



# Research Journal of Pharmaceutical, Biological and Chemical Sciences

## Comparative Studies of H5N1 Gene Segments with other Subtypes of Influenza- A Virus.

TC Venkateswarulu<sup>\*1</sup>, Kodali V Prabhaker<sup>2</sup>, A Ranganadha Reddy<sup>1</sup>, D John Babu<sup>1</sup>, Rahul R Nair<sup>1</sup>,  
M Indira<sup>1</sup>, and A Venkatanarayana<sup>1</sup>

<sup>1</sup>School of Biotechnology, Vignan University, Vadlamudi-522213, Andhra Pradesh, India.

<sup>2</sup>Department of Biotechnology, Vikrama Simhapuri University, Nellore-524003, Andhra Pradesh, India.

### ABSTRACT

Comparative genomics of H5N1 virus will help to analyze the evolution of subtypes of Influenza A viruses and to know the differences in their genes and surface proteins which could be useful in understanding their mode of action. In this study we have compared all gene segments and proteins of H5N1 with other types of influenza-A viruses such as H1N1, H2N2, H3N2 and H9N2. The comparative analysis revealed that H5N1 was phylogenetically closer to H9N2. Homology percentages of eight genes of H1N1, H2N2, H3N1, H5N1 and H9N2 were calculated. Average homology of genes of aligned segments in subtypes of H5N1, H9N1 was found to be 81.5%.

*\*Corresponding author*

E-mail: venki\_biotech327@yahoo.com

## INTRODUCTION

An influenza pandemic is a global outbreak of disease that occurs when a new influenza A virus appears or “emerges” in the human population, causes serious illness, and then spreads easily from person to person worldwide [1]. Past influenza pandemics have led to high levels of illness, death, social disruption, and economic loss [2]. Influenza viruses come under the class of orthomyxo-viridae family, having a single stranded negative sense, segmented RNA genome in an enveloped virion [3]. There are three types of Influenza viruses: A, B and C, differentiated by antigenic differences in their nucleoprotein and matrix protein. The genome of influenza consists 8 RNA segments and encodes 10 proteins including the internal structural proteins, nucleocapsid protein (NP), and the two matrix proteins (M1 & M2). The segmented nature of the genome also allows for the exchange of entire genes between different viral strains when they cohabit the same cell [4].

The surface proteins hemagglutinin (HA) and neuraminidase (NA) extensively studied that are essential to the virus life cycle and the antigenic variations in the surface glycoproteins are used to subtype influenza A [5]. In addition, three of the influenza polypeptides are associated with RNA polymerase activity (PA, PB1, and PB2). The RNA binding non-structural protein (NS) contributes to viral pathogenicity and plays a central role in the prevention of interferon mediated antiviral response [6].

The H5N1 virus, an avian influenza virus (Bird flu) can cause illness in many species including humans by undergoing the process of Antigenic shift in some intermediate host, generally a pig or chicken[7]. H5N1 can be considered as a pandemic threat as it can move from areas where it is endemic to other parts of the world through migrating water fowl[8]. H5N1 virus has mutated through antigenic drift into dozens of highly pathogenic varieties, but all currently belonging to genotype Z of avian influenza virus H5N1[9-10].

This infection of humans coincided with an epizootic (an epidemic in nonhumans) of H5N1 influenza in Hong Kong’s poultry population and was capable of emerging in to a zoonotic disease [11]. Populations tend to have more resistance to influenza B and C, because they only undergo antigenic drift and have more similarity with previous strains[12]. Sixteen H subtypes and nine N subtypes have been identified for influenza A virus. Only one H subtype and one N subtype have been identified for influenza B virus. The most common antigenic variants of influenza A virus are H1N1 and H3N2 [13].

This study consists of comparing all the gene segments and some important proteins of H5N1 virus with other types of Influenza A viruses like H1N1, H2N2, H3N2 and H9N2. Comparative genomics of H5N1 virus will help to analyze the evolution of these subtypes and to understand the differences in their genes and surface proteins which could be useful to comprehend their mode of action and genetic drifts among these viruses.



## METHODS

### Data Mining

All the gene sequences and amino acid sequences of all proteins of H5N1, H1N1, H2N2, H3N2 and H9N2 should be collected from the NCBI Public Database, which is maintaining a separate public genome and proteome database for all the Influenza viruses' i.e. A, B and C types of Influenza viruses. All the subtypes studied in our project work belong to Influenza A virus. All the genes should be sorted according to their corresponding virus subtypes[14].

### Comparative study of genes and amino acid sequences

The analyzed genes and proteins should be compared to have a basic idea of their evolution and know the differences in their genes and proteins and how these differences are responsible for the change in their virulence or pathogenicity. By that we can know in which aspects H5N1 is different from other subtypes and which aspects it is similar. This comparative study was done using various tools available in JEMBOSS like PLOTCON, Pretty plot, Clustal W, BLAST P, BLAST N[15-17].

