Genetic characterization of influenza viruses from influenza-related hospital admissions in the St. Petersburg and Valencia sites of the Global Influenza Hospital Surveillance Network during the 2013/14 influenza season

F. Xavier López-Labrador [a,b,c], Angels Natividad-Sancho [a], Maria Pisareva [d], Andrey Komissarov [d], Karina Salvatierra [a,b,1], Artem Fadeev [d], Andrés Moya [b], Mikhail Grudinin [d], Javier Díez-Domingo [a], Olga Afanasieva [d], Nadezhda Konovalova [d], Anna Sominina [d], Joan Puig-Barberà [a,*]

[a] Fundación para el Fomento de la Investigación Sanitaria y Biomédica de la Comunidad Valenciana (FISABIO), Avda. de Catalunya, 21, 46020 Valencia, Spain
[b] Joint Units of Infection and of Genomics and Health, FISABIO/Cavanilles Institute for Biodiversity and Evolutionary Biology, University of Valencia, Spain
[c] Centro de Investigación Biomédica en Red de Epidemiología y Salud Publica (CIBER-ESP), Instituto de Salud Carlos III, Spain
[d] Research Institute of Influenza, Ministry of Health, Prof. Popov Str. 15/17, St. Petersburg, Russian Federation

ARTICLE INFO

ABSTRACT

Background: Continuous surveillance for genetic changes in circulating influenza viruses is needed to guide influenza prevention and control.

Objectives: To compare intra-seasonal influenza genetic diversity of hemagglutinin in influenza A strains isolated from influenza hospital admissions collected at two distinct sites during the same season.

Study design: Comparative phylogenetic analysis of full-length hemagglutinin genes from 77 isolated influenza A viruses from the St. Petersburg, Russian Federation and Valencia, Spain sites of the Global Influenza Hospital Surveillance Network (GIHSN) during the 2013/14 season.

Results: We found significant variability in A(H3N2) and A(H1N1)pdm09 viruses between the two sites, with nucleotide variation at antigenic positions much lower for A(H1N1)pdm09 than for A(H3N2) viruses. For A(H1N1)pdm09, antigenic sites differed by three to four amino acids from the vaccine strain, two of them common to all tested isolates. For A(H3N2) viruses, antigenic sites differed by six to nine amino acids from the vaccine strain, four of them common to all tested isolates. A fifth amino acid substitution in the antigenic sites of A(H3N2) defined a new clade, 3C.2. For both influenza A subtypes, pairwise amino acid distances between circulating viruses and vaccine strains were significantly higher at antigenic than at non-antigenic sites. Whereas A(H1N1)pdm09 viruses clustered with clade 6B and 94% of A(H3N2) with clade 3C.3, at both study sites A(H3N2) clade 3C.2 viruses emerged towards the end of the season, showing greater pairwise amino acid distances from the vaccine strain compared to the predominant clade 3C.3.

Conclusions: Influenza A antigenic variants differed between St. Petersburg and Valencia, and A(H3N2) clade 3C.2 viruses were characterized by more amino acid differences from the vaccine strain, especially at the antigenic sites.

1. Background

As a consequence of error-prone replication of the viral RNA genome and host immune pressure, the antibody-binding sites of influenza A hemagglutinins (HAs) continuously accumulate mutations [1,2]. The emergence of variants escaping natural or vaccine-induced immune responses (antigenic drift) leads to annual influenza epidemics in humans [3,4], a major cause of morbidity and mortality and a significant economic burden on health services [5]. Continuous surveillance of circulating influenza strains is needed to identify new mutations in HA that can escape from host antibody responses [6–8] and thereby guide influenza prevention and control [9].

2. Objectives

Here, we compared the intra-seasonal influenza HA genetic diversity and differences from the vaccine strains in the A(H3N2) and A(H1N1)pdm09 influenza viruses collected at two sites (Valencia, Spain and St. Petersburg, Russian Federation) of the Global Influenza Hospital Surveillance Network (GIHSN) during the 2013/14 Northern Hemisphere influenza season. The GIHSN focuses exclusively on severe cases of respiratory infection requiring hospitalization and conducts a prospective, active surveillance, hospital-based epidemiological study over consecutive seasons [10].

3. Study design

3.1. Samples

Clinical samples for the study were collected from a prospective surveillance scheme conducted at St. Petersburg, Russian Federation and Valencia, Spain during the 2013/14 Northern Hemisphere influenza season. Study details and inclusion criteria are described elsewhere [10]. A nasopharyngeal and a pharyngeal swab were obtained from each patient using flocked swabs, submerged in transport media, and frozen at −20 °C until shipment to the reference virology laboratory for analysis (Virology Laboratory, FISABIO-Public Health, Valencia, Spain and World Health Organization National Influenza Centre (RII NIC), St. Petersburg, Russian Federation).

3.2. Molecular diagnosis, typing, and sequencing of influenza viruses

Molecular diagnosis, typing, and sequencing of influenza viruses complied with the World Health Organization recommendations [11]. In Valencia, total nucleic acid was extracted using an automated method (Nuclisens Easy-Mag, bioMérieux, Lyon, France). Influenza A and B viruses were screened by multiplex real-time reverse-transcription polymerase chain reaction (RT-PCR) in a Lightcycler 480II (Roche Applied Science, Penzberg, Germany) as described [12,13]. A typing real-time RT-PCR assay was performed following the World Health Organization protocol [11] for all influenza A-positive samples.

