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

Kim B. Westgeest,1 Miranda de Graaf,1,2 Mathieu Fourment,3 Theo M. Bestebroer,1 Ruud van Beek,1 Monique I. J. Spronken,1 Jan C. de Jong,1 Guus F. Rimmelaarwaan,1 Colin A. Russell,2,4 Albert D. M. E. Osterhaus,1 Gavin J. D. Smith,3 Derek J. Smith1,2,4 and Ron A. M. Fouchier1

1Department of Virology, Erasmus Medical Center, 3000 CA, Rotterdam, The Netherlands
2Department of Zoology, University of Cambridge, Cambridge CB2 3EJ, UK
3Program in Emerging Infectious Diseases, Duke-NUS Graduate Medical School, 8 College Road, Singapore 169857, Singapore
4Fogarty International Center, National Institutes of Health, Bethesda, MD 20892, USA

Each year, influenza viruses cause epidemics by evading pre-existing humoral immunity through mutations in the major glycoproteins: the haemagglutinin (HA) and the neuraminidase (NA). In 2004, the antigenic evolution of HA of human influenza A (H3N2) viruses was mapped (Smith et al., Science 305, 371–376, 2004) from its introduction in humans in 1968 until 2003. The current study focused on the genetic evolution of NA and compared it with HA using the dataset of Smith and colleagues, updated to the epidemic of the 2009/2010 season. Phylogenetic trees and genetic maps were constructed to visualize the genetic evolution of NA and HA. The results revealed multiple reassortment events over the years. Overall rates of evolutionary change were lower for NA than for HA1 at the nucleotide level. Selection pressures were estimated, revealing an abundance of negatively selected sites and sparse positively selected sites. The differences found between the evolution of NA and HA1 warrant further analysis of the evolution of NA at the phenotypic level, as has been done previously for HA.

INTRODUCTION

Influenza epidemics affect approximately 5–15% of the world population, resulting in an estimated 3–5 million hospitalizations and between 250,000 and 500,000 deaths annually (Stöhr, 2002; WHO, 2003). Influenza type A viruses can be divided into subtypes based on the antigenic properties of the major surface glycoproteins, haemagglutinin (HA) and neuraminidase (NA). To date, 17 HA (Fouchier et al., 2005; Tong et al., 2012) and nine NA (Schild et al., 1980) subtypes have been found in nature. Subtypes A (H1N1) and A (H3N2) are currently the causative agents of influenza A virus epidemics in humans, of which the influenza A (H3N2) viruses are the most recurring and virulent (Barr et al., 2010; Taubenberger & Morens, 2006) and show the strongest antigenic drift (Rambaut et al., 2008).

HA, a homotrimeric type I integral membrane protein, mediates viral entry into the cell by binding to sialic acids (Sauter et al., 1989). NA, a tetrameric type II integral membrane protein with sialidase activity, allows virus release from the cell (Liu et al., 1995; Palese & Compans, 1976; Palese et al., 1974). Both HA and NA are located on the surface of the viral membrane and are the main targets for antibodies. Antibodies against influenza viruses result in protective immunity, but mutations in HA and NA allow the virus to escape host immunity. This process, known as antigenic drift (Schild et al., 1974), is responsible for the recurrence of influenza epidemics almost every winter.

Vaccines have proven effective but need to be updated frequently due to antigenic drift. Since 1999, the influenza A (H3N2) virus component has been updated six times (Barr et al., 2010). Thorough surveillance by the National
Influenza Centers and WHO Collaborating Centers within the World Health Organization’s Global Influenza Surveillance Network is required to identify the most suitable strains to use in vaccines for the next epidemic (Barr et al., 2010; Russell et al., 2008). Vaccine strain selection depends on three aspects: epidemiological information, HA and NA gene sequence phylogeny and serological analysis using an HA inhibition assay. The main focus of genetic and antigenic surveillance is on HA, and official influenza vaccine formulations prescribe the amount of HA (Fiore et al., 2010).