### Phylogenetic analysis:

The phylogenetic analysis of the various subtypes of Influenza A virus will give the idea of ancestral relationships among the subtypes. The phylogenetic analysis of the 5 subtypes selected in our study sophisticated tool like NJplot (Neighbor Joining plot) used for a better understanding. Using this tool phylogenetic tree can be constructed for all the eight gene segments of the selected 5 subtypes.

## RESULTS AND DISCUSSION

### Multiple alignments of gene segments of Influenza-A subtypes

The eight genes of H1N1, H2N2, H3N1, H5N1 and H9N2 have been aligned by using the multiple alignment tool Clustal W and their percentage of homology is presented in Table (1). Based on homology percentages it was observed that H5N1 showed highest similarity with H9N2. The similarity regions of nucleotides in segment 2 of H1N1, H2N2, H3N2, H5N1 and H9N2 were analyzed by Pretty plot and was found that segment-2 Of H5N1 similar to H9N2 (figure 1). The aligned regions of these gene segments were merged using Plotcon for identification of conserved regions. Earlier it was reported as 87 conserved regions were identified in the influenza genome and in avian host strains, the segments NS, M show the most conservation (94.5%), followed by the PB1, NP, PA, and finally the PB2 segment. In the swine strains the M segment shows the most conservation followed by the segments NP, PB2, PA, and finally PB1[18].

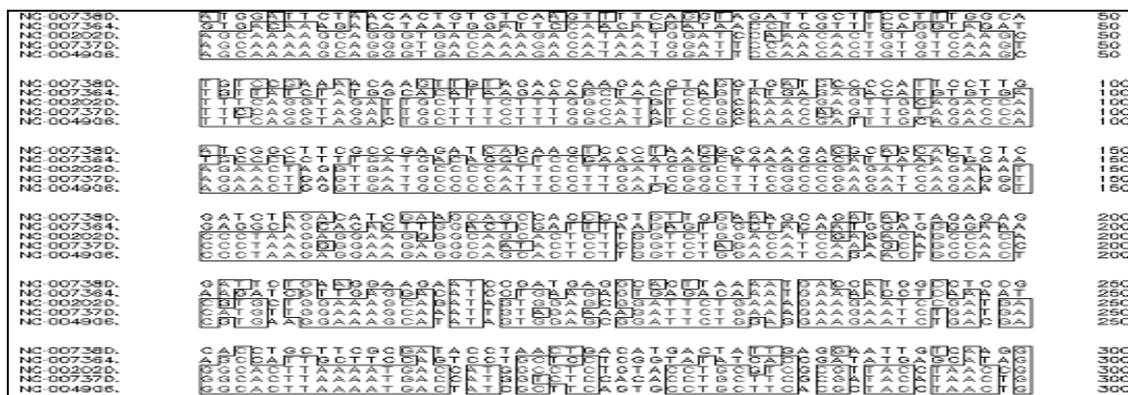
**Table 1: Percentage homology of all the eight genes of H1N1, H2N2, H3N1, H5N1 and H9N2 calculated using Clustal W (1.82)**

Aligned Subtypes	Percentage homology of genes								Avg.(%)
	PB2	PB1	PA	HA	NP	NA	M	NS	
H1N1 : H2N2	92 %	83 %	93 %	70 %	91 %	55 %	94 %	92 %	83.7
H1N1 : H3N2	89 %	82 %	89 %	58 %	87 %	60 %	92 %	89 %	80.7
H1N1 : H5N1	83 %	83 %	85 %	66 %	86 %	81 %	90 %	74 %	81.0
H1N1 : H9N2	83 %	83 %	85 %	63 %	84 %	57 %	88 %	89 %	79.0
H2N2 : H3N2	93 %	89 %	93 %	60 %	92 %	88 %	94 %	93 %	87.7
H2N2 : H5N1	83 %	90 %	84 %	71 %	84 %	56 %	89 %	72 %	78.6
H2N2 : H9N2	83 %	88 %	83 %	63 %	82 %	84 %	88 %	87 %	82.2
H3N2 : H5N1	83 %	86 %	83 %	58 %	83 %	56 %	87 %	75 %	76.4
H3N2 : H9N2	82 %	86 %	82 %	59 %	81 %	82 %	87 %	85 %	80.5
H5N1 : H9N2	87 %	93 %	90 %	63 %	94 %	60 %	92 %	73 %	81.5

The nucleotide sequences of all segments of influenza-A virus subtypes similar regions were aligned by Clustal W. Based on the aligned regions of virus subtypes, it was found that segment-6 of H5N1 showed maximum similarity with H1N1, where as other segments (2, 3, 4, 5, 7 and 8) showed similarity with H9N2 (figure 2).