In St. Petersburg, viral RNA was extracted using a RIBO-prep kit (InterLabService, Russian Federation) or an RNeasy Mini kit (Qiagen, Hilden, Germany). RT-PCR for influenza A and B viruses was performed using AmpliSens Influenza virus A/B-FL, AmpliSens Influenza virus H1/swine-FL, and AmpliSens Influenza virus A type-FL kits (InterLabService, Russian Federation) with reverse transcription using an AmpliSens Reverta-L kit (InterLab-Service, Russian Federation) or an AgPath-ID[TM] One-Step RT-PCR kit (Ambion, USA) with US Centers for Disease Control and Pre-vention primers and probes to determine the lineage (Yamagata or Victoria) for influenza B viruses.

HA nucleotide sequencing was performed on every fourth sample positive for influenza A (Valencia) or on convenience samples (cycle threshold values <25) taken at the beginning, middle, and end of the influenza season (St. Petersburg). At both sites, the complete HA gene (HA1/HA2 regions) was amplified by long-range RT-PCR and sequenced by standard Sanger chemistry using gene-specific primers for the corresponding virus subtype [11].

3.3. Influenza variability and phylogenetic analysis

3.3.1. Sequence data

Nucleotide alignments were obtained using the ClustalW algorithm integrated in BioEdit ver.7.2.5 (Ibis Biosciences, Carlsbad, CA, USA) [14], including reference sequences (defining known phylogenetic clades) from the Global Initiative on Sharing Avian Influenza Data (GISAID) database (Annex 1 of Supplementary information). Maximum-likelihood phylogenetic trees were reconstructed using the whole HA gene sequence with the online PhyML platform (http://www.atgc-montpellier.fr/phyml). The best-fitting nucleotide substitution model (general time-reversible + gamma distribution among the sites) was estimated using Modeltest [15] integrated in MEGA 5.0 [16]. Branch reliability was evaluated by an approximate likelihood-ratio test (Chi-squared statistic-based) with an interior branch cut-off value of 0.9. The trees were rooted on the 2013/14 recommended vaccine strains A/California/7/2009 for A(H1N1)pdm09 or A/Texas/50/2012 for A(H3N2) viruses [17].

3.3.2. Genetic complexity and diversity

Genetic polymorphism and diversity [18] between Valencia and St. Petersburg viruses were estimated with DnaSP version 5 [19]. All parameters were calculated for the whole HA gene region and for the HA1 and HA2 subunits separately, as well as for non-antigenic sites only or for antigenic sites only (A–E of the H3 scheme and their equivalents Sa, Sb, Ca1, Ca2 and Cb for H1; reviewed in [20]). Nucleotide diversity (haplotypes) was calculated either considering only antigenic sites (aHap) or considering the full-length HA gene (Hap), to estimate the fraction of individual viral isolates for each haplotype. Pairwise amino acid distances between individual viral isolates and the corresponding vaccine strain were calculated using the p-distance model implemented in MEGA 6 [21]. Non-parametric Kruskal–Wallis tests were used to compare pairwise amino acid distances from vaccine strains. All statistical analyses were performed using Stata version 12 (StataCorp, College Station, TX).

4. Results

4.1. Samples

Samples were collected from early December, 2013 to mid-March, 2014, and half were collected during the seasonal peak of activity (January, 2014 for A(H1N1)pdm09 and February, 2014 for A(H3N2) viruses). Complete HA gene sequences could be determined for 34 A(H1N1)pdm09 and 26 A(H3N2) isolates collected in Valencia and for eight A(H1N1)pdm09 and nine A(H3N2) isolates collected in St. Petersburg. For A(H1N1)pdm09-positive samples, patient ages ranged from 0.1 to 83 years, with a median of 62 years for Valencia and 5 years for St. Petersburg. For A(H3N2)-positive samples, patient ages ranged from 2 to 94 years, with a median of 73 years for Valencia and 18 years for St. Petersburg.

Table 1
Genetic polymorphism estimates for HA domains in each viral population for influenza A(H1N1)pdm09 and A(H3N2) viruses.

Subtype	Variable	HA total		Antigenic sites	
		Valencia	St. Petersburg	Valencia	St. Petersburg
A(H1N1)pdm09	Total number of sites	1647		150	
	Number of sequences	34	8	34	8
	H	21	6	6	2
	Hd	0.911	0.929	0.41	0.25
	S	65	20	4	1
	Eta (η)	67	20	5	1
	Singleton variable sites	46	12	3	1
	k	6.752	6.893	0.533	0.25
	Pi(π)	0.0041	0.00419	0.00355	0.00167
	Theta (θ) per site	0.00965	0.00468	0.00815	0.00257
	Ss	45	15	3	0
	πs	0.01248	0.01407	0.01186	0
	θsyn per site	NA	0.01541	0.02086	0
	Sa	22	5	2	1
	πa	0.00163	0.00126	0.00101	0.00219
	θa per site	NA	0.00152	0.00426	0.00336
A(H3N2)	Total number of sites	1650		393	
	Number of sequences	26	9	26	9
	H	21	7	13	6
	Hd	0.966	0.917	0.788	0.833
	S	61	42	17	19
	Eta (η)	61	42	17	19
	Singleton variable sites	45	32	13	14
	k	6.991	12.111	2.123	5.556
	Pi(π)	0.00424	0.00734	0.00540	0.01414
	Theta (θ) per site	0.00969	0.00937	0.01134	0.01779
	Ss	39	24	12	8
	πs	0.01255	0.01967	0.0218	0.02934
	θsyn per site	0.02808	0.02425	0.03945	0.03532
	Sa	22	18	5	11
	πa	0.00188	0.00385	0.00123	0.01005
	θa per site	0.00448	0.00515	0.00418	0.01307

