Article

Molecular Epidemiology and Spatio-Temporal Dynamics of the H3N8 Equine Influenza Virus in South America

Cecilia Olguin Perglione [1], Marcelo D. Golemba [2], Carolina Torres [3,4] and Maria Barrandeguy [1,5,*]

[1] Instituto de Virología CICVyA, Instituto Nacional de Tecnología Agropecuaria (INTA), Dr. Nicolás Repetto y De Los Reseros s/n Hurlingham B1686LQF, Buenos Aires, Argentina; olguin.cecilia@inta.gob.ar

[2] Hospital de Pediatría S.A.M.I.C. "Prof. Dr. Juan P. Garrahan", Combate de los Pozos 1881, Ciudad Autónoma de Buenos Aires C1245AAM, Argentina; mgolemba@garrahan.gov.ar

[3] Cátedra de Virología, Facultad de Farmacia y Bioquímica, Universidad de Buenos Aires, Junín 956, Ciudad Autónoma de Buenos Aires C1113AAD, Argentina; ctorres@ffyb.uba.ar

[4] CONICET, Av. Rivadavia 1917, Ciudad Autónoma de Buenos Aires C1033AAJ, Argentina

[5] Escuela de Veterinaria, Universidad del Salvador, Viamonte 1856, Ciudad Autónoma de Buenos Aires C1056ABB, Argentina

* Correspondence: barrandeguy.maria@inta.gob.ar; Tel.: +54-11-4621-9050

Abstract: Equine influenza virus (EIV) is considered the most important respiratory pathogen of horses as outbreaks of the disease lead to substantial economic losses. The H3N8 EIV has caused respiratory disease in horses across the world, including South American countries. Nucleotide and deduced amino acid sequences for the complete haemagglutinin gene of the H3N8 EIV detected in South America since 1963 were analyzed. Phylogenetic and Bayesian coalescent analyses were carried out to study the origin, the time of the most recent common ancestors (tMRCA), the demographic and the phylogeographic patterns of the H3N8 EIV. The phylogenetic analysis demonstrated that the H3N8 EIV detected in South America grouped in 5 well-supported monophyletic clades, each associated with strains of different origins. The tMRCA estimated for each group suggested that the virus was circulating in North America at least one year before its effective circulation in the South American population. Phylogenetic and coalescent analyses revealed a polyphyletic behavior of the viruses causing the outbreaks in South America between 1963 and 2012, possibly due to the introduction of at least 4 different EIVs through the international movement of horses. In addition, phylodynamic analysis suggested South America as the starting point of the spread of the H3N8 EIV in 1963 and showed migration links from the United States to South America in the subsequent EIV irruptions. Further, an increase in the relative genetic diversity was observed between 2006 and 2007 and a subsequent decline since 2009, probably due to the co-circulation of different lineages and as a result of the incorporation of the Florida clade 2 strain in vaccines, respectively. The observed data highlight the importance of epidemiological surveillance and the implementation of appropriate quarantine procedures to prevent outbreaks of the disease.

Keywords: equine influenza; South America; haemagglutinin; phylodynamics; phylogeography

1. Introduction

Equine influenza (EI) is one of the major infectious respiratory diseases of horses and is considered severely damaging in economic terms. This is due to the highly contagious nature of the virus which facilitates rapid spread among susceptible horses, with the consequent interruption of equestrian activities [1,2]. EI is usually a self-limiting disease characterized by pyrexia, coughing and

nasal discharge [3]. Two subtypes of influenza A virus are known to infect horses, the H7N7 and the H3N8. The H7N7 was identified in 1956 during an outbreak of respiratory disease in horses in Czechoslovakia [4], and the H3N8 subtype was first detected in 1963 during a major epizootic of respiratory disease in horses in Florida, United States (USA) [5,6]. Both subtypes of viruses co-circulated for around 25 years; however, the H7N7 has not been isolated since 1979 and is considered to be extinct [7], and since that time, all outbreaks of EI worldwide have been caused by the H3N8 subtype.

Phylogenetic analysis of the equine influenza virus (EIV) haemagglutinin (HA) gene revealed that the H3N8 EIV evolved as a single lineage for at least two decades [8] and diverged during the mid-1980s, into the American and the European lineages, named according to the geographic origin of the isolates [9,10]. Subsequently, strains within the American lineage diverged into South American, Kentucky and Florida sub-lineages [11]. Nowadays, the Florida sub-lineage is predominant and has evolved into two antigenically different clades: Florida clade 1 and Florida clade 2 [9]. Florida clade 1 strains have been isolated in North America since 2003 [9] and have been associated with EI outbreaks in South Africa, Japan, Australia, Europe, Dubai and South America [1,12-17]. Florida clade 2 predominates in Europe, but has been also identified in some countries of Asia [18-20]. In addition to this evolutionary pattern, in a genome-scale analysis, Murcia et al. described that, according to the H3N8 EIV HA gene characteristics, 12 phylogenetically distinct clades (including Florida clade 1 and Florida clade 2) could be distinguished, each of them comprised by a group of viruses from a particular geographical region and sharing a common ancestor [21].

In South America, the first reported EI outbreak occurred in Chile in 1963 due to EIV the H3N8 infection [22,23]; a further occurrence, in 1977, was caused by the H7N7 subtype infection [24]. The H3N8 EIV re-emerged in Chile in several outbreaks that occurred in 1985 and 1992 [25,26]. Another episode of EI, which affected both vaccinated and unvaccinated horses, occurred in 2006, caused by a virus belonging to the H3N8 American lineage. The last reported outbreak of EI in Chile was in 2012 [27].

In Brazil, there has been evidence of EIV circulation since 1963 (A/eq/SaoPaulo/6/1963, GenBank accession number CY032293). Subsequent outbreaks of EIV occurred in 1969, 1976, 1986 and 1988 [28]. An extensive onset of acute respiratory disease in horses in 2012 was the last reported occurrence of EI in this country [29].

Regarding the EI situation in Uruguay, the extensive outbreak occurred in 2012 was the only one reported [16]. Nevertheless, previous evidence of the circulation of the virus is the HA gene nucleotide sequences of two strains available at the GenBank database, the H3N8, A/eq/Uruguay/1/1963 (GenBank accession number CY032421) and the H7N7, A/eq/Uruguay/1063/1976 (GenBank accession number CY036887).

With respect to the EI in Argentina, the first reported outbreak was in 1976 and the virus belonged to the H7N7 subtype [30]. Since then, all the reported EI outbreaks have been caused by the H3N8 subtype infection. The first detection occurred in the summer of 1985–1986, and is described as one of the major outbreaks of EI due to the rapid spread of the virus all across the country [31]. In 1993, a new episode of acute respiratory disease with high morbidity occurred among thoroughbred horses in Palermo's and San Isidro's racecourses [10,32]. Sporadic H3N8 EI virus incursions, with only a few horses involved, took place between 1994 and 2001 [11]. In 2005, EI arose among horses stabled in a jumping club in Buenos Aires city, which had temporarily received horses from Chile; the morbidity was low, not exceeding 10% of the horse population, and no cases outside this location were reported [33]. All the EIVs identified between 1993 and 2005 were grouped within the South American sub-lineage of the American lineage [11,33]. The last extensive outbreak of EI registered in Argentina was in 2012 [16].