g. Segment-8

**Figure 2: Multiple Sequence alignments of different segments i.e. seg-2 to seg- 8 (a to g) of subtypes (H2N2, H5N1, H1N1, H3N1 and H9N2) of influenza-A virus. In H5N1 all the segments shows maximum aligned regions with H9N2, except segment -6 of H5N1 and it shows more aligned regions with H1N1.**

### Phylogenetic Analysis of Segments of Subtypes of Influenza-A

The gene segment-2 of H5N1 shows close resemblance compared with segment -2 of H9N2. The influenza-A subtypes of H2N2 and H5N1 are closely related and the cluster of gene segments -2 of H1N1 and H9N2 is distantly related compared with H3N2 as shown in figure (3). The unrooted phylogenetic tree of segment 3 in the selected five subtypes of Influenza A virus presented in figure (4 & 5). There is no common ancestor for all the third segments and it clearly shows the close resemblance between H5N1 and H9N2. The phylogenetic tree of segment 4 in the selected five subtypes of Influenza A virus showed the close resemblance between H3N2 and H2N2 (figure 6 & 7). The gene segment 5 in the selected five subtypes of Influenza A virus shows the close likeness between H5N1 and H9N2 in figure (8 & 9). Phylogenetic tree of segment 6 in the selected five subtypes of influenza-A virus reveals close resemblance between H5N1 and H1N1. The cluster of these gene segments are distantly related to H9N2 as shown in figure (11). The phylogenetic tree analysis of segment 7 and segment-8 in the selected five subtypes of Influenza A virus shown the close resemblance between H5N1 and H9N2 (figures 12, 13 & 14). The haemagglutinin of selected subtypes of influenza-A shows close resemblance between H5N1 and H1N2 (figure 15, 16 & 17). Similarly, neuraminidase of subtypes of influenza-A reveals close relation between H5N1 and H1N1 (figures 18, 19 & 20).

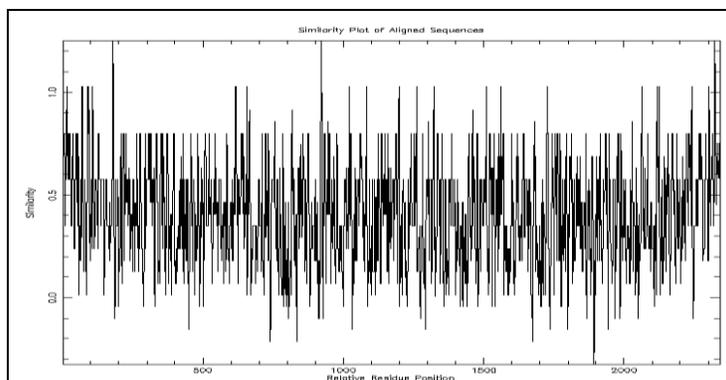


Figure 1: Similarity regions of nucleotides in segment 2 of H1N1, H2N2, H3N2, H5N1 and H9N2 by PLOTCON

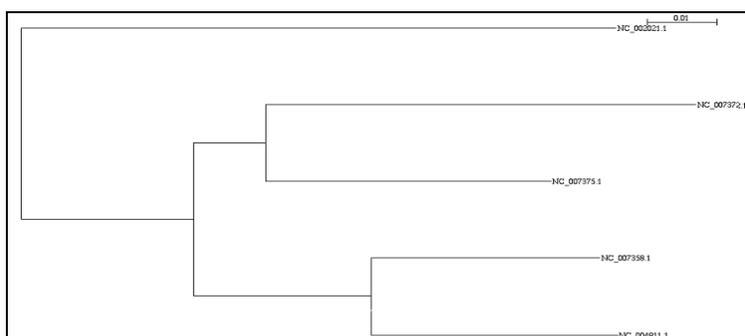


Figure 3: Phylogenetic tree of segment-2 of H1N1, H2N2, H3N2, H5N1 and H9N2

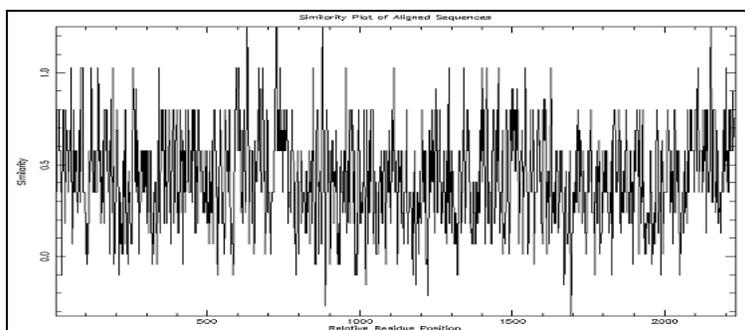


Figure 4: Similarity regions of nucleotides in segment 3 of H1N1, H2N2, H3N2, H5N1 and H9N2.by PLOTCON

