# Molecular Epidemiology and Phylogenetic Analysis of Human Influenza Viruses

Elias.T. Plakokefalos
Department of Medical Laboratories
Technological Educational Institute
Athens, Greece
plakoke@yahoo.com

Panagiotis Markoulatos[1] Efthimia Peteinaki[2]
1. Department of Biochemistry-Biotechnology,
2. Department of Medicine
University of Thessaly
Larissa, Greece

Alexandros Vontas,  Evaggelia Gramoustianou,
Kalliopi Georgantza, Anastasios Minas, Argyrios
Noulas, Vasileios Krikelis
Department of Medical Laboratories
Technological Educational Institute
Larissa, Greece
vkrikelis@teilar.gr

*Abstract*—The present study examines the genetic polymorphism of the hemagglutinin (HA) and neuraminidase (NA) gene segments of influenza type A viral strains isolated from Central Greece. Based on the molecular and phylogenetic analysis of these genes, both the pathogenicity and evolution of influenza viruses can be determined. The results from this study can significantly contribute to the better management and prevention of yearly influenza epidemics. The molecular analysis of epidemic viral strains will focus primarily on the HA1 region of the HA gene segment, which constitutes target of host neutralizing antibodies and part of the receptor binding site (RBS). Following culture propagation of viral strains in sensitive cell lines or embryonated eggs, antigenic analysis and identification of isolated viruses will take place. Isolated influenza strains from Central Greece will be compared to current and previous seasons influenza vaccine strains as well as influenza viral strains isolated from other regions of Greece during the last decades. Using RT-PCR protocols and DNA sequencing there will be determined the changes that have accumulated on the HA and NA genes which impact on virus epidemic potential and its ability to recognize and bind to host cellular receptors.

*Key words: Genetic analysis, influenza A virus, hemagglutinin, neurraminidase, influenza vaccine*

## I.   INTRODUCTION

Influenza viruses belong to the family Orthomyxoviridae and contain a single stranded segmented RNA molecule of negative polarity (-ss RNA). Based on antigenic differences on the nucleoprotein (NP) and matrix (M) proteins, influenza viruses are distinguished into three antigenic types: A, B and C. The first influenza virus was isolated in 1933 by Smith, Andrews and Laidlaw, and was named type A, while type B virus was isolated in 1940 and type C in 1947 [15, 16, 17, 19]. Types A and B are responsible for yearly influenza epidemics in humans. Type A viruses are further classified into antigenic subtypes and infect a series of animal hosts including migratory birds, horses, mink, etc. All recorded influenza pandemics have been associated with type A viruses. All viral subtypes have common NP and M antigens while they differ in their surface hemagglutinin (HA) and neuraminidase (NA) antigens, [1, 2].

Influenza viruses are usually pleiomorphic particles that contain an external host-derived membrane, are spherical in shape with a diameter of 50-120nm. The HA glycoprotein is coded by the 4th RNA segment of types A and B and binds to host cellular receptors which contain sialic acid (N-acetylneuraminic acid). The HA biological activity includes receptor recognition and viral membrane fusion with cellular membrane. The HA molecule is put under continuous pressure by the host immune system, resulting in antigenic variation in this molecule and emergence of viral variants against whom there is decreased immune coverage. The HA is a homotrimer molecule with a diameter of 135 A. Its head region is found on the outer region of the membrane and also contains a trans membrane region [20, 22]. The coding region (HAO), is cut by host photolytic enzymes into two subunits, the HA1 and HA2, which are connected via a disulfide bond. This cleavage is necessary for virus infectivity [3]. The head region of the HA includes the HA1 which contains the active center for sialic acid binding. The tail region is made up of the HA2 region and part of the HA1. The NA molecule is a membrane glycoprotein and its function includes the cutting off of neuraminic acid from glycoproteins and glycolipids. NA is a homotetramer, with a diameter of 60 A, complex and is not randomly dispersed on the viral surface, but is concentrated in small  regions of the virus membrane [11].