The HA, one of the surface glycoproteins of the EIV, is the primary target of the protective immune response [34]. The HA1 domain contains the major neutralizing antibody binding sites identified from A to E [10,35]; if sufficient changes occur at these antigenic sites, the virus may evade antibody neutralization [9]. Vaccines provide protection through the stimulation of antibody response, especially against the HA glycoprotein [1,36], and vaccination plays a major role in

controlling infection and disease, particularly for horses that travel widely, intermingle with other horses or take part in sports events [37]. Hence, the HA gene segment is of major importance and is the focus for EIV surveillance [34]. The World Organization for Animal Health OIE Expert Surveillance Panel (ESP) reviews the genetic, antigenic and epidemiological information of EI field outbreaks occurring worldwide, in order to evaluate the vaccine effectiveness and elaborate recommendations to ensure that the vaccines contain epidemiologically relevant strains [36].

In spite of the fact that viral surveillance and molecular descriptions have been performed on EIV, no studies have analyzed the origin and the spatio-temporal phylodynamics of EIV in South America. Thus, the aim of this work was to describe the molecular characteristics of the H3N8 EIV detected in South America since its first identification, to infer the potential source of infections and to study its demographic and phylogeographic patterns.

2. Results

2.1. Phylogenetic Analysis

EIV HA gene phylogenetic analyses were carried out with two different data sets: (a) partial HA1 and (b) complete HA nucleotide sequence gene. The topology of the trees obtained with maximum likelihood (ML) was similar for both data sets (Figures S1 and S2). In addition, the time-scaled maximum clade credibility (MCC) tree inferred by Bayesian coalescent analysis also displayed a similar phylogenetic pattern (Figure 1) [38]. The resulting trees grouped all the viruses included in this study into five well-supported clusters: Pre-divergent, Eurasian, American and Florida clade 1 and Florida clade 2 sub-lineages (Figure 1, Figures S1 and S2). Following the criteria introduced by Murcia, the twelve described different clades could also be distinguished in the trees obtained with the complete HA date set [21]. However, three of the twelve groups (VI, VII and IX) were not observed in the ML tree obtained with the partial HA data set, probably due to a lower phylogenetic signal in the partial HA data set compared to the complete HA data set. Moreover, two additional clades, called South America clade 1 and South American clade 2, were found in all the trees. Correlating, isolates from Groups I to VIII belong to the Pre-divergent lineage, whereas Group IX and X belong to the American and the Eurasian lineage, respectively.

All the H3N8 EIVs detected in South America were grouped into five monophyletic clades, sustained by high support values (Figure 1, Figures S1 and S2). As previously described by Murcia et al. [21], Group I is made up of strains detected in Brazil in 1963 and 1969 ($n = 2$), and Group VIII of strains from Argentina ($n = 1$), Chile ($n = 1$) and interestingly, from USA ($n = 1$) in 1985; these two groups were included in the Pre-divergent lineage. The EIV detected from 1993 to 1996 was grouped in the South American clade 1, which is composed only of Argentinian strains ($n = 6$). The EIV identified between 1997 and 2006 belonged to the South American clade 2, grouping strains detected in Argentina in 1997, 2001, 2004 and 2005 ($n = 8$) and in Chile in 2006 ($n = 1$). The South American clades 1 and 2 were included in the American lineage constituting two separate groups, independent from Murcia's Group X. The EIV detected in Chile ($n = 1$), Brazil, ($n = 1$), Uruguay ($n = 1$) and Argentina ($n = 6$) in 2012 belonged to the Florida clade 1 lineage (Figure 1, Figures S1 and S2).
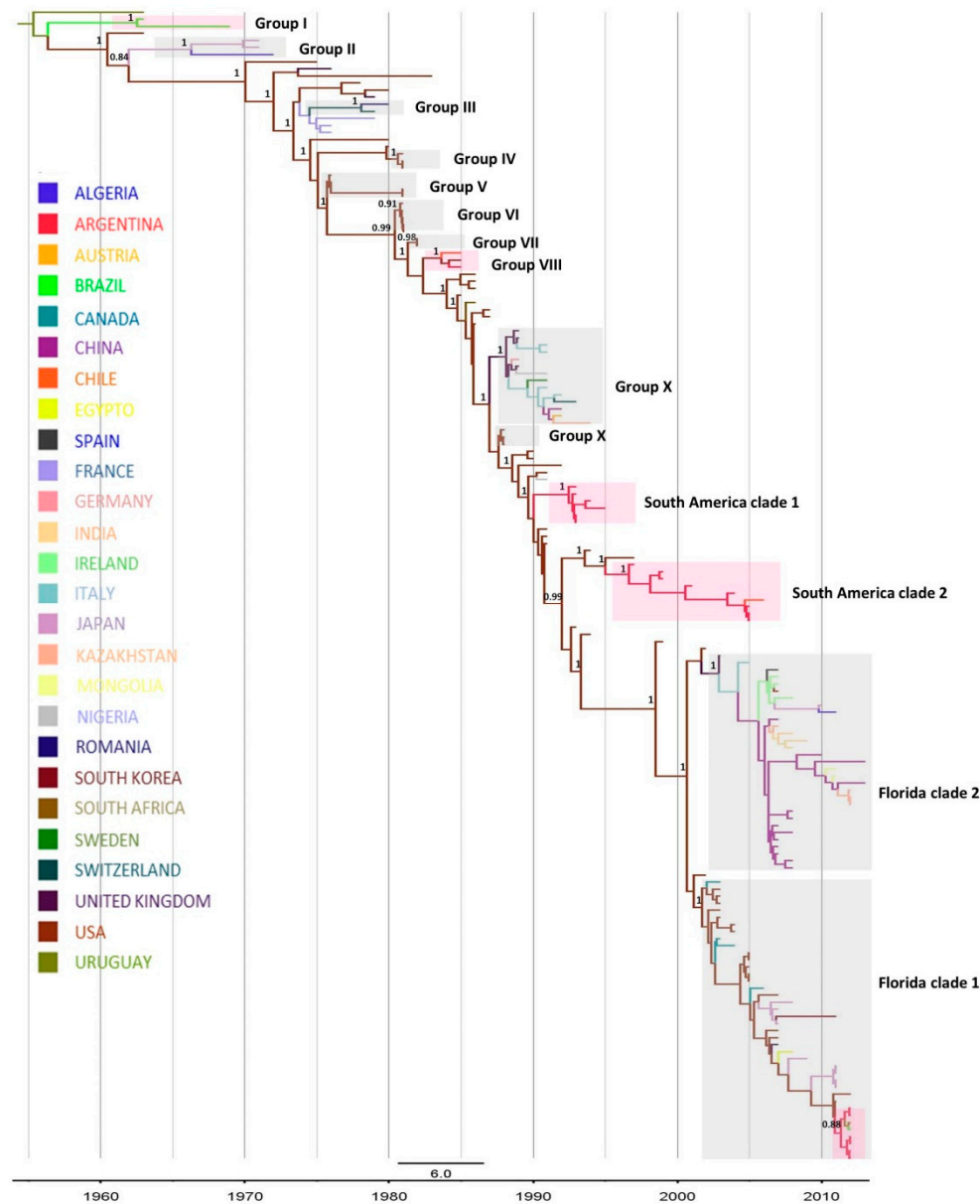
Figure 1. Maximum clade credibility tree for the complete HA gene of the H3N8 EIV. The branches are in time scale (years) and are colored on the basis of the most probable ancestor location. Posterior values for relevant groups are shown at nodes. Groups highlighted in Magenta correspond to South American clades.