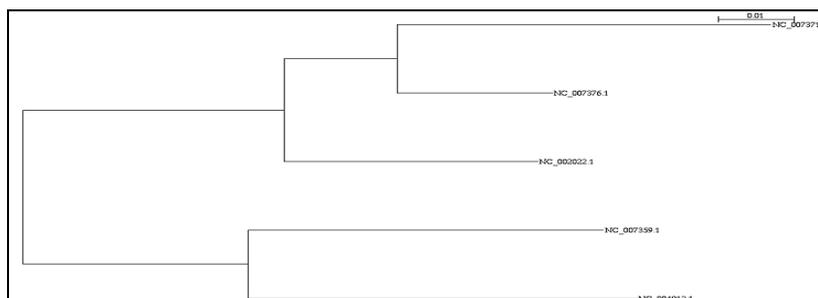


Figure 5: Phylogenetic tree of segment-3 of H1N1, H2N2, H3N2, H5N1 and H9N2

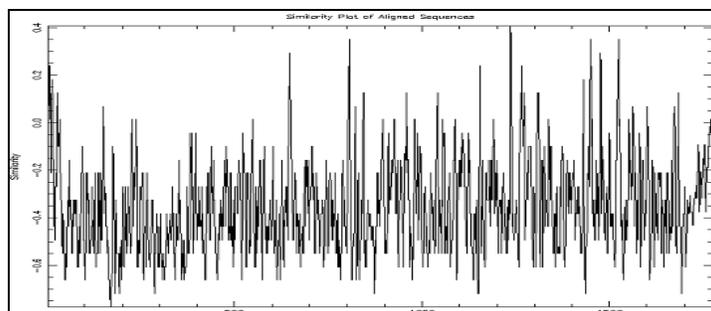


Figure 6: Similarity regions of nucleotides in segment 4 of H1N1, H2N2, H3N2, H5N1 and H9N2.by PLOTCON

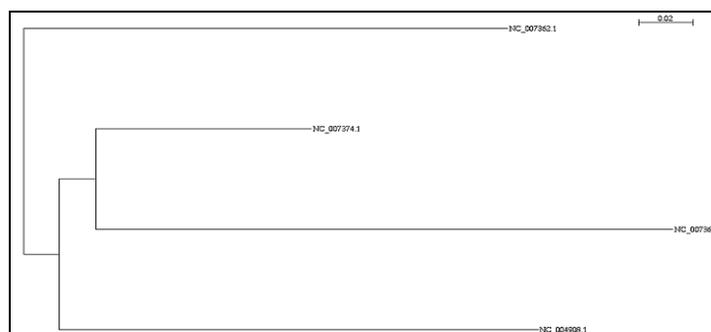


Figure 7: Phylogenetic tree of segment- 4 of H9N2, H3N2, H5N2, H2N2, and H2N2

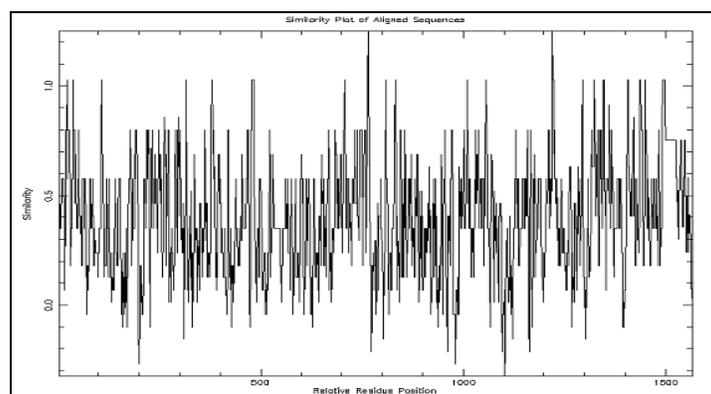


Figure 8: Similarity regions of nucleotides in segment 5 of H1N1, H2N2, H3N2, H5N1 and H9N2.by PLOTCON