Although antibodies against NA do not prevent infection, numerous pre-clinical and clinical studies indicate a role of NA immunity in reducing the severity of influenza virus infection (Brett & Johansson, 2005; Couch et al., 1974; Johansson et al., 1993; Kilbourne, 1976; Murphy et al., 1972; Schulman et al., 1968). Early crystallographic studies of NA have shown that antigenic regions surround the enzyme’s highly conserved active site (Colman et al., 1983, 1987). Antigenic sites A, B and C (Air et al., 1985) have been shown to be highly variable, probably due to antigenic drift (Laver et al., 1982; Luther et al., 1984). Studies comparing the antigenic drift of HA and NA using limited numbers of viruses have revealed that their evolution differs and is often asynchronous (Kilbourne et al., 1990; Sandbulte et al., 2011; Schulman & Kilbourne, 1969). Given these factors, understanding the patterns of evolution in NA is important.

In 2004, Smith et al. (2004) mapped the antigenic evolution of HA of human influenza A (H3N2) virus from its introduction in humans in 1968 until 2003. The study was based on an extensive dataset, comprising influenza virus isolates obtained within each influenza season. Here, we focused on the genetic evolution of NA and compared it with HA1 (the immunogenic section of the HA) using the dataset of Smith et al. (2004), updated to the epidemic of the 2009/2010 season. Reassortment events, rates of evolutionary change and selection pressures were analysed over 40 years of influenza A (H3N2) virus evolution. The differences found between the evolution of NA and HA1 warrant further analysis of the antigenic properties of NA.

RESULTS

Dataset background

Virus isolates obtained between 1968 and 2003 and used in the study of Smith et al. (2004) were combined with 19 influenza A (H3N2) viruses isolated between 2003 and 2009 into a single updated dataset consisting of 291 virus isolates. The new isolates included seven vaccine or reference strains and 12 isolates from epidemics in the Netherlands, of which the latter were chosen based on divergent placement in the updated influenza A (H3N2) antigenic map (de Jong et al., 2011). All NA gene segments of these 291 viruses were sequenced. For the study by Smith et al. (2004), only the HA1 coding regions were sequenced, as this is the immunogenic section of the HA protein (Wiley et al., 1981; Wilson & Cox, 1990). To obtain a similar dataset for HA and NA, the HA1 coding regions of the 19 recent influenza viruses were sequenced. Deduced amino acid sequence alignments showing only positions with at least one mutation are available in Figs S1 and S2 for NA and HA1, respectively (available in JGV Online). Both alignments were colour coded according to the antigenic clusters of HA (Smith et al., 2004).

Genetic evolution of NA and HA1

Maximum-likelihood (ML) phylogenetic trees were constructed to observe the genetic evolution of NA and HA1, updated from 2004, at the nucleotide level (Fig. 1). The general topology of the NA tree was similar to that of the HA1 tree, showing the typical ‘ladder-like’ gradual evolution with rapid replacement of old strains by newer
ones. NA had fewer nucleotide substitutions over 40 years of evolution compared with HA1; overall, the genetic distance from the root of the tree to the most recent cluster of CA04 strains was roughly 1.5-fold greater for HA1 than for NA. Although we did see that variants on the trunk were the ancestors of all variants in future years (Fitch et al., 1997), the HA1 phylogeny did have some variants that clustered away from the trunk in an evolutionary terminal or dead-end clade, most notably for the VI75- and the BE89-like viruses. Compared with HA1, such dead-end clades were less obvious in the NA tree. For NA, the most obvious – yet relatively small – evolutionary dead-end clades included a number of the SY97- and BE92-like viruses. Whilst the HA genes of the VI75-like and BE92-like viruses clustered continuously in the ML tree, the NA genes appeared as two separate lineages for both groups of viruses. Notably, the NA genes of HK68-like viruses were more genetically divergent than HA1 (0.044403 vs 0.03556 nucleotide substitutions per site, respectively). Larger genetic distances were found between BK79- and SI87-like viruses and within the SY97-like viruses for NA compared with HA1. In contrast, the distances between EN72- and VI75-, between TX77- and BK79-, and between SY97- and FU02-like viruses were greater for HA1.