Type A viruses can undergo two types of mutation in their genome, the first  is the accumulation of point mutations and results in minor antigenic differences on the HA and NA

genes, defined also as antigenic drift, while the other includes gene segment reassortment and results in novel antigenic subtypes and is defined as antigenic shift [13, 18]. In types B and C viruses only antigenic drift takes place [3]. With each mutation at the amino acid level there exists the potential for alteration in the epitopes with high affinity for antibody and receptor recognition [1, 3]. This can result in either increased or decreased pathogenicity and evolution of disease. The above genetic mechanisms are based on the virus segmented genome and on the nature of the RNA molecule. Antigenic shift is a type of homologous recombination. During infection of the host by different viral strains this results in random reassortment of their gene segments and the formation of new HA and NA combinations or subtypes. The epidemiological success of influenza viruses is due mainly to two types of polymorphism in their external antigens of HA and NA. Such changes frequently result in novel antigenic variants, against which there is decreased immunity in the population resulting in yearly local or global epidemics or pandemics [21, 8, 12, 5].

The techniques of hemagglutination (HA test) and hemagluttination-inhibition (HAI), first described by G. Hirst in 1941 and 1944, respectively, are employed to detect and titer viral antigen, the former, and to identify the isolated strain, the latter. HAI provides the antigenic analysis against a pane of antisera for reference and vaccine strains. Low-reacting isolates are analyzed with an increased number of antisera for proper identification of antigenic drift on the HA molecule. Neuraminidase-inhibition (NAI), is employed to analyze the antigenic drift on the NA molecule. Both assays are based on the virus' affinity to recognize, attach and strongly bind to sialic acid (SA) and other oligosaccharide containing cellular receptors on target cells. All type A viruses recognize oligosaccharides which contain terminal SA as receptors. The receptor-binding site (RBS) in influenza is highly conserved, even among the HA's from different subtypes, including avian subtypes. The HA's of human viruses have affinity for SA which are bound with galactose via an $\alpha2,6$ bond. Influenza virus HA from birds and horses are bound with type $\alpha2,3$ Gal-SA bonds. In the case of H2 and H3 subtypes, amino acids at positions 226 and 228 of the HA1 region are responsible for this specificity in affinity. The HA of human influenza viruses contain a Leucine and a Serine at positions 226 and 228, respectively, while avian HA's contain a Gln and a Glycine in the respective amino acid positions. Previous studies have shown that a Leu to Gln substitution at position 226 can modify this specificity from f $SA\alpha2,6Gal$ to $SA\alpha2,3Gal$. A further modification of a Ser to Gly at position 228, reinforces the above modified specificity. Furthermore, changes at positions 145, 186 and 225 in the HA molecule of H1 and H3 viruses, can effect HA she cellular receptor [3, 5].

Influenza type A(H3N2) and other subtypes can be isolated from a number swine, horses, sea mammals, and especially wild migratory birds, the latter being the reservoir of all known influenza subtypes. In these birds there have been isolated all subtypes of HA (H1-H15). and NA (N1-N9), which proves that these viruses are endemic in these birds. Thus, genetic reassortment in bird species can result in the generation of novel influenza variant strains with a new subtype, as was the case of the H5N1 and H9N2 viruses. In

migratory birds is currently focused the research on the new influenza vaccine and continuous influenza surveillance by the WHO now includes molecular assays for the timely detection of RNA of viral origin. Recent phylogenetic studies of type A viruses have revealed viral genes whose origin is related to certain bird species and provided evidence for virus transmission among birds in the wild [1, 14].

The experimental determination of the linear sequence of bases in viral genomes is now used as the method for subtyping. Sequence determination also reveals mutation that has occurred during the replication process of the viral genomes. The DNA sequences that are analyzed by PCR and DNA sequencing are usually 1kb in length. All RNA genomes must first be converted to cDNA prior to PCR and sequencing reactions.