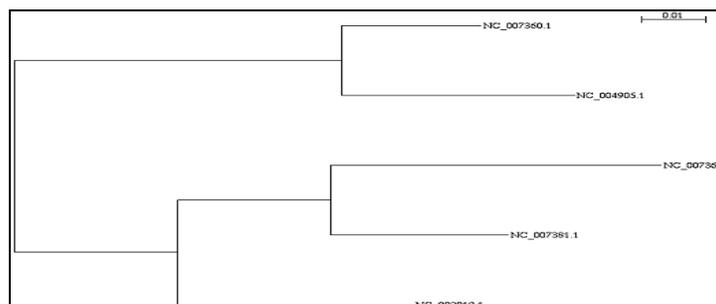


Figure 9: Phylogenetic tree of segment- 5 of H1N1, H2N2, H3N2, H5N1 and H9N2

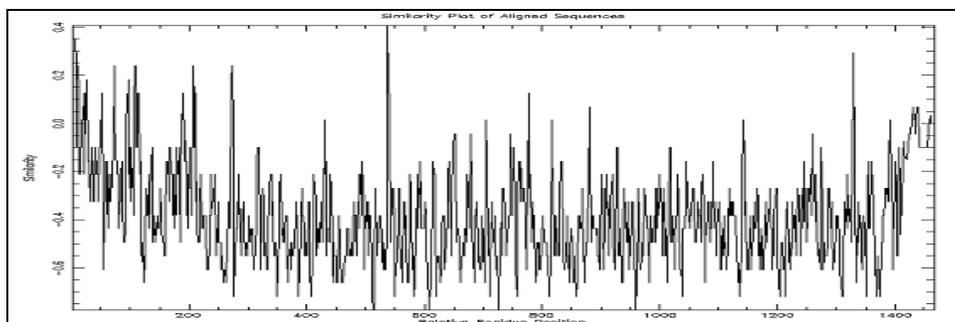


Figure 10: Similarity regions of nucleotides in segment 6 of H1N1, H2N2, H3N2, H5N1 and H9N2.by PLOTCON

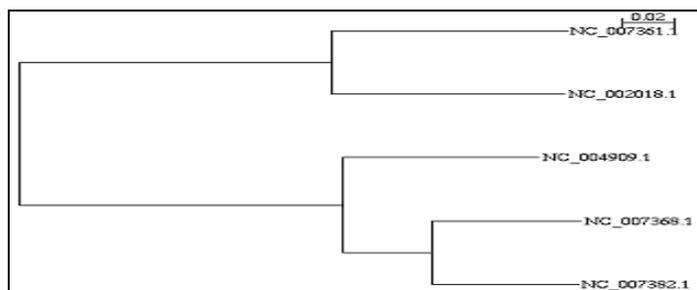


Figure 11: Phylogenetic tree of segment- 6 of H1N1, H2N2, H3N2, H5N1 and H9N2

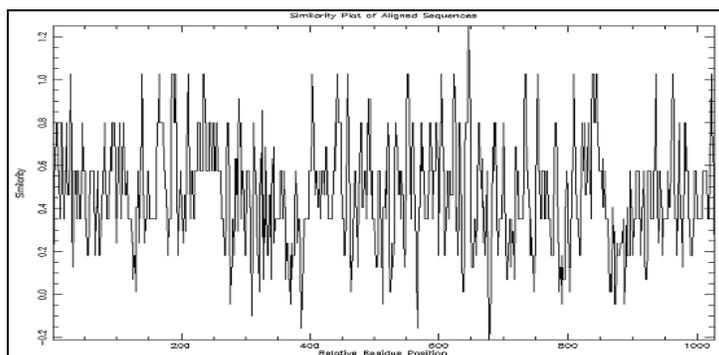


Figure 12: Similarity regions of nucleotides in segment 7 of H1N1, H2N2, H3N2, H5N1 and H9N2.by PLOTCON

