Virus Research 259 (2019) 38–45

Contents lists available at ScienceDirect

Virus Research

journal homepage: www.elsevier.com/locate/virusres

Molecular epidemiology of the hemagglutinin gene of prevalent influenza virus A/H1N1/pdm09 among patient in Iran

Atefeh Mohebbi[a], Fatemeh Fotouhi[b,*], Abbas Jamali[b], Ramin Yaghobi[c], Behrokh Farahmand[b], Reza Mohebbi[d]

[a] Department of Microbiology, College of Science Agriculture and Modern Technology, Shiraz Branch, Islamic Azad University, Shiraz, Iran
[b] Department of Influenza and other Respiratory Viruses, Pasteur Institute of Iran, Tehran, Iran
[c] Shiraz Transplant Research Center, Shiraz University of Medical Sciences, Shiraz, Iran
[d] Gastroenterology and Liver Diseases Research Center, Research Institute for Gastroenterology and Liver Diseases, Shahid Beheshti University of Medical Sciences, Tehran, Iran

ARTICLE INFO

Keywords:
Influenza virus
Genetic characterization
Amino acid substitution
Phylogenetic analysis
pdm09 strain

ABSTRACT

In 2015, the influenza virus A/H1N1/pdm09 strain outbreak became prevalent throughout the different provinces of Iran. There are relatively limited complete genetic sequences available for this virus from Asian countries. Diagnosis and virological surveillance of influenza is essential for detecting novel genetic variants causing epidemic potential. This study describes the genetic properties of HA genome of influenza A/H1N1 pdm09 viruses circulating in Iran during the 2015/2016 season. In order to investigate the genetic pattern of influenza A/H1N1 pdm09, a total of 1758 nasopharyngeal swabs were screened by real-time RT-PCR. Of those, 510 cases were found to be positive for A/H1N1/pdm09 virus. Evolution of the approximately 100 positive specimens with high virus load was conducted via genomic phylogeny. Phylogenetic analysis of the HA genes of the A/H1N1pdm09 viruses revealed the circulation of clade 6B1, characterized by amino acid substitutions S84N, S162N and I216T, where position 162 became glycosylated. The N-glycosylation of HA protein is post or co-translational modification that affect the evolution of influenza viruses. For influenza A(H1N1) pdm09 viruses, we found more mutations in the antigenic sites than in the stem region. The results of this study confirmed the necessity of constant regular antigenic and molecular surveillance of circulating seasonal influenza viruses.

1. Introduction

Influenza viruses cause yearly repetition of contagious respiratory infections in animals and humans which result in severe illness and even a fetal end. The World Health Organization (WHO) estimates that approximately 5–10% of adults and 20–30% of children are infected by influenza annually (WHO, 2017a,b). Of those, 3–5 million infected individuals experience severe illness resulting in approximately 290,000–650,000 deaths (WHO, 2018a,b,c). Inactivated vaccines have been used since1950 decades to confront influenza threat. It is necessary to replace viruses contained in influenza vaccines periodically to have an effective vaccine, due to the constant evolving nature of influenza viruses (WHO, 2017a,b).

Influenza A virus (IAV) is a negative-sense single-stranded RNA virus which belongs to Orthomyxoviridae. The two surface glycoproteins, hemagglutinin (HA) and neuraminidase (NA) carry the major antigenic determinants of the virus and are the primary targets of the humoral immune response in humans (Te Velthuis and Fodor, 2016; Wohlbold and Krammer, 2014). From year to year, gradual mutations accumulate in the HA gene that produce immunologically distinct virus strains through a process known as antigenic drift (Yuan and Koelle, 2013). HA protein is coded by RNA segment 4, and forms HA0 as homotrimer precursor protein. Each monomer has a molecular mass of approximately 60 kDa and contains 549 amino acid residues and 17 amino acids as signal peptide. Translation and N-linked glycosylation of precursor hemagglutinin occurs in endoplasmic reticulum and transported through the Golgi complex to the plasma membrane where cleaves to HA1 (327 amino acids) and HA2 (222 amino acids), which are still linked by a disulphide bond. Hemagglutinin is the major surface glycoprotein on the envelope of influenza and mediates attachment and

* Corresponding author at: No: 69, Pasteur Ave, Influenza and other Respiratory Viruses Department, Pasteur Institute of Iran, Tehran, 13164, Iran.
E-mail addresses: mohebi.atefe@gmail.com (A. Mohebbi), fotouhi@pasteur.ac.ir (F. Fotouhi), jamali@pasteur.ac.ir (A. Jamali), yaghoobir@sums.ac.ir (R. Yaghobi), b_farahmand@pasteur.ac.ir (B. Farahmand), srmohebbi@gmail.com (R. Mohebbi).

https://doi.org/10.1016/j.virusres.2018.10.001
Received 8 August 2018; Received in revised form 1 October 2018; Accepted 2 October 2018
Available online 15 October 2018
0168-1702/

entry into the host through interactions with sialylated cell receptors (Sriwilaijaroen and Suzuki, 2012; Byrd-Leotis et al., 2017).

Each monomer of the HA, consists of a globular head domain and a stem/stalk domain. The globular domain comprises of HA1 subunit only, whereas the stem domain contains HA2 and small residue of HA1 subunit (Mallajosyula et al., 2014) The receptor-binding site and five distinct antigenic sites labeled Sa, Sb, Ca1, Ca2 and Cb, locate on the globular head (Yasugi et al., 2013). The Neuraminidase (NA), or the receptor-destroying protein, is another surface glycoprotein that cleaves the sialic acid receptors from cellular membrane to allow for release of nascent viral particles. Neuraminidase is a major target for some antivirals such as oseltamivir which inhibits its enzymatic activity (Sautto et al., 2018). The HA and NA genes are particularly variable in their sequence, and at least 16 HA and 9 NA subtypes are perpetuated in the wild waterfowl natural reservoir (Fouchier et al., 2005), however, only three subtypes H1N1, H2N2, and H3N2 have adapted successfully to infect and transmit efficiently among humans (Mair et al., 2014).

Currently circulating subtypes in humans are A/H1N1 and A/H3N2 influenza viruses. The A/H1N1 is also written as A/H1N1/pdm09, as it caused the pandemic flu in 2009 and subsequently replaced the seasonal influenza A/H1N1 virus which had circulated prior to 2009. Seasonal epidemics are the result of antigenic drift, that caused antigenic variation via point mutations in two key viral genes HA and NA. Only influenza type A viruses are known to have caused pandemics (WHO, 2018a,b,c; Sleman, 2017).

