# Phylogenetic Relationship of the H5N1 Viruses Isolated in Vietnam, 2012 Based on Sequencing Analyses of HA and NA Gene Segments

Tran Hoan Xuan, Vu Cuong Chi, and Luu Minh Quang

Abstract-Vietnam is one of the countries most severely affected by H5N1 highly pathogenic avian influenza (HPAI) virus in terms of poultry and human health. However, there is little information on the diversity of H5N1 viruses circulating all over the country. In this study, 24 H5N1 virus infections occurring in provinces along Vietnam, 2012 were investigated and the phylogenetic relationships of these viruses, based on HA and NA gene sequencing were analyzed. The results indicated the high variation of nucleotides in HA and NA gene segments of 24 H5N1 viruses isolated. All H5N1 viruses isolated from the North and the Middle of Vietnam were clustered into sub-clade 2.3.2.1. A Hubei-like viruses were identified A/Duck/Vietnam/LMB139/2012 similarity with strain (Genbank accession: AB742289.1), based on HA and NA phylogenetic tree analyses, respectively. Especially, virus collected in Bac Ninh province (Northern side) was recognized as sub-clade 2.3.2.1 B Barn Swallow-like viruses while all viruses collected from the South were belonged to clade 1.1, based on HA phylogenetic analysis. In addition, NA amino acid and nucleotide sequence analyses demonstrated that viruses isolated from the South were found to be similar with two published strains including A/Duck/Vietnam/OIE -3313/211 AB716339.1) A/Muscovy (Genhank accession: and Duck/Vietnam/OIE-3313/2011 (H5N1).

Index Terms-H5N1, phylogenetic, sequencing, virus.

## I. INTRODUCTION

The highly pathogenic avian influenza (HPAI) virus belonging to the A/H5N1 subtype was detected for the first time in geese in Guangdong province of China in 1996 [1]. Since that outbreak, H5N1 has spread among birds out of East Asia across Eurasia and as far west as England and West Africa, threatening further spread into the American and Australian continents. H5N1 is primarily a pathogen of poultry; more than 200 million poultry died or have been culled because of this virus [2]. HPAI/H5N1 viruses have now spread through poultry populations in many countries. These viruses have also crossed species barriers to infect different hosts [3]. Avian influenza viruses (AIV) are influenza A viruses with an up to eight-fold segmented single-stranded RNA genome of negative polarity. Influenza viruses hold generic status in the Orthomyxoviridae family and are classified into types A [4]. The main antigenic determinants of influenza A viruses are the haemagglutinin (HA) and the neuraminidase (NA) transmembrane glycoproteins, capable of eliciting subtype specific and

Manuscript received December 4, 2013; revised April 15, 2014.

immune responses which are fully protective within, but only partially protective across, different subtypes [5], [6].

Vietnam is one of the highest frequencies of HPAI H5N1. The outbreaks of HPAI H5N1 virus was first identified in Vietnam in 2001 [7], and outbreaks in poultry have been reported in more than 59 of the 64 Vietnamese provinces since December, 2003 (OIE, 2010). The first human infection in Vietnam was reported in 2004; by August of 2012, 123 cases and 61 deaths had been reported by WHO. Six different HA clades were identified and designated according to the recently described nomenclature system for the highly pathogenic H5N1 viruses as well as by potential precursors: clade 0, HK97-like (HK/483/97); clade 1, HK821-like (Dk/HK/821/02); clade 2.3.2, E319-like (Dk/China/E319-2/03); clade 2.3.4, FJ584-like (Ck/Fujian/584/05); clade 3, GX22-like (Dk/GX/22/01); clade 5, F1-like (swine/Fujian/F1/01) in Vietnam. Although the H5N1 viruses that circulated in Vietnam from 2003-2005 had clade 1 HA genes, many viruses isolated in 2007 had newly introduced clade 2.3.4 HA genes [8].

Like most other countries, Vietnam has stockpiled oseltamivir for use in the event of a pandemic. Furthermore, Vietnam has developed a prototype reverse geneticsgenerated human H5N1 vaccine derived from a clade 1 H5N1 influenza isolate (A/Vietnam/1194/2004) [9]. Between 2006 and 2011 in Vietnam, almost of poultry and waterfowl were injected H5N1 vaccines. However, the effective protection of H5N1 vaccine in Vietnam has decreased, leading to the outbreak in northern and central Vietnam in 2011 [10]. Understanding of diversity and evolution of H5N1 viruses will bring about effective treatments of diseases caused by H5N1 influenza virus. Therefore, in this study, we aim to sequence the HA and NA genes from viruses isolated in the North, the Middle and the South of Vietnam, 2012 in order to evaluate the nucleotide variation and construct phylogenetic relationship between these viruses.

