

# Genetic Structure of Human A/H1N1 and A/H3N2 Influenza Virus on Corsica Island: Phylogenetic Analysis and Vaccine Strain Match, 2006–2010

Alessandra Falchi<sup>1,2,3,4\*</sup>, Jean Pierre Amoros<sup>1,2,3</sup>, Christophe Arena<sup>1,2,3,4</sup>, Jean Arrighi<sup>4</sup>, François Casabianca<sup>5</sup>, Laurent Andreoletti<sup>6,7</sup>, Clément Turbelin<sup>1,2</sup>, Antoine Flahault<sup>8</sup>, Thierry Blanchon<sup>1,2</sup>, Thomas Hanslik<sup>1,2</sup>, Laurent Varesi<sup>3</sup>

1 INSERM, UMR-S 707, F-75012, Paris, France, 2 University Pierre and Marie Curie, Université Paris 06, UMR-S U707, F-75012, Paris, France, 3 Laboratory of Virology, University of Corsica, Corsica, France, 4 Observatoire Régionale de la Santé, Ajaccio, France, 5 Institut national de la Recherche Agronomique, Corte, France, 6 Unité de Virologie Médicale et Moléculaire, Centre Hospitalier Universitaire, Reims, France, 7 IFR 53/EA-4303 (DAT/PPCIDH), Faculté de Médecine, Reims, France, 8 EHESP School of Public Health, Rennes, France

## Abstract

**Background:** The aim of this study was to analyse the genetic patterns of Hemagglutinin (HA) genes of influenza A strains circulating on Corsica Island during the 2006–2009 epidemic seasons and the 2009–2010 pandemic season.

**Methods:** Nasopharyngeal samples from 371 patients with influenza-like illness (ILI) were collected by General Practitioners (GPs) of the *Sentinelles* Network through a randomised selection routine.

**Results:** Phylogenetic analysis of HA revealed that A/H3N2 strains circulating on Corsica were closely related to the WHO recommended vaccine strains in each analyzed season (2006–2007 to 2008–2009). Seasonal Corsican influenza A/H1N1 isolated during the 2007–2008 season had drifted towards the A/Brisbane/59/2007 lineage, the A/H1N1 vaccine strain for the 2008–2009 season. The A/H1N1 2009 (A/H1N1pdm) strains isolated on Corsica Island were characterized by the S220T mutation specific to clade 7 isolates. It should be noted that Corsican isolates formed a separate sub-clade of clade 7 as a consequence of the presence of the fixed substitution D222E. The percentages of the perfect match vaccine efficacy, estimated by using the  $p_{\text{epitope}}$  model, against influenza viruses circulating on Corsica Island varied substantially across the four seasons analyzed, and tend to be highest for A/H1N1 compared with A/H3N2 vaccines, suggesting that cross-immunity seems to be stronger for the H1 HA gene.

**Conclusion:** The molecular analysis of the HA gene of influenza viruses that circulated on Corsica Island between 2006–2010 showed for each season the presence of a dominant lineage characterized by at least one fixed mutation. The A/H3N2 and A/H1N1pdm isolates were characterized by multiples fixation at antigenic sites. The fixation of specific mutations at each outbreak could be explained by the combination of a neutral phenomenon and a founder effect, favoring the presence of a dominant lineage in a closed environment such as Corsica Island.

**Citation:** Falchi A, Amoros JP, Arena C, Arrighi J, Casabianca F, et al. (2011) Genetic Structure of Human A/H1N1 and A/H3N2 Influenza Virus on Corsica Island: Phylogenetic Analysis and Vaccine Strain Match, 2006–2010. PLoS ONE 6(9): e24471. doi:10.1371/journal.pone.0024471

**Editor:** Malcolm Gracie Semple, University of Liverpool, United Kingdom

**Received:** April 27, 2011; **Accepted:** August 11, 2011; **Published:** September 14, 2011

**Copyright:** © 2011 Falchi et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** Corsican Economic Development Agency and the FEDER programme provided funding and support for the design and conduct of the study. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: falchi@u707.jussieu.fr

## Introduction

Type A influenza viruses are major pathogens for humans. Influenza A virus is further classified based on the antigenic properties of its surface glycoproteins, the Haemagglutinin (HA) and the Neuraminidase (NA) [1]. Influenza A/H3N2 and A/H1N1 subtypes are the major subtypes currently circulating in human populations [2]. HA is of special interest due to its role in the viral entry mechanism and immune recognition. It consists of two subunits: HA1, which contains the receptor-binding and antigenic domains, and the HA2 subunit, which is responsible for the fusion of the virion with the endosomal membrane of the host

cell [3]. The HA1 subunit undergoes a process termed positive Darwinian selection through continuous antigenic mutations that allow the virus to evade the host's humoral immune response [4].

This process is called antigenic drift. Variants that best escape the host immune response are thought to have a significant reproductive advantage [5]. Specifically, the variable antigenic regions of the HA1 domain are potential targets of neutralising antibodies, and thus amino acid substitutions at these regions (A to E) have been associated with annual epidemics in humans. In contrast, the amino acids within the receptor binding site (RBD) of the HA1 domain are relatively conserved. Variations have also been observed in the N-linked glycosylation sites of HA, with some

strains having lost or gained a glycosylation site and potentially an altered glycoprotein function. Concerning the NA, antigenic variations usually occur in the catalytic or framework sites of the protein.

Another process, called reassortment, is also considered to be a major force in the evolution of influenza virus [5]. It occurs when the virus acquires an HA of a different influenza virus subtype via reassortment of one or more gene segments, and this is thought to be the basis for the more devastating influenza pandemics [6]. New influenza pandemics may emerge through reassortment with strains from swine or avian reservoirs [7].

The first influenza pandemic of this century was declared in April of 2009, with the emergence of a novel H1N1 influenza A virus strain in Mexico and the USA [8,9]. In France, the A/(H1N1)2009 pandemic wave spanned from September 6, 2009 to December 26, 2009 ([www.sentiweb.fr](http://www.sentiweb.fr)). During the epidemic, the pandemic strain was almost exclusively found among influenza virus isolates [10].

The effectiveness of annual influenza vaccines depends on the selection of component strains that offer optimal immunity from the numerous variants in the global influenza virus circulation. Studies based on sequencing analyses of viruses can be utilised as surveillance tools and can contribute to the vaccine selection process when they are combined with classical serological antigenic analysis.

The aim of this study was to analyse the genetic patterns of HA genes of influenza A strains circulating on Corsica Island during the 2006–2009 epidemic seasons and the 2009–2010 pandemic season.

## Methods

### Ethics Statement

Based on French national laws ethical permissions (Law 1121-1-1° R. 1121-2) are not required for specific microbiological diagnostics treatment of the patients and further characterization of the viruses. All samples were coded and tested anonymously. Patient information was stored according to national regulations and access to such data was restricted (permissions CNIL 471393).

### Clinical samples and epidemic data

Three hundred and seventy one samples from patients with influenza-like illness (ILI) were collected by *Sentinelles* General Practitioners (GPs) of the *Sentinelles* Network starting with the 2006–2007 season (Fig. 1). All cases represented mild infections. Nasopharyngeal swabs and clinical data were collected through a randomised selection routine. Doctors included the first patient of each week, of any age. Corsican *Sentinelles* GPs swabbed patients presenting an ILI according to the *Sentinelles* definition: “sudden onset of fever  $>39^{\circ}\text{C}$  ( $>102^{\circ}\text{F}$ ) with respiratory signs and myalgia” [11].

The nasopharyngeal swabs were sent by mail in 2 ml of viral transport medium within two days to the virological laboratory. Detailed demographic and clinical data (the time of onset of symptoms, reported symptoms, physical findings and influenza vaccination status) were obtained from patients during the medical visit.

### RT-PCR Assays for Influenza Virus Detection and Subtyping

After processing the clinical materials by the standard laboratory procedures, the individual sample was inoculated in the Madin Canine Kidney (MDCK) cell lines for the isolation of viruses. The inoculated MDCK cell line was observed at  $34^{\circ}\text{C}$  for

cytopathic effects. The viruses were passaged three times to obtain sufficient virus titers for virus identification.

RNA was extracted from 140  $\mu\text{l}$  of clinical sample and of the supernatant using the QIAmp Viral RNA Mini kit® (QIAGEN, Courtaboeuf, France). Multiplex RT-PCR assays were performed using the One-step RT-PCR kit® (QIAGEN, Courtaboeuf, France) and using primers based on the M, HA and NA glycoprotein gene sequences of seasonal influenza A/H1N1 and A/H3N2 viruses [12–14]. Positive and negative controls were included in each multiplex RT-PCR.

In order to detect and assign the A/H1N1pdm strains isolated from patients, a real-time PCR assay was performed using a specified RT-PCR protocol provided by the French National Influenza Centre [15]. Positive and negative controls were included in each RT-PCR run.

For the detection of A/H1N1pdm, the M RT-PCR was associated with a specific RT-PCR targeting the HA gene. This strategy is pertinent in the context of a pandemic as the new virus may drift, mainly due to mutation in hemagglutinin and/or neuraminidase genes, but variants will still be detected by this universal M RT-PCR [15].

### Nucleotide sequencing

The HA sequences were amplified using the primer set for A/H1N1, A/H1N1pdm and A/H3N2 influenza virus (primers sequences are available on request). The resulting amplicons were analyzed by 2% agarose gel electrophoresis. Double-stranded sequencing of the purified PCR products was performed using an Applied Biosystems Sequencer (ABI 3700 Perkin-Elmer).

### Nucleotide sequence accession numbers

The nucleotide sequence data from this study were deposited with the following accession numbers: for A/H1N1 isolates between JF701798 and JF701831; for A/H3N2 isolates between JF701832 and JF711877 and for A/H1N1pdm isolates between JF01878 and JF701906.

