viruses because they can mask antibody epitopes (61, 62), but changes in glycosylation sites were not involved in the major antigenic changes of clade 2.1 viruses. The location of the amino acid positions involved in antigenic change of clade 2.1 viruses exclusively in the periphery of the RBS implies that clade 2.1 viruses, like seasonal human influenza viruses, evolved to specifically evade the antibodies directed at the RBS region. Follow-up studies looking into the effect of substitutions involved in antigenic change on receptor binding could help to understand potential limitations of influenza virus antigenic change.

Evolutionary patterns that become noticeable after extensive periods of virus circulation may not yet be clear in the relatively short period of HPAI H5N1 virus circulation in Indonesia. However, we observed some distinct differences in the antigenic evolution of clade 2.1 H5N1 and seasonal human influenza viruses. Human influenza viruses accumulate antigenically important substitutions over time. To avoid neutralization by antibodies raised to an earlier strain, antigenic evolution is typically away from previously circulating antigenic variants and shows a somewhat linear pattern (31). Moreover, antigenic variants are periodically replaced by newly emerging variants that can better evade population immunity. In contrast, the substitutions responsible for antigenic change of clade 2.1 viruses did typically not accumulate over time, and antigenic evolution did not follow a clear pattern away from earlier viruses. Additionally, multiple antigenic variants co-circulated in the same geographical region during multiple consecutive years. A plausible explanation for the different evolutionary patterns of seasonal human influenza viruses and clade 2.1 viruses is the difference in selective pressures acting on these viruses. The short life span of chickens, approximately 28 – 35 days for broilers and 18 months for breeders and layers in commercial poultry farms, prevents build up of immunity to multiple antigenic variants in the poultry population. Similar to what was described by de Jong

et al. for H3N2 viruses in pigs (30), the necessity to evade population immunity that was acquired through infection or vaccination with consecutive antigenic variants is absent in poultry. Population immunity is therefore unlikely to be a major driving force of clade 2.1 virus antigenic evolution, in contrast to seasonal human influenza virus. Moreover, whereas antigenic variants of seasonal human influenza viruses can spread around the globe in a matter of months (81), the limited spatial and temporal mixing of chicken populations limits competition among antigenic variants, and therefore may allow co-existence of multiple antigenically different viruses in the same geographic region of Indonesia.

Vaccination can be an efficient method to reduce the burden of HPAI H5N1 viruses in poultry (82-84). The current poultry vaccine strain recommendation for Indonesia is A/Chicken/West Java/30/07 or an antigenically similar strain. Viruses of this antigenic phenotype circulated until at least 2009. However, our analyses indicate that A/Chicken/East Java/121/10-like viruses have become the major antigenic variant in more recent years. Jadhao *et al*. (83) showed that chickens vaccinated with an antigenically poorly matched vaccine are fully protected from morbidity and mortality, but shed virus at much higher levels than those vaccinated with an antigenically matched strain. Reduced virus shedding decreases the probability of further spread between poultry flocks and zoonotic events. It is therefore advantageous to pursue the use of vaccine strains that match the antigenic variants in circulation, if vaccination is chosen as one of the approaches to stop or limit the spread of HPAI viruses.

Antigenic analysis of clade 2.1.3.2 viruses isolated from humans in 2011 and 2012 indicated that these viruses were antigenically different from A/Indonesia/5/05, which is the current candidate vaccine strain for human use. An updated vaccine based on an A/Indonesia/NIHRD11771/2011-like virus, which has the 183DN and 189RM substitutions typical for the A/Chicken/East Java/121/10-like viruses, is currently in preparation (85).

Updating the vaccines for poultry or human use to a contemporary strain that matches the major antigenic variant is an important measure that helps to ensure optimal vaccine efficacy. Given that one or two amino acid substitutions near the RBS can already substantially decrease recognition by ferret and chicken antibodies, the long-term efficacy of vaccines that target a single antigenic variant is questionable. We here show that vaccination of chickens with a recombinant virus that only contains two key substitutions for antigenic change can be sufficient to substantially change the reactivity of chicken antisera to antigenic variants. Recombinant viruses can be designed to contain the key amino acids necessary to match an antigenic variant while maintaining the growth characteristics endowed by other parts of the virus genome, and can be constructed without multi basic cleavage site. Although the efficacy of such vaccines should be evaluated in *in vivo* models, antigenically engineered viruses may provide a means to quickly adapt the vaccine to emerging antigenic variants. However, antigenic variants have co-circulated in the same geographic region of Indonesia from 2005 onwards. These variants were

isolated from chickens, thus excluding that they circulated exclusively in different host-species. Therefore, any vaccine that is designed to neutralize a single antigenic variant may be of only limited use. The identification of persistent co-circulation of antigenic variants stresses the need for more potent, broadly reactive, next generation vaccines. At the same time, this study underscores the need for continued monitoring of circulating influenza viruses in poultry, certainly when vaccines are employed to contribute to control outbreaks.

Analysis of the amino acid sequence variation immediately adjacent to the RBS region of HA can be applied as a proxy to identify the most prominent antigenic phenotypes, and could provide an early indication of emerging antigenic variants. Such genetic data should be complemented with antigenic data and vaccination-challenge experiments with the circulating viruses, to ensure that the vaccines continue to match the circulating viruses associated with outbreaks.

MATERIALS AND METHODS

Sample collection and virus isolation

Samples were collected during the investigations of outbreaks in both commercial and backyard poultry flocks from 2003 onwards. From 2008 onwards, samples were collected during outbreaks in commercial farms and backyard poultry throughout Indonesia as part of ongoing surveillance studies. Tracheal and oropharyngeal swabs and organ samples—typically the brain, trachea, lung, spleen, pancreas, and intestines—were used to inoculate specific pathogen free embryonated chicken eggs and allantoic fluids were tested for viral presence by hemagglutination assay. Subsequently, virus isolates were subject to H5N1 virus specific reverse-transcription polymerase chain reaction followed by sequencing of the full HA gene.

Construction of phylogenetic trees

The HA sequences of 96 H5N1 viruses isolated in Indonesia that were obtained as described above were combined with the publically available clade 2.1 HA sequences (17-19, 78). Phylogenetic tree construction of HA sequences was performed using the maximum likelihood method available from the PAUP* package (version 4.0b10) (86) under the GTR + I + Γ_4 model (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) as determined by ModelTest (87). Global optimization of the tree topology was performed by tree bisection–reconnection branch swapping. The tree was rooted to A/Goose/Guangdong/1/1996.

Cells

293T cells were cultured in DMEM (Cambrex, Heerhugowaard, The Netherlands) supplemented with 10% fetal calf serum (FCS), 100 µg/ml streptomycin, 100 IU/ml

penicillin, 2mM glutamine, 1mM sodiumpyruvate, and non-essential amino acids (MP Biomedicals, Europe, Illkirch, France). Madin Darby Canine Kidney (MDCK) cells were cultured in EMEM (Cambrex) supplemented with 10% FCS, 100 IU/ml penicillin, 100 µg/ml streptomycin, 2mM glutamine, 1.5 mg/ml sodiumbicarbonate (Cambrex), 10mM Hepes (Cambrex), and non-essential amino acids (MP Biomedicals).

Plasmid construction and site-directed mutagenesis

A/Indonesia/5/05 was isolated from a human case of HPAI virus infection (71), and passaged once in embryonated chicken eggs followed by a single passage in MDCK cells before amplification of the HA segment by reverse transcription polymerase chain reaction. The HA gene of A/Chicken/North Sumatra/27/09 was synthesized by Life Technologies Corporation (Carlsbad, CA, USA), while full HA genes of other representative viruses (Table S1) were synthesized by BaseClear (Leiden, The Netherlands). All HA genes were synthesized without multi basic cleavage site and were cloned into the modified pHW2000 expression plasmid as described previously (64, 65). We were unable to construct the HA of A/Chicken/West Java/6-1/08. Mutations were introduced into the HA gene of A/Indonesia/5/05 using the QuickChange multi-site directed mutagenesis kit (Agilent Technologies, Santa Clara, CA, USA) according to the manufacturers instructions.

Generation of recombinant viruses

Plasmids containing wildtype or modified HA genes were used to generate recombinant viruses by reverse genetics as described elsewhere (64). Briefly, one day prior to transfection 293T cells were seeded in 100 mm gelatinized culture dishes. Cells were transfected overnight with 40 µg of plasmid DNA. Subsequently, transfection medium was replaced by fresh medium supplemented with 2% FCS. After incubation for 72 hours at 37°C / 5% CO_2 the supernatant was harvested. Virus stocks were propagated by inoculation of MDCK cells with 2mL 293T transfection supernatant. Inoculum was removed after 2 hours and replaced by MDCK infection medium, consisting of EMEM (Cambrex), 100 IU/ml penicillin, 100 µg/ml streptomycin, 2mM glutamine, 1.5 mg/ml sodium bicarbonate, 10mM Hepes, non-essential amino acids, and 25 µg/ml trypsin. Viruses were harvested after incubation at 37°C / 5% CO_2 for 72 hours. All viruses were rescued as 6:2 viruses and contained six segments of A/Puerto Rico/8/1934 (A/H1N1), the neuraminidase (NA) gene of A/Hong Kong/156/1997 (H5N1), and the HA gene of the appropriate clade 2.1 viruses, except for A/Indonesia/5/05 and A/Chicken/West Java/30/07 that contained the NA from A/Indonesia/5/05. Absence of undesired mutations and presence of introduced substitutions was confirmed by sequence analysis.

Antisera

Ferret antisera were prepared by intranasal inoculation with 500 µl recombinant virus stock. Fourteen days post inoculation ferrets were boosted using 500 µL of a 1:1 mix of concentrated virus (>2048 HAU) and Titermax Gold Adjuvant (Sigma-Aldrich,

St. Louis, MO, USA). Antisera were collected four weeks later. Chicken antisera were prepared by subcutaneous vaccination of 4-week old SPF White Leghorn chickens (Charles River Laboratories, Wilmington, MA) with 0.5 mL of a 1:1 mix of 512 HAU of virus and Montanide ISA50V (Seppic, France). Antisera were collected three weeks post-vaccination. All antisera were pre-treated overnight at 37°C with receptor destroying enzyme (*Vibrio cholerae* neuraminidase) followed by inactivation for 1h at 56°C before use in hemagglutination inhibition (HI) assays.

Serological assays

HI assays were performed following standard procedures (37). Briefly, two-fold serial dilutions of the antisera, starting at a 1:20 dilution, were mixed with 25 µL of a virus stock containing 4 hemagglutinating units and were incubated at 37°C for 30 minutes. Subsequently, 25 µL 1% turkey erythrocytes was added and the mixture was incubated at 4°C for 1h. The HI titer is expressed as the reciprocal value of the highest serum dilution that completely inhibited hemagglutination.

Antigenic cartography

Analysis of antigenic properties was performed using antigenic cartography methods as described previously (31). Briefly, antigenic cartography is a method to increase the resolution of, solve paradoxes in, and visualize HI assay data or other binding assay data. In an antigenic map, the distance between antigen point A and antiserum point S corresponds to the difference between the $\log_2$ of the maximum observed titer to antiserum S from any antigen and the titer of antigen A to antiserum S. The titers in an HI table can be thought of as specifying target distances between the antigens and the antisera. Modified multidimensional scaling methods are then used to arrange the antiserum and antigen points in an antigenic map to best satisfy the target distances specified by the HI data. The distances between the points in an antigenic map represent antigenic distance as measured by the HI assay, in which the distances between antigens and antisera are inversely related to the $\log_2$ HI titer. Because antigens are tested against multiple antisera, and antisera are tested against multiple antigens, many measurements can be used to determine the position of the antigens and antisera in an antigenic map, thus improving the resolution of HI data.

Structural analysis

The amino acid positions responsible for antigenic change were plotted on the crystal structure of the A/Indonesia/5/2005 virus HA (PDB accession code 4K62 (88)). MacPyMOL (The PyMOL Molecular Graphics System, Version 1.3, Schrödinger, LLC) was used to visualize the trimer.

Ethics Statement

The surveillance program from which virus samples were derived was initiated following the recommendations of the Government of Indonesia to monitor vaccinated flocks and the environment. No special permits were required for this

program as it only included flocks and the environment of farms associated with the author's organizations. Ferrets were housed and experiments were conducted in strict compliance with European guidelines (EU directive on animal testing 86/609/EEC) and Dutch legislation (Experiments on Animal Act, 1997). The protocol was approved by the independent animal experimentation ethical review committee 'stichting DEC consult' (Erasmus MC permit number EMC 2114) and was performed under ABSL-3 conditions. Animal welfare was monitored daily and all animal handling was performed under light anesthesia (ketamine) to minimize animal discomfort. Studies that involved the use of chickens were conducted under ABSL-3 conditions approved by USDA and performed according to the protocol R-12-51 "Inactivated vaccines against antigenic variants of highly pathogenic H5N1 influenza viruses" approved by the Institutional Animal Care and Use Committee of the University of Maryland. Chicken studies adhered strictly to the U.S. Animal Welfare Act (AWA) laws and regulations.

ACKNOWLEDGEMENTS

We thank M. de Graaf and K.B. Westgeest for their assistance in phylogenetic analyses. DFB and DJS acknowledge the use of the CamGrid distributed computing resource. EEB, IWY, IH, BML and TYP are employed by JAPFA Comfeed. ADMEO (on behalf of Viroclinics Biosciences BV) has advisory affiliations with GSK, Novartis, and Roche. ADMEO and GFR are consultants for Viroclinics Biosciences BV. This project was initiated by OFFLU and continued under National Institute of Allergy and Infectious Diseases-NIH contract HHSN266200700010C and a ZonMw VICI grant.

SUPPLEMENTAL MATERIALS

Antigenic Variation of Clade 2.1 H5N1 Virus is Determined by Few Amino Acid Substitutions Immediately Adjacent to the Receptor Binding Site

Björn F. Koel, Stefan van der Vliet, David F. Burke, Theo M. Bestebroer, Eny E. Bharoto, I. Wayan W. Yasa, Inna Herlian, Brigitta M. Laksono, Kemin Xu, Eugene Skepner, Colin A. Russell, Guus F. Rimmelzwaan, Daniel R. Perez, Albert D. M. E. Osterhaus, Derek J. Smith, Teguh Y. Prajitno, Ron A. M. Fouchier

Figs. S1 to S4
Tables S1 to S5

Zoomable pdf files are available from:
http://mbio.asm.org/content/5/3/e01070-14.full.html#SUPPLEMENTAL

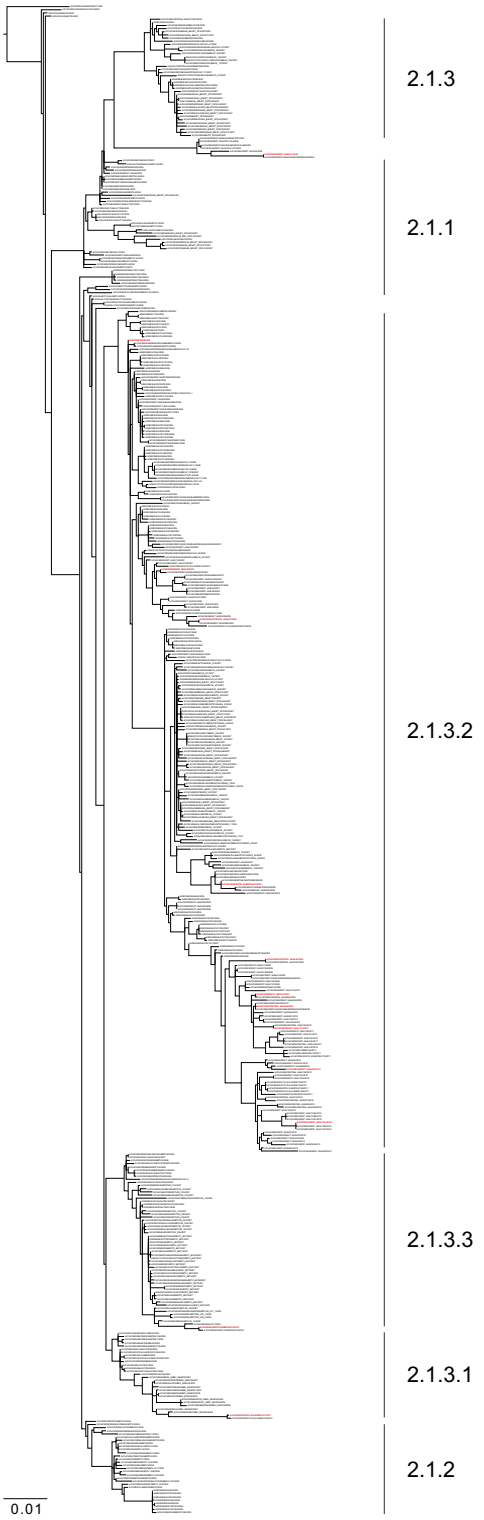


Fig. S1. Phylogenetic tree for the HA gene of clade 2.1 viruses. The tree is the same nucleotide maximum-likelihood tree as shown in Fig. 1 of the main text, but shows full virus names. Representative viruses are indicated in red. The clades as defined previously are specified in the tree (17-19).

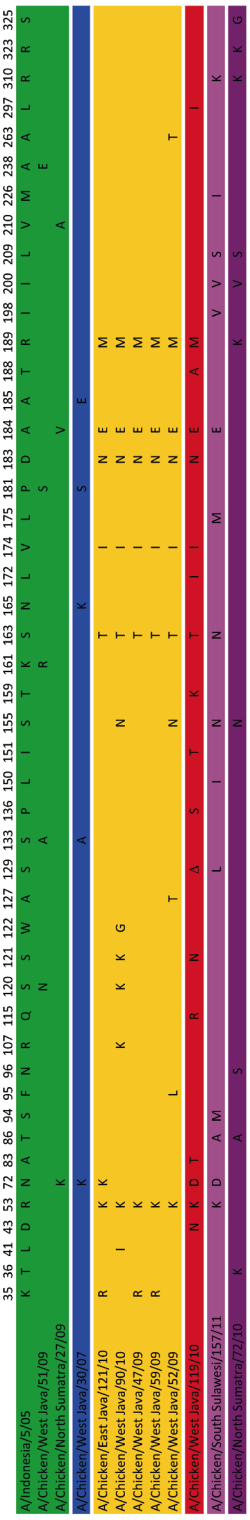


Fig. S2. HA1 amino acid sequence alignment of the representative viruses selected for mapping the antigenic diversity of clade 2.1 viruses. The sequences are grouped and color-coded according to their antigenic phenotype (Fig. 2). The top row indicates the amino acid positions where substitutions from A/Indonesia/5/05 occurred. Consensus amino acid positions are excluded from the alignment.

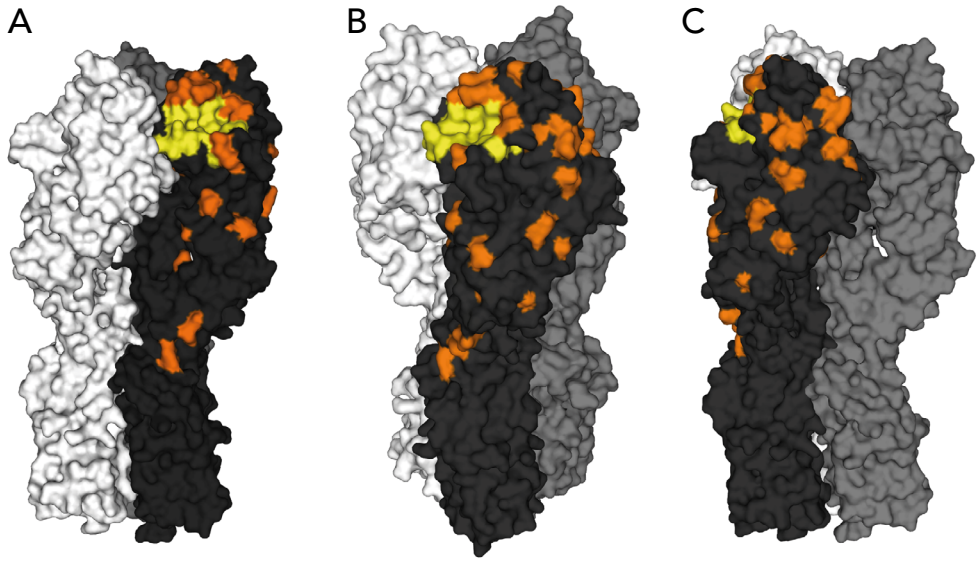
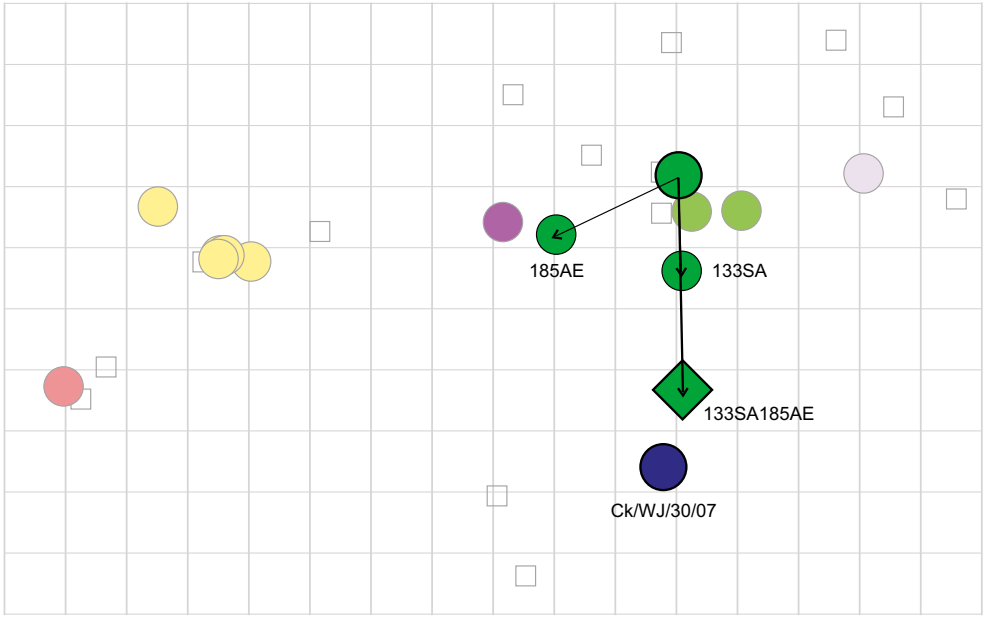


Fig. S3. Amino acid sequence variation of representative clade 2.1 viruses plotted on an A/Indonesia/5/05 crystal structure. Monomers are shown in black, grey and white, the receptor binding site in yellow. The amino acid positions on which differences from A/Indonesia/5/05 were observed in the representative viruses (Fig. S2) are indicated in orange. Panels A, B and C show different angles of the same crystal structure.

Fig. S4.

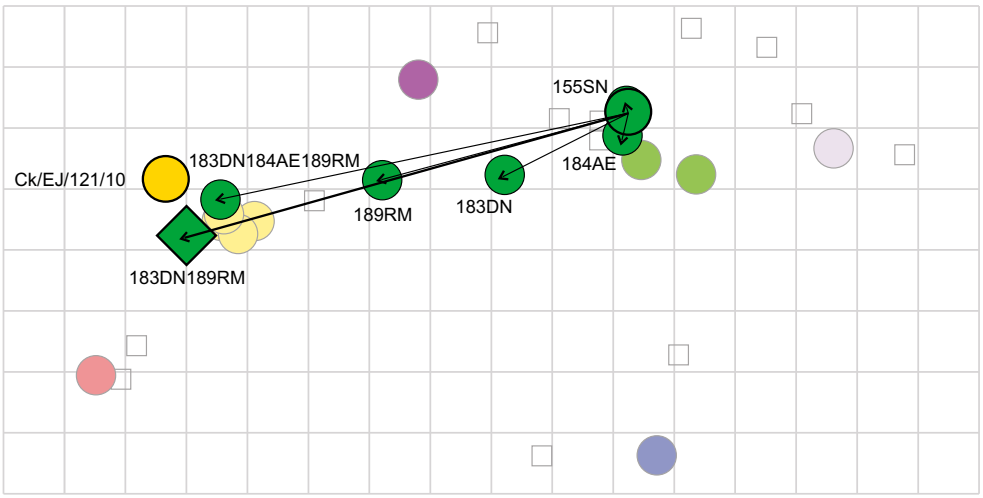
Antigenic maps for the analysis of mutants constructed to identify the molecular basis for antigenic differences from A/Indonesia/5/05. The antigenic maps are similar to those shown in Fig. 2, but in addition the positions of all tested mutants is indicated. A representative virus for each antigenic variant is shown as a larger colored circle, A/Indonesia/5/05 is indicated as a larger green circle, and the remaining representative viruses from Fig. 2 are shown as smaller faint colored circles. A diamond shape indicates the mutant with the minimal set of substitutions that determined the antigenic difference between A/Indonesia/5/05 and the respective representative virus. The remaining mutants are indicated as smaller green filled circles. Antisera are represented by open squares. The spacing between gridlines is one antigenic unit distance, which equals a two-fold difference in the HI assay. A table below each map shows the HI titers for the representative viruses and mutants. Antisera are named according to the viruses that were used to prepare the antisera. Suffices A, B, C, etcetera, in the antisera names (e.g. ID/5B/05) indicate that multiple antisera were raised to the same antigen. Only those serum columns with at least one titer above the detection limit of the assay (typically a 1/10 dilution of the antiserum) are shown. HI titers below the detection limit are shown as threshold values (e.g. <10). NT indicates that the antigen – antiserum combination was not tested. Abbreviations

in the virus names are as follows: Ck, Chicken; CJ, Central Java; EJ, East Java; ID, Indonesia; NS, North Sumatra; SS, South Sulawesi; WJ, West Java.



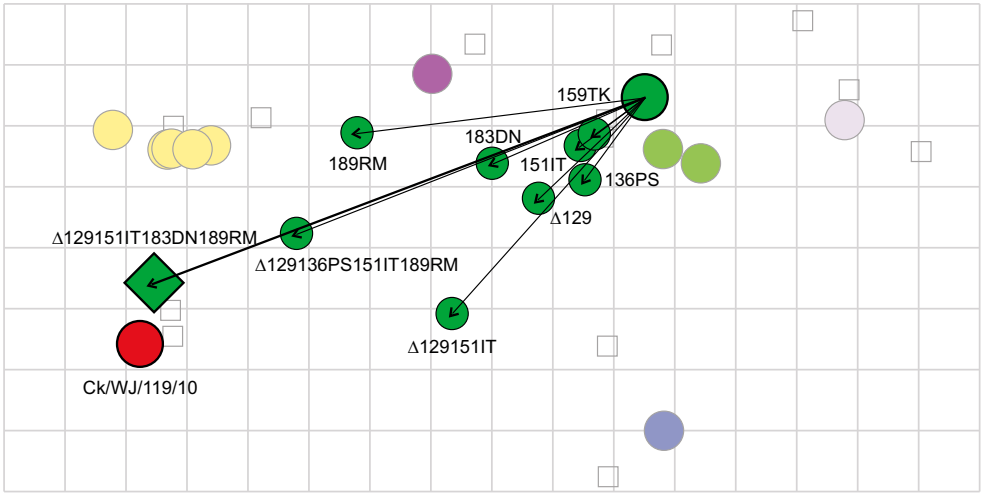
	ID/5A/05	ID/5B/05	Ck/ID/5D/05	Ck/WJ/30A/07	Ck/WJ/30B/07	Ck/CJ/51A/09	Ck/CJ/051B/09	Ck/NS/72/10	Ck/EJ/121B/10	Ck/SS/157A/11	Ck/SS/157B/11
ID/5/05	256	2136	160	<80	<80	446	1126	158	146	193	200
133SA	40	1280	70	40	<20	240	960	160	120	70	40
185AE	160	1280	NT	120	NT	480	960	120	240	NT	NT
133SA185AE	<30	320	<40	30	40	124	554	<40	<70	<10	<10
Ck/WJ/30/07	<10	108	<20	196	239	203	207	<20	<10	<10	<10

Fig. S4A. A/Indonesia/5/05 - A/Ck/West Java/30/07.



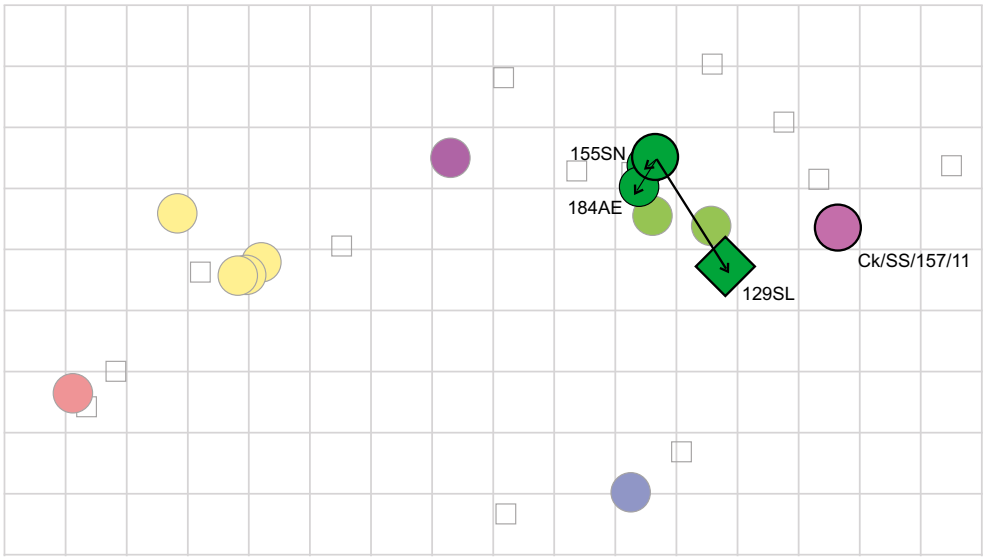
	ID/5A/05	ID/5B/05	Ck/WJ/30A/07	Ck/CJ/51A/09	Ck/CJ/051B/09	Ck/NS/72/10	Ck/WJ/119A/10	Ck/WJ/119B/10	Ck/EJ/121A/10	Ck/EJ/121B/10	Ck/SS/157A/11	Ck/SS/157B/11
ID/5/05	256	2136	<80	446	1126	158	<40	<40	<40	146	193	200
155SN	320	1810	24	480	1109	277	<10	<40	<40	240	320	640
183DN	320	2560	40	640	1280	240	80	40	240	320	NT	NT
184AE	320	1810	35	960	1109	160	<30	<40	<60	240	640	240
189RM	<10	240	60	320	160	640	30	<40	1280	2560	NT	NT
183DN189RM	<30	<60	<10	<10	<80	<40	480	640	1920	2560	<10	<10
183DN184AE189RM	<30	<60	<10	40	<80	<40	320	280	1280	2240	NT	NT
Ck/EJ/121/10	<10	<40	<10	<40	<80	<20	196	197	796	1091	<10	<10

Fig. S4B. A/Indonesia/5/05 - A/Ck/East Java/121/10.