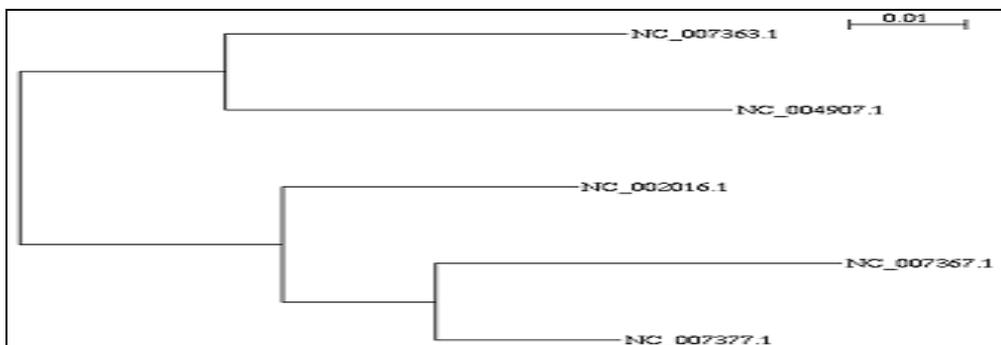


Figure 13: Phylogenetic tree of segment- 7 of H1N1, H2N2, H3N2, H5N1 and H9N2

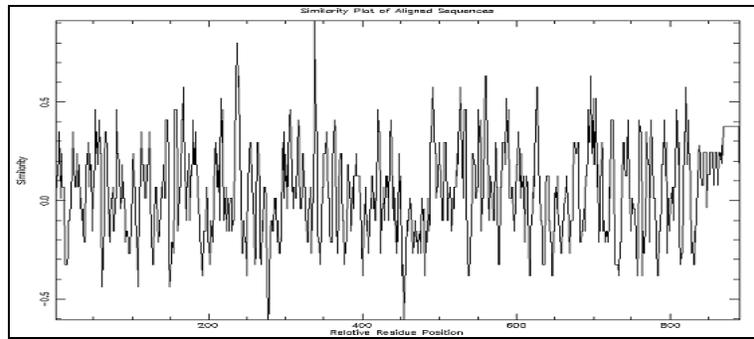


Figure 14: Similarity regions of nucleotides in segment- 8 of H2N2, H5N1, H1N1, H3N1 and H9N2 by PLOTCON

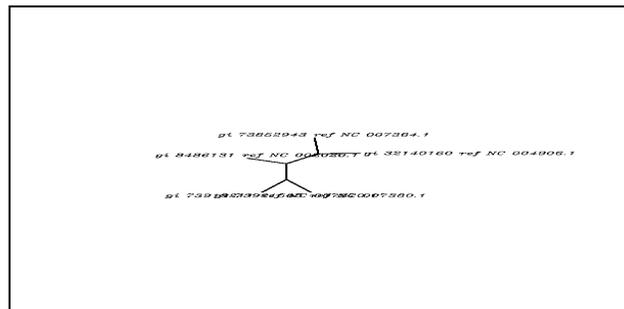


Figure 15: Phylogenetic tree of segment- 8 of H2N2, H5N1, H1N1, H3N1 and H9N2

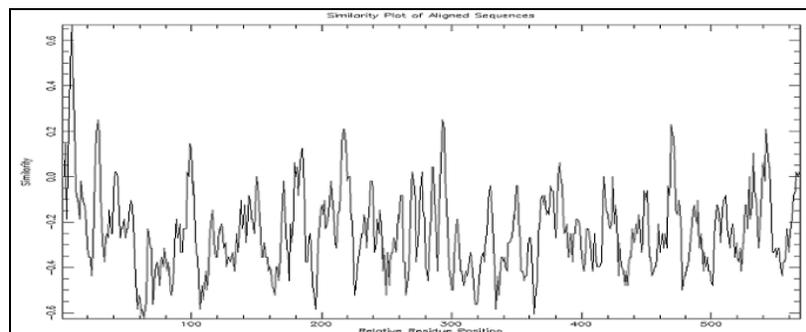


Figure 16: Similarity regions of aminoacids in Hemagglutinin of H1N1, H2N2 and H5N1

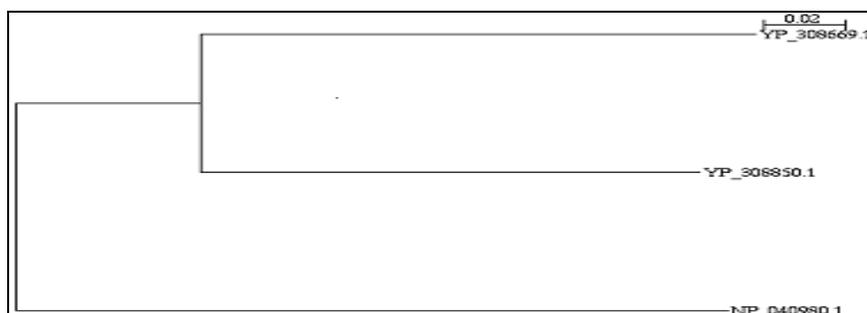


Figure 17: Phylogenetic tree of Hemagglutinin of H1N1, H2N2, and H5N1

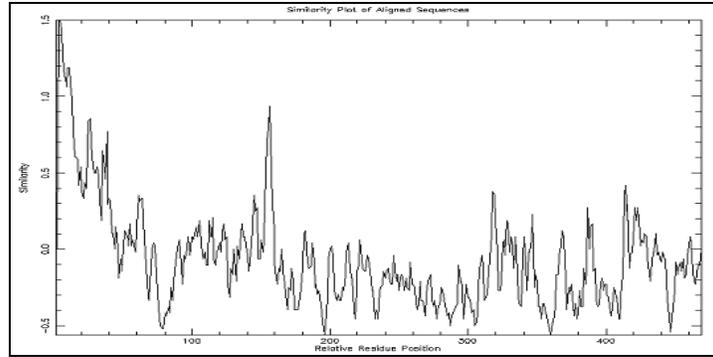


Figure 18: Similarity regions of amino acids in Neuraminidase of H5N1, H1N1, and H2N2