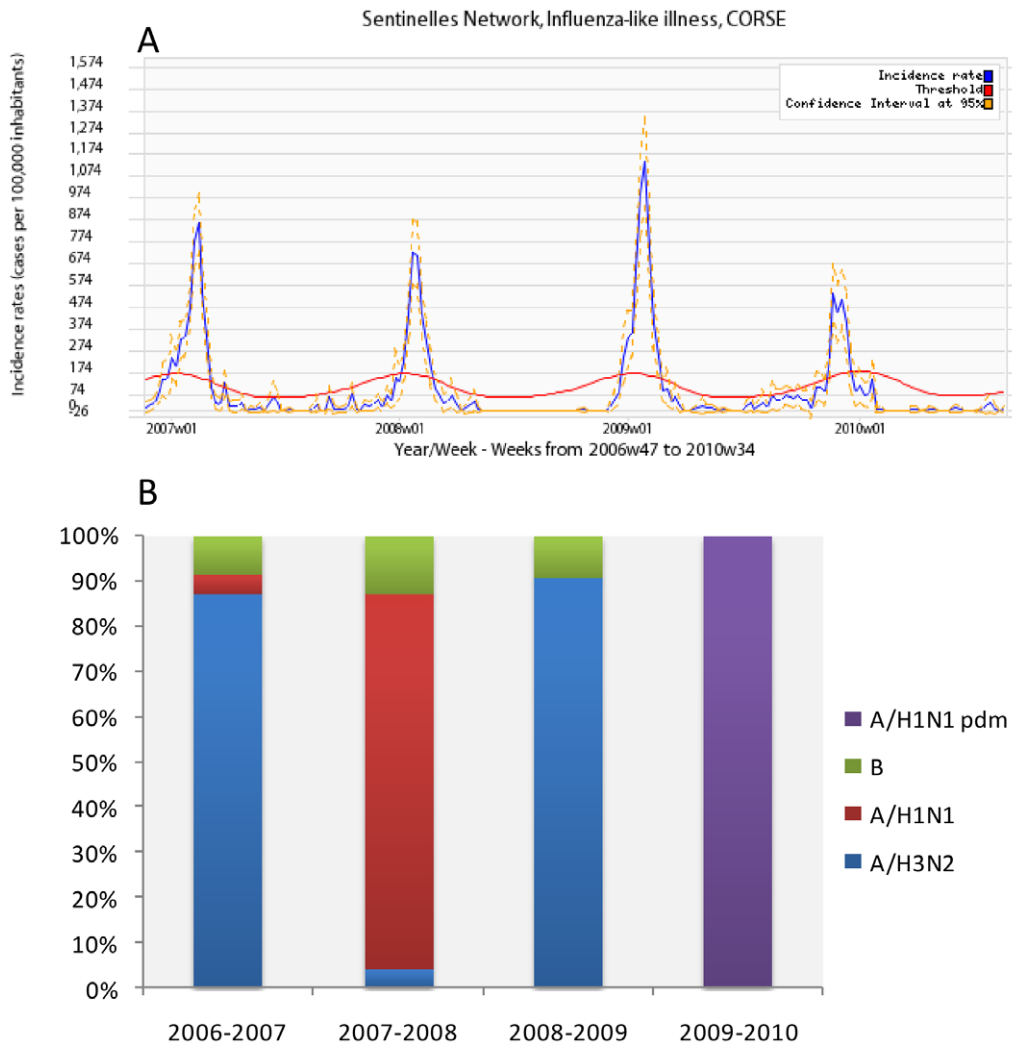
### Phylogenetic analysis

Nucleotide sequences were aligned with the CLUSTAL X [16] program for the major coding regions of the three segments: HA (875 bp) for A/H3N2 isolates from 2007–2008 and 2008–2009, HA (1056 bp) for A/H1N1 isolates from 2006–2007 and 2007–2008 and HA (1018 bp) for A/H1N1pdm from 2009–2010. The Find Model program [17] was used to identify the optimal evolutionary model that best fitted our sequence dataset. Akaike Information Criteria and the hierarchical likelihood ratio test indicated that the HKY+ $\Gamma$  model was the best fit to the data. To infer the evolutionary relationship for the influenza virus analyzed, we employed the Maximum likelihood (ML) trees constructed under the HKY+ $\Gamma$  model using software from the PhyML program [18].

The genomic sequences of the vaccine strains and the other strains used in this study were obtained from the Influenza Sequences Database (<http://www.flu.lanl.gov>).

### Nucleotide diversity

The average pairwise nucleotide diversity among sequences was estimated for the A/H3N2, the A/H1N1 seasonal influenza viruses and for the A/H1N1 pdm virus population. For this analysis, a maximum likelihood (ML) method was employed, available in the MEGA4 package [19], with the standard deviation estimated using 100 bootstrap replicates.



**Figure 1. Weekly influenza epidemic time series.** (A). Influenza seasons in Corsica Island. Incidence rate (cases per 100,000 inhabitants) in blue. In red, the epidemic threshold since 2006–2007 season calculated by a periodic regression model applied to the former observed data (Serfling's method). In yellow confidence interval at 95%. (B) Proportions of influenza viruses A/H1N1 pdm; A/H1N1; A/H3N2 and B in Corsica islands. doi:10.1371/journal.pone.0024471.g001

### Prediction of N-Glycosylation

Potential N-glycosylation sites (amino acids Asn-X-Ser/Thr, where X is not Asp or Pro) were predicted using nine artificial neural networks with the NetNGlyc server 1.0 [20]. A threshold value of 0.5 average potential score was set to predict glycosylated sites.

### Estimation of vaccine efficacy (VE) using the $p_{\text{epitope}}$ model

Antigenic distance is a quantity that should define difference of viral strains, as determined by the human immune system. The  $p_{\text{epitope}}$  sequence-based method has been shown to be an effective antigenic distance measure between two influenza viral strains [21–24]. The meta-analysis of epidemiological human VE studies shows that the single dominant epitope is the critical region that determines the epidemiological VE [21–24]. There are five non-overlapping epitopes on the surface of H3 and H1 HA molecule, namely epitopes A–E, to which different sets of antibodies bind. In each epitope, the p value is defined as the fraction of mutated amino acids [21]. The dominant epitope is defined as the epitope with the greatest p value. The greatest p value is  $p_{\text{epitope}}$  [21].

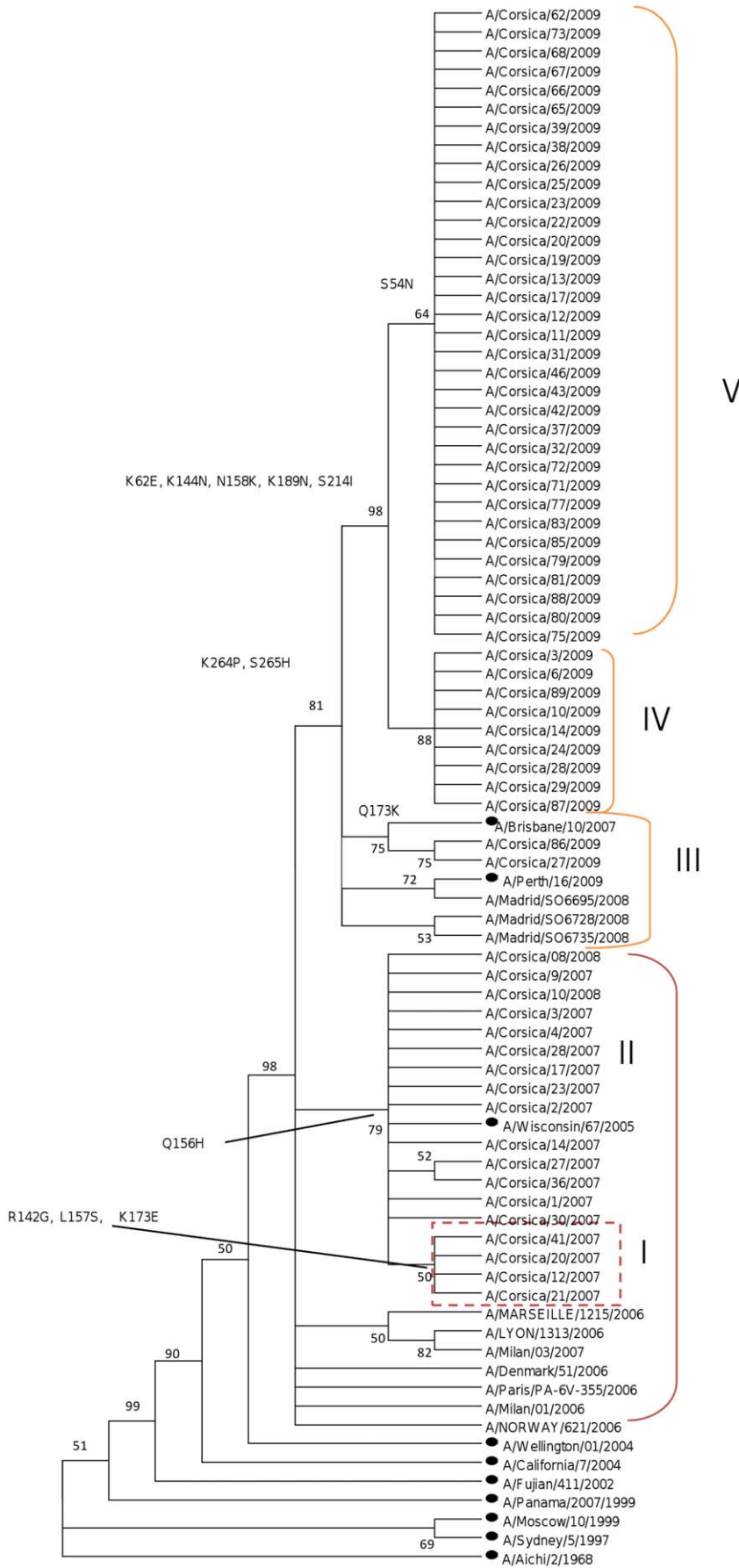
Epidemiological data on the vaccine efficacies in previous flu seasons when A/H3N2 or A/H1N1 subtype was dominant were collected [21–24]. The identities of the vaccine strains and dominant circulating strains were also obtained to calculate  $p_{\text{epitope}}$ .