HA is the primary antigen for the host immune response located on the viral surface and subjected to high selective pressure due to innate and adaptive host immunity. HA acquires mutations to bypass the host defenses. These mutations predominantly occur in the antigenic sites on HA that are proximal to the receptor binding site which is the primary target for neutralizing antibodies or in the regions that increase protein stability or changes glycosylation pattern (Hensley et al., 2009). The N-linked glycosylation sites have also undergone changes over the years such as addition, loss or conversion, which might affect the antigenic properties or biological function of hemagglutinin.

Although there are commercial vaccines todays, the occurrence of mutations causing changes in the structure of antigenic sites might influence influenza vaccines efficacy.

Molecular and phylogenetic analyses of the viral genome are necessary to determine the antigenic and pathogenic characteristics of influenza viruses that cause severe outcomes of the disease (Kobayashi and Suzuki, 2012; Arellano-Llamas et al., 2017).

The National Collaborating Laboratory of Influenza, Pasteur Institute of Iran, investigates the frequency of influenza type/subtype of circulating viruses in three provinces, annually. We received a huge number of respiratory samples confirmed with A/H1N1 viruses in 2015–2016 flu season. In this study, possible changes in the genetic structure of HA of influenza (A/H1N1) viruses isolated from clinical samples during 2015–2016 seasonal flu that have led to acute illnesses are investigated. This study provides a better understanding of currently circulating flu viruses in the country compared to 2009 strain which is critical in terms of health care, vaccination and also efficient treatment.

2. Materials and methods

2.1. Study population and specimen collection

During the 2015–2016 influenza season 1758 nasopharyngeal swab samples received from the inpatients with severe acute respiratory illness from 3 provinces of Iran (Markazi, Semnan and Zanjan) were referred to the National Collaborating Laboratory of Influenza, Pasteur Institute of Iran. The sample collection methodology was approved by the research ethics committee of this Institute.

2.2. Viral RNA extraction and real time RT-PCR

RNA extraction from clinical samples was performed using the High Pure Viral RNA Kit (Roche Molecular Systems, Inc.) according to the manufacturer's instructions. The Real Time RT-PCR technique was used as a rapid and sensitive method for the detection and typing/subtyping of influenza viruses with the use of SuperScript III Platinum® One-Step qRT-PCR System (Invitrogen, USA) according to WHO guide line (WHO, 2009). Those samples that were positive for influenza A virus were subsequently screened for A/H1N1/pdm09 and A /H3N2.

2.3. Virus isolation and propagation

Samples confirmed positive for A/H1N1 pdm09 were grown in Madin-Darby Canine Kidney (MDCK) cell cultures for two or three passages (Eisfeld et al., 2014); Cells were cultured in Dulbecco modified Eagle medium (DMEM) containing 10% fetal bovine serum (FBS) and 1% penicillin-streptomycin (Gibco, Invitrogen). Cultures were incubated at 35 °C in a 5% CO_2 atmosphere and observed daily for 3 days for evidence of cytopathology (Abdoli et al., 2013).

2.4. Gene amplification and sequencing

HA gene of influenza A/H1N1 viruses were amplified by RT-PCR and sequenced. Viral RNA was extracted from propagated viruses and transcribed into cDNA with using the universal primer (U12: AGCGA AAGCAGG) by ThermoScript RT-PCR System for First-Strand cDNA Synthesis (Hoffmann et al., 2001).

Full HA was amplified as two overlapping fragment using designed primer as follows. PCR was performed by Platinum® Taq DNA Polymerase High Fidelity kit (Invitrogen, UK) using specific primers as mentioned in Table 1 at a final concentration of 0.5 µM. Primer design carried out by primer blast in NCBI. The PCR program was optimized in different condition (temperature and cycling) to obtain a strong and clear product without any nonspecific result. Finally, the thermal cycle was programmed for amplification as follow: incubation at 94 °C for 2 min, and then 35 cycles of denaturation at 95 °C for 30 s, 43 °C for 1 min for hybridization, 68 °C for 30 s. This was followed by a final elongation step at 68 °C for 10 min. PCR products were gel purified and subjected for sequencing using an ABI Sequence Genetic Analyzer (Applied Biosystems, Foster City, CA) at Sequence Laboratories of First Base Company, Malaysia.

2.5. Multiple-sequence alignments and phylogenetic analyses

The HA sequences of isolated A/H1N1 were aligned using Bio Edit software version 7.0 (https://www.bioedit.com/) and edited using Mega 6 (MEGA, version 6.0; http://www.megasoftware.net/). All 100 sequences of the viruses isolated in the present study were submitted in GenBank. Amino acid sequences corresponding to complete HA from A/California/07/2009(H1N1) (FJ966974) and other reference sequences

Table 1
Sequences and positions of the HA specific primers used to amplify two overlapping fragments of A/H1N1pdm09.

Influenza subtype	primer sequence (5'-3')	Position (nt)	PCR product (bp)
A/H1N1	F1: GAA GGC AAT ACT AGT AGT TCT R1: AGG CTG GTG TTT ATA GCA	3–25 920–939	937
A/H1N1	F2: AAA GTG AGG GAT CAA GAA G R2: TAC ACT GTA GAG ACC CA	706–725 1687–1704	1000

were downloaded from the NCBI Influenza Resource Database (http://www.ncbi.nlm.nih.gov/genomes/FLU) and the Global Initiative on Sharing All Influenza Data (GISAID; http://platform.gisaid.org/) databases (Korsun et al., 2017; Liechti et al., 2010). Phylogeny tree for HA amino acid was constructed using the maximum likelihood method inferred on the basis of the best fit amino acid substitution model for HA genes within MEGA software. The reliability measure of tree topology was assessed by bootstrap analysis with 1000 replications. Clades were analyzed based on the clustering patterns in the HA phylogeny (Korsun et al., 2017; Tamura et al., 2007; Jiménez-Alberto et al., 2013).

2.6. Prediction of potential N-glycosylation sites and homology modeling on HA

Full-length amino acid sequences of HA from patient and refseq from NCBI were used to predict N-glycosylation sites by the NetNGlyc 1.0. This program is available online (http://www.cbs.dtu.dk/services/NetNGlyc/). The NetNGlyc 1.0 analyzes the sequence context of Asn-Xaa-Ser/Thr sequons (Gupta and Brunak, 2002). The number of glycosylation sites was obtained for a single monomer of HA and locations of the glycosylation sites were numbered according to the full- length HA sequence of A/California/04/2009.

