The Journal of Microbiology (2011) Vol. 49, No. 2, pp. 274-279
DOI 10.1007/s12275-011-0311-2


# Genome Sequence Analysis of H5N1 Influenza A Virus Isolated from a Vietnamese in 2007[§]

**Dieu Linh Tran[1†], Kangmo Kim[2,5†], Jae Yoo Choi[1], Hyun Dong Paik[1], Si-Woo Choi[3], Jin Yeul Ma[4], Sung-Soon Kim[2], Sung Joon Ahn[5*], and Young Bong Kim[1*]**

*[1]College of Animal Bioscience and Technology, Konkuk University, Seoul 143-701, Republic of Korea*
*[2]Center for Immunology and Pathology, Korean National Institute of Health, Seoul 122-701, Republic of Korea*
*[3]Department of Genetic Engineering, Faculty of Life Science and Technology, Sungkyunkwan University, Suwon 440-746, Republic of Korea*
*[4]Center for Herbal Medicine Improvement Research, Korea Institute of Oriental Medicine, Daejeon 305-811, Republic of Korea*
*[5]School of Information and Communication Engineering, Sungkyunkwan University, Suwon 440-746, Republic of Korea*

(Received August 10, 2010 / Accepted December 6, 2010)

**Highly pathogenic H5N1 avian influenza A virus (AIV) crossed the species barrier and caused a number of deaths in humans in Vietnam and 14 other countries. Since the last report of human H5N1 infection in November 2005, the first documented H5N1 human infection was reported in June 2007 in Vietnam and was followed by 7 more cases, including 5 fatalities. In this study, we isolated and analyzed the full length of the H5N1 genome from a sample from the first patient in 2007. Phylogenetic analysis of eight genomic segments of the H5N1 virus strain (A/Vietnam/HN/2007, VNH07) revealed that this strain appears to be of genotype V and contains the HA gene, which is classified into clade 2.3.4. The deduced amino acid sequence of the HA protein has a typical affinity sequence for α2,3 linkage (SAα2,3-Gal) receptors and typical multibasic cleavage sequences. Compared with other H5N1 isolates, VNH07 showed that the possible reassortments for the NA and NP segments occurred between A/goose/Guangxi/3017/2005-like isolates (2.3.2) and A/human/Zhejiang/16/2006-like isolates (2.3.4).**

***Keywords*:** avian influenza, H5N1, human infection, phylogenetic analysis, reassortment

Influenza A viruses belong to the *Orthomyxoviridae* family, containing 8 segmented, negative-sense RNAs that encode 10 proteins. They are classified into 16 H subtypes and 9 N subtypes based on the nature of two surface glycoproteins: hemagglutinin (HA) and neuraminidase (NA), respectively.

Since the first human case of H5N1 was reported in Hong Kong (De Jong *et al.*, 1997; Claas *et al*., 1998), the number of cases increased to 18 in 1999, raising concerns about a new influenza pandemic. Four years later, human tropic avian influenza re-emerged, leading to outbreaks in 2003. According to the cumulative number of confirmed human cases of H5N1, as reported by the WHO, highly pathogenic avian influenza A (H5N1) viruses have infected 502 people and caused 298 deaths in 15 countries. Indonesia ranks first with 167 infections (138 deaths) and Vietnam occupies the second rank with 119 cases (59 deaths). The number of human Vietnamese H5N1 cases has increased rapidly during the first epidemic period between 2004 and 2005. Since then, no additional cases have been reported up to the first half of 2007. On 29 June 2007, two new human cases of avian influenza A (H5N1) were confirmed. By the end of 2007, the total number of human cases in Vietnam was reported to be 8, including 5 fatalities (WHO, 2010).