The A/H3N2 and A/H1N1 VE correlates with  $p_{\text{epitope}}$  with  $R^2 = 0.81$  and  $R^2 = 0.68$ , respectively. This correlation shows that  $p_{\text{epitope}}$  defined by the single dominant epitope is a quantitative definition of antigenic distance.

The relation between VE and  $p_{\text{epitope}}$  is given by  $E = -2.47 \times p_{\text{epitope}} + 0.47$  (1) for A/H3N2 influenza virus and by  $E = -1.19 \times p_{\text{epitope}} + 0.53$  (2) for A/H1N1 strains. These two fitted models predict a vaccine effectiveness of 47% when  $p_{\text{epitope}} = 0$  for A/H3N2 and of 53% when  $p_{\text{epitope}} = 0$  for A/H1N1 [21,23].

The comparison between A/H3N2 and A/H1N1 vaccine effectiveness illustrates that A/H1N1 vaccine has higher effectiveness than the A/H3N2 vaccine as a function of  $p_{\text{epitope}}$ . This observation suggests that the host immune system is more effective at recognizing and eliminating the A/H1N1 virus ( $p_{\text{epitope}} = 0$ ), and that humoral cross-immunity is stronger for H1 HA ( $p_{\text{epitope}} > 0$ ).

By considering only the substitutions that occur in the epitopes, we calculated the value of the  $p_{\text{epitope}}$  (as the number of mutations



**Figure 2. Phylogenetic tree of the HA domain of the A/H3N2 strains from Corsica Island.** Maximum likelihood phylogenetic tree analysis of HA genes of A/H3N2 strains circulating in Corsica Island from 2006 to 2009 (\* represents vaccine strains). doi:10.1371/journal.pone.0024471.g002

within an antibody antigenic site divided by the number of amino acids defining the site), between the influenza virus strains representatives of the dominant lineage isolated in Corsica Island with respect to the vaccine strain of each corresponding season. The antigenic epitope which has the greatest percentage of mutations was considered as dominant. Percentages of the perfect match VE were estimated by using the equation 1 and 2 described previously.

## Results

### Prevalence of influenza viruses on Corsica Island from 2006 to 2010

A total of 371 samples were collected and analyzed at the *Sentinelles* site on Corsica Island in the four-year period from 2006–2010. Influenza-like illness activity and the relative prevalence of influenza virus subtypes from season to season on Corsica Island are shown in Fig. 1 (A) and (B), respectively.

In 2006–2007 [25], 134 samples were collected from February 2<sup>nd</sup> to March 16<sup>th</sup> 2007. Out of 134 samples, 93 (69.4%) were positive and 85 (91.4%) were Influenza type A, whereas influenza B was detected in 8 (6%) of 93 patients. Of the 85 influenza A strains identified, 81 (95%) were sub-typed as A/H3N2 and 4 (5%) were sub-typed as A/H1N1.

In 2007–2008, 85 samples were collected from January 30<sup>th</sup> to February 28<sup>th</sup> 2008. Out of 85 samples, 70 (83%) were positive; 61 (87%) were influenza type A and 9 (13%) were typed as influenza B. Of the 61 influenza A strains identified, 58 (95%) were sub-typed as A/H1N1 and 3 (5%) were sub-typed as A/H3N2.

In 2008–2009, 98 samples were collected from November 2<sup>nd</sup> 2008 to February 18<sup>th</sup> 2009. Out of 98 samples, 64 (65%) were positive; 58 (91%) were Influenza A and 6 (9%) were influenza B. Of the 58 influenza A strains identified, 100% were sub-typed as A/H3N2.

In 2009–2010, 54 samples were collected from November 3<sup>rd</sup> 2009 to January 11<sup>th</sup> 2010. Out of the 54 samples, 33 (61%) were positive. All positive samples were A/H1N1pdm influenza viruses.

### Influenza A/H3N2 viruses

The HA sequences of the A/H3N2 viruses circulating on Corsica Island during the 2006–2007 (N = 16), 2007–2008 (N = 2) and the 2008–2009 seasons (N = 45) were compared with the vaccine strains.

Phylogenetic analysis showed that HA sequences of the A/H3N2 isolates formed clusters representative for the 2006–2007, 2007–2008 and 2008–2009 seasons.

As previously reported [25], the HA sequences of the A/H3N2 isolated during the 2006–2007 season appeared to be closely related to the A/Wisconsin/67/2005-like lineage (vaccine strain for 2006–2007 and 2007–2008 in the Northern Hemisphere) (Fig. 2). All isolates were characterized by the substitution H156Q (Fig. 2) with respect to A/Wisconsin/67/2005. Several strains were characterized by H156Q associated with at least one of these three additional substitutions: R142G, L157S and K173E (clade I; Fig. 2). The nucleotide diversity of the A/H3N2 2006–2007 population was  $0.002 \pm 0.001$ . The nucleotide identities between the 2006–2007 strains and the A/Wisconsin/67/2005-like lineage were 99.7% and 99.4% based on amino acids.

The A/H3N2 viruses isolated on Corsica Island during the 2007–2008 season clustered with the A/H3N2 viruses isolated

during the 2006–2007 season and were related to A/Wisconsin/67/2005. These two strains were not characterized by the mutation H156Q (Fig. 2; Table 1). The nucleotide identities between the 2007–2008 strains and the A/Wisconsin/67/2005-like lineage were 99.6% and 99.4% based on amino acids. The nucleotide identities between the 2007–2008 strains and the A/Brisbane/10/2007-like lineage (vaccine strain for the 2008–2009 season in the Northern Hemisphere) were 99.8% and 88.6% based on amino acids.

The HA sequences of A/H3N2 viruses isolated on Corsica Island during the 2008–2009 season had drifted from the A/Wisconsin/67/2005-like lineage (98.1% and 97% similarity based on nucleotide and amino acids sequences, respectively). These strains were closely to the A/Brisbane/10/2007-like lineage (vaccine strain for the 2008–2009 season in the Northern Hemisphere) (Fig. 2; Table 1). The nucleotide identities between the 2009 strains and the A/Brisbane/10/2007-like lineage were 99.3% and 98.9% based on amino acids. The nucleotide identities between the 2009 strains and the A/Perth/16/2009-like lineage (vaccine strain for the 2009–2010 season in the Northern Hemisphere) were 98.2% and 96.2% based on amino acids.

The H3 nucleotide sequences of the 2009 virus isolates clustered into three sub-clades (III, IV and V). All groups were characterized by two variations (K264P and S265H) compared to the vaccine strain A/Brisbane/10/2007 (vaccine strain for the 2008–2009 season in the Northern Hemisphere) and by five altered amino acids (K62E, K144N, N158K, K189N and S214I) with respect to A/Perth/16/2009 (vaccine strain for the 2009–2010 season in the Northern Hemisphere). The sub-clade V was characterized by the supplementary mutation S54N when compared to A/Brisbane/10/2007 and A/Perth/16/2009 (Fig. 2). The sub-clade III was characterized by the Q173K mutation with respect to A/Perth/16/2009. The sub-clade IV showed one supplementary variation compared to A/Perth/16/2009 (Q173K and L194D). The K144N mutation was located at antigenic site A, while the N158K and K189N mutations were located at site B, the S241I and Q173K mutations were at site D and the K264P and S265H mutations were at site E.

The nucleotide diversity of this A/H3N2 2008–2009 population was  $0.005 \pm 0.002$ . The amino acids at the terminal SA of all H3 isolates were Ile (I)-226 and Ser (S)-228.

Nine potential N-glycosylation sites were predicted in HA1 of the H3 subtype (8, 22, 38, 63, 126, 133, 165, 246 and 285) with a threshold  $>0.5$ . These sites were found to be conserved in all isolates obtained during the 2006–2009 seasons. Strains isolated in 2007–2008, 2008–2009 and ten strains (4.2%; 10/42) isolated in 2006–2007 had lost glycosylation at position 144.

### Influenza A/H1N1 virus

A/H1N1 isolates from the 2006–2007 (N = 1) and 2007–2008 (N = 34) season were compared with vaccine strains for phylogenetic analysis (Fig. 3; Table 2). The HA sequences of influenza viruses isolated on Corsica Island during the 2007–2008 season were separated into two sub-clusters (I and II) as shown in Fig. 3. Both clades were characterized by the N39D and I189D (antigenic site B) with respect to A/Brisbane/59/2007 (vaccine strain of the 2008–2009 season in the Northern Hemisphere) (Table 2). Clade II was characterized by the A62G, S125N (antigenic site B), W127G, C139S and S140C (antigenic site A), Y209F (at antigenic site D) and E246Q (Table 2). The isolate A/Corsica/7/2007 was