The complete HA amino acid (includes 18-566aa) sequence of 2015–2016 A/H1N1/pdm09 isolate from the present study, was used for three-dimensional (3D) structure analysis. The 3D structure of the 2015–2016 A/H1N1/pdm09 HA protein was generated by SWISS-MODEL (http://swissmodel.expasy.org/). SWISS-MODEL is a fully automated protein structure homology-modelling server. The crystal structure of A/California/07/ 2009 HA (3LZG) were downloaded from PDB database (http://www.rcsb.org). Then glycan was added onto the potential N-glycosites of HA using with the Glyprot Server (http://www.glycosciences.de/modeling/glyprot/). The figures were created and rendered by PyMOL. (Bohne-Lang and von der Lieth, 2005; Sun et al., 2012).

3. Results

During the outbreak of 2015–2016 flu season, from Oct 2015 until May 2016, a total of 1836 specimens from patients hospitalized with acute respiratory illness in three provinces, were referred to the National Collaborating Laboratory of Influenza, Pasteur Institute of Iran.

The early influenza detection was reported at the end of October 2015 and peaked in November and December. Influenza viruses were detected in 566 (30.8%) patient samples of which, 504 cases (89%) were found to be positive for A/H1N1/pdm09 and 3 cases (0.5%) for H3N2 viruses. Within the January 2016, influenza A/ H1N1/pdm09 virus was characterized as moderate intensity and influenza type B appears at the first of February with a lower incidence rate (Fig. 1).

During the two consecutive flu seasons 2013–2014 and 2014–2015,

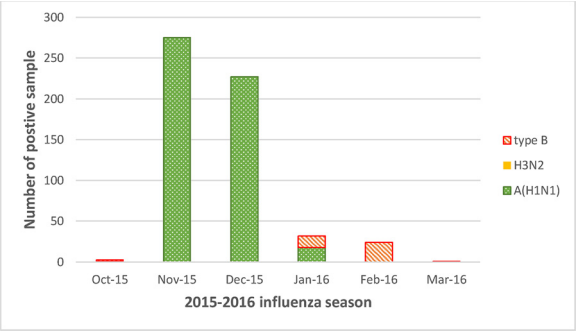

Fig. 1. Influenza virus detection, during the 2015–2016 influenza season.

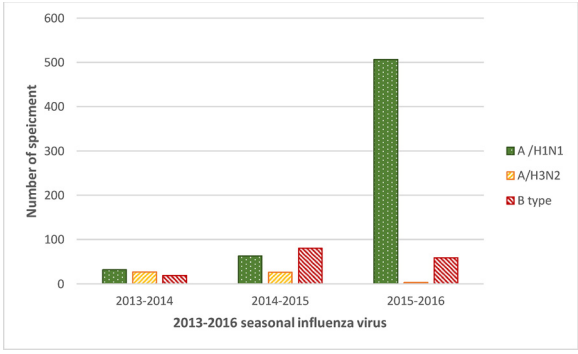

Fig. 2. Seasonal distribution of influenza virus (type/subtype) during the 2013–2016 flu seasons in our region.

totally 905 nasopharyngeal swab samples were received from the mentioned area, of those just 21.5% contained influenza viruses and only 18.4% were detected as A/H1N1/pdm09. Accordingly, a remarkable increase in A/H1N1/pdm09 was found in the circulating viruses within 2015–2016 season in our region (Fig. 2).

In order to deep understanding of the antigenic evolution of objective pandemic A/H1N1/pdm09 viruses, positive samples were propagated in MDCK cell culture for two or three passages. The HA gene was amplified as mention in M&M. Finally, 100 PCR products from each segment were submitted for sequencing. Sequencing results of the two segments assembled and refined to obtain full HA sequence. All sequences of the viruses isolated in the present study are available in GenBank from MG230334 to MG230433.

Phylogenetic tree was drawn to determine the genetic diversity and relationships among our isolates and viruses circulating in neighboring region in the same period. The HA phylogenetic tree was shown in Fig. 3. Based on the sequence analysis of HA genes isolated here, and compared to A/California/07/2009 reference virus, the HA molecules at the amino acid level had 97% identity and clustered in 6B.1 clade with the reference virus A/Slovenia/2903/2015.

HA gene analysis illustrated that they differed from A/California/07/2009 by 15 amino acid substitutions P83N, S84N, D97N, S162N, K163Q, S185T, S203T, I216T, A256T, K283E, I321V in HA1 and E47K, S124N, S215L and R219K in HA2. Generally, the amino acid substitutions were most frequently found in the globular domain, including substitutions near the antigenic sites or the receptor binding site. The other mutations were found in stalk domain including fusion peptide site in HA2. According to ECDC and WHO report, A/California/07/2009 located in clade 1 and amino acid changes in HA define different cluster from 2009 pandemic. Amino acid changes in HA1 substitutions consist of D97N, S185T and S203T in HA1 and E47K and S124N in HA2 defining clade 6. Amino acid substitutions K163Q, A256T and K283E in HA1 and E172 K in HA2 referred to group 6B. Worldwide new genetic subclusters of viruses within the 6B clade have emerged, with two being designated as subclades 6B.1 and 6B.2. Subclade 6B.1 is defined by HA1 amino acid substitution S84N, S162N and I216T and 6B.2 by HA1 amino acid substitution V152T and V173I (Table 2) (European Centre for Disease Prevention and Control, 2017; Santos et al., 2017).

Among the specified amino acid substitutions, five were located in antigenic sites: S203T in site Ca1; A141T in Ca2; S162N and K163Q in Sa and S185T in Sb. Only the S185T substitution fell within a domain defining the receptor binding site (RBS) and I321V in the non-antigenic sites of the HA1 subunit. Our result is comparable with others (Korsun et al., 2017).