#### II. MATERIALS AND METHODS

#### A. Sample Collection and Diagnostic Tests

Swab samples were collected from avian infected by H5N1 virus in provinces along Vietnam where the virus spreads significantly in 2012. The samples were tested for the presence of viral RNA using real-time reverse transcription polymerase chain reaction (rRT-PCR) for type A influenza and H5 subtype. Furthermore, the H5 virus was also identified by haemagglutination inhibition (HI) assay with H5-specific antiserum. These works were done in The

The authors are with the Key Laboratory of Animal Cell Technology – National Institute of Animal Sciences, Vietnam (e-mail: hoantranvcn@vahoo.com).

National Center for Veterinary Diagnostics (NCVD, Hanoi, Vietnam).

#### B. Virus Isolation, RNA Extraction and cDNA Synthesis

The pathogenic H5N1 viruses were propagated from the positive samples by infection of either 10-11 day old embryonated chicken eggs (ECEs) or MDCK cells. A total of 24 H5N1 viruses were isolated in 2012. Viral RNA was extracted, using QIAamp Viral Mini kit (QIAGEN) and subsequently, the protocol of Goscript Reverse Transcription System (Promega) was used to synthesize viral cDNA. All infectious viruses were handled in biosafety level 3 containment, including enhancements required by the U.S. Department of Agriculture.

## C. Nucleotide Sequencing of the Virus Isolated

Sequencing of the HA and NA gene segments were conducted on all 24 virus isolated. Polymerase chain reactions (PCRs) in 1x PCR buffer supplemented with 1.5 mM MgCl<sub>2</sub>, 200 µM dNTPs, 1 µM primers, 0.5 U GoTag Hot Start Polymerase (Promega) and 15 µl cDNA as templates were performed to amplify the HA and NA amplicons, using specific primers (Le TQM et al., 2008, PLoS ONE 3 (10): e3339. doi: 10.1371/ journal. pone. 0003339 - primer sequences available upon request). Conditions of the PCRs were 95°C for 2 min, followed by 40 cycles of 95°C for 30 sec, 60°C for 45 sec and 72°C for 30 sec and a final extension at 72°C for 5 min. PCR products were then extracted from an agarose gel using QIAquick Gel Extraction Kit (QUIAGEN), and each purified amplicon was used directly for cycle sequencing using BigDye Cycle v3.1 Sequencing Terminator Kit (Applied Biosystems), to the manufacturer's instructions. Post sequencing products were purified using BigDye XT Terminator Purification Kit (Applied Biosystems) prior to running on the ABI PRISM 3130 Genetic Analyzer (Applied Biosystems).

## D. Phylogenetic and Sequence Analysis

Sequencing Analysis software (v5.2.0) and Genetyx-Win software (v4.0) were used for raw sequence data assembly and editing. Virus gene sequences were aligned using ClustalW program within the Bioedit (v7.0.5.3) software to compare with representative H5N1 influenza A virus sequences that have been published and available on GenBank database. Multiple sequence alignments of the HA segment (about 1700 bp) and NA segment (about 616 bp) were used for phylogenetic analysis. To determine the evolution relationship of the virus, phylogenetic analysis was conducted using the Neighbour-Joining (NJ) method provided in the MEGA 5.1 software with 1000 bootstrap replicates. Clustering within H5N1 clades was investigated by pairwise analysis of both HA and NA sequence pairs between and within groups using the same MEGA programe.

## III. RESULTS AND DISCUSSION

## A. Amplification of HA and NA Amplicons

The presence of H5 HA RNA from the collected swab samples was detected by rRT-PCR (Fig. 1) and H5 virus was identified by haemagglutination inhibition (HI) assay with H5-specific antiserum (data not shown). The rate of positive samples ranged from 15 - 25 %, depending on the techniques and sampling methods used. All 24 samples that were identified as H5-positive by both rRT-PCR and HI assay were selected for virus isolation in ECEs or MDCK cells. After synthesis of cDNA from viral RNA extracted, the PCR was applied to amplify HA and NA amplicons (Fig. 2). PCR products separated in 1.5 % agarose gel show that the expected sizes of 1700 bp and 616 bp corresponding to HA and NA amplicons, respectively were well amplified.





Fig. 2. Amplificons of the HA (1-3) and NA (4-6) gene from cDNA synthesized from viral RNA isolated. DNA was size separated on 1.5% agarose gels, stained with ethidium bromide and photographed under UV-light. SL: smart ladder was used.