Genetic maps were generated to visualize the genetic evolution of NA at the amino acid level (Fig. 2a). The HA1 genetic map was updated with the newly added isolates (Fig. 2b). In Fig. 2(c), the two maps are compared with arrows that point from the position of a virus in the HA1 map to the position of the corresponding virus in the NA map. Overall, the genetic map of the NA appeared less clustered and more gradual compared with the HA1 genetic map. Although the genetic clustering of strains based on HA1 amino acid sequences was in good agreement with the antigenic properties of the HA proteins as described previously (Smith et al., 2004), the NA amino acid sequences did not strictly follow the same pattern of clustering. Other features of the NA genetic map, including the less pronounced evolutionary dead ends (yellow and

---

**Fig. 2.** Genetic maps of the major glycoproteins of human influenza A (H3N2) viruses. Genetic maps were generated with 291 aa sequences and multidimensional scaling (MDS) algorithms for NA (a) and HA1 (b). The vertical and horizontal axes correspond to the number of amino acid substitutions; one square represents five amino acid substitutions. The orientation of the HA1 and NA maps was chosen to match the orientation of the antigenic map of human influenza A (H3N2) viruses (Smith et al., 2004). The colour coding of viruses is based on the antigenic clusters of HA and is consistent among all maps (see legend to Fig. 1). The map in panel (c) depicts the HA genetic map with superimposed arrows pointing towards the corresponding data points in the NA genetic map.
red clusters in Fig. 2b compared with Fig. 2a) and differences in genetic distance within and between clusters compared with HA1, were in agreement with the ML trees.

**Reassortment events**

In 2005, Holmes *et al.* (2005) demonstrated for a limited number of influenza seasons that multiple lineages of influenza A (H3N2) viruses co-circulate, persist and reassort in epidemiologically significant ways. By applying the antigenic cluster colours of HA (Smith *et al.*, 2004) to both the HA1 and NA trees, it was noted that the clusters of NA sequences did not strictly coincide with those of HA (Fig. 1). The HA1 and NA trees shown in Fig. 1 were used in TreeMap to generate a tanglegram. A tanglegram enables visualization of the location of particular isolates within both ML trees. In the absence of reassortment, the twines should connect both trees, in theory, in a seamlessly horizontal way. This was not the case for the isolates used in this study, suggesting frequent reassortment between the NA and HA gene segments. However, this was based on only one phylogeny per gene segment, with clades and branches ordered by branch length in a fixed fashion. To improve on this, the computational software package Graph-incompatibility-based Reassortment Finder (GiRaF) was used to identify reassortments in multiple trees per gene segment (Nagarajan & Kingsford, 2011). This method compares large collections of Markov chain Monte Carlo (MCMC)-sampled trees for groups of incompatible splits to identify sets of taxa with differential phylogenetic placement, whilst accounting for uncertainties in the inferred phylogenies. We performed 100 independent GiRaF analyses on ten independent MrBayes HA runs compared with ten independent MrBayes NA runs. Each run consisted of two independent tree files each with 1000 MCMC-sampled trees. GiRaF removed the first 500 of these trees as burn-in, and the remaining 500 were used for the actual analysis. All reassortment events are reported in Table S1 and those with a support value of ≥50 % of the GiRaF runs are depicted on the tanglegram made from the ML trees as bold twines (Fig. 3). A 50 % cut-off was chosen arbitrarily.

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

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

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

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

**Evolutionary rates**

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

**Selection pressures**

The degree of natural selection acting on NA and HA1 was estimated by looking at the mean number of amino acid-changing (non-synonymous or d_<n>) and silent (synonym- ous or d_<s>) substitutions per site (d_<n>/d_<s> ratio) using Codeml. The overall d_<n>/d_<s> and the mean d_<n>/d_<s> ratios for the internal and external branches were estimated (Table 2). Both genes were under strong selection (NA d_<n>/d_<s>

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

<table>
<thead>
<tr>
<th>Segment</th>
<th>Mean rate of nucleotide substitution (×10⁻³ substitutions per site per year)</th>
<th>Time of MRCA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>95 % HPD</td>
</tr>
<tr>
<td></td>
<td>Lower</td>
<td>Upper</td>
</tr>
<tr>
<td>NA</td>
<td>3.15</td>
<td>2.81</td>
</tr>
<tr>
<td>HA1</td>
<td>5.15</td>
<td>4.62</td>
</tr>
</tbody>
</table>
Table 2. Global, inner and outer $d_S/d_S$ values for NA and HA1