The highly pathogenic avian H5N1 virus is a panzoonotic in birds and has shown the ability to infect other species. Fortunately, it has no ability for human to human transmission. However, a recent outbreak of novel influenza A/H1N1 virus (now known as 2009 H1N1 influenza) showed high infectivity among humans (Taubenberger and Morens, 2010). It is not surprising that we are considering the question of whether the H5N1 virus and the H1N1 virus could reassort to create a highly pathogenic human to human transmissible H5N1 virus.

Elucidation of the genetic characteristics of the newly identified human H5N1 virus is important. In order to clarify the origin of the VNH07 strain, we analyzed the full genome of an H5N1 strain isolated from a patient in Vietnam in 2007.

## Materials and Methods

### Virus

The A/Vietnam/HN1/2007 virus strain (VNH07) was isolated from throat and nasal swabs of a 28-year-old female patient in Ha Nam Province, Vietnam. She developed symptoms, including high fever (40°C), sore throat, and double pneumonia. She was admitted to the hospital on 6 June 2007 and died 15 days later. This case was confirmed as an H5N1 infection by the National Institute of Hygiene and Epidemiology, Vietnam, and WHO according to the amended WHO criteria.

### Full-length cDNA synthesis and sequencing

First-strand cDNAs of VNH07 were kindly provided by the Vietnam National Institute of Hygiene and Epidemiology, Hanoi, Vietnam.

† These authors contributed equally to this work.
* For correspondence. (S.J. Ahn) E-mail: finger@skku.edi; Tel: +82-31-290-7977 / (Y.B. Kim) E-mail: kimera@konkuk.ac.kr; Tel: +82-2-450-4208; Fax: +82-2-450-3667
§ Supplemental material for this article may be found at http://www.springer.com/content/120956

Using specific primer pairs, the two genes coding for glycoproteins, HA and NA, were generated for further study. Double strand cDNAs of the other six gene segments were synthesized by PCR. using a universal primer set. All PCR-generated gene segments were sequenced and deposited in GenBank (EU294369-EU294373; GQ232439-GQ232444).

### Phylogenetic analysis

Sequences of the eight gene segments were aligned with both HPAIV and H5N1 representatives from Vietnam, China, Japan, and Korea using the MEGA4 program (Tamura *et al.*, 2007). Evolutionary history was inferred using the Neighbor-Joining method with 1,000 bootstrap replicates. Phylogenetic trees were constructed using a Maximum composite Likelihood method. Protein sequences coding for the eight proteins of VNH07 were analyzed for genetic and molecular features.

## Results and Discussion

### Phylogenetic analysis of eight gene segments

VNH07 is the first H5N1 AIV human isolate found in Vietnam since November 2005. Phylogenetic analysis of all eight segments (Figs. 1-2, Supplementary data Figs. 1-6) re-