In addition, three viruses (MG230334, MG230335 and MG230347) related to the early period carried N129D and V199D substitutions in HA which were clustered in 6B clade as shown in phylogenic tree (Fig. 4). Thirty-six viruses carried additional single amino acid changes in their HA sequence that was not in antigenic site. One of the isolated

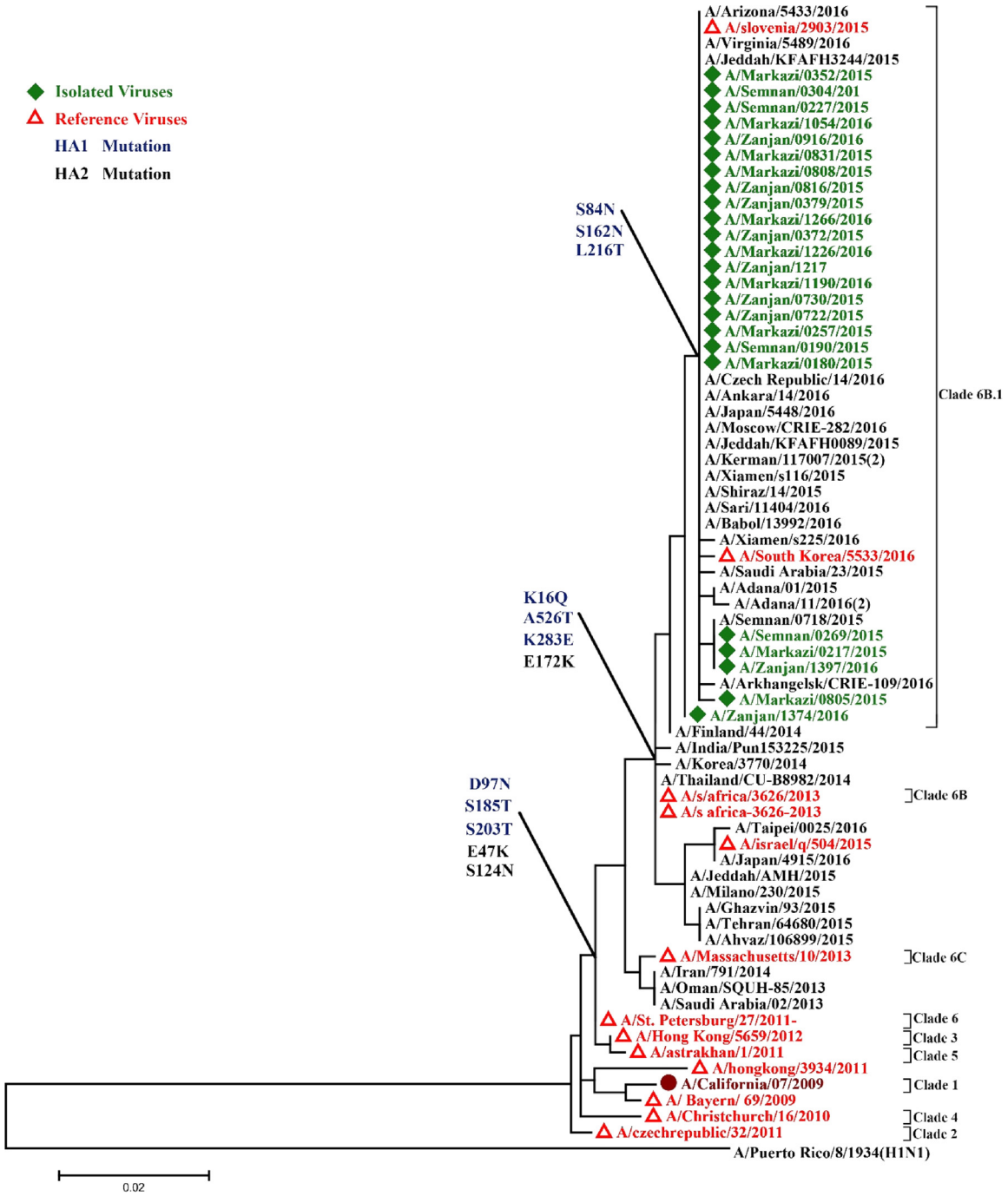

Fig. 3. The HA phylogenetic tree viruses isolated in this study, using Mega 6 (MEGA, version 6.0). The trees were constructed using the maximum likelihood (ML) method with bootstrap analysis of 1000 replicates. A/Puerto Rico/8/1934 was used as the outgroup for the tree analysis.

viruses, (MG230338) carries D 222G substitution that was associated with virus pathogenicity.

N-glycosylation motifs in the HA were predicted using the NetNGlyc 1.0 web Server to identify potential glycosylation sites. N-linked glycosylation defined as regions on glycoprotein that undergo post-translational modifications by adding some oligosaccharides through N-glycoside linkages to the Asn residue of the glycosylation motif Asn-X-Ser/Thr, where X is any amino acid except proline, N–X–S/T (sequon) (Tate et al., 2014; Kim et al., 2018).

The structure-based in silico prediction study further suggests that the studied samples had common glycosylation in the globular region and the stalk/ stem region of hemagglutinin as compared to the reference strain.

Our isolated sequences showed 6 common glycosylation sites at positions 11, 23, 87, 287, 481 and 540. In addition, positions 84, 97, 162 and 451 in these isolates were identified as glycosylation sites. Among them, position 162 was located in the Sa antigenic site, while positions 84 and 97 were not in the antigenic area of the globular head. S162N substitution, specific to subclade 6B.1 virus, generated a new potential n-glycosylation motif within the Sa antigenic site, which were found in 97% of our samples. This S162 N substitution have been found in other HA sequences reported from 2015 to 2016 (Korsun et al., 2017; European Centre for Disease Prevention and Control, 2016). Also, position 451 in the stalk domain which is a glycosylation region might be important in protein folding (Fig. 4).

Table 2
Amino acid substitutions on the HA of influenza A/H1N1/pdm09, compared to A/California/7/2009.

Clade	References Strain	GISAID accession No	83	84	97	162	163	185	203	216	256	283	321	47	124	172	219
clade 1	A/California/7/2009	EPI177294	P	S	D	S	K	S	S	I	A	K	I	E	S	E	R
clade 2	A/czechrepublic/32/2011	EPI319447	S	*	*	N	*	*	T	*	*	*	V	K	N	*	*
clade 3	A/hongkong/3934/2011	EPI326206	*	*	*	S	*	*	T	*	*	*	V	E	*	*	*
clade 4	A/Christchurch/16/2010	EPI278607	*	*	*	S	*	*	T	*	*	*	V	K	*	*	*
clade 5	A/astrakhan/1/2011	EPI1170231	*	*	N	S	*	*	T	V	*	*	V	K	*	*	*
clade 6	A/ST Petersbrg/27/2011	EPI319527	*	*	N	S	*	T	T	I	*	*	V	K	N	*	*
clade 6B	A/South Africa/3626/2013	EPI498431	*	*	N	S	Q	T	T	*	T	E	V	K	N	K	*
clade 6B.1	A/Michigan/45/2015	EPI1170231	*	N	N	N	Q	T	T	T	T	E	V	K	N	K	*
	Our strain		*	N	N	N	Q	T	T	T	T	E	V	K	N	K	k
Clade 6C	A/Massachusetts/10/2013	EPI464862	*	s	S	N	Q	T	T	I	A	E	V	K	N	K	*

The HA1 position numbered from 18 to 343 and HA2 position number from 345 to 566 in HA0. Amino acid abbreviations: A (Alanine), R (Arginine), N (Asparagine), D (Aspartic acid), E (Glutamic acid), Q (Glutamine), I (Isoleucine), K (Lysine), L (Leucine), P (Proline), S (Serine), T (Threonine), V (Valine).