<table>
<thead>
<tr>
<th>Segment</th>
<th>Sequence length (codons)</th>
<th>$d_S/d_S$</th>
<th>Overall</th>
<th>Internal</th>
<th>External</th>
</tr>
</thead>
<tbody>
<tr>
<td>NA</td>
<td>470</td>
<td>0.249</td>
<td>0.250</td>
<td>0.248</td>
<td></td>
</tr>
<tr>
<td>HA1</td>
<td>329</td>
<td>0.362</td>
<td>0.350</td>
<td>0.371</td>
<td></td>
</tr>
</tbody>
</table>

$=0.249$ and $\text{HA1 } d_S/d_S=0.362$); however, $\text{HA1 was less}
constrained. The $d_S/d_S$ values for the internal and external
branches were similar for NA (0.250 and 0.248) and HA1
(0.350 and 0.371), suggesting no difference between
selection pressures on NA and HA1. In theory, a higher
$d_S/d_S$ value for the external relative to the internal branches
indicates an excess of non-synonymous mutations that are
eventually removed from the virus population by purifying
selection, suggesting that these mutations are deleterious
(Pybus et al., 2007). This was not the case for NA and HA1.

To determine the positively and negatively selected sites
along internal branches only, where advantageous muta-
tions are more likely to fall, the internal fixed effects
likelihood (IFEL) method was used (Kosakovsky Pond
et al., 2006). IFEL detected three positively selected sites in
NA: codons 43, 267 and 370 (Table S2). In HA1, nine
positively selected sites were reported, all located in anti-
genic sites (Table S3).

Sites subject to episodic diversifying selection were iden-
tified using mixed effects model of evolution (MEME)
(Kosakovsky Pond et al., 2011). MEME detected five sites
(codons 43, 148, 199, 338 and 465) in NA (Table S2) and
14 in HA1 (Table S3), of which all but two (codons 31 and
185) were found in antigenic sites.

Directional positive selection was detected with the direc-
tional evolution in protein sequences (DEPS) test imple-
mented in the HyPhy package (Kosakovsky Pond et al.,
2005). The DEPS test identifies directional evolution
towards residues at sites within an amino acid alignment
in combination with a nucleotide tree and is particularly
useful for the detection of selective sweeps (Kosakovsky
Pond et al., 2008). For NA, 11 sites (Tables S2 and S4) were
identified to be involved in this directional evolution (see
Table S4 for the inferred amino acid substitution pat-
terns). For HA1, nine sites were involved (Tables S3 and
S5). In contrast, this method detected more sites under
directional selection in NA compared with HA1 (11 and
nine sites, respectively).

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

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

DISCUSSION

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

The general topology of the NA and HA1 trees were
similar, showing the typical ‘ladder-like’ gradual evolution
with rapid replacement of old strains by newer ones, as
described previously for HA1 (Fitch et al., 1997). NA
evolved more slowly at the nucleotide level and more
gradually, with overall shorter branches and fewer and
smaller evolutionary dead-end clades. The genetic distances
between clades in the ML trees and in the genetic maps
for NA and HA1 were often discordant, suggesting an
asynchronous genetic evolution of the two genes. This
observation is in agreement with the reported asynchron-
ous evolution at the phenotypic level, with discordant
patterns of antigenic drift of HA and NA, as measured by
serological assays (Kilbourne et al., 1990; Sandbulte et al.,
2011; Schulman & Kilbourne, 1969).