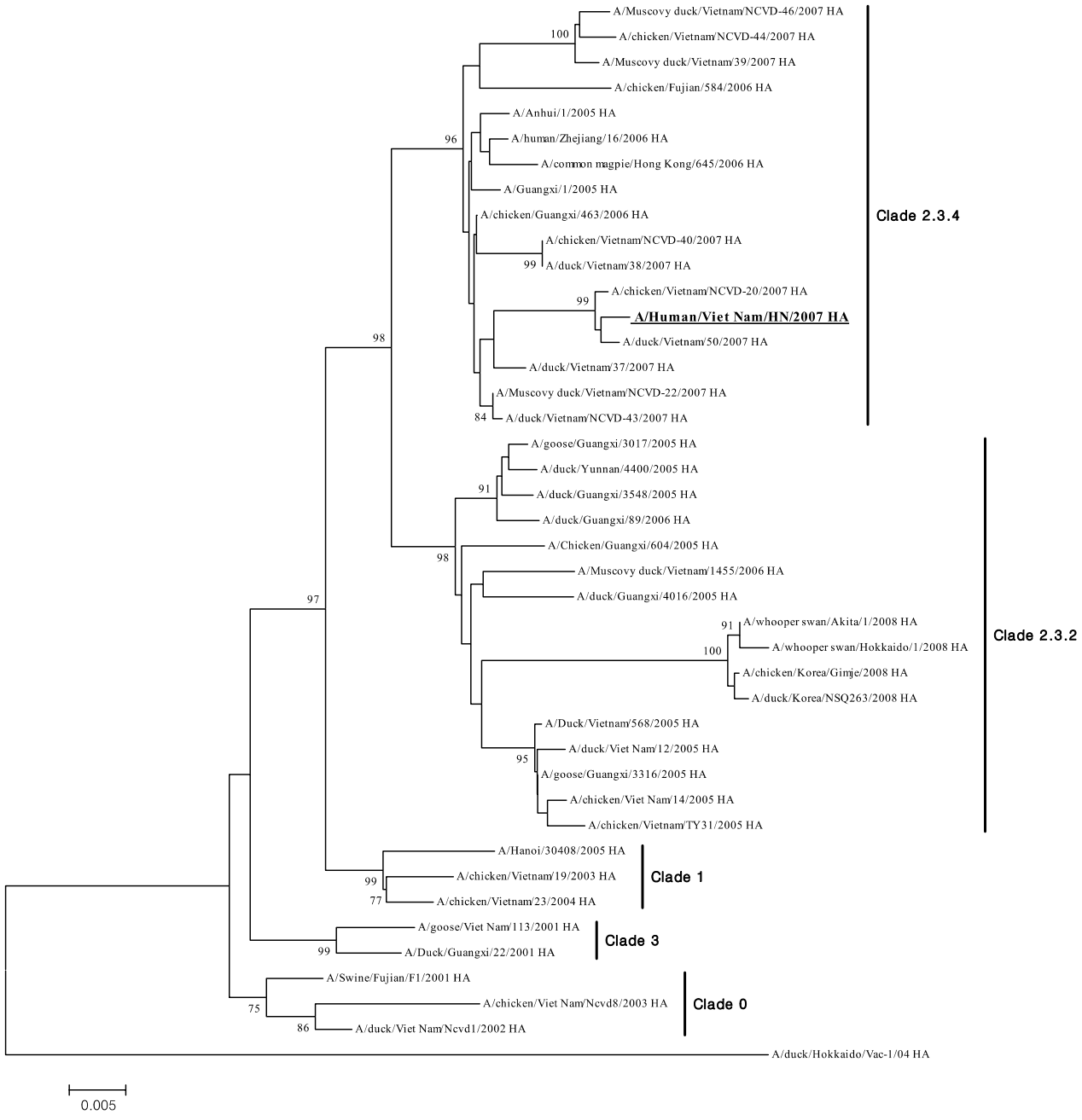


**Fig. 1.** Phylogenetic distribution of the hemagglutinin gene (HA) of VNH07 in the unified nomenclature system for highly pathogenic H5N1 avian influenza. The tree is rooted to A/duck/Hokkaido/Vac-1/04. VNH07 is underlined and in bold.