2.2. Phylodynamic Analysis

A Bayesian coalescent analysis of the complete HA gene was carried out to study the demographic and phylogeographic pattern of EIV, and particularly, to estimate the ancestral times associated with South American viruses. The time of the most recent common ancestors (tMRCA) was estimated for the different groups: Group I, 1962; Group VIII, 1984; South American clade 1, 1992 and South American clade 2, 1997. As for the viruses detected in South America in 2012, which grouped within the Florida clade 1, the tMRCA was estimated in 2011 (Table 1) [38].

Table 1. Estimated time of the most recent common ancestor (tMRCA) for the five clades detected in South America by Bayesian Coalescent analysis.

Clade	Period	tMRCA (HPD 95%) *
Group I	1963–1969	1962 (1961–1963)
Group VIII	1985	1984 (1983–1984)
South American Clade 1	1993–1996	1992 (1992–1993)
South American Clade 2	1997–2005	1997 (1996–1997)
Florida Clade 1	2012	2011 (2011–2012)

* HPD95%: Highest posterior density interval (95%).

In addition, the geographic spread pattern of the H3N8 EIV in South America was studied. The most probable ancestral locations of different groups were estimated and are represented in different colors in the MCC tree (Figure 1). Particularly, despite the fact that the analysis estimated Brazil as the most probable ancestral location of the strains belonging to Group I (p = 0.98), the location of the ancestral virus originating this group could not be defined with certainty. For Group VIII, the most probable origin appears to be USA (p = 0.97). The South American clade 1 EIV, identified between 1993 and 1996, would derive from strains circulating in USA (p = 1). The most probable ancestral location of South American clade 2 and Florida clade 1 was also USA, with high probabilities (p = 0.98 and p = 0.97, respectively). In accordance with this observation, the South American clade 2 was closely related to A/eq/California/4537/97 and A/eq/Kentucky/1/94 strains, while Florida clade 1 strains from South America were closely related to A/eq/Florida/146609/11 strain; these were identified as probable ancestors of these groups.

To gain insight into the spatio-temporal dynamics of the viral diffusion process, the MCC tree can be visualized over time (Figure 2). The earliest divergence events were observed from Uruguay to USA and Brazil between 1963 and 1969. After that, in 1971, the virus spread from USA into Asia and subsequently from Asia into Africa in 1972. In 1973, the H3N8 EIV entered from USA into France, and during the next few years the virus has spread within the European continent. The first link from USA to South America occurred between 1983 and 1985, when the virus spread into Argentina and Chile. Later, the EIV spread from USA into South Africa in 1986, and circulated in Europe during the period 1989 and 1992. A new entrance from USA into Africa occurred in 1991 and from Europe into Asia in 1992. A second migration wave from USA to South America was inferred between 1991 and 1993, and the virus was identified in Argentina in 1993. A third migration link between USA and South America was inferred in 1996 and after its introduction, the virus spread across Argentina (A/eq/Argentine/E-433/1997) and Chile (A/equine/Lonquen/1/2006). Between 2002 and 2007, new introductions from USA to Europe and Asia occurred, and subsequently from Asia into Europe. The last migration link between USA and South America was observed in 2011, after which the virus circulated in Uruguay, Argentina and Brazil (Figure 2).

In addition, demographic reconstruction revealed that genetic diversity has remained constant between 1960 and 1970, followed by an abrupt drop in the early 1980s, and an increase and stabilization from about 1990 onwards. In 2006, viral diversity shows a remarkable increase with a maximum in 2007, and then a slight decrease from 2007 to 2009 followed by an abrupt decrease from then on (Figure 3) [38].
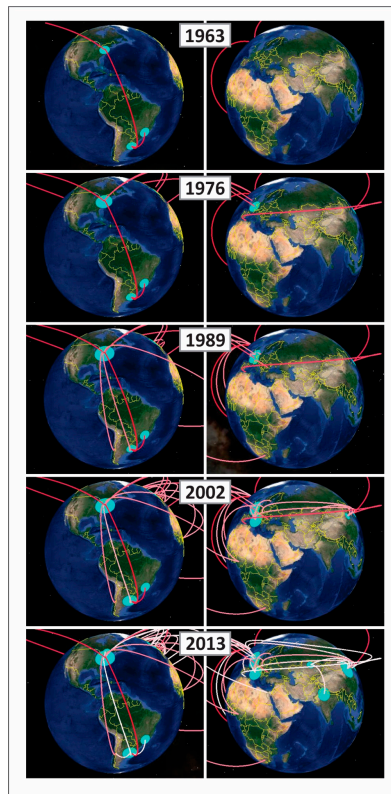
Figure 2. Spatio-temporal diffusion process for EIV in America (left side) and Europe, Asia and Africa (right side). Lines represent the maximum clade credibility tree projected on the surface. The red-white colour gradient represents the relative age of the dispersal pattern (older-recent). The size of light blue polygons represents lineage density for each location.
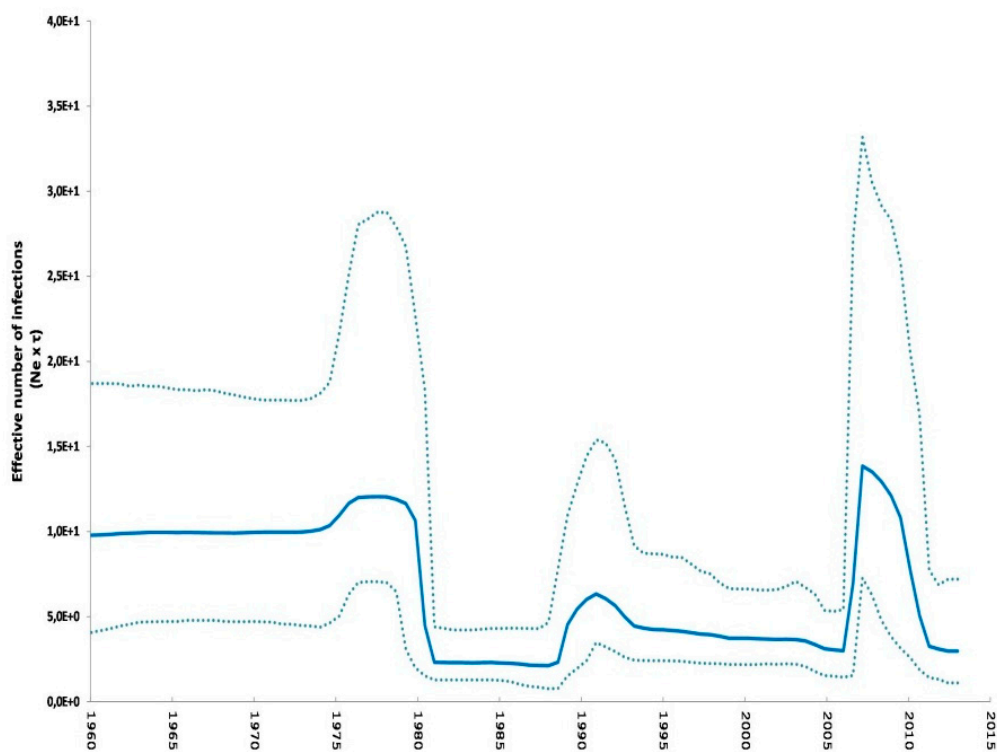
Figure 3. Phylodynamic reconstruction for complete HA data set showing the relative genetic diversity through time (Ne x t). The thick solid line represents the mean value and the dash lines the HPD 95% values.