Reassortment is a common feature in influenza viruses
(Webster et al., 1992). The Asian, Hong Kong and 2009
swine flu pandemics of 1957, 1968 and 2009, respectively,
all emerged following reassortment of swine, avian and/or
human influenza viruses (Scholtissek et al., 1978; Smith
et al., 2009). In addition to the emergence of pandemic
strains, it has been suggested that the emergence of
new virus lineages during seasonal epidemics of the past
decade may have been associated with reassortment events
(Holmes et al., 2005), perhaps resulting from improved
matching of the activity of HA and NA in relation to HA
antigenic change (Kaverin et al., 2000; Mitinaul et al., 2000;
Wagner et al., 2002). However, when amino acid substitu-
tions in the active sites of HA and NA were analysed, no
evidence for a ‘matched’ evolution of HA and NA was
found, as there were only isolated cases of substitutions in the catalytic site of NA (specifically at position 151) that seemed to provide no increased fitness as their occurrences were transient.

For our influenza A virus dataset spanning >40 years of influenza A (H3N2) virus epidemics, differences in clustering of HA and NA gene segments within phylogenetic trees indeed suggested that reassortment events had occurred relatively frequently over time. The large genetic distances between NA sequences of the HK68-like viruses, shown in both the phylogenetic ML tree (Fig. 1a) and the genetic map (Fig. 2a), could be explained by a reassortment event involving an early influenza A (H3N2) virus, in which a heterologous N2 gene of H2N2 virus origin was reintroduced (Lindstrom et al., 2004). Indeed, in an ML tree generated with NA genes of all early influenza A (H3N2) viruses and additional H2N2 strains, the NA of the late H2N2 viruses clustered together with HK68-like viruses of 1969 and 1970 (data not shown). GiRaF analyses provided further support that several HK68-like strains (A/Bilthoven/17938/1969, A/Bilthoven/2668/1970 and A/Bilthoven/93/1970) were reassortants. Apart from these late HK68-like strains, GiRaF provided evidence for reassortment events between 1971 and 1977. Co-circulation of antigenically distinct viruses has been reported in this period (Kendal et al., 1978; Pereira & Chakraverty, 1977; Schild et al., 1973), providing opportunities for such reassortment events. In agreement with a previous study (Xu et al., 1996), there was also support for reassortment events between BE89- and BE92-like viruses. GiRaF analyses further suggested that reassortment events between HA and NA of influenza A (H3N2) viruses occurred particularly frequently during the circulation of BE92- and WU95-like viruses. It remains unclear whether the frequency of reassortment events over time is significant and whether these reassortment events have been of epidemiological significance. Analyses with all genomic sequences would provide us with a better insight into reassortment patterns.

Some concerns arise from the use of GiRaF in the intra-H3N2-evolution context. The evolutionary distances sampled in the present dataset are not the same as those GiRaF was originally benchmarked against, and the performance of GiRaF was shown to depend on the distance distribution. Moreover, GiRaF seemed to perform slightly better for single-taxa reassortments compared with larger events. It should further be noted that the low false-positive rate of GiRaF was only estimated on synthetic datasets assuming neutral evolution models (Nagarajan & Kingsford, 2011).

Here, we used a tanglegram based on robust HA and NA ML trees to visually inspect the results reported by GiRaF.
as an independent test. The reassortment events reported by GiRaF with >50% support could all be confirmed by visual inspection of the tanglegram. To test whether the small genetic distances between the sets of HA and NA sequences were problematic for GiRaF, we also ran GiRaF on HA trees only (comparing one set of HA trees with another set of HA trees) and on NA trees only, and using alignments in which the HA and NA sequences were split into half. One would not expect to detect any reassortment events in these runs. GiRaF analysis reported only three events for the 5'HA–3'HA analysis, and no reassortment events for the analysis of 5'NA–3'NA, HA–HA and NA–NA trees. Thus, we concluded from this that the 22 HA–NA reassortment events reported in the present work are supported by GiRaF and visual inspection.

The evolutionary rates of NA and HA1 were high (Hanada et al., 2004; Jenkins et al., 2002) and of the expected level for influenza viruses (Bhatt et al., 2011). The evolutionary rate of NA of 3.15 × 10⁻³ nucleotide substitutions per site per year was slightly higher than a previously reported rate (2.3 × 10⁻³; Xu et al., 1996). Although the rate of nucleotide substitution of NA was lower than that of HA (5.15 × 10⁻³ nucleotide substitutions per site per year), it is important to note that only HA1 was analysed here. As HA1 is the most variable part of the HA gene, this part would probably yield higher evolutionary rates than the full HA gene (Bhatt et al., 2011).

