ORIGINAL ARTICLE

Molecular characterization and phylogenetic analysis of human influenza A viruses in three consecutive seasons with different epidemiological profiles

E. PARIANI[*], E.R. FRATI[*], A. AMENDOLA[*], A. ZAPPA[*], S. BIANCHI[*], D. COLZANI[*], M. CANUTI[*], D. BRAMBILLA[*], A. ZANETTI[* **], E. TANZI[* **]

[*] Department of Public Health-Microbiology-Virology, University of Milan, Italy; [**] CIRI-IV, Department of Health Sciences, University of Genoa, Italy

Key words

Influenza A viruses • Molecular characterization • Phylogenetic analysis

Summary

Introduction. *Influenza activity and influenza virus circulation were observed in Lombardy (northern Italy) during three consecutive seasons and the molecular characteristics of circulating viruses analysed to control for introduction of new variants.*

Methods. *The molecular characterization of 38 isolates, namely 20 A/H3N2 and 18 A/H1N1 influenza strains from the 2005/06, 2006/07 and 2007/08 seasons, was performed by sequence analysis of the globular head region of the HA protein (HA1 subunit), specific for influenza virus A/H3 and A/H1.*

Results and discussion. *The last three influenza seasons in the study region were characterized by medium-low activity. A typical co-circulation of several variants was shown for A/H3 viruses for approximately two years and were subsequently almost entirely substituted by new emerging variants. Vice versa, A/H1 viruses had a more homogeneous circulation with a single lineage clearly dominating each season. The HA sequences of the A/H3 and the A/H1 viruses isolated in the last three seasons fell into 4 and 3 principal phylogenetic groups, respectively. No evidence of positive or negative selection in the sequence alignments was observed.*

Conclusions. *Molecular characterization of the influenza viruses in three consecutive seasons highlighted considerable heterogeneity in their HA sequences. A careful surveillance of genetic changes in the HA1 domain during seasonal influenza epidemics may reveal immune escape and provide early information on newly emerging strains with epidemiologic inference.*

Introduction

Influenza, a major cause of mortality and morbidity worldwide, is an acute infection caused by a group of negative-stranded RNA viruses of the *Orthomyxoviridae* family [1]. Influenza viruses are characterized by remarkable biological dynamism and responsible for their rapid, unpredictable antigenic variation [2, 3]. Haemagglutinin (HA) is the major membrane-bound glycoprotein on the viral surface responsible for receptor-binding and membrane fusion, and is the target for neutralizing antibodies elicited by both infection and vaccination. HA is synthesized as a single polypeptide that is subsequently cleaved into two polypeptides: HA1 and HA2, linked by a disulphide bond [4]. The HA1 polypeptide, the principal target of antibody-mediated immunity, mutates more frequently than HA2 and plays a crucial role in natural selection.

It has been suggested that antibody binding on the HA protein of A/Aichi/2/68 (H3N2) occurs in five antigenic sites (A to E) located on the protein's three-dimensional structure [5, 6]. Approximately one-third of the HA1 amino acids lies in proximity to these sites, although the importance of their positions is unclear [7, 8]. Hence, amino acid substitutions may impair the neutralizing ability of antibodies through interference with either antibody binding or an associated process (e.g. receptor binding). As a result, antisera against one virus often display only limited effectiveness against future strains [9].

The rapid evolution of influenza viruses represents a major obstacle to the ability to timely recognise or to predict current and future epidemiological threats. Sequence-based studies of viral evolution to evade immune response yielded some interesting clues on the possible mechanisms of influenza seasonality [10].

As part of the Italian Influenza Surveillance Network [11-13], influenza activity and influenza viruses circulation were observed in Lombardy (a region with approximately 9 million inhabitants) during three consecutive seasons (2005/06, 2006/07, 2007/08) characterized by different epidemiological pictures. The molecular characteristics of circulating influenza A viruses were analysed in order to evaluate the introduction of new variants in the territory. A phylogenetic analysis was carried out to investigate the evolution of A/H3N2 and A/H1N1 viruses in such different epidemiological scenarios.