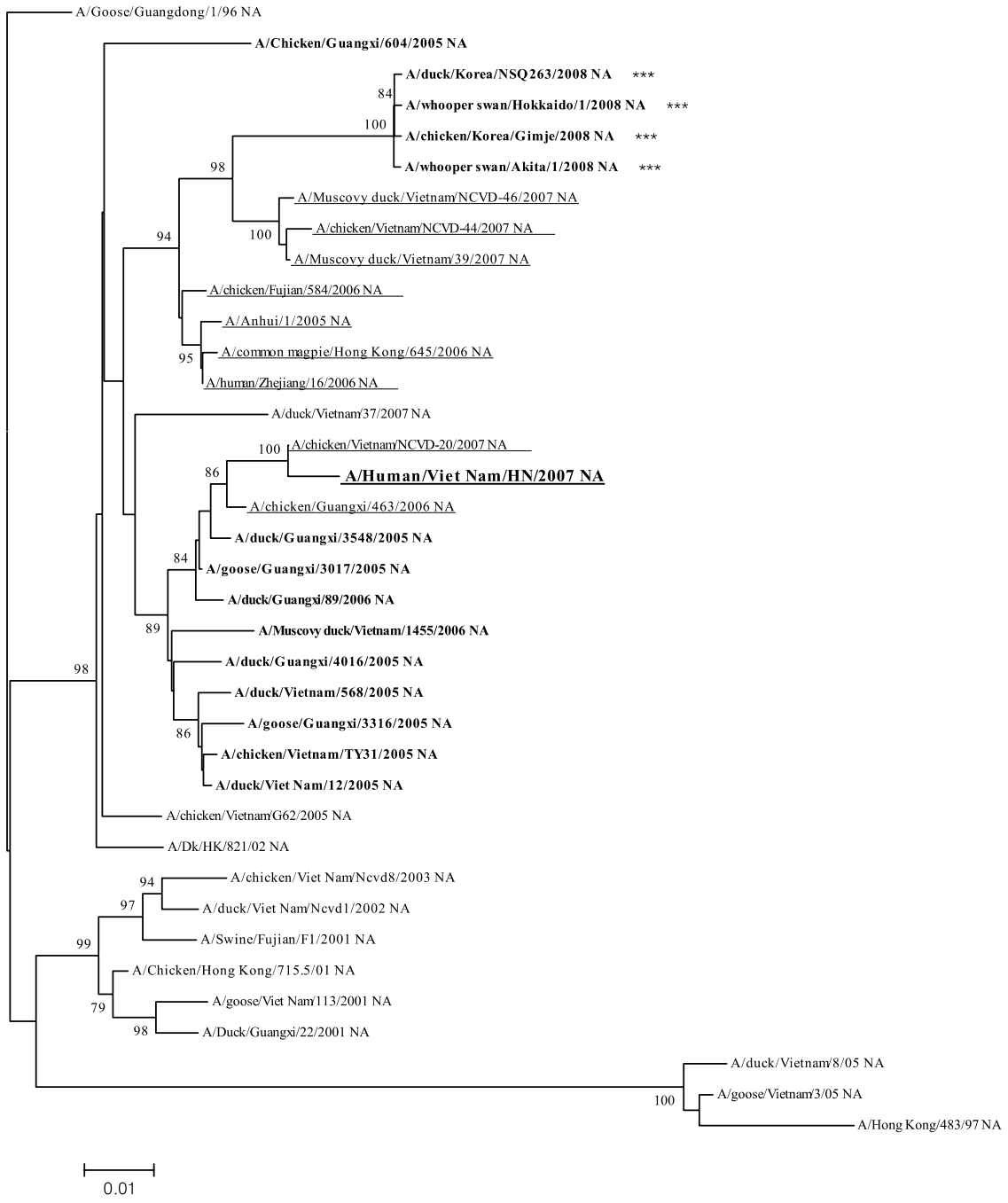


**Fig. 2.** Phylogenetic tree of the neuraminidase (NA) gene of VNH07. VNH07 is underlined and in bold. Korean and Japanese isolates are denoted by three asterisks. All clade 2.3.2 isolates are in bold characters, while clade 2.3.4 isolates are underlined.

vealed that VNH07 is most closely related to the A/duck/Vietnam/50/2007 and A/chicken/Guangxi/463/2006 strains with 99% and 98% similarity, respectively (Table 1). While the NA and NP segments of VNH07 indicate that the virus is related to clade 2.3.2 strains, HA and other internal genes were clustered within clade 2.3.4. In addition to evidence of the patient's direct contact with chickens and living in Ha Nam province, northern Vietnam, the transmission route appears to have been via exposure to H5N1-infected poultry from southern China. Among Chinese isolates, A/chicken/Guangxi/463/2006 is the strain most closely related to VNH07, in accordance with another finding (Nguyen *et al.*, 2008) showing that clade 2.3.4 H5N1 virus was introduced into northern Vietnam from southern China in 2007.

