Predominance of influenza A(H3N2) viruses during the 2016/2017 season in Bulgaria

Neli Korsun,1,∗ Svetla Angelova,1 Ivelina Trifonova,1 Iren Tzotcheva,2 Sirma Mileva,2 Silvia Voleva,1 Irina Georgieva1 and Penka Perenovska2

Abstract

Purpose. Influenza viruses are characterised by high variability, which makes them able to cause annual epidemics. The aim of this study is to determine the antigenic and genetic characteristics of influenza viruses circulating in Bulgaria during the 2016/2017 season.

Methodology. The detection and typing/subtyping of influenza viruses were performed using real time RT-PCR. Results of antigenic characterisation, phylogenetic and amino acid sequence analyses of representative influenza strains are presented herein.

Results. The 2016/2017 season was characterised by an early start, an exclusive dominance of A(H3N2) viruses accounting for 93 % of total influenza virus detections, and a low circulation of A(H1N1)pdm09 (4.2 %) and type B (2.5 %) viruses. The analysed A(H3N2) viruses belonged to subclades 3C.2a (52 %) and 3C.2a1 (48 %); all studied A(H1N1)pdm09 and B/Victoria-lineage viruses belonged to subclades 6B.1 and 1A, respectively. The amino acid sequence analysis of 56 A(H3N2) isolates revealed the presence of substitutions in 18 positions in haemagglutinin (HA) as compared to the A/Hong Kong/4801/2014 vaccine virus, seven of which occurred in four antigenic sites, together with changes in 23 positions in neuraminidase (NA), and a number of substitutions in internal proteins PB2, PB1, PB1-F2, PA, NP and NS1. Despite the many amino acid substitutions, A(H3N2) viruses remained antigenically similar to the vaccine strain. Substitutions in HA and NA sequences of A(H1N1)pdm09 and B/Victoria-lineage strains were also identified, including in antigenic sites.

Conclusion. The results of this study confirm the genetic variability of circulating influenza viruses, particularly A(H3N2), and the need for continued antigenic and molecular surveillance.

INTRODUCTION

Influenza is a highly contagious viral infection in humans caused by three types of influenza viruses - A, B and C. The first two types cause seasonal flu epidemics, which are associated with high morbidity and significant mortality. Worldwide, these annual epidemics are estimated to result in about three to five million cases of severe illness, and about 250 000 to 500 000 deaths [1]. Type C viruses cause mild sporadic diseases and have no epidemiological significance. Influenza A viruses are classified into subtypes based on the antigenic properties of surface glycoproteins: haemagglutinin (HA) and neuraminidase (NA). To date, 18 HA (H1-H18) and 11 NA (N1-N11) subtypes have been identified [2]. Since 1977 influenza type A viruses with combinations H1N1 and H3N2, along with type B viruses, have been circulating in the human population, and they are the antigens in influenza vaccines. During the last influenza pandemic in 2009/2010, a novel swine-origin A(H1N1)
virus, designated influenza A(H1N1)pdm09, completely replaced the seasonal A(H1N1) viruses circulating previously and continues to circulate on a global scale as a seasonal influenza virus together with A(H3N2) and type B viruses. Influenza type B viruses are divided into two main groups (lineages) represented by the strains B/Victoria/2/87 and B/Yamagata/16/88, which have been co-circulating in variable proportions since the late 1980s.

The genomes of influenza A and B viruses consist of eight segments of single stranded RNA with negative polarity, each of them encoding at least one protein with certain functions. Surface glycoproteins HA and NA play essential roles in viral infection and are strongly subjected to the immune pressure of the macro-organism [3, 4]. The antibodies induced by a natural infection or vaccination bind to certain antigenic domains, called antigenic sites, located on the globular head of the HA1 subunit, and prevent binding to sialic acid receptors. Historically, the major antigenic sites designated Sa, Sb, Ca 1/2, Cb in the A(H1N1) subtype, A-E in the A(H3N2) subtype and 120 loop, 150 loop, 160 loop, 190 helix in type B have been mapped by an isolation of escape mutants [5–8]. Influenza A viruses can circumvent host immune response by an accumulation of amino acid substitutions in the antigenic sites of HA, a process known as ‘antigenic drift’. A limited number of substitutions that occurred near the receptor binding site (RBS) can produce antigenically drifted variants with potential to cause epidemics [9, 10]. Gradually accumulated minor molecular changes in HA and NA of circulating influenza viruses lead to an increase of the antigenic distances to current vaccine strains and to a reduction of vaccine effectiveness. The glycosylation of HA and NA is another important viral mechanism to evade pre-existing humoral immunity. It has been speculated that N-linked glycans shield antigenic epitopes in the globular head region of HA, masking the access of antibodies towards antigenic sites, thus block immune recognition [11].