Methods

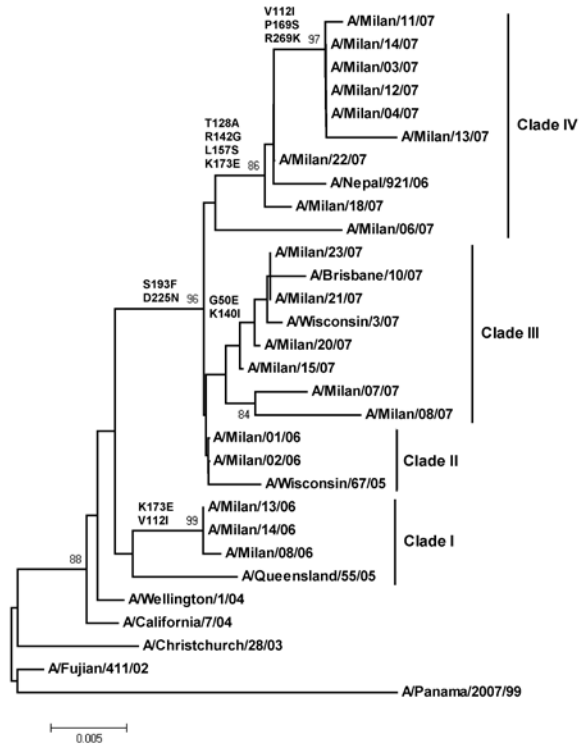
38 viral isolates, namely 20 A/H3N2 and 18 A/H1N1 influenza strains from 2005/06, 2006/07, and 2007/08

seasons underwent molecular characterization by sequence analysis of the globular head region of the HA protein (HA1 subunit), specific for influenza virus A/H3 (nt. 174-1056) [14] and A/H1 (nt. 76-1090) [15].
Viral RNA was extracted from respiratory samples of outpatients with clinical evidence of influenza-like illness (ILI) by QIAmp Viral RNA kit (QIAGEN GmbH, Germany). Following the RT-PCR of the HA1 gene, amplicons were purified using NucleoSpin® Extract II (Macherey-Nagel GmbH, Germany) and nucleotide sequences obtained from automated DNA sequencing on the ABI PRISM 3100 genetic analyzer (Applied Biosystem, CA, USA). Multiple sequence alignment was conducted using ClustalX, version 2.0. Phylogenetic trees from A/H1 and A/H3 HA1 sequences were constructed by means of the Neighbor-Joining method [16] and Kimura 2-Parameter model [17], using the MEGA package, version 4.0 [18]. A bootstrap analysis (N=1,000) was performed and major branches with bootstrap values > 70% were identified as distinct groups [19]. The HA gene sequences of the study viral strains were deposited into NCBI Influenza Virus Sequence Database [20], under accession numbers: EU400232-EU400235, EU400237-EU400246, EU400248-EU400256, EU400258, EU400259, EU400261, EU400263-EU400267 and GQ246463-GQ246470. The reference viral strains used for the construction of phylogenetic trees were obtained from the NCBI Influenza Virus Sequence Database [20] (EU100702, EU199366, EF473424, CY017611, EF473341, EF566035, EF541397, DQ487340, EU199273, CY012104, AY289929, EU100594, EU124177, EU199352, CY030230, DQ265706).
To estimate the selection pressures acting on the HA gene, codon-specific non-synonymous (d_N) and synonymous (d_S) substitutions were inferred using the Nei-Gojobori method [21] and Jukes-Cantor model [22] by MEGA [18], and the Single Likelihood Ancestor Counting (SLAC), Fixed Effects Likelihood (FEL) and Random Effects Likelihood (REL) methods, all incorporating the HKY85 substitution models with phylogenetic trees inferred using the Neighbor-Joining method [16] available at the Datamonkey facility [23].