Especially for influenza viruses a somewhat recent application of this method aimed to genetically characterize emerging vial strains in the bird population, such as the H5N1 subtype from the lethal Hong Kong incidents in 1997 and 2003. The isolated virus was originated from a tracheal sample. The HA and NA genes were studied following RT-PCR in order to confirm the genotype. DNA sequencing revealed the presence polybasic amino acid mutation at the site of trypsin activity. This particular mutation was found in previously studied highly pathogenic viruses of avian origin and is believed that it contributes to systematic virus spread, allowing proteases similar to trypsin to cleave HA0 into HA1 and HA2 regions.

## II.     MATERIALS AND METHODS

**Clinical samples**: Nasal and pharyngeal swabs are collected from individuals presenting with influenza-like illness (ILI), 3-4 days following the onset of symptoms. Sampling takes place during the influenza season from October to April by health professionals in hospital clinics and private medical offices in Central Greece. Sample storage and transport is accomplished in virus transport medium (24,9 gr bacto-tryptose, 5 gr gelatin, 0,5 gr streptomycin, 500,000 U penicilin and 50 mg fungizone) per liter. Storage at $+4^0C$ for up to 7 days. Long-term storage of clinical samples and virus isolates takes place at $-70^oC$.

**Influenza virus isolation and identification**: Following an initial screening of all processed clinical samples, virus propagation is performed in both embryonated hen's eggs and cell culture. MDCK and HeLa cell lines are used as sensitive substrates for virus isolation. All samples are inoculated into the amniotic and allantoic cavities of 7 day old embryonated eggs. Incubation in eggs occurs at $33\text{-}35^oC$ for 48 hrs and in cells at $37^oC$ for 7-10 days. A second and more passages will be performed, if necessary, in order to increase the HA titer. Antisera will be provided by the World Health Organization (WHO), and used according to standard protocol. Viral HA antigen detection and titration was performed with an HA test using 0.5% avian or human red blood cells. Identification of isolated viral strains takes place using the HAI method, against a panel of antisera, according to WHO protocol.

**Immunocapture ELISA**: This assay was used for initial screening of samples. This in-house enzyme immunoassay detects the presence of influenza type A and B viral antigens in clinical samples and following propagation in culture. We use polyclonal antibodies raised against a variety of type A and B strains, in a modified indirect sandwich ELISA assay. Following initial incubation of the samples, we add diluted monoclonal antibody against the NP antigen of type A and B viruses, followed by a secondary antibody labeled with an enzyme. All reactions included positive and negative controls.

**Viral RNA isolation and cDNA synthesis**: A volume of 200-250 µl of sample or cultured virus is used in the RNA isolation protocol, (TRIzol, Gibco BRL) in the presence of 20 µg glycogen used as RNA carrier molecule. Reference influenza viral RNA A (H1N1), (H3N2) and type B are included as positive controls. The reaction mixture for the cDNA synthesis or reverse transcription step, included 5µl of RNA, 2,5µl random primer, (40µg/µl), 20U/µl Rnasin, 5µl 5X reaction buffer and 100U/tube reverse transcriptase enzyme. The integrity and quality of the isolated RNA was determined by spectroscopy and gel electrophoresis.

**Polymerase Chain Reaction (PCR):** In the PCR we use primers specific for the conserved regions of the M gene of types A and B viruses and subtypes H1, H2, H3 and N1, N2. A nested PCR protocol will be used for increased sensitivity [9, 10]. The first reaction includes 25 cycles and the second 40 cycles. Analysis of the PCR products is performed in 2% agarose gel in 1X TBE with 1 µg/ml bromidium ethide. Prior to sequencing reaction, all PCR products are cleaned in order to remove the primers and non-specific molecules, and analyzed in 1% agarose gel electrophoresis. Following PCR product cleansing the concentration DNA in band was determined by a DNA ladder.