The Global Influenza Surveillance and Response System (GISRS) established by the WHO in 1952, monitors the circulation of influenza viruses and their continuous evolution, and provides recommendations concerning vaccine composition and antiviral susceptibility [1]. The data from antigenic and molecular analyses of influenza viruses are essential for identifying the emergence of influenza virus variants with a potential for avoiding host immune defense, with enhanced virulence or reduced susceptibility to antivirals. The main aims of this study are to investigate the circulation pattern of influenza viruses in Bulgaria during the 2016/2017 season, to characterise their antigenic and genetic features and to perform an amino acid sequence analysis of their surface glycoproteins and internal proteins.

**METHODS**

**Influenza surveillance system**

In Bulgaria, an acute respiratory infections (ARI) surveillance system is used to monitor influenza. The national sentinel network consists of general practitioners and pediatricians working in 211 outpatient care facilities in all 28 major cities - regional centers and serving 382 104 people from all age groups (5.3 % of the population in the country). During the period from November 1 to March 31, primary care physicians report the daily number of new cases of ARI by age group, and between April and October, the data are reported weekly (http://grippe.gateway.bg). Sentinel physicians are requested to collect respiratory specimens from patients with symptoms of influenza-like illness (ILI) or ARI and to send them to the National Reference Laboratory (NRL) for virological evaluation. ILI and ARI are defined in compliance with the European Centre for Disease Prevention and Control (ECDC) (https://ecdc.europa.eu/en/infectious-diseases-public-health/surveillance-and-disease-data/eu-case-definitions). The NRL is recognised as a WHO National Influenza Center. It performs testing of clinical samples from the sentinel network and from severely ill patients hospitalised in different regions of the country.

**Study population and specimen collection**

From week 40/2016 to week 20/2017, 913 patients from different regions of Bulgaria treated for ILI or ARI in primary care settings or hospitals were enrolled within the national influenza surveillance programme. Combined nasal and pharyngeal specimens from the enrolled patients were collected with the help of commercial polyester collection swabs (Deltalab). Swabs were stored at 4°C for up to 72 h before shipment to the laboratory. Specimens were processed immediately or stored at −80°C before testing.

**Extraction of nucleic acids and real time RT-PCR**

Viral nucleic acids were extracted automatically from respiratory specimens using a commercial ExiPrep Dx Viral DNA/RNA kit (Bioneer) in accordance with the manufacturer’s instructions. Detection and typing/subtyping of influenza viruses were carried out by a real time RT-PCR method and the SuperScript III Platinum One-Step qRT-PCR System (Invitrogen, ThermoFisher Scientific). Primers and probes were provided by the International Reagent Resource (IRR), USA. All samples were first tested for the presence of influenza A and B viruses. Those that were positive for influenza A were subsequently screened for A (H1N1)pdm09 and A(H3N2). The genetic lineage of detected influenza B viruses was also determined by real time RT-PCR. Amplification was performed using a Chromo 4 thermal cycler (Bio-Rad Laboratories) in accordance with the following protocol from CDC-Atlanta: reverse transcription (RT) at 50°C for 30 min, Taq inhibitor inactivation at 95°C for 2 min, 45 cycles of denaturation at 95°C for 15 s and annealing at 55°C for 30 s [12]. Samples with a cycle threshold \( C_t \) value \(<38\) were considered positive. The detection of respiratory-synctial virus (RSV), human metapneumovirus (HMPV), parainfluenza viruses (PIV) 1/2/3, rhinoviruses (RV), adenoviruses (AdV) and bocaviruses (BoV) was performed using singleplex real time PCR assays and an AgPath-ID One Step RT-PCR kit (Applied Biosystems, ThermoFisher Scientific). The
primers, probes and thermocycling conditions used in the study were identical to those previously described [13, 14]. For influenza type A and type B viruses, positive controls were provided by IRR, USA; for other targets, AmpliRun DNA/RNA Amplification Controls (Vircell) were used.