2.3. Amino Acid Alignment

The HA1 derived amino acid (aa) sequences from the South American strains were aligned with the prototype and the inferred ancestor strains of each group (Figure 4 and Table 2).

Figure 4. Amino acid changes in the antigenic sites.

STRAINS / Position	SITE A *132-137	SITE B *155-163	SITE B *187-199	SITE C *48-55	SITE C *273-278	SITE D *201-218	SITE D *241-248	SITE E *62-63	SITE E *78-83
A/eq/Uruguay/1/1963		^TKSGNSYPT	^TNNEQTKLYVQAS	^TGKICNNP			^DVLMINSN		
A/eq/SaoPaulo/6/1963		^····S····	^·········	^··R····			^·I·····		
A/eq/SaoPaulo/1/1969		^····S····	^···D·······	^··R····			^·I·····		
A/equine/Fontainebleau/1/1979			^TNNEQTKLYVQEL						
A/equine/Cordoba/18/1985			^S········I··S						
A/equine/SantaFe/1/1985			^S···········S						
A/Equine/Santiago/1/1985			^S·········I··S						
A/eq/Newmarket/1/93			^SNQQQTELYIQES			^VTVSTKRSQQTIIPNIG		^RN	^VFQYEN
A/eq/Argentine/E-37-2/93			^··NE··K·····			^·············		^··	^······
A/eq/Argentine/E-37-18/93			^··NE··K·····			^·············		^··	^······
A/eq/Argentine/E-41-2/93			^··NE··K·····			^·············		^··	^······
A/eq/Argentine/E-131/94			^··NE··K·····			^·············		^··	^······
A/eq/Argentine/E-226/95			^··NE··K·····			^·············		^··	^······
A/eq/Kentucky/1/94			^···········			^·············		^··	^······
A/eq/California/4537/97			^···········			^·········V···		^··	^······
A/eq/Argentine/E-433/97			^···········			^·········V···		^··	^······
A/eq/Argentine/E-652/99			^···E··K·····			^·········V···		^··	^······
A/eq/Argentine/E-750/99			^···E··K·····			^·········V···		^··	^······
A/eq/Argentine/E-1134-6/01			^···K·······			^·········V···		^··	^······
A/eq/Argentine/E-4581-3/04			^···K·······			^·········V···		^··	^······
A/eq/Argentine/E-6189-3/05			^···K·······			^·········V···		^··	^·I····
A/eq/Argentine/E-6424-2/05			^···K·······			^·········V···		^··	^·I····
A/eq/Argentine/E-6424-1/05			^···K·······			^·········V···		^··	^·I····
A/eq/Lonquen/1/2006			^···K·······			^·········V···		^··	^·I····
A/eq/Ohio/1/03								^RN	
A/eq/Florida/146609/11								^K·	
A/eq/Argentine/E-2345-1/12								^K·	
A/eq/Argentine/E-2345-5/12								^K·	
A/eq/Argentine/E-2357/12								^K·	
A/eq/Argentine/E-2397-2/12								^K·	
A/eq/Argentine/E-2397-3-12								^K·	
A/eq/Argentine/E-3146-1/12								^K·	
A/eq/Rio Grande do Sul/1/12								^K·	

Table 2. Amino acid substitution characterizing the South American groups.

Group	Aa Substitution
Group I	K50R; T131N; S149N; S265G; L431I; A476T; L496V
Group I and A/eq/Uruguay/1/1963	G82; R323; I347; G381
South American clade 1	N312K
South American clade 2	S92N
Florida clade 1	K-14A; M70V

Strains belonging to Group I were compared with A/eq/Uruguay/1/1963, as the H3N8 EIV is suggested to have originated in South America. Eleven aa substitutions were found, three of them at antigenic sites: K50R (site C), which is exclusive to this group, N159S (site B) and V242I (site D). Three additional aa changes, all of them outside of antigenic sites and exclusive to this group, were observed: T131N, S149N and S265G. The remaining aa substitutions were: G6S, G7D, N291E, Q327R and I328L, which are also found in other strains. The strain detected in Sao Paulo in 1969 possesses an additional aa substitution at the antigenic site B (N189D), which is unique to this strain. The aa G and R at positions 82 and 323 respectively, are observed in Group I and Uruguay/1/1963 but not in other H3N8 EIV strains.

Compared with A/eq/Fontainebleau/79 (Pre-divergent prototype strain), the strains belonging to Group VIII identified in South America presented seven aa substitutions: T46I, A93T, R140K, V223I, I267V, T187S and L199S, the last two at the antigenic site B.

In comparison with A/eq/Newmarket/1/93 (American lineage prototype strain), viruses within the South American clade 1 carried four aa substitutions at the HA1, three of them at the antigenic site B: Q189N, Q190E and E193K. The remaining aa substitution, N312K, is present only in the strains of this lineage. The South American clade 2 strains showed three aa substitutions; one of these substitutions, I214V, is located at the antigenic site D and is also present in one of the possible

ancestors, A/eq/California/4537/97. The remaining aa changes are S92N, which is exclusive to this group, and D104N. In addition to the described aa substitutions, the strains detected in 1999 possess the two aa changes at antigenic site B, Q190E and E193K, which were also observed in strains from the South American clade 1. The EIV strains detected between 2001 and 2006 have evolved with a new aa change, Q190K, in the antigenic site B, which is also exclusive to this group. The strains detected in 2005 and 2006 possess an additional and distinctive aa change, V78I.