4. Discussion

Influenza A viruses are known to cause severe acute respiratory tract infections and represent a significant public health threat (Klenk et al., 2002). The influenza A/H1N1 virus caused pandemic flu in 2009 and is still circulating during the winter season in many countries (Jhung et al., 2011). The prevalence and influenza activity is monitored annually in our country. In the period of this study, the distribution and frequency of influenza virus increased dramatically compared to the two consecutive flu season 2013–2014 and 2014–2015, from the mentioned area in Iran. The study season was characterized by a dominant spread of A/H1N1/pdm09 viruses accounting for 92% of the detected influenza viruses and by low circulation of A/H3N2 viruses (0.5%), and influenza B viruses (0.1%). Our data showed that prevalence of influenza virus in Iran during the 2015–2016 season peaked in November and December to similar proportions of circulating seasonal influenza viruses in most European countries. Stored data for the WHO European region showed that within type A, the A/H1N1/pdm09 subtype predominated (91%) over A/H3N2 viruses (9%) (Korsun et al., 2017).

An unexpected increase in the outbreak of A/H1N1 virus in 2015–2016 season was observed despite the same flu vaccine coverage over the recent years, probably due to declined vaccine effectiveness. It has been shown that the presence of several mutations in the antigenic sites of HA can lead to a rapid decrease in vaccine efficacy (Castelán-Vega et al., 2014).

In this study, we detected the mutations in HA sequence that might be associated with viral evolution and might influence the antigenicity of the virus. We found a clear variation in the circulating viruses in Iran in 2015–2016 season based on amino acid substitutions and in accordance with previous reports (Wedde et al., 2015; Mukherjee et al., 2016). Antigenic variations of the HA show the molecular epidemiology among circulating influenza strains, which can be used for improvement of influenza virus vaccine.

The HA protein is the major antigenic target and has undergone significant genetic changes. From 2009 until now, molecular analyses of HA gene sequences obtained from A/H1N1/pdm09 strains revealed that at least nine genetic groups/clades and several subgroups/subclades circulated in the world (European Centre for Disease Prevention and Control, 2016). The ML tree topology based on the sequence analysis of the HA genes showed a clear temporal structure with co-circulation of phylogenetic groups and subgroups.

It has been shown that several mutations occurred in different antigenic and non-antigenic sites of HA. Based on the phylogenetic tree plotted for the hemagglutinin gene, we detected circulation of 6B.1 subclade. In this subclade, amino acid substitutions of S84N, S162N and I216T in the globular head of hemagglutinin were observed compared to the previous year (WHO, 2017a,b).

Substitution S162N, specific to subclade 6B.1 virus, generated a new potential n-glycosylation motif within the Sa antigenic site. It was located in the globular head and found in 97% of our samples during 2015–2016 season. This substitution has not been reported in the previous flu season in this region (Moasser et al., 2017). Viruses isolated in these parts of Iran during the previous season, 2014–2015, clustered in subgroup 6C, represented by A/Massachusetts/ 10/2013 and characterized by amino acid substitution V234I and K283E in HA1 and E172K in HA2, as compared to the representative virus of group 6 (Korsun et al., 2017; European Centre for Disease Prevention and Control, 2016).

Five antigenic sites are located at the globular head of HA designated as Sa (residues 128–129, 156–160, 162–167) and Sb (residues 187–198) near the spike tip, Ca1 (residues 169–173, 206–208,

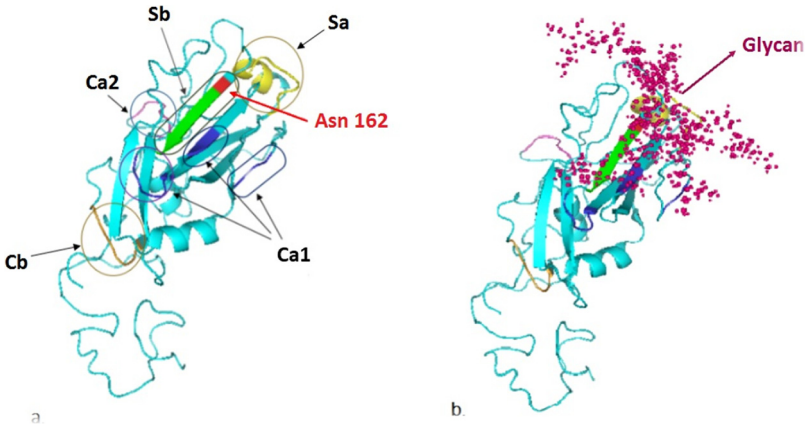

Fig. 4. Structural overviews of the glycans attached to Asn 162 and their shielding regions on HA of human seasonal influenza H1N1 viruses, a: The position of antigenic sites on HA1, b: Insertion of glycans attached to new glycosylation site (Asn162).

238–240) and Ca2 (residues 140–145, 224–225) between adjacent HA monomers, and Cb (residues 74–79), that is inside the vestigial esterase domain near the base of the globular head (Sriwilaijaroen and Suzuki, 2012; Mallajosyula et al., 2014).

Mutation in P83S in antigenic site Cb has been found in influenza A/H1N1/ pdm09 viruses from 2011 in clades 6 and 7 compared to A/California/07/2009(H1N1) (European Centre for Disease Prevention and Control, 2016; Potdar et al., 2010). This amino acid substitution in the HA1 globular head has been observed in all samples in this study. Residue change S84 N which was observed in strains from 2014 to 2015 during the pandemic period, were identified in our samples as well. Amino acid substitution D97N was circulated from 2011 worldwide. Four mutations (S162N, K163Q, S185T and S203T) were observed at the antigenic epitopes within the receptor-binding site (Arellano-Llamas et al., 2017; Korsun et al., 2017; Mukherjee et al., 2015). Substitution S162N & S203T shown in clade 2 was represented by A/czechrepublic/32/2011. The latter, S203T, still continued, but S162T disappeared until clade 6B.1 in 2015 by A/Michigan/45/2015 reference. Also Substitution K163Q was found in antigenic site Sa from 2013 that was defined in clade 6 with A/South Africa/3626/2013. The amino acid substitution of S185T that was represented by A/ST Petersburg/27/2011 in clade 6 was observed in all HA sequences analyzed in this study located in the190helix (184–195 region) of predicted antigenic site *Sb*. This substitution has side chain exposed to the solvent and does not interact with other amino acids (Castelán-Vega et al., 2014; Community Network of Reference Laboratories (CNRL) for Human Influenza in Europe, 2012).