**Table 1.** Amino acid substitution in A/H3N2 influenza viruses isolated in Corsica Island from 2008 to 2009.

Amino acid position	50	54	62	122	138	140	144	156	158	173	186	189	194	214	223	264	265
<b>Antigenic sites</b>	<b>C</b>			<b>A</b>	<b>A</b>	<b>A</b>	<b>A</b>	<b>B</b>	<b>B</b>	<b>D</b>	<b>B</b>	<b>B</b>	<b>D</b>	<b>D</b>		<b>E</b>	<b>E</b>
<b>A/Perth/16/2009</b>	<b>E</b>	<b>S</b>	<b>K</b>	<b>N</b>	<b>A</b>	<b>I</b>	<b>K</b>	<b>H</b>	<b>N</b>	<b>Q</b>	<b>G</b>	<b>K</b>	<b>L</b>	<b>S</b>	<b>V</b>	<b>K</b>	<b>S</b>
<b>A/Brisbane/10/2007</b>	.	.	<b>E</b>	.	.	.	<b>N</b>	.	<b>K</b>	<b>K</b>	.	<b>N</b>	.	<b>I</b>	.	.	.
<b>A/Wisconsin/67/2005</b>	<b>G</b>	.	<b>E</b>	<b>D</b>	<b>S</b>	<b>K</b>	<b>N</b>	<b>H</b>	<b>K</b>	<b>K</b>	<b>V</b>	<b>N</b>	.	<b>I</b>	<b>I</b>	.	.
A/Corsica/3/2009	.	.	<b>E</b>	.	.	.	<b>N</b>	.	<b>K</b>	.	.	<b>N</b>	.	<b>I</b>	.	<b>P</b>	<b>H</b>
A/Corsica/79/2009	.	.	<b>E</b>	.	.	.	<b>N</b>	.	<b>K</b>	.	.	<b>N</b>	.	<b>I</b>	.	<b>P</b>	<b>H</b>
A/Corsica/85/2009	.	.	<b>E</b>	.	.	.	<b>N</b>	.	<b>K</b>	.	.	<b>N</b>	.	<b>I</b>	.	<b>P</b>	<b>H</b>
A/Corsica/86/2009	.	.	<b>E</b>	.	.	.	<b>N</b>	.	<b>K</b>	<b>K</b>	.	<b>N</b>	<b>P</b>	<b>I</b>	.	.	.
A/Corsica/6/2009	.	.	<b>E</b>	.	.	.	<b>N</b>	.	<b>K</b>	.	.	<b>N</b>	.	<b>I</b>	.	<b>P</b>	<b>H</b>
A/Corsica/89/2009	.	.	<b>E</b>	.	.	.	<b>N</b>	.	<b>K</b>	.	.	<b>N</b>	.	<b>I</b>	.	<b>P</b>	<b>H</b>
A/Corsica/83/2009	.	.	<b>E</b>	.	.	.	<b>N</b>	.	<b>K</b>	.	.	<b>N</b>	.	<b>I</b>	.	<b>P</b>	<b>H</b>
A/Corsica/77/2009	.	.	<b>E</b>	.	.	.	<b>N</b>	.	<b>K</b>	.	.	<b>N</b>	.	<b>I</b>	.	<b>P</b>	<b>H</b>
A/Corsica/71/2009	.	<b>N</b>	<b>E</b>	.	.	.	<b>N</b>	.	<b>K</b>	.	.	<b>N</b>	.	<b>I</b>	.	<b>P</b>	<b>H</b>
A/Corsica/72/2009	.	.	<b>E</b>	.	.	.	<b>N</b>	.	<b>K</b>	.	.	<b>N</b>	.	<b>I</b>	.	<b>P</b>	<b>H</b>
A/Corsica/32/2009	.	<b>N</b>	<b>E</b>	.	.	.	<b>N</b>	.	<b>K</b>	.	.	<b>N</b>	.	<b>I</b>	.	<b>P</b>	<b>H</b>
A/Corsica/37/2009	.	.	<b>E</b>	.	.	.	<b>N</b>	.	<b>K</b>	.	.	<b>N</b>	.	<b>I</b>	.	<b>P</b>	<b>H</b>
A/Corsica/42/2009	.	<b>N</b>	<b>E</b>	.	.	.	<b>N</b>	.	<b>K</b>	.	.	<b>N</b>	.	<b>I</b>	.	<b>P</b>	<b>H</b>
A/Corsica/43/2009	.	<b>N</b>	<b>E</b>	.	.	.	<b>N</b>	.	<b>K</b>	.	.	<b>N</b>	.	<b>I</b>	.	<b>P</b>	<b>H</b>
A/Corsica/46/2009	.	<b>N</b>	<b>E</b>	.	.	.	<b>N</b>	.	<b>K</b>	.	.	<b>N</b>	.	<b>I</b>	.	<b>P</b>	<b>H</b>
A/Corsica/31/2009	.	<b>N</b>	<b>E</b>	.	.	.	<b>N</b>	.	<b>K</b>	.	.	<b>N</b>	.	<b>I</b>	.	<b>P</b>	<b>H</b>
A/Corsica/10/2009	.	.	<b>E</b>	.	.	.	<b>N</b>	.	<b>K</b>	.	.	<b>N</b>	.	<b>I</b>	.	<b>P</b>	<b>H</b>
A/Corsica/11/2009	.	.	<b>E</b>	.	.	.	<b>N</b>	.	<b>K</b>	.	.	<b>N</b>	.	<b>I</b>	.	<b>P</b>	<b>H</b>
A/Corsica/12/2009	.	.	<b>E</b>	.	.	.	<b>N</b>	.	<b>K</b>	.	.	<b>N</b>	.	<b>I</b>	.	<b>P</b>	<b>H</b>
A/Corsica/17/2009	.	<b>N</b>	<b>E</b>	.	.	.	<b>N</b>	.	<b>K</b>	.	.	<b>N</b>	.	<b>I</b>	.	<b>P</b>	<b>H</b>
A/Corsica/13/2009	.	<b>N</b>	<b>E</b>	.	.	.	<b>N</b>	.	<b>K</b>	.	.	<b>N</b>	.	<b>I</b>	.	<b>P</b>	<b>H</b>
A/Corsica/14/2009	.	.	<b>E</b>	.	.	.	<b>N</b>	.	<b>K</b>	.	.	<b>N</b>	.	<b>I</b>	.	<b>P</b>	<b>H</b>
A/Corsica/19/2009	.	<b>N</b>	<b>E</b>	.	.	.	<b>N</b>	.	<b>K</b>	.	.	<b>N</b>	.	<b>I</b>	.	<b>P</b>	<b>H</b>
A/Corsica/20/2009	.	<b>N</b>	<b>E</b>	.	.	.	<b>N</b>	.	<b>K</b>	.	.	<b>N</b>	.	<b>I</b>	.	<b>P</b>	<b>H</b>
A/Corsica/22/2009	.	<b>N</b>	<b>E</b>	.	.	.	<b>N</b>	.	<b>K</b>	.	.	<b>N</b>	.	<b>I</b>	.	<b>P</b>	<b>H</b>
A/Corsica/23/2009	.	<b>N</b>	<b>E</b>	.	.	.	<b>N</b>	.	<b>K</b>	.	.	<b>N</b>	.	<b>I</b>	.	<b>P</b>	<b>H</b>
A/Corsica/24/2009	.	.	<b>E</b>	.	.	.	<b>N</b>	.	<b>K</b>	.	.	<b>N</b>	.	<b>I</b>	.	<b>P</b>	<b>H</b>
A/Corsica/25/2009	.	.	<b>E</b>	.	.	.	<b>N</b>	.	<b>K</b>	.	.	<b>N</b>	.	<b>I</b>	.	<b>P</b>	<b>H</b>
A/Corsica/26/2009	.	.	<b>E</b>	.	.	.	<b>N</b>	.	<b>K</b>	.	.	<b>N</b>	.	<b>I</b>	.	<b>P</b>	<b>H</b>
A/Corsica/27/2009	.	.	<b>E</b>	.	.	.	<b>N</b>	.	<b>K</b>	<b>K</b>	.	<b>N</b>	<b>P</b>	<b>I</b>	.	.	.
A/Corsica/28/2009	.	.	<b>E</b>	.	.	.	<b>N</b>	.	<b>K</b>	.	.	<b>N</b>	.	<b>I</b>	.	<b>P</b>	<b>H</b>
A/Corsica/29/2009	.	.	<b>E</b>	.	.	.	<b>N</b>	.	<b>K</b>	.	.	<b>N</b>	.	<b>I</b>	.	<b>P</b>	<b>H</b>
A/Corsica/38/2009	.	.	<b>E</b>	.	.	.	<b>N</b>	.	<b>K</b>	.	.	<b>N</b>	.	<b>I</b>	.	<b>P</b>	<b>H</b>
A/Corsica/39/2009	.	.	<b>E</b>	.	.	.	<b>N</b>	.	<b>K</b>	.	.	<b>N</b>	.	<b>I</b>	.	<b>P</b>	<b>H</b>
A/Corsica/65/2009	.	.	<b>E</b>	.	.	.	<b>N</b>	.	<b>K</b>	.	.	<b>N</b>	.	<b>I</b>	.	<b>P</b>	<b>H</b>
A/Corsica/66/2009	.	.	<b>E</b>	.	.	.	<b>N</b>	.	<b>K</b>	.	.	<b>N</b>	.	<b>I</b>	.	<b>P</b>	<b>H</b>
A/Corsica/67/2009	.	<b>N</b>	<b>E</b>	.	.	.	<b>N</b>	.	<b>K</b>	.	.	<b>N</b>	.	<b>I</b>	.	<b>P</b>	<b>H</b>
A/Corsica/68/2009	.	<b>N</b>	<b>E</b>	.	.	.	<b>N</b>	.	<b>K</b>	.	.	<b>N</b>	.	<b>I</b>	.	<b>P</b>	<b>H</b>
A/Corsica/62/2009	.	<b>N</b>	<b>E</b>	.	.	.	<b>N</b>	.	<b>K</b>	.	.	<b>N</b>	.	<b>I</b>	.	<b>P</b>	<b>H</b>
A/Corsica/73/2009	.	<b>N</b>	<b>E</b>	.	.	.	<b>N</b>	.	<b>K</b>	.	.	<b>N</b>	.	<b>I</b>	.	<b>P</b>	<b>H</b>
A/Corsica/81/2009	.	<b>N</b>	<b>E</b>	.	.	.	<b>N</b>	.	<b>K</b>	.	.	<b>N</b>	.	<b>I</b>	.	<b>P</b>	<b>H</b>
A/Corsica/88/2009	.	<b>N</b>	<b>E</b>	.	.	.	<b>N</b>	.	<b>K</b>	.	.	<b>N</b>	.	<b>I</b>	.	<b>P</b>	<b>H</b>
A/Corsica/87/2009	.	.	<b>E</b>	.	.	.	<b>N</b>	.	<b>K</b>	.	.	<b>N</b>	.	<b>I</b>	.	<b>P</b>	<b>H</b>
A/Corsica/80/2009	.	.	<b>E</b>	.	.	.	<b>N</b>	.	<b>K</b>	.	.	<b>N</b>	.	<b>I</b>	.	<b>P</b>	<b>H</b>
A/Corsica/75/2009	.	.	<b>E</b>	.	.	.	<b>N</b>	.	<b>K</b>	.	.	<b>N</b>	.	<b>I</b>	.	<b>P</b>	<b>H</b>
A/Corsica/08/2008	<b>G</b>	.	<b>E</b>	<b>D</b>	<b>S</b>	<b>K</b>	<b>N</b>	.	<b>K</b>	<b>K</b>	<b>V</b>	<b>N</b>	.	<b>I</b>	<b>I</b>	.	.