**DNA sequencing of the HA1 region of HA and NA gene segments and phylogenetic analysis**

Sequencing reactions for both the HA1 and N gene segments included specific primer pairs and the reactions are performed in an automated sequencer. Molecular analysis performed at our laboratory on a number of influenza viruses isolated in Greece in previous years, for both the HA and NA gene segments was submitted to the influenza database and Genbank. DNA sequences were aligned with all relevant sequences in the influenza database, using the BLAST program. This program was used to align and compare our viruses with vaccine viruses. Following BLAST, all sequences were analyzed to determine their genetic relatedness in a phylogenetic analysis. The CLUSTAL X, program which is based on the neighbor-joining analysis was used to construct both the nucleotide and amino acid based phylogenetic trees [4].

## III.   RESULTS AND DISCUSSION

Molecular analysis of the HA and NA genes of archival influenza viruses, isolated isolated from previous influenza epidemics,  was performed and they will be compared with future isolates. These sequences were submitted to Genbank (Accession numbers: EU642547-EU642551, EU650318, AF316817- AF316821, EU835536-EU835538, EU502463, AF315570, AF315571 and AF315559-AF315569). More specifically, in antigenic sites A and B of the HA1 region of the HA, there were a Val at residue 124, a Lys at residue 133 and a Val at residue 134 and a Phe at residue 142 substitutions for the majority (96%) of our isolates. Continuous evolution of surface antigens hemagluttinin (HA) and neuraminidase (NA) of influenza viruses through the mechanisms of antigenic drift and shift is directly related with the epidemic potential of the virus during a certain season. A more effective yearly influenza surveillance requires among others, direct genetic analysis at of the HA and NA genes and identification of variant strains, including the early identification of antiviral resistance. Molecular analysis of the HA1 is now included in the routine analysis of influenza viruses in many virology laboratories[9,10,6,7,11,14]. This analysis should focus primarily on important antigenic sites of the HA1 which are involved in receptor recognition and binding and have affinity for the SA of rbc's from different animals [7]. RT-PCR and DNA sequencing provide maximum sensitivity and specificity for the timely detection of circulating influenza variants and identify strains that have developed resistance to antivirals. Molecular analysis of different gene segments of influenza viruses or even whole genome analysis provides a global understanding of influenza virus evolution and adaptation to new hosts and environments. The PCR products can be used to run a series of other molecular and diagnostic tests and to determine variation of the HA and NA antigens by performing analysis of the amino acid sequence. Virus isolation is the method of reference for vaccine and antiviral development by providing important antigenic information on isolated viral variants [18, 20].

Studies relating to the genetic variation of influenza viruses provide valuable information for the epidemic potential, the antigenic profile and evolutionary lineage, by phylogenetic analysis, for each viral strain under study. With DNA sequencing and at the amino acid level, changes at key amino acid positions crucial for virus infectivity and pathogenicity, are revealed. These amino acids become targets of neutralizing antibodies in the immune response. Point mutations on the HA and NA segments result primarily from the infidelity of the RNA polymerase enzyme and from host immune pressure[7,11]. These changes constitute the antigenic drift detected by the HAI method, but key substitutions for examples at sites 226 and 228 are revealed by molecular analysis. These amino acid changes can significantly alter the degree of influenza virus ability to recognize and bind to red cells from different animals thereby not being detected by the HA test, following virus growth in eggs or cells [5, 6]. Especially, in type A viruses this ability has been shown to be influenced by changes at amino acid position 186 and 226 which are found in the HA1 region of the HA. These

variations can alert the virus phenotype for causing disease and its long-term evolutionary potential.

## ACKNOWLEDGMENT

## REFERENCES

[1] Lamb, RA, Krug RM, "Orthomyxoviridae: The viruses and their replication' In 'Virology' (Fields, BN, Knipe, DM, and Howley PM., Eds.), pp. 1353-1395. Lipincott-Raven, Philadelphia, 1996.