It has been confirmed that HA1 D222G substitution was found more frequently in patients with severe disease (Chan et al., 2009). This substitution allows the receptor specificity shift from α2-6 sialic acid to mixed α2,3/α2,6-sialic acid and let virus reach the lower respiratory tract, causing an increase severity of the disease (Liu et al., 2012). One of the isolated viruses here (accession no: MG230338) showed D222G substitution (1%), although we had no more information about the patient destination.

Substitution E47K in HA2 (E374K numbering from HA1) has been observed in the most of the H1N1pdm strains isolated after the year 2010. The E47K results in some structural changes, causes acid stability of the HA trimers involving in membrane fusion, and potentially affects the viral antigenicity and adaptation (Castelán-Vega et al., 2014; Cotter et al., 2014). All viruses isolated in this epidemic wave contained this substitution.

The other mutations detected in HA2 subunit i.e. S124N and E172K might increase the HA stability. Mutation S124N made intrachain interactions with nearby residue R123 and E132. Also, mutation E172K located at the base of the HA, created contact with N169 residue, which have stabilized the bottom of the HA stem (Castelán-Vega et al., 2014; Cotter et al., 2014).

Some of the mutations described here, might have occurred during MDCK propagation. However, other researchers studied influenza diversity, have used samples propagated in MDCK cells as well (Korsun, N et al, 2017). Notably, our finding is in accordance with WHO reports that claimed viruses circulating during 2015–2016 season in our region, have been located in 6B.1 clade with amino acid substitutions which mentioned above (WHO 2017).

4.1. The N- glycosylation of HA protein

The N- glycosylation of HA protein is post or co-translational modification that affect the evolution of influenza viruses (Klenk et al., 2002). Glycosylation of HA is thought to be an important mechanism contributing to folding and stability of the protein, and, in some cases, effects on HA biological functions, including receptor binding activity and the HA0 cleavability which influence the virulence and antigenicity of the virus. Furthermore, the presence of glycan on antigenic site can stimulate virus to evasion of host immunity by filling or masking the antigenic regions of protein and escape from antibody-mediated neutralization, thereby promoting virus survival in the face of widespread vaccination and/or infection. Thus, it played an important role in the host adaptation of the viruses (Kim et al., 2018; Klenk et al., 2002).

The number and distribution of N-glycosylation sites over the viral proteome can therefore be computationally analyzed by NetNGlyc 1.0 Server. This method was based on the assumption that all potential N-glycosylation sites were occupied (Gupta and Brunak, 2002; Joshi and Gupta, 2015). In this study, a series of bioinformatics tools were used to maximize the reliability of this strategy. First, all of the sequons were found from 100 full-length amino acid sequences of isolated HA. Then, the structures of representative HA proteins from human influenza A/H1N1 viruses were modeled, including their different patterns of potential N-glycosylation sites. Finally, glycans were added in silico onto each glycosylation site using the Glyprot server to confirm that these potential glycosylation sites could be glycosylated, and the structures of the obtained glycoproteins were further used to determine the location of glycosylation sites.

Previous reports had showed that the 2009 H1N1 pandemic strain and the 1918 pandemic strain have similar glycosylation pattern (Sun et al., 2011). The human seasonal influenza H1N1 viruses before 2009 had more glycosylation sites on the head of HA than 1918 and 2009 pandemic viruses. There were highly conserved potential glycosites such as 11, 23, 87, 287, 481, 540 in all strains isolated, that are essential for biological function of HA, while some other glycosites such as 125, 129 and162 on the top of HA head, glycosites 84 and 97 on the side of HA head and glycosite 451 in stalk domain of HA appeared during seasonal influenza H1N1 viruses (Sun et al., 2012).

The general hypothesis suggests that the acquisition or migration of glycosylation on HA might change immune interaction between virus and pre-exiting antibodies and enable virus to escape from the host immunity (Kim et al., 2018, Sun et al., 2011).

After the 2009 pandemic, however, the top of HA head remained completely conserved until 2015–2016 seasonal flu, when position 162 became glycosylated (Korsun et al., 2017; Sun et al., 2012). Similarly, our results showed that all studied samples had glycosite at position 162 which might mask the antigenic site Sa.

It is interesting to note that the HA of swine H1N1 viruses has been glycosylated in position 162 in recent years. Glycosite 162, appearing on the top of the HA head in human influenza H1N1 viruses in 1933, was replaced by glycosite 160 in 1951 until 2009 (Sun et al., 2012). In 2009 pandemic, there was no glycosylation site on the top of HA head including glycosite 160, and in early period of 2015, glycosite 162 appeared. The glycans attached to glycosites 162 and 87 of HA, may function to stabilize the trimeric structures of HA. Glycosylation site at position Asn84 and Asn97 have been reported since 2015 and 2011, respectively. The latter, was defined as group 6, but none of the mentioned positions were at an antigenic point.

The other mutations in this study that can increase HA stability, located in HA2 subunit were S124N, and E172K (S468N, and E516K numbering from HA1). Mutation S124N made intrachain interactions with nearby residues R123 and E132. Also, mutation E172K located at the base of the HA created contact with N169 residue, which have stabilized the bottom of the HA stem. Stabilizing effect of mutations S124N and E172K was evaluated according to the free-energy analysis (Castelán-Vega et al., 2014). Glycosylation Asn124 (468), in the HA2 region, was observed in the majority of our samples (97%), which was absent in the reference strain of 2009, but identified as glycosylation region from 2011 (Kim et al., 2018). According to these results, we suggest that further study of HA N-glycosylation of the H1N1 virus may guide future pandemic planning and lead to a better understanding of mechanistic roles of glycosylation as a viral strategy for human adaptation and might be useful for improving pandemic and seasonal influenza vaccines.

The viruses in influenza vaccines have to be evaluated and updated regularly. Hemagglutinin of circulating influenza viruses continuously

undergo mutations in antigenic sites that enable the virus to escape from pre-existed immunity. When the vaccine strains do not match the circulating epidemic strains, vaccine effectiveness is considerably reduced. Therefore, WHO provides a guide to vaccine manufacturers for the development and production of vaccines for the next influenza period.