**Table 1.** Cont.

Amino acid position	50	54	62	122	138	140	144	156	158	173	186	189	194	214	223	264	265
A/Corsica/10/2008	G		E	D	S	K	N		K	K	V	N		I	I		

doi:10.1371/journal.pone.0024471.t001

characterized by the substitutions I189D (antigenic site B) and G240R with respect to A/Brisbane/59/2007 (Fig. 3; Table 2).

The nucleotide identities between the Corsican A/H1N1 2008 strains and the A/Salomon/3/2006-like lineage were 98.9% and 98.7% based on amino acids. The nucleotide identities between the 2008 strains and the A/Brisbane/59/2007-like lineage were 97.3% and 96.9% based on amino acids.

Seven positions of potential glycosylation with a threshold value of >0.5 were predicted (at positions 15, 27, 58, 91, 129, 164 and 290). These sites were found to be conserved among all isolates in this study.

The nucleotide diversity of this the A/H1N1 2008–2009 population was  $0.005 \pm 0.002$ .

### Influenza A/H1N1 pdm virus

The phylogenetic tree of the HA nucleotide sequences of the Corsican pandemic strains and the vaccine strain is shown in Fig. 4. The H1 phylogeny showed that the amino acid substitutions P83S and S203T were found in all of the Corsican isolates. The substitution D222E was identified in 99% of the A/H1N1pdm viruses (Table 3). The substitution P83S were assigned to epitope E and the substitutions S203T and D222E to epitope D. Eight potential N-glycosylation sites are found in the HA molecule of the A/California/07/2009 virus, six of which reside in the HA1 and the remaining in the HA2 region. Four of six potential glycosylation sites residing in the HA1 region, with a threshold value of >0.5 at positions 28, 40, 104 and 304, were predicted among the A/H1N12009 Influenza virus isolated on Corsica Island.

The nucleotide diversity of the A/H1N1pdm isolates in 2009–2010 was  $0.002 \pm 0.001$ . The nucleotide identities between the 2009 strains and the A/California/7/2009-like lineage were 99.2% and 99.5% based on amino acids.

### Antigenic distance and vaccine efficacy

Vaccine efficacy has a linear correlation with the antigenic distance between the vaccine strain and the circulating virus strains [21]. The accumulation of epitope substitutions over time in influenza viruses strains isolated in Corsica Island are shown in Fig. 5. We calculated the  $p_{\text{epitope}}$  antigenic distance between the 2006–2007 vaccine strain A/Wisconsin/67/2005 and the A/H3N2 Corsican isolates circulating in Corsica Island during 2006–2007 winter season. The largest  $p_{\text{epitope}}$  value was 0.095 (epitope B; substitutions Q156H and L157S), suggesting that two mutations in dominant epitope B would lead to a rate of worst–case VE against these strains of 50.1% ( $E = 23.5\%$  of 47%,  $p_{\text{epitope}} = 0$ ) of that of a perfect match between vaccine and virus. We calculated the antigenic distance from the 2008–2009 vaccine strain A/Brisbane/10/2007 (H3N2) and the A/H3N2 Corsican strains isolated in 2008–2009. Using the  $p_{\text{epitope}}$  method, we found the largest  $p_{\text{epitope}}$  value was 0.045 (dominant epitope = E; substitutions K264P and S265H), suggesting a rate of worst–case VE against these strains of 76% ( $E = 35.9\%$  of 47%,  $p_{\text{epitope}} = 0$ ) of that of a perfect match ( $p_{\text{epitope}} = 0$ ) vaccine.

We calculated the antigenic distance from the 2007–2008 vaccine strain A/Salomon Island/3/2006 (H1N1) and A/H1N1

Corsican strains characterized by the mutation S125N (epitope B). Using the  $p_{\text{epitope}}$  method, we found the largest  $p_{\text{epitope}}$  value was 0.045 (dominant epitope = B), suggesting a rate worst–case VE of 90% ( $E = 47.6\%$  of 53%,  $p_{\text{epitope}} = 0$ ) compared of that of a perfect match ( $p_{\text{epitope}} = 0$ ) vaccine.

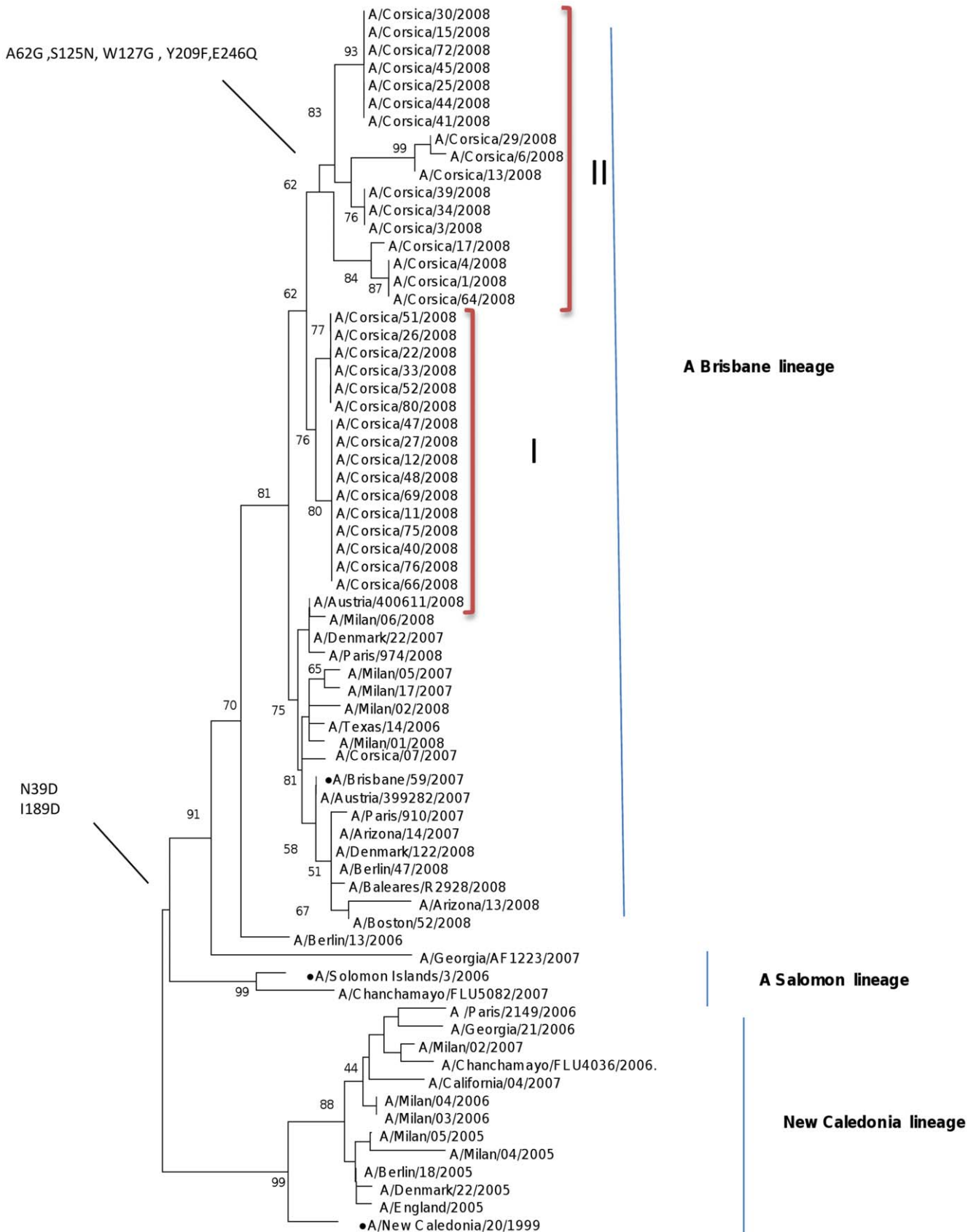
We calculated the antigenic distance from the vaccine strain of the Northern Hemisphere in 2009–2010, A/California/7/2009 and the A/H1N1pdm strains having the substitution P83S, S183T and D222E. The  $p_{\text{epitope}}$  value was 0.042 (dominant epitope D), suggesting a worst–case VE against these strains of 91% ( $E = 48\%$  of 53%,  $p_{\text{epitope}} = 0$ ) of that of a perfect match between vaccine and virus

### Discussion

Genetic and phylogenetic analysis of influenza A viruses circulating on Corsica Island during the four-year period from 2006–2010 confirmed that the genetic make-up of influenza A viruses may change from year to year.

Our results showed that the strains of influenza A/(H3N2) viruses isolated on Corsica Island belonged to two or more phylogenetic clades co-circulating from the 2006–2007 [25] to 2008–2009 seasons. During the 2006–2007 and 2007–2008 seasons, one clade was dominant. However, in 2006–2007, a sub-clade of A/Nepal/921/06-like circulating strains accumulated variations [25], altering their antigenicity as confirmed by HI assays on representative strains that were conducted by the WHO Influenza Reference Centre (Mill Hill, London). In 2008–2009, isolated viruses had drifted from A/Wisconsin/67/2005-like strains towards the vaccine strain A/Brisbane/10/2007 (vaccine strain for the 2008–2009 season in the Northern Hemisphere). These strains showed a higher value of genetic diversity with respect to the influenza A/H3N2 viruses isolated on Corsica during the 2006–2007 and 2007–2008 seasons.