[2] Cox, NJ, Subbarao K, 'Influenza' The Lancet. 354: 1277-1282, 1999

[3] Steinhauer DA, 'Role of the hemagglutinin cleavage for the pathogenicity of influenza virus', Virology 258:1-20, 1999.

[4] Higgins DG, Bleasby AJ, Fuchs R, 'CLUSTAL V: improved software for multiple sequence alignment', Cabios 8(2):189-191, 1992.

[5] Lin YP, Gregory V, Bennett M, Hay A, 'Recent changes among human influenza viruses', Virus Res 103:47-52, 2004.

[6] Isabelle Hardy, Yan Li, Michael B. Coulthart, Nathalie Goyette , Guy Boivin, 'Molecular evolution of influenza A/H3N2 viruses in the province of Quebec (Canada) during the 1997–2000 period, Virus Research 77, 89–96, 2001

[7] E. T. Plakokefalos, P. Markoulatos, N. Spyrou, T. A. Bei, N. C. Vamvakopoulos,'Molecular and phylogenetic analysis of haemagglutinin and neuraminidase sequences from recent human influenza type A (H3N2) viral isolates in Southern Greece', Arch Virol, 146: 1899–1918

[8] World Health Organization, 'Recommended composition of influenza vaccines for use in the 1998–1999 season', Wkly Epidemiol Rec 73: 56–65, 1998.

[9] Magnard C, Valette M, Aymard M, Lina B (1999). Comparison of two nested PCR, cell culture, and antigen detection for the diagnosis of upper respiratory tract infections due to influenza viruses. J Med Virol 59:215-220.

[10] Bressoud A, Whitcomb J, Pourzand C, Haller O, Cerruti P (1990). Rapid detection of influenza virus H1 by the polymerase chain reaction. Biochem Biophys Res Commun 167:425-430.

[11] Kilbourne DE, Johansson BE, Grajower B (1990). Independent and disparate evolution in nature of influenza A virus hemagglutinin and neuraminidase glycoproteins. Proc Natl Acad Sci USA 87:786-790.

[12] Monto AS (1997). Prospects for pandemic influenza control with currently available vaccines and antivirals. J Infect Dis 176:32-37.

[13] Both GW, Sleigh MJ, Cox NJ, Kendal AP (1983). Antigenic drift in influenza virus H3 hemagglutinin from 1968 to 1980: multiple evolutionary pathways and sequential amino acid changes at key antigenic sites. J Virol 48:52-60.

[14] Webster RG, Bean WJ, Gorman OT, Chambers TM, Kawaoka Y (1992).Evolution and ecology of influenza A viruses. Microbiol rev 56:152-179.

[15] Shope RE (1931). Swine influenza 1. Experimental transmission and pathology. Journal of experimental medicine. 54:349-360.

[16] Smith W, Andrews CH, Laidlaw PP (1933). A virus from influenza patient. Lancet, ii 666-668.

[17] Smith W, Stuart-Harris CH (1936). Influenza infection of man from the ferret. Lancet ii 121-123. Wiley DC, Skehel JJ (1987).

[18] Strauss JH, Strauss EG (1988). Evolution of RNA viruses. Ann Rev Microbiol 42: 666-699.

[19] Taylor RM (1949). Studies on survival of influenza virus between epidemics and antigenic variants. Am J public health 39:171-178.

[20] Wilson IA, Cox NJ (1990). Structural basis of immune recognition of influenza virus hemagglutinin. Ann rev Immunol 8:737-771.

[21] Elton D, Medcalf L, Bishop K, Harrison D, Digard P (1999). Identification of amino acid residues of influenza virus nucleoprotein essential for RNA binding. J Virol 73:7357-7367.

[22] Wiley DC, Skehel JJ (1987). The structure and function of the haemagglutinin membrane glycoprotein of influenza virus. Ann Rev Biochem 56:365-394.