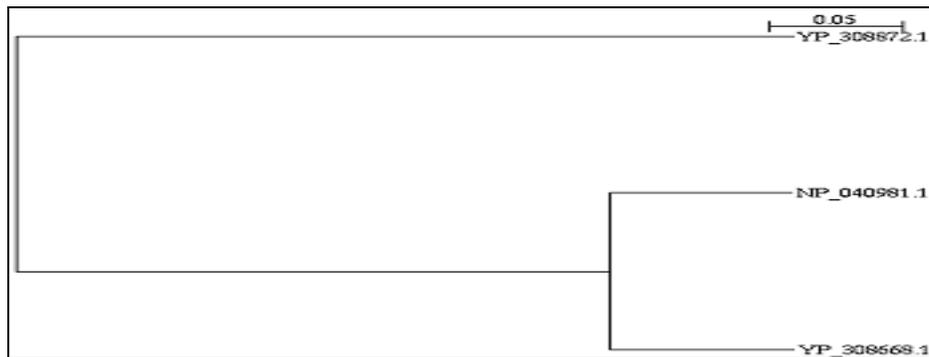


Figure 19: Phylogenetic tree of Neuraminidase of H5N1, H1N1, and H2N2

NP_040980. YP_308850. YP_308669.	M A N L V L L C A L A A A D A D T D I G F A N N S T D T V D T L K N V T V T H S F N L M K I V L L L F T R G D D I C G Y K A N F E L V D T I E R N V T V T H A K D L E L M K I V L L A T V L Y K S D G I C T G E H A N N T F V D T E F E N V T V T H A G D L E L	50 50 50
NP_040980. YP_308850. YP_308669.	D I N G K L C R L K G I A P Q L G K C N I A G W L G N P F C P P P V R E V S Y V E T E P H I G K L C K L N G I P P L E L G D C S I A G W L G N P E C D R L L S L P E W S L E E N P I I G K L C D L N O V K P L L R D C S V A G W L L G N P M C D F F N V P E V S Y L E E K A S	100 100 100
NP_040980. YP_308850. YP_308669.	N E F G I C Y P G F I V E L R E Q L S V S S F F E F P F E S W P D H K G V T R Y F C Y P G S F I V E L H L F E M K T F E V I P K D W L G I T T G G L W F C A P L I I C Y P G D F I V E S L H L E F E F V I O P K S S W S H A S R V F C A	150 150 150
NP_040980. YP_308850. YP_308669.	A F S H P B K S F Y R N L W L E K E G Y P E L K N K Y V N C K E V L W G L H P P V I B K P F F R R M V W T R K G E N Y P V K G S Y N T F G E I O R L W G V H H P I D E A E C P Y H G R S S F F R N V L W E K K N S A Y P T I E R F I F I N Q P L V L V G I H P P D	200 200 200
NP_040980. YP_308850. YP_308669.	N S D D N T Y Q N E A Y F S V I T S I N R R F T P E I A E P K V F D Q A G R M N V W L Q R A L Y N G T Y V L A F E L Y K R S I P E I A A R P V N G L R R R M E F S T L L D W A E D F L Y G N P T T T S V L S L E D R L V P E I A T F P Y N S S G R M E F D F I K	250 250 250
NP_040980. YP_308850. YP_308669.	L K P G D T I I F E A N G L I A P Y F A L S R F P G S G I I T S N A S M H E C N T P O T P L D T I N F E S T G R L V A P E Y G F I K R G S S I M K T F D I L E N C E T L C G T P L G A I N P N D A I N P E S N G N P I A P E L Y K I V K K S S A I M S E L E M G N C N T C P F P M G A	300 300 300
NP_040980. YP_308850. YP_308669.	G A F S S L P F O H P I I G E C P K Y V R S A K R F V T G L N I P S I S S G F G A T T L P F H N V P L T I G E C P K Y V K S K L V L A T G L R N V P I S R G F G A L A G F I N S M P F H P I I G E C P K Y V S N R L V A T G L R N T P I S R R F R K R O L E P	350 350 350
NP_040980. YP_308850. YP_308669.	A G F I C D W T S M I D G W G R H D P E G G E Y A D P K A T Q N A I N G I N K V N S I E I S W G M V D S M Y G Y H L S N D G S Y A L P K E T E K A F F S I T T N K L N V L K I A J A G F T E G G W G G M V D G W G P H L S N E G E S G Y A A D K E L F Q K A I D G V T R L V N I	400 400 400