Variations in the A/H3N2 viruses isolated on Corsica Island were predominantly detected at four antigenic sites A, B, D and E (Table 1). During the 2006–2007 season, a sub-group of isolates was characterized by four mutations previously described in the HA1 sequence of the A/Nepal/921/2006 non-vaccine reference strain and located at antigenic sites A, B and D. Wilson and Cox [26] have proposed that epidemiologically important drift variants usually display four or more amino acid substitutions located at two or more antigenic sites on the HA1 protein. Shih et al. [27] proposed that new antigenic variants are created when more than two mutations occur in antigenic sites or when one variation occurs in one antigenic site and one in a sialic acid RBD. The HA sequences of A/H3N2 viruses isolated on Corsica Island during the 2008–2009 season had drifted from the A/Wisconsin/67/2005-like lineage. These strains, which were characterized by two fixed mutations located at site E (K264E and K265E) compared to the vaccine strain A/Brisbane/10/2007, were closely related to the A/Brisbane/10/2007-like lineage (vaccine strain for the 2008–2009 season in the Northern Hemisphere). Recently, Huang et al. [28] showed that a pair of viruses are often “similar viruses” if the epitope A and B, which are closed to the RBD, are not changed.





**Figure 3. Phylogenetic tree of the HA domain of the A/H1N1 strains from Corsica Island.** Maximum likelihood phylogenetic tree analysis of HA genes of A/H1N1 strains circulating in Corsica Island during 2007–2008 season (\* represents vaccine strains). doi:10.1371/journal.pone.0024471.g003

Few mutations in the A/H1N1 seasonal isolates during the 2007–2008 season that differentiate them from the A/SalomonIslands/3/2006 (S125N at antigenic site B) and A/Brisbane/59/2007-like (S189D fixed at antigenic site B) lineages were located at the antigenic sites (Table 2).

**Table 2. Amino acid substitution in A/H1N1 influenza viruses isolated in Corsica Island from 2007 to 2008.**

Amino acid positions	39	62	125	127	139	140	189	209	240	246
Antigenic sites	B			A			B	D	D	
A/Brisbane/59/2007	N	A	S	W	C	S	I	Y	G	E
A/SalomonIslands/3/2006	D	A	S	W	C	S	D	Y	G	E
A/Corsica/30/2008	D	G	.	.	.	.	D	.	.	Q
A/Corsica/51/2008	D	.	.	.	.	.	D	.	.	.
A/Corsica/75/2008	D	.	N	.	.	.	D	.	.	.
A/Corsica/4/2008	D	G	.	G	.	.	D	F	.	.
A/Corsica/17/2008	D	G	.	G	.	.	D	.	.	.
A/Corsica/29/2008	D	G	.	G	S	C	D	.	.	.
A/Corsica/34/2008	D	G	N	.	.	.	D	.	.	.
A/Corsica/69/2008	D	.	N	.	.	.	D	.	.	.
A/Corsica/15/2008	D	G	.	.	.	.	D	.	.	Q
A/Corsica/13/2008	D	.	.	.	.	.	D	.	.	.
A/Corsica/12/2008	D	.	N	.	.	.	D	.	.	.
A/Corsica/6/2008	D	G	.	G	S	C	D	.	.	.
A/Corsica/3/2008	D	G	N	.	.	.	D	.	.	.
A/Corsica/47/2008	D	.	N	.	.	.	D	.	.	.
A/Corsica/41/2008	D	G	.	.	.	.	D	.	.	Q
A/Corsica/52/2008	D	.	.	.	.	.	D	.	.	.
A/Corsica/76/2008	D	.	N	.	.	.	D	.	.	.
A/Corsica/1/2008	D	G	.	G	.	.	D	F	.	.
A/Corsica/25/2008	D	G	.	.	.	.	D	.	.	Q
A/Corsica/27/2008	D	.	N	.	.	.	D	.	.	.
A/Corsica/13/2008	D	G	.	G	S	C	D	.	.	.
A/Corsica/26/2008	D	.	.	.	.	.	D	.	.	.
A/Corsica/39/2008	D	G	N	.	.	.	D	.	.	.
A/Corsica/40/2008	D	.	N	.	.	.	D	.	.	.
A/Corsica/44/2008	D	G	.	.	.	.	D	.	.	Q
A/Corsica/22/2008	D	.	.	.	.	.	D	.	.	.
A/Corsica/11/2008	D	.	N	.	.	.	D	.	.	.
A/Corsica/64/2008	D	G	.	G	.	.	D	F	.	.
A/Corsica/66/2008	D	.	N	.	.	.	D	.	.	.
A/Corsica/72/2008	D	G	.	.	.	.	D	.	.	Q
A/Corsica/80/2008	D	.	.	.	.	.	D	.	.	.
A/Corsica/45/2008	D	G	.	.	.	.	D	.	.	Q
A/Corsica/33/2008	D	.	.	.	.	.	D	.	.	.
A/Corsica/48/2008	D	.	N	.	.	.	D	.	.	.
A/Corsica/7/2007	.	.	.	.	.	.	D	.	R	.

doi:10.1371/journal.pone.0024471.t002

These findings seem to indicate that non-synonymous changes in the HA1 gene, especially in antigenic regions, have been more frequent in A/H3N2 than in seasonal A/H1N1 in influenza virus isolated in Corsica Island.

Effectively, since the reintroduction of seasonal A/H1N1 into human populations in 1977, this subtype has exhibited a lower case fatality rate than A/H3N2, particularly when these subtypes co-circulate [29]. In addition, A/H1N1 undergoes less severe seasonal genetic bottlenecks than A/H3N2, resulting in less pronounced or frequent reductions in genetic diversity [30].

This hypothesis was supported by subtype-specific differences in adaptation rate with higher rates of adaptation for A/H3N2 than A/H1N1 in HA and NA genes [31].

The A/H1N1pdm isolates were characterized by one fixed mutation at site E and two fixed mutations at site D with respect to A/California/07/2009 (Table 3). All Corsican A/H1N1pdm strains analyzed were characterized by the S203T mutation specific to clade 7 isolates [32]. The clade 7 marker variation S203T, observed in all sampled Corsican viruses, is a A/H1N1pdm virus site under positive selection and also involved in antigenicity [33].

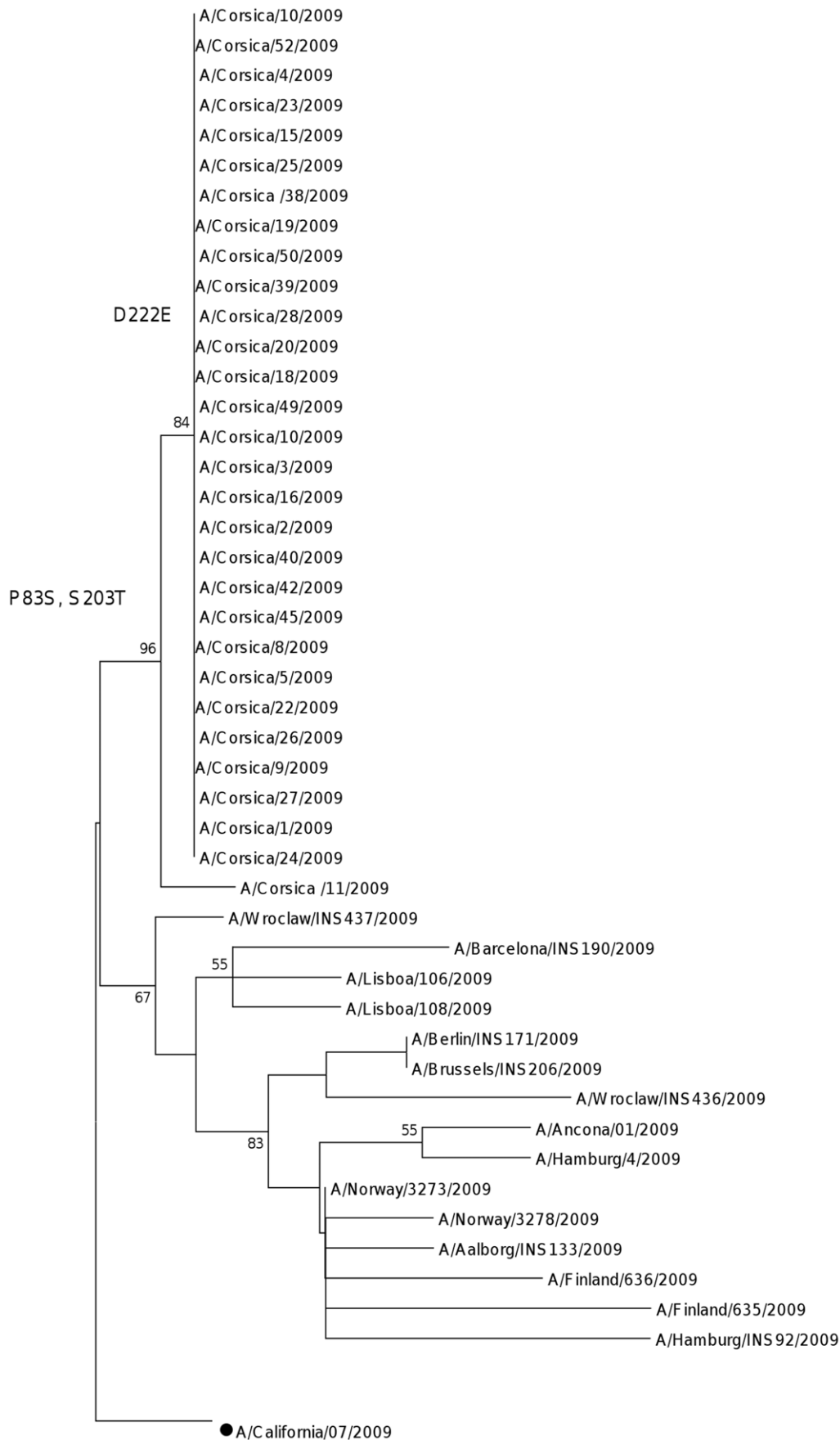
The viruses constituting this clade were therefore responsible for most of the pandemic burden worldwide. Following its origin, which remains obscure, the clade 7 viruses have been subjected to strong purifying selection, with the exception of the earliest phases of its evolution, behaving later as a well-fit virus, similar to viruses circulating in swine or seasonal influenza in humans.

