Analysis of Neuraminidase Inhibitors Resistance to H5N1 & H1N1 of Influenza Virus A from 2005-2007, 2010 & 2013

Submitted By

Hannan Mir

ID: 2013-3-79-006

July 2015

East West University

Aftabnagar

Dhaka, Bangladesh

Analysis of Neuraminidase Inhibitors

Resistance to H5N1 & H1N1 of Influenza Virus A from 2005-2007, 2010 & 2013

A dissertation submission for the partial fulfillment of the requirement for the degree of Master of Pharmacy (M. Pharm)

Submitted By

Hannan Mir

ID: 2013-3-79-006

July 2015

Research Invigilator

Dr. Repon Kumer Saha

Assistant Professor

East West University

Aftabnagar

Dhaka, Bangladesh

DEDICATED

TO

MY BELOVED

PARENTS & TEACHERS

Department of Pharmacy
East West University
Aftabnagar
Dhaka, Bangladesh
www.ewubd.edu

Declaration by the Candidate

I, Hannan Mir hereby declare that this dissertation, entitled "**Analysis of Neuraminidase Inhibitors Resistance to H5N1 & H1N1 of Influenza Virus A from 2005-2007, 2010 & 2013**" Submitted by me to the department of Pharmacy, East West University, in the partial fulfillment of the requirement for the degree of Masters of Pharmacy is a genuine & authentic research work carried out by me during Spring 2014 under the guidance of Dr. Repon Kumer Saha, Assistant professor, Department of Pharmacy, East West University, Aftabnagar, Dhaka.

Hannan Mir
ID: 2013-3-79-006
Department of Pharmacy
East West University
Aftabnagar
Dhaka, Bangladesh

Department of Pharmacy
East West University
Aftabnagar
Dhaka, Bangladesh
www.ewubd.edu

Certificate by the Supervisor

This is to certify that the dissertation, entitled "**Analysis of Neuraminidase Inhibitors Resistance to H5N1 & H1N1 of Influenza Virus A from 2005-2007, 2010 & 2013**", submitted to the Department of Pharmacy, East West University, Aftabnagar, Dhaka, in partial fulfillment of the requirement for the degree of Master of Pharmacy (M. Pharm) was carried out by Hannan Mir (ID# 2013-3-79-006) under our guidance and supervision and that no part of the thesis has been submitted for any other degree. We further certify that all the sources of information availed in this connection are duly acknowledged.

Dr. Repon Kumer Saha
Assistant professor
Department of Pharmacy
East West University
Aftabnagar
Dhaka, Bangladesh

Department of Pharmacy
East West University
Aftabnagar
Dhaka, Bangladesh
www.ewubd.edu

Endorsement by the Chairperson

This is to certify that the dissertation, entitled "**Analysis of Neuraminidase Inhibitors Resistance to H5N1 & H1N1 of Influenza Virus A from 2005-2007, 2010 & 2013**" is a bonafide research work done by Hannan Mir (ID: 2013-3-79-006) is partial fulfillment of the requirements for the degree of Master of Pharmacy, under the guidance of Dr. Repon Kumer Saha, Assistant Professor, Department of Pharmacy, East West University, Dhaka. We further certify that all the sources of information availed in this connection is duly acknowledged.

Dr. Chowdhury Faiz Hossain
Professor & Chairperson
Department of Pharmacy
East West University
Aftabnagar
Dhaka, Bangladesh

Acknowledgement

At first I want to express my praises to almighty Allah who enabled me to undertake and finish this research work and finally write up the outcome project work leading towards the fulfillment of the degree of Masters in Pharmacy. I would like to thanks the people, who have made valuable contribution towards this project work.

This is my pleasure and proud privilege to express my heartiest regards and gratitude to my respected mentor Dr. Repon Kumer Saha, Assistant Professor, Department of Pharmacy, East West University, Dhaka for his expert supervision, constructive criticism, valuable advice, optimistic counseling, constant support, continuous backup and encouragement throughout every phase of the project as well as to prepare this dissertation.

I would like to express my deep sense of gratitude and warm regards to honorable Dr. Chowdhury Faiz Hossain, Dean, Department of Science & technology and Professor, Department of Pharmacy, East West University, Dhaka for extending his helping hand and affectionate attitude whenever I needed.

Special thanks are extended to my classmates of the Department of Pharmacy, East West University, Dhaka for their helping hand, continuous support and cooperation during this research.

I remember here the inspiring words of my family members and to all my well-wishers. I say many thanks to them for their wholehearted inspiration during my research work.

I am also grateful towards those who helped me in one way or other during the period of my study.

Rational and Objective of the work

Influenza A virus is a negative-sense, single-stranded RNA virus, with an eight-segment genome encoding 10 proteins. It belongs to the family *Orthomyxoviridae* which includes the genera of influenza virus A, B, and C. Generally, influenza A virus is associated with more severe disease in humans. Influenza A virus is further subtyped by two surface proteins - Hemagglutinin (H) which attaches the virion to the host cell for cell entry, and neuraminidase (N) which facilitates the spread of the progeny virus by cleaving the host sialic acid receptors attaching the progeny virus.

Influenza pandemics, defined as global outbreaks of the disease due to viruses with new antigenic subtypes, have exacted high death tolls from human populations. Now pandemics are causing by hybrid viruses, or reassortants, that are a combination of avian and human viral genes. Avian influenza viruses are therefore key contributors to the emergence of human influenza pandemics. Influenza pandemics, defined as global out-breaks of the disease due to viruses with new antigenic subtypes, have exacted high death tolls from human populations.

Mutations occur unpredictably and result in minor changes to the surface proteins. Antigenic drift produces new virus strains that may not be recognized by antibodies to earlier influenza strains. This process works as follows: a person infected with a particular influenza virus strain develops antibody against that strain. As newer virus strains appear, the antibodies against the older strains might not recognize the "newer" virus, and infection with a new strain can occur. This is one of the main reasons why people can become infected with influenza viruses more than one time. Human influenza, the seasonal affliction that causes symptoms such as fever, cough, sore throat and headaches, is caused by human influenza A viruses. Three human influenza A subtypes: H1N1, H1N2 and H3N2 have caused major outbreaks in humans.

The objective of the study is to understand the genetic patterns of avian influenza virus (H5N1 & H1N1) in the year of 2005, 2006, 2007, 2010 and 2013, to identify the mutations in amino acids and binding sites of H5N1 & H1N1 to discover the antigenic shift and drug resistance of seasonal and pandemic influenza A virus in H5N1 strain and H1N1 strain.

Abstract

There are three types of influenza viruses which are A, B and C. Among these three viruses only influenza A virus can infect wide variety of animals like horse, pigs, birds etc. even human. There are different subtypes of Influenza A virus which are HA (Hemagglutinin) and NA (Neuraminidase). Hemagglutinin and Neuraminidase are proteins which are located on the surface of the virus. 18 different HA subtypes (H1 to H18) and 11 different NA subtypes (N1 to N11) are found. In most of the cases infected birds exhibits no symptoms. These infected birds may act as silent reservoir of the virus which may affect the other birds by transmitting the virus. Low pathogenic viruses can be mutated to Highly Pathogenic Avian Influenza (HPAI) strains during replication. Variants from unique HPAI viruses could cause infection and has the ability to replicate in humans. Phylogenetic analysis helps to find out the evolution of influenza viruses. Phylogenetic tree helps to find out the diverse strains from a couple of strains. Phylogenetic tree shows the evolutionary interrelations of a group of organisms derived from a common ancestral form. It is also used to conceptualize, visualize, and analyze the relationships among biological lineages. In most of the cases phylogenetic analysis is carried out with protein sequence because proteins are the fundamental building blocks of life. While DNA sequences consist of only four bases A, G, C, and T, the functional properties of proteins are determined by a sequence of 20 possible amino acids, which leads to a much higher resolution at large evolutionary distances.

This study brings the analysis of phylogenetic tree and amino acid sequences of Neuraminidase (NA) from the influenza A virus that can infect a wide variety of birds and mammals. We have analyzed strains of three different years (2005, 2006, 2007, 2010 & 2013) of H5N1 & H1N1 from different country to see the drugs resistance pattern with respect to reported mutant position of amino acid. We did not find the exact location where reported mutations are occurred. But we found similar amino acid near the reported mutated position. We have analyzed around (before and after the mutation point) twenty positions with respect to the reported mutation point. We found same mutations around the mutated position that may cause drugs resistance.

Contents

Part-One (1)

Chapter 3: Units/scores Used to Interpret Phylogenetic Analysis, their Importance and Limitations. **Page (38-48)**

Chapter 4: Some Reported Phylogenetic Tree of NA Protein and their Interpretation. **Page (50-53)**

Chapter 5: Objectives, Download Links of Software, Version, Operation Manual, Methods, Applications and Limitations.

Page (55 - 80)

Chapter 6: Background Study and Virology of Influenza Viruses, Reason for Neuraminidase Inhibitors Resistance of Influenza Virus A.

Page (82-86)

Chapter 7: Molecular Mechanisms of Avian H5N1 Influenza A Virus Replication in Mammalian Cells. Page (88-91)

Chapter 11: Discussion & Conclusion. **Page (119-138)**

Chapter 12: References. **Page (171-193)**

List of Figures. **Page (24- 115)**

List of Tables. Page (108- 134)

Part-Two (2)

Chapter 1: Introduction. **Page (141-148)**

Chapter 2: Some Reported Phylogenetic Tree of NA Protein and their interpretation. **Page (150-152)**

List of Tables. Page (161-167)

Part-One (1)

Analysis of Neuraminidase Inhibitors Resistance to H5N1 of Influenza Virus A from 2005-2007

Chapter 1

Introduction

1.1 Phylogenetics

Phylogenetics is the science of estimating the evolutionary past, in the case of molecular phylogeny, based on the comparison of DNA or protein sequences. The idea of representing these hypotheses as trees probably dates back to Darwin, but the numerical calculation of trees using quantitative methods is relatively recent (Sneath & Sokal 1973), and their application to molecular data even more (Zuckerkandl & Pauling 1965).

1.2 Some basics (Sandra 2003)

1.2.1 Terminology

A phylogenetic tree is composed of branches (edges) and nodes. Branches connect nodes; a node is the point at which two (or more) branches diverge. Branches and nodes can be internal or external (terminal). An internal node corresponds to the hypothetical Last Common Ancestor (LCA) of everything arising from it. Terminal nodes correspond to the sequences from which the tree was derived (also referred to as operational taxonomic units or 'OTUs'). Trees can be made up of multigene families (gene trees) or a single gene from many taxa (species trees, at least theoretically) or a combination of the two. In the first case, the internal nodes correspond to gene duplication events, in the second to speciation events.

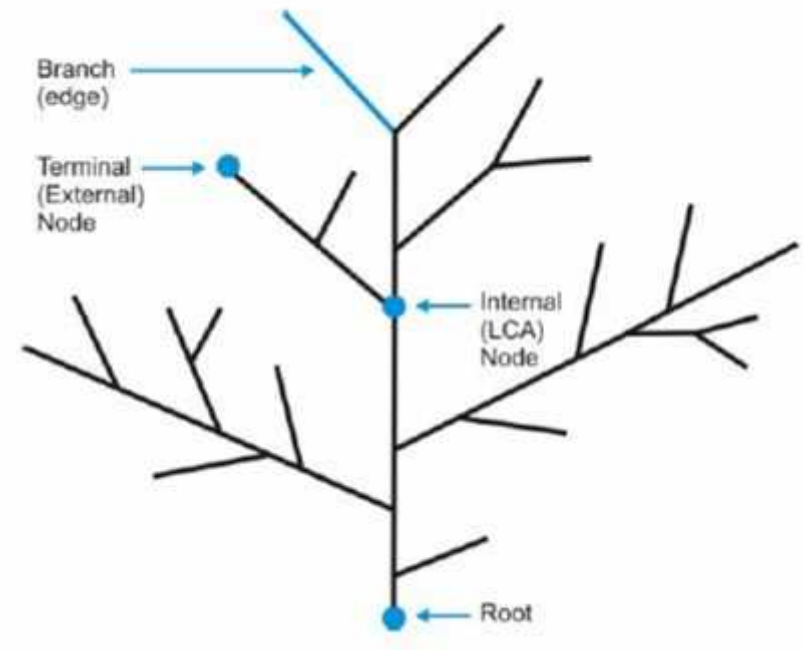


Figure: Basic elements of a phylogenetic tree.

1.2.2 **Groups**

Trees are about groupings.

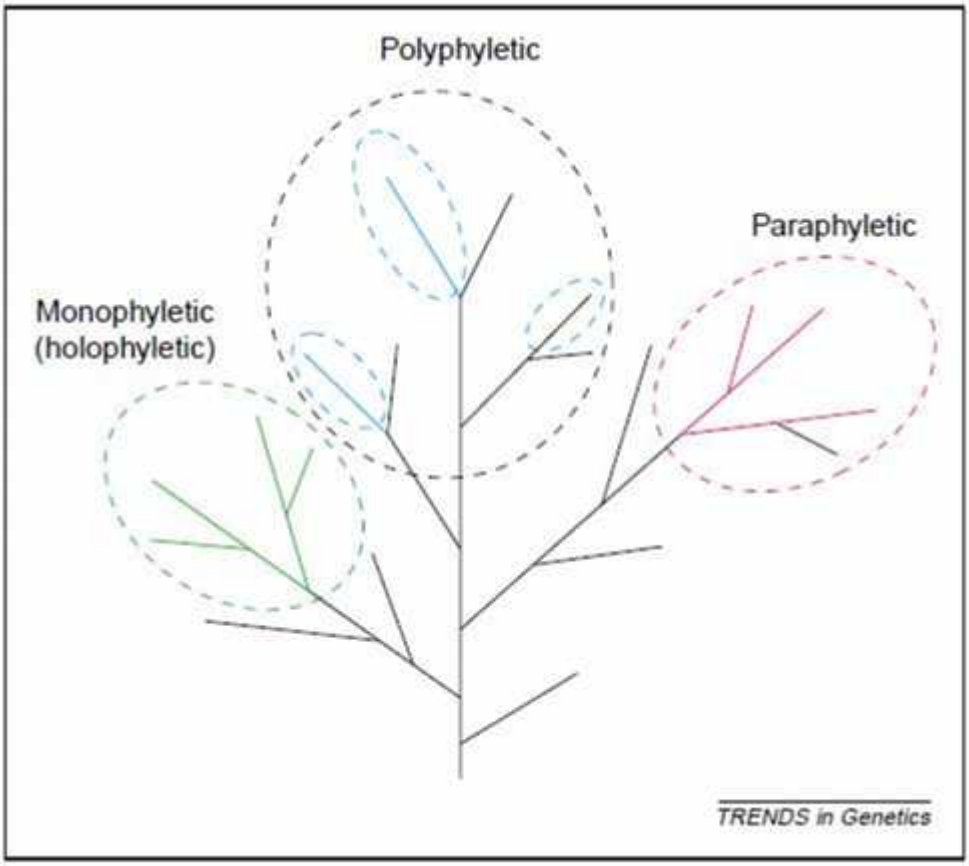


Figure: Trees are about groups: monophyletic, paraphyletic and polyphyletic.

A node is a point of divergence, with two branching lines of descendance, indicating evolutionary divergence from a common ancestor. A node and everything arising from it is a 'clade' or a 'monophyletic group'. A monophyletic group is a natural group; all members are derived from a unique common ancestor (with respect to the rest of the tree) and have inherited a set of unique common traits (characters) from it. A group excluding some of its descendents is a paraphyletic group (e.g. animals excluding humans). A hodge-podge of distantly related OTUs, perhaps superficially resembling one another or retaining similar primitive characteristics, is polyphyletic; that is, not a group at all.

1.2.3 Trees

Intuitively trees are drawn from the ground up like real trees (Fig. a). However, as these trees get larger and more complex, they can become cluttered and difficult to read. As an alternative we can expand the nodes (Fig. b) and turn the tree on its side (Fig. c). Now the tree grows left to right, and all the labels are horizontal. This makes the tree connect (Fig. a–f). Thus, the longer the branches the more relatively divergent (highly evolved) are the sequences attached to them. Alternatively, trees can be drawn to display branching patterns only ('cladograms'), in which case the lengths of the branches have no meaning (Fig. g), but this is rare done with molecular sequence trees.

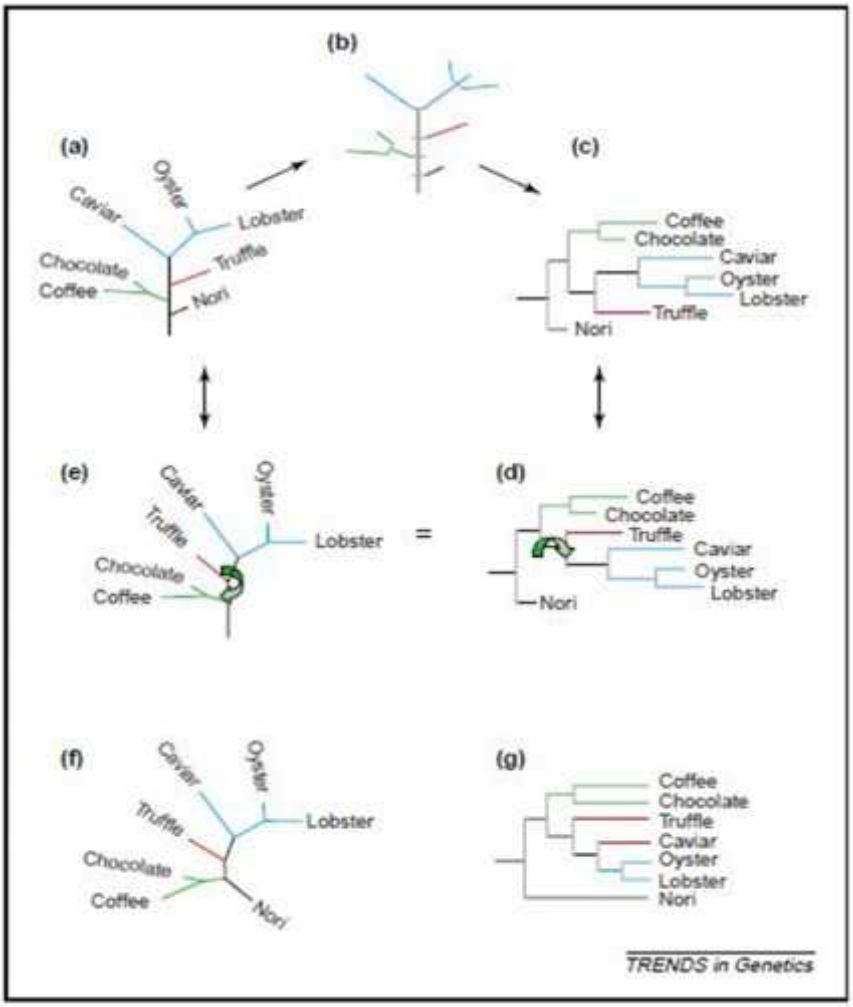


Figure: Phylogenetic tree styles. All these trees have identical branching patterns. The only differences are (f), which is unrooted (g) is a cladogram, so the branch lengths are right justified and not drawn to scale (i.e. they are not proportional to estimated evolutionary difference).

A phylogenetic tree, also known as a phylogeny, is a diagram that depicts the lines of evolutionary descent of different species, organisms, or genes from a common ancestor. Phylogenies are useful for organizing knowledge of biological diversity, for structuring classifications, and for providing insight into events that occurred during evolution (Baum 2008).

Phylogenetic tree, also called Dendrogram, a diagram showing the evolutionary interrelations of a group of organisms derived from a common ancestral form. The ancestor is in the tree "trunk"; organisms that have arisen from it are placed at the ends of tree "branches." The distance of one group from the other groups indicates the degree of relationship; *i.e.,* closely related groups are located on branches close to one another. Phylogenetic trees, although speculative, provide a convenient method for studying phylogenetic relationships.

Phylogenetic trees are used to conceptualize, visualize, and analyze the relationships among biological lineages (Stephen *et al.* 2013). Biologists use cladograms and phylogenetic trees to illustrate relationships among organisms and evolutionary relationships for organisms with a shared common ancestor. Both cladograms and phylogenetic trees show relationships among organisms, how alike, or similar, they might be.

First, a cladogram can look at trees that may have been derived from a common ancestor to arrange organisms on different branches. But those branches used aren't representative of the relative amount of change or evolutionary time that has occurred between organisms. Plus, a cladogram doesn't necessarily show exact relationships between ancestors and descendants. From assumption that all species have evolved from common ancestor, all species must have evolved through a series of events in which ancestral species gave rise to new species, each giving rise to a new linage.

On the other hand, the branches on a phylogenetic tree can be proportional to mount of change or evolutionary time. So, we can also track how species have changed over time. Species are still grouped according to similarities and physical or genetic characteristics - for example, the presence or absence of gills. But, a phylogenetic tree describes an evolutionary history by showing how ancestors are related to their descendants and how much those descendants have changed over time (Laura).

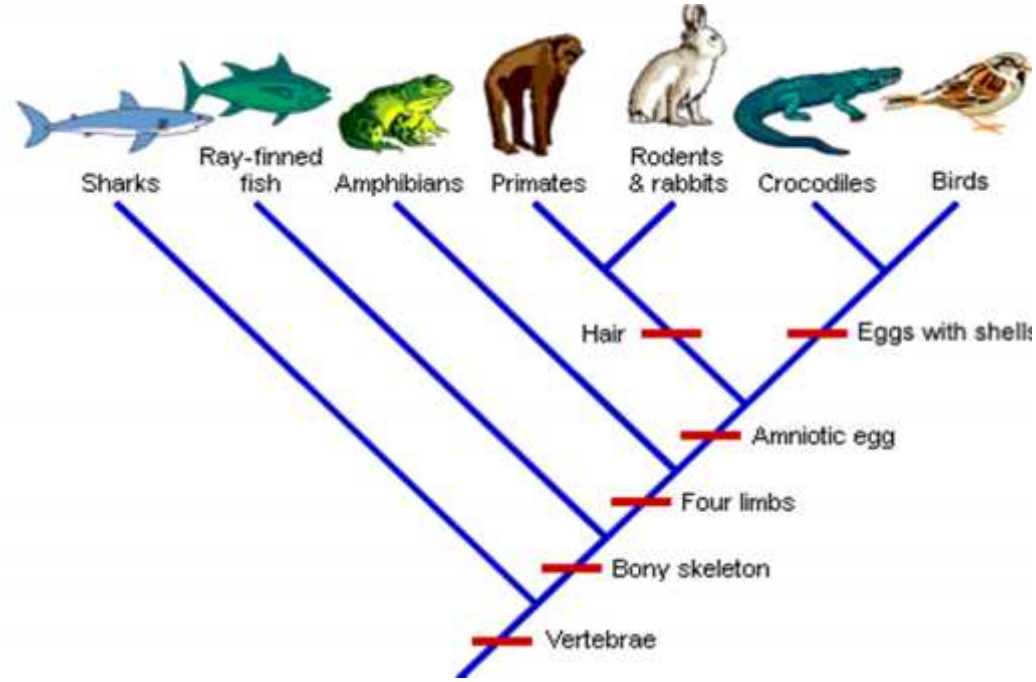


Figure: Evolution of Phylogenetic tree.