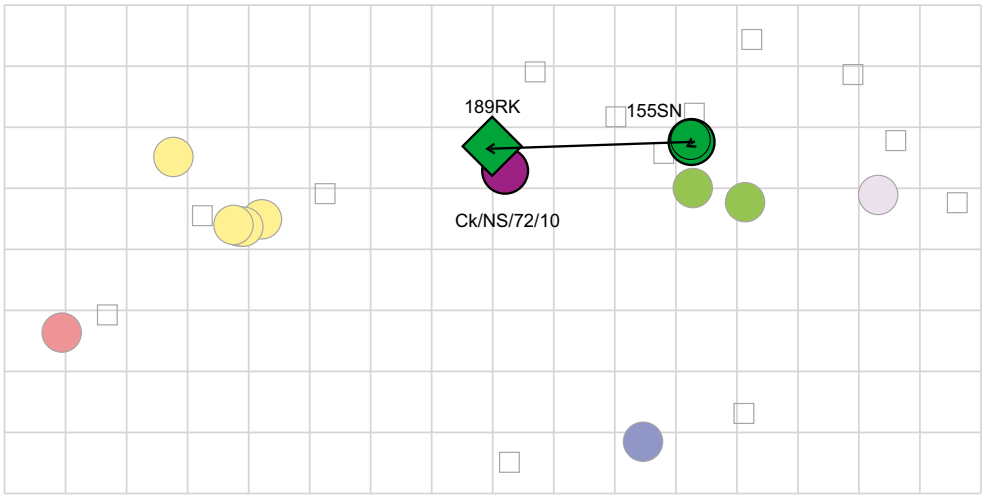
	ID/5A/05	ID/5B/05	Ck/WJ/30A/07	Ck/CJ/51A/09	Ck/CJ/051B/09	Ck/NS/72/10	Ck/WJ/119A/10	Ck/WJ/119B/10	Ck/EJ/121A/10	Ck/EJ/121B/10	Ck/SS/157A/11	Ck/SS/157B/11
ID/5/05	256	2136	<80	446	1126	158	<40	<40	<40	146	193	200
136PS	320	1920	40	480	1280	160	<10	<40	30	160	NT	NT
151IT	480	2560	80	640	1280	320	20	<40	10	80	NT	NT
159TK	640	3840	60	640	1920	320	10	<40	30	60	NT	NT
Δ129	160	1280	30	320	640	80	160	<40	10	60	NT	NT
Δ129151IT	NT	320	<10	120	240	80	320	320	<10	<10	NT	NT
Δ129136PS151IT189RM	<30	80	<10	60	<80	320	1280	320	160	320	NT	NT
Δ129151IT183DN189RM	NT	<20	<10	<10	<60	<80	2240	1600	480	640	<10	<10
Ck/WJ/119/10	<10	<120	<10	<10	<10	<20	2537	1674	177	73	<10	<10

Fig. S4C. A/Indonesia/5/05 - A/Ck/West Java/119/10



	ID/5A/05	ID/5B/05	Ck/WJ/30A/07	Ck/CJ/51A/09	Ck/CJ/051B/09	Ck/NS/72/10	Ck/EJ/121B/10	Ck/SS/157A/11	Ck/SS/157B/11
ID/5/05	256	2136	<80	446	1126	158	146	193	200
155SN	320	1810	24	480	1109	277	240	320	640
184AE	320	1810	35	960	1109	160	240	640	240
129SL	80	423	<40	200	113	24	120	453	599
Ck/SS/157/11	<80	80	<10	126	124	<20	45	1016	784

Fig. S4D. A/Indonesia/5/05 - A/Ck/South Sulawesi/157/11.



	ID/5A/05	ID/5B/05	Ck/WJ/30A/07	Ck/CJ/51A/09	Ck/CJ/051B/09	Ck/NS/72/10	Ck/EJ/121B/10	Ck/SS/157A/11	Ck/SS/157B/11
ID/5/05	256	2136	<80	446	1126	158	146	193	200
155SN	320	1810	24	480	1109	277	240	320	640
189RK	40	160	40	160	640	640	80	<10	<10
Ck/NS/72/10	47	111	<40	91	234	447	46	<10	20

Fig. S4E. A/Indonesia/5/05 - A/Ck/North Sumatra/72/10.

Table S1. Clade 2.1 viruses sequenced for the present study. Accession numbers refer to sequences in the GISAID database (http://gisaid.org) and were shared with kind permission of the government of the Republic of Indonesia. Representative viruses as defined in Figs. 1 and S1 are indicated in bold.

Strain	Accession number	URL
A/Chicken/Indonesia/2A/2003	EPI533371	http://gisaid.org/EPI/533371
A/Chicken/Indonesia/7/2003	EPI115476*	http://gisaid.org/EPI/115476
A/Chicken/Indonesia/BL/2003	EPI533370	http://gisaid.org/EPI/533370
A/Chicken/Indonesia/PA/2003	EPI533369	http://gisaid.org/EPI/533369
A/Chicken/West Java/A/2003	EPI533400	http://gisaid.org/EPI/533400
A/Chicken/West Java/028/2006	EPI533426	http://gisaid.org/EPI/533426
A/Chicken/West Java/12/2006	EPI533442	http://gisaid.org/EPI/533442
A/Chicken/West Java/8/2006	EPI533443	http://gisaid.org/EPI/533443
A/Chicken/West Java/CSLK-EC/2006	EPI533451	http://gisaid.org/EPI/533451
A/Chicken/West Java/PWT-WIJ/2006	EPI224605*	http://gisaid.org/EPI/224605
A/Chicken/Lampung/007/2007	EPI533455	http://gisaid.org/EPI/533455
A/Chicken/West Java/129/2007	EPI533420	http://gisaid.org/EPI/533420
A/Chicken/West Java/130/2007	EPI533419	http://gisaid.org/EPI/533419
A/Chicken/West Java/132/2007	EPI533418	http://gisaid.org/EPI/533418
A/Chicken/West Java/20/2007	EPI533414	http://gisaid.org/EPI/533414
A/Chicken/West Java/29-002/2007	EPI533441	http://gisaid.org/EPI/533441
A/Chicken/West Java/30/2007	EPI533406	http://gisaid.org/EPI/533406
A/Chicken/Lampung/7/2008	EPI533454	http://gisaid.org/EPI/533454
A/Chicken/West Java/028/2008	EPI533425	http://gisaid.org/EPI/533425
A/Chicken/West Java/32-031/2008	EPI533440	http://gisaid.org/EPI/533440
A/Chicken/West Java/34/2008	EPI533439	http://gisaid.org/EPI/533439
A/Chicken/West Java/6-1/2008	EPI533450	http://gisaid.org/EPI/533450
A/Chicken/West Java/X2/2008	EPI533438	http://gisaid.org/EPI/533438
A/Chicken/Banten/019/2009	EPI533412	http://gisaid.org/EPI/533412
A/Chicken/Central Java/51/2009	EPI533397	http://gisaid.org/EPI/533397
A/Chicken/Central Java/52/2009	EPI533395	http://gisaid.org/EPI/533395
A/Chicken/Central Java/053/2009	EPI533394	http://gisaid.org/EPI/533394
A/Chicken/Central Java/056/2009	EPI533387	http://gisaid.org/EPI/533387
A/Chicken/East Java/013/2009	EPI533381	http://gisaid.org/EPI/533381
A/Chicken/East Java/029/2009	EPI533383	http://gisaid.org/EPI/533383
A/Chicken/East Java/035/2009	EPI533379	http://gisaid.org/EPI/533379
A/Chicken/East Java/6/2009	EPI533380	http://gisaid.org/EPI/533380
A/Chicken/North Sumatra/27/2009	EPI533461	http://gisaid.org/EPI/533461
A/Chicken/Riau/071/2009	EPI533368	http://gisaid.org/EPI/533368
A/Chicken/South Kalimantan/070/2009	EPI533378	http://gisaid.org/EPI/533378
A/Chicken/West Java/59/2009	EPI533448	http://gisaid.org/EPI/533448
A/Chicken/West Java/015/2009	EPI533407	http://gisaid.org/EPI/533407
A/Chicken/West Java/034/2009	EPI533410	http://gisaid.org/EPI/533410
A/Chicken/West Java/47/2009	EPI533424	http://gisaid.org/EPI/533424
A/Chicken/West Java/049/2009	EPI533423	http://gisaid.org/EPI/533423
A/Chicken/West Java/068/2009	EPI533437	http://gisaid.org/EPI/533437
A/Chicken/West Java/127/2009	EPI533399	http://gisaid.org/EPI/533399
A/Chicken/West Java/128/2009	EPI533411	http://gisaid.org/EPI/533411

Table S1. (continued)

Strain	Accession number	URL
A/Chicken/West Java/2/2009	EPI533435	http://gisaid.org/EPI/533435
A/Chicken/West Java/20/2009	EPI533413	http://gisaid.org/EPI/533413
A/Chicken/West Java/59A/2009	EPI533444	http://gisaid.org/EPI/533444
A/Chicken/Central Java/59/2009	EPI533396	http://gisaid.org/EPI/533396
A/Chicken/West Java/6-2/2009	EPI533449	http://gisaid.org/EPI/533449
A/Chicken/West Java/X3/2009	EPI533436	http://gisaid.org/EPI/533436
A/Chicken/Central Java/075/2010	EPI533390	http://gisaid.org/EPI/533390
A/Chicken/Central Java/099/2010	EPI533388	http://gisaid.org/EPI/533388
A/Chicken/Central Java/123/2010	EPI533393	http://gisaid.org/EPI/533393
A/Chicken/Central Java/139/2010	EPI533392	http://gisaid.org/EPI/533392
A/Chicken/Central Java/75/2010	EPI533391	http://gisaid.org/EPI/533391
A/Chicken/Central Java/99/2010	EPI533389	http://gisaid.org/EPI/533389
A/Chicken/East Java/121/2010	EPI533384	http://gisaid.org/EPI/533384
A/Chicken/Lampung/092/2010	EPI533453	http://gisaid.org/EPI/533453
A/Chicken/North Sumatra/072/2010	EPI533459	http://gisaid.org/EPI/533459
A/Chicken/North Sumatra/72/2010	EPI533460	http://gisaid.org/EPI/533460
A/Chicken/South Sulawesi/140/2010	EPI533377	http://gisaid.org/EPI/533377
A/Chicken/West Java/081/2010	EPI533405	http://gisaid.org/EPI/533405
A/Chicken/West Java/084/2010	EPI533409	http://gisaid.org/EPI/533409
A/Chicken/West Java/090/2010	EPI533404	http://gisaid.org/EPI/533404
A/Chicken/West Java/091/2010	EPI533434	http://gisaid.org/EPI/533434
A/Chicken/West Java/093/2010	EPI533422	http://gisaid.org/EPI/533422
A/Chicken/West Java/094/2010	EPI533447	http://gisaid.org/EPI/533447
A/Chicken/West Java/097/2010	EPI533402	http://gisaid.org/EPI/533402
A/Chicken/West Java/119/2010	EPI533433	http://gisaid.org/EPI/533433
A/Chicken/West Java/120/2010	EPI533432	http://gisaid.org/EPI/533432
A/Chicken/West Java/125/2010	EPI533421	http://gisaid.org/EPI/533421
A/Chicken/West Java/125A/2010	EPI533417	http://gisaid.org/EPI/533417
A/Chicken/West Java/126/2010	EPI533446	http://gisaid.org/EPI/533446
A/Chicken/West Java/131/2010	EPI533416	http://gisaid.org/EPI/533416
A/Chicken/West Java/134/2010	EPI533431	http://gisaid.org/EPI/533431
A/Chicken/West Java/138/2010	EPI533415	http://gisaid.org/EPI/533415
A/Chicken/West Java/142/2010	EPI533398	http://gisaid.org/EPI/533398
A/Chicken/West Java/145/2010	EPI533430	http://gisaid.org/EPI/533430
A/Chicken/West Java/84/2010	EPI533408	http://gisaid.org/EPI/533408
A/Chicken/West Java/90/2010	EPI533403	http://gisaid.org/EPI/533403
A/Chicken/West Java/97/2010	EPI533401	http://gisaid.org/EPI/533401
A/Chicken/West Java/X4/2010	EPI533428	http://gisaid.org/EPI/533428
A/Chicken/Central Java/146/2011	EPI533386	http://gisaid.org/EPI/533386
A/Chicken/East Java/151/2011	EPI533382	http://gisaid.org/EPI/533382
A/Chicken/Jambi/184/2011	EPI533457	http://gisaid.org/EPI/533457
A/Chicken/Lampung/153/2011	EPI533452	http://gisaid.org/EPI/533452
A/Chicken/North Sumatra/198/2011	EPI533458	http://gisaid.org/EPI/533458
A/Chicken/South Sulawesi/156/2011	EPI533376	http://gisaid.org/EPI/533376
A/Chicken/South Sulawesi/157/2011	EPI533375	http://gisaid.org/EPI/533375

Table S1. (continued)

Strain	Accession number	URL
A/Chicken/South Sulawesi/188/2011	EPI533374	http://gisaid.org/EPI/533374
A/Chicken/South Sulawesi/195/2011	EPI533373	http://gisaid.org/EPI/533373
A/Chicken/South Sulawesi/196/2011	EPI533372	http://gisaid.org/EPI/533372
A/Chicken/South Sumatra/170/2011	EPI533456	http://gisaid.org/EPI/533456
A/Chicken/West Java/145/2011	EPI533429	http://gisaid.org/EPI/533429
A/Chicken/West Java/148/2011	EPI533445	http://gisaid.org/EPI/533445
A/Chicken/West Java/200/2011	EPI533427	http://gisaid.org/EPI/533427
A/Chicken/Yogyakarta/159/2011	EPI533385	http://gisaid.org/EPI/533385

Table S2. Ferret antisera used in this study. Antisera were prepared as described in Materials and Methods. Antisera prepared to the same virus were labeled A, B, C, etcetera, additional to the isolation number in the virus name (e.g. A/Indonesia/5A/05, A/Indonesia/5B/05). Data obtained from antisera prepared to the same virus were treated as individual antisera, i.e. HI titers from these sera were not averaged.

Antigenic variant	Clade	Antisera	
A/Indonesia/5/05	2.1.3.2	A/Indonesia/5A/05	A/Indonesia/5B/05
		A/Indonesia/5D/05	A/Chicken/Central Java/51A/09
		A/Chicken/Central Java/51B/09	
A/Chicken/West Java/30/07	2.1.3.2	A/Chicken/West Java/30A/07	A/Chicken/West Java/30B/07
A/Chicken/North Sumatra/72/10	2.1.3.3	A/Chicken/North Sumatra/72/10	
A/Chicken/West Java/119/10	2.1.3.2	A/Chicken/West Java/119A/10	A/Chicken/West Java/119B/10
A/Chicken/East Java/121/10	2.1.3.2	A/Chicken/East Java/121A/10	A/Chicken/East Java/121B/10
A/Chicken/South Sulawesi/157/11	2.1.3.1	A/Chicken/South Sulawesi/157A/11	A/Chicken/South Sulawesi/157B/11
N.A.	0	A/Hong Kong/156A/97	A/Hong Kong/156B/97
	1	A/Vietnam/1194A/04	A/Vietnam/1194B/04
	2.2	A/Turkey/Turkey/1A/2005	A/Turkey/Turkey/1B/2005
	2.3	A/Anhui/1A/05	A/Anhui/1A/05
	Classic	A/Mallard/Netherlands/3B/99	

Table S3. Amino acid changes and antigenic distances from A/Indonesia/5/05. Antigenic distance is expressed as antigenic units in the antigenic maps, where 1 unit equals a two-fold difference in the HI titer. * Number of HA1 amino acid differences from A/Indonesia/5/05. ** Not including a deletion of position 129.

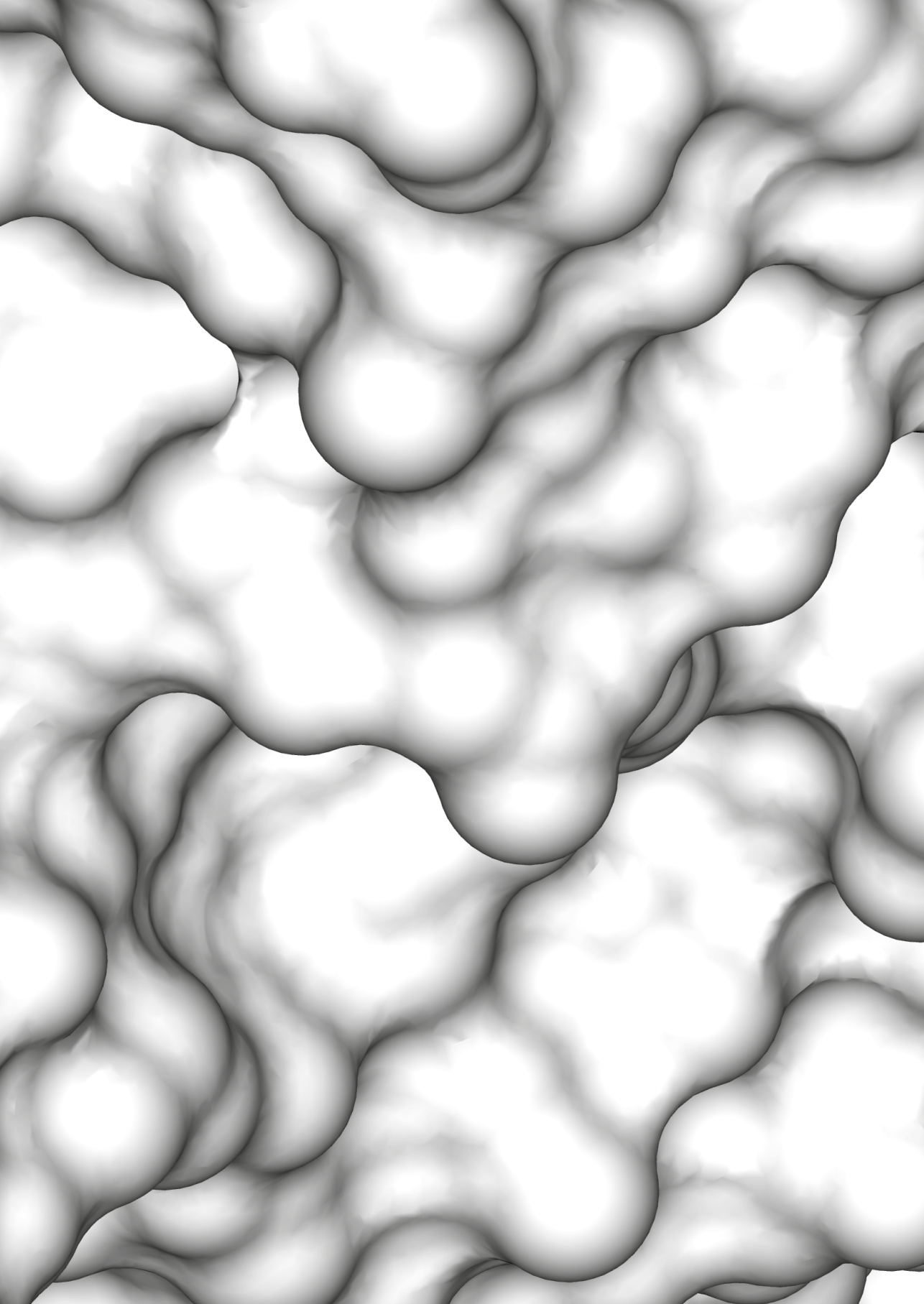
Strain	Number of substitutions from A/Indonesia/5/05*	Antigenic distance from A/Indonesia/5/05 (units)
A/Ck/North Sumatra/27/09	3	1.3
A/Ck/West Java/30/07	4	5.0
A/Ck/Central Java /51/09	5	0.8
A/Ck/Central Java/59/09	7	7.4
A/Ck/West Java/47/09	7	7.5
A/Ck/East Java/121/10	8	8.5
A/Ck/North Sumatra/72/10	10	3.0
A/Ck/Central Java/52/09	10	7.1
A/Ck/West Java/90/10	12	7.5
A/Ck/South Sulawesi/157/11	15	3.2
A/Ck/West Java/119/10	17**	10.6

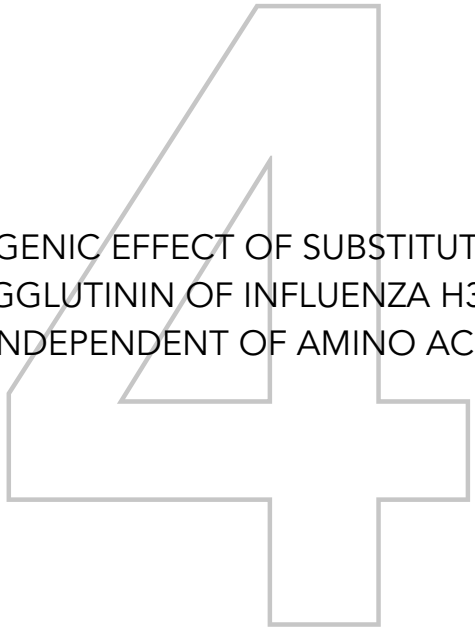
Table S4. Amino acid substitutions responsible for the antigenic differences from A/Indonesia/5/05 and antigenic distance of representative and mutant viruses from A/Indonesia/5/05. The antigenic distances are derived from the antigenic maps shown in Fig. 3. One antigenic unit equals a two-fold difference in HI titer. We considered viruses to be substantially antigenically different if the distance between them is at least two antigenic units.

		Antigenic distance (units)		
Antigenic change from A/Indonesia/5/05 to	Substitution(s)	A/Indonesia/5/05 to representative	A/Indonesia/5/05 to mutant	Mutant to representative
A/Ck/West Java/30/07	133SA 185AE	4.7	3.4	1.3
A/Ck/East Java/121/10	183DN 189RM	8.4	7.7	0.7
A/Ck/West Java/119/10	Δ129151IT 183DN 189RM	10.6	9.5	1.2
A/Ck/North Sumatra/72/10	189RK	3.0	3.2	0.5
A/Ck/South Sulawesi/157/11	129SL	3.6	2.6	1.3

Table S5. Biophysical properties of the amino acid substitutions responsible for antigenic differences from A/Indonesia/5/05. Amino acid charges are at physiological pH. [V] Substantial change in amino acid volume after substitution.

Antigenic change to	Substitutions	Biophysical change
A/Ck/West Java/30/07	133SA	hydrophilic – hydrophobic
	185AE	neutral hydrophobic – negative hydrophilic [V]
A/Ck/East Java/121/10	183DN	negative – neutral
	189RM	positive hydrophilic – neutral hydrophobic
A/Ck/West Java/119/10	S129Δ	loss of neutral hydrophilic
	151IT	hydrophobic – hydrophilic
	183DN	negative – neutral
	189RM	positive hydrophilic – neutral hydrophobic
A/Ck/North Sumatra/72/10	189RK	–
A/Ck/South Sulawesi/157/11	129SL	hydrophilic – hydrophobic





ANTIGENIC EFFECT OF SUBSTITUTIONS IN HEMAGGLUTININ OF INFLUENZA H3N2 VIRUS IS LARGELY INDEPENDENT OF AMINO ACID CONTEXT

ABSTRACT

Single amino acid substitutions at seven positions in hemagglutinin were recently shown to determine major antigenic change of influenza H3N2 virus. Here, the impact of two such substitutions was tested in eleven prototype H3 hemagglutinins to investigate context-dependence effects. The data indicate that the amino acid context in which a substitution occurs may determine the magnitude of the antigenic effect, but not the ability of the substitution to cause antigenic change.

Björn F. Koel,[1] David F. Burke,[2,3] Stefan van der Vliet,[1] Theo M. Bestebroer,[1]
Guus F. Rimmelzwaan,[1] Albert D.M.E. Osterhaus,[1]
Derek J. Smith,[2,3,1] Ron A.M. Fouchier[1]

[1]Department of Viroscience, Erasmus MC, Rotterdam, The Netherlands,
[2]Center for Pathogen Evolution, Department of Zoology, University of Cambridge, Cambridge, United Kingdom,
[3]WHO Collaborating Centre for Modeling Evolution and Control of Emerging Infectious Diseases, University of Cambridge, Cambridge, United Kingdom.

Submitted

Influenza viruses of the H3N2 subtype have been circulating in humans since 1968 and are a major cause of annual epidemics. Antibodies against the hemagglutinin (HA) surface glycoprotein can neutralize the virus and are a critical component of our immune defense against influenza viruses (40). However, the HA changes over time to escape from such neutralizing antibodies present in the human population. Smith et al. used hemagglutination inhibition (HI) assay data spanning a 35-year period to map H3N2 virus antigenic evolution (31). During this period, 11 antigenic clusters that comprise viruses of high antigenic similarity emerged, each of which was consecutively replaced by a new cluster of antigenically distinct viruses. Antigenic cluster transitions, the major antigenic changes between clusters, were subsequently shown to be predominantly caused by single amino acid substitutions on seven key positions adjacent to the receptor binding site (RBS), and most positions were involved in cluster transitions multiple times ((77), and Fig. 1A and 1B). An important question that remained is whether these single substitutions could have comparable impact on antigenic properties independent of the HA context.

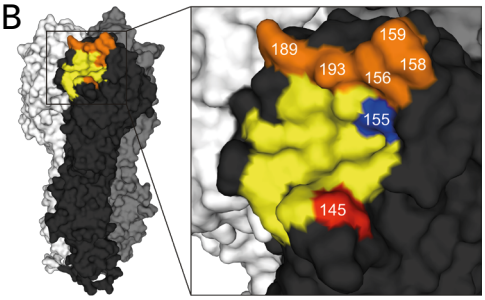
Here we investigated if the amino acid context in which a substitution occurs determines its antigenic effect. Cluster-transition substitutions 155TY and 145NK were introduced in the HA genes of viruses representing the 11 antigenic clusters (Fig. 1C). Both substitutions substantially change the biophysical properties of the amino acids on these positions and were, as single substitutions, responsible for antigenic cluster transitions during H3N2 virus antigenic evolution (Fig. 1A). Substitution 155TY was responsible for the first antigenic cluster transition of H3N2 virus (77). Substitution 145NK caused an antigenic cluster transition after 21 years of H3N2 virus evolution and was responsible for another cluster transition six years later (31, 77). The substitutions were introduced as single mutations into the 11 HA genes. Depending on the amino acid at position 155 or 145 we introduced either 155Y or T, or 145K or N in the HA genes (Fig. 1C). When representative viruses had 145S we introduced 145K, but not 145N as this results only in a small biophysical change. For viruses with 155H two modified HA genes, containing 155T or 155Y, were constructed. All substitutions resulted in substantial changes in the biophysical properties of the amino acids. Plasmids containing wildtype or modified HA genes were used to generate recombinant viruses consisting of the (modified) HA gene and remaining genes of A/Puerto Rico/8/34 by reverse genetics (64). Subsequently, the antigenic properties of recombinant viruses were analyzed in HI assays using a previously defined panel of ferret antisera (37, 77).

Most mutant HAs with substitutions at position 155 were substantially antigenically different from the respective wildtype HAs, with up to 64-fold differences in HI titers (Fig. 2A). A subset of the antisera yielded similar HI titers (i.e. < 4-fold different) to viruses with wildtype and mutant HA. In the context of some HAs (e.g. SI87, BE89, BE92), substitutions at position 155 had a modest effect, with up to 4-fold differences in HI titers. For the BE92 mutants, only a single antiserum showed 4-fold different HI titers compared to the wildtype virus. Thus, although the magnitude of the antigenic

A

Antigenic cluster transition	Cluster-transition substitution
HK68 - EN72	155TY
EN72 - VI75	189QK
VI75 - TX77	158GE 193DN
TX77 - BK79	156KE
BK79 - SI87	155YH 159SY 189KR
SI87 - BE89	145NK
SI87 - BE92	156EK
BE92 - WU95	145NK
WU95 - SY97	156KQ 158EK
SY97 - FU02	156QH

B



C

Antigenic Cluster	Cluster Representative	Substitution	
		145	155
HK68	BI/16190/68	S K	T Y
EN72	BI/21793/72	S K	Y T
VI75	BI/1761/76	N K	Y T
TX77	BI/2271/76	N K	Y T
BK79	NL/233/82	N K	Y T
SI87	NL/620/89	N K	H Y
			H T
BE89	NL/823/92	K N	H Y
			H T
BE92	NL/179/93	N K	H Y
			H T
WU95	NL/178/95	K N	H Y
			H T
SY97	NL/427/98	K N	H Y
			H T
FU02	NL/213/03	K N	T Y

Fig. 1. Experimental background and viruses used in this study. (A) Substitutions responsible for antigenic cluster transitions as defined in (77). Letters and digits in antigenic cluster names refer to the location and year of isolation of the first vaccine strain in that cluster (HK, Hong Kong; EN, England; VI, Victoria; TX, Texas; BK, Bangkok; SI, Sichuan; BE, Beijing; WU, Wuhan; SY, Sydney; FU, Fujian). (B) Amino acid positions responsible for major antigenic change during H3N2 virus antigenic evolution plotted on an A/Aichi/2/68 HA trimer (PDB accession code 5HMG). Monomers are shown in black, grey, and white, the RBS in yellow. Amino acid positions 145 and 155 are indicated in red and blue, while the remaining key positions are indicated in orange. (C) Mutants constructed for this study. Cluster representative viruses had the amino acid consensus sequence of all viruses in that cluster (described in (77)). BI: Bilthoven, NL; The Netherlands.

change caused by the substitution of T/Y at position 155 was variable, this substitution caused a large antigenic change in the majority of HA contexts that were tested.

Viruses with a substitution on position 145 of HA were each antigenically distinct from the respective wildtype viruses (Fig. 2B), with HI titer differences up to 128-fold. Also here, a subset of the antisera yielded similar HI titers to viruses with wildtype and mutant HA. The antigenic change caused by the substitution of K/N at position

145 in HA was even less affected by the HA context as compared to substitution T/Y at position 155.