**Virus isolation and antigenic characterisation**

All real time RT-PCR positive clinical specimens with Ct values <28 were inoculated onto Madin Darby canine kidney (MDCK) and MDCK-SIAT1 (that express increased levels of α2,6-sialyltransferase) [15] cell cultures. Cultures were incubated at 35°C in a 5% CO2 atmosphere and observed on a daily basis for 7 days for evidence of cytopathology. The presence of a virus in a culture was confirmed by haemagglutination assay following standard protocols using a 1% suspension of guinea pig red blood cells. The antigenic characterisation of isolates was performed by the haemagglutination inhibition (HI) assay, in accordance with the WHO Manual, using vaccine viruses/antigens and their corresponding antisera provided by the WHO Collaborating Centres (WHO-CC) in London and Atlanta [12].

**Genetic characterisation**

Clinical specimens positive for influenza viruses and virus isolates were sequenced using Sanger method at the WHO-CC, London, and using Next Generation Sequencing (NGS) at the WHO-CC, Atlanta. Sequences were deposited in the Global Initiative on Sharing All Influenza Data (GISAID) (http://www.gisaid.org). For the purposes of phylogenetic analyses, the study sequences, together with sequences of reference viruses whose genetic group identities were known and of viruses circulating in different countries of Europe during the 2016/2017 season, were retrieved from the GISAID. They were all aligned using the MUSCLE program embedded in the Molecular Evolutionary Genetics Analysis (MEGA, version 6.06; http://www.megasoftware.net/) software [16]. Best nucleotide substitution models for phylogenetic analysis of HA (Hasegawa-Kishino-Yano model with a gamma distribution, HKY+G) and NA (Tamura 3-parameter) were selected using the Maximum Likelihood method within the MEGA 6.06. Phylogenetic trees were constructed using the Maximum Likelihood method within the MEGA 6.06. The reliability of the tree topology was assessed by bootstrapping with 1000 replications.

Pairwise nucleotide distances ($p$-distances), the number of pairwise base differences divided by the total number of nucleotides in the sequenced segment, were calculated using the MEGA $p$-distance method. Strains with 100% nucleotide sequence identity were excluded from the analyses.

**Deduced amino acid sequence analysis and prediction of N-glycosylation motifs**

The amino acid sequences were generated by translating nucleotide sequences with the standard genetic code using the MEGA software. The deduced amino acid sequences of the study strains were compared to those of vaccine strains and other reference strains to identify amino acid substitutions. The amino acid identity was calculated using flusurver (http://flusurver.bii.a-star.edu.sg). The potential N-glycosylation motifs in the HA and NA were predicted using the NetNGlyc 1.0 web Server (http://www.cbs.dtu.dk/services/NetNGlyc) to identify sequence motifs N–X–S/T (sequon), where X can be any amino acid except proline.

**Antiviral susceptibility testing**

The screening of A(H1N1)pdm09 viruses for the presence of point mutation conferring H275Y oseltamivir resistance was carried out using real time RT-PCR assay that allowed discrimination of a single nucleotide difference between oseltamivir sensitive and resistant viruses. Primer/probe sequences and protocol were kindly provided by Public Health England, London.

**RESULTS**

The influenza epidemic in the 2016/2017 season in Bulgaria started in week 49 of 2016 and continued until the 4th week of 2017. The epidemic peak was reached in the first week of 2017 with an incidence rate of 215.52 per 10 000 inhabitants. The highest consultation rate for ARI was observed in the 0–4 years age group, followed by the 5–14 age group with maximum values of 699.68 and 570.37 per 10 000 population, respectively (http://grippe.gateway.bg).

**Influenza virus detection**

Influenza viruses were detected in 283 (31%) patient samples. Of these, 276 (97.5%) were positive for influenza type A virus and seven (2.5%) for type B. Among the influenza A viruses, 264 (95.7%) were A(H3N2) and 12 (4.3%) - A (H1N1)pdm09 (Fig. 1). Of patients with influenza A(H3N2) infection, 96 (36.4%) were outpatients and 168 (63.6%) were hospitalised. Four detected influenza type B viruses belonged to the Victoria-lineage and three to the Yamagata-lineage.