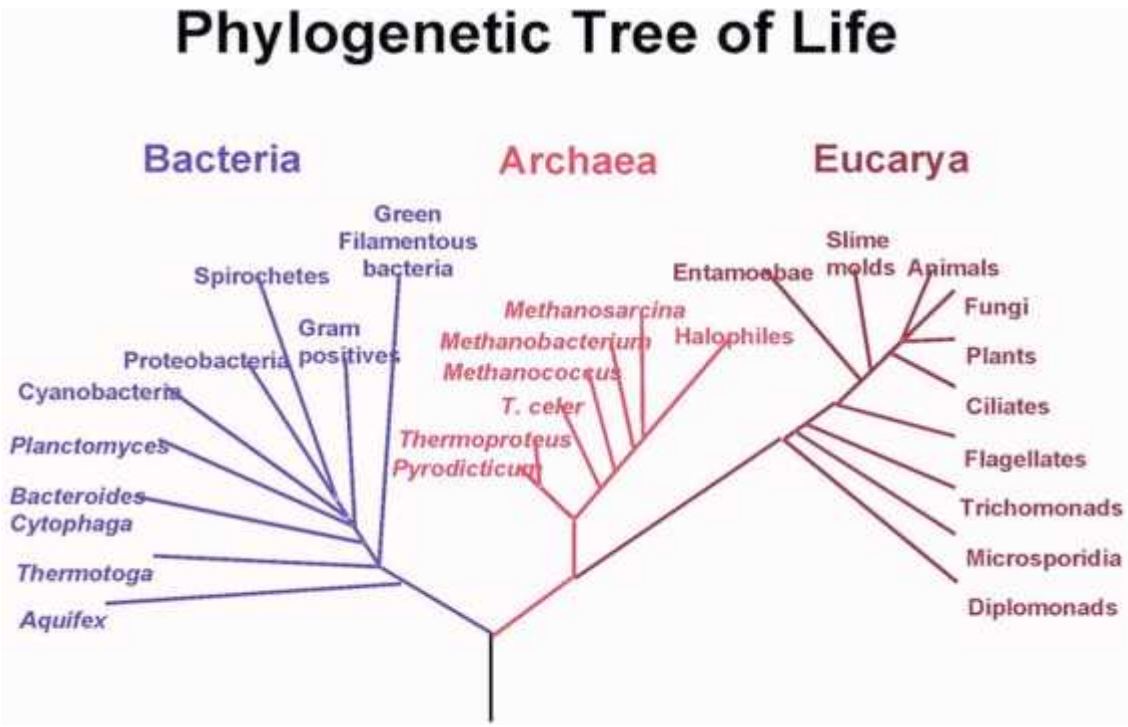


Figure: A phylogenetic tree of living things with scientific names based on RNA data and proposed by Carl Woese, showing the separation of bacteria, archaea and eukaryotes.

1.3 Molecular Evolution: Beyond Darwin

Evolution is a process by which the traits of a population change from one generation to another. In *On the Origin of Species by Means of Natural Selection,* Darwin proposed that, given overwhelming evidence from his extensive comparative analysis of living specimens and fossils, all living organisms descended from a common ancestor. The book's only illustration (the following figure) is a tree-like structure that suggests how slow and successive modifications could lead to the extreme variations seen in species today (Hartwell *et al.* 2008; Laszlo 1999).

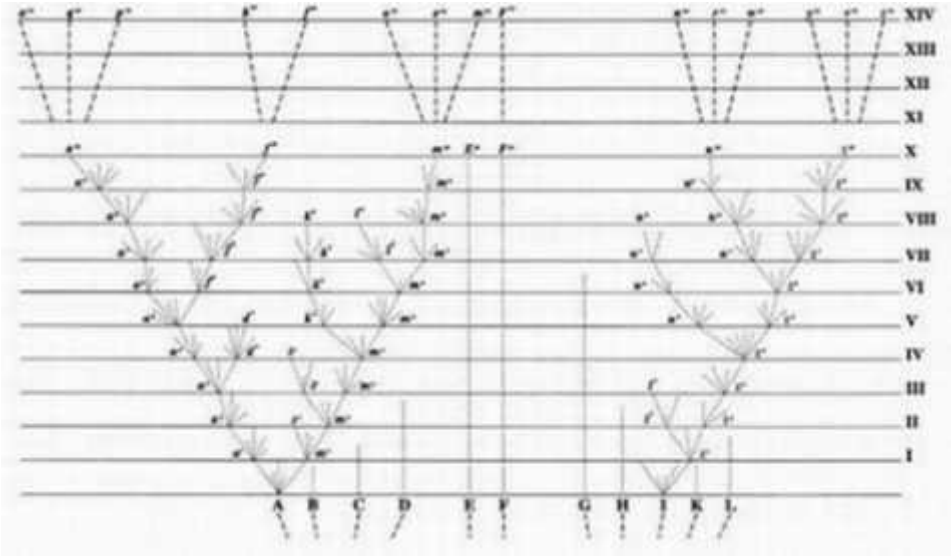


Figure: Evolution Defined Graphically. The sole illustration in Darwin's *Origin of the Species* uses a tree-like structure to describe evolution. This drawing shows ancestors at the limbs and branches of the tree, more recent ancestors at its twigs, and contemporary organisms at its buds.

Darwin's theory of evolution is based on three underlying principles: variation in traits exist among individuals within a population, these variations can be passed from one generation to the next via inheritance, and that some forms of inherited traits provide individuals a higher chance of survival and reproduction than others (Hartwell *et al.* 2008).

Although Darwin developed his theory of evolution without any knowledge of the molecular basis of life, it has since been determined that evolution is actually a molecular process based on genetic information, encoded in DNA, RNA and proteins. At a molecular level, evolution

is driven by the same types of mechanisms Darwin observed at the species level. One molecule undergoes *diversification* into many variations. One or more of those variants can be *selected* to be reproduced or *amplified* throughout a population over many generations. Such variations at the molecular level can be caused by mutations, such as deletions, insertions, inversions or substitutions at the nucleotide level, which in turn affect protein structure and biological function (Hartwell *et al.* 2008; Laszlo 1999).

1.4 Example (A simple Phylogenetic Tree)

Phylogenetic tree-building methods presume particular evolutionary models. For a given data set, these models can be violated because of occurrences such as the transfer of genetic material between organisms. Thus, when interpreting a given analysis, we should always consider the model used and its assumptions and entertain other possible explanations for the observed results. In the following example an investigation of organismal relationships in the tree suggests the eukaryote 1 is more related to the bacteria than to the other eukaryotes. Because the vast majority of other cladistic analyses, including those based on morphological features, suggest that eukaryote 1 is more related to the other eukaryotes than to bacteria; we suspect that for this analysis the assumptions of a splitting pattern of evolution are incorrect. We suspect that horizontal gene transfer from an ancestor of the bacteria 1, 2, and 3 to the ancestor of eukaryote 1 occurred because this would most simply explain the results.

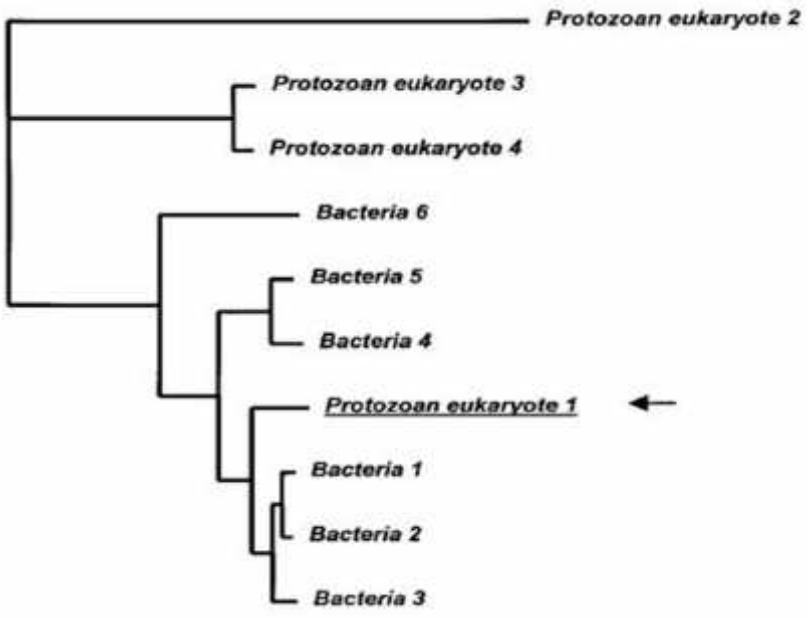


Figure: Phylogenetic tree example.

Example of a phylogenetic tree based on genes that do not match organismal phylogeny, suggesting horizontal gene transfer has occurred. The ancestor of protozoan eukaryote 1 (underlined and marked with an arrow) appears to have obtained the gene from the ancestor of Bacteria 1, 2, and 3, as this is the simplest explanation for the results. This unexpected result is not without precedent: there have been a number of reported phylogenetic analyses that suggest that protozoa have taken up genes from bacteria, most likely from bacteria that they have ingested.

Chapter 2

Molecular Basis of Phylogenetic Analysis

2.1 Nucleotide sequences vs. Amino Acid (AA) sequences

Nucleotide sequences may be coding or non-coding. In the former case, they may code for structural or catalytic RNA species but more commonly for proteins. In the case of protein-coding genes, the alignment can be accomplished based on the nucleotide or amino acid sequences. This choice may be biased by the type of analysis to be carried out after alignment; for example, silent changes in closely related sequences may be counted. In this case, an amino acid alignment will not be of much use for later analysis. By contrast, if the sequences are only distantly related, an analysis of amino acid or of nucleotide differences can be performed. Regardless of the end analysis desired, amino acid alignments are easier to carry out and less ambiguous than nucleotide alignments, which is also true for sequence database searching. Another disadvantage of nucleotide alignment is that most programs do not recognize a codon as a unit of sequence and can break up the reading frame during alignment. Particular two-sequence alignment and database search programs can be exceptions (Birney *et al.* 1996). A typical approach is to carry out the alignment at the amino acid level and to then use this to generate a corresponding nucleotide sequence alignment.

If the sequences are not protein coding, then the only choice is to carry out a nucleotide alignment. If the sequences code for structural RNA (e.g. small sub-unit ribosomal RNA [SSU rRNA]), these will be constrained by function to conserve primary and secondary structure over at least some of their lengths. Typically, there are regions of clear nucleotide identity interspersed by regions that are free to change rapidly. The performance of automatic software depends on the circumstances, but specific algorithms are being developed (Kiryu *et al.* 2007). With rRNA, the most commonly used programs manage to align the large conserved core sections, which are conserved across very broad phylogenetic distances. These core alignment blocks, however, will be interspersed by the highly variable expansion segments; most programs have difficulties with them. Consideration of the secondary structures, perhaps using a dedicated RNA editor, can help but it may still be difficult to find an unambiguous alignment. Excluding these regions from further analysis should be seriously considered; if the alignment is arbitrary, further analysis would be meaningless anyway. Fortunately, there are usually enough clearly conserved blocks of alignment to make a phylogenetic analysis possible. If the nucleotide sequences are non-coding (e.g. SINES or introns), then alignment may be difficult once the sequences diverge beyond a certain level.

Sequences that are highly unconstrained can accumulate substitutions in all positions; these rapidly become unalienable. There is no algorithmic solution and the situation will be especially difficult if the sequences have small repeats. Even if a “somewhat optimal” alignment is obtained using a particular alignment score or parsimony criterion, there may be no reason to believe it is biologically reasonable. Such scores are based on arbitrary assumptions about the details of the evolutionary processes that have shaped the sequences. Even if the details of the assumptions are justifiable, the alignment may be so hidden that it has become unrecoverable. Caution should be taken if one alignment cannot be justified over an alternative one.

2.2 The reason to analyze protein sequence

The sequence of a coding gene contains all the necessary information to create functional proteins, and its nucleotides directly incorporate the mutations that result from replication errors, radiation damage, oxidative stress, or chemical modification. Therefore, it is often advocated that DNA rather than protein sequences should be used in molecular evolution studies. However, there are many reasons why it may be more appropriate to use protein sequences. Although DNA contains all the necessary information to create a protein, it is generally not the DNA itself that is subject to natural selection. The catalysts of virtually all of the chemical transformations in the cell are proteins. Proteins are the fundamental building blocks of life and they are the units of life on which natural selection acts. While DNA sequences consist of only four bases A, G, C, and T, the functional properties of proteins are determined by a sequence of 20 possible amino acids, which leads to a much higher resolution at large evolutionary distances. The underlying DNA sequence, which is related to the amino acid sequence of the proteins via the genetic code, reflects this selection process in combination with species-specific pressures on the DNA sequence. Thus, while there is sufficient room for the DNA sequence to change and respond to requirements of GC content, the protein sequence may remain almost unaltered and not reflect such adaptations.

2.3 Choosing DNA or Protein

If possible, it is recommended to analyze a data set both ways (DNA and Protein); however, with the limitation that, for very distantly related taxa, nucleic acid sequences have probably lost most, if not all, phylogenetic signal. For a group of taxa that are relatively closely related (like different isolates of the same or similar viruses); DNA-based analysis is probably a good recommendation, since there may be fewer problems like differences in codon bias or saturation of the third position of codons. It is nevertheless strongly recommended to carry out an analysis in parallel on the protein sequence data. Moreover, where there is ambiguity in the alignment of gene sequences, it is recommended translating the sequences first to their corresponding protein sequences, then aligning the protein sequences, and determining the position of gaps in the DNA sequences according to the more reliable protein alignment.

2.4 Phenetics and Cladistics

2.4.1 Phenetics: Pheneticists argued that classifications should encompass as many variable characters as possible, these characters being analysed by rigorous mathematical methods **(**Charles & Robert 1957**)**.

Phenetics is the study of relationships among a group of organisms on the basis of the degree of similarity between them, be that similarity molecular, phenotypic, or anatomical. A tree-like network expressing phenetic relationships is called a phenogram. Such methods (exp. distance based) place a greater emphasis on the relationships among data sets than the paths they have taken to arrive at their current states.

2.4.2 Cladistics: Cladistics emphasizes the need for large datasets but differs from phenetics in that it does not give equal weight to all characters. Cladists are generally more interested in evolutionary pathways than in relationships (exp. maximum parsimony). While a phenogram may serve as an indicator of cladistic relationships, it is not necessarily identical to the cladogram. If there is a linear relationship between the time of divergence and the degree of genetic (or morphological) divergence, the two types of trees may become identical to each other **(**Hennig 1966).

The maximum parsimony method (Discussed in chapter 3) is a typical representative of the cladistic approach, whereas the UPGMA (Discussed in chapter 3) method is a typical phenetic method. The other methods, however, cannot be classified easily according to the above criteria. For example, the transformed distance method and the neighbor's relation method have often been said to be phenetic methods, but this is not an accurate description. Although these methods use similarity (or dissimilarity, i.e., distance) measures, they do not assume a direct connection between similarity and evolutionary relationship, nor are they intended to infer phenetic relationships.

The maximum likelihood (Discussed in chapter 3) method is a phenetic method that is statistically well founded. It has often lower variance than other methods (i.e. it is frequently the estimation method least affected by sampling error) and tends to be robust to many violations of the assumptions in the evolutionary model. Even with very short sequences maximum likelihood tends to outperform alternative methods such as parsimony or distance methods. Different tree topologies are evaluated. An important disadvantage is that is very CPU intensive and thus time consuming and not appropriate for large datasets.

Chapter 3

Units/scores Used to Interpret Phylogenetic Analysis, their Importance and Limitations

3.2 UPGMA

UPGMA is an abbreviation of *unweighted pair group method with arithmetic means*. The method is simple and intuitive (Durbin *et al.* 1999) which makes it appealing. The method works by clustering nodes at each stage and then forming a new node on a tree.

- Unweighted – all pairwise distances contribute equally.
- Pair – group – groups are combined in pairs.
- Arithmetic mean – pairwise distances to each group (clade) are mean distances to all members of that group.

Clustering is done by searching for the smallest value in the pairwise distance matrix. The newly formed cluster replaces the operational taxonomic units (OTUs). It represents in the distance matrix, and distances between the newly formed cluster and each of the remaining OTUs are calculated. This process is repeated until all OTUs are clustered. In UPGMA the distance of the newly formed cluster is the average of the distances of the original OTUs. This process of averaging assumes that the evolutionary rate from the node of the two clustered OTUs to each of the two OTUs is identical. The whole process of clustering thus assumes that the evolutionary rate is the same in all branches, which is frequently violated. Therefore, UPGMA tends to give the wrong tree when evolutionary rates differ along the different branches.

This process continues from the bottom of the tree and in each step a new node is added, and the tree grows upward. The length of the branch at each step is determined by the difference in heights of the nodes at each end of the branch. UPGMA has built in assumptions that the tree is additive and that all nodes are equally distance from the root. Since a "molecular clock" hypothesis assumption poses biological issues, UPGMA is not used much today, but gave way to a very common approach now termed "Neighbor Joining" (Saitou & Nei 1987).

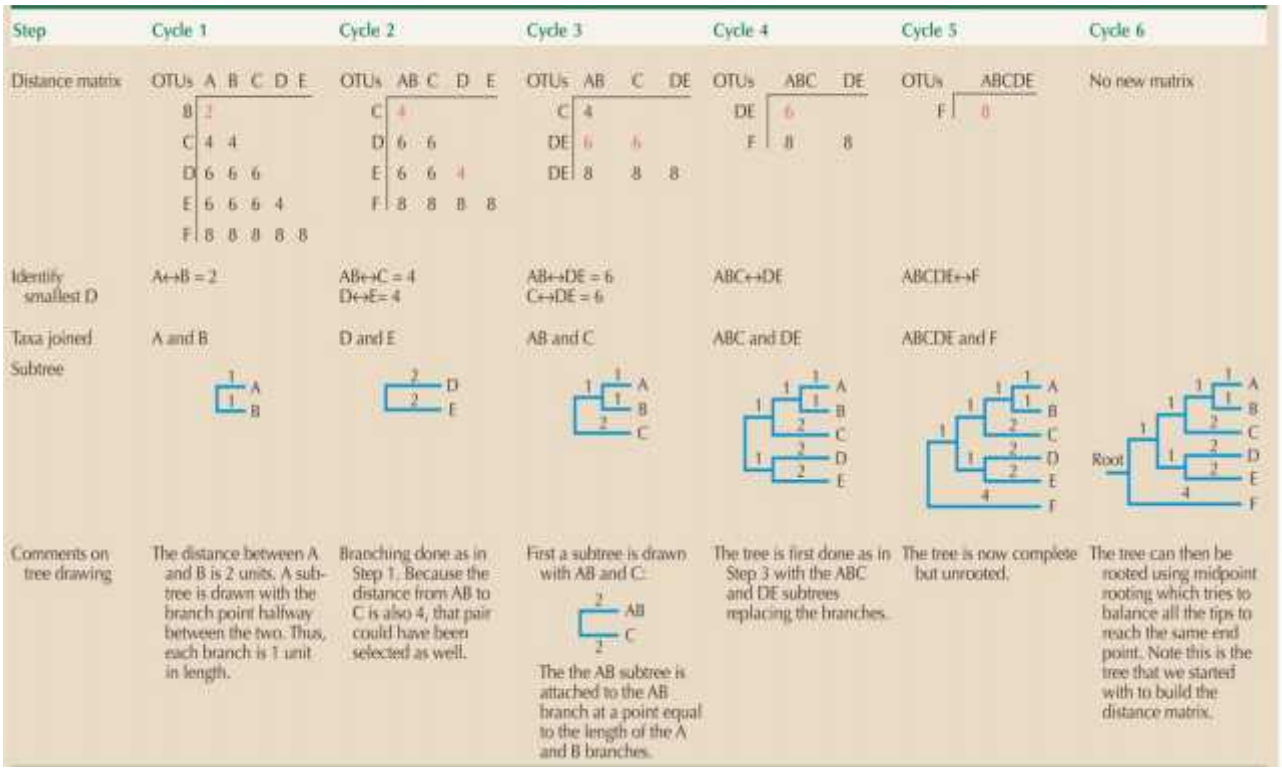


Figure: Example of UPGMA tree construction.

3.2 Neighbor Joining

Neighbor Joining (NJ) works like UPGMA in that it creates a new distance matrix at each step, and creates the tree based on the matrices. The difference is that NJ does not construct clusters but directly calculates distances to internal nodes. The first step in the NJ algorithm is to create a matrix with the Hamming distance between each node or taxa. The minimal distance is then used to calculate the distance from the two nodes to the node that directly links them. From there, a new matrix is calculated and the new node is substituted for the original nodes that are now joined. The advantage here is that there is not an assumption about the distances between nodes since it is directly calculated.

Since The Neighbor Joining approach to tree construction takes advantage of common clustering techniques, it is efficient to execute and easy to understand. It produces an unrooted tree that shows the relationship between sequences without assigning a root node from which all other sequences have been derived. In order to construct a tree with a common

ancestor node, an outgroup species is chosen that is distantly related to the remaining sequences. The location where the new species connects to the recently constructed tree is a good indicator for the most likely location for the root of the tree. If it is not easy or possible to find an outgroup, other strategies allow one to locate a root of the tree such as using the midpoint of the longest chain of consecutive edges, which would indicate the root if the tree followed the molecular clock within reason (Durbin *et al.* 1999).

The Neighbor-joining (NJ) method constructs a tree by sequentially finding pairs of neighbors, which are the pairs of OTUs connected by a single interior node. The clustering method used by this algorithm is quite different from the UPGMA, because it does not attempt to cluster the most closely related OTUs, but rather minimizes the length of all internal branches and thus the length of the entire tree. The NJ algorithm starts by assuming a star-like tree that has no internal branches. In the first step, it introduces the first internal branch and calculates the length of the resulting tree. The algorithm sequentially connects every possible OTU pair and finally joins the OTU pair that yields the shortest tree.