Regarding genotype, genotype V viruses share all of their genes with genotype Z viruses, except for PA, which was thought to be a reassorted form of genotype Z with other aquatic avian viruses. VNH07 and several other Vietnamese

**Table 1.** Nucleotide similarity among VNH07 and other isolates

| | | A/Vietnam/HN/2007 | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Clade | segment | PB2 | PB1 | PA | HA | NP | NA | M | NS |
| 2.3.4 | VNA-07 | **99%** | **99%** | **99%** | **99%** | **99%** | **99%** | **99%** | **99%** |
| | GX-463 | **98%** | **98%** | **98%** | **98%** | **99%** | **98%** | **99%** | **98%** |
| | Zhe-16 | 96% | **98%** | **98%** | **98%** | 97% | 96% | **98%** | 97% |
| 2.3.2 | KRA-08 | 97% | **98%** | 97% | 95% | 97% | 93% | **98%** | **98%** |
| | Hok-01 | 97% | 97% | 97% | 95% | **98%** | 93% | **98%** | **98%** |
| | GX-3017 | 91% | 96% | 92% | 96% | **98%** | **98%** | **98%** | 95% |
| 1 | VNA-04 | 96% | 97% | 92% | 95% | 97% | 95% | **98%** | 97% |
| | VNA-05 | 96% | 96% | 92% | 95% | 97% | 95% | 97% | 96% |

VNA-07, A/duck/Vietnam/50/2007; GX-463, A/chicken/Guangxi/463/2006; Zhe-16, A/human/Zhejiang/16/2006; KRA-08, A/duck/Korea/NSQ263/2009; Hok-01, A/whooper swan/Hokkaido/1/2008; GX-3017, A/goose/Guangxi/3017/2005; VNA-04, A/chicken/Vietnam/23/2004; VNA-05, A/human/Vietnam/CL105/2005. Highly homologous segments with VNH07 (more than 98%) are in bold characters.

avian isolates share seven out of eight genes (PB1, PB2, HA, NA, NP, M, and NS) with genotype Z strains, whereas their PA sequences are related to those of genotype V isolates (Supplementary data Fig. 1) (Vijaykrishna *et al.*, 2008).

## Deduced amino acid sequence analysis

**Hemagglutinin cleavage sequence:** H5N1 isolates in various clades have a multibasic motif in HA cleavage sequences, a feature of HPAIV. All isolates in clade 2.3.4, including VNH07, share the "PLRERRRK-R" pattern, which differs from the other clades. Although one Lysine deletion was observed in this motif, no information has been reported regarding the effect of such a deletion on cleavage ability or on the pathogenicity of the H5N1 virus. This can be explained by the fact that although the cleavage site bears the gap, its adjacent motif still retained the polybasic nature and, consequently, also retained the high level of pathogenicity (Li *et al.*, 2004).

**Receptor binding sites:** Without a doubt, receptor binding sites play a pivotal role in determination of the host tropism of AIV. For influenza A viruses, a switch in receptor specificity from α2-3 linkages (preferred by avian influenza A viruses) to α2-6 linkages (preferred by human influenza A viruses) can be considered a prerequisite to crossing the species barrier and attaining the ability to initiate a human pandemic (Yamada *et al.*, 2006). The receptor binding site-related residues: 91Y, 130-134 (GVSS), 149W, 151I, 179H, 186E, 190L, 191Y, and 220-224 (NGQSG) are conserved in VNH07, especially positions 222Q and 224G, which are presumably responsible for preferential binding of AIV to the SAα2,3-Gal receptor (Matrosovich *et al.*, 1997; Vines *et al.*, 1998; Ha *et al.*, 2001).