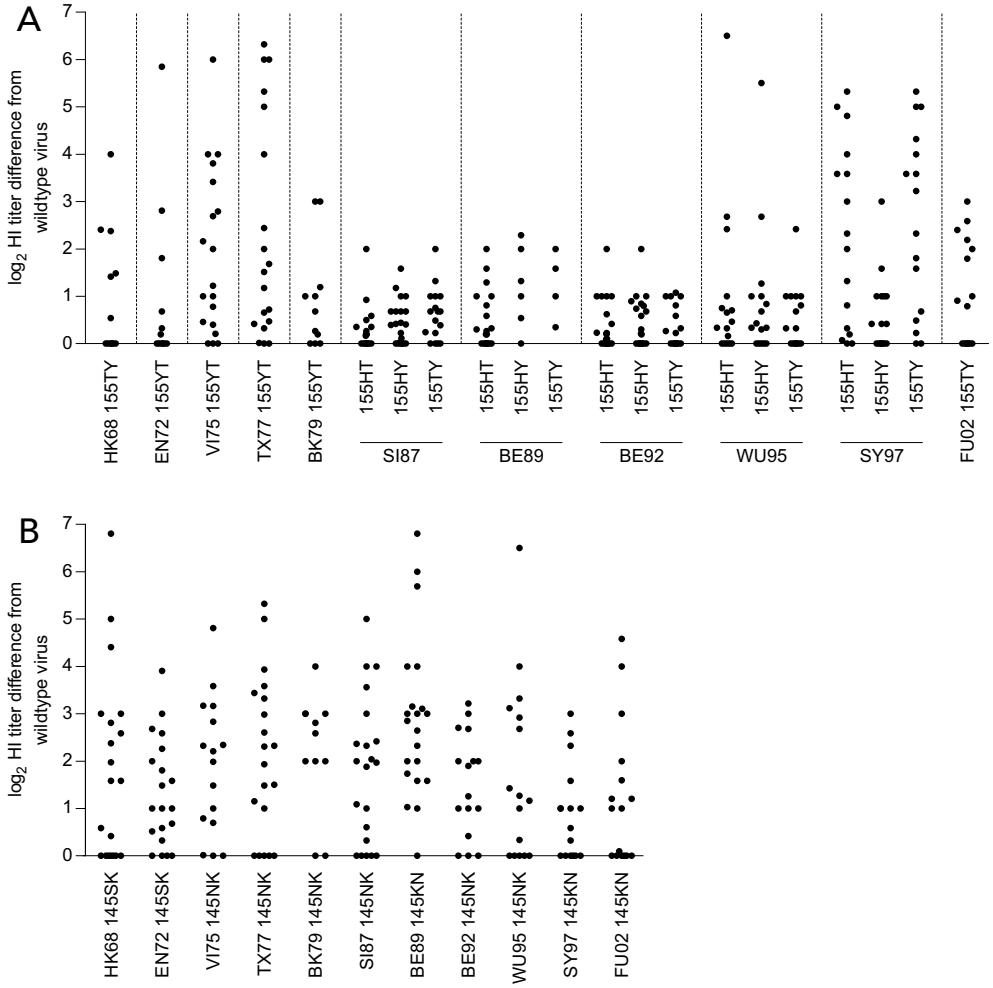


Fig. 2. HI titer differences between viruses with wildtype and mutant HAs. (A) Each symbol represents the log2 HI titer difference for an individual antiserum between a representative virus and a mutant with 155TY or 155YT, or between mutants with 155HT and 155HY (indicated as 155TY for SI87, BE89, BE92, WU95, and SY97). (B) Log2 HI titer differences between viruses with wildtype HA and 145K or 145N mutants. Symbols as in panel (A).

Substitution 145NK was responsible for the cluster transition from SI87 to BE89 (Fig. 1A). When 145K was introduced in viruses representing the antigenic clusters that circulated prior to the SI87 cluster, this caused similar escape from inhibition by antisera to contemporary or previously circulating strains as 145K in the SI87 representative virus (Fig. 2B). We next tested if 145K affected the antigenic properties

to similar levels as the cluster-transition substitutions that were taken naturally before 1989 (Fig. 1A). In this analysis only antisera to strains from the same or previous antigenic clusters as the representative virus were included, thus testing escape from antibodies induced to previously circulating strains. Antigenic differences caused by 145K were similar to those caused by the substitutions that were responsible for the "natural" antigenic cluster transitions (Fig. 3). Thus, if viruses with 145K had appeared before the BE89 antigenic cluster, they would have been sufficiently antigenically different from earlier H3N2 viruses to provide antibody escape. These results suggest that antibody escape only partially determines the evolutionary success of an emerging antigenic variant.

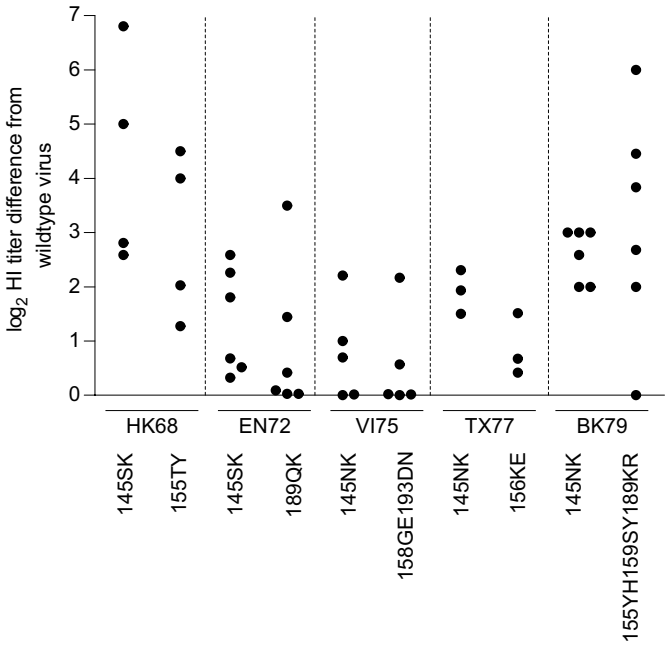


Fig. 3. HI titer differences between cluster representatives, 145K mutants, and cluster-transition mutants. Each symbol represents the log2 HI titer difference for an individual antiserum between viruses with wildtype and 145K mutant HA or between the wildtype and cluster-transition mutant virus. Only antisera to strains from the same or preceding antigenic clusters as the representative virus were included in the analyses.

The differences in magnitude of the antigenic effect of 155T and 155Y substitutions versus the context independent antigenic change caused by the 145N and 145K substitutions in different amino acid backgrounds may be due to differences in local HA structure. Position 155 is located in the depression between the 190-helix that contains conserved position 195Y and a loop that contains conserved position 153W (Fig. 4), which are fundamental components of the RBS (24, 33). Although adjacent to

RBS positions 135 - 137, position 145 is located on a protruding loop that may have fewer structural constraints. The substitutions introduced at position 145 therefore possibly have a larger impact on the local HA structure than the substitutions introduced at position 155, resulting in the more pronounced antigenic changes of the mutants with a substitution at position 145 observed here.

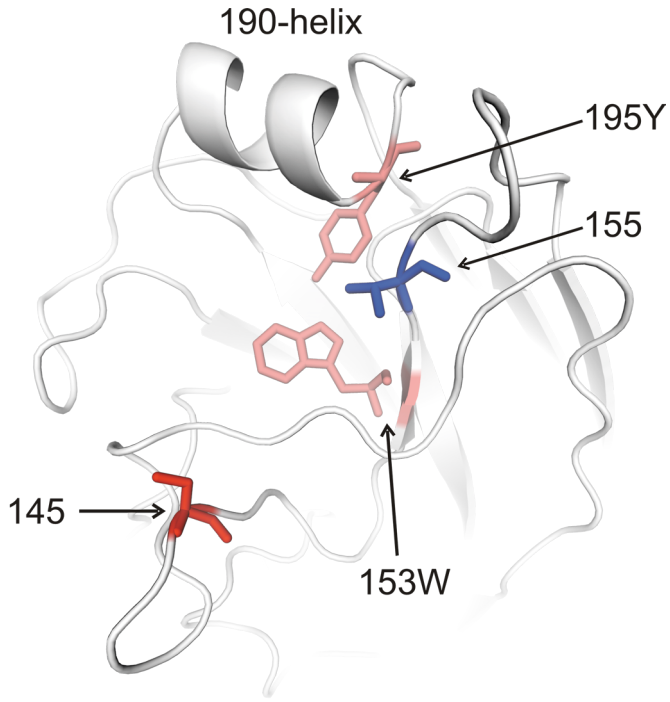


Fig. 4. Cartoon representation of the A/Aichi/2/68 RBS area. Positions 155 and 145 are indicated in blue and red, respectively. Positions 195Y and 153W, which are conserved among influenza A virus subtypes (24, 33), are indicated in pink.

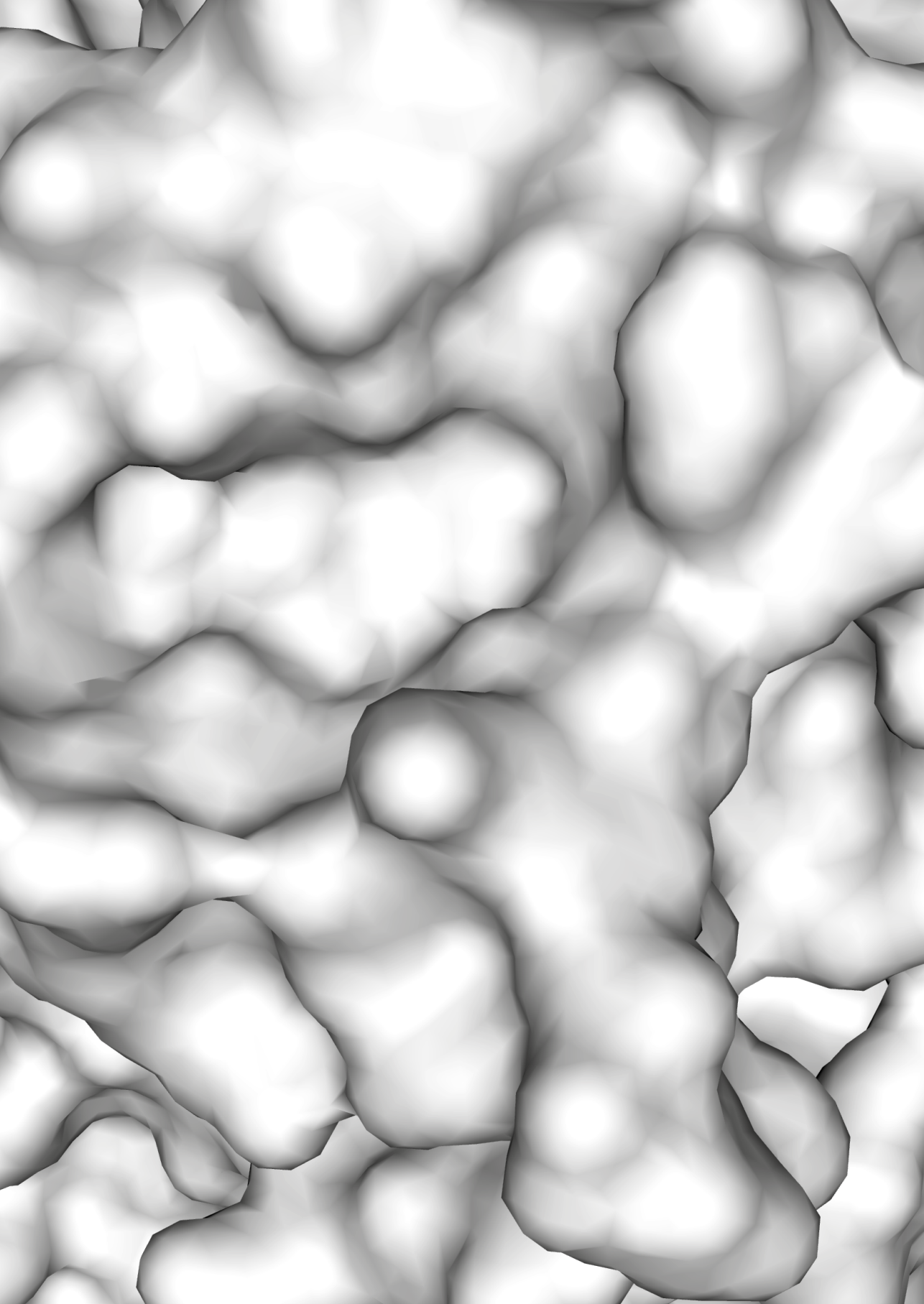
In summary, the ability of cluster-transition substitutions to cause antibody escape was found to be largely independent of the amino acid context, although the magnitude of antigenic change may be context-dependent. In agreement with the observed repeated substitution at the 7 key antigenic positions during influenza H3N2 virus antigenic evolution (77), these results emphasize the potential importance of the key positions for future antigenic change.

ACKNOWLEDGEMENTS

This work was supported by a ZonMW VICI grant and NIH contracts HHSN266200700010C and HHSN272201400008C, NIH Director's Pioneer Award DP1-OD000490-01, European Union FP7 program EMPERIE (223498), European

Union FP7 program ANTIGONE (278976), and program grant P0050/2008 from the Human Frontier Science Program.

The authors gratefully thank J.C. de Jong, B.M. Laksono, N.S. Lewis, and C.A. Russell for stimulating discussions and technical assistance.



5

IDENTIFICATION OF AMINO ACID SUBSTITUTIONS SUPPORTING ANTIGENIC CHANGE OF INFLUENZA A(H1N1)PDM09 VIRUSES

ABSTRACT

The majority of currently circulating influenza A(H1N1) viruses are antigenically similar to the virus that caused the 2009 influenza pandemic. However, antigenic variants are expected to emerge as population immunity increases. Amino acid substitutions in the hemagglutinin protein can result in escape from neutralizing antibodies, affect viral fitness, and change receptor preference. Here we constructed mutants with substitutions in the hemagglutinin of A/Netherlands/602/09 in an attenuated backbone to explore amino acid changes that may contribute to emergence of antigenic variants in the human population. Our analysis revealed that single substitutions affecting the 151 – 159 loop located adjacent to the receptor binding site caused escape from ferret and human antibodies elicited after primary A(H1N1)pdm09 virus infection. The majority of these substitutions resulted in similar or increased replication efficiency *in vitro* compared to the virus carrying the wildtype hemagglutinin, and did not result in a change of receptor preference. However, none of the substitutions was sufficient to escape from the antibodies in sera from individuals that experienced both seasonal and pandemic A(H1N1) virus infections. These results suggest that antibodies directed against epitopes on seasonal A(H1N1) viruses contribute to neutralization of A(H1N1) pdm09 antigenic variants, thereby limiting the number of possible substitutions that could lead to escape from population immunity.

Björn F. Koel,[1] Ramona Mögling,[1,2] Salin Chutinimitkul,[1] Pieter L. Fraaij,[1,3] David F. Burke,[2,4] Stefan van der Vliet,[1] Emmie de Wit,[1] Theo M. Bestebroer,[1] Guus F. Rimmelzwaan,[1] Albert D. M. E. Osterhaus,[1] Derek J. Smith,[1,2,4] Ron A. M. Fouchier,[1] Miranda de Graaf[1,2]

[1]Department of Viroscience, Erasmus MC, Rotterdam, The Netherlands,
[2]Center for Pathogen Evolution, Department of Zoology, University of Cambridge, Cambridge, United Kingdom,
[3]Department of Pediatrics, Erasmus MC-Sophia, Rotterdam, The Netherlands,
[4]WHO Collaborating Centre for Modeling Evolution and Control of Emerging Infectious Diseases, University of Cambridge, Cambridge, United Kingdom

J Virol. 2015; 89(7): 3763-75

IMPORTANCE

Influenza A viruses can cause significant morbidity and mortality in humans. Amino acid substitutions in the hemagglutinin protein can result in escape from antibody-mediated neutralization. This allows the virus to re-infect individuals that have acquired immunity to previously circulating strains through infection or vaccination. To date, the vast majority of A(H1N1)pdm09 strains remain antigenically similar to the virus that caused the 2009 influenza pandemic. However, antigenic variants are expected to emerge as a result of increasing population immunity. We show that single amino acid substitutions near the receptor binding site were sufficient to escape from antibodies specific for A(H1N1)pdm09 viruses, but not from antibodies elicited in response to infections with seasonal A(H1N1) and A(H1N1)pdm09 viruses. This study identifies substitutions in A(H1N1)pdm09 viruses that support escape from population immunity, but also suggests that the number of potential escape variants is limited by previous exposure to seasonal A(H1N1) viruses.

INTRODUCTION

Influenza pandemics occur when a novel influenza A virus is introduced in the human population and spreads around the globe. Since existing antibody responses are typically not cross-reactive with the antigenically novel virus, the virus encounters little pre-existing humoral immunity and can cause severe outbreaks. Three influenza pandemics occurred during the 20th century: A(H1N1) virus in 1918, A(H2N2) virus in 1957, and A(H3N2) virus in 1968 (89). In each case the newly introduced subtype replaced the previous subtype. In 1977 an A(H1N1) virus that caused epidemics in the early 1950s was reintroduced in the human population (90), and continued to co-circulate with A(H3N2) until 2009. In April 2009 a swine-origin A(H1N1) virus A(H1N1) pdm09) caused the first influenza A virus pandemic of the 21th century (91). It replaced the previously circulating seasonal A(H1N1) virus, but continues to co-circulate with seasonal A(H3N2) virus (92).

A prerequisite for the influenza virus to infect the host cell is the binding of the hemagglutinin (HA) surface protein to sialylated glycan receptors on the host cell through its receptor-binding site (RBS). HA is the main target of neutralizing antibodies and is therefore a critical component of influenza vaccines (40). Influenza viruses continually escape antibody-mediated neutralization by variation of the amino acids in the HA protein. This process is referred to as antigenic drift, and allows the virus to infect individuals that are immune to contemporary or previously circulating antigenic variants.

Studies from the 1980s identified four immunodominant antigenic regions within the HA of A(H1N1) virus (34, 93). Similar antigenic regions were identified for A(H3N2) (33) and A(H5N1) viruses (35, 36). Amino acid substitutions in these so called antigenic sites, which cover much of the HA globular head, can result in escape from antibody recognition. More recently it was shown that major antigenic change during

evolution of A(H3N2) and A(H5N1) viruses and recent antigenic change of seasonal A(H1N1) and influenza B viruses were predominantly caused by single substitutions that occurred near the RBS (77, 94).

Antigenic change may also be a secondary effect of substitutions in HA that facilitate more efficient replication in the human host. HA is pivotal in adaptation of zoonotic influenza A viruses to a new host because of its function in receptor binding (95). Human influenza viruses bind to sialic acids (SAs) linked to the galactose in an α2,6-linkage, avian influenza viruses have a preference for α2,3-linked SAs, while swine viruses bind either α2,6 or both α2,3 and α2,6 linked SAs (27). We hypothesized that substitutions that modify or fine-tune receptor specificity, thereby altering host range and tissue tropism, may result in escape from antibodies directed at the RBS area. Finally, addition or removal of carbohydrate side chains on HA has been associated with changes in the antigenic properties of influenza viruses (61, 62).

The HA of A(H1N1)pdm09 viruses is antigenically most similar to that of recent classical and triple reassortant swine A(H1N1) viruses (96, 97). The HA of these swine viruses descended from the 1918 pandemic influenza virus but in contrast to the human lineage did not undergo extensive antigenic drift, as was reported for A(H3N2) swine viruses (30). Structural analyses suggested high antigenic similarity between the A(H1N1)pdm09 virus and A(H1N1) viruses that circulated in the first decades after the 1918 pandemic (98, 99). Accordingly, age groups that experienced A(H1N1) virus infection before 1950 were partially immune to the A(H1N1)pdm09 virus (100).

During the five years after its emergence multiple antigenic variants of A(H1N1) pdm09 virus have been detected, but the vast majority of recently isolated viruses remain antigenically similar to the A/California/7/2009 vaccine virus (92). However, as population immunity to A(H1N1)pdm09 virus builds up it becomes beneficial for the virus to be antigenically different from the pandemic strain. Therefore, the goal of the present study was to explore molecular changes that contribute to antibody escape of A(H1N1)pdm09 virus.

We selected 25 single and five double substitutions based on substitutions that were shown to be important for antigenic change in other subtypes, changes in receptor specificity, or genetic differences between A(H1N1)pdm09, swine A(H1N1) and seasonal A(H1N1) viruses. The substitutions were introduced into the HA gene of influenza virus A/Netherlands/602/09 in an attenuated virus backbone, and their antigenic effect was tested in hemagglutination inhibition (HI) assays using a ferret antiserum prepared to A/Netherlands/602/09. Mutants that displayed altered antigenic properties were further tested to a larger panel of ferret antisera and human sera. In addition, the impact of these substitutions on replication kinetics and receptor specificity was evaluated.

MATERIALS AND METHODS

Cells

293T cells were cultured in DMEM (Lonza, Breda, The Netherlands) supplemented with 10% FCS (Sigma-Aldrich, Zwijndrecht, The Netherlands), 100 IU/ml penicillin, 100 µg/ml streptomycin, 2 mM glutamine, 1mM sodiumpyruvate, non-essential amino acids (Lonza, Breda, The Netherlands), and 500 µg/mL geneticin (Life Technologies, Bleiswijk, The Netherlands). Madin-Darby Canine Kidney (MDCK) cells were cultured in EMEM (Lonza) supplemented with 10% FCS, 100 IU/ml penicillin, 100 µg/ml streptomycin, 2mM glutamine, 1.5 mg/ml sodiumbicarbonate (Lonza), 10mM Hepes (Lonza), and non-essential amino acids.

Plasmid construction

A/Netherlands/602/09 was isolated from a patient in the Netherlands during the 2009 influenza pandemic (101), and was used in this study to represent the antigenic properties of A(H1N1)pdm09 viruses. The full HA gene was cloned in the modified pHW2000 expression plasmid as previously described (64, 65, 102). Mutations were introduced with the minimal number of nucleotide substitutions necessary to change the amino acid. When more than one single nucleotide change could lead to the desired substitution, the codon change observed in naturally occurring amino acid substitutions was selected (e.g. genetic differences between human A(H1N1) viruses). If the desired amino acid substitution did not occur previously in A(H1N1) viruses, the mutation was introduced using a codon observed in A(H3N2) viruses. Mutations were introduced in the HA gene using the QuickChange multi-site directed mutagenesis kit (Agilent Technologies, Amstelveen, The Netherlands) according to the manufacturers instructions. The presence of introduced mutations and absence of undesired additional mutations was confirmed by sequence analysis of the modified HA gene.

Construction of recombinant virus stocks

Plasmids containing wildtype or modified A/Netherlands/602/09 HA genes were used to generate recombinant viruses consisting of the (modified) HA gene and seven remaining gene segments of A/Puerto Rico/8/34, as described previously (64). Briefly, 293T cells were seeded in 100mm dishes one day prior to transfection. Cells were transfected overnight with 40 µg of plasmid DNA. Transfection medium was subsequently replaced with medium containing 2% FCS. Cells were incubated for 72 hours at 37°C and 5% CO_2 before harvesting the supernatant. The virus stocks were propagated by inoculation of MDCK cells with 2 mL of supernatant of transfected cells. After 2 hours the inoculum was replaced by MDCK infection medium consisting

of EMEM, 100 IU/ml penicillin, 100 µg/ml streptomycin, 2mM glutamine, 1.5 mg/ml sodium bicarbonate, 10mM Hepes, non-essential amino acids, and 25 µg/ml trypsin. Viruses were harvested after incubation for 72 hours at 37°C and 5% CO_2. The culture supernatants were subjected to ultracentrifugation to increase the viral particle concentration if hemagglutination titers were below 12 hemagglutinating units (HAU). The presence of introduced mutations and absence of unwanted additional mutations was confirmed by sequencing of the HA gene. Work with recombinant viruses was performed under biosafety level 2 conditions under a permit from the Ministry of Infrastructure and the Environment.

Antisera

Ferret antisera were prepared by intranasal inoculation with 500 µL virus stock. Antisera were collected fourteen days after inoculation. Ferrets were housed and experiments were conducted in strict compliance with European guidelines (EU directive on animal testing 86/609/EEC) and Dutch legislation (Experiments on Animal Act, 1997). The protocol was approved by an independent animal experimentation ethical review committee 'stichting DEC consult'. Animal welfare was monitored daily and all animal handling was performed under light anesthesia (ketamine) to minimize animal discomfort.

Human serum samples were selected from the serum bank of the Viroscience department. Sera from patients in whom the possibility of non-naturally obtained antibody responses to A(H1N1)pdm09 virus (e.g. vaccination, intravenous immunoglobulin administration) existed or sera from patients with immune deficiencies (e.g. auto-immune disease, HIV-positive status, use of immunosuppressive medication) were excluded from use in this study. In addition, only sera of patients who, or of whom the caregivers, did not object to scientific use of leftover materials were included in this study. The study protocol was reviewed and approved by the medical ethics board of the Erasmus University Medical Center (study number MEC-2012-181). Informed consent was waived because patient inclusion was performed retrospectively and anonymously.

HI assay

HI assays were performed using standard procedures (37). Briefly, ferret antisera or human sera were pre-treated overnight with the receptor destroying enzyme Vibrio cholerae neuraminidase (VCNA) at 37°C followed by inactivation for 1 hour at 56°C. 2-fold serial dilutions of the pre-treated sera, starting at a 1:20 dilution, were mixed with 25 µL virus stock containing 4 HAU and the mixture was incubated at 37°C for 30 minutes. Subsequently, 25 µL 1% turkey erythrocytes was added and hemagglutination

patterns were read after a 1 hour incubation at 4°C. The HI titer is expressed as the reciprocal value of the highest serum dilution that completely inhibited agglutination of turkey erythrocytes.

Plaque assay

The assay was performed as described (67). In brief, MDCK cells were seeded in a 6 well plate to reach 90% confluency the following day. One hour after inoculation with virus, the inoculum was replaced with a 1:1 mixture of 2.4% Avicel (FMC biopolymers, Brussels, Belgium) with 2xEMEM infection medium. After 36 hours, cells were washed with PBS and incubated with 80% Aceton for at least 30 minutes at -20°C. Fixed and permeabilized cells were washed 3 times with PBS, and incubated for 1 hour at 37°C with mouse-anti-NP monoclonal antibody (1 mg/ml HB65, ATCC). Following three washes with PBS, cells were incubated for 1 hour at 37°C with rabbit-anti-mouse-FITC (Life Technologies, Bleiswijk, The Netherlands). Cells were washed with PBS and allowed to air dry. Plaques were scanned on a Typhoon 9410 variable mode imager (GE Healthcare, Diegem, Belgium). This data was analysed with ImageQuant TL Colony counter & Image feature measurement (Amersham Biosciences, Freiburg, Germany). Plaque size was plotted as the radius of the plaques.

Modified turkey red blood cell hemagglutination assay

The modified turkey red blood cell (TRBC) assay was performed as described previously (103) with modifications. Briefly, SAs were removed from the turkey erythrocyte surface by incubating 20% turkey erythrocytes in a total volume of 62,5 µL phosphate-buffered saline supplemented with 50mU of VCNA (Roche, Almere, The Netherlands) in 8 mM calcium chloride at 37°C for 1 hour. The removal of SAs was confirmed by absence of hemagglutination of the treated turkey erythrocytes by control viruses. Turkey erythrocytes stripped of SAs were resialylated to contain only α2,3- or α2,6-linked SAs using either 0.5 mU α2,3-(N)-sialyltransferase (Calbiochem, CA) or 2 mU of α2,6-(N)- sialyltransferase (Japan Tobacco, Inc., Shizuoka, Japan), and 1.5 mM cytidine monophospho-N-acetylneuraminic (CMP) sialic acid (Sigma-Aldrich, Zwijndrecht, Netherlands) at 37°C for 2 hours in a total volume of 75 µl. Subsequently, resialylated turkey erythrocytes were washed with PBS and were resuspended to a final concentration of 0.5% in PBS containing 1% bovine serum albumin. Standard hemagglutination assays were performed to confirm correct resialylation using control viruses with known receptor specificities for α2,3- and α2,6-linked SAs. The receptor specificity of mutant viruses was tested by conventional hemagglutination assay with the modified TRBCs using cell culture supernatants.

Structural analysis

Amino acid positions were plotted on the HA crystal structure of the A/California/04/09 virus [PDB accession code 3LZG (42)] using MacPyMOL (The PyMOL Molecular Graphics System, Version 1.3, Schrödinger, LLC) to visualize the trimer.

RESULTS

Selection of substitutions

The antigenic regions of seasonal A(H1N1) virus have previously been mapped (34, 93), but the substitutions that caused antigenic change during A(H1N1) virus evolution remain almost entirely unknown. Additionally, fine-grain differences in receptor specificity beyond the coarse distinction for α2,3-SA or α2,6-SA preference that allow more efficient replication in the human host and that may change the virus antigenically are currently incompletely understood (95, 104). Because of the limited insight in the molecular basis for previous antigenic change of A(H1N1) viruses we selected substitutions to introduce into the A/Netherlands/602/09 representative virus HA gene based on the following three approaches.

First, we aligned the HA sequences of 28 viruses representing pre-1957 and post-1977 seasonal A(H1N1) viruses, swine A(H1N1) viruses and A(H1N1)pdm09 viruses, and selected substitutions based on genetic differences between these viruses (Table 1). We hypothesized that part of the substitutions that became fixed during A(H1N1) virus evolution had a selective advantage because of their role in immune escape. In addition, amino acid differences between swine or A(H1N1) pdm09 viruses and pre-2009 human A(H1N1) viruses may contribute to adaptation to the human host and potentially change the virus antigenically as a secondary effect. Previous work indicated a critical role for antibodies targeting the RBS area in virus neutralization (77, 94) and we therefore focused primarily on amino acid differences in this region of the HA. Substitution 84SN is located away from the RBS, but was consistently different between A(H1N1)pdm09 viruses and all swine and human A(H1N1) viruses (H1 numbering is used throughout, unless indicated otherwise).

A second approach was based on information regarding the molecular basis for antigenic change during evolution of other influenza virus subtypes. Substitutions at seven positions (145, 155, 156, 158, 159, 189, and 193, H3 numbering) were entirely responsible for the major antigenic changes during evolution of A(H3N2) virus (77), three of which (151, 185, and 189, H5 numbering) were also identified as key positions for antigenic change of A(H5N1) clade 2.1 virus (94). The corresponding A(H1N1)pdm09 virus positions were identified and where possible the substitution responsible for the observed antigenic change was introduced (Table 1). Multiple mutants were made if it was not possible to introduce the desired substitution. For example, 156QH was responsible for a major antigenic change during A(H3N2) virus evolution and we therefore generated mutants with 153KQ and 153KH.

Third, we selected substitutions shown or predicted to affect receptor binding specificity of influenza viruses of the A(H1N1), A(H3N2) or A(H5N1) subtypes (Table 1). A(H1N1)pdm09 viruses containing the substitutions 222DE and 222DG were previously tested for their effects on receptor binding and antigenic properties by Chutinimitkul *et al.* (102). That work additionally suggested that amino acid differences at positions 142, 187, 222, and 224 were responsible for differences in receptor binding between a 1918 A(H1N1) virus and A/Netherlands/602/09 (26, 102). Mutants with these substitutions, or combinations thereof, were included in the current more extensive analyses. Substitutions 222QL and 224GS (223QL and 225GS in H1 numbering) were previously shown to affect the receptor specificity of A(H5N1) virus (105).

Recombinant viruses with the modified A/Netherlands/602/09 HA gene and the seven remaining gene segments of A/Puerto Rico/8/34 were rescued by reverse genetics. We were unable to rescue the 127DN and 186AK mutants. The mutant with 187DE did not agglutinate turkey erythrocytes even upon concentration of the virus and was therefore omitted from the antigenic analyses. After MDCK passaging, we found addition of substitution 153KE to mutant 222DG224EA and this mutant was therefore excluded from this study. Figure 1A indicates the positions of introduced substitutions on an A/California/04/09 HA crystal structure.