Overrepresentation of certain mutations among geographically and temporally related samples needs to be carefully controlled for possible founder effects which could be identified as homogenous clusters in phylogenetic analyses, as was observed to be the case for the D222E mutation. Phylogenetic analysis of A/H1N1pdm influenza viruses showed that viruses isolated in Corsica Island formed a separate sub-clade of clade 7 as a consequence of the presence of the D222E substitution. While founder effect mutations cannot automatically be linked to phenotypes simply by increased occurrence, they may nevertheless alter the virus fitness for which even tiny changes could result in advantages shifting selection to their favor [34]. Recent studies showed that this position was positively selected for human strains [35–36]. Moreover, this codon, located in antigenic site D, is also associated with RBD indicating a positive selection from the hosts, caused maybe by vaccination and mass use of antiviral drugs.

This mutation, whose biological meaning is still unknown, has been isolated in other countries but at lower frequencies (e.g.; Italy, Turkey, Sweden and Finland). Recently, a cluster of D222E viruses among school children was isolated in Italy, confirming human-to-human transmission of viruses mutated at amino acid position 222 [37].

The D222G mutation was not observed among A/H1N1pdm isolated on Corsica Island. This substitution was observed more frequently in viruses isolated from patients with fatal outcomes [38].

Several studies showed that the A/H1N1pdm virus has a high genome-wide evolutionary rate (3.6661023 substitutions/site/year) [39] and a rate for the HA segment (0.961023 substitutions/site/year) that represents the lowest of all of the viral HAs of seasonal A/H1N1 (1918–1957; 2.961023), seasonal A/H1N1 (1977–2009; 1.761023), and swine H1 (1930–2009; 1.961023) [33].



**Figure 4. Phylogenetic tree of the HA domain of the A/H1N1 pdm strains from Corsica Island.** Maximum likelihood phylogenetic tree analysis of HA genes of A/H1N1 strains circulating in Corsica Island during 2009–2010 season (• represents vaccine strains). doi:10.1371/journal.pone.0024471.g004

All Corsican A/H1N1pdm isolates possessed D204 in the RBD, which confers binding of H1 viruses to human receptors, supporting efficient transmission of these viruses in humans. Key residues in the RBD [40] predicted to have a role in binding to human receptors (T98Y, S136T, 153W and 183H) were found to be Y98, S136, W153 and H183 in the Corsican isolates as in other A/H1N1pdm viruses isolated elsewhere.

Among the A/H3N2 Corsican isolates, the residues mainly for NeuAc $\alpha$ 2,6Gal linkage specific for the H3 subtype were Ile(I)226 and Ser (S)228, similar to other studies [41].

N-linked glycosylation is conserved among various HA subtypes of influenza A viruses. Its presence or absence can cause an increase or loss in function of the glycoprotein because N-linked

glycosylation can initiate and maintain folding, stability, solubility, antigenicity and immunogenicity of the protein. Most of the currently circulating viruses have six or seven N- glycosylation sites in the HA globular head region. Interestingly, the predicted N-linked glycosylation site at position 144 of the HA antigenic site A has been observed in the majority of strains isolated on Corsica in 2006–2007 and was definitively lost during the 2007–2008 and 2008–2009 seasons. Other studies have reported the loss of this predicted N-linked glycosylation site in all strains isolated since 2006–2007 [42].

In this study, we tried to predict a rate of VE for A/H3N2 and A/H1N1 seasonal influenza viruses and of the A/H1N1pdm virus using the  $p_{\text{epitope}}$  model. This model has been developed to provide researchers and health authorities with a new tool to quantify antigenic distance and to help with vaccine design. These results showed that the rate of VE compared to that of a perfect match against influenza viruses circulating on Corsica Island varied substantially across the four seasons analyzed, and that tend to be highest for A/H1N1 compared with A/H3N2 vaccines, suggesting that cross-immunity seems to be stronger for H1 HA gene [23].

Even if, as precedently described [21], the vaccine effectiveness estimated by several epidemiologic studies seems to be supported by the  $p_{\text{epitope}}$  model, the rate of VE estimated in this study need to be boosted with real VE data from the same patients. Nevertheless, studying the relationship between epitopes and vaccine efficiency [21] could be useful for studying influenza virus evolution and to consolidate studies based on epidemiological data.

One important limit of this study is that we have not analyzed the genetic evolution of the influenza virus outside HA proteins. Suzuki (2008) [43] applied a dN/dS approach to 100 complete H3N2 genomes and concluded that negative selection dominated in all proteins; significant positive selection was observed only in a handful of codons in the HA, NA and NP genes. Pond et al., (2008) [44] analyzed the same data using a technique more sensitive to the detection of individual selective sweeps and reported more evidence for adaptation in 5 genes: PB2 (2 codons), PB1 (5 codons), PA (3 codons), HA (81 codons), and NA (4 codons).

A recent study [31] showed that the rate of adaptation (per codon per year) is higher in surface residues of the viral NA than in HA1, indicating strong antibody-mediated selection on the former. They also observed high rates of adaptive evolution in several non-structural proteins, which may relate to viral evasion of T-cells and innate immune responses [31].

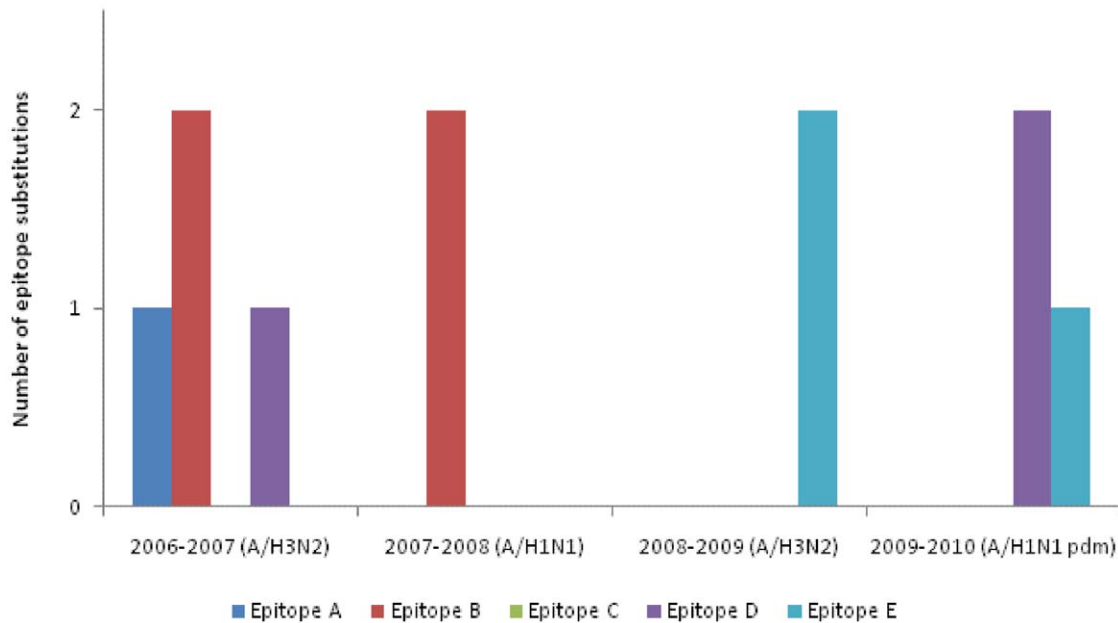
Another limit of the study is based on the sensitivity and specificity of molecular techniques used to identify strains and on the quality of specimens analyzed. Even if, the M RT-PCR used in this study to identify influenza A virus is specific and sensitive, there is a limit of detection between 10 and 100 copies/reaction [15].

The molecular analysis of the HA gene of influenza viruses that circulated on Corsica Island between 2006–2010 showed for each season the presence of a dominant lineage characterized by at least one fixed mutation. It must be noted that A/H3N2 and A/H1N1pdm were characterized by multiples fixation at antigenic sites. Subsequent mutations at antigenic sites often significantly contribute additional effects, so that multiple fixations, possibly including hitchhikers, occur rapidly and (almost) simultaneously. The fixation of specific mutations at each outbreak could be

**Table 3. Amino acid substitution in A/H1N1 pdm influenza viruses isolated in Corsica Island during 2009–2010.**

Amino acid position	32	83	203	222
<b>Antigenic sites</b>		E	D	<b>D</b>
<b>A/California/07/2009</b>	<b>L</b>	<b>P</b>	<b>S</b>	<b>D</b>
A/Corsica/01/2009	.	S	T	E
A/Corsica/02/2009	.	S	T	E
A/Corsica/03/2009	.	S	T	E
A/Corsica/04/2009	.	S	T	E
A/Corsica/05/2009	.	S	T	E
A/Corsica/08/2009	.	S	T	E
A/Corsica/09/2009	.	S	T	E
A/Corsica/10/2009	.	S	T	E
A/Corsica/11/2009	I	S	T	
A/Corsica/15/2009	.	S	T	E
A/Corsica/16/2009	.	S	T	E
A/Corsica/18/2009	.	S	T	E
A/Corsica/19/2009	.	S	T	E
A/Corsica/20/2009	.	S	T	E
A/Corsica/22/2009	.	S	T	E
A/Corsica/23/2009	.	S	T	E
A/Corsica/24/2009	.	S	T	E
A/Corsica/25/2009	.	S	T	E
A/Corsica/27/2009	.	S	T	E
A/Corsica/28/2009	.	S	T	E
A/Corsica/36/2009	.	S	T	E
A/Corsica/38/2009	.	S	T	E
A/Corsica/39/2009	.	S	T	E
A/Corsica/40/2009	.	S	T	E
A/Corsica/42/2009	.	S	T	E
A/Corsica/45/2009	.	S	T	E
A/Corsica/48/2009	.	S	T	E
A/Corsica/49/2009	.	S	T	E
A/Corsica/50/2009	.	S	T	E
A/Corsica/52/2009	.	S	T	E

doi:10.1371/journal.pone.0024471.t003



**Figure 5. Epitope substitutions over time.** Accumulation of epitope substitutions over time in influenza viruses isolated in Corsica Island since 2006.

doi:10.1371/journal.pone.0024471.g005

explained by the combination of compensatory mutations, neutral phenomenon and a founder effect, favoring the presence of a dominant lineage in a closed environment such as Corsica Island.