**Glycosylation sites:** Six potential glycosylation sites were detected on HA1 and two sites were identified on the HA2 domain of VNH07 hemagglutinin. In contrast to the isolates of clade 2.3.2, VNH07 clades 2.3.4 acquired N-linked glycosylation at residues 154-156, which was predicted to reduce the affinity for the sialic receptor (data not shown) (Matrosovich *et al.*, 1997). Although N-linked addition lowers the affinity between virus and sialyloligosaccharide receptors, it may provide influenza viruses with increased ability to escape from neutralizing antibodies and decreased ability to elicit neutralizing antibodies from the host immune system (Schulze, 1997; Kaverin *et al.*, 2002; Li *et al.*, 2004).

## Neuraminidase (NA) and other proteins

The NA protein of VNH07 has a 20-amino acid deletion at aa 49 to 68 in the stalk region (Fig. 3). A 20-aa deletion retained in the NA protein may alter the efficiency of enzymatic cleavage of various substrates, which may be the reason for a shift in host tropism (Luo *et al.*, 1993; Zhou *et al.*, 2009). Moreover, it was reported that such a deletion could not only increase retention of virions at the plasma membrane, but

| H5N1 Isolates | Amino Acid Squences | | | | |
|---|---|---|---|---|---|
| | 40 | 50 | 60 | 70 | 80 |
| A_duck_VietNam_Ncvd1_2002 | IQTGNQHQAEPCNQSIITYENNTWVNQTYVNISNTNFLTEKT | | | | |
| A_chicken_VietNam_Ncvd8_2003 | IQTGNQHQAGPCNQSIITYENNTWVNQTYVNISNTNFLTEKT | | | | |
| A_Ck_Indonesia_BL_2003 | IQTGNQHQAES----------ISNTNPLTEKA | | | | |
| A_chicken_Vietnam_27_2003 | IHTGNQHQAEP----------ISNTNFLTEKA | | | | |
| A_Ck_VietNam_38_2004 | IHTGNQHQAEP----------ISNTNFLTEKA | | | | |
| A_chicken_Vietnam_23_2004 | IHTGNQHQAEP----------ISNTNFLTEKA | | | | |
| A_duck_Vietnam_543_2005 | IHTGNQHQAEP----------ISNTNFLTEKA | | | | |
| A_goose_Guangxi_3017_2005 | IQTGNQNQVEP----------ISNTNFLTEKA | | | | |
| A_human_Zhejiang_16_2006 | IQTGNQHQAEP----------IRNTNFLTENA | | | | |
| **A_Human_Vietnam_HN_2007** | IQTGNQNQVEP----------IINTNFLTEKA | | | | |
| A_whooperswan_Hokkaido_2008 | IQTGNQHQAEP----------IRNTNFLTENA | | | | |
| A_duck_Korea_NSQ263_2008 | IQTGNQHQAEP----------IRNTNFLTENA | | | | |

**Fig. 3.** Alignment comparisions of deduced amino acid sequences of Neuraminidase Stalk-Motifs. The dash means the amino acid was deleted at that site.

might also be involved in adaptation of H5N1 from wild aquatic birds to domestic chickens as viral hosts (Matrosovich *et al.*, 1997).

Like other Southeast Asian isolates, excluding clade 0 isolates, VNH07 contains a five-amino-acid deletion in the nonstructural (NS) protein. Such deletions in NA and NS proteins are properties of genotype V (Vijaykrishna *et al.*, 2008), supporting our finding that VNH07 is a genotype V virus.