Abbreviations: H, number of haplotypes; Hd, haplotype gene diversity; S, total number of segregating (polymorphic) sites; Eta (η), total number of mutations; k, average number of nucleotide differences; NA, not applicable; Pi (π), nucleotide diversity (average number of nucleotide differences per site between two sequences); theta (θ), heterozygosis per site expected in a population in mutation-drift equilibrium given the observed S value; Ss, number of polymorphic synonymous changes; πs, nucleotide diversity at synonymous sites; θsyn, heterozygosis per site at synonymous loci; Sa, number of polymorphic non-synonymous changes; πa, nucleotide diversity at non-synonymous sites; θa, heterozygosis per site at non-synonymous loci.

4.2. Overall variability of influenza viruses between study sites

Across the full HA gene, A(H1N1)pdm09 and A(H3N2) HA viral sequences from Valencia showed more nucleotide polymorphic sites than those from St. Petersburg (Table 1). No single nucleotide site differentiated all isolates from St. Petersburg from all isolates from Valencia, or the converse (data not shown). A(H3N2) viruses shared more mutations across study sites than A(H1N1)pdm09 viruses. Nucleotide diversity was higher for A(H3N2) isolates from St. Petersburg than for those from Valencia. At both locations, estimated nucleotide diversity for A(H3N2) viruses was higher in antigenic sites than in other regions of HA, whereas for A(H1N1)pdm09, diversity was lower in antigenic sites than in other HA regions.

4.3. Phylogenetic analysis and differences in antigenic sites

4.3.1. A(H1N1)pdm09 viruses

Influenza A(H1N1)pdm09 viruses from Valencia and St. Petersburg clustered in monophyletic subgroup 6B, represented by A/South Africa/3626/2013 (Fig. 1a). Up to 93% (39/42) of the viruses belonged to the same antigenic-site amino acid haplotype (aHap1, Fig. 2a) defined by substitutions K163Q (Sa), S185T (Sb), and S203T (Ca1) relative to the vaccine strain A/California/7/2009. One virus (St. Petersburg, February 2014) had a substitution K163R in antigenic site Sa. Two remaining isolates (Valencia, December 2013–January 2014) had a fourth substitution at S74 in antigenic site Cb.

Analysis of non-antigenic sites revealed seven additional mutations: P83S, D97N, A256T, and K283E in HA1; and E374K, S451N, and E499K in HA2 (Fig. 2b). Only one isolate (Valencia, Jan. 2014) contained the D97S substitution also present in the 2012/13 viruses. In addition, 38/42 (90%) of all viruses had the I321V mutation and 22/42 (52%) carried the A315V mutation (HA1). The remaining substitutions (n = 21) were mostly singletons in a few positions of HA1 or scattered across HA2. Overall, variants outside the antigenic sites divided antigenic aHap1 into 11 sub-haplotypes, but only two, Hap3 and Hap6 (differing at residue A315V), accounted for 64% of all viruses (Fig. 2b). Viruses from St. Petersburg had, on average, the same number of amino acid changes as those from Valencia (12.3 vs. 12.5 for the full-length HA, 9.0 vs. 8.9 for HA1, and 3.3 vs. 3.6 for HA2).