Analysis of antigenic properties using ferret antisera

HI assays with a ferret antiserum prepared to A/Netherlands/602/09 were performed to test if the mutant viruses could escape recognition by antibodies against the wildtype virus. When comparing HI titers obtained to an individual antiserum, viruses were considered substantially antigenically different from the reference virus if the HI titer was at least 4-fold (2 $\log_2$) lower. Ten of the 27 mutants were substantially antigenically different from A/Netherlands/602/09 (Fig. 2): 224EA, 127DT, 155GE, 156ND, 156NG, 153KE, 156NY, 156NS, 152VT156NS, and 155GE224EA.

Two groups of mutants were further tested in HI assays to a panel of ferret A(H1N1)pdm09 antisera. One group contained the ten viruses that were substantially antigenically different from A/Netherlands/602/09, the second group contained seven mutants; 84SN, 152VT, 190SN, 222DE, 222DG, 223QL, and 223QR that were antigenically similar to A/Netherlands/602/09 (Fig. 1B and 2). These groups are hereafter referred to as escape mutants and non-escape mutants, respectively. The panel of 14 ferret antisera was prepared to seven A(H1N1)pdm09 viruses isolated between 2009 and 2011. Of these, six wildtype viruses were antigenically similar to A/Netherlands/602/09, in agreement with the fact that these viruses did not have any of the substitutions present in the mutants. However, A/Netherlands/219/11 contains the 155GE substitution and had a more than 8-fold (3 $\log_2$) lower mean HI titer (data not shown).

The mean HI titers of the non-escape mutants and 224EA were less than 2-fold lower than that of A/Netherlands/602/09 to the panel of ferret antisera (Fig. 3). Escape mutants 156ND, 127DT, 155GE, and 156NG had 3 to 4-fold lower mean

Table 1. Rationale of selected substitutions.

		Genetic difference between[a]			Position involved in antigenic change of other influenza viruses[b]			
	Subtype number conversion	A(H1N1)pdm09 viruses – swine A(H1N1) viruses	A(H1N1)pdm09 viruses – seasonal A(H1N1) viruses	Swine A(H1N1) viruses – seasonal A(H1N1) viruses	A(H3N2) virus	A(H5N1) virus	Influenza B virus	Substitutions associated with changes in receptor binding specificity
84SN		✓	✓					
127DE	126[c]	✓				✓[f]		
127DN	126[c]	✓				✓[f]		
127DT	126[c]		✓			✓[f]		
142KN	145[d]				(77)			
142KS	145[d]		✓		(77)			
152VT	151[c], 155[d]			✓	(77)	(94)		
153KE	156[d]			✓	(77)			
153KH	156[d]				(77)			
153KQ	156[d]				(77)			
155GE	158[d], 165[e]				(77)		(77)	
156ND	159[d]				(77)			
156NG	159[d]		✓	✓	(77)			
156NS	159[d]			✓	(77)			
156NY	159[d]				(77)			
186AK	185[c], 189[d]				(77)	(94)		
186AQ	185[c], 189[d]				(77)	(94)		
190SD	189[c], 193[d]				(77)	(94)		
190SN	189[c], 193[d]				(77)	(94)		
222DE								(102)
222DG								(102)
223QL	222[c]							(105)
223QR	222[c]							(108)

Table 1. (continued)

	Subtype number conversion	Genetic difference between[a]			Position involved in antigenic change of other influenza viruses[b]			Substitutions associated with changes in receptor binding specificity
		A(H1N1)pdm09 viruses – swine A(H1N1) viruses	A(H1N1)pdm09 viruses – seasonal A(H1N1) viruses	Swine A(H1N1) viruses – seasonal A(H1N1) viruses	A(H3N2) virus	A(H5N1) virus	Influenza B virus	
224EA		✓	✓	✓				(102)
225GS	224[c]							(105)
152VT156NS				✓				
187DE222DG								(102)
155GE224EA								SP
222DG224EA								(102)
223QL225GS								(105)

[a] Genetic differences were identified from an amino acid alignment of A(H1N1) viruses that circulated between 1918 and 2009, including 11 pre-1957 human and five swine A(H1N1) viruses (data not shown). Subtype number conversion is indicated where appropriate;
[b] Numbers in brackets refer to previous studies listed in the References. SP, substitution predicted to change receptor binding specificity based on structural modeling (data not shown).
[c] H5.
[d] H3.
[e] influenza B.
[f] Unpublished results.

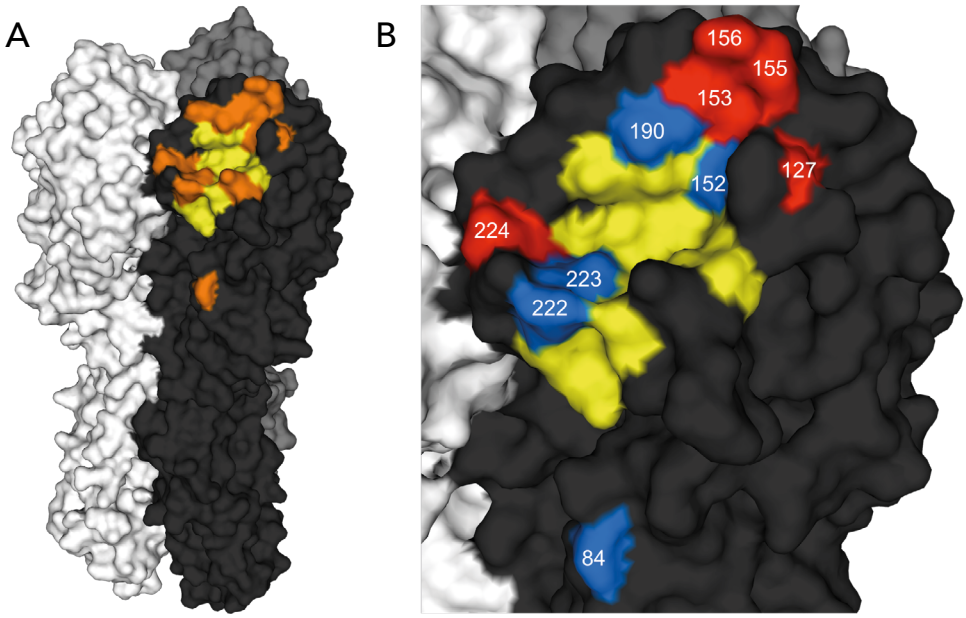


Fig. 1. Position of amino acid substitutions indicated on an A/California/04/09 HA crystal structure. (A) The three HA monomers are indicated in white, gray, and black; the RBS is in yellow. Amino acid positions that were mutated in this study are indicated in orange. (B) Zoom image of the globular head of HA. Amino acid substitutions in mutants that were substantially antigenically different from A/Netherlands/602/09 (escape mutants) are indicated in red; substitutions in mutants that were antigenically similar to A/Netherlands/602/09 and that were included in further antigenic analyses (non-escape mutants) are indicated in blue.

HI titers than A/Netherlands/602/09 (4-fold for 155GE when antisera prepared to A/Netherlands/219/11 were not included). The mean HI titers of 156NS, 152VT156NS, 156NY, 155GE224EA, and 153KE were up to 14-fold (3.8 $\log_2$) lower than A/Netherlands/602/09 HI titers. These results indicate that the amino acid substitutions in all escape mutants except 224EA caused evasion of recognition by antisera raised to this panel of A(H1N1)pdm09 viruses. Mutants 156NS and 152VT156NS were antigenically similar, as were 152VT and A/Netherlands/602/09, thus indicating that 156NS was solely responsible for the antigenic change of this double mutant. Introduction of 224EA as single substitution had only minor effects on HI titers to most antisera. The 155GE224EA double mutant displayed a mean HI titer decrease of 12-fold (3.6 $\log_2$), indicating that these substitutions had a cumulative antigenic effect.

Analysis of antigenic properties using human infant sera

Antibodies in sera obtained from infants that experienced a primary A(H1N1)pdm09 virus infection were elicited in response to a single antigenic variant of influenza virus, as is the case for the antibody repertoire of inoculated ferrets. We next examined if

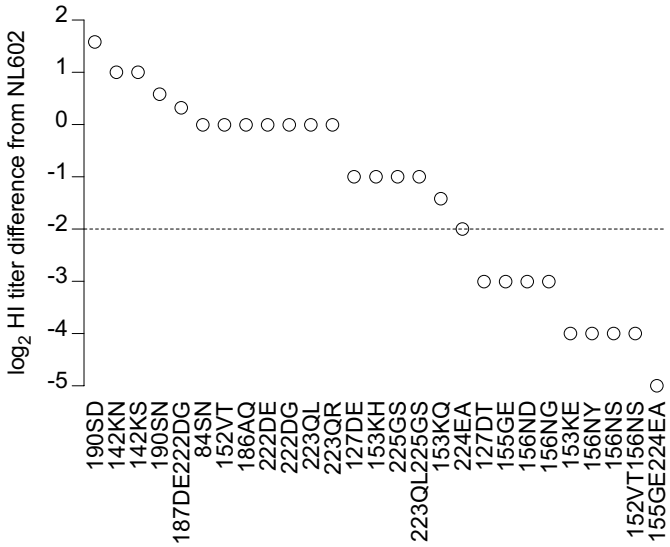
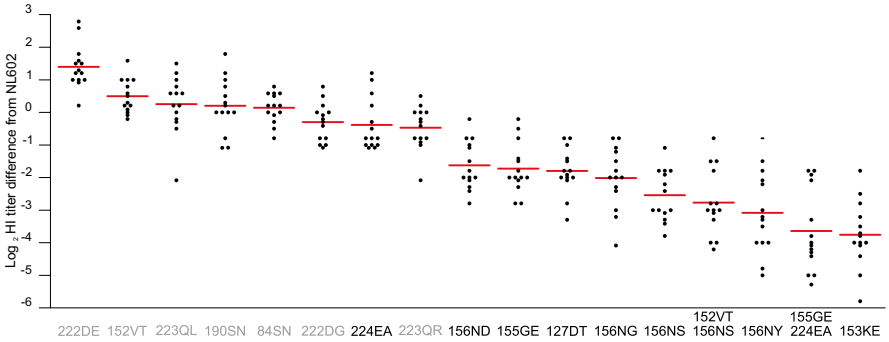


Fig. 2. HI titer differences between viruses with wild-type or mutant HAs against an A/Netherlands/602/09 ferret antiserum. Viruses with either wild- type or mutant A/Netherlands/602/09 HAs were tested in HI assays with a ferret antiserum prepared against the A/Netherlands/602/09 wild-type virus. Each point represents the log2 HI titer difference between a mutant and A/Netherlands/602/09. Mutants with HI titers at least 4-fold (2 log2) lower than that of A/Netherlands/602/09 (dashed line) were considered substantially antigenically different. The viruses are ordered by the log2 HI titer difference from A/Netherlands/602/09.

the substitutions that led to antigenic variation as tested by ferret antisera were sufficient to escape recognition by human antibodies. 49 surplus sera from unvaccinated infants were tested in HI assays for the presence of antibodies to A(H1N1)pdm09 virus. Six sera from infants born in 2009 or 2010 that were between 6 and eleven months of age at the time of sampling had detectable HI titers to A/Netherlands/602/09. HI titers to A/Brisbane/59/07 (seasonal A(H1N1)) were below the detection limit of the HI assay at a starting dilution of 1/40. This suggested that the infants experienced a primary infection and that maternal antibodies potentially present in the sera did not influence our results. Escape and non-escape mutants were tested in HI assays to sera 6, 11, 14 and 16. Non-escape mutants were not included in HI assays with sera 7 and 25 because of insufficient material available for these sera.

Escape and non-escape mutants had 3.5 – 14-fold lower mean HI titers than A/Netherlands/602/09 (Fig. 4A). The variation in HI titers between the different infant sera was larger than for ferret antisera and contradicting results were obtained with different sera. For example, mutant 222DE was antigenically similar to A/Netherlands/602/09 using ferret sera and infant serum 14, yet the HI titer to serum 16 was more than 10-fold lower. Sera 6, 7 and 16 poorly discriminated between the mutants. Serum 7 had a low HI titer of 160 to A/Netherlands/602/09 and HI titers of the mutants tested to this serum were similarly low or up to 4-fold higher. The HI

A



B

	NL602	222DE	152VT	223QL	190SN	84SN	222DG	224EA	223QR	156ND	155GE	127DT	156NG	156NS	152VT 156NS	156NY	155GE 224EA	153KE
NL/151B/11	2217	7680	2715	2560	5120	1280	3840	2560	1920	1920	1280	640	1280	640	1280	1280	640	320
NL/219A/11	2217	2560	1920	1920	2560	2560	1920	1280	1280	1280	1920	1280	1280	640	640	640	640	640
NL/2001B/09	1109	7680	2217	2560	3840	1280	1280	2560	1280	640	320	640	480	320	160	240	80	120
NL/151A/11	1810	5120	1920	1280	2560	1280	2560	1280	960	640	640	480	640	480	640	640	480	160
NL602A	2217	5120	3840	3840	3840	2560	1280	1280	1280	640	320	640	640	480	320	240	113	160
NL/219B/11	1358	2560	1280	320	640	1280	1280	640	320	640	960	480	640	640	480	320	320	240
NL/007A/10	2560	5120	2560	5120	2560	2560	2560	3840	1280	640	960	640	640	480	320	320	160	160
NL/2001A/09	1280	7680	3840	1280	2560	1280	1280	2560	960	320	320	640	320	160	80	80	60	80
NL/007B/10	1568	3840	1810	1280	1920	1280	1280	1280	1280	320	320	160	320	160	160	160	160	120
NL602B	2560	5120	5120	3840	2560	3840	1280	1280	2560	640	640	640	640	320	320	160	139	160
CAL/004A/09	2560	5120	3840	3840	2560	3840	1280	1280	2560	1280	640	960	320	320	320	160	80	120
CAL/004B/09	2560	7680	5120	3840	2560	3840	1920	1280	2560	480	640	640	480	240	160	80	80	80
CAL/007A/09	2715	7680	3840	7680	1280	3840	1280	1280	3840	640	640	640	160	320	320	240	160	160
CAL/007B/09	2217	5120	2560	2560	1280	3840	1280	1280	1280	320	320	320	240	160	120	80	57	40

Fig. 3. HI titer differences between viruses with wild-type or mutant HAs in ferret antisera. Each point in panel A represents the log2 HI titer difference between a mutant and A/Netherlands/602/09 for an individual ferret antiserum. The viruses are ordered by the mean log2 HI titer difference from A/Netherlands/602/09, which is indicated as red horizontal lines. Names of escape and nonescape mutants are shown in black and gray, respectively. Ferret antisera are indicated in the leftmost column of panel B and are ordered from top to bottom by a decreasing ability of the serum to inhibit the test viruses in the HI assay. Two antisera (labeled A and B) were prepared against each virus. HI titers are color-coded for the difference from A/Netherlands/602/09 (NL602): orange, equal to or higher than that of A/Netherlands/602/09; yellow, up to 2-fold lower; green, 2- to 4-fold lower; cyan, 4- to 8-fold lower; blue, 8- to 16-fold lower; purple, 16- to 32-fold lower; and magenta, at least 32-fold lower.

titers to sera 6 and 16 were at least 8-fold lower than that of A/Netherlands/602/09 for all mutants.

To test if escape mutants had an antigenic advantage over non-escape mutants in evasion of the infant sera that discriminated between the mutants, the results from sera 7 (that had similarly high titers to all viruses), and 6 and 16 (that had low titers to all mutants) were omitted and the analysis was repeated. Non-escape mutants and 156ND had mean HI titers that ranged from similar to 2-fold lower than A/Netherlands/602/09, individual HI titers to the different sera were up to 3-fold lower (Fig. 4B, 4C). The remaining mutants, which were all escape mutants, had

mean HI titers 2 to 10-fold lower than A/Netherlands/602/09 HI titers, and had one or more individual HI titers that was at least 4-fold lower than HI titers to A/Netherlands/602/09. Therefore, substitutions responsible for antigenic change as measured by ferret antisera can also mediate escape from recognition by human antibodies elicited in response to a primary infection with an A(H1N1)pdm09-like virus. However, in contrast to the ferret sera, the infant sera were sometimes non-discriminative (serum 7) or reacted low across the board with all mutants (sera 6 and 16). Thus, mutants 127DT, 153KE, 155GE, 156ND, 156NG, 156NS, 156NY, 152VT156NS, and 155GE224EA escaped from recognition by antibodies in ferret sera and human infant sera (Figs. 3 and 4).

Analysis of virus replication

Next, plaque assays were performed to test the effect of the introduced substitutions on replicative fitness. Plaque sizes were determined at a fixed time-point after inoculation and were used as a proxy for replication efficiency, i.e. larger plaques indicate more efficient replication.

The majority of the mutants displayed plaque sizes similar to A/Netherlands/602/09 (Fig. 5). Single mutants 187DE, 156NY, 224EA, 84SN, and 222DG displayed reduced plaque sizes. Mutants 127DT, 153KH, 153KE, 155GE, 153KQ, and 156ND had predominantly larger plaques. Of the ten mutants that had an antigenic effect in HI assays with ferret or infant sera, only 156NY and 224EA had substantially smaller plaques. The 155GE224EA double mutant displayed plaque sizes similar to A/Netherlands/602/09, indicating that 155GE compensated for the adverse effect of 224EA. Thus, eight mutants (127DT, 155GE, 156ND, 156NG, 153KE, 156NS, 152VT156NS, and 155GE224EA) were antigenically different from A/Netherlands/602/09 with no apparent loss of replication efficiency (Figs. 3, 4, and 5).

Effect of substitutions on receptor binding specificity

To test if substitutions that caused antigenic change in the HI assays had altered receptor binding specificity, the escape mutants were tested in hemagglutination assays using normal TRBCs or TRBCs resialylated to contain either α2,3- or α2,6-linked SAs. Removal of SAs from the TRBC surface and correct resialylation was confirmed in hemagglutination assays using the avian A/Vietnam/1194/2004 (H5N1) and A/Netherlands/213/03 (H3N2) viruses. A/Netherlands/602/09 and all mutants yielded hemagglutination titers to TRBCs reconstituted with α2,6-linked SA (Table 2). Only mutants 153KE, 155GE and 156ND showed a very weak binding to TRBCs with α2,3-linked SAs. Mutants 156NG and 156NS showed markedly reduced binding to α2,6-reconstituted TRBCs as compared to unmodified TRBCs. Although some mutations that affect antigenic properties appeared to influence receptor

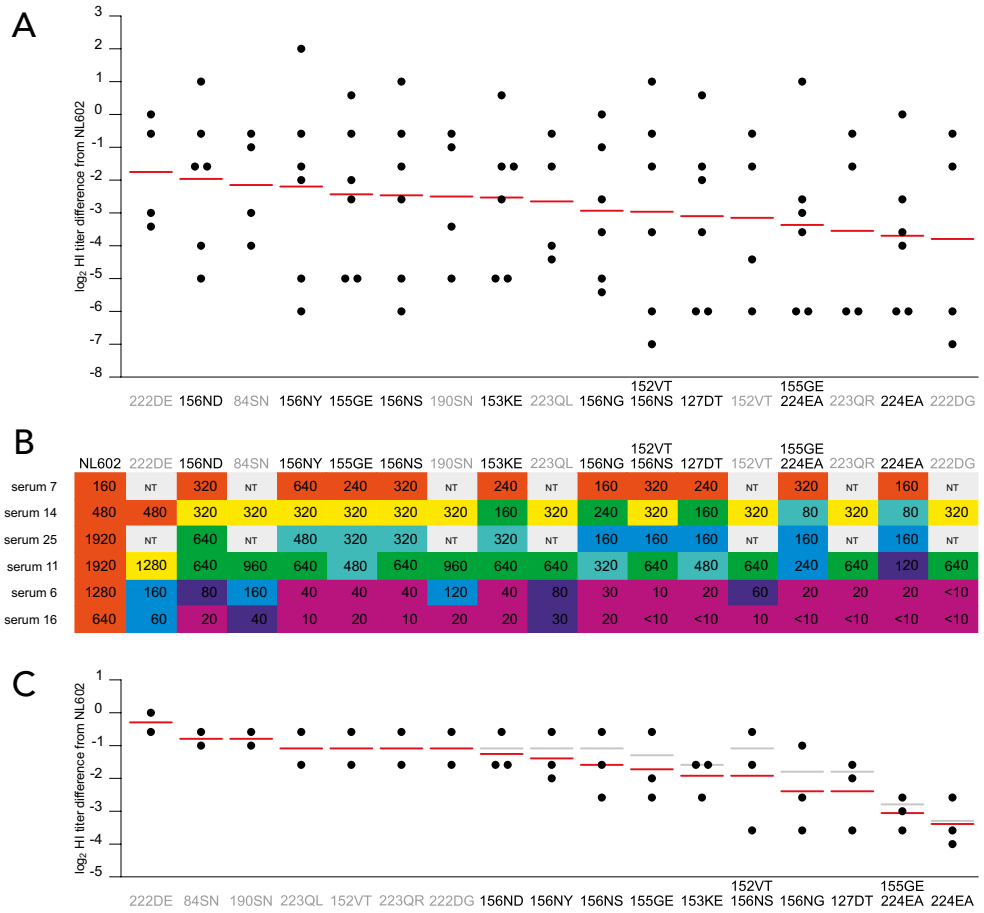


Fig. 4. HI titer differences between viruses with wild-type or mutant HAs against human infant sera. (A) Symbols, order, and nomenclature are as in Fig. 3. (B) HI titers are color-coded for the difference from A/Netherlands/602/09: orange, equal to or higher than that of A/Netherlands/602/09; yellow, up to 2-fold lower; green, 2- to 4-fold lower; cyan, 4- to 8-fold lower; blue, 8- to 16-fold lower; purple, 16- to 32-fold lower; and magenta, at least 32-fold lower. NT, the virus-serum combination was not tested. (C) The analysis was repeated with inclusion of only the sera that differentiated between the mutants (sera 6, 7, and 16 were omitted). Non-escape mutants were not tested with serum 25 because of insufficient material available for this serum. Gray horizontal lines indicate the mean log2 HI titer difference from A/Netherlands/602/09 when this serum was also omitted for the other mutants.

specificity to some extent, a clear tendency towards reduced affinity for α2,6-linked SAs was not demonstrated.

Frequency of substitutions in natural isolates

We analyzed the frequencies of the tested substitutions in 10,422 A(H1N1)pdm09 sequences from GenBank submitted between April 2009 and February 2014

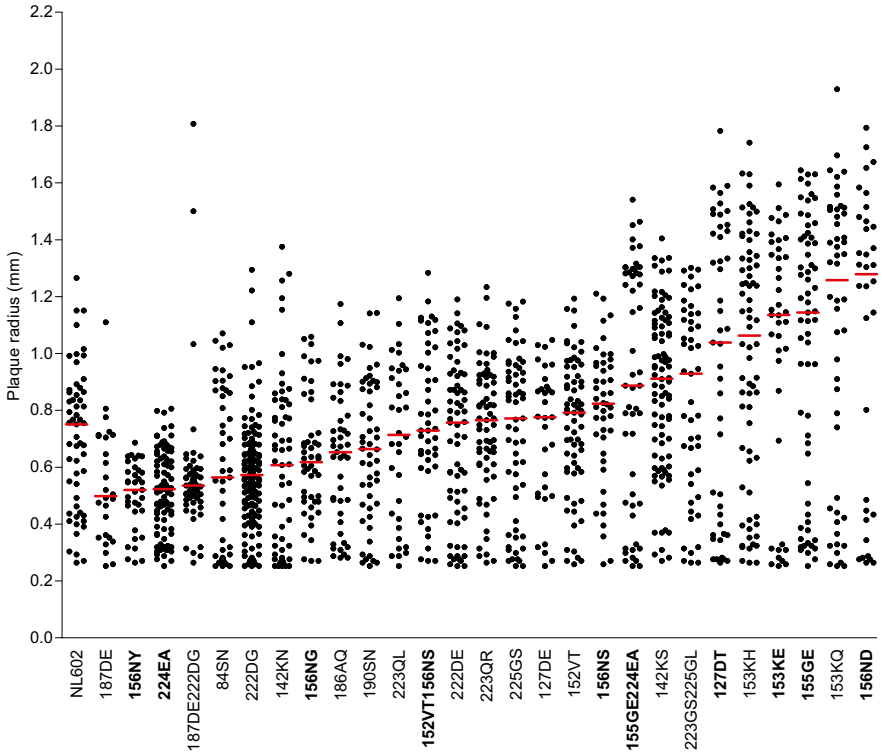


Fig. 5. Effects of substitutions on virus replication. MDCK cells were inoculated with viruses containing wild-type or mutant A/Netherlands/602/09 HAs. After 36 h, the plaque sizes were determined as a measure of replication efficiency. Each point indicates the size of a single plaque. The mutants are ordered by increasing median plaque size, which is indicated by red horizontal lines. Escape mutants are in bold.

(data not shown). Four substitutions were detected in more than 1% of the sequences: three non-escape mutants; 222DE, 222DG, and 223QR (4.80, 1.94, and 1.60%, respectively), and escape mutant 155GE (1.46%). Of the mutants with substantially increased plaque sizes, substitutions 155GE, 156ND, 153KE were found in 1.46, 0.36 and 0.17%, of the sequences, respectively. Substitutions 127DT and 153KH were not detected. Substitution 156NS, which had a large antigenic effect but did not affect plaque size was detected in 0.18% of the sequences. Other substitutions in escape mutants that did not affect plaque size or resulted in substantially smaller plaques were not detected.

Analysis of antigenic properties using human sera

The majority of the human population has a more broadly reactive antibody repertoire than influenza-virus inoculated ferrets or infants after primary infection due to previous infections or vaccination with seasonal A(H1N1) viruses (106, 107). Therefore we also

Table 2. Agglutination of TRBCs by viruses with wildtype or mutant HAs[a]

	HA titer			
Antigen	**TRBC**	**VCNA**	**α2,3-TRBC**	**α2,6-TRBC**
A/Netherlands/602/09	512	0	0	32
127DT	128	0	0	32
153KE	128	0	1	64
155GE	512	0	1	16
156ND	512	0	2	128
156NG	256	0	0	4
156NS	256	0	0	2
156NY	64	0	0	32
224EA	128	0	0	64
152VT156NS	128	0	0	8
155GE224EA	1024	0	0	1024
A/Vietnam/1194/2004	128	0	256	0
A/Netherlands/213/2003	256	0	0	256

[a] Hemagglutination titers are expressed as the HAU with unmodified TRBCs, TRBCs stripped from SAs using VCNA, or TRBCs resialylated to contain either α2,3- or α2,6- SA. A/Vietnam/1194/2004 and A/Netherlands/213/2003 served as typical avian and human viruses with α2,3- and α2,6-SA preferences, respectively.

tested the mutants in HI assays to human sera obtained from individuals anticipated to have experienced both seasonal A(H1N1) virus and A(H1N1)pdm09 virus infections. We first tested the reactivity of a panel of pre-2009 swine (-origin) viruses (A/swine/Shope/56 and A/New Jersey/8/76) and human vaccine strains (A/USSR/92/77, A/Chile/1/83, A/Taiwan/1/86, A/New Caledonia/20/99, and A/Brisbane/59/07) to a ferret antiserum prepared to A/Netherlands/602/09. HI titers were below the detection limit of the HI assay at a starting dilution of 1/20 (data not shown), indicating that these viruses were antigenically distinct from A(H1N1)pdm09 viruses. 21 sera obtained from individuals aged 7 to 85 at the time of serum collection (October 2009 – May 2011) were selected based on their ability to inhibit A/Netherlands/602/09 in HI assays. The sera were tested in HI assays to ten seasonal A(H1N1) viruses covering the period from 1977 to 2009 (Fig. 6A). All sera had detectable HI titers to five or more seasonal A(H1N1) viruses, demonstrating a broadly reactive antibody repertoire. Subsequently, the escape and non-escape mutants and A/Netherlands/602/09 were tested in HI assays using the selected sera (Figs. 6B and C). The mean HI titers of escape as well as non-escape mutants were less than 1.5-fold (0.54 $\log_2$) lower than A/Netherlands/602/09 HI titers, suggesting that none of the mutants escaped recognition by this panel of sera. The majority of individual HI titers to the different sera were also less than 2-fold lower than A/Netherlands/602/09 HI titers. Six mutants had a single titer that was more than 2-fold lower than the A/Netherlands/602/09 HI titer: 155GE, 155GE224EA, 223QL, 222DG displayed a 4-fold lower HI titer; 224EA and 127DT had 8 and 12-fold lower HI titers, respectively.

The antigenic effect of the substitutions in neither non-escape nor escape mutants was therefore sufficient to escape recognition by this panel of human sera.

DISCUSSION

In this study we attempted to identify amino acid substitutions that contribute to antibody escape of A(H1N1)pdm09 viruses. Substitutions introduced into the HA of A/Netherlands/602/09 were selected based on genetic and antigenic changes of influenza A viruses that circulated in the past and on substitutions associated with changes in receptor binding specificity. We show that at least nine mutants were antigenically distinct from A/Netherlands/602/09 in HI assays using ferret antisera and found that the substitutions that caused escape from ferret antibodies also evaded antibodies in some human sera collected after primary infection.

Substitution 127DT introduces a glycosylation pattern (N-X-S/T-X) that potentially adds a carbohydrate side chain at position 125. Carbohydrate side chains can mask antibody epitopes and have been shown to change the antigenic properties of influenza viruses (61, 62). Seasonal A(H1N1) viruses that circulated since the 1930s until 1986 had a glycosylation site starting at position 127, viruses that circulated from 1986 onwards had a glycosylation site starting at position 125. Both glycosylation sites were absent from the 1918 A(H1N1) and A(H1N1)pdm09 viruses. Absence of glycosylation at these positions has been suggested to contribute to the antigenic difference between seasonal A(H1N1) viruses and A(H1N1)pdm09 viruses (108). Our results indicate that re-introduction of a potential glycosylation site in this region was sufficient to escape from antibodies generated in response to A(H1N1)pdm09 virus.

Substitutions 153KE and 155GE in A(H1N1)pdm09 virus are analogous to substitutions 156KE and 158GE that were responsible for major antigenic changes during evolution of A(H3N2) viruses (77). The observation that these substitutions changed the antigenic properties of an A(H1N1)pdm09 virus is also in agreement with previous studies that reported escape of A(H1N1)pdm09 virus from monoclonal antibodies (109, 110), and from polyclonal antibody responses in ferrets (108, 111, 112).

The four different substitutions introduced at position 156 (ND, NG, NY and NS) all had a large antigenic effect. Two additional substitutions on this position, 156NK and 156NE, were previously reported to escape antibody neutralization of A(H1N1) pdm09 virus (110, 111). Interestingly, the large changes in biophysical properties of the substituted amino acids introduced by 156ND or 156NY (charge and volume differences, respectively) had a similar antigenic effect to the small difference in biophysical properties introduced by 156NS. This finding emphasizes the potential importance of this position for antigenic change of A(H1N1)pdm09 viruses.