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

The overall mean $d_S/d_N$ values of 0.249 (NA) and 0.362 (HA1) estimated by the PAML program (Table 2) were in agreement with previously reported values (Chen & Holmes, 2008). There was an abundance of negatively selected sites and a limited number of positively selected sites in both HA1 and NA (Tables 2, S2 and S3). Generally, the number of positively selected sites was lower in NA compared with HA1. In contrast, DEPS detected more sites under directional selection in NA compared with HA1, which suggests that positive selection in HA1 is more random, or it could reflect toggling of antigenic sites due to antibody pressure.

The positively selected codons 43, 46 and 52 are all located in the highly variable stalk region (Blok & Air, 1982). Residue 199, detected by MEME, was reported previously as an antigenic site (Gulati et al., 2002). Residue 267, detected by IFEL, is not an antigenic site but has been reported previously as being positively selected (Kosakovsky Pond et al., 2008). Two documented antigenic sites, 328 (Air et al., 1985) and 334 (Air et al., 1985; Laver et al., 1982), were detected by DEPS. Residue 370, a reported antigenic residue (Air et al., 1985; Laver et al., 1982; Lentz et al., 1984), was found by IFEL and DEPS as being positively selected. By IFEL, MEME and DEPS, the number of positively selected sites found to be within antigenic sites in NA was lower than in HA1.

In summary, we have analysed the evolutionary genetics of NA of influenza A (H3N2) viruses isolated from 1968 to 2009 and compared them with those of HA. High-level similarities were observed between the evolution of HA and NA, although notable differences were apparent. Future research on the evolution of NA should focus at the phenotypic level, using serological tests and antigenic cartography methods, as described recently (Kilbourne et al., 1990; Sandbulte et al., 2011; Schulman & Kilbourne, 1969), on extensive virus datasets. Such analysis would allow not only the side-by-side comparison of HA and NA evolution at the genetic level as carried out here but also its relationship with immune escape, the major driver of evolution of the surface glycoproteins of influenza A virus. Such analyses, along with full virus genome data, ultimately may lead to a better understanding and increased predictability of the evolution of influenza A (H3N2) virus.

**METHODS**

**Viruses.** Human influenza A (H3N2) viruses, isolated over 35 years of influenza virus surveillance between 1968 and 2003 that were used in the study of Smith et al. (2004), served as the basis for this study. Two viruses within this dataset were no longer available (A/Victoria/7/87 and A/Netherlands/440/93) and were thus excluded from the dataset. The dataset was updated with 19 influenza A (H3N2) viruses circulating between 2003 and 2009, including seven vaccine or reference strains and 12 isolates from epidemics in the Netherlands. The latter 12 isolates were chosen based on divergent placement in the updated A (H3N2) antigenic map (de Jong et al., 2011). This led to a total of 291 human influenza A (H3N2) viruses. Human influenza A (H3N2) viruses were propagated in Madin–Darby canine kidney (MDCK) cells or, if unsuccessful, in 11-day-old embryonated chicken eggs.

**Sequence analysis.** MDCK supernatant or allantoic fluid was used for RNA extraction, using a High Pure RNA isolation kit (Roche Applied Science) and cDNA was synthesized with SuperScript III reverse transcriptase (Invitrogen). The complete NA and HA1 gene segments were amplified by PCR using AmpliTaq Gold DNA polymerase (Applied Biosystems) and purified by gel extraction with a QIAquick Gel Extraction kit (Qiagen). Sequencing was performed with NA- or HA-specific primers using a BigDye Terminator Version 3.1 Cycle Sequencing kit (Applied Biosystems) and a 3130 XL Genetic Analyzer (Applied Biosystems), according to the instructions of the manufacturer.