#### B. Nucleotide Sequence Analyses

After purified from an agarose gel, the amplicons was used for cycle sequencing as mentioned in the methods. The sequence data then were analyzed by Sequencing Analysis (v5.2.0), Genetyx-Win (v4.0) and Bioedit (v7.0.5.3) softwares.

To confirm the authentic HA and NA sequences, the similarity between the sequences studied and HA and NA published nucleotide sequences was compared, using the Basic Local Alignment Search Tool – BLAST (blast.ncbi.nlm.nih.gov). What BLAST results show the high consensus between compared sequences demonstrated the authentication of the HA and NA sequences studied (data not shown).

Sequencing analysis of HA gene segment amplified from 24 viruses isolated presents a high variation of nucleotides in comparison with HA gene sequence from strain A/goose/Vietnam/3/05 (H5N1) published previously. The variation regions were mainly located in 406-476, 639-720, 935-996, 1385-1443 sites of gene coding for HA (data not shown). Based on the nucleotide variations, 74 amino acid

substitutions were detected. Similarly, the high variation of nucleotides in 24 NA gene segments was also presented in comparison with NA gene sequence from strain A/Hong Kong/483/1997 (H5N1) that have been published. However, the variation regions were found to be located only in 445-756 site of NA coding gene HA (data not shown). Subsequently, 30 amino acid substitutions were reported, including X158V, T169I, I173T, I176V, G182E, K198R,

M205L, Y218S, T213A, L222V, D229N, <u>G230E</u>, N246K, I285V, R294Q, T295I, M295I, E234Y, S336N, M238K, <u>K243R</u>, I244V, I346T, D250L, <u>I270T</u>, <u>N306S</u>, <u>K313T</u>, <u>P321L</u>, <u>H347S</u>, <u>X363E</u>, in which the underlines (12 substitutions) mark the specific differences between H5N1 viruses isolated in Vietnam, 2012 and the comparative strains that have been reported and the rest (18 substitutions) are newly found.



Fig. 3. Phylogenetic relationship of the HA gene segment from 24 H5N1 viruses isolated in Vietnam, 2012.

The HA phylogenetic trees were constructed by the NJ method provided in the MEGA 5.1 software with 1000 bootstrap replicates. Bootstrap values under 50% are not shown. Tree rooted at the clade 0 node A/goose /Vietnam/3/2005 (H5N1).

## C. Phylogenetic Relationship

To analyze the phylogenetic relationship between H5N1 viruses isolated in Vietnam, 2012 and the H5N1 viruses published by NCBI and WHO, based on the HA and NA sequence data, the Neighbour-Joining (NJ) method provided in the MEGA 5.1 software with 1000 bootstrap replicates was applied.



0.04 0.03 0.02 0.01 0.00

Fig. 4. Phylogenetic relationship of the NA gene segment from 24 H5N1 viruses isolated in Vietnam, 2012.

The NA phylogenetic trees were constructed by the NJ method provided in the MEGA 5.1 software with 1000 bootstrap replicates. Bootstrap values under 50% are not shown. Tree rooted at the clade 0 node A/Hong Kong/483/1997 (H5N1). Mien Bac = the North of Vietnam; Mien Nam = the South of Vietnam; Mien Trung = the Middle of Vietnam.

The phylogenetic analysis of the HA gene (Fig. 3) showed that all the H5N1 viruses in the present study differ

from the ancient strains such as A/Chicken/Vietnam/NCVD-016/2008 (H5N1), A/goose/Vietnam/113/2001 (H5N1), A/Goose/Guangdong/1/96 (H5N1). The result also indicated that most of H5N1 viruses isolated from both the North and the Middle of Vietnam belonged to sub-clade 2.3.2.1 A Hubei-like viruses while only viruses isolated from infected ducks in Bac Ninh province (Northern side) were recognized to be sub-clade 2.3.2.1 B Barn Swallow-like viruses. In the South of Vietnam, all viruses isolated were clustered into clade 1.1.

In general, in each group, the viruses shared 95-98% nucleotide similarities in HA gene segment analyzed and 90-95% in the deduced amino acid sequences.

Similarity, the phylogenetic analysis of the NA gene (Fig. 4) indicated that all the H5N1 viruses studied differ from the ancient strains such as A/HK/483/97 (H5N1) and A/Dk/HK/821/02 (H5N1). Since both nucleotide and amino acid sequence identities were high (96-100%), indicating the genetic homogeneity, the H5N1 viruses isolated from the North and the Middle of Vietnam can be classified into the same sub-lineage closed to strain A/Duck/Vietnam/LMB139/2012 (Genbank accession: AB742289.1). Based on NA phylogenetic relationship analysis, viruses isolated from the South of Vietnam were mainly recognized as strain A/Duck/Vietnam/OIE 3313/211 (Genbank accession: AB716339.1), in which viral samples (number 9, 10, 11, 13, 14, 16) collected in Hau Giang, Bac Lieu, Dong Thap, Soc Trang share a high similarity with strain A/Muscovy Duck/Vietnam/OIE-3313/2011 (H5N1).