According to our result in phylogenetic tree for the HA genes, mutation in HA of circulating A/H1N1 pdm09 viruses showed that these isolates were in subclade 6B.1 represented by A/Michigan/45/2015. Based on the reports of severe disease associated with A/H1N1/pdm09 in some countries, evidence of vaccine failure in a number of cases and also the dominance of subclade 6B.1 virus, World Health Organization recommended including a subclade 6B.1 virus in the vaccine recommendation for 2017 influenza season. Therefore, A/Michigan/45/2015 was added to vaccine strains thereafter (European Centre for Disease Prevention and Control, 2016).

5. Conclusions

In conclusion, the results of our study give us a better understanding of the molecular and phylogenetic analysis of circulating influenza A (H1N1) pdm09 viruses in Iran, which is useful for improving seasonal and consequently pandemic influenza vaccines.

Moreover, we suggest that further study of HA N-glycosylation of the H1N1 virus may guide future pandemic planning and lead to a better understanding of mechanistic roles of glycosylation as a viral strategy for human adaptation.

Acknowledgement

We would like to thank Dr P. Mehrbod for kindly editing the manuscript. We are grateful to all co-workers in Pasteur Institute of Iran, department of Influenza and other Respiratory Viruses that support us in conducting this project. This work was funded by grant No 942 from Pasteur Institute of IRAN.

References

Abdoli, A., Soleimanjahi, H., Tavassoti Kheiri, M., Jamali, A., Jamaati, A., 2013. Determining influenza virus shedding at different time points in madin-darby canine kidney cell line. Cell J. (Yakhteh) 15 (2), 130–135.

Arellano-Llamas, R., Alfaro-Ruiz, L., Arriaga Canon, C., Imaz Rosshandler, I., Cruz-Lagunas, A., Zúñiga, J., Vázquez-Pérez, J.A., 1804. Molecular features of influenza A (H1N1) pdm09 prevalent in Mexico during winter seasons 2012-2014. PLOS ONE 12 (7), e019.

Bohne-Lang, A., von der Lieth, C.-W., 2005. GlyProt: in silico glycosylation of proteins. Nucleic Acids Res. 33 (Web Server issue), W214–W219.

Byrd-Leotis, L., Cummings, R.D., Steinhauer, D.A., 2017. The interplay between the host receptor and influenza virus hemagglutinin and neuraminidase. Int. J. Mol. Sci. 18 (7), 1541.

Castelán-Vega, J.A., Magaña-Hernández, A., Jiménez-Alberto, A., Ribas-Aparicio, R.M., 2014. The hemagglutinin of the influenza A(H1N1) pdm09 is mutating towards stability. Adv. Appl. Bioinf. Chem.:AABC 7, 37–44.

Chan, P.K.S., Lee, N., Joynt, G.M., Choi, K.W., Cheung, J.L.K., Yeung, A.C.M., Hui, D.C.S., 2009. 2011. Clinical and virological course of infection with haemagglutinin D222G mutant strain of pandemic influenza A (H1N1) virus. J. Clin. Virol. 50 (4), 320–324.

Community Network of Reference Laboratories (CNRL) for Human Influenza in Europe 2012.

Cotter, C.R., Jin, H., Chen, Z., 2014. A single amino acid in the stalk region of the H1N1pdm influenza virus ha protein affects viral fusion, stability and infectivity. PLOS Pathog. 10 (1), e1003831.

Eisfeld, A.J., Neumann, G., Kawaoka, Y., 2014. Influenza a virus isolation, culture and identification. Nature Protoc. 9 (11), 2663–2681.

European Centre for Disease Prevention and Control, 2017. Influenza virus characterisation, summary Europe. ECDC, Stockholm September 2017.

European Centre for Disease Prevention and Control, 2016. Influenza virus characterisation, summary Europe. ECDC, Stockholm September 2016.

Fouchier, R.A., Munster, V., Wallensten, A., Bestebroer, T.M., Herfst, S., Smith, D., Osterhaus, A.D., 2005. Characterization of a novel influenza A virus hemagglutinin subtype (H16) obtained from black-headed gulls. J. Virol. 79 (5), 2814–2822.

Gupta, R., Brunak, S., 2002. Prediction of glycosylation across the human proteome and the correlation to protein function. Pac. Symp. Biocomput. 310–322.

Hensley, S.E., Das, S.R., Bailey, A.L., Schmidt, L.M., Hickman, H.D., Jayaraman, A., Yewdell, J.W., 2009. Hemagglutinin receptor binding avidity drives influenza A virus antigenic drift. Science 326 (5953), 734–736.

Hoffmann, E., Stech, J., Guan, Y., Webster, R.G., Perez, D.R., 2001. Universal primer set for the full-length amplification of all influenza A virus. Arch. Virol. 146 (12), 2275–2289.

Jhung, M.A., Swerdlow, D., Olsen, S.J., Jernigan, D., Biggerstaff, M., Kamimoto, L., Finelli, L., 2009. Epidemiology of 2009 pandemic influenza A (H1N1) in the United States. Clin. Infect. Dis. 52 (suppl 1).

Jiménez-Alberto, A., Alvarado-Facundo, E., Ribas-Aparicio, R.M., Castelán-Vega, J.A., 2013. Analysis of adaptation mutants in the hemagglutinin of the influenza A(H1N1) pdm09 virus. PLOS ONE 8 (7), e70005.

Joshi, H.J., Gupta, R., 2015. Eukaryotic glycosylation: online methods for site prediction on protein sequences. Methods Mol. Biol. 1273, 127–137.

Kim, P., Jang, H.Y., Kwon, B.S., Lee, M.C., Han, G., Seong, L.B., 2018. Glycosylation of hemagglutinin and neuraminidase of influenza A virus as signature for ecological spillover and adaptation among influenza reservoirs. Viruses 10 (4), 183.

Klenk, H.D., Wagner, R., Heuer, D., Wolff, T., 2002. Importance of hemagglutinin glycosylation for the biological functions of influenza virus. Virus Res. 82 (1-2), 73–75.

Kobayashi, Y., Suzuki, Y., 2012. Evidence for n-glycan shielding of antigenic sites during evolution of human influenza A virus hemagglutinin. J. Virol. 86 (7), 3446–3451.