The antigenic effect of 155GE224EA in HI assays with ferret antisera was larger than the sum of the effects of the single substitutions. A possible explanation for how substitutions bordering opposite sides of the RBS can amplify each other's effect on

antibody escape is given by the greater surface area of an antibody footprint in relation to the size of the RBS (113), which allows more efficient escape from antibodies that cover the RBS in addition to the proportion of antibodies evaded by the individual substitutions. Interestingly, substitution 224EA had a small antigenic effect in HI assays with ferret antisera but was substantially different from A/Netherlands/602/09 when tested with infant sera. In tests with infant sera 155GE224EA and 224EA were equally different from A/Netherlands/602/09, indicating that 155GE did not add to the antigenic effect of 224EA. These results suggest that position 224 plays a more prominent role in escape from human antibodies than in escape from ferret antibodies. Antigenic variants with a substitution on this position may therefore not be readily detected in conventional HI assays using ferret antisera.

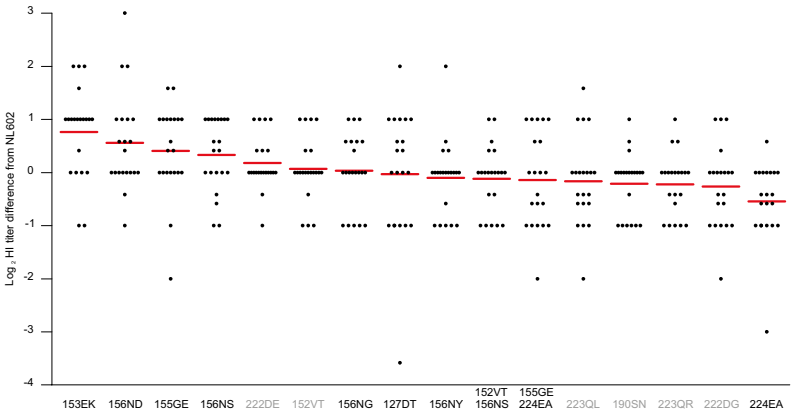
The high mutation rate of influenza A viruses and the observation that single substitutions caused substantial antigenic change during evolution of A(H3N2), A(H1N1), and influenza B viruses contradicts the relatively slow rate at which influenza viruses have changed antigenically (77). One possible explanation for this paradox is that substitutions responsible for escape from antibodies targeting the RBS have an adverse effect on HA function, necessitating the co-occurrence of compensatory substitutions, which slows down the emergence of new antigenic variants. Substitutions in and surrounding the A(H1N1)pdm09 RBS may affect receptor binding and consequently change replication efficiency (114). Surprisingly, we found that most of the substitutions that caused antigenic change in HI assays with ferret and human sera did not result in less efficient virus replication. Substitutions at positions 153 and 156 often had large effects on replication efficiency. Interestingly, change of the lysine at position 153 for three biophysically diverse amino acids all resulted in increased replication efficiency. Replacing the asparagine at position 156 with either an aspartic acid or tyrosine had opposite effects on replication. These results suggest that positions 153 and 156 are key determinants of replication efficiency. It should be noted that replication assays are a surrogate measure of viral replication *in vivo*, and that results may vary with the model system used. In addition, differences in replication efficiency may in part be caused by the use of an A/Puerto Rico/8/34 backbone. However, other studies have also demonstrated that A(H1N1)

Fig. 6. HI titer differences between viruses with wild-type or mutant HAs against human sera. (A) HI titers of human sera against antigenic variants of seasonal A(H1N1) viruses isolated between 1977 and 2009 (USSR77, A/USSR/90/77; NL78, A/Netherlands/3075/78; TA89, A/Taiwan/1/89; NL87, A/Netherlands/414/87; NC99, A/New Caledonia/20/99; NL99, A/Netherlands/271/99; NL03, A/Netherlands/02/03; NL06, A/Netherlands/364/06; SS06, A/Solomon Islands/03/06; and NL09, A/Netherlands/1005/09). (B) HI titer differences between viruses with wild-type or mutant NL602 HAs against human sera. Symbols, order, and nomenclature are as in Fig. 3. (C) HI titers are color-coded for the difference from A/Netherlands/602/09: orange, equal to or higher than A/Netherlands/602/09; yellow, up to 2-fold lower; green, 2- to 4-fold lower; cyan, 4- to 8-fold lower; and blue, 8- to 16-fold lower. The first two digits of the serum number indicate the age of the individual at the time of sampling.

A

	USSR77	NL78	TA89	NL87	NC99	NL99	NL03	NL06	SS06	NL09
55 2273	960	960	640	1280	160	1280	160	320	160	640
34 25621	640	1280	640	1280	<10	640	<10	160	10	160
18 12105	20	40	80	160	320	160	320	640	160	160
19 13519	<10	<10	320	640	80	640	160	160	160	160
24 4717	240	160	640	5120	1280	1920	640	1280	1280	480
40 23692	160	640	160	320	80	160	80	80	80	80
50 24601	40	40	160	160	80	80	80	160	40	40
07 10613	80	160	320	640	320	640	320	640	320	160
29 4404	40	80	640	1280	320	640	160	480	320	120
37 1493	1280	1280	2560	5120	<60	5120	160	2560	640	640
46 26710	640	640	80	1280	<10	320	<10	320	40	160
32 26442	<10	20	640	1280	160	1280	160	320	160	40
43 1077	2560	1920	1280	5120	30	5120	<10	1280	160	1280
43 1849	960	640	1280	2560	<10	2560	<10	640	80	640
50 2824	320	480	320	1280	<10	640	20	160	<10	240
37 25032	640	1280	<30	40	<30	80	<30	<30	<30	640
07 2965	<10	10	<10	60	120	40	80	160	40	40
41 2453	160	160	10	40	<10	20	<10	<10	<10	10
85 3763	<10	10	10	40	40	<10	20	40	<10	<10
18 21877	40	60	320	640	320	640	160	640	320	80
52 24131	80	320	10	160	640	160	240	80	<10	320

B



C

	NL602	153EK	156ND	155GE	156NS	222DE	152VT	156NG	127DT	156NY	152VT 156NS	155GE 224EA	223QL	190SN	223QR	222DG	224EA
55 2273	1280	3840	1920	3840	2560	2560	2560	1920	1280	1920	2560	1920	3840	2560	1920	1280	1920
34 25621	640	1280	1280	1280	960	640	1280	640	640	640	640	640	640	640	640	1280	640
18 12105	80	80	80	160	160	160	80	160	160	320	80	160	160	80	120	160	80
19 13519	40	160	320	80	80	80	80	80	80	40	80	80	80	40	80	80	40
24 4717	160	640	640	320	320	160	80	320	640	160	160	320	160	160	160	160	160
40 23692	160	320	320	160	240	160	320	160	80	160	240	160	160	160	160	160	160
50 24601	80	160	80	80	80	80	80	80	40	80	80	120	80	80	80	80	80
07 10613	160	320	240	160	160	160	80	160	320	160	160	320	160	160	160	160	120
29 4404	80	160	160	160	160	160	80	120	160	80	80	160	160	40	80	40	80
37 1493	3840	7680	3840	5120	5120	5120	5120	5120	5120	5120	3840	2560	2560	3840	3840	2560	2560
46 26710	480	960	640	640	640	640	640	480	640	640	640	320	320	640	320	640	320
32 26442	160	640	640	320	320	160	160	160	320	160	160	80	160	80	160	80	80
43 1077	2560	5120	2560	5120	5120	2560	2560	3840	3840	2560	1920	2560	1280	3840	1920	1920	1920
43 1849	1280	2560	2560	3840	2560	1280	1280	1920	1280	1280	960	1280	640	1280	960	960	960
50 2824	640	1280	960	960	640	640	640	640	960	640	640	160	160	320	320	640	80
37 25032	3840	5120	3840	3840	2560	5120	3840	3840	320	2560	5120	2560	2560	3840	3840	2560	2560
07 2965	640	640	640	640	640	640	640	320	320	320	320	480	480	480	480	640	320
41 2453	320	320	320	320	320	320	320	160	160	160	160	160	240	160	160	160	160
85 3763	80	40	40	20	40	80	80	40	80	40	40	40	80	80	40	40	40
18 21877	640	640	640	640	480	320	480	320	320	320	320	320	480	320	320	320	320
52 24131	2560	1280	1920	1280	1280	1920	1280	1280	1280	1280	1280	1280	1280	1280	1280	640	1280

pdm09 viruses with 153KE, 155GE, and 156ND had increased replication efficiency in eggs or MDCK cells (99, 108), and they were often associated with cell culture adaptation (99, 108, 111).

The substitutions that caused antigenic change in HI assays with ferret and human infant sera were, with the exception of 224EA, located in or near the 151 - 159 loop. Substitutions in this region have previously been shown to affect the antigenic properties of A(H3N2) and A(H5N1) viruses (77, 94). In addition, these positions often affected replication efficiency of the mutants constructed in this study. Amino acid substitutions at other positions tested in this study that caused similarly large changes in biophysical properties did not substantially change the antigenic properties or replication efficiency. However, it should be noted that it is possible that other amino acid changes outside the 151- 159 loop, that were not included in this study, can result in phenotypic change (115) . Our results are in agreement with previous studies that reported mice or human monoclonal antibodies (109, 110) or antibodies in ferret antisera (111) that target the 151 - 159 loop. We here show that several substitutions in or affecting the 151 - 159 loop caused antigenic escape while retaining replicative fitness. These results substantiate the importance of this region for antigenic evolution of A(H1N1) viruses.

We found large differences in the reactivity of infant antisera. Surprisingly, three sera (from the youngest infants) had almost uniform HI titers to all mutants that were either the same as that of A/Netherlands/602/09, or much lower. Infection in the presence of maternally derived antibodies or mechanisms associated with the transformation of the neonatal immune system to a more mature immunological phenotype may explain the difference from the other infant sera (116, 117). Nevertheless, the different reactivities of these three sera warrant further investigation. However, the remaining three sera had reactivities similar to that observed with ferret antisera. This indicates that substitutions that caused escape from ferret antibodies also promote escape from human antibodies elicited in response to infection with an A(H1N1)pdm09-like virus. In contrast, the substitutions responsible for escape from ferret and infant sera were not sufficient to escape recognition by antisera from individuals that experienced infection with seasonal and pandemic A(H1N1) viruses. Thus, the antibody repertoire of the individuals that had a more experienced immune system included antibodies that were absent in the sera of infants. These results suggest that antibodies directed to epitopes on seasonal A(H1N1) viruses are cross-reactive with epitopes on the mutants tested in this study, and thereby complement the antibody repertoire elicited in response to A(H1N1)pdm09 infection alone. The presence of such antibodies in much of the population may also explain why mutations that gave an *in vitro* replication advantage have not been detected at higher frequencies. Carter *et al.* showed that sera from ferrets sequentially infected with antigenically diverse seasonal A(H1N1) viruses can efficiently neutralize A(H1N1) pdm09 virus, whereas sera from ferrets infected with a single seasonal A(H1N1) virus were not cross-reactive (118). Additionally, a recent study by Miller *et al.* indicated

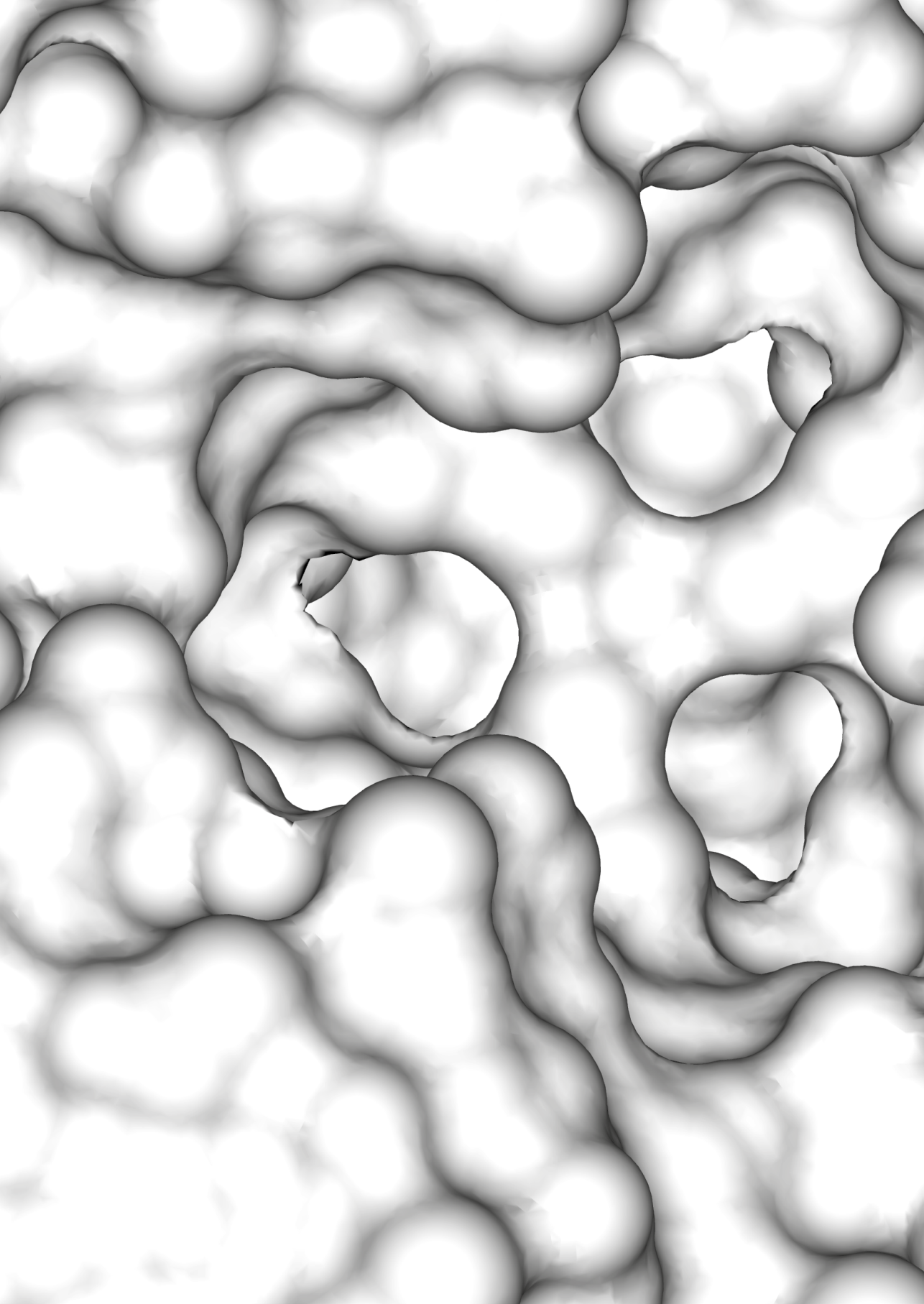
that antibodies against previously encountered influenza A viruses are periodically boosted upon natural exposure to drift variants of the same subtype (106), possibly resulting in the more broadly reactive antibody repertoire observed here. This study and other studies (112, 115) suggest that the complexity of the antibody repertoire to A(H1N1) viruses in much of the human population cannot be represented by antigenic analyses based on single-infection ferret antisera. This limits the use of ferret antisera to test the ability of an antigenic variant to escape from population immunity.

In conclusion, substitutions in or near the RBS can influence the antigenic properties of A(H1N1)pdm09 viruses. Based on the current and previous studies into antigenic change of influenza A viruses (77, 94), it is probable that emerging antigenic variants of A(H1N1)pdm09 viruses will escape from population immunity because of substitutions in or near the RBS. However, our results also suggest that the presence of antibodies directed to epitopes on seasonal A(H1N1) and A(H1N1)pdm09 viruses in much of the population limits the number of antigenic variants that can emerge to cause new epidemics.

ACKNOWLEDGEMENTS

This work was supported by a ZonMW VICI grant and NIH contracts HHSN266200700010C and HHSN272201400008C, NIH Director's Pioneer Award DP1-OD000490-01, European Union FP7 program EMPERIE (223498), European Union FP7 program ANTIGONE (278976), and program grant P0050/2008 from the Human Frontier Science Program. PF receives funding from the EU FP7 project PREPARE (602525). MdG was funded by a Marie Curie fellowship under contract PIEF-GA-2009-237505. EdW is supported by the Intramural Research Program of NIAID, NIH.

We are grateful to G. van Amerongen, R. van Beek, B. Laksono, N. Lewis, G. de Mutsert, and C. A. Russell for stimulating discussions and technical assistance.



SUMMARIZING DISCUSSION

SUMMARIZING DISCUSSION

Substitutions near the receptor binding site shape antigenic evolution

It is thought that the necessity to re-infect their hosts in the face of population immunity drives antigenic change of seasonal human influenza viruses. Influenza viruses escape population-level antibody-mediated neutralization by accumulating amino acid substitutions in the globular head domain of the hemagglutinin (HA) surface protein (32-34). Early studies on influenza virus antigenic change identified four (H1) or five (H3) regions of potentially important amino acid positions that could cause antibody escape upon mutation (32-34). It was postulated that at least four amino acid substitutions in two or more antigenic sites are necessary for an antigenic variant to become epidemically relevant (32). The studies described in this thesis demonstrate that antigenic change sufficiently large to escape recognition from antibodies to previously circulating antigenic variants never required more than three amino acid substitutions and that it was single amino acid substitutions that caused the majority of large antigenic changes during influenza virus evolution. For both influenza A and B virus and for the human and avian viruses, 15 of 19 major antigenic changes for which the molecular basis is known were primarily caused by a single amino acid change (chapters 2 and 3, and (31, 49)). Single amino acid substitutions introduced into a virus representing A(H1N1)pdm09 viruses were also sufficient to escape recognition by antibodies specific to A(H1N1)pdm09 viruses that circulated between 2009 and 2011 (chapter 5), and are thought to cause increased susceptibility of people in certain age groups to contemporary A(H1N1) viruses (112, 115).

The antigenic sites of A(H3N2) virus contain 131 amino acid positions that cover much of the HA globular head (32, 33). Each of these positions potentially contributes to antigenic change. The 20 amino acids that could theoretically occupy each of the 131 positions imply a plethora of amino acid substitutions that may lead to escape from neutralizing antibodies. In chapter 2 we showed that major antigenic change during evolution of A(H3N2) virus was due to substitutions on key positions located immediately adjacent to the receptor binding site (RBS). Major change during A(H5N1) clade 2.1 virus evolution, and large antigenic differences between recent pre-2009 A(H1N1) viruses and between influenza B viruses, was also caused exclusively by amino acid changes adjacent to the RBS (chapters 2 and 3, and (49)). The most striking example indicating the importance of positions adjacent to the RBS for antigenic change was an experiment in which the amino acids on the seven key positions from an A(H3N2) virus that circulated in 2003 were introduced in the HA of a virus isolated in 1968. Although there were 54 amino acid differences between the HAs of these viruses, introduction of only the five amino acid differences on the key antigenic positions in the 1968 virus resulted in a mutant that had antigenic properties characteristic of viruses isolated in 2002 and 2003. Thus, the area immediately surrounding the RBS determines the antigenic phenotype of seasonal human influenza viruses and HPAI A(H5N1) clade 2.1 viruses.

In chapter 2 we showed that substitutions at only seven key positions dictated 35 years of A(H3N2) virus antigenic evolution, the work for chapter 3 showed that substitutions at six key positions explained the antigenic changes of A(H5N1) clade 2.1 virus in chickens. Five of seven key positions of A(H3N2) virus and three of six key positions of A(H5N1) clade 2.1 virus were involved in major antigenic at least twice, suggesting that key positions were selected over other bordering positions to cause major antigenic change. Interestingly, the key positions identified for A(H3N2), A(H1N1), A(H5N1) and influenza B viruses were often analogous between influenza viruses. Six key positions shared between two or more influenza A subtypes and B viruses map to equivalent positions on the hemagglutinin (Table 1, Fig. 1A).

Different influenza viruses, similar constraints

The results from chapter 4 suggest that the ability of key substitutions to cause large antigenic change of A(H3N2) viruses is largely independent of the amino acid context, emphasizing the potential of key positions to cause antigenic change. Based on chapters 2 and 3, we hypothesized that substitutions on key positions identified for other subtypes may cause antigenic change of A(H1N1)pdm09 viruses, which have to date remained antigenically similar to the viruses that caused the 2009 influenza

Table 1. Amino acid positions responsible for major antigenic change of influenza viruses studied for this thesis.

Influenza A virus subtypes				Influenza B virus lineage	
H1	H1pdm09	H3	H5	B/Vic	B/Yam
130[a]		133[b]	129		
		135[b]			
		137	133		
140		144			
		145			
		155	151		
	153	156			
	155	158			
	156	159		165	166
		187	183		
		189	185		
		193[c]	189		

Numbering of the amino acid positions is according to Burke and Smith (144). The positions listed for all influenza viruses but A(H1N1)pdm09 viruses were responsible for major antigenic change during influenza virus evolution. The positions indicated for A(H1N1)pdm09 were identified to support antibody escape, but as A(H1N1)pdm09 viruses have remained antigenically stable since 2009 they were not involved in observed antigenic change (chapter 5). [a]McDonald *et al.*(49) [b]For three antigenic cluster transitions during evolution of A(H3N2) virus an accessory amino acid substitution was required to change the antigenic phenotype to that characteristic of the subsequent antigenic cluster (chapter 2). [c]A substitution on position 193 was responsible for the antigenic cluster transition between the VI75 and TX77 antigenic clusters, and was involved in an antigenic cluster transition a second time as accessory substitution for the SI87 to BE89 cluster transition (chapter 2). H3 positions indicated in grey are indicated for numbering conversion, but were not involved in antigenic change.

pandemic. Introducing single substitutions at key positions of A(H3N2), A(H5N1), and influenza B viruses into an A/(H1N1)pdm09 representative virus resulted in mutants that escaped recognition from A(H1N1)pdm09-specific antibodies (chapter 5). Collectively these results suggest that vastly different influenza viruses have similar constraints to escape from antibody-mediated neutralization.

The work for this thesis suggests another constraint to antibody escape. Twenty-four of 26 substitutions responsible for major antigenic change involved substantial alterations of the biophysical properties of the amino acids. Twenty-one amino acid substitutions resulted in a charge change, and five substitutions resulted in a large change in amino acid volume. These findings indicate that only a limited number of the possible amino acid substitutions were selected to cause major antigenic change (chapters 2 and 3).

However, many substitutions that would cause a similarly large change in the biophysical properties of the amino acids and that may result in antibody escape were not observed in wildtype viruses. We found that key substitutions for antigenic change of A(H3N2) viruses never required three nucleotide changes and were always caused by the smallest number of nucleotide changes possible. Some amino acid changes may be absent because they require multiple nucleotide substitutions and are therefore more difficult to access than the substitutions that have caused antigenic change. Others may revert the antigenic properties to a phenotype against

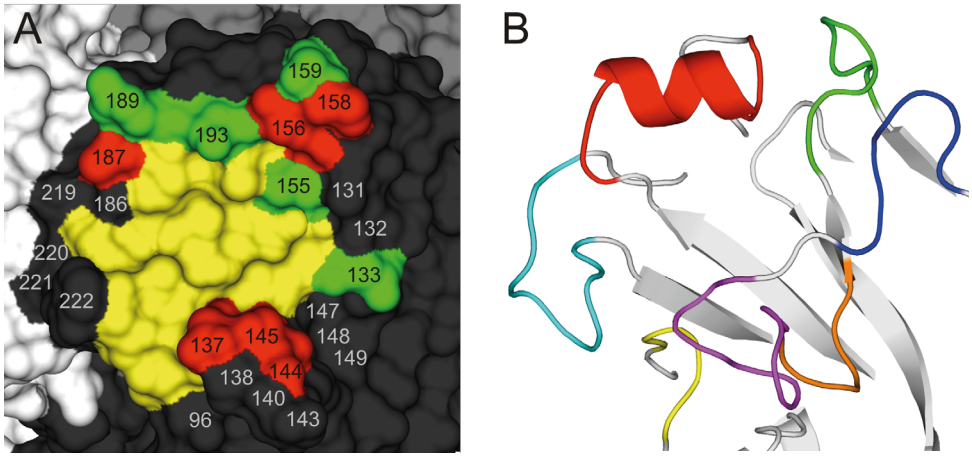


Fig. 1. Amino acid positions bordering the RBS of an A/Aichi/2/68 A(H3N2) virus HA. (A) Structural representation of the RBS area. The monomers are indicated in white, grey, and black. Key positions identified in a single subtype are indicated in red, key positions shared between multiple subtypes and influenza B virus are indicated in green (defined in Table 1). Non-key positions bordering the RBS are numbered in grey fonts. The RBS is indicated in yellow. (B) Cartoon representation of the RBS area. The structural elements surrounding the RBS are arbitrarily defined and color-coded for reference: 100-loop, yellow; 130-loop, blue; 140-loop, magenta; 150-loop, orange; 160-loop, green, 190-helix, red; 220-loop, cyan. Key positions are located in the 130-, 140-, and 160-loops and the 190 α-helix.

which there is already a high level of population immunity, or may simply not result in antigenic change. Finally, some substitutions that have a low genetic barrier and a large antigenic effect away from immunity to previous strains may have a deleterious effect on viral intrinsic fitness.

The majority of amino acid substitutions responsible for antigenic change of A(H5N1) clade 2.1 viruses were not seen in seasonal human influenza viruses and vice versa. Although antigenic change of both subtypes was often caused by substitutions that affected amino acid charge, almost all antigenically important substitutions in A(H5N1) clade 2.1 virus also involved changes in amino acid hydrophilicity. Except for a single threonine to tyrosine substitution, hydrophilicity changes were absent from substitutions that caused antigenic change of seasonal human influenza viruses. Perhaps this difference is due to the different receptor specificities of A(H5N1) viruses and seasonal human influenza viruses. Additional work looking into the effects of HA amino acid substitutions on intrinsic fitness is necessary to test this hypothesis. Nonetheless, the use of limited sets of amino acids for antigenic change suggests that seasonal human influenza viruses and A(H5N1) clade 2.1 viruses are similarly constrained to escape antibody recognition, but both may utilize distinct amino acid subsets to couple the necessity for antibody escape with optimal HA functionality.

The small number of key positions and strong biophysical limitations to the substitutions at key positions (chapters 2 and 3), the limited influence of the amino acid context on the effect of key substitutions (chapter 4), and the finding that the antigenic effect of substitutions may be translated to other influenza viruses (chapters 2 and 5) suggest that antigenic change of influenza viruses may be much more predictable than previously anticipated.

Contribution of substitutions away from the RBS to antigenic change

We found that substitutions away from the RBS never caused substantial antigenic change (chapter 2). Almost 90% of the naturally occurring amino acid substitutions that delineated between the antigenically different groups of viruses, including some on positions adjacent to the RBS, had a small or undetectable antigenic effect. This result does not imply that substitutions away from the RBS are irrelevant for antibody escape because they may collectively add to a decreased recognition by antibodies raised to strains from earlier antigenic clusters. We showed in chapter 2 that hemagglutination inhibition (HI) assay data and virus neutralization (VN) assay data correlated well, and that substitutions away from the RBS had no effect on the antigenic properties when tested in VN assays. Nevertheless, the antigenic analyses for this thesis were almost entirely based on HI assay data that rely on receptor binding interference. Other modes of antibody-mediated neutralization cannot be detected using HI assays. Antibodies that prevent fusion of the viral and endosomal membranes, that interfere with the release of progeny virus, or that activate other arms of the immune system by opsonization or antibody dependent cell cytotoxicity are not considered in the HI assay but may significantly contribute to viral clearance

(119-124). However, most such antibodies target conserved regions of HA and may therefore be less relevant for past antigenic change of influenza viruses.

Antisera prepared to pointmutants with a substitution responsible for antigenic change had substantially increased reactivity with antigenically advanced wildtype viruses, suggesting that a single amino acid change close to the RBS is biologically relevant (chapters 2 and 3). However, it has not yet been tested if a cumulative effect of key substitutions and substitutions on other positions are required *in vivo* to achieve a magnitude of antibody escape approximating that of wildtype viruses. To specifically study the contribution of substitutions that lead to antibody evasion in an immune host, the interference of innate and T cell immunity to viral clearance should be minimized. Passive transfer of antisera to naive animals followed by challenge with wildtype or mutant viruses is a possible method to determine if mutants with substitutions on key positions escape antibody neutralization as efficiently as wildtype viruses. Similar challenge studies are also necessary to test if recombinant viruses designed to contain the key amino acids to match an antigenic variant provide levels of protection comparable to vaccination or infection with wildtype viruses.

Intrinsic viral fitness

The rates of antigenic change for the A(H3N2) and A(H1N1) subtypes and influenza B/Victoria and B/Yamagata lineages are estimated at 1.01, 0.62, 0.42 and 0.32 units per year, respectively (125). For A(H3N2) virus the rate of antigenic change translates to antigenic cluster transitions occurring approximately once every 3.3 years (31). Major antigenic change occurred even less frequently for A(H1N1) and influenza B viruses (unpublished observations). Koelle *et al*. hypothesized that every position in an epitope is potentially important and that the amino acid context in which a substitution occurs determines if it results in antigenic change (56). In this model, influenza viruses evolve along antigenically neutral networks, thereby accumulating mutations that create the right context for a substitution to cause antibody escape. The neutral networks model accurately describes the hallmarks of A(H3N2) virus evolution, but suggests that the existence of a small set of key influential positions for antigenic change is unlikely. The results from chapters 2 and 4 are not in agreement with this suggestion.

Because single substitutions were often sufficient to cause major antigenic change and every possible point mutation and many double mutations are generated with each viral replication cycle [chapter 2, (49, 126)], factors other than the occurrence of substitutions that caused large biophysical changes must have limited the frequency at which new antigenic clusters emerged. To explain this, we hypothesize that major antigenic change may come at an intrinsic fitness cost (chapter 2). A substitution that causes antibody escape but adversely affects HA functionality may only have a selective advantage if the adverse effect is compensated by co-mutations. In this scenario, the rate at which influenza viruses change antigenically is impeded by the necessity for co-mutations that offset intrinsic fitness loss. Thus, rather than

the necessity to shape an epitope for antibody escape the neutral networks model may have captured the selection of antigenic variants that have the right combination of substitutions that cause antigenic change and that offset intrinsic fitness loss. Further studies are required to obtain a better insight into the balance between antibody escape and intrinsic virus fitness. First, an assay that is sufficiently sensitive to detect the potentially minute differences in replication efficiency caused by single amino acid substitutions should be developed. A plaque assay-based method that may allow such sensitivity is currently near completion (R. Mögling *et al.*, in preparation). Second, it should be determined if the hypothesized intrinsic fitness loss is present in mutants carrying key substitutions and if the paradox between high mutation rate and slow antigenic evolution can be explained by the necessity to compensate for intrinsic fitness loss. Third, the amino acid substitutions that offset intrinsic fitness should be mapped to provide important new insights in the dynamics of influenza virus antigenic evolution. Multiple studies have identified codons under positive selection during influenza virus evolution, and many positively selected substitutions occurred almost simultaneously and sped up fixation in the viral population (44-47, 127-129). The positively selected substitutions that occur alongside key substitutions may provide clues to which amino acid changes may offset intrinsic fitness loss.