Results and discussion

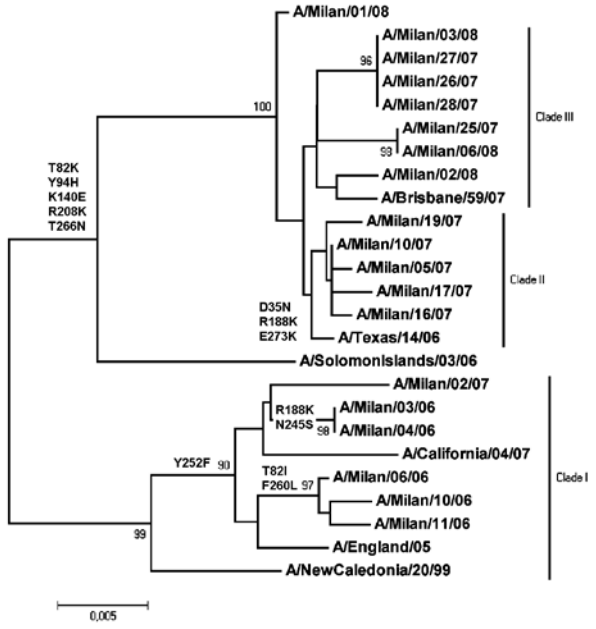
The last three influenza seasons in the study area were marked by medium-low influenza activity but with distinct epidemiological features. The 2005/06 winter season was characterized by a patchy pattern of influenza activity almost exclusively sustained by influenza A viruses, which accounted for 80.5% of total detections (51.7% A/H1N1 and 48.3% A/H3N2). The 2006/07 season was dominated by influenza A/H3N2 viruses, accounting for 93.6% of total detections and the 2007/08 epidemic was upheld by both A/H1N1 and B viruses (40% and 60% of total detections, respectively).
As shown in Figure 1 the HA sequences of the A/H3 viruses isolated in the last three seasons fell into four distinct phylogenetic groups. Three viruses isolated

Fig. 1. A/H3 HA1 phylogenetic tree. Sequences from 2005/06 are labelled in light red and the ones from 2006/07 in dark red. Major amino acid changes are reported in block letters.



during the 2005/06 season were characterized as belonging to the older phylogenetic group (Clade I) represented by A/Queensland/55/05 and characterized by the V112I and K173E amino acid changes. The viral variants to this clade isolated in the present study were distinguished by an additional S199P substitution. The remaining A/H3 viruses isolated in 2005/06 were represented by A/Wisconsin/67/05 (Clade II) and presented S193F and D225N changes in HA1.
In the 2006/07 season, the majority of HA sequences from the isolated A/H3 viruses fell within two principal clades, distinguished by amino acid changes on the A/Wisconsin/67/05 strain. Several viruses fell within Clade III represented by A/Brisbane/10/07 and presented the G50E and K140I amino acid changes.
Finally, the majority of the A/H3 viruses isolated during the 2006/07 season fell within Clade IV represented by A/Nepal/921/06 and were characterized by the amino acid changes T128A - which resulted in the loss of a potential N-linked glycosilation site -, R142G, L157S, and K173E. Variants were further distinguished by amino acid substitutions V112I, P169S, and R269K.
Seasons 2005/06 and 2006/07 were characterized by the co-circulation of A/H3 viruses belonging to distinct phylogenetic groups while no A/H3 viruses were detected in 2007/08 season. The HA sequences of the analyzed A/H3 viruses presented heterogeneity suggesting the fixed presence of several amino acid mutations in the viral population and the co-circulation of a mis-

Fig. 2. A/H1 HA1 phylogenetic tree. Sequences from 2005/06 are labelled in light blue, the ones from 2006/07 in dark blue, and the ones from 2007/08 in grey. Major amino acid changes are reported in block letters.



cellaneous set of variants notwithstanding the mid-low activity of influenza.

There was no evidence of positive or negative selection in the sequence alignment of A/H3 viruses. The [d_N-d_S (± Standard Error)] value was [-0.016 (± 0.006)]. The integrative selection analysis (SLAC p-value = 0.1; FEL p-value = 0.1; and REL BF = 50) found one positive selected codon in position 186 (epitope B) and three negative selected codons in positions 222, 267 and 302, respectively.