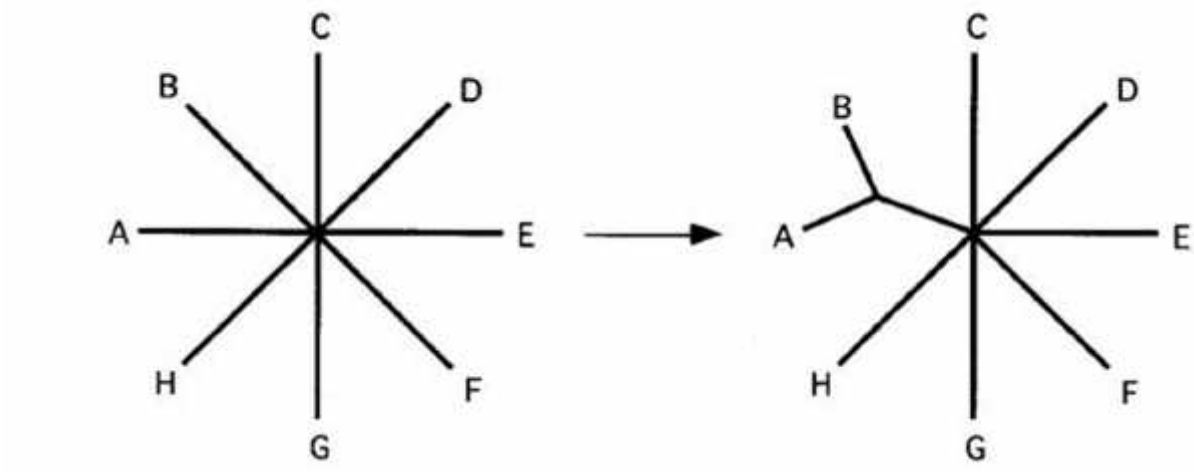


Figure: Star decomposition. This is how tree-building algorithms such as neighbor joining (NJ) work. The most similar terminals are joined, and a branch is inserted between them and the remainder of the star. Subsequently, the new branch is consolidated so that its value is a mean of the two original values, yielding a star tree with n-1 terminals. The process is repeated until only one terminal remains.

3.3 Maximum parsimony (MP)

Maximum parsimony (MP) aims to find the tree topology for a set of aligned sequences that can be explained with the smallest number of character changes (i.e. substitutions). For a particular topology, the MP algorithm infers for each sequence position the minimum number of character changes required along its branches to explain the observed states at the terminal nodes. The sum of this score for all positions is called the *parsimony length* of a tree and this is computed for different tree topologies. When a reasonable number of topologies have been evaluated, the tree that requires the minimum number of changes is selected as the maximum parsimony tree. Maximum parsimony is very sensitive to the "***long-branch attraction***" problem since there is no way to correct for multiple hits. MP always assumes that a common character is inherited directly from a common ancestor and thus always underestimates the real divergence between distantly related taxa (Philippe *et al.* 2009).

Parsimony is based on the assumption that the mostly likely tree is the one that requires the fewest number of changes to explain the data in the alignment (Hall 2007).

Maximum Parsimony is also different from the other methods in that it does not find branch lengths but rather the total overall length in terms of the number of changes. Often MP finds two or more trees that it deems equal and does not provide a definite answer in how to distinguish which tree represents the actual evolutionary tree. In most cases a strict (majority rule) consensus is used to solve this dilemma (Hall 2007).

An artifact called long-branch attraction sometimes occurs in parsimony. The branch length indicates the number of substitutions between two taxa or nodes. Parsimony assumes that all taxa evolve at the same rate and contribute that same amount of information. Long-branch is the phenomenon in which rapidly evolving taxa are placed together on a tree because they have many mutations. Anytime two long branches are present, they may be attracted to one another (Pevsner 2003).

3.3.1 Maximum Parsimony (Positive points)

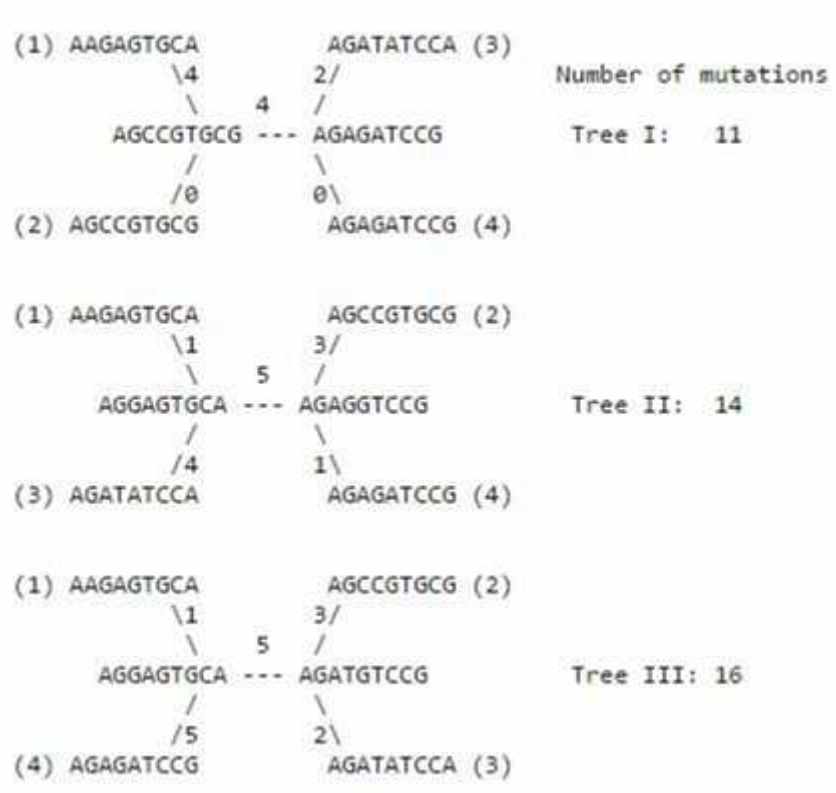
- Is based on shared and derived characters. It therefore is a cladistic rather than a phenetic method,
- Does not reduce sequence information to a single number,
- Tries to provide information on the ancestral sequences,
- Evaluates different trees.

3.3.2 Maximum Parsimony (Negative points)

- Is slow in comparison with distance methods,
- Does not use all the sequence information (only informative sites are used),
- Does not correct for multiple mutations (does not imply a model of evolution)
- Does not provide information on the branch lengths,
- Is notorious for its sensitivity to codon bias.
- An example of the maximum parsimony method for a dataset of 4 nucleic-acid sequences is given below.
- We can consider the following set of homologous sequences:

	Site								
Sequence	1	2	3	4	5	6	7	8	9
1	A	A	G	A	G	T	G	C	A
2	A	G	C	C	G	T	G	C	G
3	A	G	A	T	A	T	C	C	A
4	A	G	A	G	A	T	C	C	G

- For four OTUs there are three possible unrooted trees. The trees are then analyzed by searching for the ancestral sequences and by counting the number of mutations required to explain the respective trees as shown below:



Tree I have the topology with the least number of mutations and thus is the most parsimonious tree.

3.4 Maximum Likelihood (ML)

Maximum likelihood (ML) is similar to the MP method in that it examines different tree topologies and evaluates the relative support by summing over all sequence positions. ML algorithms search for the tree that maximizes the probability of observing the character states, given a tree topology and a model of evolution. For a particular tree, the likelihood calculation involves summing over all possible nucleotide (or amino acid) states in the ancestral (internal) nodes. Numerical optimization techniques are used to find the combination of branch lengths and evolutionary parameters that maximizes the likelihood. Depending on the search algorithm, the likelihood of a number of tree topologies is searched with this criterion, and the tree yielding the highest likelihood is chosen as the best tree. Unfortunately, obtaining the likelihood of a tree can be computationally very demanding (Philippe *et al.* 2009).

One of the big advantages to ML is the ability to make statistical comparisons between topologies and data sets. ML can return several equally likely trees – a pro and con depending on the study (Egan & Crandall 2006). Maximum Likelihood makes assumptions that the model used is accurate and if the model does not accurately reflect the underlying data set, the method is inconsistent. ML is designed to be robust, but breaching is assumptions can cause problems (Jeffrey & Eric 2007).

A disadvantage of ML is the extensive computation as well as new evidence that suggest there can be multiple maximum likelihood points for a given phylogenetic tree (Dorigo & Caro 1999).

3.5 Distance, Parsimony and Maximum Likelihood: The differences

Distance matrix methods simply count the number of differences between two sequences. This number is referred to as the evolutionary distance and its exact size depends on the evolutionary model used. The actual tree is then computed from the matrix of distance values by running a clustering algorithm that starts with the most similar sequences (i.e., those that have the shortest distance between them) or by trying to minimize the total branch length of the tree.

The principle of maximum parsimony searches for a tree that requires the smallest number of changes to explain the differences observed among the taxa under study. A maximum-likelihood approach to phylogenetic inference evaluates the probability that the chosen evolutionary model has generated the observed data. The evolutionary model could simply mean that one assumes that changes between all nucleotides (or amino acids) are equally probable. The program will then assign all possible nucleotides to the internal nodes of the tree in turn and calculate the probability that each such sequence would have generated the data (if two sister taxa have the nucleotide ‘‘A,’’ a reconstruction that assumes derivation from a ‘‘C’’ would be assigned a low probability compared with a derivation that assumes there already was an ‘‘A’’). The probabilities for all possible reconstructions (not just the more probable one) are summed up to yield the likelihood for one particular site. The likelihood for the tree is the product of the likelihoods for all alignment positions in the data set.

3.6 Bayesian Markov Chain Monte Carlo (MCMC) methods

Bayesian methods are character-state methods that use an optimality criterion, but they are conceptually very different from MP and ML in that they do not attempt to search only for the single best tree. Bayesian methods also employ the concept of likelihood, but by targeting a probability distribution of trees, they search for a set of plausible trees or hypotheses for the data. This posterior distribution of trees inherently holds a confidence estimate of any evolutionary relationship. Bayesian methods require the researcher to specify a prior belief, which is formalized as a prior distribution on the model parameters, i.e. substitution model parameters, branch lengths and tree topology. The relative evidence present in the data is then used to evaluate how one should update his/her prior belief. If no appropriate biological information is available, the prior belief is preferably vague or uninformative. In Bayesian phylogenetic inference, a uniform prior on topology is the most objective (e.g. every tree topology is assumed to be equally likely before looking at the data). Posterior probabilities are obtained by exploring tree space using a sampling technique, called Markov Chain Monte Carlo (MCMC). This sampling method starts by simulating a random set of parameters and proposes a new "state," which is a new set of parameters, by changing the parameters to some extent using random operators. In each step, the likelihood ratio and prior ratio is calculated for the new state relative to the current state. When the combined product is better, the parameters are accepted and a next step is proposed; if the outcome is worse, the probability that the state is rejected is inversely proportional to how much worse the new state is. After an initial convergence to a set of probable model/tree solutions ("burn-in", which needs to be discarded), it is hoped that this stochastic algorithm samples from the "posterior" probability distribution (Philippe *et al.* 2009).

3.7 Bootstrap

Bootstrapping is a resampling tree evaluation method that works with distance, parsimony, likelihood and just about any other tree derivation method. It was invented in 1979 (Efron 1979) and introduced as a tree evaluation method in phylogenetic analysis by Felsenstein (1985) (Felsenstein 1985). The result of bootstrap analysis is typically a number associated with a particular branch in the phylogenetic tree that gives the proportion of bootstrap replicates that supports the monophyly of the clade. Bootstrapping can be considered a two-step process comprising the generation of (many) new data sets from the original set and the computation of a number that gives the proportion of times that a particular branch (e.g., a taxon) appeared in the tree. That number is commonly referred to as the bootstrap value. New data sets are created from the original data set by sampling columns of characters at random from the original data set with replacement. ‘‘With replacement’’ means that each site can be sampled again with the same probability as any of the other sites. As a consequence, each of the newly created data sets has the same number of total positions as the original data set, but some positions are duplicated or triplicated and others are missing. It is therefore possible that some of the newly created data sets are completely identical to the original set - or, on the other extreme, that only one of the sites is replicated, like, 500 times, whereas the remaining 499 positions in the original data set are dropped.

Although it has become common practice to include bootstrapping as part of a thorough phylogenetic analysis, there is some discussion on what exactly is measured by this method. It was originally suggested that the bootstrap value is a measure of repeatability (Felsenstein 1985). In more recent interpretations, it has been considered to be a measure of accuracy - a biologically more relevant parameter that gives the probability that the true phylogeny has been recovered. On the basis of simulation studies, it has been suggested that, under favorable conditions (roughly equal rates of change, symmetric branches), bootstrap values greater than 70% correspond to a probability of greater than 95% that the true phylogeny has been found (Hillis & Bull 1993). By the same token, under less favorable conditions, bootstrap values greater than 50% will be overestimates of accuracy (Hillis & Bull 1993). Simply under certain conditions, high bootstrap values can make the wrong phylogeny look good; therefore, the conditions of the analysis must be considered. Bootstrapping can be used in experiments in which trees are recomputed after internal branches are deleted one at a

time. The results provide information on branching orders that are ambiguous in the full data set (Leipe *et al.* 1994).

The following figure illustrates the basic steps in any bootstrap analysis. Sample datasets are automatically generated from an original dataset. Trees are then estimated from each sample dataset. The results are compiled and compared to determine a bootstrap consensus tree.

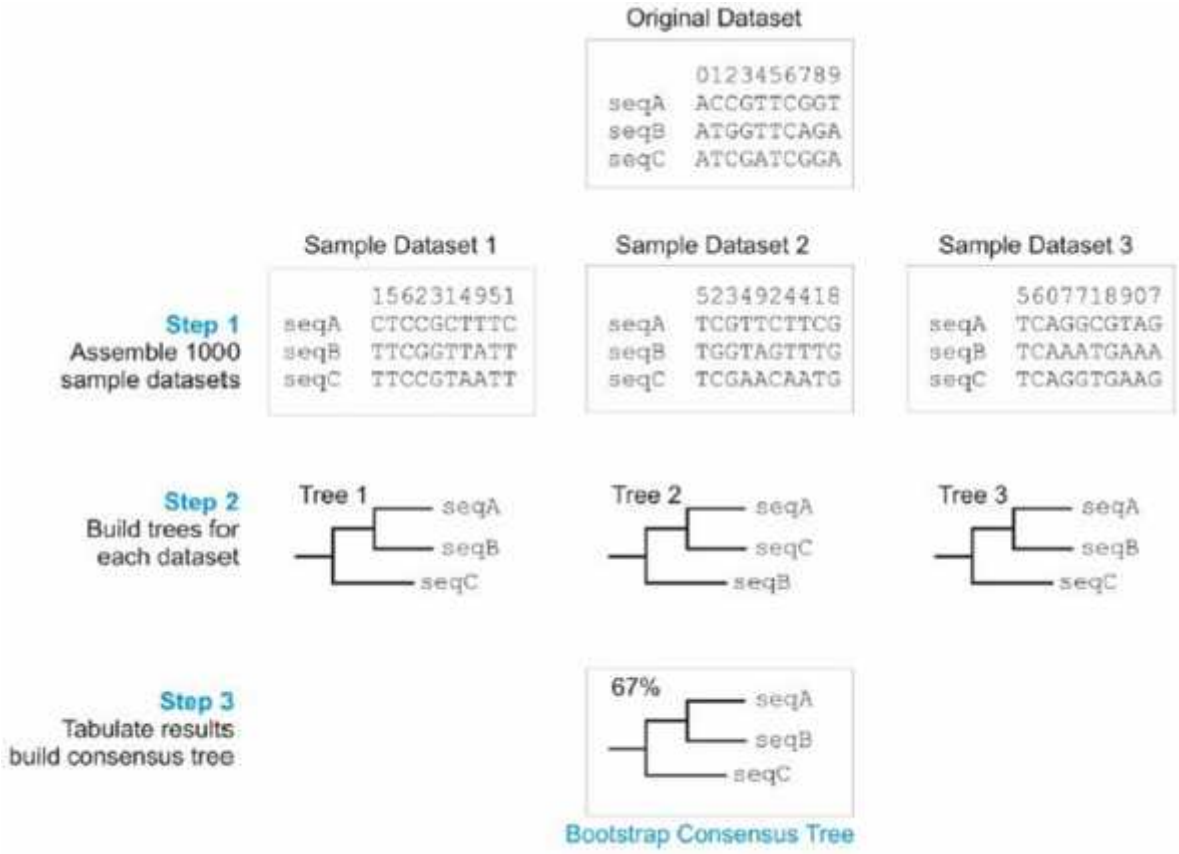


Figure: Steps in a phylogenetic tree bootstrap analysis (Baldauf 2003).

3.8 Jackknifing

An alternative resampling technique often used to evaluate the reliability of specific clades in the tree is the so-called delete-half jackknifing or jackknife. Jackknife randomly purges half of the sites from the original sequences so that the new sequences will be half as long as the original. This resampling procedure typically will be repeated many times to generate numerous new samples. Each new sample (i.e. new set of sequences) - no matter whether from bootstrapping or jackknifing - will then be subjected to regular phylogenetic reconstruction. The frequencies of subtrees are counted from reconstructed trees. If a subtree appears in all reconstructed trees, then the jackknifing *value* is 100%; that is, the strongest possible support for the subtree. As for bootstrapping, branches supported by a jackknifing *value* less than 70% should be treated with caution.

Chapter 4

Some Reported Phylogenetic Tree of NA Protein and their Interpretation

4.1 Phylogenetic relationships between representative H5N1 influenza A virus genes, NA, nucleotide positions 1 to 1011.

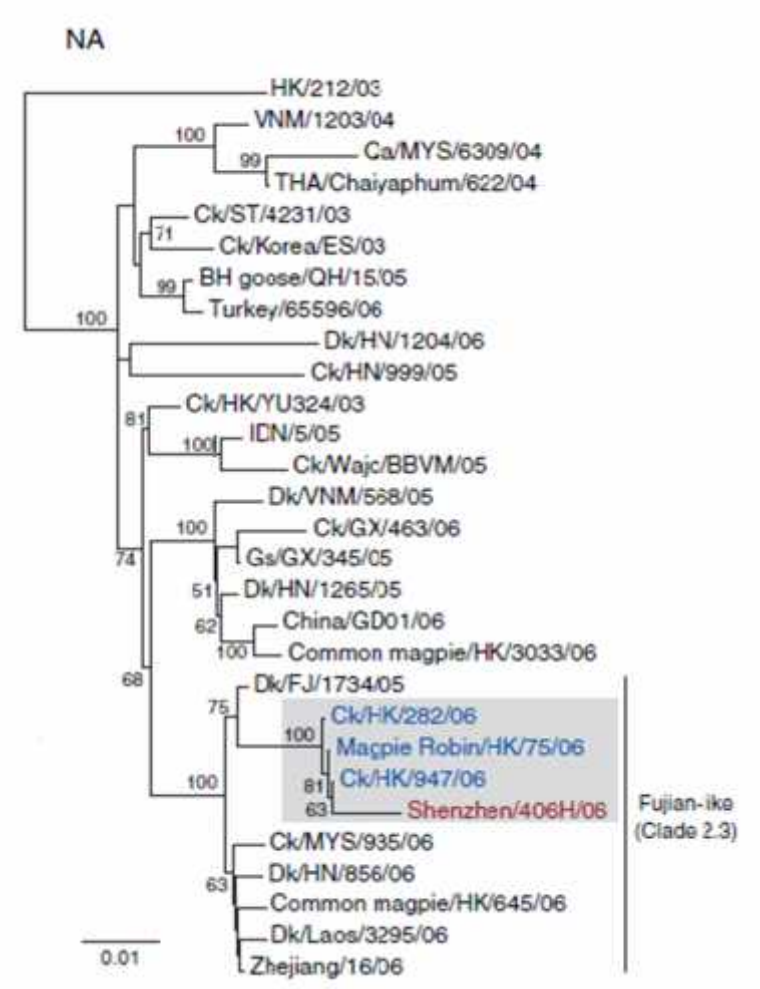


Figure: Phylogenetic relationships between representative H5N1 influenza A virus genes. NA, nucleotide positions 1 to 1410.

Phylogenetic analysis showed that the NA genes of the virus Shenzhen/406H/06 fell into the Fujian-like (Clade 2.3) sublineage. The virus was most closely related to those viruses isolated from wild birds and smuggled chicken in Hong Kong in February 2006. These viruses formed a distinguishable subgroup within the Fujian-like sublineage. This Fujian-like sublineage was in turn most closely related to other viruses from throughout southern China. It was noted that in the NA gene tree, the H5N1 influenza viruses isolated from several other human cases in different provinces since November 2005, including Anhui/1/05, Zhejiang/16/06 and Guangdong/1/06, all clustered into the Fujian-like sublineage (Chen *et al.* 2008).

4.2 Phylogenetic relationships among N1 neuraminidase (NA) genes of H5N1 influenza viruses.

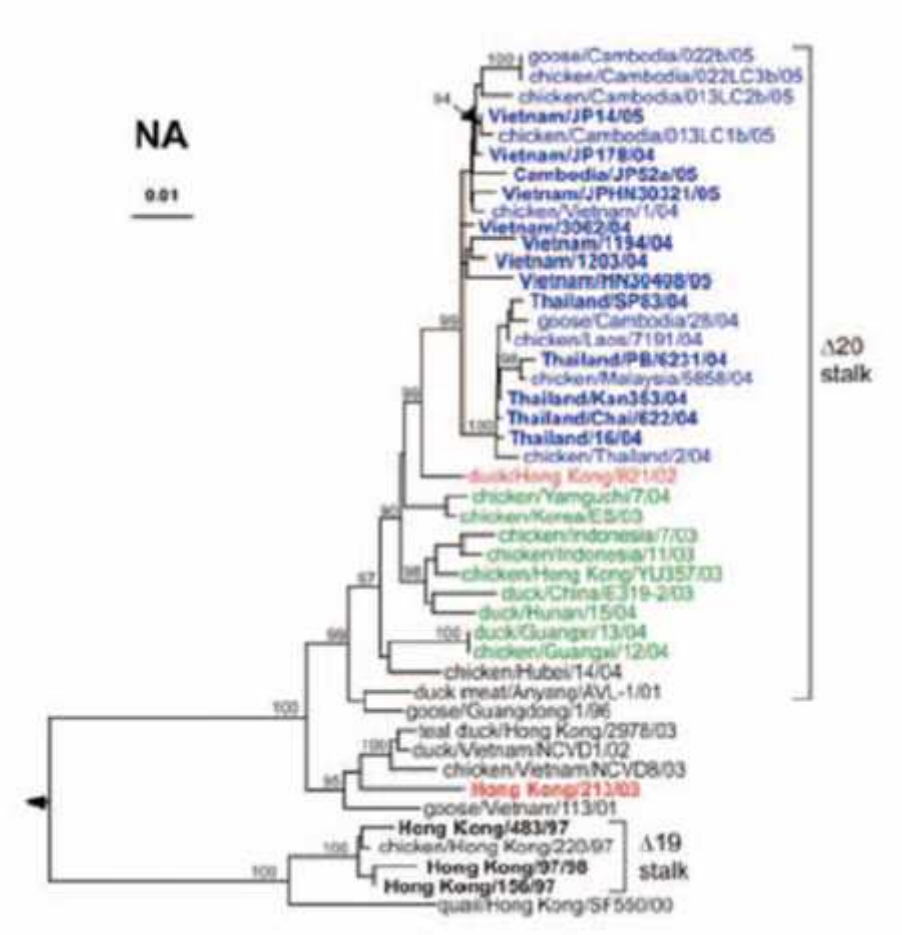


Figure: Phylogenetic relationships among N1 neuraminidase (NA) genes of H5N1 influenza viruses.