Overall, haplotypes were population-specific (P = 0.008 by Chi-squared test) and only Hap3 was shared at both sites (3/8 [37%] from St. Petersburg, 5/34 [15%] from Valencia). The peak of haplotype diversity occurred in February in St. Petersburg and in late January 2014 in Valencia (Fig. 3a and b). Hap3, common to both study sites, was detected from mid-late January 2014 until the end of the season in both locations, whereas the Valencia-specific Hap6, present in 19/34 (56%) viruses, circulated between December 2013 and March 2014. Based on the sampling date, we deduced that intra-seasonal antigenic sites haplotypes co-circulated contemporaneously in each study site and did not displace one another over the course of the season (data not shown).
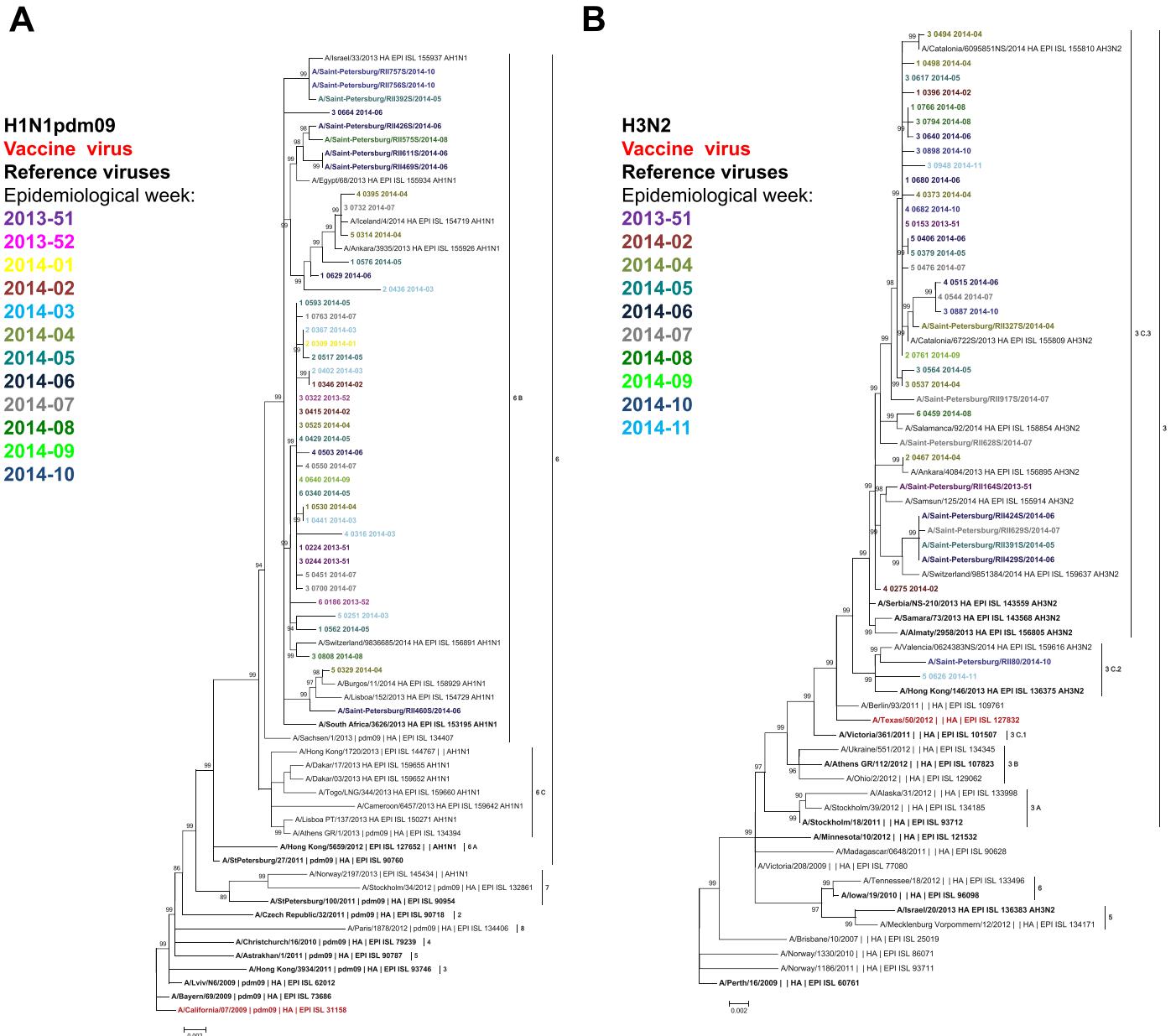
Fig. 1. Maximum-likelihood phylogenetic trees for nucleotide sequences of the complete HA gene from influenza A(H1N1)pdm09 and A(H3N2) isolates collected from the Valencia and St. Petersburg sites during the 2013/14 season.

4.3.2. A(H3N2) viruses

Most of the A(H3N2) viruses (94%; 33/35) belonged to clade 3C.3 (A/Samara/73/2013-like) and only a few (6%; 2/35) to clade 3C.2 (A/Hong Kong/146/2013-like) (Fig. 1b). Each clade had a median of 8 (range, 6–9) substitutions at antigenic sites relative to the vaccine reference strain A/Texas/50/2012. Both clades had the same number of amino acid substitutions relative to the vaccine strain and diverged at the clade-defining residue N128 (site B), resulting in the loss of a potential glycosylation site (substitution N128T at clade 3C.2 and N128A at clade 3C.3). In addition to N128A/T, for all isolates, differences between circulating and vaccine viruses included N145S (site A), V186G (site B), P198S (site B), and F219S (site D) (Fig. 4a). The following additional substitutions were present in antigenic sites: 69% (24/35) of viruses from clade 3C.3 carried the N122D substitution (site A), 91% (32/35) carried A142G (site A), and 89% (31/35) carried L157S (site B). Seven additional single substi-

tutions scattered throughout antigenic sites A–E were unique to clade 3C.3 and five were unique to clade 3C.2 (Fig. 4).

Consistent with substantial genetic diversity, clade 3C.3 had 13 aHaps (Fig. 4a), which were site-specific ($P = 0.007$ by Chi-squared test). Overall, five dominant and three high-frequency substitutions were present in the antigenic sites A, B and D, whereas singleton substitutions (n = 12) were present mostly in sites C and E. Over the whole season, aHap6 was present in 54% (19/35) of all viruses (Fig. 3c), 69% (18/26) of viruses from Valencia, and 11% (1/9) of viruses from St. Petersburg (Fig. 4a). In St. Petersburg, 44% (4/9) of the isolates had the St. Petersburg-specific aHap3 (Fig. 4a), which circulated during the peak of the influenza season (Fig. 3c) and differed from aHap6 at two residues within antigenic sites A and E. The remaining haplotypes were collected throughout the season, except for clade 3C.2 aHap9 from Valencia and aHap4 from St. Petersburg, which were collected towards the end of the season (mid- to late March). Overall, intra-seasonal haplotypes circulated

A

Antigenic site				Cb	Cb	Cb	Ca2	Sa	Ca1	Sb	Sb	Ca1	Ca1
Codon position				72	73	74	141	163	168	185	186	203	205
A/California/07/2010				T	A	S	A	K	D	S	A	S	R
aHap	Clade	N VLC	N StP	pw/dist*									
1	6B	32	7	0.06	.	.	.	.	Q	.	T	.	T	.
2	6B	0	1	0.06	.	.	.	.	R	.	T	.	T	.
3	6B	1	0	0.08	.	.	I	.	Q	.	T	.	T	.
4	6B	1	0	0.08	.	.	N	.	Q	.	T	.	T	.