The aa substitutions presented in the Uruguayan, Brazilian and Argentinian strains detected in 2012, when compared with the Florida clade 1 prototype strain (A/eq/Ohio/01/03) are: G7D, R62K (in antigenic site E), D104N, A138S and V223I, which are also present in the possible origin of the group, A/eq/Florida/146609/11 strain. In addition, two different sub-populations can be recognized among them, one with a substitution at K-14A at the predicted peptide signal sequence; and the other one at K-14T also at the predicted peptide signal sequence and M70V, the latter being present only in Argentinian strains. With the available data for the strain detected in Chile in 2012 (deduced aa sequence from position 1 to 159), the same four aa substitutions can be observed. Except for K-14A and M70V, the Florida clade 1 South American strains have the same aa substitutions as A/eq/Florida/146609/11.

3. Discussion

Since 1963, the EIV H3N8 subtype has been the cause of numerous outbreaks of respiratory disease in horses worldwide, including South American countries where the horse industry is highly significant. In this study, comprehensive phylogenetic and phylodynamic analyses of the EIV South American strains were carried out to study their diversification, to infer their origin and to estimate the dynamics of their geographical spread. This information could contribute to a better understanding of the epidemiology of the H3N8 EIV infections worldwide and particularly in South America, and could also allow Animal Health Authorities to improve the preventive and control measures currently in use.

The spatio-temporal analysis demonstrated that the H3N8 EIV strains detected in different outbreaks of acute respiratory disease in horses in South America were grouped into five well supported monophyletic clades. As was previously described, Group I is the oldest clade, and is made up of strains detected between 1963 and 1969 also grouped into the Pre-divergent lineage. It is believed that EIV of the H3N8 subtype crossed the species barrier from birds in the early 1960s and has subsequently spread worldwide [1,3,36,39]. It has been suggested that the H3N8 EIV subtype originated in South America, since it was identified after the arrival of thoroughbred horses imported from Argentina to Miami in 1963 [5,6]. Even though there are no reports of EIV circulation in Argentina at that time, there is evidence of H3N8 EIV infection and disease in horses in Chile, Brazil and Uruguay [22,23]. Group VIII, also in the Pre-divergent lineage, consists of strains circulating in 1985 [21], which comprise the first isolation of an H3N8 EIV in Argentina [31] and a reintroduction of this virus in Chile, apparently after the entry of horses from Argentina [25,26]. Surprisingly, a strain detected in USA in 1985 belonged to this group, which could be due, as in 1963, to subclinically infected horses traveling from South America to USA. The South America clade 1, described by Lai in 2001 [11], groups only Argentinian strains. This lineage circulated for a period of 3 years (between 1993 and 1996) and it has not been detected since that time. The South American clade 2, that was previously reported by Miño et al. [33], includes Argentinian and Chilean EIV strains detected between 1997 and 2006, and it has been previously suggested that based on the close resemblance of the viruses, it entered Chile from Argentina [22]. The South American clade 1 and 2 are clustered into the American lineage. These clades are made up of strains that circulated exclusively in South America between 1993 and 2006; yet, the clades would have different ancestors. Finally, the causative virus of the multiple large-scale outbreak of EI in 2012 in South America belongs to the Florida clade 1, and spread among several countries, yielding important economic losses [1,16,29]. In addition, the same virus was also detected in Dubai in 2012, in an outbreak in a quarantine facility after the entry of a group of endurance horses from Uruguay [1].

The phylogenetic patterns of EIVs identified in South America are consistent with previous studies that show that viruses from the same lineage are usually isolated in a particular geographical region [11,21]. Interestingly, the viruses of each clade detected in South America have circulated only for a few years, with no evidence of reemergence. This could be interpreted as a consequence of the development of wide population immunity due to natural infections and/or efficacious vaccination programs. The polyphyletic behavior of the viruses causing the outbreaks of equine influenza, which took place in South America between 1963 and 2012, seem to be due to the introduction of at least four different viruses, probably through the movement of subclinically infected horses. These findings are in agreement with previous consideration that incursion of the influenza virus into new equine populations can occur through subclinically infected horses facilitated by short quarantine periods when horses travel internationally for competitions or sales [10,34].

Although the complete HA data set included all the HA complete sequences available at the Influenza Research Database (n = 144), the most related sequences of the Group I, Group VIII and South America clade 1 could not be estimated. Possibly, the analysis of other genome segments could bring up more information and help to estimate the most related strains of these old South American viruses. Additionally, the spatial and temporal dynamics for the geographic spread provided more information on the migration patterns of the H3N8 EIV. The ancestral location for Group I could not be determined with confidence; this could be due to a possible biased sampling of the EIV H3N8 during that period for which few isolates were sequenced. Despite being unable to accurately infer the possible ancestors for Group VIII and South America clade 1, migration links were established from USA to South America, so there is a high probability of the former being the ancestral location of these groups, as well as of South America clade 2 and Florida clade 1. In agreement with previous reports [5,6], the discrete phylogeographic model suggested South America as the starting point of the spread of the EIV H3N8. In the map, migration links were established repeatedly between North and South America, which is consistent with the frequent movement of horses that exists between USA and South America and vice versa. Thoroughbred shuttle stallions, for example, travel from USA to South America every breeding season and several thoroughbred mares are bought for genetic improvement every year; furthermore, polo horses travel from Argentina for the polo season in USA, coming back to Argentina when the season is over. Additionally, the diffusion process also suggested USA as an important primary origin for the spread of the EIV H3N8 into Asia and Europe. This is consistent with a report of the first detection of this virus in Europe, in an important outbreak of EI in Switzerland in 1965, which was considered a major transcontinental pandemic of the EIV H3N8 [21,40]. However, a sub notification of EIV outbreaks could exist in other countries that could act as important spreaders of the virus. The tMRCA estimated for the five clades detected in South America suggests that the ancestor strains had been circulating in the Northern hemisphere at least one year before their effective circulation in South America, with the exception of Miami 63, that apparently moved from South America to North America [5,6].

As an aid in the understanding of the evolutionary pathways and population dynamics of the H3N8 EIV, a demographic reconstruction was used to examine the changes of the effective number of viral population size. The relative genetic diversity observed between 1963 and 2008 was similar to that previously described by Murcia et al. [21]. In our work, the analyzed time was extended from 2008 onwards. Since 2006, an accelerated increase in the relative genetic diversity was observed, with a peak in 2007, followed by an abrupt decline in 2009. The co-circulation of the South American clade 2, Florida clade 1 and Florida clade 2 sub-lineages strains [41] could be the reason of the observed increase in the relative genetic diversity. The decrease observed since 2009 could have resulted from the introduction of the Florida clade 2 strain in the vaccines, as recommended by the ESP [42]. However, not all the manufacturers have updated their vaccines to contain Florida clade 2 strain. For this reason, another possible explanation may be the lack of active surveillance in some regions of the globe, or the lack of sequences being made publicly available. The limited viral diversity found in the H3N8 EIV strains is noteworthy, and it was homologated to the one observed with Influenza B virus in humans, with only two lineages circulating [11,34].