Nucleotide sequences of the NA ORFs and HA1 coding regions were aligned using the CLUSTAL W program running within the BioEdit software package, version 7.0.9.0 (Hall, 1999). Seven HA1 and 14 NA sequences contained one or more degenerate nucleotide positions. In all HA1 and ten of the 14 NA sequences, the degenerate nucleotide led to degenerate amino acid positions. One NA sequence contained an insertion of 3 nt between codons 221 and 222, and one sequence had a 3 nt deletion at codon 154. Newly sequenced and previously published NA and HA1 domain accession numbers are provided in the Supplementary data.
Phylogenetic analysis. With the nucleotide sequence alignments, initial ML trees were inferred using the PhyML package version 3.0 (Guindon & Gascuel, 2003), by means of a full heuristic search and the subtree pruning and regrafting (SPR) method. As the TVM+I+F model of nucleotide substitution was not an option within the PhyML package, the GTR+I+F model was chosen for all datasets. GARLI version 0.951 (Zwickl, 2006) was run on the best nucleotide tree from PhyML for 2 million generations to optimize tree topology and branch lengths. Additionally, ML trees were estimated using the TVM+I+F model of base substitution in combination with tree-bisection-reconnection (TBR) searches using the PAUP* version 4.0b10 package (Swofford, 2003). For each of the trees, the reliability of all phylogenetic groupings was determined through a non-parametric bootstrap resampling analysis: either 500 replicates of ML trees using the GTR+I+F model and the SPR method using PhyML, or 1000 replicates of neighbour-joining trees estimated under the ML substitution model using PAUP*. All trees are available from the authors upon request. Trees were visualized with the FigTree program version 1.3.1 (http://tree.bio.ed.ac.uk/software/figtree/). Trees were rooted on the outgroup strains (A/Duck/Hokkaido/33/80 for HA1 and A/Japan/305/1957 for NA) or on the 1968 pandemic strain (A/Hong Kong/1/1968).

Amino acid alignments were used in combination with the HIVw+F4 (HAI) or HIVw+I+F4+F (NA) model of protein evolution to infer ML trees using the PhyML package version 3.0 (Guindon & Gascuel, 2003) by means of a full heuristic search and the SPR method.

Genetic maps. Amino acid sequence alignments were used to calculate a distance matrix with the number of amino acid substitutions between pairs of strains to produce ‘genetic maps’, as described previously (Smith et al., 2004). Genetic mapping is a way to facilitate a quantitative analysis and visualization of genetic data. In a genetic map, the distance between isolates A and B corresponds to the number of amino acid substitutions between a particular protein-coding region of isolates A and B. Thus, each difference in an amino acid alignment can be thought of as specifying a target distance for the points in a genetic map. Modified MDS methods are then used to arrange the points between two isolates in a genetic map to best satisfy the target distances specified by the amino acid alignment distance matrix. The result is a map in which the distance between points represents the number of amino acid substitutions in a particular protein-coding region between isolates. To avoid underestimating genetic distances due to sequential mutations at the same location, a threshold needs to be applied. For example, when a threshold of 30 is applied to the alignment, the software is free to use a distance of at least 30 aa between two points that have a 30 aa difference in order to retrieve the lowest error function. For HAI, the previously determined threshold of 30 was used (Smith et al., 2004). Threshold considerations for NA were based on adequate correlation between observed distances and genetic map distances and sufficient correlation between ML amino acid tree distances and genetic map distances. The threshold that approached both considerations best for NA was 20 (data not shown). The correlation between the number of amino acid substitutions and the corresponding distances between strains in the genetic map was 0.93 (data not shown), indicating that the two-dimensional genetic map is a reasonable representation of the target amino acid distance matrix. For software, see http://www.antigenic-cartography.org.

Detection of reassortment. ML trees of NA and HAI were used in TreeMap version 1.0 (http://taxonomy.zoology.gla.ac.uk/rod/treemap.html). TreeMap, originally designed for comparing host and parasite trees, was applied solely to display a tanglegram between the HA and NA phylogenies. The twines were colour coded according to the HA antigenic clusters (Smith et al., 2004).

Reassortment events were identified by the GiRaF program (Nagarajan & Kingsford, 2011). Nucleotide alignments of HAI or NA were used as input for MrBayes (Huelsenbeck & Ronquist, 2001; Ronquist & Huelsenbeck, 2003) to sample 1000 unrooted candidate trees with the GTR+I+F4 substitution model, a burn-in of 100 000 and sampling every 200 iterations. These trees were subsequently used to model the phylogenetic uncertainty for each segment with the GiRaF program under default settings. This procedure was repeated 100 times with ten independent MrBayes runs for NA and ten independent MrBayes runs for HAI.