These findings corroborate the dynamic evolution of A/H3 viruses, as previously reported by others [24-29]. A/H3 viruses present a characteristic co-circulation of several variants for up to two years, which are subsequently almost completely substituted by new emerging variants with different antigenic features [30]. This continuous evolution of A/H3 viruses results in the appearance of new viral variants that can elude the human immune response, causing clinical outbreaks [30].

The HA sequences of the A/H1 viruses isolated during the three considered seasons were separated into three phylogenetic groups as shown in Figure 2. All A/H1 isolated during 2005/06 fell within Clade I represented by A/New Caledonia/20/99 and presenting the Y252F amino acid change. Three sequences presented the amino acid changes T82I and F260L whilst the rest was characterized by the R188K and N245S substitutions. The A/Milan/10/06 sequence was characterized by the amino acid change S161F which resulted in the loss of a potential N-linked glycosilation site. HA sequences within Clade II were represented by A/Solomon Islands/3/06, characterized by 5 amino acid changes (T82K, Y94H, K140E, R208K, and T266N), and exhibited a greater diversity. All the HA sequences, but one, of the viruses isolated during the 2006/07 season fell into this phylogenetic group and presented three additional amino acid changes, i.e. D35N, R188K, and E273K. All the A/H1 viruses isolated during the 2007/08 season fell within Clade III and were represented by A/Brisbane/59/07. These sequences did not present any amino acid changes.

The circulating A/H1 viruses were more homogeneous than the A/H3 viruses. Most A/H1 viruses isolated during the 2005/06 season were closely related to A/New Caledonia/20/99, while the HA sequences of the A/H1 viruses from the 2006/07 were A/Solomon Islands/3/06-like though exhibiting a greater diversity. The last season was characterized by the exclusive circulation of A/H1 viruses, all closely related to A/Brisbane/59/07.

The sequence alignment of A/H1 viruses reported a negative [d_N-d_S (± Standard Error)] value [-0.060 (± 0.012)]. The integrative selection analysis found four negative selected codons in positions 100, 113, 157, and 218. These findings confirm that the evolution of A/H1 viruses is marked by two main lineages circulating in the human population in two distinct times [25]. A/H1 viruses are also characterized by a low frequency of amino acid substitutions that reflects their neutral evolution [25].

Conclusion

Despite the mid-low clinical activity of influenza during the three analysed seasons, the molecular characterization of the isolated viruses highlighted a considerable heterogeneity in their HA sequences suggesting the co-circulation of different variants. This co-circulation of multiple lineages may condition the seasonal evolution of influenza viruses even more than the antigenic drift. Indeed, periodic and selection-driven cluster jumps could result in major changes in the antigenic phenotype of these viruses [9].

Both molecular characterization and the phylogenetic analysis were shown to be important means to better understand the evolution of influenza A viruses and their dynamic nature.

Therefore, a careful surveillance of genetic changes in the HA1 domain during seasonal influenza epidemic may provide early information on newly emerging strains with epidemiologic inference.

References

[1] Palese P. *Influenza: old and new threats*. Nat Med 2004;10:82-7.

[2] Nicholson KG, Wood JM, Zambon M. *Influenza*. Lancet 2003;362:1733-45.

[3] Zambon MC. *The pathogenesis of influenza A and B in humans*. Rev Med Virol 2001;11:227-41.

[4] Bush RM, Bender CA, Subbarao K, Cox NJ, Fitch WM. *Predicting the evolution of human influenza A*. Science 1999;286:1921-5.

[5] Lee MS, Chen JS. *Predicting antigenic variants of influenza A/H3N2 viruses*. Emerg Infect Dis 2004;10:1385-90.

[6] Ghedin E, Sengamalay NA, Shumway M, Zaborsky J, Feldblyum T, Subbu, et al. *Large-scale sequencing of human influenza reveals the dynamic nature of viral genome evolution*. Nature 2005;437:1162-6.