Among the 404 patients younger than 5 years, tested for eight non-influenza viruses, the A(H3N2) virus was the second (21%) most frequently identified virus after RSV (36%). The remaining respiratory viruses were detected significantly less frequently: A(H1N1)pdm09 (1%), B/Victoria-lineage (0.5%), B/Yamagata-lineage (0.5%), HMPV (6.4%), PIV-1 (0%), PIV-2 (1.2%), PIV-3 (3.7%), RV (8.4%), AdV (9.9%) and BoV (0.7%) (Fig. 2).

Influenza virus detections started in week 49/2016 when the first A(H3N2) virus of the season was detected, and reached a peak in the first week of 2017. Influenza B virus detection started at the end of February. The last influenza virus (type B) was detected in week 14/2017 (Fig. 1).

**Age distribution**

Influenza infections occurred in all studied age groups. The highest incidence rate for influenza viruses was found in the age group 15–29 years (44%), followed by the age group 5–14 years (40%). The incidence of A(H3N2) infection was highest amongst the same age groups - 15–29 years (44%) and 5–14 years (35%).
Virus antigenic characterisation

The first 18 detected influenza viruses (clinical specimens) were sent to the WHO-CCs in London and Atlanta for further characterisation. In the course of the 2016/2017 season, a total of 63 influenza viruses were isolated on MDCK-SIAT1 cell culture. In total, 51 representative influenza isolates were sent to the WHO-CC, London, where they were characterised in detail. Out of the study A(H3N2) viruses, nine strains (six in subclade 3C.2a and three in subclade 3C.2a1) retained sufficient HA titre and were antigenically characterised at the WHO-CC, London, by HI assay using post-infection ferret antisera raised against the reference viruses shown in Table 1. All test viruses but two were recognised well (reacting at titers of less than or equal to four-fold of the homologous virus titer) by the antisera raised against the reference viruses A/Stockholm/6/2014, cell culture-propagated A/Switzerland/9715293/2013, A/Hong Kong/5738/2014 and A/Hong Kong/4801/2014, the A (H3N2) component of the 2016/2017 Northern Hemisphere (NH) vaccine. The study viruses showed low reactivity with antisera raised against the egg-propagated viruses A/Texas/50/2012, A/Samara/73/2013, A/Switzerland/9715293/2013 and cell culture-propagated virus A/Georgia/532/2015. All six A(H1N1)pdm09 isolates analysed were well-inhibited by...
post-infection ferret antisera raised against A/California/7/2009 and A/Michigan/45/2015 strains, the A(H1N1)pdm09 components of the 2016/2017 and 2017/2018 NH vaccines, indicating antigenic similarity to both vaccine strains. Two Bulgarian B/Victoria-lineage viruses analysed showed low reactivity with antisera raised against egg-propagated B/Brisbane/60/2008 strain, a recommended influenza B component of the 2016/2017 NH trivalent and quadrivalent influenza vaccines.

Genetic characterisation

Phylogenetic trees based on HA and NA genes were constructed to determine the genetic relationships of Bulgarian isolates with reference viruses and viruses circulating in other countries in the same period. The HA and NA genes of 29 (52 %) of the 56 A(H3N2) viruses analysed belonged to genetic clade 3C.2a together with the 2016/2017 NH vaccine virus, A/Hong Kong/4801/2014, and the remaining 27 (48 %) belonged to the new genetic subclade 3C.2a1. The 3C.2a genetic subclade includes an emerging subclade known as 3C.2a1 (Fig. 3). All six A(H1N1)pdm09 viruses analysed fell within genetic subclade 6B.1, and clustered with A/Slovenia/2903/2015 (representative of this subclade) and A/Michigan/45/2015, the recommended A(H1N1)pdm09 vaccine strain for the upcoming 2017/18 season (Fig. 4). Two B/Victoria-lineage viruses analysed fell into genetic clade 1A represented by the B/Brisbane/60/2008 vaccine virus.

Amino acid sequence analysis

The complete HA and NA amino acid sequences of the Bulgarian influenza isolates were compared to those of vaccine viruses in order to identify substitutions that could lead to reduced vaccine effectiveness (Table 2).