The phylogenetic tree of the NA genes resembled that of the HA genes, which indicates coevolution of these 2 envelope genes (Figure 2). NA genes of isolates from Thailand seem to have diverged to form a group distinct from that of genes from Vietnam viruses. As reported Previously, the NA of HK/213/03 did not co-evolve with the HA genes (Guan *et al*. 2004). NA genes from human and related avian H5N1 isolates from 2003–2005 as well as clade 3 isolates were characterized by deletions in the stalk region of the protein (positions 49–68 for clades 1–2 and 54–72 for clade 3) (Bender *et al*. 1999). Deletions in the stalk of the NA are thought to increase retention of virions at the plasma membrane (Matrosovich *et al*. 1999), to balance weaker binding of sialic acid receptors by the HA with newly acquired N154 glycosylation.

4.3 Phylogenetic tree constructed by Bayesian analysis of the neuraminidase gene segment of representative infl uenza viruses A (H5N1)

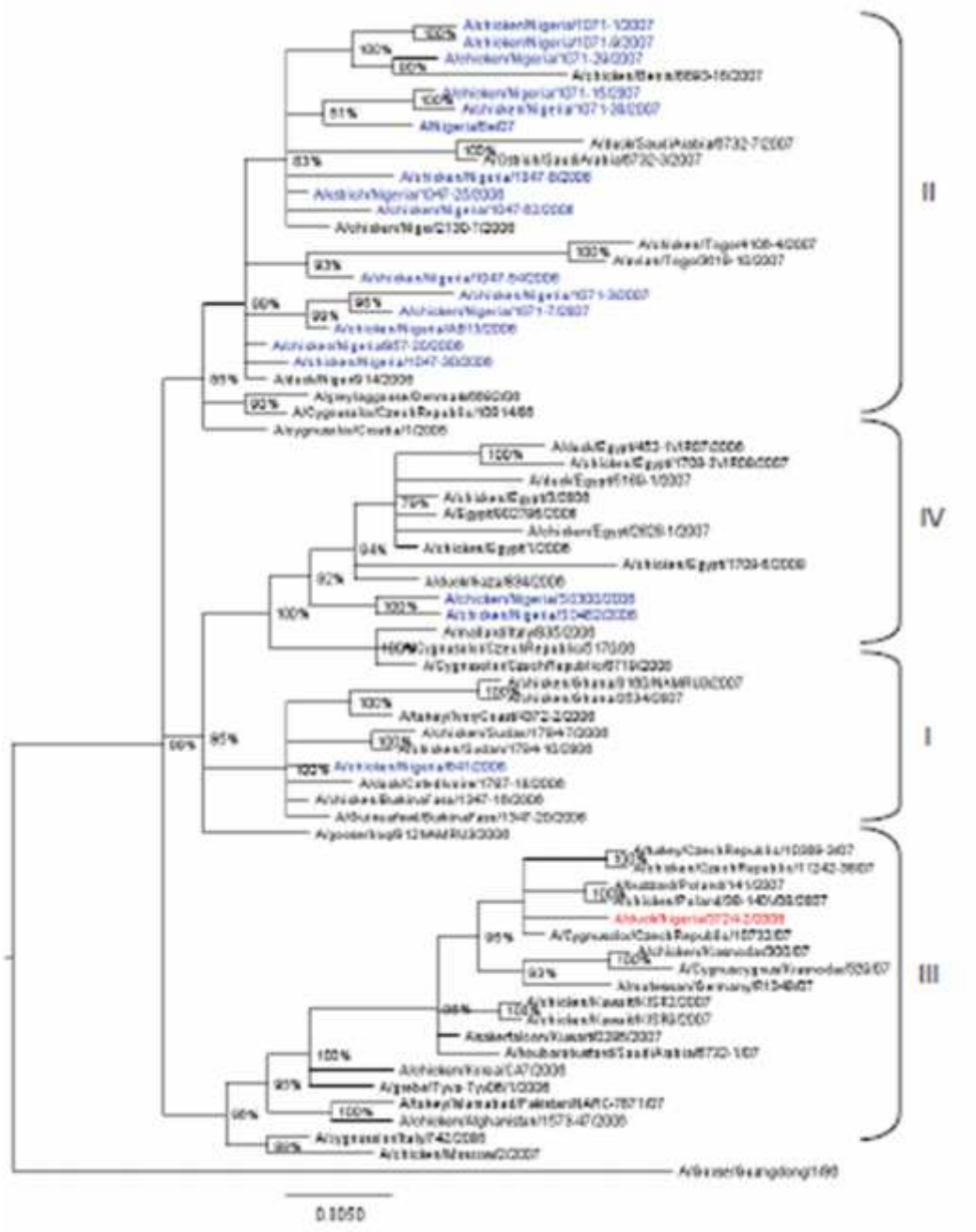


Figure: Phylogenetic tree constructed by Bayesian analysis of the neuraminidase gene segment of representative infl uenza viruses A (H5N1) from Africa, Europe, and the Middle East.

Phylogenetic analysis of the NA gene segment of A/duck/Nigeria/3724-2/2008 supported these results (Figure). Sequence analysis of the 8 gene segments of A/duck/ Nigeria/3724-2/2008 showed the highest similarity at the nucleotide level with the HPAI (H5N1) virus isolate A/*Cygnus olor*/Czech Republic/10732/2007 (99.3% for HA, 99.8% for NA, 99.7% for nonstructural protein, 99.4% for polymerase basic protein 1 [PB1], 99.7% for polymerase basic protein 2 [PB2], 99.8% for nucleoprotein [NP], 99.3% for polymerase acidic protein [PA], and 100% for matrix [MA] protein) and with the HPAI (H5N1) strains from Romania collected in 2007 (99.3%, only HA gene segment is available). For the HA protein, only 3 amino ac- ids differences were observed between the Nigerian isolate and the strain A/*Cygnus olor*/Czech Republic/10732/2007. Lower similarities (ranging from 96.9% to 98% for the HA gene) were shown with previous isolates from Nigeria. No molecular changes were associated with increased affi nity toward α 2,6 linkage sialic acid substrates in the HA receptor- binding domain (Matrosovich *et al*. 1997) or mutations related to resistance to NA inhibitors and to adamantanes were observed in the HA, NA, and M2 genes of the Nigerian isolate. Analysis of the amino acid sequences of the internal proteins of A/ duck/Nigeria/3724-2/2008 virus showed the amino acid lysine at position 627 of the PB2 gene known to be associated with increased virulence of HPAI (H5N1) virus in Mice (Subbarao *et al*. 1993) and 1 amino acid mutation at position 33 (V33I) of the NP gene, which is described as genetic signature of human infl uenza A virus (Chen *et al*. 2010). The PB1-F2 protein had 1 mutation at position 66 (N66S), previously observed only in the Hong Kong 1997 subtype H5N1 viruses and in the 1918 pandemic strain (A/Brevig Mission/18) and is associated to high pathogenic phenotype in mice (Conenello *et al*. 2007).

Chapter 5

Objectives, Download Links of Software, Version, Operation Manual, Methods, Applications and Limitations

5.1 Objectives of the study

- To understand the genetic patterns of avian influenza virus (H5N1) in the year of 2005, 2006 and 2007,
- To identify the mutations in amino acids and binding sites of H5N1,
- To discover the antigenic shift and drug resistance of seasonal and pandemic influenza A virus in H5N1 strain.

5.2 Download links & Version

Geneious (Smith 2014) is free to download from http://geneious.com/download. At the starting time of this study there was Geneious version 7.1.3. So we have used Geneious version 7.1.3 in this study.

5.3 Geneious setup (Operational Manual)

5.3.1 User preferences

User preferences can be changed by going to **Tools →Preferences**. This window can also be opened using the shortcut keys Ctrl+Shift+P. In the user preferences we can change data storage, memory and connection settings, install plugins, customize the appearance and behavior of Geneious, define shortcut keys and set up sequencing profiles.

The tabs in the Preferences window are as follows:

5.3.2 General

This tab contains general setup options:

- **Data storage location** - Shows the location of Geneious database.
- **Search history** - Allows to clear search history
- **Max memory available to Geneious** - Allows changing the RAM allocated to Geneious.

5.3.3 Plugins and Features

Install plugins and customize features in this tab. This tab can also be access via **Tools→ Plugins.**

5.3.4 Appearance and Behavior

In the **Appearance** panel we can change the way the main toolbar and the document table look, and also show or hide tips and the memory usage bar.

In the **Behavior** panel we can change the way newly created documents are handled, such as where they should be saved to and whether they are selected straight away. We can choose whether to store document history, and create active parent/descendant links.

5.3.5 Keyboard

This section contains a list of Geneious functions and allows us to define keyboard shortcuts to them. Shortcuts that are already defined are highlighted in blue. Setting shortcuts can help us quickly navigate through Geneious without using the mouse and also allows us to redefine shortcuts to ones we may be familiar with from other programs.

Double click on a function to bring up a window to enter new keyboard shortcut. If we use one that is already assigned, Geneious will tell what function currently has that shortcut.

5.3.6 NCBI

Here we can set the URL for the NCBI BLAST database and specify which field of the GenBank document should be copied to the ”Name” field in Geneious.

5.3.7 Choosing where to store data

Geneious stores data in a folder called Geneious X.Y Data (where X and Y are the version of Geneious we are using), which is stored separately from the application itself. When Geneious first starts up we will be asked to choose a location for this folder. The default location in the user’s home directory is normally the best option. Although it’s possible to store data on a network or USB drive so we can access it from other computers, this is not recommended because it can have adverse effects on performance.

To store data somewhere different to the default, simply we need to click the ‘Select’ button in the welcome window and choose an empty folder on drive where we would like to store our data. The data location can also be changed later by going to the **General** tab under **Tools → Preferences** in the menu and changing the **Data Storage Location** option. Geneious will offer to copy our existing data across to the new location if appropriate.

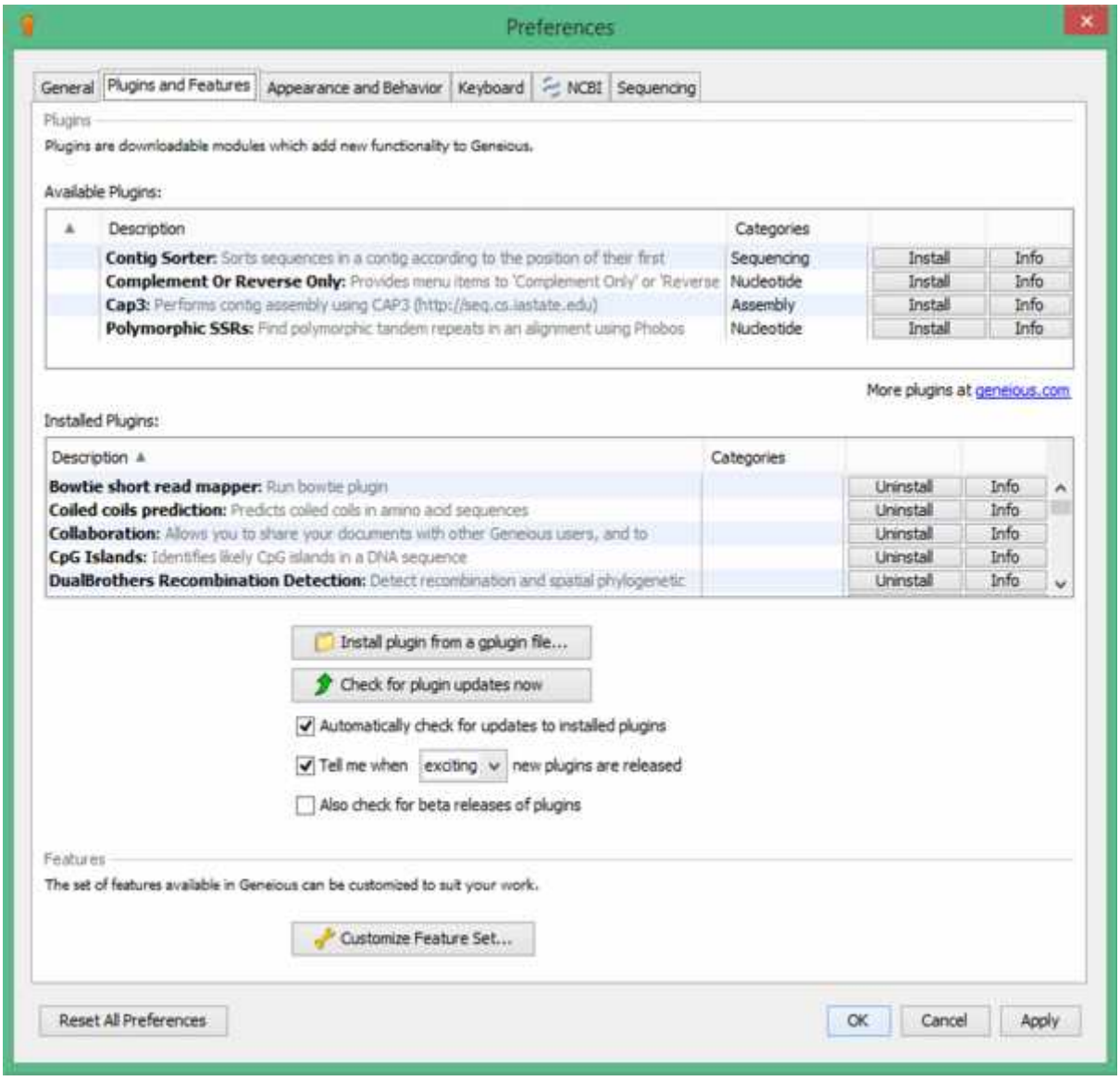


Figure: The plugins preferences in Geneious.

The Geneious main window

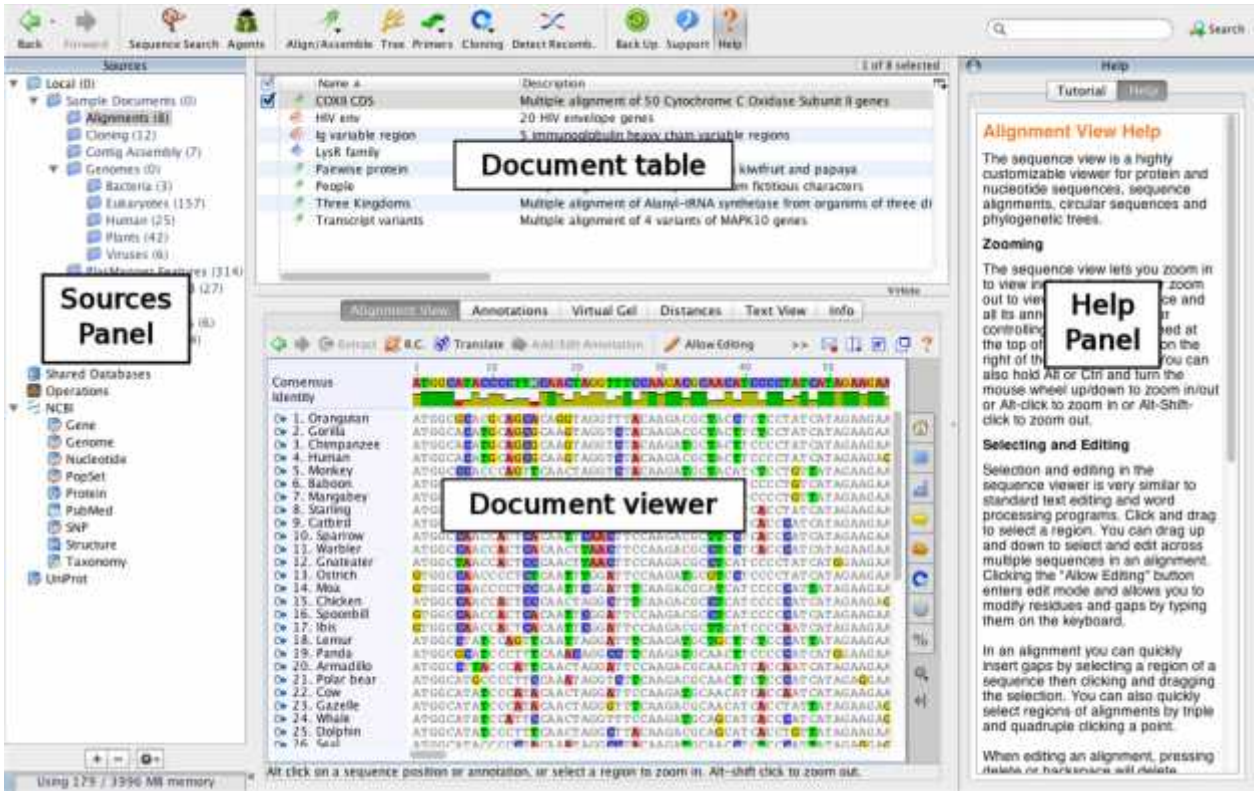


Figure: The main window in Geneious.

5.3.8 The Sources Panel

The Sources Panel displays us stored documents and contains the services Geneious offers for storing and retrieving data. The plus (+) symbol indicates that a folder contains sub-folders. A minus (-) indicates that the folder has been expanded, showing its sub-folders.

Geneious Sources Panel allows us to access:

- Local Documents.
- NCBI databases - Gene, Genome, Nucleotide, PopSet, Protein, Pubmed, SNP, Structure and Taxonomy.
- An EMBL database - Uniprot.
- Shared databases, if set up.

5.3.9 The Document Table

The Document Table displays summaries of downloaded and imported data such as DNA sequences, protein sequences, journal articles, sequence alignments and trees. This information is presented in table form. By clicking on the search icon we can search data for text or by sequence similarity (BLAST). We can enter a search string into the “Filter” box located at the right side of the toolbar; this will hide all documents that do not contain the search string.

While search results usually contain documents of a single type, a local folder may contain any mixture of documents, whether they are sequences, publications or other types.

Selecting a document in the Document Table will display its details in the Document View Panel. Selecting multiple documents will show a view of all the selected documents if they are of similar types, e.g. selecting two sequences will show both of them side-by-side in the sequence view.

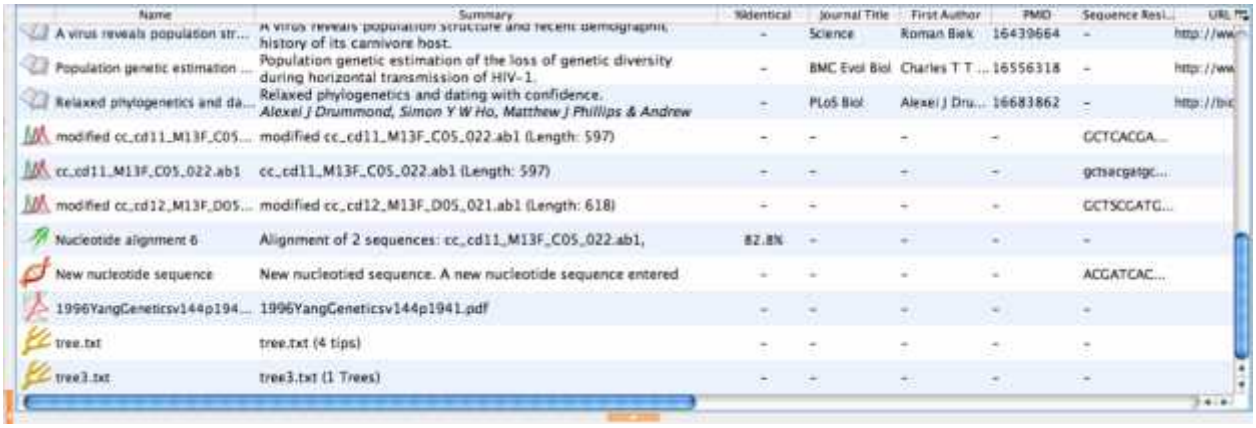


Figure: The document table, when browsing the local folders.

The easiest way to select multiple documents is by clicking on the checkboxes down the left-hand side of the table. Double-clicking a document in the Document Table displays the same view in a separate window.

The Document Table has some useful features.

Editing. Values can be typed into the columns of the table. This is a useful way of editing the information in a document.

Copying. Column values can be copied. This is a quick method of extracting searchable information such as an accession number.

Sorting. All columns can be alphabetically, numerically or chronologically sorted, depending on the data type.

Managing Columns. We can reorder the columns to suit by clicking on the column header and dragging it to the desired horizontal position.

5.3.10 The Document Viewer Panel

The Document Viewer Panel shows the contents of any document clicked on in the document table, allowing us to view sequences, alignments, trees, 3D structures, journal article abstracts and other types of documents in a graphical or plain text view. Many document viewers allow us to customize settings such as zoom level, color schemes, layout and annotations (nucleotide and amino acid sequences); three different layouts, branch and leaf labeling (tree documents); and many more. When viewing journal articles, this panel includes direct link to Google Scholar.

To view large documents, it is sometimes better to double click on them. This opens a view in a new window. In the document viewer panel there are two tabs that are common to most types of documents: “Text view” and “Info”. “Text view” shows the document’s information in text format. The exception to this rule occurs with PDF documents where the user needs to either click the “View Document” button or double-click to view it.

Most viewers have own small toolbar at the top of the document viewer panel. This always has five buttons on the far right:

Figure: Document Viewer Panel.

- "Split View" which opens a second viewer panel of the same document. Selection is synchronized between these two views.
- "Expand Document View" which expands the viewer panel out to fill the entire main window. Clicking again will return the viewer to normal size.
- "Open Document in New Window" will open a new view of the selected document in a new, separate window.
- "Help" opens the Help Panel and displays some short help for the current viewer.

5.3.11 The Help Panel

The Help Panel has a "Help" tab and a "Tutorial" tab.

The Help tab provides information about the service we are currently using or the viewer we are currently viewing. The help displayed in the help tab changes as we click on different services and choose different viewers.

The Tutorial is aimed at first-time users of Geneious and has been included to provide a feel for how Geneious works.