Multiple amino acid substitutions have occurred in the HA of the viruses detected in South America since the first description in 1963. The strains belonging to the Group I, Group VIII, South American clade 1 and South American clade 2, have aa changes at antigenic site B compared with the prototype strains. This observation is important because the fact that antigenic site B is positioned at the top of HA1 implies that aa changes occurring in this site could compromise the viral antigenicity [43,44]. In addition, some amino acid substitutions such as N312K and S92N present in the South American clade 1 and 2, respectively, seem to be exclusive to South American strains. These aas could be hence used as molecular markers to identify the reemergence of those viruses in the future. On the contrary, no aa changes in the antigenic site B were observed in Florida clade 1 strains detected in South America in 2012 when comparing it with the prototype strain A/eq/Ohio/2003. Nevertheless, two subpopulations characterized by the aa substitutions K-14A and M70V (not in antigenic sites) were observed indicating the occurrence of some degree of evolutionary change. These aa changes are only found in South American strains, and M70V was observed only in Argentinian strains [16]. Nevertheless, these changes, outside the antigenic sites, do not imply the necessity to replace the vaccine strain in regard to the Florida clade 1 strain, A/eq/Ohio/2003, recommended by the ESP. This is additionally supported by the antigenic analysis carried out with 2012 Brazilian and Argentinian isolates, that showed a good recognition by sera raised against other members of Florida clade 1 [16,29]. It is important to highlight that at the time of the last outbreak of EI in South America, the vaccines on the market were not in line with the ESP recommendations; fortunately, after the outbreak many vaccine manufacturers in Uruguay and Argentina updated their vaccine strains [16,29].

4. Experimental Section

4.1. HA Gene Sequencing

The complete HA nucleotide sequence data of 19 viruses detected in different outbreaks of EI in Argentina during the period 1993 and 2012, and one strain detected in Uruguay in 2012, were obtained (Table S1). The sequence data of 13 of these viruses were derived from RNA extracted from nasopharyngeal swabs; the remaining 7 viruses (A/eq/Argentine/E131/1994, A/eq/Argentine/E226/1995, A/eq/Argentine/E433/1997, A/eq/Argentine/E652/1999, A/eq/Argentine/E750/1999, A/eq/Argentine/E1134-7/2001 and A/eq/Argentine/E6189-3/2005) required amplification in embryonated hen's eggs prior to sequencing.

The complete HA gene was amplified in two separate reactions by the OneStep RT-PCR kit (Qiagen) using primers previously described by Gildea et al. [12]. The PCR products were purified using the ExoStar (Illustra™, GE Healthcare UK Limited, Amersham, UK), according to manufacturer's recommendations, and sequenced by the "Unidad de Genómica, Instituto de Biotecnología, INTA, Hurlingham".

4.2. Sequence Analysis and Phylogenetic Trees

The obtained HA nucleotide sequences were edited with the BioEdit software v7.0.9.0 [45] and were aligned using ClustalW in the software package of BioEdit, along with 144 reference HA complete sequences (complete HA data set) available at the Influenza Research database (IRD) up to January 2015 and one sequence from Brazil obtained from the GISAID EpiFlu™ Database. In addition, a partial HA1 nucleotide sequence (from position 52 to 545) from a strain detected in Chile in 2012 available at the IRD, was also analyzed in the partial HA1 data set (Table S2).

Phylogenetic trees for partial HA1 and complete HA nucleotide sequences were inferred by the maximum likelihood (ML) method using PhyML v3.1 software [46]. The best fit evolutionary models of nucleotide substitution were estimated according to the Akaike Information Criterion (AIC) statistics obtained with the jModelTest v2.1.6 software [47]. Branch connecting sequence A/equine/Uruguay/1/1963 detected in Uruguay in 1963 with the others was selected for rooting the phylogenetic trees inferred from both data sets.

4.3. Phylodynamic Analysis

The nucleotide sequences including the information about location and the year of isolation from the complete HA data set were introduced in the Bayesian coalescent analysis to estimate the time of the most recent common ancestor (tMRCA), the population dynamics and the spatio-temporal diffusion process. The analysis was carried out using the best fit evolutionary model of nucleotide substitution (see above), the Bayesian Skyline Plot (BSP) model and the relaxed uncorrelated lognormal molecular clock. A uniform prior to nucleotide substitution rate from 1×10^{-3} to 3×10^{-3} s/s/year was used to carry out the analysis, which included the nucleotide substitution rate for HA gene of the EIV H3N8 previously calculated by Murcia et al. [21], corresponding to 1.7×10^{-3} s/s/year. To infer the ancestor location and viral migration events, we grouped the H3N8 isolates into 26 states (countries) using the discrete phylogeographic asymmetric model implemented in BEAST v1.8.3 software [48], which uses a continuous time Markov Chain (CTMC) over discrete sampling locations and allows different rates of diffusion between each pair of locations [49] Analyses were run using BEAGLE library [50,51] at the CIPRES Science Gateway server [52]. The convergence of the parameters to a stationary distribution was assessed with the TRACER v1.6 software [53]. Two independent analyses were performed and results were combined. The trees were summarized in a maximum clade credibility (MCC) tree with temporal and spatial annotation using the Tree Annotator v1.8.3 software after discarding the burn-in samples. To visualize the geographic migration of the virus over the time, a keyhole markup language (KML) file was generated using SpreaD3 v0.9.6 [54], which can be visualized via Google Earth V. Monophyly was constrained for all sequences except for sequence A/eq/Uruguay/1/1963 detected in Uruguay in 1963.

4.4. Amino Acid Analysis

The deduced amino acid sequences of HA1 were aligned with representative strains and the most related sequences of each group. Alignments were analyzed using the BioEdit v7.0.9.0 software [45].

5. Conclusions

As a conclusion, EIV outbreaks that have taken place in South American countries were caused by at least four different and independent introductions of the virus, presumably brought in through the international movement of horses. Afterwards, these viruses circulated and induced outbreaks of acute respiratory disease at a regional level, as the exchange of horses for competition amongst South American countries is extremely frequent. Unfortunately, there is no information available related to EIV infections in other South American countries; thus, we may infer that either the disease has not occurred, or there are no surveillance programs, or there is an unavailability of diagnostic laboratories or a sub-notification of EIV outbreaks, or a combination of these. Finally, taking into account that the spread of EIV worldwide is facilitated by the international movement of horses,

adequate epidemiological surveillance programs, appropriate quarantine procedures and the maintenance of high levels of population immunity through vaccination with updated and efficacious vaccines, are critical aspects in the control of EI.

Author Contributions: All the authors contributed to conceive and design the experiments. Cecilia Olguin Perglione, Marcelo Dario Golemba and Carolina Torres performed the experiments and analyzed the data. Cecilia Olguin Perglione wrote the paper under the supervision of Maria Barrandeguy, while Marcelo Dario Golemba and Carolina Torres assisted in the writing. All the authors commented on manuscript drafts.

Conflicts of Interests: The authors declared no conflict of interest.