These findings confirmed that careful surveillance of genetic changes in the HA1 domain during the influenza epidemic season may provide early information on virus variants and improve the influenza vaccine.

## Acknowledgments

This work was made possible thanks to the participating general practitioners of the *Sentinelles* Network: Agostini, Albertini, Andrei, Appietto, Arfi, Atlan,

Ballejos, Benadda, Calendini, Campana, Casamatta, Castellucci, Ciabrini, Corneille, Cucchi, Dahan, Delaunay, Denis, Dolovici, Erminj-Arifi, Geromini, Ghionga, Giancarli, Grossi, Guidicelli, Leccia, Leoni, Mattei, Miniconi, Moretti, Moretti-Ciabrini, Nozze, Ottavi, Paoletti, Patriarche, Petroni Renard, Rossi, Simeoni, Soreau, Spadoni and Vincensini.

## Author Contributions

Conceived and designed the experiments: A.Falchi JPA LV LA TH A.Flahault TB. Performed the experiments: A.Falchi. Analyzed the data: A.Falchi CT CA. Contributed reagents/materials/analysis tools: LV TB AL FC A.Falchi TH TB JA. Wrote the paper: A.Falchi TH CA.

## References

- Webster RG, Bean WJ, Gorman OT, Chambers TM, Kawaoka Y (1992) Evolution and ecology of influenza A viruses. *Microbiol Rev* 56: 152–79.
- Lin YP, Gregory V, Bennett M, Hay A (2004) Recent changes among human influenza viruses. *Virus Res* 103: 47–52.
- Bush RM, Bender CA, Subbarao K, Cox NJ, Fitch WM (1999) Predicting the evolution of human influenza A. *Science* 286: 1921–5.
- Hampson AW, Mackenzie JS (2006) The influenza viruses. *Med J Aust* 185: S39–43.
- De Jong JC, Rimmelzwaan GF, Fouchier RA, Osterhaus AD (2000) Influenza virus: a master of metamorphosis. *J Infect* 40: 218–228.
- Ferguson NM, Galvani AP, Bush RM (2003) Ecological and immunological determinants of influenza evolution. *Nature* 422: 428–433.
- Kilbourne ED, Smith C, Brett I, Pokorny BA, Johansson B, et al. (2002) The total influenza vaccine failure of 1947 revisited: major intrasubtypic antigenic change can explain failure of vaccine in a post-World War II epidemic. *Proc Natl Acad Sci USA* 99: 10748–10752.
- World Health Organization (2009) Pandemic (H1N1). Influenza-like illness in the United States and Mexico. 24 April 2009. Available: [http://www.who.int/csr/don/2009\\_04\\_24/en/index.html](http://www.who.int/csr/don/2009_04_24/en/index.html).
- Centers for Disease Control and Prevention (2009) Update: infections with a swine-origin influenza A (H1N1) virus—United States and other countries, April 28th, 2009 MMWR Mortal. *Wkly Rep* 58: 431–433.
- Rousset D, Bouscambert-Duchamp M, Enouf V, Valette M, I G, et al. (2010) Influenza A/(H1N1)2009 epidemic in France: virological parameters. *Bulletin épidémiologique hebdomadaire* 24-25-26.
- Valleron AJ, Bouvet E, Garnerin P, Menares J, Heard I, et al. (1986) A computer network for the surveillance of communicable diseases: the French experiment. *Am J Public Health* 76: 1289–1292.
- Wright KE, Wilson GA, Novosad D, Dimock C, Tan D, et al. (1995) Typing and subtyping of influenza viruses in clinical samples by PCR. *J Clin Microbiol* 33: 1180–84.
- Schweiger B, Zadow I, Heckler R, Timm H, Pauli G (2000) Application of a fluorogenic PCR assay for typing and subtyping of influenza viruses in respiratory samples. *J Clin Microbiol* 38: 1552–58.
- Chi XS, Li F, Tam JS, Rappaport R, Cheng SM (2007) Semiquantitative one-step RT-PCR for simultaneous identification of human influenza and respiratory syncytial viruses. *J Virol Methods* 139: 90–92.
- Duchamp MB, Casalegno JS, Gillet Y, Frobert E, Bernard E, et al. (2010) Pandemic A(H1N1)2009 influenza virus detection by real time RT-PCR: is viral quantification useful? *Clin Microbiol Infect* 16: 317–21.
- Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The Clustal X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research* 24: 4876–82.
- Posada D, Crandall KA (2001) Selecting the best-fit model of nucleotide substitution. *Syst Biol* 50: 580–601.
- Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* 52: 696–704.
- Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* 24: 1596–1599.
- Gupta R, Jung E, Brunak S (2004) Prediction of N-glycosylation sites in human proteins. Available: <http://www.cbs.dtu.dk/services/NetNGlyc/>.
- Gupta V, Earl DJ, Deem MW (2006) Quantifying influenza vaccine efficacy and antigenic distance. *Vaccine* 24: 3881–3888.
- Deem MW, Lee H-Y (2003) Sequence space localization in the immune system response to vaccination and disease. *Phys Rev Lett* 91: 068101.
- Pan K, Subieta KC, Deem MW (2011) A novel sequence-based antigenic distance measure for H1N1, with application to vaccine effectiveness and the selection of vaccine strains. *Protein Eng Des Sel* 24: 291–299.
- Deem MW, Pan K (2009) The epitope regions of H1-subtype influenza A, with application to vaccine efficacy. *Protein Eng Design Select* 9: 543–546.

25. Falchi A, Varesi L, Arena C, Leveque N, Renois F, et al. (2009) Co-circulation of two genetically distinct sub-groups of A/H3N2 influenza strains during the 2006–2007 epidemic season in Corsica Island, France. *J Clin Virol* 45: 265–268.
26. Wilson IA, Cox NJ (1990) Structural basis of immune recognition of influenza virus hemagglutinin. *Annu Rev Immunol* 8: 737–771.
27. Shih AC, Hsiao TC, Ho MS, Li WH (2007) Simultaneous amino acid substitutions at antigenic sites drive influenza A hemagglutinin evolution. *Proc Natl Acad Sci USA* 104: 6283.
28. Huang, Yang (2011) Changed epitopes drive the antigenic drift for influenza A (H3N2) viruses. *BMC Bioinformatics* 12: S31.
29. Wright PF, Thompson J, Karzon DT (1980) Differing virulence of H1N1 and H3N2 influenza strains. *Am J Epidemiol* 112: 814–819.
30. Rambaut A, Pybus OG, Nelson MI, Viboud C, Taubenberger JK, et al. (2008) The genomic and epidemiological dynamics of human influenza A virus. *Nature* 453: 615–619.
31. Bhatt S, Holmes EC, Pybus OG (2011) The genomic rate of molecular adaptation of the human influenza A virus. *Mol Biol Evol*;2011 Mar 16.
32. Nelson M, Spiro D, Wentworth D, Beck E, Fan J, et al. (2009) The early diversification of influenza A/H1N1pdm. *PLoS Curr* 1: RRN1126.
33. Furuse Y, Shimabukuro K, Odagiri T, Sawayama R, Okada T, et al. (2010) Comparison of selection pressures on the HA gene of pandemic (2009) and seasonal human and swine influenza A H1 subtype viruses. *Virology* 405: 314–321.
34. Lee RT, Santos CL, de Paiva TM, Cui L, Sirota FL, et al. (2010) All that glitters is not gold—founder effects complicate associations of flu mutations to disease severity. *Virol J* 7: 297.
35. Li W, Shi W, Qiao H, Ho SY, Luo A, et al. (2011) Positive selection on hemagglutinin and neuraminidase genes of H1N1 influenza viruses. *Virol J* 8: 183.
36. Mullick J, Cherian SS, Potdar VA, Chadha MS, Mishra AC (2011) Evolutionary dynamics of the influenza A pandemic (H1N1) 2009 virus with emphasis on Indian isolates: Evidence for adaptive evolution in the HA gene. *Infect Genet Evol* 11: 997–1005.
37. Puzelli S, Facchini M, De Marco MA, Palmieri A, Spagnolo DI, et al. (2010) Influnet Surveillance Group for Pandemic A(H1N1) 2009 Influenza Virus in Italy. Molecular surveillance of pandemic influenza A(H1N1) viruses circulating in Italy from May 2009 to February 2010: association between haemagglutinin mutations and clinical outcome. *Euro Surveill* 28: 19696.
38. Kilander A, Rykkvin R, Dudman SG, Hungnes O (2010) Observed association between the HA1 mutation D222G in the 2009 pandemic influenza A(H1N1) virus and severe clinical outcome, Norway 2009–2010. *Euro Surveill* 15: 19498.
39. Smith GJ, Vijaykrishna D, Bahl J, Lycett SJ, Worobey M, et al. (2009) Origins and evolutionary genomics of the 2009 swine-origin H1N1 influenza A epidemic. *Nature* 459: 1122–5.
40. Stevens J, Blixt O, Tumpey TM, Taubenberger JK, Paulson JC, et al. (2006) Structure and receptor specificity of the hemagglutinin from an H5N1 influenza virus. *Science* 312: 404–10.
41. Lindstrom S, Sugita S, Endo A, Ishida M, Huang P, et al. (1996) Evolutionary characterization of recent human H3N2 influenza A isolates from Japan and China: novel changes in the receptor binding domain. *Arch Virol* 141: 1349–1355.
42. Suwannakarn K, Chieochansin T, Thongmee C, Makkoch J, Praianantathavorn K, et al. (2010) Molecular evolution of human H1N1 and H3N2 influenza A virus in Thailand, 2006–2009. *PLoS One* 16: e9717.
43. Suzuki Y (2008) Positive selection operates continuously on hemagglutinin during evolution of H3N2 human influenza A virus. *Gene* 427: 111–116.
44. Pond SLK, 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: 1809–24.