Estimation of nucleotide substitution rates and times of divergence. To identify potential errors in sequence data annotation that might affect the clock estimation, the ML nucleotide HAI and NA trees were exported to Path-O-Gen version 1.3 (http://tree.bio.ed.ac.uk/software/pathogen; Drummond et al., 2003) and a linear regression plot for the years of sampling versus root-to-tip distance was generated (Fig. S4). No anomalies were seen in the NA and HAI datasets, which both behaved in a clock-like manner ($R^2=0.977$ and $R^2=0.967$, respectively).

Overall rates of evolutionary change (nucleotide substitutions per site per year), MRCA in years and relative genetic diversity were estimated using the BEAST program version 1.6.2 (http://beast.bio.ed.ac.uk/; Drummond & Rambaut, 2007). For all analyses, the uncorrelated log-normal relaxed molecular clock was used to accommodate variation in molecular evolutionary rate among lineages in combination with the SRD06 codon position model, with a different rate of nucleotide substitution for the first plus second versus the third codon position, and the HKY85 substitution model (Shapiro et al., 2006). Isolation dates were added to calibrate the molecular clock. This analysis was conducted with a time-aware linear Bayesian skyline coalescent tree prior (Minin et al., 2008) over the unknown tree space, with relatively uninformative priors on all model parameters. Two independent Bayesian MCMC analyses, performed for HAI and NA for 100 million states, sampling every 1000 states, were performed. Convergences and effective sample sizes of the estimates were checked using Tracer version 1.5 (http://tree.bio.ed.ac.uk/software/tracer/) and the first 10% of each chain was discarded as burn-in. Uncertainty in parameter estimates is reported as values of the 95% HPD.

Selection pressures. To determine the degree of natural selection acting on HAI and NA, the mean number of $d_s$ and $d_{is}$ substitutions per site ($d_{is}/d_s$ ratio or $w$) was estimated using the ‘one-ratio’ model of the Codeml program implemented in the PAML package (Yang, 2004).
2007). Codeml uses the codon substitution model of Goldman & Yang (1994) for protein-coding DNA sequences and was used in combination with the ML tree rooted on A/Hong Kong/1/68. Selection pressures on the internal and external branches of the HA1 and NA ML tree were estimated with the ‘two-ratio’ model of the Codeml program.

To determine the positively and negatively selected sites along internal branches, the IFEL method was used (Kosakovsky Pond et al., 2006). Sites subject to episodic diversifying selection were identified using MEME (Kosakovsky Pond et al., 2011). Both methods were accessed through the Datamonkey interface (http://www.datamonkey.org) and the GTR (HA1) or TVM (NA) nucleotide substitution model was used. Sites with a P value <0.05 were reported as positively selected sites.

Directional positive selection was detected with the DEPS test implemented in the HyPhy package (Kosakovsky Pond et al., 2005). For this procedure, the ML trees with the A/Hong Kong/1/68 pandemic strain as outgroup were used in combination with the amino acid alignment and the HIVw model of protein evolution.

Positively selected sites were shown on the monomeric and tetrameric NA structure constructed using MacPyMOL (PyMOL Molecular Graphics System, version 13) on subtype N2 (PDB code 2BAT; Varghese et al., 1992).

ACKNOWLEDGEMENTS

The authors would gratefully like to thank Dr Mathilde Richard for constructive discussions, Stefan van der Vliet and Oanh Vuong for excellent technical assistance and Ton Marzec for providing part of the influenza A (H3N2) virus isolates. We are grateful for the outstanding training received at the 15th and 16th International Bioinformatics Workshop on Virus Evolution and Molecular Epidemiology, September 2009, Rotterdam and 2010, Baltimore (http://regaweb.med.kuleuven.be/workshop/). This work was supported by an NWO-VICI grant and NIH contract no. HHSN26620070001OC.

REFERENCES


Zwickl, D. J. (2007). Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. PhD thesis, The University of Texas at Austin, TX, USA.