[7] van Nimwegen E. *Epidemiology. Influenza escapes immunity along neutral networks*. Science 2006;314:1884-6.

[8] Lee MS, Chen MC, Liao YC, Hsiung CA. *Identifying potential immunodominant positions and predicting antigenic variants of influenza A/H3N2 viruses*. Vaccine 2007;25:8133-9.

[9] Hay AJ, Gregory V, Douglas AR, Lin YP. *The evolution of human influenza viruses*. Phil Trans R Soc Lond 2001;356:1861-70.

[10] Blackburne BP, Hay AJ, Goldstein RA. *Changing selective pressure during antigenic changes in human influenza H3*. PLoS Pathog 2008;4:e1000058. doi:10.1371/Journal.ppat.1000058

[11] http://www.influciri.it

[12] http://www.iss.it/iflu/

[13] http://www.ministerosalute.it

[14] Ellis JS, Chakraverty P, Clewley JP. *Genetic and antigenic variation in the haemagglutinin of recently circulating human influenza A (H3N2) viruses in the United Kingdom*. Arch Virol 1995;140:1889-904.

[15] Ellis JS, Alvarez-Aguero A, Gregory V, Lin TP, Hay A, Zambon MC. *Influenza AH1N2 viruses and their impact during the 2001-02 influenza season in the United Kingdom*. Emer Infect Dis 2003;9:304-10.

[16] Saitou N, Nei M. *The neighbor-joining method: A new method for reconstructing phylogenetic trees*. Mol Biol Evol 1987;4:406-25.

[17] Kimura M. *A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences*. J Mol Evol 1980;16:111-20.

[18] Tamura K, Dudley J, Nei M, Kumar S. *MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0*. Mol Biol Evol 2007;24:1596-9.

[19] Felsenstein J.*Confidence limits on phylogenies: An approach using the bootstrap*. Evolution 1985;39:783-91.

[20] http://www.ncbi.nlm.nih.gov/Genbank/index.html

[21] Nei M, Gojobori T. *Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions*. Mol Biol Evol 1986;3:418-26.

[22] Jukes TH, Cantor CR. *Evolution of protein molecules*. In: Munro HN. *Mammalian Protein Metabolism*. New York, USA: Editor Academic Press 1969, pp. 21-132.

[23] www.datamonkey.com

[24] Nelson MI, Holmes E. *The evolution of epidemic influenza*. Nature 2007;8:196-204.

[25] Wolf YI, Viboud C, Holmes EC, Koonin EV, Lipman DJ. *Long intervals of stasis punctuated by bursts of positive selection in the seasonal evolution of influenza A virus*. Biol Direct 2006;1:1-19.

[26] Nelson IM, Simonsen L, Viboud C, Miller MA, Taylor J, George KS, et al. *Stochastic processes are key determinants of short-term evolution in influenza A virus*. PLoS Pathog 2006;2:1144-51.

[27] Zhai W, Slatkin M, Nielsen R. *Exploring variation in the dN/dS ratio among sites and lineages using mutational mappings applications to the influenza virus*. J Mol Evol 2007;65:340-8.

[28] Plotkin JB, Dushoff J, Levin SA. *Hemagglutinin sequence clusters and the antigenic evolution of influenza A virus*. Proc Natl Acad Sci USA 2002;99:6263-8.

[29] Bush RM, Fitch WM, Bender CA, Cox NJ. *Positive selection on the H3 hemagglutinin gene of human influenza virus A*. Mol Biol Evol 1999;16:1457-65.

[30] Holmes EC, Ghedin E, Miller N, Taylor J, Bao Y, St. George K, et al. *Whole-genome analysis of human influenza A virus reveals multiple persistent lineages and reassortment among recent H3N2 viruses*. PLoS Biol 2005;3:1579-88.

■ Received on March 5, 2009. Accepted on May 19, 2009.

■ Correspondence: Elena Pariani, Department of Public Health-Microbiology-Virology, University of Milan, via Pascal 36/38, 20133 Milan, Italy - Tel. +39 02 50315122 - E-mail: elena.pariani@unimi.it