**A(H3N2)**

The HA amino acid sequence identity of the studied 56 A(H3N2) isolates ranged from 97.88 to 99.115 % as compared to the vaccine strain, A/Hong Kong/4801/2014. Like the A/Hong Kong/4801/2014 strain, the Bulgarian isolates possessed a number of amino acid substitutions as compared to the previous vaccine virus, A/Texas/50/2012: N145S in HA1 and D160N in HA2, which define clade 3C.2, plus L3I, N144S (resulting in the loss of a potential glycosylation site), F159Y, K160T (resulting in the gain of a potential glycosylation site), N225D and Q311H in HA1, which define subclade 3C.2a. Substitutions N171K in HA1, I77V and G155E in HA2 as compared to the A/Hong Kong/4801/2014 vaccine strain define the new subclade 3C.2a1 [17]. Most Bulgarian subclade 3C.2a1 sequences fell in a group of viruses that also carried substitution N121K in HA1 (Table 3). In general, the study 3C.2a and 3C.2a1 sequences showed variations at seven and 12 HA positions, respectively as compared to the A/Hong Kong/4801/2014 vaccine strain. Substitutions N122D and T135K resulted in the loss of N-glycosylation motifs. Amino acid changes N122D, T135K, R142G and S144K are located in antigenic site A; G49S - in C; S219Y - in D and K29R - in E; S144K and T135K are located adjacent to the RBS. Thirteen potential N-glycosylation motifs in HA (HA1 positions 8, 22, 38, 45, 63, 122, 126, 133, 158, 165, 246, and 285, and HA2 position 154) were identified. All subclade 3C.2a1 viruses possessed the variation NGT →NET in position 155 of HA2 glycosylation site.

The NA sequences of the Bulgarian A(H3N2) viruses differed from the NA of the A/Hong Kong/4801/2014 strain by 23 amino acid substitutions: 13 in subclade 3C.2a and 20 in subclade 3C.2a1 viruses. Substitution S245N resulted in the formation of a potential N-glycosylation motif; G247T maintained the glycosylation motif; N329S/T resulted in the loss of a glycosylation motif. Nine potential N-glycosylation motifs in NA (61, 70, 86, 146, 200, 234, 245, 329 and 367) were identified, with two (146 and 367) being located around the enzymatic active site [18].

The amino acid sequences of the internal proteins of five Bulgarian A(H3N2) viruses were compared to those of the A/Hong Kong/4801/2014 vaccine virus and other reference viruses. The Bulgarian sequences possessed the following amino acid substitutions: V63I, R299K and I588T in PB2;
T152S and Y166F in PB1; R25Q, H75P and K85R in PB1-F2; N272S in PA; G101D, D454E and T472A in NP; A56S, E71G, N74D and V182I in NS1. Most of these changes were also observed in subclade 3C.2a1 viruses detected in other countries. No substitutions were found in the proteins M1, M2 and NEP.

Fig. 3. HA gene phylogeny of influenza A(H3N2) viruses detected during the 2016/2017 season. Reference viruses are indicated in bold and vaccine virus A/Hong Kong/4801/2014 in red. Bulgarian viruses detected in December 2016, January, February and March 2017 are indicated in pink, blue, green and purple, respectively. The tree is rooted at A/Texas/50/2012. Amino acid substitutions defining particular nodes are indicated and bootstrap values greater than 70% are shown.
A(H1N1)pdm09

The HA amino acid sequence identity of the studied six A(H1N1)pdm09 isolates as compared to vaccine strains, A/California/7/2009 and A/Michigan/45/2015, was 97.35 and 99.823 %, respectively. The HA sequences contained variations in 11 HA1 and 3 HA2 positions as compared to the A/California/7/2009 vaccine virus. HA1 amino acid substitutions S84N, S162N (resulting in the formation of a new potential glycosylation motif) and I216T define the subclade 6B.1, which emerged at the beginning of the 2015/2016 season and became dominant worldwide [19]. Two substitutions were located in antigenic sites: S185T in Ca1 and S203T in Sb [20]. The S185T substitution falls within a domain defining the

Fig. 4. HA gene phylogeny of influenza A(H1N1)pdm09 viruses detected during the 2016/2017 season. Reference viruses are indicated in bold, vaccine viruses A/California/7/2009 and A/Michigan/45/2015 – in red. Bulgarian viruses detected in December 2016, January and February 2017 are indicated in pink, blue and green, respectively. The tree is rooted at A/California/7/2009. Amino acid substitutions defining particular nodes are indicated and bootstrap values greater than 70 % are shown.