**Reassortment**

While the NA and NP segments of VNH07 indicate that the virus is related to clade 2.3.2 strains, HA and the other internal genes cluster within clade 2.3.4. To assess the possible reassortment in HP H5N1 AIVs, we compared the sequence homology with regional H5N1 isolates (Table 1). As expected, the phylogenetic data showed that VNH07 is closely related to the A/duck/Vietnam/50/2007 and A/chicken/Guangxi/463/2006 strains with more than 98% similarity in all eight segments. Compared with other H5N1 isolates, VNH07 showed that possible reassortments for the NA and NP segments occurred between the A/goose/Guangxi/3017/2005 (2.3.2)-like NA and NP and the A/human/Zhejiang/16/2006 (2.3.4)-like remaining segments, except PB2. Therefore, it is also possible that the reassortment between clades 2.3.2 and 2.3.4 appeared first in China, prior to introduction of clade 2.3.4 into Vietnam. The distinctive difference between the Vietnamese 2005 strain and VNH07 supports that VNH07 most likely originated in China, rather than in circulating Vietnamese domestic sources, in which clade 1 viruses are dominant (Nguyen *et al.*, 2008).

Of particular interest, the 2008 Korean strains appeared to be a reassortment between clade 2.3.2-like HA and clade 2.3.4-like remaining segments. Six internal genes of Korean and Japanese strains, including PA, PB1, PB2, M, NP, and NS, share greater than 97% similarities with those of VNH07 (Table 1), whereas a large divergence in HA and NA genes exists among them. It was reported that these 2008 Korean strains may derive from a reassortment between two groups of HPAI H5N1 viruses (clade 2.3.2 and 2.3.4) isolated from southeast Asia via wild bird migrations rather than from domestic fowl (Kim *et al.*, 2010). In addition, Japan, Korea, Vietnam, and some parts of China (including Guangxi) lie in the East Asia/Australia bird migration flyway (Boere and Stroud, 2006). Therefore, it is possible that viruses from China spread to southeast Asia or east Asia and that they are circulated by migrating birds.

Genetic reassortment affects the infectivity and host range in avian influenza viruses. Although no human case of clade 2.3.2 has been reported, clade 2.3.4 viruses introduced to Vietnam have infected humans, with a high fatality rate (Le *et al.*, 2008). Virologists are concerned about the potential for reassortant of a HPAIV H5N1 with a 2009 H1N1 influenza virus, which could result in a new highly pathogenic pandemic strain. To prepare for the pandemic potential and to reduce concern, we will need to compare more of the genetic characteristics with all of the human HPAIV isolates.

## Acknowledgements

This work was supported by: a Top Brand Project grant from the Korea Research Council of Fundamental Science & Technology and KRIBB Initiative program (KGM3110912), the BioGreen 21 Program in Rural Development Administration, an Agricultural R&D Promotion Center (20070501034009), the Korea Institute of Oriental Medicine, Republic of Korea (K09040) and Konkuk University. We also thank Mai Phuong Hoang Vu and Khanh Hang Nguyen Le, the National Institute of Hygiene and Epidemiology, Hanoi, Vietnam, for isolating the virus. Dieu Linh Tran received a foreigner scholarship from the National Research Foundation of Korea.

## References

Boere, G.C. and D.A. Stroud. 2006. The flyway concept: what it is and what it isn't, pp. 40-47. The Stationery Office, Edinburgh, UK.

Claas, E.C., A.D.M.E. Osterhaus, R.V. Beek, J.C. De Jong, G.F. Rimmelzwaan, D.A. Senne, S. Krauss, K.F. Shortridge, and R.G. Webster. 1998. Human influenza A H5N1 virus related to a highly pathogenic avian influenza virus. *Lancet* 35, 472-477.

De Jong, J.C., E.C. Claas, A.D. Osterhaus, R.G. Webster, and W.L. Lim. 1997. A pandemic warning? *Nature* 389, 554.

Ha, Y., D.J. Stevens, J.J. Skehel, and D.C. Wiley. 2001. X-ray structures of H5 avian and H9 swine influenza virus hemagglutinins bound to avian and human receptor analogs. *Proc. Natl. Acad. Sci. USA* 98, 11181-11186.