B

HA					4	11	17	30	48	74 (Cb)	77	83	84	97	109	151	163 (Sa)	185 (Sb)	203 (Ca1)	209	213	223	238	256	272	283	286	315	321	374	399	443	449	451	474	479	493	499
A/California/07/2010					C	N	D	V	A	S	S	P	S	D	S	L	K	S	S	K	E	Q	D	A	V	K	I	A	I	E	H	K	V	S	T	V	A	E
Hap	Clade	N VLC	N StP	pw/dist*																																		
1	6B	2	0	0.0219	.	.	.	.	.	.	.	S	.	.	N	.	Q	T	T	.	.	.	.	.	.	T	.	K	.	.	.	N	M	.	.	K	.	.	.	.
2	6B	1	0	0.02	.	.	.	.	.	.	.	S	.	.	N	.	Q	T	T	.	.	.	.	.	.	T	.	K	.	N	.	N	.	.	.	K	.	.	.	.
3	6B	5	3	0.02	.	.	.	.	.	.	.	S	.	.	N	.	Q	T	T	.	.	.	.	.	.	T	.	K	.	.	.	N	.	.	.	K	.	.	.	.
4	6B	0	2	0.0237	.	.	.	.	.	.	.	S	.	.	N	.	Q	T	T	.	.	.	.	.	.	T	.	K	.	.	.	N	.	.	I	S	.	.	.	.
5	6B	0	1	0.0219	.	.	.	.	.	.	.	S	.	.	N	.	Q	T	T	.	.	.	.	.	.	T	.	K	Y	.	.	N	.	.	.	K	.	.	.	.
6	6B	19	0	0.0219	.	.	.	.	.	.	.	S	.	.	N	.	Q	T	T	.	.	.	.	.	.	T	.	K	.	.	.	N	.	.	.	K	.	.	.	.
7	6B	1	0	0.0237	.	.	.	.	.	.	.	S	.	.	N	.	Q	T	T	.	.	.	.	.	.	T	.	K	.	.	I	N	M	.	.	K	.	.	.	.
8	6B	1	0	0.0219	.	.	.	.	.	.	.	S	.	.	N	.	Q	T	T	.	.	.	.	T	A	K	.	.	.	N	.	.	.	K	.	.	.	.
9	6B	1	0	0.0237	.	.	.	.	.	.	.	S	.	.	N	.	Q	T	T	N	.	.	.	.	.	T	.	K	.	.	.	N	.	.	.	K	.	.	.	.
10	6B	1	0	0.0273	.	.	.	.	.	.	.	S	.	.	N	.	R	T	T	.	.	.	.	.	.	T	.	K	.	.	.	N	.	.	.	K	.	.	.	.
11	6B	0	1	0.02	.	.	.	.	.	.	.	S	.	.	N	I	Q	T	T	.	.	.	.	.	.	T	.	K	.	.	.	N	.	.	.	K	.	.	.	.
12	6B	1	0	0.0237	.	.	.	5	I	.	.	S	.	.	N	.	Q	T	T	.	.	.	.	.	.	T	.	K	.	.	.	N	.	.	.	K	.	.	.	.
13	6B	1	0	0.0237	.	.	.	A	.	.	.	S	.	.	N	.	Q	T	T	.	.	.	.	.	.	T	.	K	.	.	.	N	.	.	.	K	.	.	.	.
14	6B	0	1	0.0219	G	K	N	.	.	.	N	.	.	S	.	P	.	Q	T	T	.	.	.	.	.	.	T	.	K	.	.	.	N	.	.	.	K	.	.	.	.
15	6B	1	0	0.0291	.	.	.	.	.	A	S	.	S	.	.	Q	T	T	.	D	.	L	H	T	.	.	.	.	N	.	.	.	K	.	.	.	.	