Korsun et al., Journal of Medical Microbiology 2018;67:228–239
Table 2. Amino acid substitutions in HA and NA of influenza viruses, detected in Bulgaria during the 2016/2017 season

<table>
<thead>
<tr>
<th></th>
<th>A(H1N1)pdm09 6B.1 sequences (n=6) compared to A/California/7/2009</th>
<th>A(H3N2) 2C.2a sequences (n=29) compared to A/Hong Kong/4801/2014</th>
<th>A(H3N2) 2C.2a1 sequences (n=27) compared to A/Hong Kong/4801/2014</th>
<th>B-Victoria sequences (n=2) compared to B/Brisbane/60/2008</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Amino acid changes in HA and NA</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>HA2</strong></td>
<td>F47K (6)</td>
<td>S124N (6)</td>
<td>E172K (6)</td>
<td></td>
</tr>
</tbody>
</table>

Only amino acid substitutions found in two or more viruses are presented in this table. The number of Bulgarian influenza sequences possessing the substitution is indicated within parentheses. Substitutions within antigenic epitopes are highlighted in dark gray; antigenic sites are identified in bold italic. (+/-CHO) - gain/loss of N-glycosylation motif.
RBS 190-helix (residues 184–191) [21, 22]. The HA sequences of the A(H1N1)pdm09 viruses circulating in Bulgaria during the 2015/2016 and 2016/2017 seasons were very similar. However, in the analysed 2016/2017 viruses, there were found no HA substitutions A141T and I418V, which were present in 28 and 16%, respectively, of the analysed 2015/2016 viruses [23]. The NA sequences of the studied six A(H1N1)pdm09 viruses differed from that of A/California/7/2009 by substitutions in 16 positions. All study strains carried nine conserved potential N-glycosylation motifs in HA (HA1 positions 10, 11, 23, 87, 162, 276, 287 and HA2 positions 154 and 213) and eight in NA (positions 42, 50, 63, 68, 88, 146 and 235).

**B/Victoria**

The amino acid analysis revealed that the HA and NA sequences of the studied two B/Victoria 2016/2017 strains were similar to those of strains circulating during the previous epidemic season [23]. Two HA amino acid substitutions, both (I117V and N129D) in the 120-loop (positions 116–137) antigenic site, and five NA amino acid substitutions were identified as compared to the B/Brisbane/60/2008 vaccine virus. No deletions at positions 162–163 and 162–164 in HA1 gene were found. Both Bulgarian isolates possessed the same twelve conserved N-glycosylation motifs in HA (HA1 positions 25, 59, 145, 166, 197, 233, 304 and 333, and HA2 positions 145, 171, 184 and 216) and four in NA (positions 56, 64, 144 and 284) identified in B/Brisbane/60/2008.

**Antiviral susceptibility testing**

All 12 detected A(H1N1)pdm09 viruses were analysed by real-time RT-PCR with respect to the H275Y mutation—none of the isolates harbored the H275Y neuraminidase substitution that confers oseltamivir resistance. No other specific NA amino acid substitutions, markers of resistance or reduced susceptibility to NA inhibitors oseltamivir or zanamivir were present in Bulgarian influenza sequences [24]. The M2 proteins of all studied viruses carried S31N substitution conferring resistance to M2-channel blockers (amantadine and rimantadine) [25].

**DISCUSSION**

The 2016/2017 season in Bulgaria was characterised by an early start as compared to previous seasons, an exclusive dominance of A(H3N2) viruses accounting for a total of 93% of influenza virus detections and a low circulation of A (H1N1)pdm09 (4.2%) and type B (2.5%) viruses. Similar proportions of circulating seasonal influenza viruses were observed in most European countries [17, 26]. Influenza A (H3N2) viruses emerged in the human population during the 'Hong Kong Flu' pandemic of 1968 and since then have been circulating on a global scale. In recent years, a prevalent circulation of these viruses in Bulgaria and other countries was observed during the 2011/2012, 2014/2015 and 2016/2017 seasons [27–30]. Since 2009, the HA genes of A (H3N2) viruses have undergone considerable genetic diversification and evolved in seven genetic groups and multiple clades/subclades. At present, viruses in clade 3C, which has three subdivisions: 3C.1, 3C.2 and 3C.3, are the dominant group. Subclade 3C.2a emerged during the 2013/2014 NH influenza season and has been predominating since the 2014/2015 season. It was characterised by inability of 90% of the viruses to agglutinate RBCs, so it was difficult or not possible to assess the correlation between amino acid/glycosylation changes and the antigenic properties of these viruses using HI assay [17]. Subclade 3C.2a1 emerged as variant of subclade 3C.2a in 2016.