The Help panel can be closed at any time by clicking the button in its top corner, or by toggling the 'Help' button in the Toolbar.

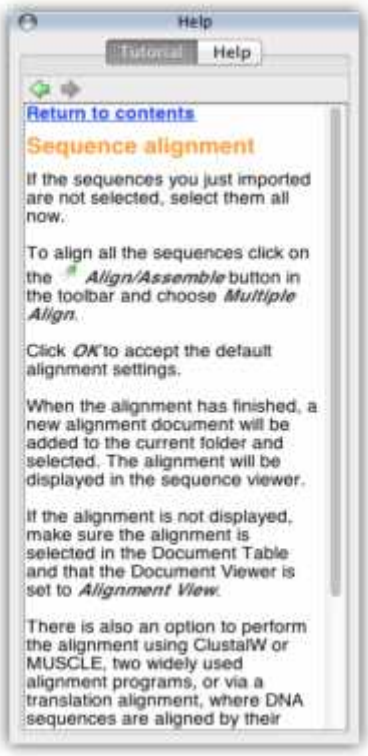


Figure: The Help Panel.

5.3.12 The Toolbar

The toolbar contains several icons that provide shortcuts to common functions in Geneious, including Sequence Search (eg. BLAST), Agents that search databases for new content even while we sleep, Align/Assemble, Tree building, and Help. We can alter the contents of the toolbar to suit our own needs. The icons can be displayed small or large, and with or without their labels. The Help icon is always available.

The “Back” and “Forward” options help us to move between previous views in Geneious and are analogous to the back and forward buttons in a web browser.

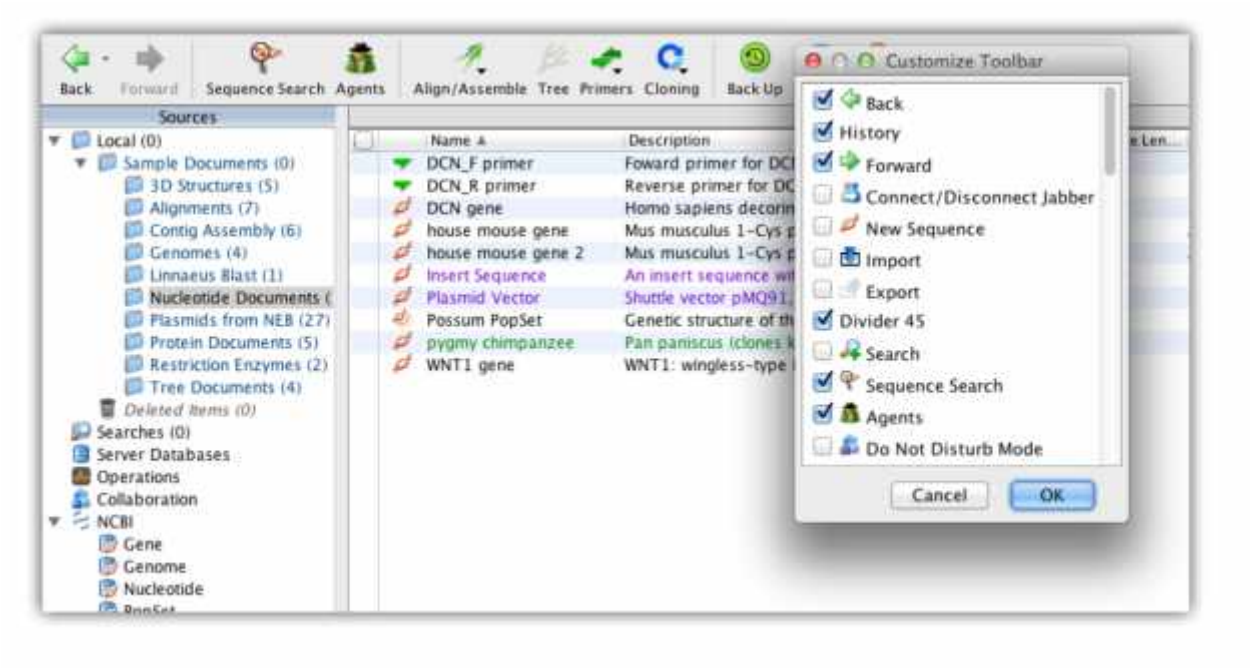


Figure: The Toolbar.

The toolbar can be customized by right-clicking on it. This gives a popup menu with the following options:

- “Show Labels” Turn the text labels on or off.
- “Large Icons” Switch between large and small icons.
- “Customize” which lists all available toolbar buttons. Selecting/deselecting buttons will show/hide the buttons in the toolbar.

5.3.13 Status bar

Below the Toolbar, there is a grey status bar. This bar displays the status of the currently selected service. For example, when we are running a search, it displays the number of matches, and the time remaining for the search to finish.

5.3.14 The Menu Bar

The Menu Bar has seven main menus “File”, “Edit”, “View”, “Tools”, “Sequence”, “Annotate & Predict” and “Help”.

5.3.15 File Menu

This contains some standard “File” menu items including printing and “Exit” on Windows. It also contains options to create, rename, delete, share and move folders and Import/Export options.

5.3.16 Edit Menu

Here we will find common editing functions including “Cut”, “Copy”, “Paste”, “Delete” and “Select All”. These are useful when transferring information from within documents to other locations, or exporting them. This menu also contains “Find in Document”, “Find Next” and “Find Previous” options. Find can be used to find text or numbers in a selected document. This is useful when looking for annotated regions or a stretch of bases in a sequence. This opens a “Find Dialog”. The shortcut to this is Ctrl+F. Next finds the next match for the text specified in the “Find” dialog. The shortcut keys are F3 or Ctrl+G. Geneious then allows us to choose another document and continue searching for the same search word. Prev finds the previous match. The shortcut keys for this are Ctrl+Shift+G or Shift+F3. There are also the useful “Find Duplicates...” and “Batch Rename...” features in this menu.

5.3.17 View Menu

This contains several options and commands for changing the way of viewing data in Geneious:

- **Back**, **Forwards** and **History** allow us to return to documents we had selected previously.
- **Next unread document** selects the next document in the current folder which is unread.
- **Table Columns** contains the same functionality as the popup menu for the document table header.
- **Open document in new window** opens a new window with a view of the currently selected document(s).
- **Expand document view** expands the document viewer panel in the main window out to fill the entire main window. Selecting this again to return to normal.
- **Split Viewer Left/Right** creates a second copy of the document viewer with the two views laid out side by side.
- **Split Viewer Top/Bottom** creates a second copy of the document viewer with one on top of the other.
- **Document Windows** Lists the currently open document windows. Selecting one from this menu will bring that document window to the front.

5.3.18 Tools Menu

- **Align/Assemble**
- **Tree**
- **Primers**
- **Cloning**
- **Sequence Search** - Perform a sequence search (such as NCBI Blast) using the currently selected sequence as the query.
- **Add/Remove Databases**
- **Extract Annotations** - Searching the selected sequences or alignments for annotations which match certain criteria then extract all of the matching annotations to separate sequence documents or a sequence list. Includes the option to concatenate all matches in each sequence into one sequence document. Useful for extracting a certain gene from a group of genomes.

- **Strip Alignment Columns** - Creates a new alignment document with some columns (for example all identical columns or all columns containing only gaps) stripped.
- **Concatenate Sequences or Alignments** - Joins the selected sequences or alignments end-on-end, creating a single sequence or alignment document from several. After selecting this operation we are given the option to choose the order in which the sequences or alignments are joined. We can also choose whether the resulting document is linear or circular, and, if one or more of the component sequences was an extraction from over the origin of a circular sequence, we can choose to use the numbering from that sequence, thus producing a circular sequence with its origin in the same place as the original circular sequence. Overhangs will be taken into account when concatenating.
- **Generate Consensus Sequence** - Generates a consensus sequence for the selected sequence alignment and saves it to a separate sequence document.
- **Plugins** - Takes us to the Plugins menu where we can install or uninstall plugins.
- **Preferences.**

5.3.19 Sequence Menu

This contains several operations that can be performed on Protein and Nucleotide sequences as well as Sequence Alignments in some cases.

- **New Sequence:** Creates a new nucleotide or protein sequence from residues that we can paste or type in.
- **Extract**: Extract the selected part of a sequence or alignment into a new document.
- **Reverse Complement**: Reverse sequence direction and replace each base by its complement. This is available only for nucleotide sequences. When only part of a sequence is selected, we can choose to either reverse-complement only the selected region and extract it to a new sequence document, or reverse complement the entire sequence. On alignment or contig documents can reverse complement individual sequences within the alignment or assembly by selecting that sequence, and choosing “reverse complement selected sequence”.
- **Translate**: Creates a new protein document from the translated DNA. Clicking on this choice brings up a list of genetic codes that can be used, and a choice of translation frame. This is available only for nucleotide sequences.

- **Back Translate** creates nucleotide version of the selected protein document. We can choose to create either an ambiguous version, or use codon usage tables to create an unambiguous sequence.
- **Circular Sequences** sets whether the currently selected sequences are circular. These effects the way the sequence view displays them as well as how certain operations deal with the sequences (eg. digestion).
- **Free End Gaps Alignment** sets whether the currently selected alignment has free end gaps. This effects calculation of the consensus sequences and statistics.
- **Change Residue Numbering...** changes the "original residue numbering" of the selected sequence. On a linear sequence, this is used to indicate that a sequence is a subsequence of some larger sequence. On a circular sequence, this is used to shift the origin of the sequence.
- **Convert between DNA and RNA** changes all T's in a sequence to U's or vice versa, depending on the type of the selected sequence.
- **Set Paired Reads** sets up paired reads for assembly.
- **Set Read Direction** marks sequences as forward or reverse reads so the correct reads are reverse complemented by assembly.
- **Separate Reads by Barcode** separates multiplex or barcode data (e.g. 454 MID data).
- **Group Sequences into a List** creates a sequence list containing copies of all of the selected sequences. Lists can make it easier to manage large numbers of sequences by keeping related ones grouped in a single document.
- **Extract Sequence from List** copies each sequence out of a sequence list into a separate sequence document.

5.3.20 Help Menu

This consists of the standard Help options offered by Geneious.

Geneious version 7.1.3 can import the following file formats:

5.3.21 CLUSTAL format

The Clustal format is used by ClustalW and ClustalX, two well-known multiple sequence alignment programs.

Clustal format files are used to store multiple sequence alignments and contain the word Clustal at the beginning.

5.3.22 FASTA format

The FASTA file format is commonly used by many programs and tools, including BLAST, TCoffee and ClustalX. Each sequence in a FASTA file has a header line beginning with a ">" followed by a number of lines containing the raw protein or DNA sequence data. The sequence data may span multiple lines and these sequences may contain gap characters. An empty line may or may not separate consecutive sequences.

Here is an example of three sequences in FASTA format (DNA, Protein, Aligned DNA):

>Orangutan
ATGGCTTGTGGTCTGGTCGCCAGCAACCTGAATCTCAAACCTGGAGAGTGCCTTC
GAGTG

>gi|532319|pir|TVFV2E|TVFV2E envelope protein
ELRLRYCAPAGFALLKCNDADYDGFKTNCSNVSVVHCTNLMNTTVTTGLLLNGSYS
ENRTQIWQK

>Chicken
CTACCCCCCTAAAACACTTTGAAGCCTGATCCTCACTA----------
CTGTCATCTTAA

5.3.23 GenBank files

Records retrieved from the NCBI website (http://www.ncbi.nlm.nih.gov) can be saved in a number of formats. Records saved in GenBank or INSDSeq XML formats can be imported into Geneious.

5.3.24 Geneious format

The Geneious format can be used to store all local documents, meta-data types and program preferences. A file in Geneious format will usually have a *.geneious* extension or a *.xml* extension. This format is useful for sharing documents with other Geneious users and backing up your Geneious data.

5.3.25 MEGA format

The MEGA format is used by MEGA (Molecular Evolutionary Genetics Analysis).

5.3.26 Nexus format

The Nexus format was designed to standardize the exchange of phylogenetic data, including sequences, trees, distance matrices and so on. The format is composed of a number of blocks such as TAXA, TREES and CHARACTERS. Each block contains pre-defined fields. Geneious imports and exports files in Nexus format and can process the information stored in them for analysis.

5.3.27 Geneious document types

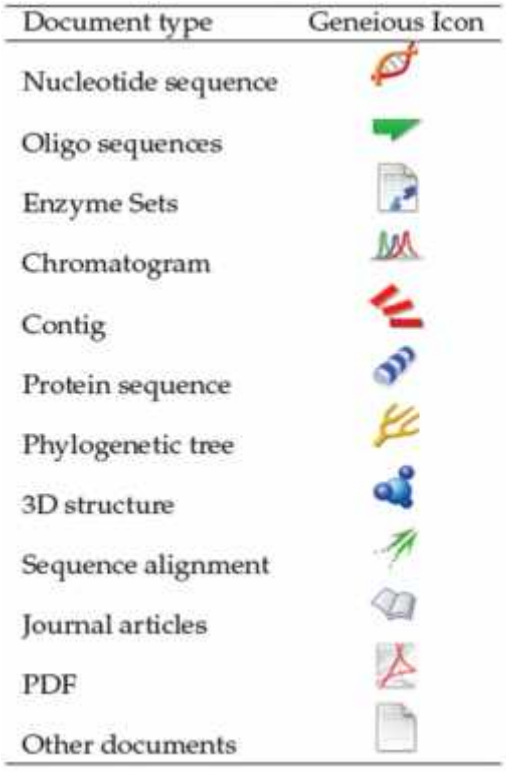


Figure: Geneious document types.

5.4 The Sequence (and alignment) Viewer

The “Sequence view” tab in the Document Viewer panel is available for Nucleotide sequences, Protein sequences, Alignments and 3D structure documents. If an alignment is selected, this will be called “Alignment View” or “Contig View” if a contig is selected. The options available vary with the kind of sequence data being viewed.

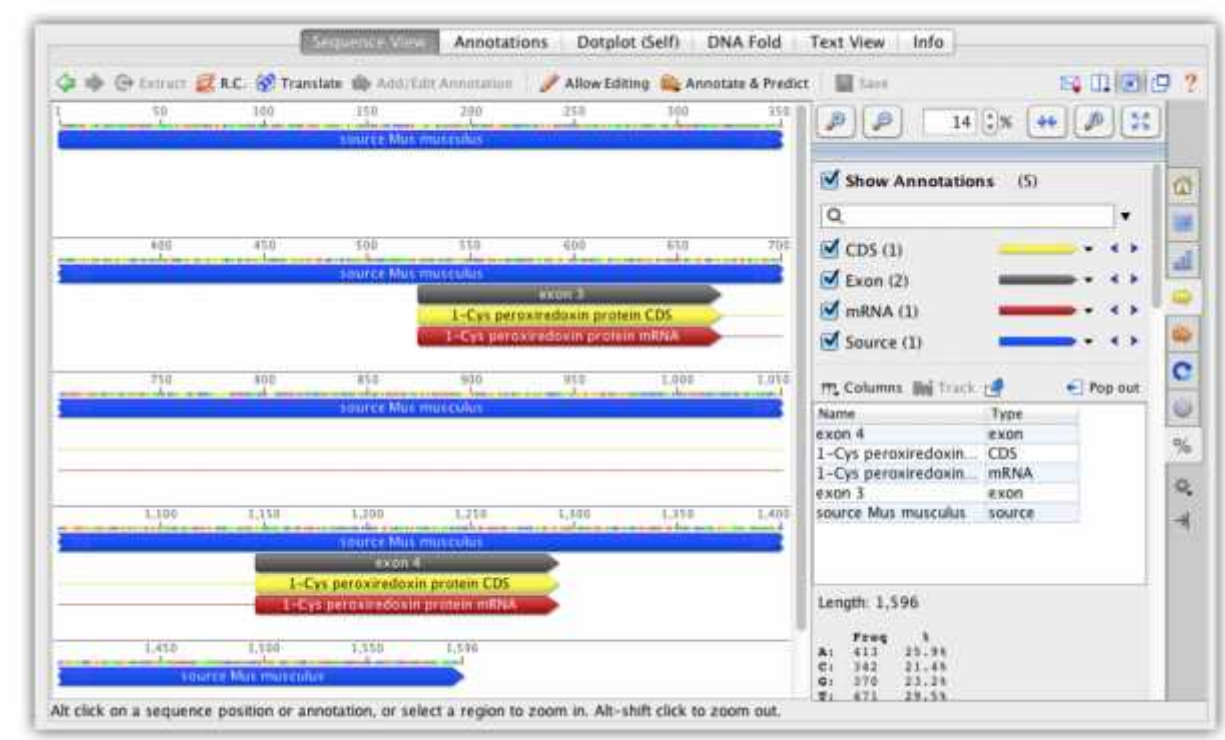


Figure: A view of an annotated nucleotide sequence in Geneious.

5.5 Sequence Alignments

Over evolutionary time, related DNA or amino acid sequences diverge through the accumulation of mutation events such as nucleotide or amino acid substitutions, insertions and deletions.

A sequence alignment is an attempt to determine regions of homology in a set of sequences. It consists of a table with one sequence per row, and with each column containing homologous residues from the different sequences, e.g. residues that are thought to have evolved from a common ancestral nucleotide/amino acid. If it is thought that the ancestral

nucleotide/amino acid got lost on the evolutionary path to one descendant sequence, this sequence will show a special gap character "–" in that alignment column.

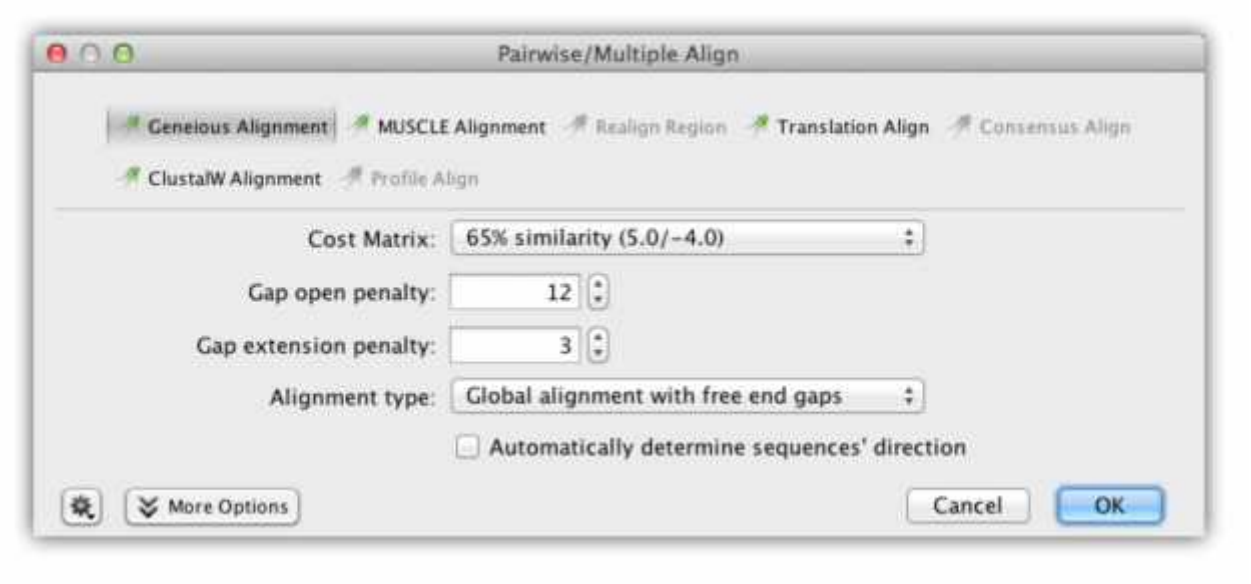


Figure: Options for nucleotide pairwise alignment.

5.6 Building Phylogenetic trees

Geneious provides inbuilt algorithms for Neighbour-joining and UPGMA methods of tree reconstruction, which are suitable for preliminary investigation of relationships between newly acquired sequences. For more sophisticated methods of phylogenetic reconstruction such as Maximum Likelihood and Bayesian MCMC, external plugins for specialist software are available. These can be downloaded from the plugins page on geneious website or within Geneious by going to "Plugins" under the "Tools" menu.

5.7 Phylogenetic tree representation

A phylogenetic tree describes the evolutionary relationships amongst a set of sequences.

Branch length. A measure of the amount of divergence between two nodes in the tree. Branch lengths are usually expressed in units of substitutions per site of the sequence alignment. Nodes or internal nodes of a tree represent the inferred common ancestors of the sequences that are grouped under them.

Tips or leaves of a tree represent the sequences used to construct the tree.

Taxonomic units. These can be species, genes or individuals associated with the tips of the tree.

A phylogenetic tree can be rooted or unrooted. A rooted tree consists of a root, or the common ancestor for all the taxonomic units of the tree. An unrooted tree is one that does not show the position of the root. An unrooted tree can be rooted by adding an outgroup (a species that is distantly related to all the taxonomic units in the tree).

5.8 Tree building in Geneious

To build a tree, we need to select an alignment or a set of related sequences (all DNA or all protein) in the Document table and clicking the "Tree" icon or choosing this option from the Tools menu.

If we are building a simple tree (Neigbour joining or UPGMA) using the Geneious tree builder, the tree can be built directly from a set of unaligned sequences, as the alignment will be built as part of the tree-building process. We can also select an existing tree document (which contains an alignment) and build another tree from that, as the alignment will simply be extracted from the existing tree and used to build the new tree.