Fig. 2. Amino acid substitutions in the antigenic sites (a) and the full-length HA gene (b) for influenza A(H1N1)pdm09. Sites are numbered according the previously published sequences for the vaccine and reference strains. Dots indicate amino acids identical to the reference sequence. Abbreviations: N, number of viruses collected; StP, St. Petersburg; pw/dist, pairwise distance; V(n/T), number of vaccinated patients in the group/total number of patients in the group; VLC, Valencia. *Pairwise amino acid distances between vaccine and each patient's virus were calculated by using the p-distance model implemented in MEGA6.
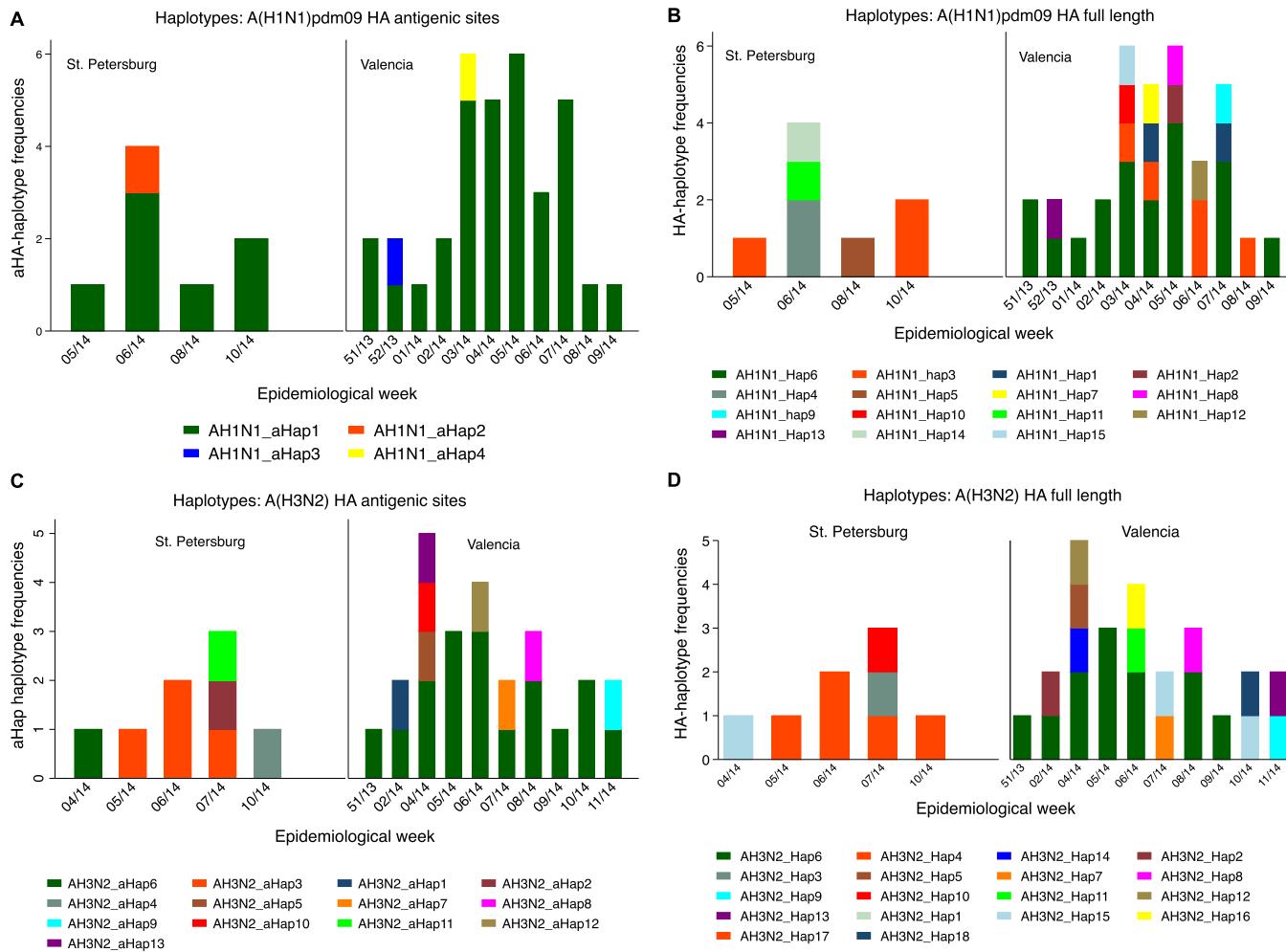
Fig. 3. Distribution of haplotypes for antigenic sites and full-length HA by epidemiological week for A(H1N1) pdm09 and A(H3N2) viruses from Valencia and St. Petersburg during the 2013/14 season.

concurrently and did not displace one another over the course of the season (Fig. 3c).

Additional amino acid changes at non-antigenic sites (n = 12; two of them at frequencies of 50% or higher) divided antigenic site aHap6 into several sub-haplotypes (Fig. 4b). Only one of the sub-

haplotypes, Hap6, (late December, 2013—early March, 2014) was found at a high frequency (50%) in Valencia (Fig. 3d), bearing all of the substitutions of aHap6 plus the two common mutations V529I (66%) and V347M (94%) in the HA2 region. Singleton mutations scattered across HA1 and HA2 non-antigenic sites accounted for

Fig. 4. Amino acid substitutions in the antigenic sites (a) and the full-length HA gene (b) for influenza A(H3N2). Sites are numbered according the previously published sequences for the vaccine and reference strains. Dots indicate amino acids identical to the reference sequence. Abbreviations: N, number of viruses collected; pw/dist, pairwise distance; StP, St. Petersburg; V(n/T), number of vaccinated patients in the group/total number of patients in the group; VLC, Valencia. *Pairwise amino acid distances between vaccine and each patient's virus were calculated by using the p-distance model implemented in MEGA6.

most of the additional variability. The peak of amino acid changes occurred in late January, 2014 in Valencia and mid-February, 2014 in St. Petersburg. During these peaks, only four full-length HA haplotypes co-circulated in Valencia and three co-circulated in St. Petersburg (Fig. 3d).

4.4. Pairwise amino acid distances between circulating viruses and vaccine strains

The median pairwise amino acid distances between the vaccine and A(H1N1)pdm09 circulating viruses was 0.06 (5th to 95th percentile range, 0.06–0.08) for antigenic sites and 0.022 [0.020–0.024] for all HA sites. Pairwise distances from the corresponding vaccine strain were lower for A(H3N2) isolates (0.023 [0.023–0.038] for antigenic sites and 0.018 [0.014–0.021] for all HA sites). Amino acid distances between vaccine and circulating A(H1N1)pdm09 and A(H3N2) viruses tended to be lower for common haplotypes than for singleton haplotypes (Figs. 2b and 4b). For A(H3N2), median amino acid distances from vaccine strains at antigenic sites were significantly higher for the most recent clade 3C.2 viruses than for the predominant 3C.3 clade viruses (median 0.034 [0.030–0.038] vs. 0.026 [0.023–0.038]; P=0.0423 by Kruskal–Wallis test). By location, median pairwise amino acid distances at antigenic sites between A(H3N2) and the vaccine strain were significantly lower for viruses collected from Valencia than for those collected from St. Petersburg (0.023 [0.023–0.030] vs. 0.038 [0.023–0.038]; P=0.0003 by Kruskal–Wallis test).