Mechanisms of antigenic change

The work by Smith *et al.* indicated that antigenic change of A(H3N2) viruses is punctuated, thus providing an alternative for gradual antigenic change as the absolute basis for escape from population immunity (31). Antigenic evolution of A(H1N1), A(H2N2), and influenza B viruses also appears to be punctuated (chapter 2, (125), and unpublished observations). The identification of key positions in HA and single amino acid substitutions responsible for major antigenic change provides a molecular explanation for the punctuated nature of influenza virus antigenic evolution. The current understanding of influenza virus antigenic evolution thus no longer supports the minor and gradual antigenic change principle presumed by the antigenic drift model.

Another alternative for gradual antigenic change was provided by a study that looked into the contribution of reassortment to genetic variation. Multiple influenza virus lineages can persistently co-circulate and reassortment of the HA segment between viruses of different lineages may influence major antigenic change (130). Indeed, the SY97 to FU02 antigenic cluster transition during A(H3N2) virus evolution was associated with a reassortment event that involved the HA segment that carried the 156QH substitution (130-132). However, an elaborate study analyzing 286 influenza virus full genome sequences that covered the entire period of A(H3N2) virus circulation showed that no other persistent reassortment events coincided with antigenic cluster transitions (132). Although reassortment could still play an important role in providing the right gene constellation to support the emergence of new epidemic strains, its direct contribution to antigenic change seems limited.

An interesting hypothesis based on the observation that passaging of A/Puerto Rico/8/34 (H1N1) escape variants through immune and non-immune mice leads to differences in receptor binding proposed that antigenic change is a by-product of changes in receptor binding avidity (57). A higher receptor binding avidity, the combined affinity that results from multivalent interactions, may be advantageous for the virus when infecting (partially) immune hosts because sialic acids and antibodies compete for binding to HA. When mutants with increased receptor binding avidities where passaged in non-immune mice avidity returned to the original state, a trade-off between intrinsic fitness and immune escape, and some substitutions responsible for this decrease in binding occurred in antigenic sites. The avidity hypothesis postulates that increased positive charge enhances avidity by increasing charge attraction with negatively charged cell surfaces and that substitutions that alter avidity frequently modify antigenicity of seasonal human influenza A viruses (57, 133). During the 35 years of antigenic evolution of A(H3N2) viruses studied in chapter 2, the number of amino acid substitutions leading to an increase in positive charge were more prevalent than decreases in negative charge. However, for both the substitutions identified as responsible for cluster transitions and the cluster-difference substitutions [the amino acids that differ between all viruses of two antigenic clusters (31)] we found that some cluster transitions coincided with a (net) increase in positive charge while others were neutral or resulted in an increase of negative charge. In at least five major antigenic changes studied in chapter 2 the charge did not change as predicted by the avidity hypothesis. Importantly, the magnitude of antigenic change observed during most cluster transitions far exceeds the antigenic change observed during the passaging experiments in mice (57). The studies that support the avidity hypothesis used erythrocytes with decreasing amounts of sialic acids to titrate influenza viruses (57, 112, 133). Avidity was then determined by quantifying agglutination of erythrocytes with the lowest amount of sialic acids that showed agglutination similar to untreated erythrocytes. A major flaw in this method is that the amount of input virus was measured by HA assay. The HA assay also measures agglutination of erythrocytes and is therefore dependent on the amount of HA present on the surface of the virus, which has not been taken into consideration. More importantly, hemagglutination is dependent on the avidity of HA for sialic acids on erythrocytes. Given that A(H3N2) viruses continue to cause annual epidemics, the inability of most contemporary A(H3N2) viruses to agglutinate erythrocytes through HA binding suggests that avidity measurements using hemagglutination assays are of limited biological relevance. Although the avidity hypothesis provides a possible mechanism for the minor antigenic changes that occur within antigenic clusters it is insufficiently supported by empirical evidence to be considered an important mechanism driving antigenic evolution of influenza viruses. Thus, natural selection of antigenic variants that most efficiently escape neutralizing antibodies present in the human population remains the most plausible mechanism for major antigenic changes during influenza virus evolution to date.

Predicting influenza virus evolution

Vaccination is the primary method to prevent complications from influenza disease in humans and is an important intervention strategy to reduce the burden of influenza in poultry. The Achilles' heel of vaccination is the highly variable nature of influenza viruses because it allows re-infection of previously immune hosts and is the cause of annually recurring influenza epidemics in humans. Optimal vaccine efficacy requires that the influenza strains included in the vaccine antigenically match the strains that circulate when the vaccine is administered. The time gap between selection of the vaccine strains and administration of the influenza vaccine demands forecasting the antigenic properties of strains that will circulate at least ten months later. Several studies therefore focused on predicting influenza virus evolution using genetic change as basis for prediction (44, 46, 47, 134). These studies often lacked insight in the phenotypic effects of the studied molecular changes. The work described in this thesis contributes to a genotype-to-phenotype translation for antigenic change and provides a rationale for sequence-based detection of escape variants. A currently ongoing follow-up study uses saturated mutagenesis on the A(H3N2) virus key positions to evaluate the effects of different amino acid substitutions on antigenic change and intrinsic fitness and will further support sequence-based identification of escape variants.

An important open question is why key positions are reused within and between subtypes while other bordering positions were not involved in major antigenic change. The key positions among the different influenza viruses are clustered on two main areas bordering the RBS; an antigenic ridge that is formed by the 190 α-helix and part of the 160-loop on the membrane distal part of the RBS, and the protruding 140-loop on the membrane proximal part of the RBS (Fig. 1, Table 1). A single key position, 133, is located in the 130-loop outside of the main areas (H3 numbering, a numbering conversion is provided in Table 1). Some bordering positions may not have contributed to major antigenic change because they are less accessible to antibody interaction than key positions. For example, positions 147 – 149 in the 150-loop and position 96 in the 100-loop are located below the concave structure of the RBS, and are conserved in nearly all A(H3N2) viruses. Positions in the 130-loop (131, 132), position 186 (preceding the 190 α-helix), and positions in the 220-loop (219 – 222) are not hidden. In particular positions 138, 140, and 143 (140-loop), and position 222 are surface exposed, located on protruding structures, and appear equally accessible for antibody interaction to key positions. In chapter 4 we hypothesized that local HA structure may reduce the magnitude of antigenic change caused by a substitution and non-key positions bordering the RBS may have restricted ability to cause structural changes. Finally, the 220-loop holds critical positions for receptor binding and the amino acids in this loop are almost entirely conserved between 1968 and 2003 (24, 31, 95). Substitutions in the 220-loop and other non-key bordering positions that change the antigenic properties may come at too large an intrinsic fitness cost to be compensated. Structural constraints and the necessity to maintain

optimal receptor binding may limit the rate of antigenic evolution as well as the number of amino acid positions that contributed to antigenic change. To address the question if amino acid positions will be added to the set of key positions identified here, future studies should test if substitutions on bordering positions not involved in past antigenic change can cause antibody escape and evaluate their effect on intrinsic fitness compared to substitutions at key positions.

The ferret sera used in HI assays are usually obtained after inoculation of previously naive animals with a single strain of influenza virus. Most humans experienced multiple influenza virus infections and have a more complex antibody repertoire. The current vaccine strain selection procedure uses HI assay data to identify viruses that best represent circulating viruses, but does not take prior immunity into consideration when selecting the vaccine strains. Fonville *et al*. recently introduced a method for the quantitative analysis of antibody-mediated immunity to antigenically variable pathogens (107). The results are presented as antibody landscapes, a visualization of serum antibody titers plotted as a function of antigenic relations between influenza viruses. Antibody landscapes thus provide a means to survey antibody immunity in the human population, allowing a more rational selection of vaccine strains and vaccine composition. Taking population immunity into account may be particularly beneficial for selecting the A(H1N1)pdm09 virus component of the influenza vaccine. We found that mutants that escaped recognition by A(H1N1)pdm09-specific antisera from ferrets and infants did not escape from recognition by antisera collected from individuals that experienced both seasonal and pandemic A(H1N1) virus infections (chapter 5). This observation suggests that antibodies to pre-2009 A(H1N1) and A(H1N1)pdm09 viruses act synergistically and may limit the possibilities of antigenic variants to evade population-level immunity. The absence of cross-reactivity between ferret antisera prepared to seasonal A(H1N1) viruses and A(H1N1)pdm09 viruses poses a challenge to identify antigenic variants that can escape from population-level immunity, because the more broadly reactive antibody response of adult individuals can currently not be represented easily by ferret antisera prepared to A(H1N1)pdm09 viruses (96, 112, 115).

The longitudinal analysis of human sera using antibody landscapes by Fonville *et al*. showed that preemptively updating the vaccine may have the dual advantage of inducing antibodies against both advanced and previously circulating antigenic variants of A(H3N2) virus, which is anticipated to improve vaccine efficacy in much of the population (107). However, their results emphasize that the accuracy of predicting upcoming epidemic strains can be improved by taking the human antibody repertoire into consideration.

The evolutionary dynamics that govern emergence of new antigenic clusters are incompletely understood. We hypothesized that antibody escape comes at an intrinsic fitness cost and that antigenic change is only part of the equation that determines the development of novel epidemic strains (chapter 2). The balance between antibody escape and intrinsic fitness should be further explored to determine if the hypothesized

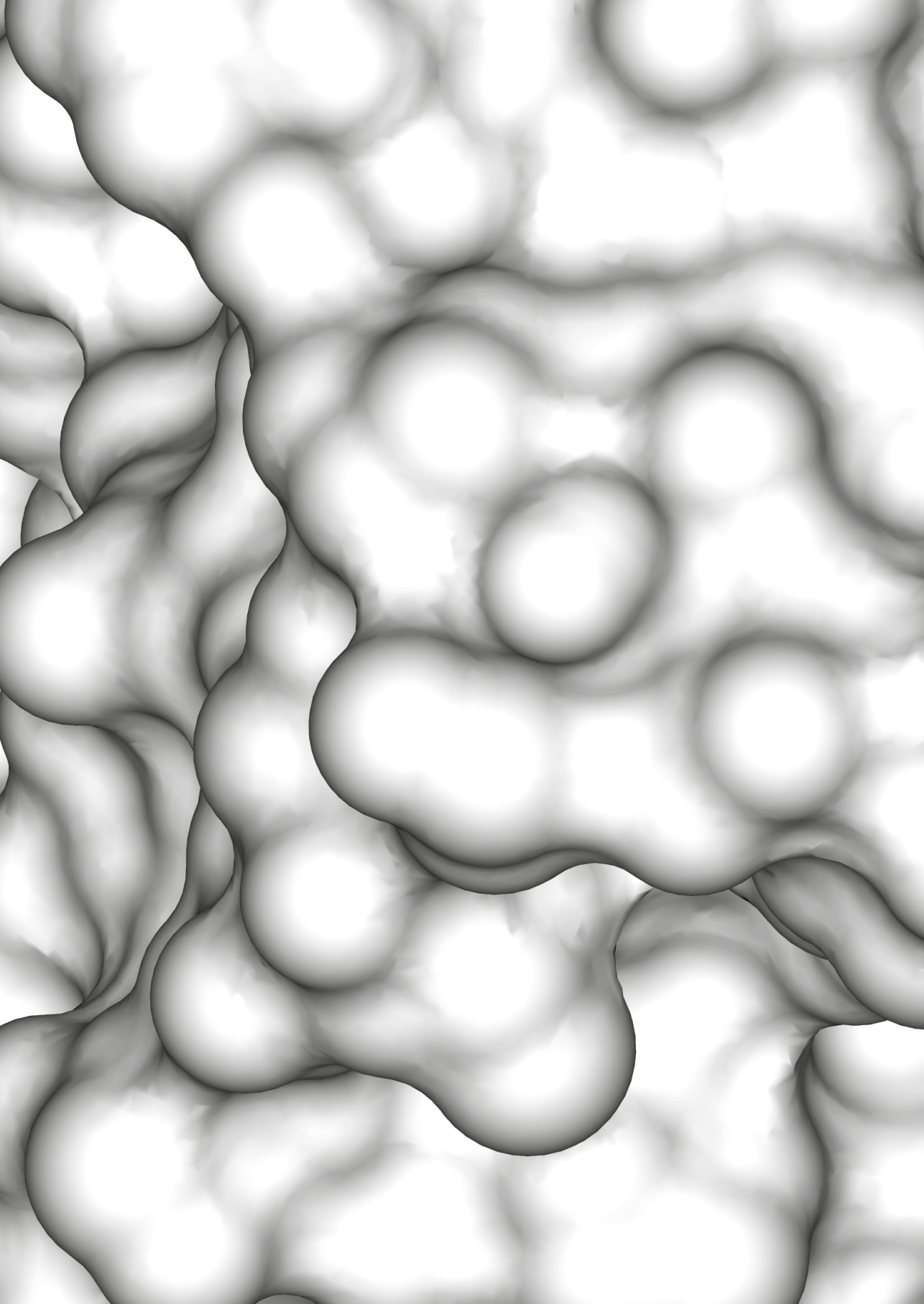
intrinsic fitness cost explains the paradox between fast genetic evolution and slow antigenic change. If true, mapping the HA amino acid changes that offset intrinsic fitness loss caused by antigenic change will help answer the question when a vaccine update should be issued. The opposite functions of HA and NA during replication require that the binding properties of HA are balanced with NA activity (135, 136). Amino acid substitutions in NA or reassortment of HA or NA gene segments from other lineages may compensate for decreased receptor binding caused by antigenic change. Although HA is the most important antigen for neutralizing antibody responses, NA is similarly subject to antigenic change. Although the timing of NA antigenic change is often different from HA, single point mutations can also cause major antigenic change of NA and human sera react differently with strains that have antigenically different NAs (137). Influenza vaccine efficacy could therefore benefit from selecting vaccine strains that have both HAs and NAs that antigenically match epidemic strains.

The genetic composition of seasonal human influenza viruses is usually represented by consensus sequences obtained using Sanger sequencing of clinical isolates propagated on cell cultures or embryonated hen's eggs. Such traditional genetic analyses lack the sensitivity to identify minority variants in the genetically highly diverse viral population in clinical samples and the resulting sequences may be biased towards genetic variants selected for optimal replication in the model system. Next generation sequencing (NGS) methods permit detection of minority variants with population frequencies below 1% and provide a powerful tool to address fundamental questions regarding the development of genetic traits associated with phenotypic change (15, 138-140). NGS methods have already been used to study genetic bottlenecks, immune escape, and intra-host evolutionary dynamics of other RNA viruses (141-143). The high throughput nature of NGS technologies allows full genome sequencing on large clinical sample sets and can be used to study the intra- and inter host genetic variation during and between influenza epidemics. Integrating the knowledge of global circulation patterns (81), population-level immunity, genotype-to-phenotype translations, and large-scale genetic surveillance may eventually allow detection and potentially prediction of variants with epidemic potential before these can be detected using conventional antigenic analyses or Sanger sequencing, with obvious advantages to public health.

SUMMARY

The results presented in this thesis provide new insights in the antigenic evolution of influenza viruses and will help to improve selection of vaccine strains. Rather than the seemingly endless number of amino acid substitutions that influenza viruses can access to escape from antibody-mediated neutralization, we showed that the possibilities for influenza viruses to change antigenically are limited and that genetically vastly different influenza viruses have analogous ways to change antigenically. The small number of critical sites and restricted amino acid usage decimates the number of

possible escape variants from thousands to dozens and thus substantially increases the possibility to predict the antigenic component of influenza virus evolution. Follow-up studies should further define the limitations of influenza virus antigenic change, and should focus on the *in vivo* relevance of the findings presented here. The work for this thesis could form the basis for efforts towards identification of escape variants based on sequence information alone and provides a platform for studies into antigenic evolution of other antigenically variable pathogens.



7

NEDERLANDSE SAMENVATTING

NEDERLANDSE SAMENVATTING

In mensen is influenza (griep) een luchtweginfectie die gepaard gaat met onder andere spierpijn, hoofdpijn, hoge koorts, en hoesten. Complicaties als gevolg van influenza komen het meest voor bij jonge kinderen, ouderen, en mensen met onderliggende medische problemen. Elk jaar ondervinden wereldwijd ongeveer 350 miljoen tot een miljard mensen problemen als gevolg van influenza. Influenzavirussen (griepvirussen) behoren tot de *Orthomyxoviridae* familie van virussen die een gesegmenteerd genoom bezitten dat bestaat uit negatief enkelstrengs RNA en een gesegmenteerd genoom. Influenzavirussen zijn onderverdeeld in types A, B en C. Influenza A virussen zijn verder onderverdeeld in subtypes op basis van de genetische en antigene eigenschappen van de oppervlakte eiwitten hemagglutinine (HA) en neuraminidase (NA). Momenteel zijn er 18 verschillende HAs en 11 verschillende NAs bekend. De combinatie van HA en NA bepaalt het influenza A virus subtype, bijvoorbeeld A(H3N2) virus of A(H5N1) virus. Het overgrote deel van de influenza A virus subtypes circuleert in wilde vogels, maar komt ook voor in onder andere mensen, varkens, paarden, honden, katten, vleermuizen en zeezoogdieren. Alleen van subtypes A(H1N1), A(H2N2), en A(H3N2) is bekend dat ze jaarlijkse griepepidemieën in mensen kunnen veroorzaken. Andere influenza A virus subtypes worden ook in mensen aangetroffen, maar deze zijn tot op heden niet of zeer beperkt tussen mensen overdraagbaar. Meestal zijn dat aviaire influenzavirussen (vogelgriepvirussen) die in pluimvee circuleren zoals A(H5N1) en A(H7N9) virussen. Influenza B virussen zijn in principe specifiek voor de humane gastheer hoewel zij ook in zeehonden zijn aangetroffen. Influenza C virussen infecteren alleen mensen. Deze infectie verloopt meestal mild.

Wanneer een aviair influenza A virus, direct of via een tussengastheer zoals varkens, in de humane populatie terechtkomt en goed tussen mensen overdraagbaar is, kan deze een wereldwijde uitbraak veroorzaken. Een dergelijke wereldwijde uitbraak wordt een pandemie genoemd. In de afgelopen 100 jaar zijn er vier influenza virus pandemieën geweest: A(H1N1) virus in 1918, A(H2N2) virus in 1957, A(H3N2) virus in 1968, en A(H1N1) virus in 2009. Deze laatste wordt aangeduid als A(H1N1)pdm09 virus om onderscheid te maken met de A(H1N1) virussen die voor 2009 circuleerden. Na de pandemieën bleven deze virussen aanwezig in de humane populatie als epidemische influenzavirussen ("seizoens-griepvirussen) en zijn samen met influenza B virussen verantwoordelijk voor de jaarlijks terugkerende griepepidemieën. Momenteel veroorzaken alleen A(H3N2) virussen, A(H1N1)pdm09 virussen, en influenza B virussen griepepidemieën. Sinds 1918 zijn ongeveer evenveel mensen aan de gevolgen van infectie met pandemische en epidemische influenzavirussen overleden.

Het HA oppervlakte eiwit van influenzavirussen is verantwoordelijk voor het binnendringen in de gastheercel. De opname van het virus door receptor-gemedieerde endocytose wordt geïnitieerd door interactie van aminozuren in de receptor bindingsplaats (RBS) van HA met siaalzuren op het oppervlak van de cel. Als onderdeel van de afweerreactie van de gastheer tegen influenzavirussen worden antistoffen gemaakt die gericht zijn tegen de aminozuren op het HA van

het virus waarmee iemand besmet of gevaccineerd is. Antistoffen gericht tegen HA kunnen een influenzavirus neutraliseren, onder andere door de binding van het virus aan de cel te blokkeren en daardoor vermenigvuldiging van het virus voorkomen. Bij een volgende infectie met hetzelfde of een antigeen vergelijkbaar influenzavirus kunnen deze antistoffen zeer efficiënt de vermenigvuldiging van het virus remmen. HA is daarom de belangrijkste component van het influenzavaccin (de "griepprik"). Doordat influenzavirussen regelmatig muteren ontstaan antigene varianten van het virus die kunnen ontsnappen aan immuniteit die is opgebouwd tegen influenzavirussen die eerder circuleerden. Daarom is het noodzakelijk het influenzavaccin periodiek aan te passen zodat de virussen in het vaccin overeenkomen met de virussen die op dat moment circuleren. Het proces van opeenstapeling van mutaties dat leidt tot verandering van de antigene eigenschappen van influenzavirussen wordt antigene drift genoemd.

Om inzicht te krijgen in de veranderingen die bijdragen aan antigene drift zijn al ruim 30 jaar geleden de aminozuurposities in HA in kaart gebracht die geassocieerd worden met ontsnapping aan antistoffen. Deze zeer invloedrijke onderzoeken waren gebaseerd op experimenten waarin influenzavirussen werden vermenigvuldigd in aanwezigheid van monoclonale antistoffen. Alleen virussen met mutaties die leiden tot ontsnapping aan neutralisatie door de monoclonale antistoffen kunnen zich in zulke omstandigheden efficiënt vermenigvuldigen. Een tweede manier om de antigeen belangrijke aminozuurposities van HA te identificeren was het vergelijken van de aminozuren van influenzavirussen die in opeenvolgende jaren epidemieën hadden veroorzaakt. Deze onderzoeken suggereerden dat voor influenza A(H3N2) virussen de aminozuren in vijf antigene regio's verspreid over het HA bijdragen aan antigene variatie. De antigene regio's op het HA van influenza A(H3N2) virussen bestaan uit totaal 131 aminozuurposities. Op basis van deze bevindingen werd aangenomen dat minimaal vier aminozuren verdeeld over twee of meer antigene regio's moeten veranderen om een antigene variant te krijgen die een nieuwe epidemie zou kunnen veroorzaken.

Een in 2004 gepubliceerde studie toonde aan dat influenza A(H3N2) virussen die circuleerden tussen 1968 en 2003 niet geleidelijk antigeen veranderden maar sprongsgewijs. Deze zogenaamde antigene "clusterovergangen" gebeurden gemiddeld elke 3,3 jaar. Gedurende deze periode hebben 11 antigene clusters gecirculeerd en waren er dus tien clusterovergangen. Het onderzoek beschreven in hoofdstuk 2 richtte zich op het identificeren van de aminozuurveranderingen die verantwoordelijk waren voor de clusterovergangen tijdens de evolutie van influenza A(H3N2) virussen. Omdat de aminozuurvolgorde van het HA van virussen die behoorden tot de elf antigene clusters bekend was, konden gericht aminozuurveranderingen in het HA van referentievirussen voor de verschillende antigene clusters worden aangebracht, waardoor de bijdrage van deze veranderingen aan de clusterovergangen kon worden bepaald. Voor de meeste clusterovergangen was een enkele aminozuurverandering voldoende om de antigene verschillen tussen influenzavirussen van opeenvolgende antigene clusters te verklaren. Voor de clusterovergangen waar twee of drie

aminozuurveranderingen voor nodig waren, werd het grootste deel van het antigene verschil ook veroorzaakt door slechts één aminozuurverandering. Alle veranderingen vonden plaats op slechts zeven aminozuurposities vlak naast de RBS. Vijf van de zeven aminozuurposities waren meerdere keren (mede)verantwoordelijk voor het verschil tussen antigene clusters. Antigene varianten met aminozuurveranderingen op de zeven "sleutelposities" werden dus steeds opnieuw door de natuur geselecteerd in plaats van varianten met veranderingen op de overige posities. Deze studie toont aan dat antistoffen gericht tegen de RBS een cruciale rol spelen voor de neutralisatie van influenzavirussen en dat een aminozuurverandering op één van deze posities al voldoende kan zijn om aan antistoffen gericht tegen eerdere virussen te ontsnappen.

Vrijwel alle aminozuurveranderingen verantwoordelijk voor een clusterovergang resulteerden in grote verschillen in de eigenschappen van de aminozuren betrokken bij de verandering. Meestal resulteerde de verandering in een andere lading of een ander volume van de aminozuren. Een specifieke groep aminozuren, de hydrofobe aminozuren, was op één uitzondering na zelfs helemaal niet betrokken bij clusterovergangen. Tussen 1968 en 2003 heeft dus maar een klein deel van de mogelijke aminozuurveranderingen in HA geleid tot grote antigene verschillen, terwijl de overige aminozuurveranderingen (die mogelijk ook kunnen zorgen voor antigene variatie) beperkt effect hadden. De fouten in nieuwe kopieën van de genetische informatie die tijdens influenzavirus-vermenigvuldiging worden gemaakt, leiden ertoe dat gedurende een infectie vrijwel elke mogelijke aminozuurverandering voorkomt. De bevinding dat een enkele aminozuurverandering kan leiden tot een groot antigeen verschil suggereert dus dat antigene varianten van influenza A(H3N2) virussen weliswaar heel vaak ontstaan maar dat de mogelijkheden om uit te groeien tot een epidemische variant beperkt wordt door andere factoren dan alleen de eigenschap om te kunnen ontsnappen aan antistofherkenning.

Ook voor influenza A(H1N1) virussen en influenza B virussen werden de meest recente clusterovergangen veroorzaakt door één aminozuurverandering direct naast de RBS en ook hier resulteerden de aminozuurveranderingen in grote verschillen in de eigenschappen van de aminozuren betrokken bij de antigene verandering. Deze resultaten suggereren dat de bevindingen voor het A(H3N2) virus subtype ook van toepassing zijn op de andere seizoens-griepvirussen.

Indien grote antigene verschillen tussen influenzavirussen veroorzaakt kunnen worden door aminozuurveranderingen op de 131 aminozuurposities in de antigene regio's en elk van de 20 mogelijke aminozuren zou kunnen bijdragen aan de verandering, zouden er vele duizenden mogelijkheden zijn om te ontsnappen aan het immuunsysteem. Het zeer kleine aantal aminozuurposities verantwoordelijk voor grote antigene veranderingen en beperkte gebruik van de potentiële aminozuurvariatie toont aan dat de mogelijkheden voor antigene veranderingen van het HA die leiden tot een antigene clusterovergang beperkt zijn. De antigene evolutie van humane influenzavirussen lijkt dus veel minder gecompliceerd dan algemeen werd aangenomen.

Sinds 2003 zorgen hoogpathogene vogelgriepvirussen van het A(H5N1) subtype in Indonesië voor grote uitbraken in pluimvee. Net als seizoens-griepvirussen muteren A(H5N1) virussen in hoog tempo en is er veel genetische variatie. De A(H5N1) virussen die behoren tot de genetische lijn die wordt aangeduid als clade 2.1 komen uitsluitend voor in Indonesië. Op basis van genetische verschillen zijn clade 2.1 virussen onderverdeeld in clades 2.1.1 – 2.1.3, en is clade 2.1.3 verder onderverdeeld in clades 2.1.3.1 – 2.1.3.3. De snelheid van verandering van A(H5N1) virussen is zo hoog dat deze indeling zeer regelmatig moet worden aangepast. Ondanks de enorme genetische variatie was er weinig bekend over de antigene verschillen tussen clade 2.1 virussen. Om de antigene variatie in kaart te brengen hebben we, zoals beschreven in hoofdstuk 3, eerst de genetische variatie bepaald en referentievirussen geselecteerd die representatief waren voor de genetische variatie. Vervolgens zijn de antigene eigenschappen van de referentievirussen bepaald. Hieruit kon worden geconcludeerd dat er vanaf 2003 minimaal zes antigeen verschillende varianten verantwoordelijk zijn geweest voor uitbraken in pluimvee. Wanneer een nieuwe antigene variant van een seizoens-griepvirus ontstaat in mensen en epidemieën veroorzaakt, vervangt deze normaal gesproken de voorgaande antigene variant van dat subtype. Bij A(H5N1) clade 2.1 virussen in pluimvee bleek dit niet het geval te zijn. Tussen 2003 en 2011 werden pluimvee uitbraken door virussen uit de verschillende genetische lijnen binnen clade 2.1 veroorzaakt en vaak bleven oude en nieuwe antigene varianten naast elkaar circuleren.

De antigene afstand tussen de referentievirussen (een maat voor de antigene verschillen tussen virussen) correleerde slecht met de genetische afstand. Kennis van genetische variatie van clade 2.1 virussen geeft dus zeer beperkt informatie over antigene verschillen. Om inzicht te krijgen in antigene verschillen op basis van genetische informatie is het noodzakelijk de genetische veranderingen te kunnen vertalen naar het effect op de antigene eigenschappen van een virus zoals in hoofdstuk 2 voor seizoens-griepvirussen gedaan is. Om dit te bereiken werden gericht aminozuurveranderingen aangebracht in referentievirussen die de genetische variatie binnen clade 2.1 reflecteerden. Eén tot maximaal vier aminozuurveranderingen vlak naast de RBS waren verantwoordelijk voor de antigene verschillen tussen de referentievirussen. De sleutelposities voor antigene verschillen tussen clade 2.1 virussen vertoonde grote gelijkenis met de sleutelposities van A(H3N2) virussen, A(H1N1) virussen en influenza B virussen. Deze studie impliceert dat de aminozuurveranderingen in het HA van influenzavirussen van verschillende gastheren die kunnen leiden tot ontsnapping aan neutraliserende antistoffen op vergelijkbare manier beperkt wordt.