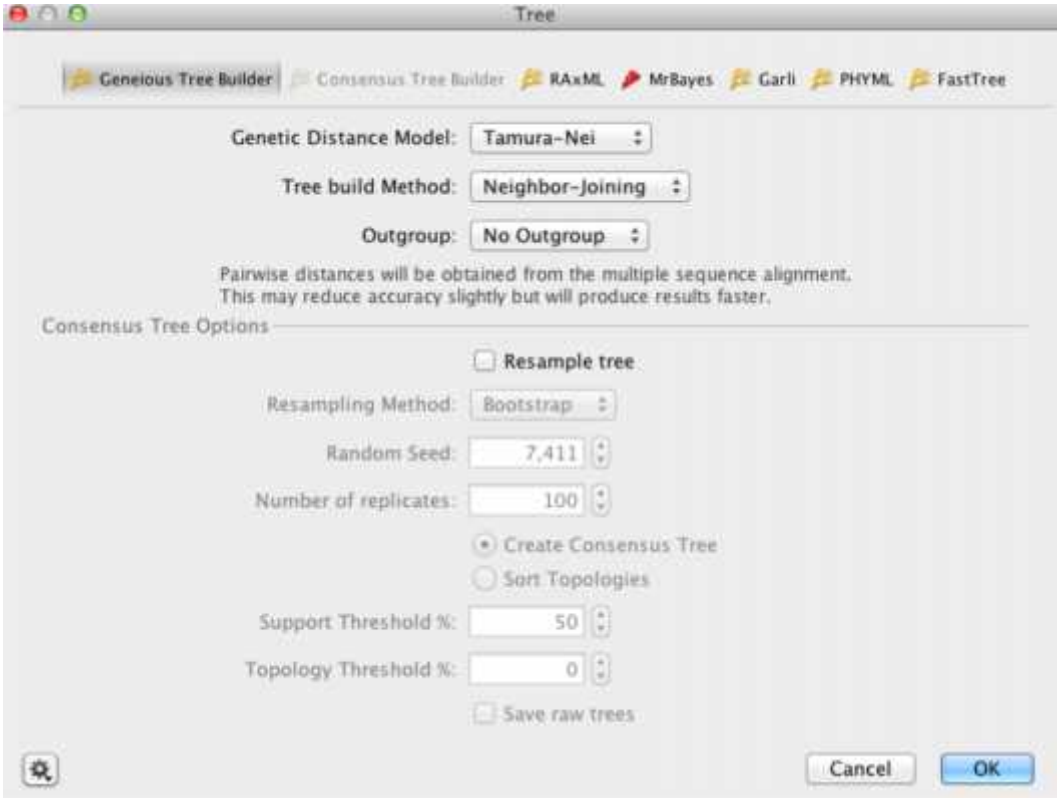


Figure: Tree building options in Geneious.

5.9 Tree building methods and models

5.9.1 Neighbor-joining

In this method, neighbors are defined as a pair of leaves with one node connecting them. The principle of this method is to find pairs of leaves that minimize the total branch length at each stage of clustering, starting with a star-like tree. The branch lengths and an unrooted tree topology can quickly be obtained by using this method without assuming a molecular clock.

5.9.2 UPGMA

This clustering method is based on the assumption of a molecular clock. It is appropriate only for a quick and dirty analysis when a rooted tree is needed and the rate of evolution is does not vary much across the branches of the tree.

5.10 Advanced Tree Building methods

The following plugins are available for running maximum likelihood or Bayesian phylogenetic analyses in Geneious:

5.10.1 MrBayes For Bayesian estimation of phylogenies, runs MrBayes 2.0.9 (http://mrbayes.sourceforge.net/)

5.10.2 PhyML Builds maximum likelihood trees using PhyML 3.0 (http://www.atgc-montpellier.fr/phyml)

5.10.3 GARLI (Genetic Algorithm for Rapid Likelihood Inference): Builds maximum likelihood trees from alignments of 4 or more sequences using Garli 2.0 (https://www.nescent.org/wg_garli/Main_Page).

5.10.4 RAxML (Randomized Axelerated Maximum Likelihood): Rapid maximum likelihood tree-building using RAxML 7.2.8 (http://www.exelixis-lab.org/). Allows partitioned datasets.

5.10.5 FastTree Approximately-maximum-likelihood phylogenetic trees from alignments of nucleotide or protein sequences using Fasttree 2.1.5 (http://www.microbesonline.org/fasttree/). Ideal for large alignments.

5.10.6 PAUP* Builds maximum parsimony and maximum likelihood trees.

5.11 Resampling – Bootstrapping

Resampling is a statistical technique where a procedure (such as phylogenetic tree building) is repeated on a series of data sets generated by sampling from one original data set. The results of analyzing the sampled data sets are then combined to generate summary information about the original data set.

Bootstrapping is the statistical method of resampling with replacement. To apply bootstrapping in the context of tree building, each pseudo-replicate is constructed by randomly sampling columns of the original alignment with replacement until an alignment of the same size is obtained.

A view of a phylogenetic tree in Geneious is given below:

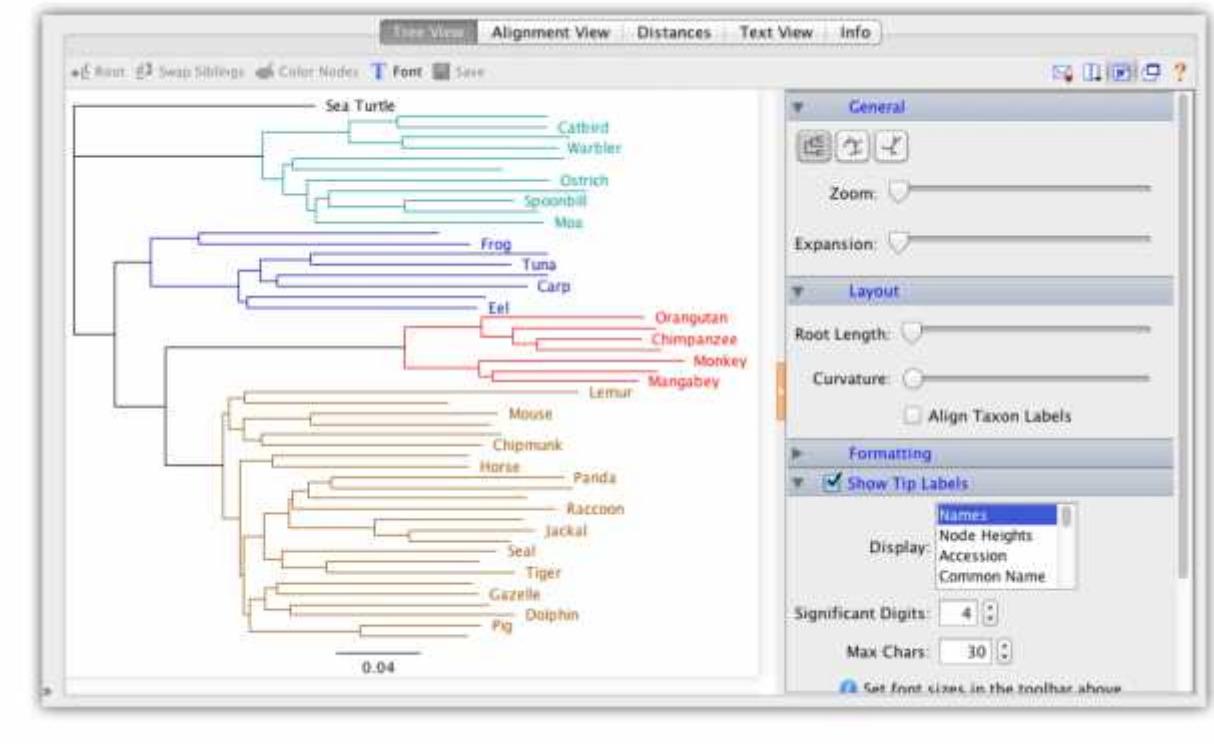


Figure: A view of a phylogenetic tree in Geneious.

5.12 Sequence Searching (BLAST)

BLAST stands for Basic Local Alignment Search Tool. It allows us to query a sequence database with a sequence in order to find entries in the database that contain similar sequences. When "BLAST-ing", we are able to specify either nucleotide or protein sequences and nucleotide sequences can be either DNA or RNA sequences. Sequences can be BLAST-ed against databases held at NCBI, or contained within local geneious database (Custom BLAST).

Geneious gives us the option of searching against a database using either currently selected sequence documents or a sequence we enter manually. If choose to enter sequence manually, then Geneious will display a large text box in which we can enter query sequence as either unformatted text or FASTA format.

5.13 NCBI BLAST

Geneious is able to BLAST to many different databases held at NCBI. These databases are listed in the following Tables and can be selected in the "Databases" drop down menu in the Sequence Search set up dialog.

Database	Nucleotide searches
nr	All non-redundant GenBank+EMBL+DDBJ+PDB sequences(no EST, STS, GSS or HTGS sequences)
genome	Genomic entries from NCBI's Reference Sequence project
est	Database of GenBank + EMBL + DDBJ sequences from EST Divisions
est_human	Human subset of est
est_mouse	Mouse subset of est
est_others	Non-Human, non-mouse subset of est
gss	Genome Survey Sequence, includes single-pass genomic data, exon-trapped sequences, and Alu PCR sequences.
htgs	Unfinished High Throughput Genomic Sequences: phases 0, 1 and 2 (finished, phase 3 HTG sequences are in nr)
pat	Nucleotide sequences derived from the Patent division of GenBank
PDB	Sequences derived from the 3D-structures of proteins from PDB
month	All new/updated GenBank+EMBL+DDBJ+PDB sequences released in the last 30 days.
RefSeq	NCBI-curated, non-redundant sets of sequences.
dbsts	Database of GenBank+EMBL+DDBJ sequences from STS Divisions
chromosome	A database with complete genomes and chromosomes from the NCBI Reference Sequence project.
wgs	A database for whole genome shotgun sequence entries.
env_nt	This contains DNA sequences from the environment, i.e all organisms put together

Database	Protein searches
env_nr	Translations of sequences in env_nt
month	All new/updated GenBank coding region (CDS) translations +PDB+SwissProt+PIR released in last 30 days
nr	All non-redundant GenBank coding region (CDS) translations+PDB+SwissProt+PIR+PRF
pat	Protein sequences derived from the Patent division of GenBank
PDB	Sequences derived from 3D structure Brookhaven PDB
RefSeq	RefSeq protein sequences from NCBI's Reference Sequence Project
SwissProt	Curated protein sequences information from EMBL

Figure: Different database of NCBI.

5.14 Method

We have collected the initial strain from NCBI database. The link is http://www.ncbi.nlm.nih.gov/pubmed.

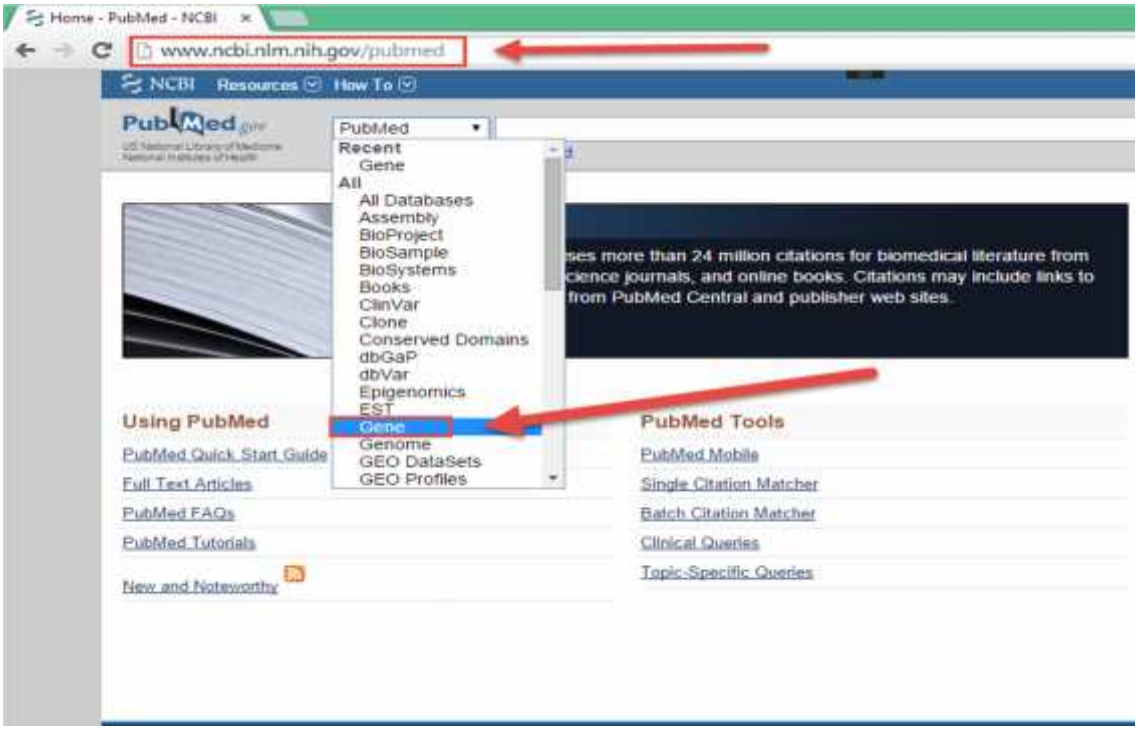


Figure: Initial data collection link (PubMed).

From the dropdown menu we have selected Gene as our desired search query. In the search box we have used the search term “Neuraminidase H5N1 NA”

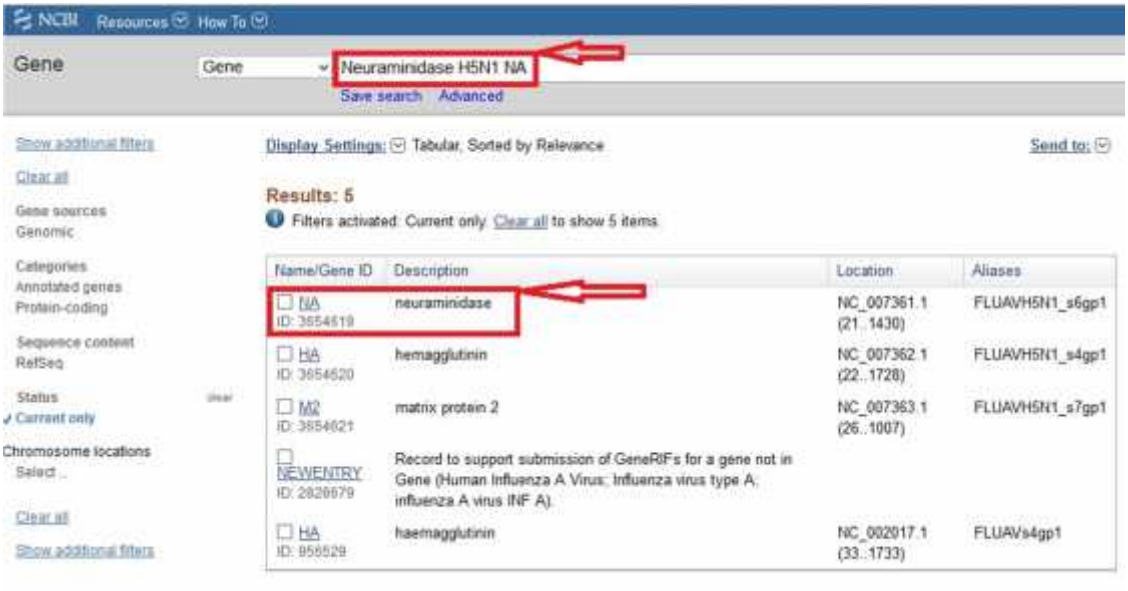


Figure: Search term used in PubMed.

We found some result. We have chosen the 1st result as it seems to us that it is our desired result. We opened the 1st search item in a new window. This should be like the following.

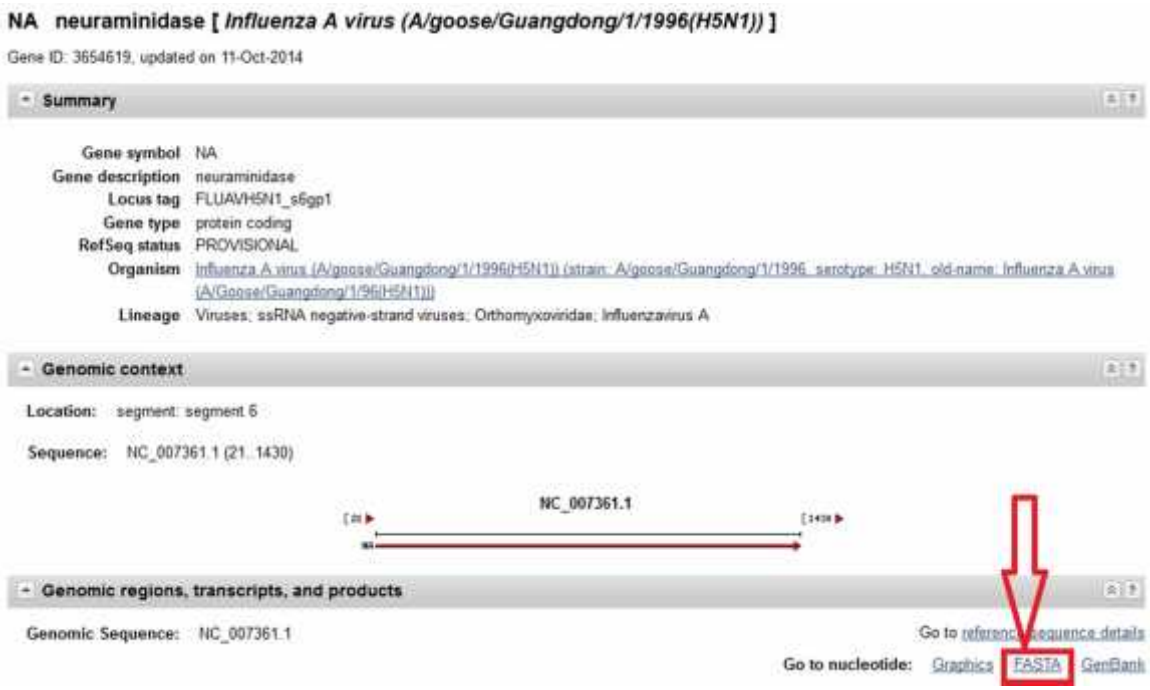


Figure: The way to find out FASTA sequence.

From the new window we will choose the FASTA options to obtain our desired FASTA sequence.

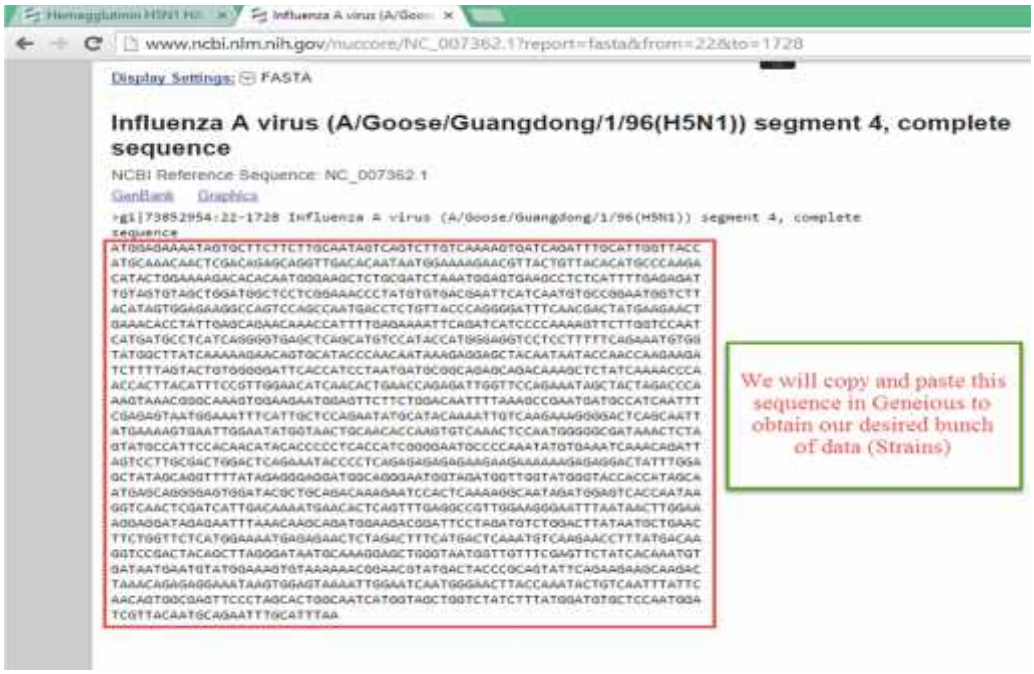


Figure: FASTA sequence.

Now we have our desired FASTA sequence. We will copy and paste this sequence in Geneious to obtain our bunch of strains.

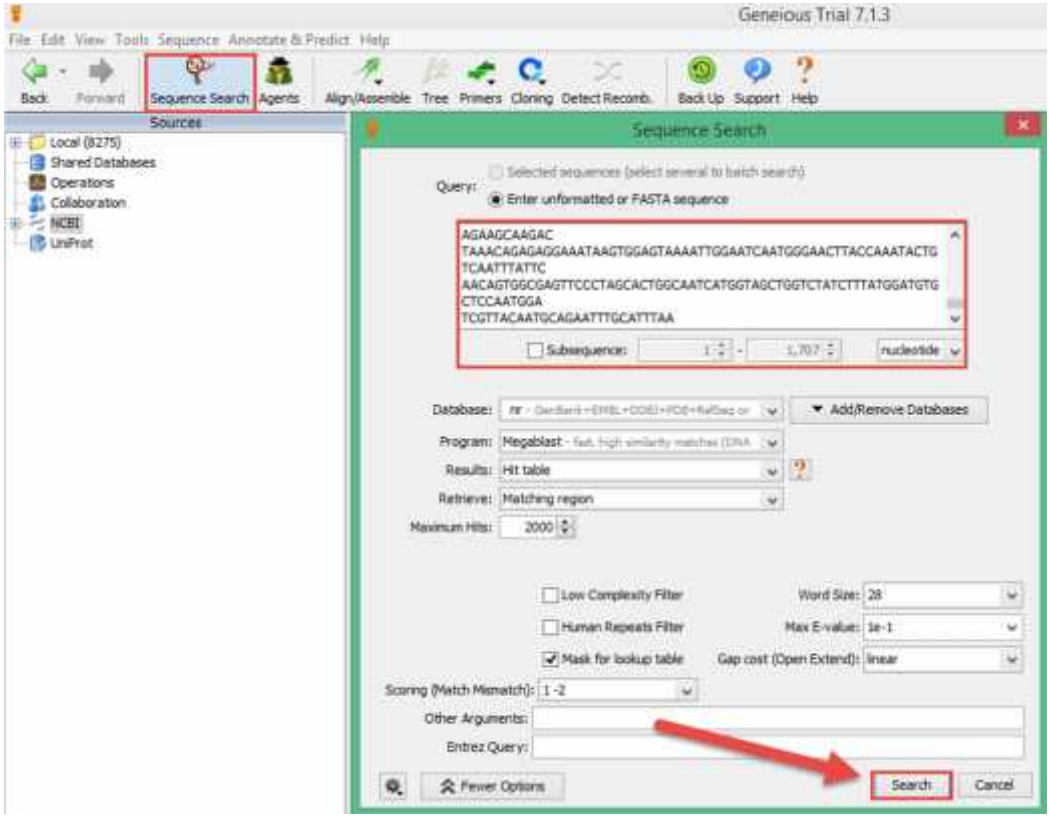


Figure: Strain search using FASTA in geneious.

After pressing search button we will get our search result in the document table.