Korsun, N., Angelova, S., Gregory, V., Daniels, R., Georgieva, I., McCauley, J., 2017. Antigenic and genetic characterization of influenza viruses circulating in Bulgaria during the 2015/2016 season. Infect. Genet. Evol. 49, 241–250.

Liechti, R., Gleizes, A., Kuznetsov, D., Bougueleret, L., Le Mercier, P., Bairoch, A., Xenarios, I., 2010. OpenFluDB, a database for human and animal influenza virus. Database: J. Biol. Databases Curation 2010 baq004.

Liu, Y., Wang, J., Ji, J., Chang, S., Xue, C., Ma, J., Xie, Q., 2012. Phylogenetic diversity and genotypic complexity of H1N1 subtype swine influenza viruses isolated in Mainland China. Virol. J. 9 (1), 289.

Mair, C.M., Ludwig, K., Herrmann, A., Sieben, C., 2014. Receptor binding and pH stability - how influenza A virus hemagglutinin affects host-specific virus infection. Biochimica et Biophysica Acta (BBA) - Biomembranes 1838 (4), 1153–1168.

Mallajosyula, V.V.A., Citron, M., Ferrara, F., Lu, X., Callahan, C., Heidecker, G.J., Varadarajan, R., 2014. Influenza hemagglutinin stem-fragment immunogen elicits broadly neutralizing antibodies and confers heterologous protection. Proceedings of the National Academy of Sciences of the United States of America 111 (25), E2514–E2523.

Moasser, E., Behzadian, F., Moattari, A., Fotouhi, F., Rahimi, A., Zaraket, H., Hosseini, S.Y., 2017. Molecular characterization and phylogenetic analysis of human influenza A viruses isolated in Iran during the 2014-2015 season. Arch. Virol. 162 (7), 1975–1984.

Mukherjee, A., Nayak, M.K., Dutta, S., Panda, S., Satpathi, B.R., Chawla-Sarkar, M., 2016. Genetic characterization of circulating 2015 A(H1N1) pdm09 influenza viruses from Eastern India. PLOS ONE 11 (12), e0168464.

Potdar, V.A., Chadha, M.S., Jadhav, S.M., Mullick, J., Cherian, S.S., Mishra, A.C., 2010. Genetic characterization of the influenza a pandemic (H1N1) 2009 virus isolates from India. PLOS ONE 5 (3), e9693.

Santos, K.C., Silva, D.B., Sasaki, N.A., Benega, M.A., Garten, R., Paiva, T.M., 2017. Molecular epidemiology of influenza A(H1N1) PDM09 hemagglutinin gene circulating in Sao Paulo State, Brazil: 2016 anticipated influenza season. Rev. Inst. Med. Trop. Sao Paulo 59, e9.

Sautto, G.A., Kirchenbaum, G.A., Ross, T.M., 2018. Towards a universal influenza vaccine: different approaches for one goal. Virol. J. 15, 17.

Sleman, S., 2017. How Influenza a causes “Epidemics and Pandemics” among the Population. Novel Targets for Anti-Influenza Molecules Vol. 05.

Sriwilaijaroen, N., Suzuki, Y., 2012. Molecular basis of the structure and function of H1 hemagglutinin of influenza virus. Proceedings of the Japan Academy Series B 88 (6), 226–249.

Sun, S., Wang, Q., Zhao, F., Chen, W., Li, Z., 2011. Glycosylation site alteration in the evolution of influenza A (H1N1) viruses. PLOS ONE 6 (7), e22844.

Sun, S., Wang, Q., Zhao, F., Chen, W., Li, Z., 2012. Prediction of biological functions on glycosylation site migrations in human influenza H1N1 viruses. PLOS ONE 7 (2), e32119.

Tamura, K., Dudley, J., Nei, M., Kumar, S., 2007. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4. 0. Mol. Biol. Evol. 24 (8), 1596–1599.

Tate, M.D., Job, E.R., Deng, Y.-M., Gunalan, V., Maurer-Stroh, S., Reading, P.C., 2014. Playing hide and seek: how glycosylation of the influenza virus hemagglutinin can modulate the immune response to infection. Viruses 6 (3), 1294–1316.

Te Velthuis, A.J.W., Fodor, E., 2016. Influenza virus RNA polymerase: insights into the mechanisms of viral RNA synthesis. Nat. Rev. Microbiol. 14 (8), 479–493.

Wedde, M., Biere, B., Wolff, T., Schweiger, B., 2015. Evolution of the hemagglutinin expressed by human influenza A(H1N1) pdm09 and A(H3N2) viruses circulating between 2008-2009 and 2013-2014 in Germany. Int. J. Med. Microbiol. 305 (7), 762–775.

WHO - Influenza update – 316, based on data up to 13 May 2018 available online http://www.who.int/influenza/surveillance_monitoring/updates/latest_update_GIP_surveillance/en.

WHO - Seasonal influenza, Available online, http://www.who.int/biologicals/vaccines/influenza/en/. 6 November 2017 15:04 CET.

WHO Consultation and Information Meeting on the Composition of Influenza Virus Vaccines for Use in the 2018 Southern Hemisphere Influenza Season, WHO Consultation, 25 - 27 September 2017, Melbourne, Australia.

WHO, up to 650 000 people die of respiratory diseases linked to seasonal flu each year, Available online http://www.who.int/news-room/detail/14-12-2017-up-to-650-000-people-die-of-respiratory-diseases-linked-to-seasonal-flu-each-year. 27 May 2018.

WHO, 2009. CDC protocol of real-time RTPCR for influenza A (H1N1). WHO, Geneva. Available: http://www.who.int/csr/resources/publications/swineflu/

CDCrealtimeRTPCRprotocol_20090428.pdf.
WHO; Jun 2018, available online; http://www.who.int/news-room/fact-sheets/detail/influenza-(seasonal).
Wohlbold, T.J., Krammer, F., 2014. In the Shadow of hemagglutinin: a growing interest in influenza viral neuraminidase and its role as a vaccine antigen. Viruses 6 (6), 2465–2494.
Yasugi, M., Kubota-Koketsu, R., Yamashita, A., Kawashita, N., Du, A., Misaki, R., Ikuta, K., 2013. Emerging antigenic variants at the antigenic site sb in pandemic A(H1N1)2009 influenza virus in Japan detected by a human monoclonal antibody. PLOS ONE 8 (10), e77892.
Yuan, H.-Y., Koelle, K., 2013. The evolutionary dynamics of receptor binding avidity in influenza A: a mathematical model for a new antigenic drift hypothesis. Philos. Trans. R. Soc. B: Biol. Sci. 368 (1614), 20120204.