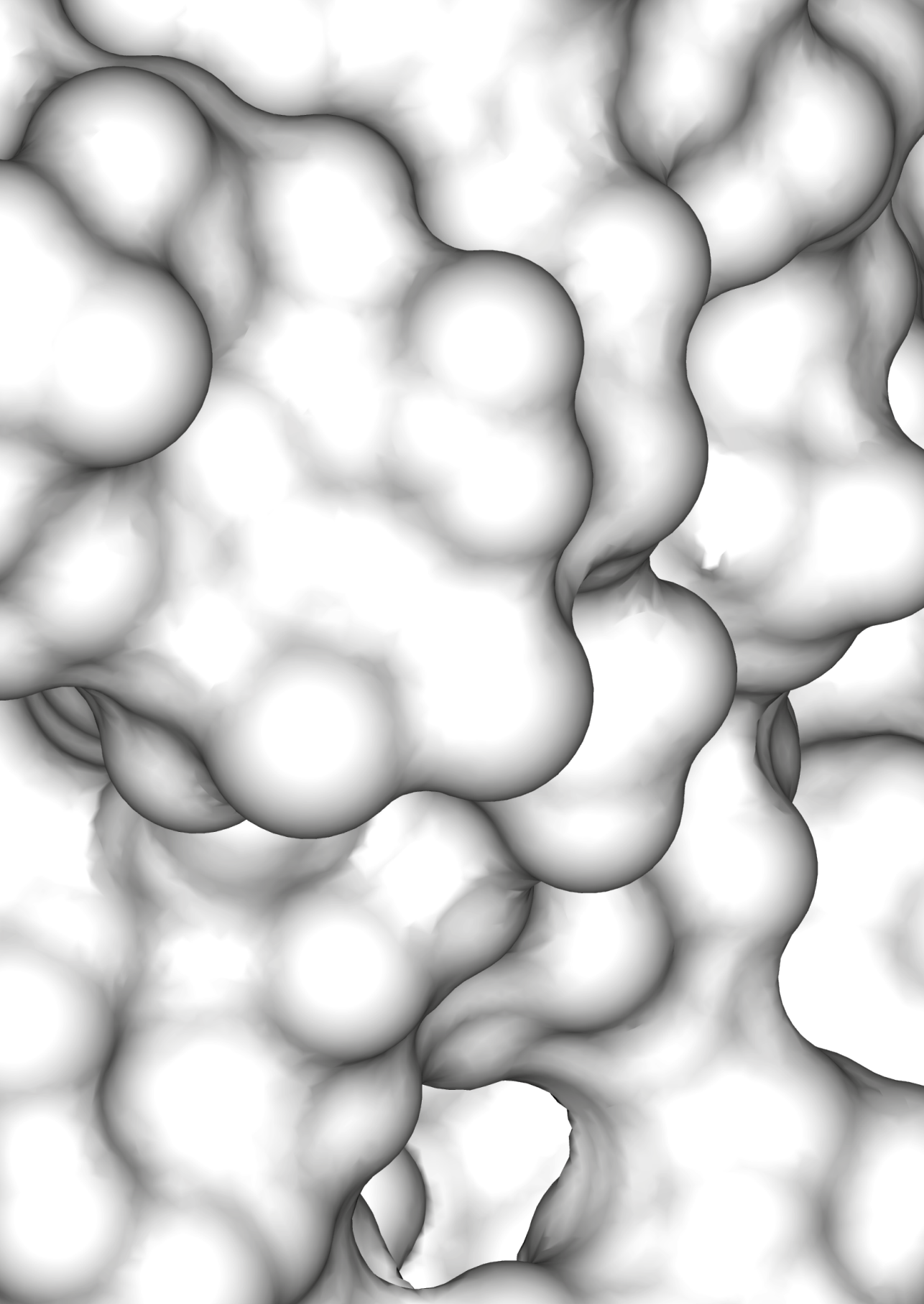
Eén of enkele aminozuurveranderingen naast de RBS kunnen dus al voldoende zijn om een influenzavirus zodanig te veranderen dat deze kan ontsnappen aan antistoffen opgewekt tegen eerder circulerende virussen. In hoofdstuk 4 hebben we getest of het antigene effect van een aminozuurverandering afhankelijk is van de aminozuurcompositie van het HA waarin de verandering plaatsvindt. Om dit te testen

zijn twee verschillende aminozuurveranderingen op verschillende aminozuurposities die verantwoordelijk waren voor clusterovergangen tijdens de evolutie van A(H3N2) virussen onafhankelijk van elkaar aangebracht in de referentievirussen voor de elf antigene clusters. Eén van de twee aminozuurveranderingen, verantwoordelijk voor een clusterovergang na meer dan twee decennia van A(H3N2) virus evolutie, veroorzaakte grote antigene veranderingen in zowel referentievirussen uit de eerste twee decennia van A(H3N2) virus evolutie als in referentievirussen uit de periode daarna. De grootte van de antigene veranderingen was vergelijkbaar met de veranderingen waargenomen voor natuurlijk voorkomende aminozuurveranderingen. Dat deze veranderingen tijdens de eerste twee decennia niet hebben geleid tot een nieuwe epidemische variant, ondersteunt de hypothese uit hoofdstuk 2 dat de mogelijkheden om uit te groeien tot een epidemische variant beperkt wordt door andere factoren dan alleen de eigenschap om te kunnen ontsnappen aan antistofherkenning. Hoewel de antigene veranderingen als gevolg van deze aminozuurverandering duidelijker waren dan de effecten van de verandering op de andere positie, waren de gemuteerde virussen in veruit de meeste gevallen antigeen verschillend van de referentievirussen. Het antigene effect van de veranderingen die hebben geleid tot clusterovergangen lijkt dus grotendeels onafhankelijk te zijn van de aminozuurcompositie van het HA waarin deze plaatsvinden. Deze resultaten benadrukken het belang van de sleutelposities voor antigene variatie van influenzavirussen.

De A(H1N1)pdm09 virussen die sinds 2009 circuleren, zijn antigeen sterk vergelijkbaar met het varkens-griepvirus dat de pandemie veroorzaakte. Omdat na eerdere pandemieën de virussen na verloop van tijd muteerden en konden ontsnappen aan immuniteit tegen de pandemische virussen wordt ook voor A(H1N1)pdm09 virussen verwacht dat er mutaties in het HA zullen plaatsvinden die een aanpassing van het vaccin noodzakelijk maken. Voor het onderzoek beschreven in hoofdstuk 5 zijn aminozuurveranderingen aangebracht in het HA van een A(H1N1)pdm09 virus om aminozuurveranderingen die mogelijk bijdragen aan het ontstaan van antigene varianten in kaart te brengen. De aangebrachte aminozuurveranderingen werden geselecteerd op basis van de kennis uit hoofdstukken 2 en 3, aminozuurveranderingen die van invloed zijn op de binding aan de gastheercel en genetische verschillen tussen A(H1N1)pdm09 virussen, varkens A(H1N1) virussen, en A(H1N1) seizoens-griepvirussen van voor 2009. Ten minste negen enkele aminozuurveranderingen in een klein gebied direct naast de RBS zorgden ervoor dat de gemuteerde virussen konden ontsnappen aan herkenning door antistoffen van fretten en kinderen gericht tegen A(H1N1)pdm09 virussen. Vrijwel al deze "escape" virussen hadden vergelijkbare bindingseigenschappen en vermenigvuldigden minimaal net zo efficiënt als het niet-gemuteerde virus. Op basis van deze resultaten zou het ontstaan van een epidemische variant met een of meer van deze mutaties verwacht kunnen worden. Geen van de mutanten gemaakt voor deze studie kon echter ontsnappen aan herkenning door antistoffen van volwassenen die infecties met zowel A(H1N1) seizoens-griepvirussen als A(H1N1)pdm09 virus hadden doorgemaakt. Deze

resultaten suggereren dat antistoffen gericht tegen de A(H1N1) virussen die voor 2009 epidemieën veroorzaakten, bijdragen aan neutralisatie van antigene varianten van de huidige A(H1N1)pdm09 virussen waardoor het aantal aminozuurveranderingen dat zou kunnen zorgen voor het ontstaan van een antigene variant die kan ontsnappen aan antistoffen in de humane populatie drastisch wordt beperkt.

De resultaten van de studies uitgevoerd voor deze promotie leveren nieuwe inzichten in de antigene evolutie van influenzavirussen. De identificatie van een klein aantal sleutelposities waartegen de meest effectieve antistoffen zijn gericht, kan helpen om de selectie van stammen voor gebruik in het influenza vaccin te verbeteren. In tegenstelling tot wat eerder werd aangenomen zijn de mogelijkheden om te ontsnappen aan antistofherkenning zeer beperkt. Het kleine aantal sleutelposities en beperkt gebruik van de mogelijke aminozuren vermindert het aantal antigene varianten waarvan werd aangenomen dat deze mogelijk kunnen uitgroeien tot een epidemisch influenza virus van vele duizenden tot slechts enkele tientallen. Bovendien hebben genetisch sterk verschillende influenzavirussen vergelijkbare manieren op te ontsnappen aan antistof herkenning. Deze kennis kan bijdragen aan de identificatie van antigene varianten op basis van alleen de genetische informatie van een influenza virus en vergroot de kans dat antigene evolutie van influenzavirussen ooit voorspeld kan worden.



ABOUT THE AUTHOR

Curriculum vitae

The author of this thesis was born on May 1st , 1978 in Den Haag, The Netherlands. After he finished high school at Christelijk Lyceum Delft in 1996, he started his Bachelor of Science (HLO) study at the Rotterdam University of Applied Sciences where he graduated in 2001. He developed an interest in virology during his work as a technician at Crucell B.V., Leiden and in 2005 he started his master study Biomedical Sciences, Immunity and Infection, at Utrecht University. As part of this study he performed a research rotation at the Department of Viroscience of the Erasmus Medical Center in Rotterdam. Here he started as a PhD student in 2007 on a joint project with the Center for Pathogen Evolution of the University of Cambridge under supervision of Prof.Dr. Ron Fouchier, Prof.Dr. Derek Smith, and Prof.Dr. Ab Osterhaus. During this project the molecular basis of influenza virus antigenic change was studied in detail, which resulted in this thesis. Since October 2014 the author of this thesis has worked at the Department of Medical Microbiology of the Academic Medical Centre of the University of Amsterdam where he studies the genetic evolution influenza viruses.

PhD portfolio

Name	Björn Ferry Koel
Erasmus MC department	Department of Viroscience
Research school	Post-graduate Molecular Medicine
PhD period	2007 – 2016
Promotors	Prof.Dr. R.A.M. Fouchier
	Prof.Dr. D.J. Smith
	Prof.Dr. A.D.M.E. Osterhaus

Education

2007 – 2016	PhD program, Erasmus Medical Centre, Rotterdam, The Netherlands. PhD thesis: The Molecular Basis of Influenza Virus Antigenic Change.
2005 – 2007	Master of Science, Utrecht University, Utrecht, The Netherlands. Study: Biomedical Sciences, Immunity and Infection
1996 – 2001	Bachelor of Science (HLO), Rotterdam University of Applied Sciences, Rotterdam, The Netherlands. Study: Chemical laboratory techniques.

Work experience

2001 – 2005	Technician, Crucell B.V., Leiden, The Netherlands

In-depth courses

Course on laboratory animal science (art. 9)	2010
Basic course on R	2010
Bioinformatics Workshop on Virus Evolution and Molecular Epidemiology	2009

Phylogeny and Genetics in microbiology and virology	2008
Course in Virology	2008
Course on Molecular Medicine	2007

Presentations

Oral

5th European Influenza Conference (ESWI), Riga, Latvia	2014
Dutch Annual Virology Symposium, Amsterdam, The Netherlands	2014
18th Molecular Medicine Day, Rotterdam, The Netherlands	2014
4th European Influenza Conference (ESWI), Malta	2011
2nd international influenza meeting, Munster, Germany	2010
Options for the Control of Influenza VII, Hong Kong SAR, China	2010
Dutch Annual Virology Symposium, Amsterdam, The Netherlands	2009
3rd European Influenza Conference (ESWI), Villamoura, Portugal	2008

Poster

6th Orthomyxovirus research conference, Montreal, Canada	2012
16th Molecular Medicine Day, Rotterdam, The Netherlands	2012
4th European Influenza Conference (ESWI), Malta	2011
14th Molecular Medicine Day, Rotterdam, The Netherlands	2010
5th Annual NIH Director's Pioneer Award Symposium, Bethesda, MD, USA	2009
15th International Bioinformatics Workshop on Virus Evolution and Molecular Epidemiology, Rotterdam, The Netherlands	2009
13th Molecular Medicine Day, Rotterdam, The Netherlands	2009

Attended meetings

International One Health Congress, Amsterdam, The Netherlands	2015
Dutch Annual Virology Symposium, Amsterdam, The Netherlands	2015
Molecular Medicine Day, Rotterdam, The Netherlands	2013
Dutch Annual Virology Symposium, Amsterdam, The Netherlands	2012
Dutch Annual Virology Symposium, Amsterdam, The Netherlands	2011
Molecular Medicine Day, Rotterdam, The Netherlands	2011
Dutch Annual Virology Symposium, Amsterdam, The Netherlands	2010

Supervision and teaching activities

Supervision BSc student (HLO)	2012
Supervision BSc student (HLO)	2010
Antigenic cartography trainings	2009 – 2011

Grants and awards

Award for best 2013 publication by a member of the Erasmus Postgraduate school Molecular Medicine	2014
Award for best poster presentation, 6th Orthomyxovirus research conference, Montreal Canada	2012
ESWI young scientist travel grant, 4th European Influenza Conference, Malta	2011
Award for best poster presentation, 14th Molecular Medicine Day, Rotterdam	2010
Award for best poster presentation, 15th International Bioinformatics Workshop on Virus Evolution and Molecular Epidemiology, Rotterdam	2009
ESWI young scientist travel grant, 3th European Influenza Conference, Villamoura, Portugal	2008

List of publications

Koel BF, Mögling R, Chutinimitkul S, Fraaij PL, Burke DF, van der Vliet S, de Wit E, Bestebroer TM, Rimmelzwaan GF, Osterhaus AD, Smith DJ, Fouchier RA, de Graaf M. Identification of amino acid substitutions supporting antigenic change of influenza A(H1N1)pdm09 viruses. *J Virol. 2015 vol. 89 (7) pp. 3763-75*

Koel BF, van der Vliet S, Burke DF, Bestebroer TM, Bharoto EE, Yasa IW, Herliana I, Laksono BM, Xu K, Skepner E, Russell CA, Rimmelzwaan GF, Perez DR, Osterhaus AD, Smith DJ, Prajitno TY, Fouchier RA. Antigenic variation of clade 2.1 H5N1 virus is determined by few amino acid substitutions immediately adjacent to the receptor binding site. *mBio. 2014 vol. 5 no. 3 e01070-14.*

Gilbert M, **Koel BF**, Bestebroer TM, Lewis NS, Smith DJ, Fouchier RA. Serological evidence for non-lethal exposures of Mongolian wild birds to highly pathogenic avian influenza H5N1 virus. *PLoS One. 2014 9(12): e113569*.

Koel BF, Burke DF, Bestebroer TM, van der Vliet S, Zondag GC, Vervaet G, Skepner E, Lewis NS, Spronken MI, Russell CA, Eropkin MY, Hurt AC, Barr IG, de Jong JC, Rimmelzwaan GF, Osterhaus AD, Fouchier RA, Smith DJ. Substitutions near the receptor binding site determine major antigenic change during influenza virus evolution. *Science. 2013 22;342(6161):976-9*

Chutinimitkul S, Herfst S, Steel J, Lowen AC, Ye J, van Riel D, Schrauwen EJ, Bestebroer TM, **Koel BF**, Burke DF, Sutherland-Cash KH, Whittleston CS, Russell CA, Wales DJ, Smith DJ, Jonges M, Meijer A, Koopmans M, Rimmelzwaan GF, Kuiken T, Osterhaus AD, García-Sastre A, Perez DR, Fouchier RA. Virulence-associated substitution D222G

in the hemagglutinin of 2009 pandemic influenza A(H1N1) virus affects receptor binding. *J Virol. 2010 vol. 84 (22) pp. 11802-13*

Gijsbers L, **Koel BF**, Weggeman M, Goudsmit J, Havenga M, Marzio G. Quantification of residual host cell DNA in adenoviral vectors produced on PER.C6 cells. *Hum Gene Ther. 2005 Mar;16(3):393-8*

Sprangers MC, Lakhai W, Koudstaal W, Verhoeven M, **Koel BF**, Vogels R, Goudsmit J, Havenga MJ, Kostense S. Quantifying adenovirus-neutralizing antibodies by luciferase transgene detection: addressing preexisting immunity to vaccine and gene therapy vectors. *J Clin Microbiol. 2003 Nov;41(11):5046-52*

Vogels R, Zuijdgeest D, van Rijnsoever R, Hartkoorn E, Damen I, de Béthune MP, Kostense S, Penders G, Helmus N, Koudstaal W, Cecchini M, Wetterwald A, Sprangers M, Lemckert A, Ophorst O, **Koel BF**, van Meerendonk M, Quax P, Panitti L, Grimbergen J, Bout A, Goudsmit J, Havenga M. Replication-deficient human adenovirus type 35 vectors for gene transfer and vaccination: efficient human cell infection and bypass of preexisting adenovirus immunity. *J Virol. 2003 vol 77 (15) pp 8263-71.*

DANKWOORD

I'm hugely privileged to have been supervised by three of the biggest "flu brains" in recent years. To say that this construction always worked out smoothly is probably an overstatement, but I think in the end it was always to my benefit. It's a no-brainer to choose to work with you again if I could start over! Ron, het gebeurt niet vaak dat ik tegen iemand opkijk. Bij jou is dat letterlijk en figuurlijk wel het geval. Je enthousiasme voor wetenschappelijk onderzoek, inzet voor academische vrijheid, enorme kennis van de moleculaire virologie en gastvrijheid hebben me enorm geïnspireerd en zullen dat in de toekomst zeker ook blijven doen. Derek, for most scientists solely "aiming for the sky" is an unrealistic approach to maintain a scientific career. It's phenomenal how you and Ron managed to tackle several of the big open questions in the flu field during the past decade. I feel very lucky that you gave me the opportunity to witness how this was done, and I'm thankful that you motivated me to "kick my ass into gear" when it was necessary. Ab, het meest geparafraseerde deel uit presentaties op de afdeling komt waarschijnlijk van de acknowledgements slide: "... and of course Ab". Punt. Geen verdere uitleg. Waarschijnlijk komt dit simpelweg omdat het bijna onmogelijk is om te benoemen wat je allemaal voor de afdeling, en dus je AIO's, betekend hebt. Ik sluit me hier graag bij deze uitspraak aan.

Stefan, toen je bij mij begon voor je HLO stage leek je een rustige en misschien zelfs enigszins verlegen student. Gelukkig is dat snel veranderd! Onze samenwerking leek voor mij bijna vanzelf te gaan, we hadden aan een half woord genoeg om elkaar te begrijpen. Het was geweldig om tussendoor ook nog wat gitaaravonden en concerten mee te pakken. Dat moeten we maar blijven volhouden. Ik gun elke AIO een Stefan!

Martin en Marina, de evolutie van collega's naar onze vriendschap is voor mij veel belangrijker dan de evolutie die in dit boekje beschreven staat. Mooie wetenschappelijke discussies, vakanties, meren in Zweden en Engeland, Vulva's en Ficussen, zondagochtendkoffie, kanoën, barbecues: ik heb er van genoten. Het was behoorlijk afkicken toen jullie naar Singapore verhuisden, maar gelukkig blijkt dat ook de fysieke locatie van vrienden ondergeschikt is aan de snelheid van hun internet verbinding. Theo, toegang hebben tot alle kennis en praktische skills van iemand als jij is goud waard. En dan heb ik het nog niet eens over je frituurkunsten. Guus (de-prof-met-de-chronische-glimlach), dank voor je hulp bij m'n manuscripten en mooie gesprekken. Miranda, mede dankzij onze goede samenwerking is Journal of Virology toch gelukt voor het H1 manuscript! Succes met je norovirus onderzoek! M'n buurtjes, Monique, Bernike, Ramona, en Josanne, dank voor alle adviezen en het aanhoren van m'n gezwets (en soms frustraties). Pascal, de beste gesprekken zijn vrijwel altijd tijdens het roken. Dat heeft zeker ook met de mensen met wie ik ging roken te maken.

Sander, Eefje, Kim, Mathilde, Debby, Dennis, Joost, Rogier, Ben, Pascal B, Salin, Sander Bo, Carolien, Jurre, Chantal, Rory, Bri, Oanh, Vincent, Emmie, Patrick, Leslie, Erhard, Marco, Lennert, het cliché dat wetenschappers saaie vakidioten zijn is voor jullie ook deels waar: het vakidioten deel. Dank voor alle hulp en samenwerkingen.

Met Sinterkerst, voetballen tussen de liften, stappen na een congresdag, barbecues in het park, mannen-weekend in Keulen, borrels, enz enz, hebben jullie (en vele anderen van de afdeling) laten zien dat het "saai" deel dikke onzin is. Myth busted!

David, Colin, Nic (the magenta spots before my eyes have finally disappeared!), Eu (aka the Eugenes), Judy, Ana, Leah, Terry, Gene, Anna, Dan, Chris: I think it's fair to say that, although all the practical work was done in Rotterdam, much of the theoretical basis for the work in this thesis comes out of Cambridge. I loved the great discussions, the atmosphere, and having a proper sized beer at the end of the day. I always looked forward coming over to Cambridge!

Jan de Jong, ik heb heel erg veel van je geleerd over de evolutie van influenzavirussen. Ik mis onze wekelijkse gesprekken! Ruud, mijn redder als ik weer eens referentievirussen of sera nodig had. Je bent een held!

Pieter, Bernadette, Tiny, Gerrie, Bart, Robert (meestal had je het gewoon wel!), Rik, Byron, Georges, Penelope, Geert, Rob, David, bedankt voor alle support en gezelligheid! En uiteraard ook de dames van het secretariaat, Maria, Simone, Loubna, Anouk, Sabine. Jullie zijn een superteam, dank voor alle hulp! Een eervolle vermelding voor The Spikes: Martin, Stefan, Gijs, Maarten, Sabine, Carel, Rob, Marleen, Monique, Esther, Mathilde, Ingrid, waarschijnlijk klonk het voor geen meter wat we deden maar daar gaat het niet om. We hebben wel lol gehad, rock on!

Piet, op dinsdagavond naar beneden lopen om te roken was eigenlijk een excuus om met jou een praatje te maken over vliegvissen. Kees, dank voor het proofreaden van m'n Nederlandse samenvatting en CV!

Robert (de noordelijke tak van de familie Koel), Sven en Kjeld (echte broers), Martijn (sinds het HLO m'n broer), Sebas, Andries, Henk, John en Gep, ontspannen na hard werk gaat het best in gezelschap van vrienden, familie, en buren, en met gitaren, goede muziek, kampvuur, en een kaartspel. Oh, en copieuze hoeveelheden alcohol natuurlijk.

Pap en Mam, ik kan kort zijn: van iedereen in dit dankwoord verdienen jullie de meeste waardering. Zonder de hulp, steun, en goede zorgen die ik altijd van jullie heb gekregen was er van dit proefschrift geen letter op papier gekomen. Jullie zijn fantastisch!

Als laatste natuurlijk nog m'n lievelingsmensen. Bregje, je kan je als geen ander inleven in anderen, een ongelooflijk mooie eigenschap waar ik heel veel van kan leren. Je bent m'n meisje, thuisbasis, raadgever, en oplader. Je bent er altijd voor me, ik voel me veilig en gelukkig bij jou. Ebbe, jochie, je hebt me vaak van m'n werk gehouden. Daar mag je nooit mee ophouden, ik geniet ervan. Ik hou van jullie!

REFERENCE LIST

1. Collin EA, Sheng Z, Lang Y, Ma W, Hause BM, Li F. 2014. Co-circulation of two distinct genetic and antigenic lineages of proposed influenza D virus in cattle. J Virol.
2. Fouchier RA, Munster V, Wallensten A, Bestebroer TM, Herfst S, Smith D, Rimmelzwaan GF, Olsen B, Osterhaus AD. 2005. Characterization of a novel influenza A virus hemagglutinin subtype (H16) obtained from black-headed gulls. J Virol 79:2814-2822.
3. Tong S, Li Y, Rivailler P, Conrardy C, Castillo DA, Chen LM, Recuenco S, Ellison JA, Davis CT, York IA, Turmelle AS, Moran D, Rogers S, Shi M, Tao Y, Weil MR, Tang K, Rowe LA, Sammons S, Xu X, Frace M, Lindblade KA, Cox NJ, Anderson LJ, Rupprecht CE, Donis RO. 2012. A distinct lineage of influenza A virus from bats. Proc Natl Acad Sci U S A 109:4269-4274.
4. Tong S, Zhu X, Li Y, Shi M, Zhang J, Bourgeois M, Yang H, Chen X, Recuenco S, Gomez J, Chen LM, Johnson A, Tao Y, Dreyfus C, Yu W, McBride R, Carney PJ, Gilbert AT, Chang J, Guo Z, Davis CT, Paulson JC, Stevens J, Rupprecht CE, Holmes EC, Wilson IA, Donis RO. 2013. New world bats harbor diverse influenza A viruses. PLoS Pathog 9:e1003657.
5. Webster RG, Bean WJ, Gorman OT, Chambers TM, Kawaoka Y. 1992. Evolution and ecology of influenza A viruses. Microbiol Rev 56:152-179.
6. Osterhaus AD, Rimmelzwaan GF, Martina BE, Bestebroer TM, Fouchier RA. 2000. Influenza B virus in seals. Science 288:1051-1053.
7. Rota PA, Wallis TR, Harmon MW, Rota JS, Kendal AP, Nerome K. 1990. Cocirculation of two distinct evolutionary lineages of influenza type B virus since 1983. Virology 175:59-68.
8. Johnson NP, Mueller J. 2002. Updating the accounts: global mortality of the 1918-1920 "Spanish" influenza pandemic. Bull Hist Med 76:105-115.
9. Kilbourne ED. 2006. Influenza pandemics of the 20th century. Emerg Infect Dis 12:9-14.
10. Dawood FS, Iuliano AD, Reed C, Meltzer MI, Shay DK, Cheng PY, Bandaranayake D, Breiman RF, Brooks WA, Buchy P, Feikin DR, Fowler KB, Gordon A, Hien NT, Horby P, Huang QS, Katz MA, Krishnan A, Lal R, Montgomery JM, Molbak K, Pebody R, Presanis AM, Razuri H, Steens A, Tinoco YO, Wallinga J, Yu H, Vong S, Bresee J, Widdowson MA. 2012. Estimated global mortality associated with the first 12 months of 2009 pandemic influenza A H1N1 virus circulation: a modelling study. Lancet Infect Dis 12:687-695.
11. Stohr K. 2002. Influenza--WHO cares. Lancet Infect Dis 2:517.
12. de Jong JC, Claas EC, Osterhaus AD, Webster RG, Lim WL. 1997. A pandemic warning? Nature 389:554.
13. Claas EC, Osterhaus AD, van Beek R, De Jong JC, Rimmelzwaan GF, Senne DA, Krauss S, Shortridge KF, Webster RG. 1998. Human influenza A H5N1 virus related to a highly pathogenic avian influenza virus. Lancet 351:472-477.
14. World Health Organization. 2014. Influenza at the human-animal interface: Summary and assessment as of 24 January 2014. Last update: January 2014. World Health Organization, Geneva, Switzerland. http://www.who.int/influenza/human_animal_interface/Influenza_Summary_IRA_HA_interface_24January14.pdf
15. Herfst S, Schrauwen EJ, Linster M, Chutinimitkul S, de Wit E, Munster VJ, Sorrell EM, Bestebroer TM, Burke DF, Smith DJ, Rimmelzwaan GF, Osterhaus AD, Fouchier RA. 2012. Airborne transmission of influenza A/H5N1 virus between ferrets. Science 336:1534-1541.
16. Imai M, Watanabe T, Hatta M, Das SC, Ozawa M, Shinya K, Zhong G, Hanson A, Katsura H, Watanabe S, Li C, Kawakami E, Yamada S, Kiso M, Suzuki Y, Maher EA, Neumann G, Kawaoka Y. 2012. Experimental adaptation of an influenza H5 HA confers respiratory droplet transmission to a reassortant H5 HA/H1N1 virus in ferrets. Nature 486:420-428.

17. WHO/OIE/FAO H5N1 Evolution Working Group. 2008. Toward a unified nomenclature system for highly pathogenic avian influenza virus (H5N1). Emerg Infect Dis. 14:e1.

18. WHO/OIE/FAO H5N1 Evolution Working Group. 2009. Continuing progress towards a unified nomenclature for the highly pathogenic H5N1 avian influenza viruses: divergence of clade 2.2 viruses. Influenza Other Respi Viruses. 3:59-62.

19. WHO/OIE/FAO H5N1 Evolution Working Group. 2012. Continued evolution of highly pathogenic avian influenza A (H5N1): updated nomenclature. Influenza Other Respi Viruses 6:1-5.

20. Shepard SS, Davis CT, Bahl J, Rivailler P, York IA, Donis RO. 2014. LABEL: fast and accurate lineage assignment with assessment of H5N1 and H9N2 influenza A hemagglutinins. PLoS One 9:e86921.

21. Richard M, de Graaf M, Herfst S. 2014. Avian influenza A viruses: from zoonosis to pandemic. Future Virol 9:513-524.

22. World Health Organization. 2015. Situation updates - Avian influenza. Last update: May 2015. World Health Organization, Geneva, Switzerland. http://www.who.int/influenza/human_animal_interface/avian_influenza/archive/en/

23. Pantin-Jackwood MJ, Miller PJ, Spackman E, Swayne DE, Susta L, Costa-Hurtado M, Suarez DL. 2014. Role of poultry in the spread of novel H7N9 influenza virus in China. J Virol 88:5381-5390.

24. Skehel JJ, Wiley DC. 2000. Receptor binding and membrane fusion in virus entry: the influenza hemagglutinin. Annu Rev Biochem 69:531-569.

25. Ha Y, Stevens DJ, Skehel JJ, Wiley DC. 2001. X-ray structures of H5 avian and H9 swine influenza virus hemagglutinins bound to avian and human receptor analogs. Proc Natl Acad Sci U S A 98:11181-11186.

26. Stevens J, Blixt O, Glaser L, Taubenberger JK, Palese P, Paulson JC, Wilson IA. 2006. Glycan microarray analysis of the hemagglutinins from modern and pandemic influenza viruses reveals different receptor specificities. J Mol Biol 355:1143-1155.

27. Rogers GN, Paulson JC. 1983. Receptor determinants of human and animal influenza virus isolates: differences in receptor specificity of the H3 hemagglutinin based on species of origin. Virology 127:361-373.

28. Huber VC, Lynch JM, Bucher DJ, Le J, Metzger DW. 2001. Fc receptor-mediated phagocytosis makes a significant contribution to clearance of influenza virus infections. J Immunol 166:7381-7388.

29. Renegar KB, Small PA, Jr. 1991. Passive transfer of local immunity to influenza virus infection by IgA antibody. J Immunol 146:1972-1978.

30. de Jong JC, Smith DJ, Lapedes AS, Donatelli I, Campitelli L, Barigazzi G, Van Reeth K, Jones TC, Rimmelzwaan GF, Osterhaus AD, Fouchier RA. 2007. Antigenic and genetic evolution of swine influenza A (H3N2) viruses in Europe. J Virol 81:4315-4322.

31. Smith DJ, Lapedes AS, de Jong JC, Bestebroer TM, Rimmelzwaan GF, Osterhaus ADME, Fouchier RAM. 2004. Mapping the antigenic and genetic evolution of influenza virus. Science 305:371-376.

32. Wilson IA, Cox NJ. 1990. Structural basis of immune recognition of influenza virus hemagglutinin. Annu Rev Immunol 8:737-771.

33. Wiley DC, Wilson IA, Skehel JJ. 1981. Structural identification of the antibody-binding sites of Hong Kong influenza haemagglutinin and their involvement in antigenic variation. Nature 289:373-378.

34. Caton AJ, Brownlee GG, Yewdell JW, Gerhard W. 1982. The antigenic structure of the influenza virus A/PR/8/34 hemagglutinin (H1 subtype). Cell 31:417-427.

35. Kaverin NV, Rudneva IA, Govorkova EA, Timofeeva TA, Shilov AA, Kochergin-Nikitsky KS, Krylov PS, Webster RG. 2007. Epitope mapping of the hemagglutinin molecule of a highly pathogenic H5N1 influenza virus by using monoclonal antibodies. J Virol 81:12911-12917.

36. Kaverin NV, Rudneva IA, Ilyushina NA, Varich NL, Lipatov AS, Smirnov YA, Govorkova EA, Gitelman AK, Lvov DK, Webster RG. 2002. Structure of antigenic sites on the haemagglutinin molecule of H5 avian influenza virus and phenotypic variation of escape mutants. J Gen Virol 83:2497-2505.

37. Hirst GK. 1943. Studies of antigenic differences among strains of influenza a by means of red cell agglutination. J Exp Med 78:407-423.