Patients were considered vaccinated for influenza if the vaccine had been administered ≥14 days before the onset of symptoms. In St. Petersburg, none of the patients had been vaccinated. In Valencia, 23% (8/34) of patients with A(H1N1)pdm09 infection and 46% (12/26) with A(H3N2) infection had received the corresponding season's influenza vaccine. Most vaccinated individuals (7/8) carried either the A(H1N1) antigenic haplotype aHap1 or the A(H3N2) antigenic haplotype aHap6, both also common in unvaccinated individuals. The relationship between genetic characteristics of the viral isolates and patient age or vaccination status could not be assessed because of insufficient sample sizes.

5. Discussion

The current analysis revealed some local genetic differences in influenza A viruses from Valencia and St. Petersburg. The most frequent haplotypes (combinations of amino acid substitutions) at antigenic sites were common at both locations. However, analysis of the entire HA gene region revealed population-specific haplotypes across the full length of HA, suggesting that variation across geographical regions may be shared within but not outside of antigenic epitopes. We deduced that intra-seasonal haplotypes co-circulated and did not displace one another during the 2013/14 epidemic season, and that the greatest amino acid sequence diversity occurred during the peak of influenza activity.

Nucleotide variation was lower at antigenic sites of A(H1N1)pdm09 viruses than in other HA regions, suggesting antigenic stability. Conversely, in A(H3N2) viruses, nucleotide variation was highest at HA antigenic sites. As expected, irrespective of the viral subtype, nucleotide variation was 10-fold lower at non-synonymous than at synonymous sites. These findings are in line with the known high rates of adaptive evolution in the HA1 domain and structurally mediated purifying selection at other sites in influenza HA genes [22]. The antigenically relevant amino acids varied more in influenza A(H3N2) isolates than H1N1pdm09 isolates, probably reflecting greater immunological pressure due to longer circulation and vaccine exposure.

We found a clear temporal pattern with co-circulation of phylogenetic groups and subgroups. For A(H3N2) viruses, we identified 94% of the isolates clustering with the previous season's dominant clade 3C.3. In contrast, for A(H1N1)pdm09, all isolates clustered with clade 6B, which was newly dominant in 2013/14. These results agree with findings in Canada [20]. We also report here the emergence of clade 3C.2 in A(H3N2) viruses at both study sites. Pairwise amino acid distances between antigenic sites of A(H3N2) circulating and vaccine viruses were greater for viruses belonging to this most recent clade 3C.2 than for viruses belonging to the predominant 3C.3 clade, suggesting that antigenic drift was already ongoing towards the end of the 2013/14 season.

Our study has some limitations that preclude exploring the differential variability of influenza viruses between the two study sites in full. Most importantly, sample size, especially for the Russian influenza A virus isolates, was too small to capture the full extent of viral genetic diversity, precluding reliable statements. Also, viral phylogenies were based only on the HA gene and were constructed from relatively few sequences.

Identifying influenza vaccine strains that will be antigenically matched with circulating strains the next season remains a challenge [9]. HA inhibition-based antigenic distances are currently used as a proxy for vaccine-virus relatedness [23] but their utility for predicting vaccine protection is limited [9]. Pairwise genetic distances between vaccine and circulating strains have been proposed as a proxy for antigenic similarity between vaccine and circulating viruses [24,25], but we are still a long way from correlating influenza HA genetic changes with vaccine performance, and serologic investigations (hemagglutination inhibition or neutralization tests) are needed to complement the genetic data. However, generating and reporting viral genetic data following a systematic approach will allow for additional sensitive and specific measurements to complement current epidemiological and functional data [20,26], to better correlate vaccine efficacy with specific mutation patterns in circulating influenza viruses [9,27], and ultimately improve selection of candidate vaccine strains.

Conflicts of interest

The authors report no conflicts of interest.

Author's contribution

M.P., F.X.L., M.G., A.S., and J.P. contributed in conception and design of the study. M.P., F.X.L., A.K., A.N., N.K., and K.S. participated in the acquisition and analysis of data. All authors contributed to interpretation of the data, participated in revising the article critically for important intellectual content, and approved the final version submitted to the journal. A.N. and J.P. wrote initial drafts of the article.

The authors would like to thank the patients and their families, as well as the staff of the participating study hospitals.

Scientific editing was provided by Dr. Phillip Leventhal (4 Clinics, Paris, France) and paid for by Sanofi Pasteur.

Ethical approval

The study protocol was approved by the institutional review board of both participating sites (Comité Ético de la Dirección General de Salud Pública y Centro Superior de Investigación en Salud Pública [CEIC-DGSP-CSISP] and the Ethics Committee of the Research Institute of Influenza, St. Petersburg).

Acknowledgments

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jcv.2016.09.006.

References

[1] S.E. Hensley, S.R. Das, A.L. Bailey, L.M. Schmidt, H.D. Hickman, A. Jayaraman, et al., Hemagglutinin receptor binding avidity drives influenza A virus antigenic drift, Science 326 (2009) 734–736.

[2] M. Schmolke, A. Garcia-Sastre, Evasion of innate and adaptive immune responses by influenza A virus, Cell. Microbiol. 12 (2010) 873–880.

[3] D.C. Wiley, J.J. Skehel, The structure and function of the hemagglutinin membrane glycoprotein of influenza virus, Annu. Rev. Biochem 56 (1987) 365–394.