Kaverin, N.V., I.A. Rudneva, N.A. Ilyushina, N.L. Varich, A.S. Lipatov, Y.A. Smirnov, E.A. Govorkova, A.K. Gitelman, D.K. Lvov, and R.G. Webster. 2002. Structure of antigenic sites on the haemagglutinin molecule of H5 avian influenza virus and phenotypic variation of escape mutants. *J. Gen. Virol*. 83, 2497-2505.

Kim, H.R., C.K. Park, Y.J. Lee, G.H. Woo, K.K. Lee, J.K. Oem, S.H. Kim, and *et al*. 2010. An outbreak of highly pathogenic H5N1 avian influenza in Korea, 2008. *Veterinary Microbiology* 141, 362-366.

Le, M.T.Q., H.F.L. Wertheim, H.D. Nguyen, W. Taylor, P.V.M. Hoang, C.D. Vuong, H.L.K. Nguyen, and *et al*. 2008. 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* 3(10), e3339. doi:10.1371.

Li, K.S., Y. Guan, J. Wang, G.J. Smith, K.M. Xu, L. Duan, A.P. Rahardjo, and *et al*. 2004. Genesis of a highly pathogenic and potentially pandemic H5N1 influenza virus in eastern Asia. *Nature* 430, 209-213.

Luo, G., J. Chung, and P. Palese. 1993. Alterations of the stalk of the influenza virus neuraminidase: deletions and insertions. *Virus Res*. 29, 141-153.

Matrosovich, M.N., A.S. Gambaryan, S. Teneberg, V.E. Piskarev, S.S. Yamnikova, D.K. Lvov, J.S. Robertson, and K.A. Karlsson. 1997. Avian influenza A viruses differ from human viruses by recognition of sialyloligosaccharides and gangliosides and by a higher conservation of the HA receptor-binding site. *Virology* 233, 224-234.

Nguyen, T.D., T.V. Nguyen, D. Vijaykrishna, R.G. Webster, Y. Guan, J.S.M. Peris, and G.J.D. Smith. 2008. Multiple sublineages of Influenza A Virus (H5N1), Vietnam, 2005-2007. *Emerg. Infect Dis.* 14, 632-636.

Schulze, I.T. 1997. Effects of glycosylation on the properties and functions of influenza virus hemagglutinin. *J. Infect. Dis.* 176, 24-28.

Tamura, K., J. Dudley, M. Nei, and S. Kumar. 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* 24, 1596-1599.

Taubenberger, J.K. and D.M. Morens. 2010. Influenza: The once and future pandemic. *Public Health Rep.* 125, 16-26.

Vijaykrishna, D., J. Bahl, S. Riley, L. Duan, J.X. Zhang, H. Chen, J.S.M.

Peiris, G.J.D. Smith, and G. Yi. 2008. Evolutionary dynamics and emergence of Panzootic H5N1 influenza viruses. *PLoS Pathog.* 4, e1000161.

Vines, A., K. Wells, M. Matrosovich, M.R. Castrucci, T. Ito, and Y. Kawaoka. 1998. The role of influenza A virus hemagglutinin residues 226 and 228 in receptor specificity and host range restriction. *J. Virol.* 72, 7626-7631.

WHO. 2010. Cumulative Number of Confirmed Human Cases of Avian Influenza A/(H5N1) Reported to WHO. http://www.who.int/csr/disease/avian_influenza/country/ases_table_2010_07_29/en/index.html

Yamada, S., Y. Suzuki, T. Suzuki, M.Q. Le, C.A. Nidom, Y. Sakai-Tagawa, Y. Muramoto, and *et al*. 2006. Haemagglutinin mutations responsible for the binding of H5N1 influenza A viruses to human-type receptors. *Nature* 444, 378-382.

Zhou, H., Z. Yu, Y. Hu, J. Tu, W. Zou, Y. Peng, J. Zhu, and *et al*. 2009. The special neuraminidase stalk-motif responsible for increased virulence and pathogenesis of H5N1 Influenza A virus. *PLoS ONE* 4, e6277.