38. Couch RB, Kasel JA. 1983. Immunity to influenza in man. Annu Rev Microbiol 37:529-549.

39. Hobson D, Curry RL, Beare AS, Ward-Gardner A. 1972. The role of serum haemagglutination-inhibiting antibody in protection against challenge infection with influenza A2 and B viruses. J Hyg (Lond) 70:767-777.

40. Virelizier JL. 1975. Host defenses against influenza virus: the role of anti-hemagglutinin antibody. J Immunol 115:434-439.

41. Both GW, Sleigh MJ, Cox NJ, Kendal AP. 1983. Antigenic drift in influenza virus H3 hemagglutinin from 1968 to 1980: multiple evolutionary pathways and sequential amino acid changes at key antigenic sites. J Virol 48:52-60.

42. Fitch WM, Leiter JM, Li XQ, Palese P. 1991. Positive Darwinian evolution in human influenza A viruses. Proc Natl Acad Sci USA 88:4270-4274.

43. Fitch WM, Bush RM, Bender CA, Cox NJ. 1997. Long term trends in the evolution of H(3) HA1 human influenza type A. Proc Natl Acad Sci USA 94:7712-7718.

44. Bush RM, Bender CA, Subbarao K, Cox NJ, Fitch WM. 1999. Predicting the evolution of human influenza A. Science 286:1921-1925.

45. Bush RM, Fitch WM, Bender CA, Cox NJ. 1999. Positive selection on the H3 hemagglutinin gene of human influenza virus A. Mol Biol Evol 16:1457-1465.

46. Kosakovsky Pond SL, Poon AF, Leigh Brown AJ, Frost SD. 2008. A maximum likelihood method for detecting directional evolution in protein sequences and its application to influenza A virus. Mol Biol Evol 25:1809-1824.

47. Sun H, Yang J, Zhang T, Long LP, Jia K, Yang G, Webby RJ, Wan XF. 2013. Using sequence data to infer the antigenicity of influenza virus. MBio 4.

48. Jin H, Zhou H, Liu H, Chan W, Adhikary L, Mahmood K, Lee M-S, Kemble G. 2005. Two residues in the hemagglutinin of A/Fujian/411/02-like influenza viruses are responsible for antigenic drift from A/Panama/2007/99. Virology 336:113-119.

49. McDonald NJ, Smith CB, Cox NJ. 2007. Antigenic drift in the evolution of H1N1 influenza A viruses resulting from deletion of a single amino acid in the haemagglutinin gene. J Gen Virol 88:3209-3213.

50. Karlsson Hedestam GB, Fouchier RA, Phogat S, Burton DR, Sodroski J, Wyatt RT. 2008. The challenges of eliciting neutralizing antibodies to HIV-1 and to influenza virus. Nat Rev Microbiol 6:143-155.

51. Akazawa M, Sindelar JL, Paltiel AD. 2003. Economic costs of influenza-related work absenteeism. Value Health 6:107-115.

52. Sui J, Hwang WC, Perez S, Wei G, Aird D, Chen L-m, Santelli E, Stec B, Cadwell G, Ali M, Wan H, Murakami A, Yammanuru A, Han T, Cox NJ, Bankston LA, Donis RO, Liddington RC, Marasco WA. 2009. Structural and functional bases for broad-spectrum neutralization of avian and human influenza A viruses. Nat Struct Mol Biol 16:265-273.

53. Wang TT, Tan GS, Hai R, Pica N, Petersen E, Moran TM, Palese P. 2010. Broadly protective monoclonal antibodies against H3 influenza viruses following sequential immunization with different hemagglutinins. PLoS Pathog 6:e1000796.

54. Steel J, Lowen AC, Wang T, Yondola M, Gao Q, Haye K, Garcia-Sastre A, Palese P. 2010. Influenza virus vaccine based on the conserved hemagglutinin stalk domain. MBio 1.

55. Ekiert DC, Friesen RHE, Bhabha G, Kwaks T, Jongeneelen M, Yu W, Ophorst C, Cox F, Korse HJWM, Brandenburg B, Vogels R, Brakenhoff JPJ, Kompier R, Koldijk MH, Cornelissen LAHM, Poon LLM, Peiris M, Koudstaal W, Wilson IA, Goudsmit J. 2011. A highly conserved neutralizing epitope on group 2 influenza A viruses. Science 333:843-850.

56. Koelle K, Cobey S, Grenfell B, Pascual M. 2006. Epochal evolution shapes the phylodynamics of interpandemic influenza A (H3N2) in humans. Science 314:1898-1903.

57. Hensley SE, Das SR, Bailey AL, Schmidt LM, Hickman HD, Jayaraman A, Viswanathan K, Raman R, Sasisekharan R, Bennink JR, Yewdell JW. 2009. Hemagglutinin Receptor Binding Avidity Drives Influenza A Virus Antigenic Drift. Science 326:734-736.

58. Zhai W, Slatkin M, Nielsen R. 2007. Exploring variation in the d(N)/d(S) ratio among sites and lineages using mutational mappings: applications to the influenza virus. J Mol Evol 65:340-348.

59. Information on materials and methods is available on Science Online.

60. de Wit E, Spronken MIJ, Vervaet G, Rimmelzwaan GF, Osterhaus ADME, Fouchier RAM. 2007. A reverse-genetics system for Influenza A virus using T7 RNA polymerase. J Gen Virol 88:1281-1287.

61. Skehel JJ, Stevens DJ, Daniels RS, Douglas AR, Knossow M, Wilson IA, Wiley DC. 1984. A carbohydrate side chain on hemagglutinins of Hong Kong influenza viruses inhibits recognition by a monoclonal antibody. Proc Natl Acad Sci U S A 81:1779-1783.

62. Abe Y, Takashita E, Sugawara K, Matsuzaki Y, Muraki Y, Hongo S. 2004. Effect of the addition of oligosaccharides on the biological activities and antigenicity of influenza A/H3N2 virus hemagglutinin. J Virol 78:9605-9611.

63. Blackburne BP, Hay AJ, Goldstein RA. 2008. Changing selective pressure during antigenic changes in human influenza H3. PLoS Pathog 4:e1000058.

64. de Wit E, Spronken MI, Bestebroer TM, Rimmelzwaan GF, Osterhaus AD, Fouchier RA. 2004. Efficient generation and growth of influenza virus A/PR/8/34 from eight cDNA fragments. Virus Res 103:155-161.

65. Hoffmann E, Neumann G, Kawaoka Y, Hobom G, Webster RG. 2000. A DNA transfection system for generation of influenza A virus from eight plasmids. Proc Natl Acad Sci U S A 97:6108-6113.

66. Rimmelzwaan GF, Baars M, van Beek R, de Lijster P, de Jong JC, Claas EC, Osterhaus AD. 1999. Influenza virus subtype cross-reactivities of haemagglutination inhibiting and virus neutralising serum antibodies induced by infection or vaccination with an ISCOM-based vaccine. Vaccine 17:2512-2516.

67. Matrosovich M, Matrosovich T, Garten W, Klenk HD. 2006. New low-viscosity overlay medium for viral plaque assays. Virol J 3:63.

68. Lam TT, Hon CC, Lemey P, Pybus OG, Shi M, Tun HM, Li J, Jiang J, Holmes EC, Leung FC. 2012. Phylodynamics of H5N1 avian influenza virus in Indonesia. Mol Ecol 21:3062-3077.

69. Smith GJ, Naipospos TS, Nguyen TD, de Jong MD, Vijaykrishna D, Usman TB, Hassan SS, Nguyen TV, Dao TV, Bui NA, Leung YH, Cheung CL, Rayner JM, Zhang JX, Zhang LJ, Poon LL, Li KS, Nguyen VC, Hien TT, Farrar J, Webster RG, Chen H, Peiris JS, Guan Y. 2006. Evolution and adaptation of H5N1 influenza virus in avian and human hosts in Indonesia and Vietnam. Virology 350:258-268.

70. Rahma SM. 2010. The Role of NGOs in Combating Avian Influenza in Indonesia: A Muhammadiyah Case Study,. Singapore: RSIS Centre for Non-Traditional Security (NTS) Studies.

71. Kandun IN, Wibisono H, Sedyaningsih ER, Yusharmen, Hadisoedarsuno W, Purba W, Santoso H, Septiawati C, Tresnaningsih E, Heriyanto B, Yuwono D, Harun S, Soeroso S, Giriputra S, Blair PJ, Jeremijenko A, Kosasih H, Putnam SD, Samaan G, Silitonga M, Chan KH, Poon LL, Lim W, Klimov

A, Lindstrom S, Guan Y, Donis R, Katz J, Cox N, Peiris M, Uyeki TM. 2006. Three Indonesian clusters of H5N1 virus infection in 2005. N Engl J Med 355:2186-2194.

72. World Health Organization. Cumulative number of confirmed human cases for avian influenza A(H5N1) reported to WHO, 2003-2013. Last update: July 2013 http://www.who.int/influenza/human_animal_interface/EN_GIP_20130705CumulativeNumberH5N1cases_2.pdf
73. Food and Agriculture Organization of the United Nations. 2010. OFFLU assists Indonesia to combat avian influenza. Last update: June 2010 http://www.fao.org/avianflu/en/news/indonesia_OFFLU_hpai.html
74. B.Sumiarto BA. 2008. Overview on poultry sector and HPAI situation for Indonesia with special emphasis on the island of Java 3. International Food Policy Research Institute (IFPRI).
75. World Health Organization. 2011. Updated unified nomenclature system for the highly pathogenic H5N1 avian influenza viruses. Last update: October 2011 http://www.who.int/influenza/gisrs_laboratory/h5n1_nomenclature/en/index.html
76. Cattoli G, Milani A, Temperton N, Zecchin B, Buratin A, Molesti E, Aly MM, Arafa A, Capua I. 2011. Antigenic drift in H5N1 avian influenza virus in poultry is driven by mutations in major antigenic sites of the hemagglutinin molecule analogous to those for human influenza virus. J Virol 85:8718-8724.
77. Koel BF, Burke DF, Bestebroer TM, van der Vliet S, Zondag GC, Vervaet G, Skepner E, Lewis NS, Spronken MI, Russell CA, Eropkin MY, Hurt AC, Barr IG, de Jong JC, Rimmelzwaan GF, Osterhaus AD, Fouchier RA, Smith DJ. 2013. Substitutions near the receptor binding site determine major antigenic change during influenza virus evolution. Science 342:976-979.
78. Wibawa H, Henning J, Wong F, Selleck P, Junaidi A, Bingham J, Daniels P, Meers J. 2011. A molecular and antigenic survey of H5N1 highly pathogenic avian influenza virus isolates from smallholder duck farms in Central Java, Indonesia during 2007-2008. Virol J 8:425.
79. World Health Organization. 2013. Summary of status of development and availability of A(H5N1) candidate vaccine viruses and potency testing reagents. Last update: January 2013. http://www.who.int/influenza/vaccines/virus/candidates_reagents/summary_a_h5n1_cvv_20130124.pdf
80. Watanabe Y, Ibrahim MS, Ellakany HF, Kawashita N, Mizuike R, Hiramatsu H, Sriwilaijaroen N, Takagi T, Suzuki Y, Ikuta K. 2011. Acquisition of human-type receptor binding specificity by new H5N1 influenza virus sublineages during their emergence in birds in Egypt. PLoS Pathog 7:e1002068.
81. Russell CA, Jones TC, Barr IG, Cox NJ, Garten RJ, Gregory V, Gust ID, Hampson AW, Hay AJ, Hurt AC, de Jong JC, Kelso A, Klimov AI, Kageyama T, Komadina N, Lapedes AS, Lin YP, Mosterin A, Obuchi M, Odagiri T, Osterhaus ADME, Rimmelzwaan GF, Shaw MW, Skepner E, Stohr K, Tashiro M, Fouchier RAM, Smith DJ. 2008. The global circulation of seasonal influenza A (H3N2) viruses. Science 320:340-346.
82. Bett B, McLaws M, Jost C, Schoonman L, Unger F, Poole J, Lapar ML, Siregar ES, Azhar M, Hidayat MM, Dunkle SE, Mariner J. 2013. The Effectiveness of Preventative Mass Vaccination Regimes Against the Incidence of Highly Pathogenic Avian Influenza on Java Island, Indonesia. Transbound Emerg Dis.
83. Jadhao SJ, Lee CW, Sylte M, Suarez DL. 2009. Comparative efficacy of North American and antigenically matched reverse genetics derived H5N9 DIVA marker vaccines against highly pathogenic Asian H5N1 avian influenza viruses in chickens. Vaccine 27:6247-6260.
84. Cattoli G, Fusaro A, Monne I, Coven F, Joannis T, El-Hamid HS, Hussein AA, Cornelius C, Amarin NM, Mancin M, Holmes EC, Capua I. 2011. Evidence for differing evolutionary dynamics of A/H5N1 viruses among countries applying or not applying avian influenza vaccination in poultry. Vaccine 29:9368-9375.

85. World Health Organization. 2012. Antigenic and genetic characteristics of zoonotic influenza viruses and development of candidate vaccine viruses for pandemic preparedness. World Health Organization. Last update: September 2015 http://www.who.int/influenza/vaccines/virus/201509_zoonotic_vaccinevirusupdate.pdf?ua=1

86. Swofford D. 2003. PAUP*: Phylogenetic analysis using parsimony (*and other methods). Sunderland (Massachusetts): Sinauer Associates 4.0 ed.

87. Posada D, Crandall KA. 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14:817-818.

88. Zhang W, Shi Y, Lu X, Shu Y, Qi J, Gao GF. 2013. An airborne transmissible avian influenza H5 hemagglutinin seen at the atomic level. Science 340:1463-1467.

89. Sorrell EM, Schrauwen EJ, Linster M, De Graaf M, Herfst S, Fouchier RA. 2011. Predicting 'airborne' influenza viruses: (trans-) mission impossible? Curr Opin Virol 1:635-642.

90. Nakajima K, Desselberger U, Palese P. 1978. Recent human influenza A (H1N1) viruses are closely related genetically to strains isolated in 1950. Nature 274:334-339.

91. Dawood FS, Jain S, Finelli L, Shaw MW, Lindstrom S, Garten RJ, Gubareva LV, Xu X, Bridges CB, Uyeki TM. 2009. Emergence of a novel swine-origin influenza A (H1N1) virus in humans. N Engl J Med 360:2605-2615.

92. World Health Organization. 2014. Recommended composition of influenza virus vaccines for use in the 2014-2015 northern hemisphere influenza season. Last update: Februari 2014 http://www.who.int/influenza/vaccines/virus/recommendations/201402_recommendation.pdf?ua=1

93. Gerhard W, Yewdell J, Frankel ME, Webster R. 1981. Antigenic structure of influenza virus haemagglutinin defined by hybridoma antibodies. Nature 290:713-717.

94. Koel BF, van der Vliet S, Burke DF, Bestebroer TM, Bharoto EE, Yasa IW, Herliana I, Laksono BM, Xu K, Skepner E, Russell CA, Rimmelzwaan GF, Perez DR, Osterhaus AD, Smith DJ, Prajitno TY, Fouchier RA. 2014. 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.

95. de Graaf M, Fouchier RA. 2014. Role of receptor binding specificity in influenza A virus transmission and pathogenesis. Embo J 33:823-841.

96. Garten RJ, Davis CT, Russell CA, Shu B, Lindstrom S, Balish A, Sessions WM, Xu X, Skepner E, Deyde V, Okomo-Adhiambo M, Gubareva L, Barnes J, Smith CB, Emery SL, Hillman MJ, Rivailler P, Smagala J, de Graaf M, Burke DF, Fouchier RA, Pappas C, Alpuche-Aranda CM, Lopez-Gatell H, Olivera H, Lopez I, Myers CA, Faix D, Blair PJ, Yu C, Keene KM, Dotson PD, Jr., Boxrud D, Sambol AR, Abid SH, St George K, Bannerman T, Moore AL, Stringer DJ, Blevins P, Demmler-Harrison GJ, Ginsberg M, Kriner P, Waterman S, Smole S, Guevara HF, Belongia EA, Clark PA, Beatrice ST, Donis R, Katz J, Finelli L, Bridges CB, Shaw M, Jernigan DB, Uyeki TM, Smith DJ, Klimov AI, Cox NJ. 2009. Antigenic and genetic characteristics of swine-origin 2009 A(H1N1) influenza viruses circulating in humans. Science 325:197-201.

97. Smith GJ, Vijaykrishna D, Bahl J, Lycett SJ, Worobey M, Pybus OG, Ma SK, Cheung CL, Raghwani J, Bhatt S, Peiris JS, Guan Y, Rambaut A. 2009. Origins and evolutionary genomics of the 2009 swine-origin H1N1 influenza A epidemic. Nature 459:1122-1125.

98. Xu R, Ekiert DC, Krause JC, Hai R, Crowe JE, Wilson IA. 2010. Structural Basis of Preexisting Immunity to the 2009 H1N1 Pandemic Influenza Virus. Science 328:357-360.

99. Yang H, Carney P, Stevens J. 2010. Structure and Receptor binding properties of a pandemic H1N1 virus hemagglutinin. PLoS Curr 2:RRN1152.

100. Centers for Disease Control and Prevention. 2009. Serum cross-reactive antibody response to a novel influenza A (H1N1) virus after vaccination with seasonal influenza vaccine. 58:521-524.

101. Munster VJ, de Wit E, van den Brand JM, Herfst S, Schrauwen EJ, Bestebroer TM, van de Vijver D, Boucher CA, Koopmans M, Rimmelzwaan GF, Kuiken T, Osterhaus AD, Fouchier RA. 2009. Pathogenesis and transmission of swine-origin 2009 A(H1N1) influenza virus in ferrets. Science 325:481-483.

102. Chutinimitkul S, Herfst S, Steel J, Lowen AC, Ye J, van Riel D, Schrauwen EJ, Bestebroer TM, Koel B, Burke DF, Sutherland-Cash KH, Whittleston CS, Russell CA, Wales DJ, Smith DJ, Jonges M, Meijer A, Koopmans M, Rimmelzwaan GF, Kuiken T, Osterhaus AD, Garcia-Sastre A, Perez DR, Fouchier RA. 2010. Virulence-associated substitution D222G in the hemagglutinin of 2009 pandemic influenza A(H1N1) virus affects receptor binding. J Virol 84:11802-11813.

103. Nobusawa E, Ishihara H, Morishita T, Sato K, Nakajima K. 2000. Change in receptor-binding specificity of recent human influenza A viruses (H3N2): a single amino acid change in hemagglutinin altered its recognition of sialyloligosaccharides. Virology 278:587-596.

104. de Vries RP, de Vries E, Martinez-Romero C, McBride R, van Kuppeveld FJ, Rottier PJ, Garcia-Sastre A, Paulson JC, de Haan CA. 2013. Evolution of the hemagglutinin protein of the new pandemic H1N1 influenza virus: maintaining optimal receptor binding by compensatory substitutions. J Virol 87:13868-13877.

105. Chutinimitkul S, van Riel D, Munster VJ, van den Brand JM, Rimmelzwaan GF, Kuiken T, Osterhaus AD, Fouchier RA, de Wit E. 2010. In vitro assessment of attachment pattern and replication efficiency of H5N1 influenza A viruses with altered receptor specificity. J Virol 84:6825-6833.

106. Miller MS, Gardner TJ, Krammer F, Aguado LC, Tortorella D, Basler CF, Palese P. 2013. Neutralizing antibodies against previously encountered influenza virus strains increase over time: a longitudinal analysis. Sci Transl Med 5:198ra107.

107. Fonville JM, Wilks SH, James SL, Fox A, Ventresca M, Aban M, Xue L, Jones TC, Le NM, Pham QT, Tran ND, Wong Y, Mosterin A, Katzelnick LC, Labonte D, Le TT, van der Net G, Skepner E, Russell CA, Kaplan TD, Rimmelzwaan GF, Masurel N, de Jong JC, Palache A, Beyer WE, Le QM, Nguyen TH, Wertheim HF, Hurt AC, Osterhaus AD, Barr IG, Fouchier RA, Horby PW, Smith DJ. 2014. Antibody landscapes after influenza virus infection or vaccination. Science 346:996-1000.

108. Chen Z, Wang W, Zhou H, Suguitan AL, Jr., Shambaugh C, Kim L, Zhao J, Kemble G, Jin H. 2010. Generation of live attenuated novel influenza virus A/California/7/09 (H1N1) vaccines with high yield in embryonated chicken eggs. J Virol 84:44-51.

109. O'Donnell CD, Vogel L, Wright A, Das SR, Wrammert J, Li GM, McCausland M, Zheng NY, Yewdell JW, Ahmed R, Wilson PC, Subbarao K. 2012. Antibody pressure by a human monoclonal antibody targeting the 2009 pandemic H1N1 virus hemagglutinin drives the emergence of a virus with increased virulence in mice. MBio 3.

110. Rudneva I, Ignatieva A, Timofeeva T, Shilov A, Kushch A, Masalova O, Klimova R, Bovin N, Mochalova L, Kaverin N. 2012. Escape mutants of pandemic influenza A/H1N1 2009 virus: variations in antigenic specificity and receptor affinity of the hemagglutinin. Virus Res 166:61-67.

111. Guarnaccia T, Carolan LA, Maurer-Stroh S, Lee RT, Job E, Reading PC, Petrie S, McCaw JM, McVernon J, Hurt AC, Kelso A, Mosse J, Barr IG, Laurie KL. 2013. Antigenic drift of the pandemic 2009 A(H1N1) influenza virus in A ferret model. PLoS Pathog 9:e1003354.

112. Li Y, Myers JL, Bostick DL, Sullivan CB, Madara J, Linderman SL, Liu Q, Carter DM, Wrammert J, Esposito S, Principi N, Plotkin JB, Ross TM, Ahmed R, Wilson PC, Hensley SE. 2013. Immune history shapes specificity of pandemic H1N1 influenza antibody responses. J Exp Med 210:1493-1500.

113. Bizebard T, Gigant B, Rigolet P, Rasmussen B, Diat O, Bosecke P, Wharton SA, Skehel JJ, Knossow M. 1995. Structure of influenza virus haemagglutinin complexed with a neutralizing antibody. Nature 376:92-94.

114. Soundararajan V, Tharakaraman K, Raman R, Raguram S, Shriver Z, Sasisekharan V, Sasisekharan R. 2009. Extrapolating from sequence--the 2009 H1N1 'swine' influenza virus. Nat Biotechnol 27:510-513.

115. Linderman SL, Chambers BS, Zost SJ, Parkhouse K, Li Y, Herrmann C, Ellebedy AH, Carter DM, Andrews SF, Zheng NY, Huang M, Huang Y, Strauss D, Shaz BH, Hodinka RL, Reyes-Teran G, Ross TM, Wilson PC, Ahmed R, Bloom JD, Hensley SE. 2014. Potential antigenic explanation for atypical H1N1 infections among middle-aged adults during the 2013-2014 influenza season. Proc Natl Acad Sci U S A 111:15798-15803.

116. Adinolfi M, Lessof MH. 1972. Development of humoral and cellular immunity in man. J Med Genet 9:86-91.

117. Bodewes R, de Mutsert G, van der Klis FR, Ventresca M, Wilks S, Smith DJ, Koopmans M, Fouchier RA, Osterhaus AD, Rimmelzwaan GF. 2011. Prevalence of antibodies against seasonal influenza A and B viruses in children in Netherlands. Clin Vaccine Immunol 18:469-476.

118. Carter DM, Bloom CE, Nascimento EJ, Marques ET, Craigo JK, Cherry JL, Lipman DJ, Ross TM. 2013. Sequential seasonal H1N1 influenza virus infections protect ferrets against novel 2009 H1N1 influenza virus. J Virol 87:1400-1410.

119. Corti D, Voss J, Gamblin SJ, Codoni G, Macagno A, Jarrossay D, Vachieri SG, Pinna D, Minola A, Vanzetta F, Silacci C, Fernandez-Rodriguez BM, Agatic G, Bianchi S, Giacchetto-Sasselli I, Calder L, Sallusto F, Collins P, Haire LF, Temperton N, Langedijk JP, Skehel JJ, Lanzavecchia A. 2011. A neutralizing antibody selected from plasma cells that binds to group 1 and group 2 influenza A hemagglutinins. Science 333:850-856.

120. Friesen RH, Lee PS, Stoop EJ, Hoffman RM, Ekiert DC, Bhabha G, Yu W, Juraszek J, Koudstaal W, Jongeneelen M, Korse HJ, Ophorst C, Brinkman-van der Linden EC, Throsby M, Kwakkenbos MJ, Bakker AQ, Beaumont T, Spits H, Kwaks T, Vogels R, Ward AB, Goudsmit J, Wilson IA. 2014. A common solution to group 2 influenza virus neutralization. Proc Natl Acad Sci U S A 111:445-450.

121. Sui J, Hwang WC, Perez S, Wei G, Aird D, Chen LM, Santelli E, Stec B, Cadwell G, Ali M, Wan H, Murakami A, Yammanuru A, Han T, Cox NJ, Bankston LA, Donis RO, Liddington RC, Marasco WA. 2009. Structural and functional bases for broad-spectrum neutralization of avian and human influenza A viruses. Nat Struct Mol Biol 16:265-273.

122. Ekiert DC, Bhabha G, Elsliger MA, Friesen RH, Jongeneelen M, Throsby M, Goudsmit J, Wilson IA. 2009. Antibody recognition of a highly conserved influenza virus epitope. Science 324:246-251.

123. Dreyfus C, Laursen NS, Kwaks T, Zuijdgeest D, Khayat R, Ekiert DC, Lee JH, Metlagel Z, Bujny MV, Jongeneelen M, van der Vlugt R, Lamrani M, Korse HJ, Geelen E, Sahin O, Sieuwerts M, Brakenhoff JP, Vogels R, Li OT, Poon LL, Peiris M, Koudstaal W, Ward AB, Wilson IA, Goudsmit J, Friesen RH. 2012. Highly conserved protective epitopes on influenza B viruses. Science 337:1343-1348.

124. Jegaskanda S, Reading PC, Kent SJ. 2014. Influenza-specific antibody-dependent cellular cytotoxicity: toward a universal influenza vaccine. J Immunol 193:469-475.

125. Bedford T, Suchard MA, Lemey P, Dudas G, Gregory V, Hay AJ, McCauley JW, Russell CA, Smith DJ, Rambaut A. 2014. Integrating influenza antigenic dynamics with molecular evolution. Elife 3:e01914.

126. Lauring AS, Andino R. 2010. Quasispecies theory and the behavior of RNA viruses. PLoS Pathog 6:e1001005.

127. Shih AC-C, Hsiao T-C, Ho M-S, Li W-H. 2007. Simultaneous amino acid substitutions at antigenic sites drive influenza A hemagglutinin evolution. Proc Natl Acad Sci USA 104:6283-6288.

128. Plotkin JB, Dushoff J. 2003. Codon bias and frequency-dependent selection on the hemagglutinin epitopes of influenza A virus. Proc Natl Acad Sci USA 100:7152-7157.

129. Lee M-S, Chen JS-E. 2004. Predicting antigenic variants of influenza A/H3N2 viruses. Emerging Infect Dis 10:1385-1390.

130. Holmes EC, Ghedin E, Miller N, Taylor J, Bao Y, St George K, Grenfell BT, Salzberg SL, Fraser CM, Lipman DJ, Taubenberger JK. 2005. Whole-genome analysis of human influenza A virus reveals multiple persistent lineages and reassortment among recent H3N2 viruses. PLoS Biol 3:e300.

131. Rambaut A, Pybus OG, Nelson MI, Viboud C, Taubenberger JK, Holmes EC. 2008. The genomic and epidemiological dynamics of human influenza A virus. Nature 453:615-619.

132. 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. Genomewide analysis of reassortment and evolution of human influenza A(H3N2) viruses circulating between 1968 and 2011. J Virol 88:2844-2857.

133. Li Y, Bostick DL, Sullivan CB, Myers JL, Griesemer SB, Stgeorge K, Plotkin JB, Hensley SE. 2013. Single hemagglutinin mutations that alter both antigenicity and receptor binding avidity influence influenza virus antigenic clustering. J Virol 87:9904-9910.

134. Luksza M, Lassig M. 2014. A predictive fitness model for influenza. Nature 507:57-61.

135. Mitnaul LJ, Matrosovich MN, Castrucci MR, Tuzikov AB, Bovin NV, Kobasa D, Kawaoka Y. 2000. Balanced hemagglutinin and neuraminidase activities are critical for efficient replication of influenza A virus. J Virol 74:6015-6020.

136. Wagner R, Matrosovich M, Klenk HD. 2002. Functional balance between haemagglutinin and neuraminidase in influenza virus infections. Rev Med Virol 12:159-166.

137. 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 hemagglutinin in H1N1 and H3N2 influenza viruses. Proc Natl Acad Sci U S A 108:20748-20753.

138. Quinones-Mateu ME, Avila S, Reyes-Teran G, Martinez MA. 2014. Deep sequencing: becoming a critical tool in clinical virology. J Clin Virol 61:9-19.

139. McElroy K, Thomas T, Luciani F. 2014. Deep sequencing of evolving pathogen populations: applications, errors, and bioinformatic solutions. Microb Inform Exp 4:1.

140. Linster M, van Boheemen S, de Graaf M, Schrauwen EJ, Lexmond P, Manz B, Bestebroer TM, Baumann J, van Riel D, Rimmelzwaan GF, Osterhaus AD, Matrosovich M, Fouchier RA, Herfst S. 2014. Identification, characterization, and natural selection of mutations driving airborne transmission of A/H5N1 virus. Cell 157:329-339.

141. Bull RA, Luciani F, McElroy K, Gaudieri S, Pham ST, Chopra A, Cameron B, Maher L, Dore GJ, White PA, Lloyd AR. 2011. Sequential bottlenecks drive viral evolution in early acute hepatitis C virus infection. PLoS Pathog 7:e1002243.

142. Henn MR, Boutwell CL, Charlebois P, Lennon NJ, Power KA, Macalalad AR, Berlin AM, Malboeuf CM, Ryan EM, Gnerre S, Zody MC, Erlich RL, Green LM, Berical A, Wang Y, Casali M, Streeck H, Bloom AK, Dudek T, Tully D, Newman R, Axten KL, Gladden AD, Battis L, Kemper M, Zeng Q, Shea TP, Gujja S, Zedlack C, Gasser O, Brander C, Hess C, Gunthard HF, Brumme ZL, Brumme CJ, Bazner S, Rychert J, Tinsley JP, Mayer KH, Rosenberg E, Pereyra F, Levin JZ, Young SK, Jessen H, Altfeld M, Birren BW, Walker BD, Allen TM. 2012. Whole genome deep sequencing of HIV-1 reveals the impact of early minor variants upon immune recognition during acute infection. PLoS Pathog 8:e1002529.

143. Rozera G, Abbate I, Bruselles A, Vlassi C, D'Offizi G, Narciso P, Chillemi G, Prosperi M, Ippolito G, Capobianchi MR. 2009. Massively parallel pyrosequencing highlights minority variants in the HIV-1 env quasispecies deriving from lymphomonocyte sub-populations. Retrovirology 6:15.

144. Burke DF, Smith DJ. 2014. A recommended numbering scheme for influenza A HA subtypes. PLoS One 9:e112302.