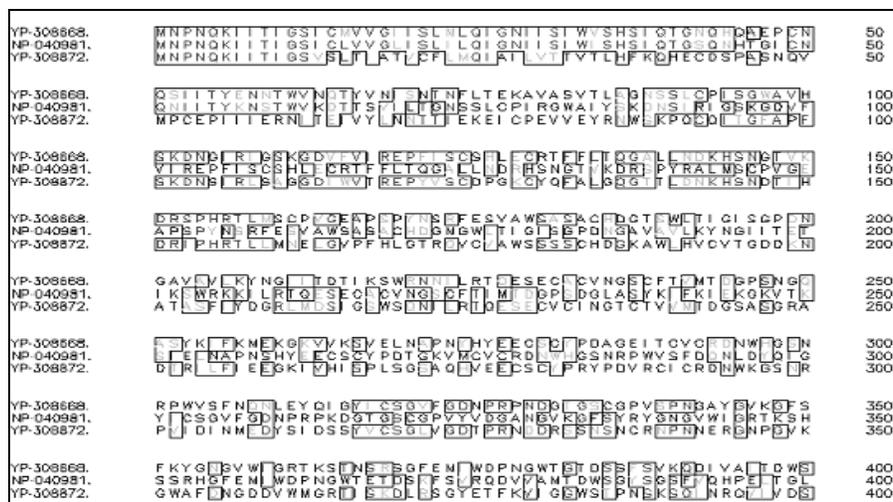


Figure 20: Multiple Sequence alignments Hemagglutinin and Neuraminidase segments of H1N1, H2N2, and H5N1

### ACKNOWLEDGEMENTS

The authors would like to thank Dr. L. Rathaiah, Chairman and Prof. M.S.C Bose, Vice Chancellor, Dr.S.Krupanidhi, HOD, School of Biotechnology, Vignan University for providing computing facilities to carry out this work.

### REFERENCES

- [1] [www.health.stste.pa.us/flu](http://www.health.stste.pa.us/flu)
- [2] DHHS: Department of Health and Human Services Pandemic Influenza Response and Preparedness Plan. August 26, 2004. ([www.dhhs.gov/nvpo/pandemicplan](http://www.dhhs.gov/nvpo/pandemicplan))
- [3] Ahn I, BJ Jeang, SE Bae, J Jung and HS Son.Genomic analysis of influenza A viruses, including avian flu(H5N1) strains.Eur.J.Epidermiol.,2006; 21: 511-519.
- [4] David A Steinhauer, John J Skehel. Annual, Review of Genetics of Influenza viruses.2002; 36: 305-332.
- [5] Bush RM, Fitch WM, Bender CA, Cox NJ: Positive selection on the H3hemagglutinin gene of human influenza virus A. Mol Biol Evol 1999; 16: 1457-1465.
- [6] Ghedin E, Sengamalay NA, Shumway M, Zaborsky J, Feldblyum T, Subbu V, Spiro DJ, Sitz J, Koo H, Bolotov P, et al: Large-scale sequencing of human influenza reveals the dynamic nature of viral genome evolution. Nature 2005; 437: 1162-1166.
- [7] De Jong MD et al. Fatal avian influenza A (H5N1) in a child presenting with diarrhea followed by coma. N Engl J Med 2005; 352: 686-91.
- [8] Ligon BL et al., Article on Avian Influenza Virus H5N1. journal of virology 2005; 32: 233-244.
- [9] Chen H, Deng G, Li Z, Tian G, Li Y, Jiao P, Zhang L, Liu Z, Webster R, G Yu K.The evolution of H5N1 influenza viruses in ducks in southern China ,Proceedings of the National Academy of Sciences of the United States of America 2004; 101: 10452-10457.



- [10] Chen W, Calvo PA, Malide D, Gibbs J, Schubert U, Bacik I, et al. Anovel influenza A virus mitochondrial protein that induces cell death. *Nat Med.* 2001; 7: 1306–12.
- [11] Tam JS. Influenza A (H5N1) in Hong Kong: an overview. *journal of Microbiology* 2002; 20: 77-81.
- [12] Agustín Portela, Paul Digard, The influenza virus nucleoprotein- a multifunctional RNA-binding protein pivotal to virus replication. *journal of general virology* 2001; 83: 723-734.
- [13] Yohannes K, Roche P, Hampson A, Miller M, Spencer J. Annual report of the National Influenza Surveillance Scheme, 2003 2004, *Commun Dis Intell* 28 (2), 160-168.
- [14] <http://www.ncbi.nlm.nih.gov/genomes/FLU/FLU.html>
- [15] [www.ebi.ac.uk/Tools/emboss](http://www.ebi.ac.uk/Tools/emboss)
- [16] [www.hpa-bioinfotools.org.uk/pise/plotcon](http://www.hpa-bioinfotools.org.uk/pise/plotcon)
- [17] [emboss.bioinformatics.nl/cgi-bin/emboss/prettyplot](http://emboss.bioinformatics.nl/cgi-bin/emboss/prettyplot)
- [18] Mahmoud ElHefnawi, Osama AlAidi, Nafisa Mohamed et al. Identification of novel conserved functional motifs across most Influenza A viral strains. *Virology Journal* 2011; 8: 44: 1- 10.