In Bulgaria, the 3C.2a and the 3C.2a1 subclades were detected in approximately equal proportions. According to a report of the WHO-CC, London [31], subclade 3C.2a1 viruses are antigenically related to the 2016/2017 A/Hong Kong/4801/2014 vaccine strain (representative of subclade 3C.2a), which was confirmed by the antigenic characterisation of three Bulgarian strains in subclade 3C.2a1. Nevertheless, amino acid variations in 13 HA and 20 NA positions were identified in the Bulgarian and other subclade 3C.2a1 viruses as compared to the A/Hong Kong/4801/2014 vaccine strain. The most frequent combination of amino acid substitutions in the study subclade 3C.2a1 viruses (N121K+N171K+I77V+G155E) was observed in viruses detected in other countries, as well [32]. The Bulgarian viruses in subclade 3C.2a have also evolved and differed

<table>
<thead>
<tr>
<th>Table 3. Frequent combinations of amino acid substitutions in HA of Bulgarian A(H3N2) viruses</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Combinations of amino acid changes</strong></td>
</tr>
<tr>
<td>N21K+S144K (-CHO)</td>
</tr>
<tr>
<td>N21K+N122D (-CHO)+S144K+S262N</td>
</tr>
<tr>
<td>G49S+N121K+N122D (-CHO)+S144K (-CHO)+S262N</td>
</tr>
<tr>
<td>E58V+N121K+S144K (-CHO)+S219Y</td>
</tr>
<tr>
<td><strong>Combinations of amino acid changes</strong></td>
</tr>
<tr>
<td>N171K+I77V+G155E</td>
</tr>
<tr>
<td>N121K+N171K+I77V+G155E</td>
</tr>
<tr>
<td>N121K+N171K+I77V+G150E+G155E</td>
</tr>
<tr>
<td>N121K+I140M+N171K+I77V+G150E+G155E</td>
</tr>
<tr>
<td>N121K+R142G+N171K+I77V+G155E</td>
</tr>
<tr>
<td>N121K+R142G+N171K+I77V+G155E</td>
</tr>
<tr>
<td>S9N+K92R+N171K+K311Q+I77V+G155E</td>
</tr>
</tbody>
</table>

Amino acid positions are numbered without considering the HA signal peptide. Substitutions in HA2 subunit are shown in italic; (+/−CHO) – gain/loss of a potential glycosylation motif.
from the vaccine strain in seven HA and 13 NA positions. These changes have been accumulated over a relatively short period of time – three seasons.

On the whole, the amino acid substitutions in the seven HA positions of the analysed A(H3N2) viruses were located within antigenic sites A, C, D, and E, and these changes could potentially alter antigenicity. In previous studies it was suggested that simultaneous occurrence of at least four substitutions across two or more antigenic sites would predict the emergence of antigenic drift variants of epidemiological significance [7, 9, 33]. Influenza A(H3N2) viruses have at least five distinct antigenic sites designated by the letters A to E, including 131 amino acid positions [5, 7]. Epitopes A, B, D, and E are located in the globular head region, while epitope C is in the stem region of HA1. It has been reported that human antibodies appear to be focused mainly on antigenic sites B and A, localised on the top of HA around the RBS [10, 34]. Koel et al. [10] found that the amino acid substitutions responsible for major antigenic changes in the A(H3N2) viruses over the period 1968–2003 were located exclusively in antigenic sites A (position 145) and B (positions 155, 156, 158, 159, 189, and 193) near the RBS, with none in sites C, D, or E. In our study no substitutions were found in any of these seven key positions.

A number of substitutions were identified in the internal proteins PB2, PB1, PB1-F2, PA, NP and NS1, but the possible biological relevance of these changes is still unknown. Internal proteins have multiple roles within the virus life cycle, pathogenesis and virulence: the first three proteins form polymerase complex; PB1-F2 is a pro-apoptotic protein; NP contains many T-cell epitopes, and substitutions in NP may affect a host’s cytotoxic immune response [35]; NS1 is involved in enhancing antagonism to interferon alpha, which increases virulence [36].