References

1. Woodward, A.L.; Rash, A.S.; Blinman, D.; Bowman, S.; Chambers, T.M.; Daly, J.M.; Damiani, A.; Joseph, S.; Lewis, N.; McCauley, J.W., *et al.* Development of a surveillance scheme for equine influenza in the uk and characterisation of viruses isolated in europe, dubai and the USA from 2010–2012. *Vet. Microbiol.* **2014**, *169*, 113–127.
2. Gildea, S.; Arkins, S.; Cullinane, A. Management and environmental factors involved in equine influenza outbreaks in ireland 2007–2010. *Equine Vet. J.* **2011**, *43*, 608–617.
3. Cullinane, A.; Elton, D.; Mumford, J. Equine influenza - surveillance and control. *Influenza Other Respir. Viruses* **2010**, *4*, 339–344.
4. Sovinova, O.; Tumova, B.; Pouska, F.; Nemec, J. Isolation of a virus causing respiratory disease in horses. *Acta Virol.* **1958**, *2*, 52–61.
5. Scholtens, R.G.; Steele, J.H.; Dowdle, W.R.; Yarbrough, W.B.; Robinson, R.Q. U.S. Epizootic of equine influenza, 1963. *Public Health Rep* **1964**, *79*, 393–402.
6. Waddell, G.H.; Teigland, M.B.; Sigel, M.M. A new influenza virus associated with equine respiratory disease. *J. Am. Vet. Med. Asso.c* **1963**, *143*, 587–590.
7. Webster, R.G. Are equine 1 influenza viruses still present in horses? *Equine Vet. J.***1993**, *25*, 537–538.
8. Kawaoka, Y.; Bean, W.J.; Webster, R.G. Evolution of the hemagglutinin of equine h3 influenza viruses. *Virology* **1989**, *169*, 283–292.
9. Daly, J.M.; MacRae, S.; Newton, J.R.; Wattrang, E.; Elton, D.M. Equine influenza: A review of an unpredictable virus. *Vet. J.* **2011**, *189*, 7–14.
10. Daly, J.M.; Lai, A.C.; Binns, M.M.; Chambers, T.M.; Barrandeguy, M.; Mumford, J.A. Antigenic and genetic evolution of equine h3n8 influenza a viruses. *J. Gen. Virol.* **1996**, *77*, 661–671.
11. Lai, A.C.; Chambers, T.M.; Holland, R.E., Jr.; Morley, P.S.; Haines, D.M.; Townsend, H.G.; Barrandeguy, M. Diverged evolution of recent equine-2 influenza (h3n8) viruses in the western hemisphere. *Arch. Virol.* **2001**, *146*, 1063–1074.
12. Gildea, S.; Quinlivan, M.; Arkins, S.; Cullinane, A. The molecular epidemiology of equine influenza in ireland from 2007–2010 and its international significance. *Equine. Vet. J.* **2012**, *44*, 387–392.
13. Gildea, S.; Fitzpatrick, D.A.; Cullinane, A. Epidemiological and virological investigations of equine influenza outbreaks in ireland (2010–2012). *Influenza Other Respir. Viruses* **2013**, *7 Suppl 4*, 61–72.
14. Ito, M.; Nagai, M.; Hayakawa, Y.; Komae, H.; Murakami, N.; Yotsuya, S.; Asakura, S.; Sakoda, Y.; Kida, H. Genetic analyses of an h3n8 influenza virus isolate, causative strain of the outbreak of equine influenza at the kanazawa racecourse in japan in 2007. *Vet. Med. Sci.* **2008**, *70*, 899–906.

15.　King, E.; Macdonal, D. Report of the board of inquiry appointed by the board of the national horseracing authority to conduct enquiry into the causes of the equine influenza which started in the western cape in early december 2003 and spread to the eastern cape and gauteng. *Aust. Vet. J.* **2004**, *23*, 139–142.

16.　Perglione, C.O.; Gildea, S.; Rimondi, A.; Mino, S.; Vissani, A.; Carossino, M.; Cullinane, A.; Barrandeguy, M. Epidemiological and virological findings during multiple outbreaks of equine influenza in south america in 2012. *Influenza Other Respir. Viruses* **2016**, *10*, 37–46.

17.　Watson, J.; Daniels, P.; Kirkland, P.; Carroll, A.; Jeggo, M. The 2007 outbreak of equine influenza in australia: Lessons learned for international trade in horses. *Rev Sci Tech* **2011**, *30*, 87–93.

18.　Yondon, M.; Heil, G.L.; Burks, J.P.; Zayat, B.; Waltzek, T.B.; Jamiyan, B.O.; McKenzie, P.P.; Krueger, W.S.; Friary, J.A.; Gray, G.C. Isolation and characterization of h3n8 equine influenza a virus associated with the 2011 epizootic in mongolia. *Influenza Other Respir. Viruses* **2013**, *7*, 659–665.

19.　Virmani, N.; Bera, B.C.; Singh, B.K.; Shanmugasundaram, K.; Gulati, B.R.; Barua, S.; Vaid, R.K.; Gupta, A.K.; Singh, R.K. Equine influenza outbreak in india (2008-09): Virus isolation, sero-epidemiology and phylogenetic analysis of ha gene. *Vet. Microbiol.* **2010**, *143*, 224–237.

20.　Qi, T.; Guo, W.; Huang, W.Q.; Li, H.M.; Zhao, L.P.; Dai, L.L.; He, N.; Hao, X.F.; Xiang, W.H. Genetic evolution of equine influenza viruses isolated in china. *Arch. Virol.* **2010**, *155*, 1425–1432.

21.　Murcia, P.R.; Wood, J.L.; Holmes, E.C. Genome-scale evolution and phylodynamics of equine h3n8 influenza a virus. *J. Virol.* **2011**, *85*, 5312–5322.

22.　Muller, I.; Pinto, E.; Santibanez, M.C.; Celedon, M.O.; Valenzuela, P.D. Isolation and characterization of the equine influenza virus causing the 2006 outbreak in chile. *Vet. Microbiol.* **2009**, *137*, 172–177.

23.　Fuschlocher, F.; Zurita, L.; Latorre, G.; Palavicino, I. Influenza equina en la provincia de santiago, In *V Convención Nacional de Medicina Veterinaria*; Boletin: Valdivia, Chile, 1963; pp. 40–46.

24.　Casanova, A.; Martínez, I.; Román, M. Aislamiento y tipificación del virus de influenza equina en chile. *Arch. Med. Vet.* **1977**, *9*, 91–93.

25.　Berríos, P. Influenza equina en chile (1963–1992). Un posible caso en un ser humano. *Rev. Chil. Infect.* **2005**, *22*, 47–50.

26.　Celedón, M.O.; De Negri, L.; Santibáñez, M.; Berríos, P. Brote de influenza equina en chile causado por el subtipo h3n8. *Agro-Ciencia* **1992**, *8*, 47–48.

27.　OIE: World animal health information database (WAHID) interface. Avaialbe online: http://www.oie.int/wahis_2/public/wahid.php/Wahidhome/Home (accessed on 26 October 2016).

28.　Mancini, D.A.; Mendonca, R.M.; Pinto, J.R.; Mori, E.; Fernandes, W. Anti-human influenza protector antibody detected in horses as a zoonotic viruses. *Braz. J. Vet. Res. Anim. Sci.* **2004**, *41*, 379–383.