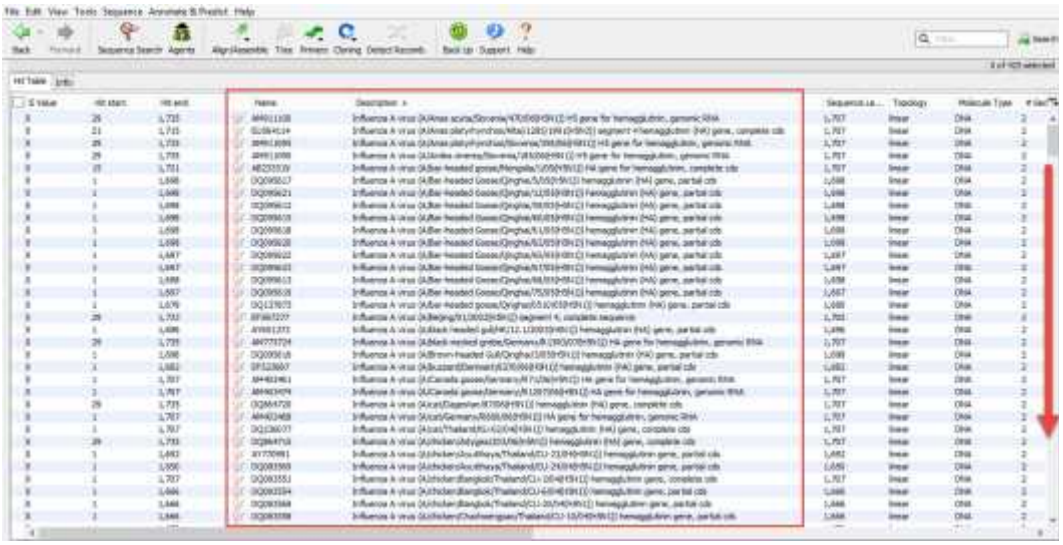


Figure: List of strains of avian Influenza in geneious.

5.15 Applications of Geneious

5.15.1 Agriculture, Plant and Crop Science

- Locating polymorphisms with multiple alignment,
- Identifying of novel bacterial strains using de-novo assembly,
- Visualizing host and virus species interaction with phylogenetic trees,
- Designing cloning PCR experiments.

5.15.2 Ecology and Conservation Genetics

- Studying biodiversity and investigate divergence,
- Identifying novel bacterial strains using de-novo assembly,
- Determining gene flow between populations,
- Designing of cloning PCR experiments,

5.15.3 Genomics

- Workflows,
- Whole Genome Alignment and Browsing,
- De Novo Assembly,
- Read Mapping,
- Variant Detection.

5.15.4 Microbiology, Virology and Infectious Diseases

- Studying diversity with 16s rRNA,
- Discovering new microbes,
- Investigating large genomic events such as horizontal gene transfer,
- Bin sequences into clusters using de-novo assembler,
- Finding novel or pathogenic mutational variants using re-sequencing,
- Finding genes in microbial DNA,
- Downloading bacterial genome annotations,
- Designing cloning PCR experiments,
- Downloading information on eukaryotic pathogens.

5.15.5 Population Genetics and Evolutionary Biology

- Easy editing of alignments,
- Studying biodiversity and investigate divergence,
- Identifying novel bacterial strains using de-novo assembly,
- Determining gene flow between populations,
- Designing cloning PCR experiments,
- Predicting protein sequences from ORFs.

5.15.6 Synthetic Biology

- Simulating Gibson Assembly, TOPO Cloning and other cloning operations in one easy step,
- Extensively testing primers at desktop before PCR,
- Built in codon optimization,
- Immediate access to Golden Gate, Biobrick and other important restriction enzymes,
- Quickly visualize, edit and share plasmids.

5.16 Limitations of Geneious

- Price of this software is high,
- Only 14 days trial.

Chapter 6

Background Study and Virology of Influenza Viruses, Reason for Neuraminidase Inhibitors Resistance of Influenza Virus A

6.1 Background

Influenza virus can infect a wide variety of birds and mammals. Natural reservoir of the virus is thought to be wild waterfowl. The other animal species (e.g. chickens, turkeys, pigs, horses) including humans are infected with influenza viruses as aberrant hosts (Suarez 2000). Of the three influenza viruses A, B and C, only certain subtypes of influenza A and the type B cause disease in humans (Zambon 1999).

Influenza virus undergoes rapid evolution by both antigenic shift and antigenic drift and this presents a significant challenge for vaccine design to best match with viruses likely to circulate in the coming influenza season (Schweiger *et al.* 2002; Stray & Pittman 2012).

The virus strains implicated in the 20th century's influenza pandemics originated directly from avian influenza viruses, either through genetic reassortment between human and avian influenza strains (1957 and 1968) or possibly through adaptation of purely avian strains to humans (1918). It was long thought that the restricted host range of avian influenza viruses precluded direct transmission to humans, and that the emergence of pandemic strains required genetic reassortment between avian and human strains. However, occurrences of direct bird-to-human transmission of avian influenza viruses have increasingly been reported in recent years, culminating in the ongoing outbreak of influenza A (H5N1) among poultry in several Asian countries with associated human infections. These unique developments have resulted in increasing global concerns about the pandemic potential of these viruses (Menno & Tran 2006).

Avian influenza A virus is playing a key role to the emergence of human influenza. Recently transmission of Avian Influenza virus from bird to human has increased in several Asian countries. Influenza A virus is a member of *Orthomyxoviridae* family is avirulent but it can be virulent by the acquisition of some genetic features which includes multibasic cleavage sites or glycosylation sites in the hemagglutinin (HA) gene can infect a wide range of species includes poultry, humans, horses, swine, quail etc. (Perdue *et al.* 1997).

6.2 Virology

Influenza is an enveloped negative-strand RNA virus with single stranded segmented RNA genome consisting of eight segments (Lamb & Choppin 1983). The eight segments of influenza A viral RNA (vRNA) encode ten recognized gene products which are PB1, PB2, PA polymerases, HA, NP, NA, Ml and M2 proteins, and NS1 and NS2 proteins (Robert *et al.* 1992).

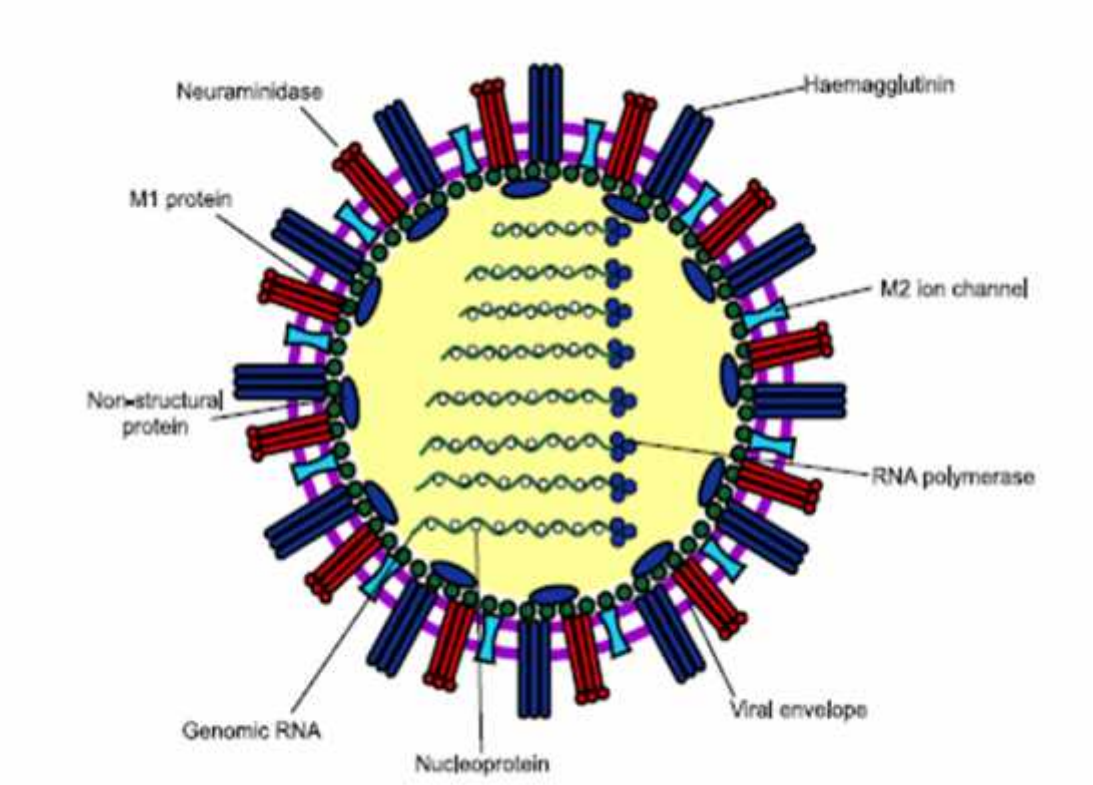


Figure: Structure of the Influenza Virus (Ludwig *et al.* 2003).

6.2.1 PB2 Polymerase

PB2 polymerase is encoded by RNA segment 1. PB2 form the trimeric vRNA dependent RNA polymerase (Engelhardt & Fodor 2006). The RNA polymerase is responsible for the transcription and replication of the viral RNA genome in the nucleus of infected cells. The PB2 protein specifically plays a role in generating 5′-capped RNA fragments from cellular pre-mRNA molecules that are used as primers for viral transcription (Engelhardt & Fodor 2006; Guilligay *et al.* 2008). Endonucleolytic cleavage of these cap structures from host mRNAs is also at least in part a function of PB2. The role of PB2 in the other virus-directed RNA synthetic processes, i.e., synthesis of full-length template cRNA and new negative-

sense viral RNA (vRNA), is not known since these processes do not require host cap priming. Newly synthesized PB2 proteins migrate to the nucleus of infected cells (Robert *et al.* 1992).

6.2.2 PB1 Polymerase

PB1 polymerase is encoded by RNA segment 2. PB1 is the best characterized functionally. Biochemical and structural analyses recognize PB1 as responsible for RNA chain elongation. PB1 contains amino acid motifs common to all RNA-dependent RNA polymerases and RNA-dependent DNA polymerases (Muller *et al.* 1994). PB1 proteins localize in the nucleus of infected cells.

6.2.3 PA polymerase

PA polymerase is encoded by RNA segment 3. The PA subunit induces a generalized proteolytic process when expressed from cloned cDNA (Sanz-Ezquerro *et al.* 1995). It also localizes in the infected cell nucleus and is a member of the RNA-dependent RNA polymerase complex along with PB1 and PB2, but its role in viral RNA synthesis is unknown. There is evidence for possible roles as a protein kinase or as a helix-unwinding protein (Robert *et al.* 1992).

Avian influenza A consists of two major glycoproteins which are Hemagglutinin (HA) and Neuraminidase (NA) (Ito *et al.* 2001). There are several subtypes of HA and NA. 18 different HA subtypes (H1 to H18) and 11 different NA subtypes (N1 to N11) are found (Tong *et al.* 2013).

6.2.4 Hemagglutinin (HA)

HA is encoded by RNA segment 4. HA glycoproteins are more prone to attach to the cell surface sialic acid receptors. HA represents a crucial component of current vaccines.

HA is a trimer-forming glycoprotein (Wilson *et al.* 1981) performing two crucial functions:

- Binding of the virus to the surface receptors of the host cell through the sialic acid and
- Releasing of the viral genome into the cytoplasm during the process of membrane fusion (Huang *et al.* 1981; Lenard & Miller 1981).

HA is synthesized in endoplasmic reticulum as a single protein. To be activated, i.e., render the virus infective and fusogenic, its cleavage is necessary by the specific host cell proteases into two chains (subunits) HA1 a HA2 (Lazarowitz *et al.* 1973). There is a difference between host surface receptors on the target cell which is believed to be the possible restrictive factor of avian influenza. HA gene of avian cell binds to Sia2-3Galactose-containing receptor which is different from human Sia2-6Galactose containing receptor (Matrosovich *et al.* 1999). Before functioning as a virus it needs post translational cleavage by host proteases (Rott 1992). HA followed by NA are important antigenic determinant from which neutralizing antibodies are directed.

6.2.5 Nucleoprotein (NP)

NP is encoded by RNA segment 5. NP binds RNA with high affinity in a sequence-independent manner (Baudin *et al.* 1994; Yamanaka *et al.* 1990), most likely through the positively charged cleft identified in the atomic structure between the head and body domains (Ye *et al.* 2006; Ng *et al.* 2008). In addition, the structure revealed the presence of a tail loop

by which NP molecules self-associate into oligomeric structures, which has been shown by mutagenesis to be required for RNP (Ribonucleoprotein) activity (Chan *et al.* 2010). NP has also been shown biochemically to interact with polymerase subunits PB1 and PB2 (Biswas *et al.* 1998), but information on the interaction domains is limited (Poole *et al.* 2004). It is generally thought that during replication the 5′ terminus of the emerging transcript is bound sequence-specifically and co-transcriptionally by “free” polymerase which then serves as a nucleation step for the sequence-independent sequential encapsidation of the transcript by NP (Ye *et al.* 2006; Tao & Ye 2010; Portela & Digard 2002).

6.2.6 Neuraminidase (NA)

NA is encoded by RNA segment 6. NA is the second major antigenic determinant for neutralizing antibodies. NA is glycosylated and possesses an amino-terminal hydrophobic sequence which functions both as signal for transport to the cell membrane and as transmembrane domain; it is not cleaved away. By catalyzing the cleavage of glycosidic linkages to sialic acid on host cell and virion surfaces, this glycoprotein prevents aggregation of virions thus facilitating the release of progeny virus from infected cells. Inhibition of this important function represents the most effective antiviral treatment strategy to date (Menno &

Tran 2006). Antiviral drugs like Oseltamivir and Zanamivir (NA inhibitors) inhibit this important function are the key to the antiviral treatment.

6.2.7 M1 Protein

M1 protein is encoded by RNA segment 7. The matrix protein Ml is the most highly conserved of all the influenza A virus proteins (McCauley & Mahy 1983) and is the smallest (252 amino acids) and most abundant of the structural proteins. The close association of MI both with the lipid envelope and with the ribonucleoprotein particles, comprising nucleoprotein (NP), polymerases and RNA, in budding and mature virions (Patterson *et al.* 1988) points to a key role in viral assembly. M1 can bind RNA directly in the absence of nucleoprotein (NP) and suggest that properties previously thought to be mediated by NP:M1 interactions may be attributable to the RNA-binding activity of M1 (Larrissa & George 1989). In the infected cell it is present in both cytoplasm and nucleus.

6.2.8 NS protein

The NS1 protein is an interferon antagonist59,60 that blocks the activation of transcription factors and IFN-b-stimulated gene products, and binds to double-stranded RNA (dsRNA) to prevent the dsRNA-dependent activation of 29-59 oligo(A) synthetase, and the subsequent activation of RNase L, an important player in the innate immune response. Recently obtained structural data are expected to help in the identification of domains that are critical for the biological functions of NS1.

The NS1 proteins of H5N1 viruses confer resistance to the antiviral effects of interferon and are associated with high levels of pro-inflammatory cytokines27,28,30,65,66; the resulting cytokine imbalance probably contributes to the high mortality of H5N1 virus infections in humans. Several amino acids in NS1 have now been shown to affect virulence
65,67,68. The S-OIVs possess the low-pathogenic-type amino acid at these positions. However, available data suggest that these amino acid changes affect virulence in a strain-specific manner, whereas a multibasic HA cleavage sequence and PB2-Lys 627 seem to be universal determinants of viral pathogenicity.

Chapter 7

Molecular Mechanisms of Avian H5N1 Influenza A Virus Replication in Mammalian Cells

7.1 Crossing the Plasma Membrane

The plasma membrane is the first cellular barrier that protects a host cell from virus infection. Influenza A viruses must recognize specific receptors (sialic acids) in order to begin the infection of a cell (Matrosovich *et al.* 2004). There are two major types of linkages between sialyloligosaccharides (SAs) and galactose (Gal): SA-α2, 6-Gal and SA-α2, 3-Gal (Matrosovich *et al.* 2004). Typically, the HA proteins of human influenza viruses preferentially bind the SA-α2, 6-Gal linkage, while avian influenza virus HA proteins have a higher affinity for SA-α2, 3-Gal (Gambaryan *et al.* 2002).

A leucine residue at position 226 of the HA receptor binding domain is necessary for recognition of the SA-α2, 6-Gal receptor, but glutamine at this site permits the recognition of SA-α2, 3-Gal (Matrosovich *et al.* 2004; Matrosovich *et al.* 2004; Rogers *et al.* 1983; Matrosovich *et al.* 2001). Fortunately, HAs from naturally isolated H5N1 influenza viruses still retain the avian-type residue that preferentially binds SA-α2, 3-Gal, suggesting a molecular explanation for lack of transmission of these viruses between humans.

However, other single amino-acid mutations, such as N154S (Chutinimitkul *et al.* 2010), N182K (Chutinimitkul *et al.* 2006), Q192R (Chutinimitkul *et al.* 2006), Q222L (Chutinimitkul *et al.* 2010), S223N (Chutinimitkul *et al.* 2010), G224S (Chutinimitkul *et al.* 2010), S227N (Gambaryan *et al.* 2006), G228S (Stevens *et al.* 2006), and combinations of mutations (Reperant *et al.* 2012), surrounding the receptor binding pocket, have been shown to increase the SA-α2, 6-Gal binding activity of recombinant H5N1 influenza viruses experimentally. It has been suggested that both respiratory and oral routes of transmission could contribute to the initial H5N1 avian-to-mammal cross-species jump (Reperant *et al.* 2012).

7.2 Nuclear Import of NP (Nucleoprotein)-Coated and Polymerase Complex (PB1, PA, and PB2)-Bound Viral RNAs (vRNPs)

There are two ways for proteins to be transported into the cell nucleus: passive diffusion (for smallb proteins of ≤60–70 kDa) and energy-dependent import (for larger proteins and complexes, such as the influenza virus vRNP) (Paine *et al.* 1975). The latter process occurs in a highly selective manner, dependent on the recognition of NLSs (nuclear localization

signals) in imported cargoes by the host factors importin-α and importin-β. Usually, importin-β requires importin-α to bind directly to the NLS of the cargo proteins, thus serving as an adaptor, before the entire importing complex is transported into the cell nucleus (Melen *et al.* 2003). In some cases, importin-β is also able to bind directly to cargoes independently of importin-α (Boivin *et al.* 2011).

As transcription and replication of the influenza viral genome occurs in the nucleus, both viral RNAs and viral proteins have to be imported into the cell nucleus. After uncoating, the incoming vRNPs are imported into the nucleus as complexes, a process likely mediated by direct binding of NP to importin-α (Wu *et al.* 2007). Subsequent newly synthesized viral polymerase and NP proteins are imported separately into the nucleus to assemble into new vRNPs during replication. NP contains two types of NLS: the unconventional NLS1 (aa1-13) and the bipartite NLS2 (aa198-216) (Ozawa *et al.* 2007; Wu *et al.* 2007). The nuclear import of NP is mainly mediated by binding of importin-α to NP NLS1, which is more exposed than NLS2 (Wu *et al.* 2007; Wu *et al.* 2009). The nuclear translocation of PB2 also depends on an interaction with importin-α (Mukaigawa *et al.* 1991; Tarendeau *et al.* 2007). In contrast, the nuclear import of monomeric PB1 and PB1-PA dimers relies on their interaction with the importin-β homologue RanBP5, which is independent of importin-α binding (Deng *et al.* 2006). Notably, interactions between viral proteins are also important for the nuclear accumulation of some viral proteins. For example, an NLS in the N-terminal domain of PA is important for the nuclear accumulation of PB1 (Fodor *et al.* 2004).

7.3 Replication in the Nucleus

After entry into the nucleus, transcription and replication of the influenza virus genome depends on viral polymerase activity, together with cellular co-factors. Since the outbreak of highly-pathogenic H5N1 avian influenza virus in 1997, avian-type viruses containing the PB2 E627K mutation have been increasingly isolated from human H5N1 infection cases, and this mutation has been identified as a mammalian adaptive marker [50]. The underlying mechanism of mammalian adaptation by PB2 E627K is still unclear, although multiple possibilities have been suggested. It has been hypothesized that ‘mammalian-type’ PB2 627K is beneficial for the replication of avian influenza viruses in the human upper respiratory tract, which generally has a temperature of only 33 °C. In contrast, the avian intestinal tract is

closer to 41 °C, a temperature at which the 'avian-like' PB2 627E polymorphism *Viruses* 2013, *5* 1436 facilitates efficient virus replication (Steel *et al.* 2009; Hatta *et al.* 2007). Additionally, NP appears to bind more efficiently to PB2 627K than 627E in mammalian cells, but not in avian cells (Labadie *et al.* 2007), which is likely responsible for enhancing polymerase activity.

The E627K substitution is not the only mammalian-adaptive mutation that can arise in the viral polymerase. Indeed, a large proportion of H5N1 strains isolated from humans possess PB2 627E, suggesting that other adaptive amino-acid substitutions must exist in order to counterbalance its generally low polymerase activity in human cells. It is reported that PB2 591K/R and PB2 701N can compensate for lack of PB2 627K, thereby ensuring efficient polymerase activity in mammalian cells (Gabriel *et al.* 2005; Yamada *et al.* 2010). There is also evidence that the PB1 gene from an avian source promotes avian polymerase activity in mammalian cells (Naffakh *et al.* 2000). Two mutations in PB1 (473V and 598P) were identified as increasing the polymerase activity of viruses carrying PB2 627E in mammalian cells (Xu *et al.* 2012). More recently, it was found that mutations in NEP are also involved in host adaptation. The adaptive mutation M16I (and others) in the NEP proteins of certain human H5N1 isolates can increase the relatively low polymerase activity of avian viruses in mammalian cells (Manz *et al.* 2012). These NEP mutations are more common in human H5N1 isolates carrying the PB2 627E mutation than in human H5N1 viruses possessing PB2 627K, which suggests that NEP can act as an important determinant of host adaptation by promoting efficient polymerase activity in human cells.

7.4 Escaping the Host Immune Response

During infection, the host innate immune response is elicited to counteract virus replication. Influenza A viruses limit this cellular defense mechanism in infected cells by producing the NS1 protein from genomic segment 8. NS1 is highly multifunctional, but is well-characterized as an antagonist of the host antiviral response, especially by repressing the production of interferons (IFNs) (Hale *et al.* 2008). NS1 acts in several ways to limit IFN production during infection. Firstly, NS1 can disrupt the cytoplasmic RIG-I-mediated signaling pathway that would otherwise lead to primary transcription of IFN mRNA. Simultaneously, NS1 protein molecules in the nucleus can inhibit the post-transcriptional

maturation of IFN mRNA by interacting with the host-cell pre-mRNA cleavage and polyadenylation specificity factor, CPSF30 (Hale *et al.* 2008).