## IV. CONCLUSIONS

In this study, the high variation of nucleotides in HA and NA gene segments of 24 H5N1 viruses isolated in Vietnam, 2012 were reported and consequently, 74 and 30 amino acid substitutions in the HA and NA peptide sequences, respectively were detected. All H5N1 viruses isolated from the North and the Middle of Vietnam were clustered into sub-clade 2.3.2.1 A Hubei-like viruses and were identified similarity with A/Duck/Vietnam/LMB139/2012 strain (Genbank accession: AB742289.1), based on HA and NA phylogenetic tree analyses, respectively. Especially, virus collected in Bac Ninh province (Northern side) was recognized as sub-clade 2.3.2.1 B Barn Swallow-like viruses while all viruses collected from the South were belonged to clade 1.1, based on HA phylogenetic analysis. In addition, NA amino acid and nucleotide sequence analyses demonstrated that viruses isolated from the South were found to be similar with two published strains including A/Duck/Vietnam/OIE -3313/211 (Genbank accession: AB716339.1) A/Muscovy Duck/Vietnam/OIEand 3313/2011 (H5N1).

#### REFERENCES

- World Health Organization, "H5N1 avian influenza: Timeline of major events," Report, 13 December 2011.
- [2] Food and Agricultural Organization, *Bird Flu Outbreaks Continue in Africa, Asia, Europe and Near East*, Geneva: FAO, 2006.
- [3] Y. K. Choi, T. D. Nguyen, H. Ozaki, R. J. Webby, P. Puthavathana *et al.*, "Studies of H5N1 influenza virus infection of pigs by using viruses isolated in Vietnam and Thailand in 2004," *J. Virol*, vol. 79, pp. 10821–10825, 2005.

- [4] G. Cattoli, A. Fusaro, I. Monne, and I. Capua, "H5N1 Virus Evolution in Europe — An Updated Overview," *Viruses*, vol. 1, pp. 1351-1363, 2009.
- [5] R. A. Fouchier, V. Munster, A. Wallensten *et al.*, "Characterization of a novel influenza A virus hemagglutinin subtype (H16) obtained from black-headed gulls," *J Virol*, vol. 79, pp. 2814-2122, 2005.
  [6] World Health Organization/World Organisation for Animal
- [6] World Health Organization/World Organisation for Animal Health/Food and Agriculture Organization H5N1 Evolution Working Group. (2008). Toward a unified nomenclature system for highly pathogenic avian influenza virus (H5N1). *Emerg. Infect. Dis.* (serial on the Internet). [Online]. Available: http://www.cdc.gov/EID/content/14/7/e1.htm
- [7] D. C. Nguyen, T. M. Uyeki, S. Jadhao, T. Maines, M. Shaw *et al.*, "Isolation and characterization of avian influenza viruses, including highly pathogenic H5N1, from poultry in live bird markets in Hanoi, Vietnam," *J. Virol*, vol. 79, pp. 4201–4212, 2001.
- [8] X.-F. Wan, T. Nguyen, C. T. Davis, C. B. Smith, and Z.-M. Zhao, "Evolution of highly pathogenic H5N1 avian influenza viruses in Vietnam between 2001 and 2007," *PLoS ONE*, vol. 3, no. 10, p. e3462, 2008.

- [9] M. T. Q. Le, H. F. L. Wertheim, H. D. Nguyen, W. Taylor, P. V. M. Hoang *et al.*, "Influenza A H5N1 Clade 2.3.4 virus with a different antiviral susceptibility profile replaced clade 1 virus in humans in Northern Vietnam," *PLoS ONE*, vol. 3, no. 10, p. e3339, 2008.
- [10] Cucthuy. [Online]. Available: http://www.cucthuy.gov.vn



**Tran Hoan Xuan** was born on January 26, 1960 in Vietnam. He received his PhD degree from College of Natural Science, Hanoi National University in 2004, and diploma's degree from Kharcop National University, The Soviet Union in 1983. He is now specialized in molecular biology. He has been a researcher in the Department of Biochemistry -National Institute of Animal Sciences (NIAS) during 1984-2001, and the deputy

head of Department of Biochemistry (NIAS) during 2001-2007, the deputy head of Key Laboratory of Animal Cell Technology (NIAS) during 2007-2009. He is now the head of Key Laboratory of Animal Cell Technology (NIAS) since 2010.