29.　Alves Beuttemmuller, E.; Woodward, A.; Rash, A.; Dos Santos Ferraz, L.E.; Fernandes Alfieri, A.; Alfieri, A.A.; Elton, D. Characterisation of the epidemic strain of h3n8 equine influenza virus responsible for outbreaks in south america in 2012. *Virol. J.* **2016**, *13*, 45.

30.　Fain Binda, J.; Martin, J. Aislamiento del agente etiológico de la epizootia de influenza equina en rosario (argentina) 1976: Myxovirus influenza a/equi 1/rosario 76. *Revista de Medicina Veterinaria Buenos Aires* **1977**, *58*, 261–268.

31.　Nosetto, E.; Pecoraro, M.; Galosi, C.m.; Massone, R.; Cid de la Paz, V.; Ando, R.; Ando, Y.; Etcheverrigaray, M.E. Isolation of an equine influenza virus strain and epizootiological study of the 1985-86 outbreak in argentina. *Rev. Sci. Tech. Off. Int. Epiz.* **1989**, *8*, 123–128.

32.　Barrandeguy, M.; Bryans, J.; San Roman, A.; Parreño, V.; Ricci, L.; Schudel, A.; Kness, V.; Fernandez, F. Hallazgos virológicos y serológicos en un brote de influenza equina en los hipódromos de palermo y san isidro ocurrido en abril de 1993. In Proceedings of the XI Jornadas de Actualización Técnico Científicas de la Asociación Argentina de Veterinaria Equina, La Plata, Buenos Aires, Argentina, 9-19 December **1993**.

33.　Miño, S.; Vissani, A.; Trono, K.; Barrandeguy, M. Evolución genética de virus de influenza equina-2 (h3n8) detectados en argentina en los últimos años. In Proceedings of the Congreso Panamericano de Ciencias Veterinarias, Santiago, Chile, 13–16 November 2006.

34.　Lewis, N.S.; Daly, J.M.; Russell, C.A.; Horton, D.L.; Skepner, E.; Bryant, N.A.; Burke, D.F.; Rash, A.S.; Wood, J.L.; Chambers, T.M., *et al.* Antigenic and genetic evolution of equine influenza a (h3n8) virus from 1968 to 2007. *J. Virol.* **2011**, *85*, 12742–12749.

35.　Wiley, D.C.; Skehel, J.J. The structure and function of the hemagglutinin membrane glycoprotein of influenza virus. *Ann. Rev. Biochem.* **1987**, *56*, 365–394.

36.　Cullinane, A.; Newton, J.R. Equine influenza—a global perspective. *Vet. Microbiol.* **2013**, *167*, 205–214.

37. Elton, D.; Cullinane, A. Equine influenza: Antigenic drift and implications for vaccines. *Equine Vet. J.***2013**, *45*, 768–769.

38. Olguin Perglione, C.; Golemba, M.D.; Barrandeguy, M. Molecular evolution of h3n8 equine influenza virus in argentina. In Proceedings of the X International Conferences on Equine Infecious Diseases, Buenos Aires, Argentina, 2016; p. S74.

39. Worobey, M.; Han, G.Z.; Rambaut, A. A synchronized global sweep of the internal genes of modern avian influenza virus. *Nature* **2014**, *508*, 254–257.

40. Paccaud, M.F.; Couard, M.; Bürki, F.; Gerber, H.; Löhrer, J. Outbreak of influenza a/equi-2 in switzerland. *Nature* **1966**, *2011*, 101–102.

41. Bryant, N.A.; Rash, A.S.; Russell, C.A.; Ross, J.; Cooke, A.; Bowman, S.; MacRae, S.; Lewis, N.S.; Paillot, R.; Zanoni, R., *et al.* Antigenic and genetic variations in european and north american equine influenza virus strains (h3n8) isolated from 2006 to 2007. *Vet. Microbiol.* **2009**, *138*, 41–52.

42. OIE. Expert surveillance panel on equine influenza vaccine composition-conclusions and recommendations. *Off. Int. Epizoot. Bull.* **2010**, 44–45.

43. Woodward, A.; Rash, A.S.; Medcalf, E.; Bryant, N.A.; Elton, D.M. Using epidemics to map h3 equine influenza virus determinants of antigenicity. *Virology* **2015**, *481*, 187–198.

44. Jin, H.; Zhou, H.; Liu, H.; Chan, W.; Adhikary, L.; Mahmood, K.; Lee, M.S.; Kemble, G. Two residues in the hemagglutinin of a/fujian/411/02-like influenza viruses are responsible for antigenic drift from a/panama/2007/99. *Virology* **2005**, *336*, 113–119.

45. Hall, T.A. Bioedit: A user-friendly biological sequence alingment editor and analysis program for windows 95/98/nt. *Nucl Acid Symp* **1999**, *41*, 95–98.

46. Guindon, S.; Dufayard, J.F.; Lefort, V.; Anisimova, M.; Hordijk, W.; Gascuel, O. New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of phyml 3.0. *Syst. Biol.* **2010**, *59*, 307–321.

47. Posada, D. Jmodeltest: Phylogenetic model averaging. *Mol. Biol. Evol.***2008**, *25*, 1253–1256.

48. Drummond, A.J.; Suchard, M.A.; Xie, D.; Rambaut, A. Bayesian phylogenetics with beauti and the beast 1.7. *Mol. Biol. Evol.* **2012**, *29*, 1969–1973.

49. Lemey, P.; Rambaut, A.; Drummond, A.J.; Suchard, M.A. Bayesian phylogeography finds its roots. *PLoS Comput. Biol.* **2009**, *5*, e1000520.

50. Suchard, M.A.; Rambaut, A. Many-core algorithms for statistical phylogenetics. *Bioinformatics* **2009**, *25*, 1370–1376.

51. Ayres, D.L.; Darling, A.; Zwickl, D.J.; Beerli, P.; Holder, M.T.; Lewis, P.O.; Huelsenbeck, J.P.; Ronquist, F.; Swofford, D.L.; Cummings, M.P., *et al.* Beagle: An application programming interface and high-performance computing library for statistical phylogenetics. *Syst. Biol.* **2012**, *61*, 170–173.

52. Miller, M.A.; Pfeiffer, W.; Schwartz, T. Creating the cipres science gateway for inference of large phylogenetic trees, In Proceedings of the Gateway Computing Environments Workshop (GCE), 2010, New Orleans, LA, 14 November 2010; pp. 1–8.

53. Rambaut, A.; Suchard, M.A.; Xie, D.; Drummond, A.J. Tracer v1.6. Available online http://beast.bio.ed.ac.uk/Tracer (14 October **2014**).

54. Bielejec, F.; Baele, G.; Vrancken, B.; Suchard, M.A.; Rambaut, A.; Lemey, P. Spread3: Interactive visualization of spatiotemporal history and trait evolutionary processes. *Mol. Biol. Evol.* **2016**, *33*, 2167–2169.