[4] A.J. Caton, G.G. Brownlee, J.W. Yewdell, W. Gerhard, The antigenic structure of the influenza virus A/PR/8/34 hemagglutinin (H1 subtype), Cell 31 (1982) 417–427.

[5] N.A. Molinari, I.R. Ortega-Sanchez, M.L. Messonnier, W.W. Thompson, P.M. Wortley, E. Weintraub, et al., The annual impact of seasonal influenza in the US: measuring disease burden and costs, Vaccine 25 (2007) 5086–5096.

[6] H. Jin, H. Zhou, H. Liu, W. Chan, L. Adhikary, K. Mahmood, et al., Two residues in the hemagglutinin of A/Fujian/411/02-like influenza viruses are responsible for antigenic drift from A/Panama/2007/99, Virology 336 (2005) 113–119.

[7] Z. Chen, H. Zhou, H. Jin, The impact of key amino acid substitutions in the hemagglutinin of influenza A (H3N2) viruses on vaccine production and antibody response, Vaccine 28 (2010) 4079–4085.

[8] B.F. Koel, R. Mogling, S. Chutinimitkul, P.L. Fraaij, D.F. Burke, S. van der Vliet, et al., Identification of amino acid substitutions supporting antigenic change of influenza A(H1N1)pdm09 viruses, J. Virol. 89 (2015) 3763–3775.

[9] S.E. Hensley, Challenges of selecting seasonal influenza vaccine strains for humans with diverse pre-exposure histories, Curr. Opin. Virol. 8 (2014) 85–89.

[10] J. Puig-Barberà, A. Tormos, S. Trushakova, A. Sominina, M. Pisareva, M.A. Ciblak, et al., The Global Influenza Hospital Surveillance Network (GIHSN): A new platform to describe the epidemiology of severe influenza, Influenza Other Respir. Viruses 9 (6) (2015) 277–286.

[11] WHO Information for Molecular Diagnosis of Influenza Virus in Humans—Update August 2011, World Health Organization Geneva, 2011.

[12] J. He, M.E. Bose, E.T. Beck, J. Fan, S. Tiwari, J. Metallo, et al., Rapid multiplex reverse transcription-PCR typing of influenza A and B virus, and subtyping of influenza A virus into H1 2, 3, 5, 7, 9, N1 (human), N1 (animal), N2, and N7, including typing of novel swine origin influenza A (H1N1) virus, during the 2009 outbreak in Milwaukee, Wisconsin, J. Clin. Microbiol. 47 (2009) 2772–2778.

[13] K. Suwannakarn, S. Payungporn, T. Chieochansin, R. Samransamruajkit, A. Amonsin, T. Songserm, et al., Typing (A/B) and subtyping (H1/H3/H5) of influenza A viruses by multiplex real-time RT-PCR assays, J. Virol. Methods 152 (2008) 25–31.

[14] T.A. Hall, BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT, Nucleic Acids Symp. Ser. 41 (1999) 95–98.

[15] D. Posada, K.A. Crandall, MODELTEST: testing the model of DNA substitution, Bioinformatics 14 (1998) 817–818.

[16] K. Tamura, D. Peterson, N. Peterson, G. Stecher, M. Nei, S. Kumar, MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods, Mol. Biol. Evol. 28 (2011) 2731–2739.

[17] World Health Organization, Recommended composition of influenza virus vaccines for use in the 2013–2014 northern hemisphere influenza season, Wkly. Epidemiol. Rec. 88 (2013) 101–114.

[18] M. Nei, Molecular Evolutionary Genetics, Columbia University Press, New York, 1987.

[19] J. Rozas, R. Rozas, DnaSP version 3: an integrated program for molecular population genetics and molecular evolution analysis, Bioinformatics 15 (1999) 174–175.

[20] D.M. Skowronski, C. Chambers, S. Sabaiduc, G. De Serres, A.L. Winter, J.A. Dickinson, et al., Integrated sentinel surveillance linking genetic, antigenic, and epidemiologic monitoring of influenza vaccine-virus relatedness and effectiveness during the 2013–2014 influenza season, J. Infect. Dis. 212 (2015) 726–739.

[21] K. Tamura, G. Stecher, D. Peterson, A. Filipski, S. Kumar, MEGA6: molecular evolutionary genetics analysis version 6.0, Mol. Biol. Evol. 30 (2013) 2725–2729.

[22] S. Bhatt, E.C. Holmes, O.G. Pybus, The genomic rate of molecular adaptation of the human influenza A virus, Mol. Biol. Evol. 28 (2011) 2443–2451.

[23] J.M. Katz, K. Hancock, X. Xu, Serologic assays for influenza surveillance, diagnosis and vaccine evaluation, Expert Rev. Anti Infect. Ther. 9 (2011) 669–683.

[24] V. Gupta, D.J. Earl, M.W. Deem, Quantifying influenza vaccine efficacy and antigenic distance, Vaccine 24 (2006) 3881–3888.

[25] K. Pan, K.C. Subieta, M.W. Deem, 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 (2011) 291–299.

[26] B.F. Koel, R. Mögling, S. Chutinimitkul, P.L. Fraaij, D.F. Burke, S. van der Vliet, et al., Identification of amino acid substitutions supporting antigenic change of A(H1N1)pdm09 viruses, J. Virol. 89 (2015) 3763–3775.

[27] M. Luksza, M. Lassig, A predictive fitness model for influenza, Nature 507 (2014) 57–61.