The accumulation of oligosaccharide chains in the globular head region of HA has been suggested to facilitate immune evasion [11, 37]. As regards A(H3N2) viruses, there has been a large increase in the number of predicted HA glycosylation motifs since their appearance in humans in 1968 [37, 38]: viruses isolated in 1968 had seven HA and eight NA N-glycosylation motifs, while those isolated in 2016/2017 had 13 and nine, respectively. The data shown in Table 4 confirm the higher levels of amino acid substitutions and of the degree of glycosylation of A/H3 viruses, as compared to A/H1 and type B viruses [39–42]. The higher variability and the faster evolution of A/H3 viruses necessitate more frequent updates to this component of the influenza vaccine. For example, from 1968 to 2010, there were 108 amino acid changes identified at 63 residue positions in HA1, resulting in 27 A(H3N2) strain alterations in the vaccine formulation, twice as many as in the other vaccine components - influenza type B or former A (H1N1) strains [34].

Despite the amino acid changes in HA and NA accumulated during the seven epidemic seasons, A(H1N1)pdm09 viruses remained antigenically similar to the vaccine virus A/California/7/2009 in studies with post-infection ferret sera. Since 2009, nine genetic clades have been defined for A(H1N1) pdm09 viruses but there has been no evidence of antigenic drift [17]. However, due to the fact that recently circulating viruses in subclade 6B.1 were poorly inhibited by some post-vaccination (A/California/7/2009) adult human serum pools, a subclade 6B.1 representative A/Michigan/45/2015-like strain was recommended as a vaccine virus for the 2017 southern hemisphere and 2017/2018 NH influenza season [43].

The amino acid sequence analyses of B/Victoria lineage viruses confirm previous observations that influenza B viruses exhibit lower rates of evolutionary changes and have limited genetic diversity as compared to type A viruses [44]. Both identified HA substitutions were located in antigenic domain 120 loop, which is one of the most frequently altered regions, and amino acid changes in this region may strongly affect virus antigenicity [45–47]. The 162–163 double deletion in the HA, observed in 78 (20 %) of the 390 B/Victoria-lineage strains in the USA and in three Norwegian strains, and the 162–164 triple deletion, detected in Hong Kong, were not present in the Bulgarian isolates [17, 48].

In conclusion, the results of this study reaffirm the phenomenal genetic plasticity and rapid evolution of seasonal influenza viruses, particularly A(H3N2), and reveal the need for continuous antigenic and molecular surveillance for the purposes of early detection of novel genetic variants of epidemiological and clinical significance. Such studies are important for the proper selection of vaccine strain candidates.

### Table 4. Number of positions with amino acid changes compared to vaccine viruses and number of potential N-glycosylation motifs in HA and NA of influenza viruses circulating in Bulgaria during the 2016/2017 season

<table>
<thead>
<tr>
<th>Influenza viruses</th>
<th>Number of positions with amino acid changes compared to vaccine virus</th>
<th>N-glycosylation motifs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total number</td>
<td>Positions in antigenic regions</td>
</tr>
<tr>
<td>A(H3N2)</td>
<td>18</td>
<td>4</td>
</tr>
<tr>
<td>A(H1N1)pdm09</td>
<td>14</td>
<td>2</td>
</tr>
<tr>
<td>B/Victoria lineage</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

Funding information

This work was supported by the Ministry of Health in Bulgaria (‘National Plan of the Republic of Bulgaria for Influenza Pandemic Preparedness’).
Acknowledgements
The authors would like to express their special gratitude to the collaborators within the Worldwide Influenza Centre, the Francis Crick Institute, London, and the collaborators within the Virus Surveillance and Diagnosis Branch, Influenza Division, CDC, Atlanta, for their contribution to antigenic and genetic characterisation of influenza viruses. The authors also gratefully acknowledge the contributions of team members in the authors’ host institutions, and of the originating and submitting laboratories of sequences from GISAI’s EpiFlu Database that were used in this study.

Conflicts of interest
The authors declare that there are no conflicts of interest.

Ethical statement
This study was approved by the local Ethics Committee of the National Centre of Infectious and Parasitic Diseases. All patients who participated in the study provided written informed consent before specimen collection and testing.

References