The interaction of NS1 with CPSF30 is also strain-dependent. With respect to avian H5N1 viruses that have sporadically entered the human population, the NS1 proteins from early viruses isolated *Viruses* **2013**, *5* **1438** around 1997 possessed an intrinsic defect in binding human CPSF30. However, since 1998, mutations selected for at NS1 residues 103 and 106 have lead to H5N1 viruses acquiring the ability to bind human CPSF30 and, thereby, better suppress the human IFN response (Twu *et al.* 2007). The exact selection pressure causing this phenomenon is unclear, but acquisition of such a function can increase the virulence of avian H5N1 viruses in mammalian hosts (Spesock *et al.* 2011).

7.5 Releases from the Cell Surface

The viral NA protein removes sialic acid from the surface of host cells, from which virus particles are budding. This minimizes any HA-mediated linkage between the virion and host cell receptors, thus enabling efficient virus release from the cell surface. NA also removes sialic acid from glycoproteins embedded within the viral membrane in order to prevent HA-mediated aggregation of virions. The NA proteins from human and avian influenza viruses can have different sialic-acid substrate specificities, which often correlate with the receptor binding specificities and affinities of their associated HAs (Wagner *et al.* 2002). The NAs from avian viruses have relatively higher enzymatic activity to cleave SA-α2, 3-Gal linkages, whereas human influenza virus NAs can more efficiently cut SA-α2, 6-Gal linkages. Notably, the avian-derived NA from human H2N2 virus, in which HA recognized SA-α2, 6-Gal, gradually drifted from hydrolyzing SA-α2, 3-Gal linkages to SA-α2, 6-Gal linkages during the period it was circulating in humans after 1957 (Kobasa *et al.* 1999; Baum *et al.* 1991). Indeed, an alignment of HA and NA specificities exists in influenza viruses that are well adapted to a particular host (Wagner *et al.* 2002).

Chapter 8

Avian Influenza (Bird Flu) in Humans

8.1 Mechanism of transmission of influenza viruses to humans

Influenza viruses are dynamic and are continuously evolving. They can change in two different ways, antigenic drift and antigenic shift. These viruses change by antigenic drift all the time, but antigenic shift happens only occasionally. Antigenic drift occurs through point mutations in two genes that contain the genetic material to produce the main surface proteins, hemagglutinin and neuraminidase. By this way antigenic drift produces new virus strains that may not be recognized by human antibodies to earlier influenza strains and infection with a new strain can occur.

Antigenic shift produces a novel influenza A virus subtypes in humans that was not currently circulation among people. Antigenic shift can occur either directly through animal-to-human transmission or through mixing of human influenza A and animal influenza A virus genes to create a new human influenza A subtype virus through genetic reassortment process. Thus influenza A viruses may responsible for a pandemics of influenza in humans (Gamblin *et al*. 2004; Taubenberger *et al*. 2005).

8.2 Threat of H5N1 in humans

Avian influenza viruses are normally highly species specific so that after infection to an individual species i.e. humans, birds, pigs, horses, seals; they stay to that species only and rarely spill over to cause infection in other species. Although avian influenza viruses usually do not infect humans, rare cases of human infection with these viruses have been reported since 1997 (Belshe *et al*. 2005) of hundreds of strains of avian influenza A viruses, only four are known to caused human infections: H5N1, H7N3, H7N7 and H9N2. Out of these four strains, H5N1 has caused the largest number of detected cases of severe diseases and death in humans (Wong *et al*. 2006; Shaw *et al*. 2002). H5N1 type influenza A virus is of greatest concern for human infection because of following reasons (Claas *et al*. 1998).

1. It can cause severe diseases in human.
2. The birds that survive infection excrete virus for at least 10 days, orally and through faces, helping spread of the virus at poultry markets and also by migratory birds.
3. It mutates rapidly and seems to acquire genes from virus infecting other animal species.
4. The more humans get infection, the people can become infected both human and bird flu strains.

Of the human cases associated with the ongoing H5N1 outbreaks in poultry and wild birds in Asia and parts of Europe, the near east and Africa, more than half of those people reported infected with the virus have died. Total 309 cases of avian influenza infection in humans have been reported between ends of December 2003 to end of May 2007; out of which 187 deaths were confirmed. But it is a matter of great assurance that, H5N1 avian influenza virus does not easily cross from birds to infect humans (World *et al*. 2007).

8.3 Mode of transmission

Investigations of all the most recently confirmed cases, in China, Turkey and Indonesia have identified that the most likely source of exposure is close contact with infected or dead poultry or their faeces. According to World Health Organization (WHO), almost all human sufferers caught the disease directly from birds and only tiny number of infected person has infected another. Other sources are flesh (even refrigerated/frozen carcasses), eggs from infected birds which can harbor the virus both outside and within the cell. Indirect transmission is also possible through contaminated clothing, footwear, vehicles and equipments. Contaminated feed, water, insects, rodents, cats, dogs can also act as vectors and transmit the disease (Bridges *et al*. 2002; Mounts *et al*. 1999).

8.4 Route of transmission

Route of transmission is mainly respiratory tract. Conjunctiva may be affected directly. The incubation period for H5N1 avian influenza ranges from 2 to 8 days and possibly as long as 17 days. WHO recommends that an incubation period of 7 days be used for field investigations and monitoring of patients' contacts (Bridges *et al*. 2003; Salgado *et al*. 2002; and van *et al*. 2006).

8.5 Antigenic Drift (Gradual, Small Changes)

Antigenic drift refers to small, gradual changes that occur through point mutations in the two genes that contain the genetic material to produce the main surface proteins, hemagglutinin (HA) and neuraminidase (NA). These point mutations occur unpredictably and result in minor changes to these surface proteins. Antigenic drift produces new virus strains that may not be recognized by antibodies to earlier influenza strains. This process works as follows: a person infected with a particular influenza virus strain develops antibody against that strain. As newer virus strains appear, the antibodies against the older strains might not recognize the "newer" virus and infection with a new strain can occur. (Centers for disease control and prevention, Department for disease control and prevention Avian Influenza (Bird Flu) November 18, 2005).

8.6 Antigenic Shift (Major, abrupt change)

The event of their replacement on a global scale is commonly referred to as an antigenic shift in the virus population (Paul *et al.* 2013). It refers to an abrupt, major change to produce a novel influenza A virus subtype in humans that was not currently circulating among people. Antigenic shift can occur either through direct animal (poultry) to human transmission or through mixing of human influenza A and animal influenza A virus genes to create a new human influenza A subtype virus through a process called genetic reassortment. Antigenic shift results in a new human influenza A subtype. A global influenza pandemic (worldwide spread) may occur if three conditions are met: (Centers for disease control and prevention, Department for disease control and prevention Avian Influenza (Bird Flu). November 18, 2005)

- A new subtype of influenza A virus is introduced into the human population.
- The virus causes serious illness in humans.
- The virus can spread easily from person to person in a sustained manner.

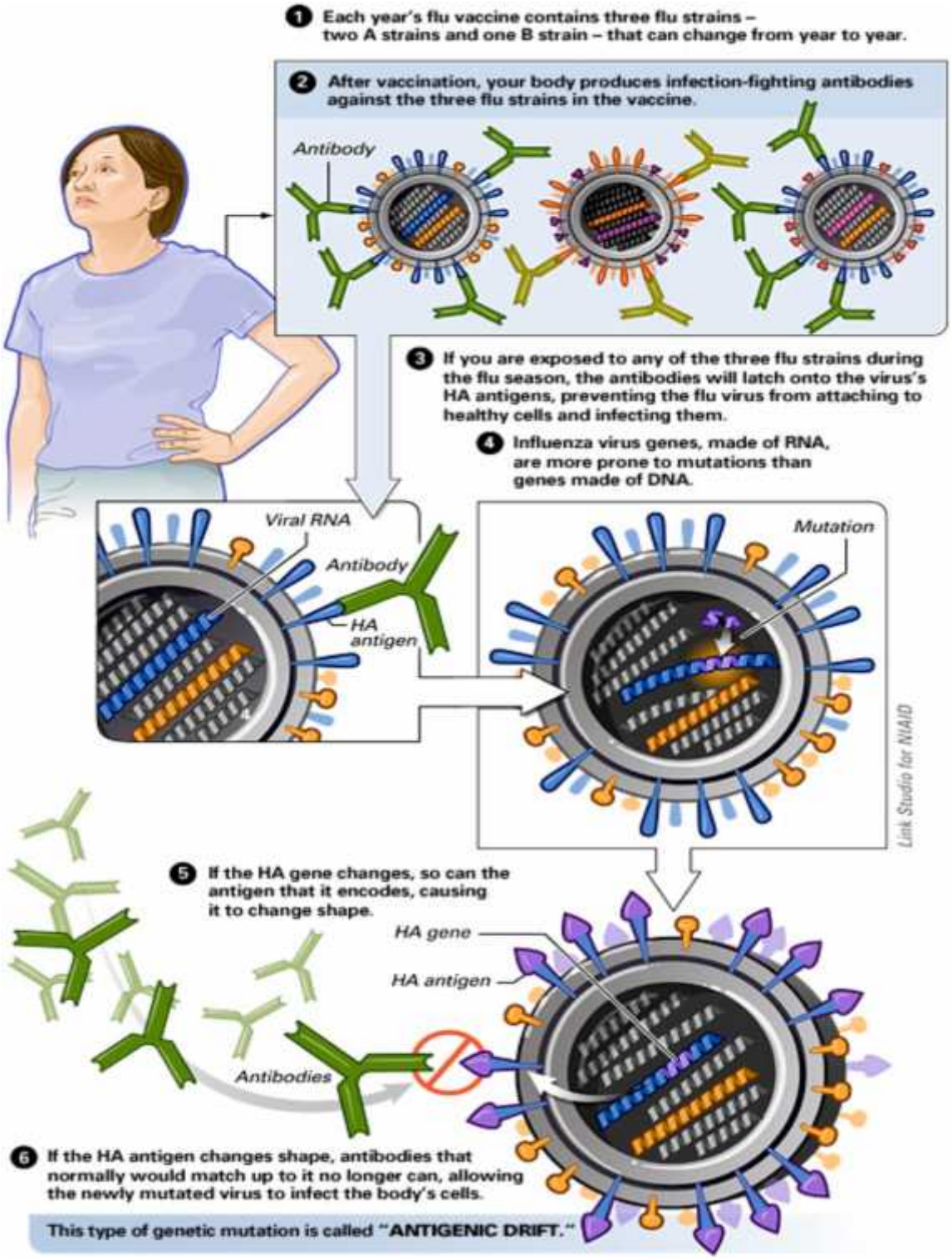


Figure: Antigenic drift © National Institute of Allergy and Infectious Diseases (NIAID).

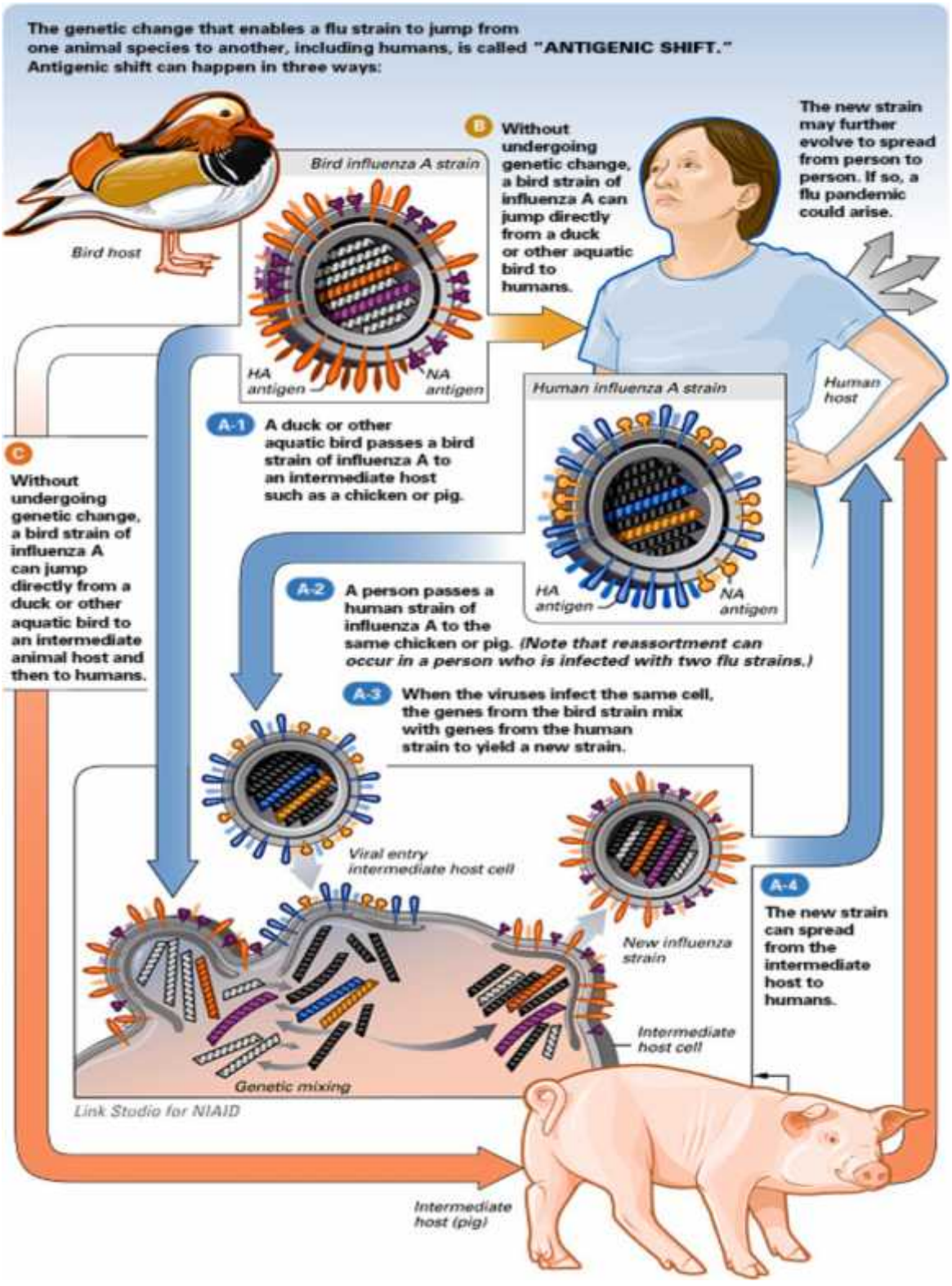


Figure: Antigenic shift © National Institute of Allergy and Infectious Diseases (NIAID).

Chapter 9

Influenza Neuraminidase Inhibitors; Antiviral Action and Mechanisms of Resistance

9.1 Major classes of anti-influenza drugs

The anti-influenza drugs are usually classified according to their target in the viral life-cycle, which is schematically depicted in Figure. Such antiviral molecules are particularly used as inhibitors of the following processes: attachment of the virus to host cell receptors, endocytosis and fusion of viral and cell membranes, replication and transcription of the viral genome, synthesis of viral proteins, assembly of the viral progeny and release of the new virions into the outside environment. The following paragraphs are focused on the description of basic classes of influenza virus inhibitors.

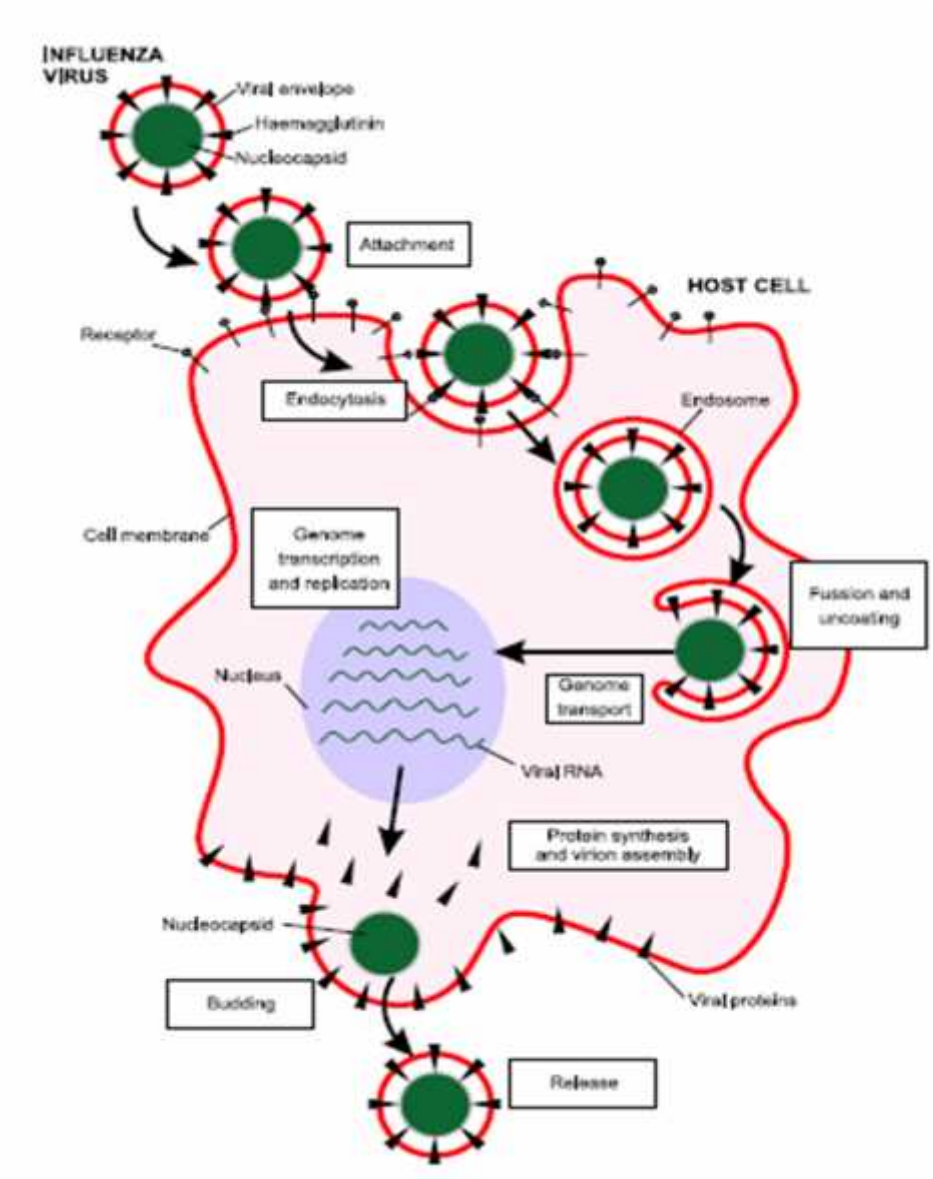


Figure: Life-cycle of influenza virus (Beigel and Bray 2008).

9.1.1 Inhibitors of haemagglutinin

Haemagglutinin is a trimeric rod-shaped glycol-protein located on the surface of influenza virions. During the initial step of infection, haemagglutinin molecules bind specifically to host cell sialic acid receptors and enable entry of the virus into the cell cytoplasm by fusion of viral and cell membranes (Skehel and Wiley 2000). The binding interaction of haemagglutinin with cellular receptors can be efficiently inhibited by synthetic macromolecules composed of multiple sialic acid residues conjugated to glycan, glycopeptide or polyacrylamide backbones (Sigal *et al*. 1996; Feng *et al*. 2010; Narla and Sun 2012). Such therapeutic constructs have been observed to exhibit higher antiviral effects than monovalent sialic acid analogues. Haemagglutinin mediated membrane fusion has been successfully blocked by a large variety of small organic compound, e,g. quinones (Larson *et al*. 2012), the antibiotic staschyflin produced by *stachybotrys* sp. (Yoshimoto *et al*. 1999) and derivatives of benzamide and podocarpic acid (Luo *et al*. 1997; Staschke *et al*. 1998). Although these structures show significance anti-influenza potency in vitro, most of them are characterized by strong strain specificity and apparent cytotoxicity.

The sialidase fusion protein DAS181 is an entirely novel broad-sprectrum haemagglutinin inhbitors that encymaticallky removes sialic acid receptors from respirator epithelium cells, preventiong virus attachment (Malakhov *et al*. 2006). This new antiviral is ctive against both A and B influenza atrains at nanomolar concentrations, causing minimal ccytolpathic effects. This substance is at prewnt being tested in phase II human trials (Moss *et al*. 2012). Recently, research has also focused on fusion- haemagglutinin RNA apatamers (Lee *et al*. 2011).

9.1.2 M2 ion channel blockers

M2 ion channel is a transmembrane viral protein that mediadtes the selective transport of proteins into the interior of the influenza virion. Conductance of the viral particle and facilitates the haemagglutinin-mediated membrane fusion which in turn results in the uncoation of the influenz nucleocapsid and tmport of the viral genome into the nucleus (Schnell and Chou 2008). Amantanes are protein m2 channel blockers, which are known as the first synthetic anti-influenza drugs described in the mid-1960s (Davies *et al*. 1964). Two adamantine derivatives, amantadine and rimantadine (Figure), have been licensed for influenza control and are commercially available under the trademarks Symmetrel® and Flumadine®, respectively (Galvao *et al*. 2012). The adamantanes are relatively cheap, highly stable in storage and show strong anti-influenza activity at micromolar concentrations. At

present, the application of adamantanes for prevention and treatment of influenza infections is, however, not recommended because of the rapid emergence of drug-resistant virus variants that retain full virulence and transmissibility (Bright *et al*. 2005; Barr *et al*. 2008). Moreover, serious gastrointestinal and neurological side effects were observed in patients undergoing adamantine therapy (Galvao *et al*. 2012). Another disadvantage of adamantanes is their strong specificity against influenza A strains only (Rosenberg and Casarotto 2010).

9.1.3 Inhibitors of viral RNA polymerase

Transcription and replication of the influenza virus genome is carried out by the influenza RNA polymerase holoenzyme, which is characterized by two catalytic activities. Polymerase activity is needed for the elongation of nascent RNA chains, whereas endonuclease activity is essential for cleavage of the 5'-capped primer sequence of the host mRNA. The cap is the terminal 7-methylguanosin bound through a triphosphate group to the host mRNA. This "cap snatching" process is needed for the initiation of viral RNA transcription (Lv *et al*. 2011). Influenza RNA polymerase is an extremely suitable target for the development of new broad-specific antivirals because of its highly conserved structure among influenza strains. It is thought that the influenza polymerase plays a crucial role in virus adaptation to humanto-human transmission and, consequently, in the formation of pandemic influenza variants (Miotto *et al*. 2008; Boivin *et al*. 2010; Aggarwal *et al*. 2011; Ping *et al*. 2011).

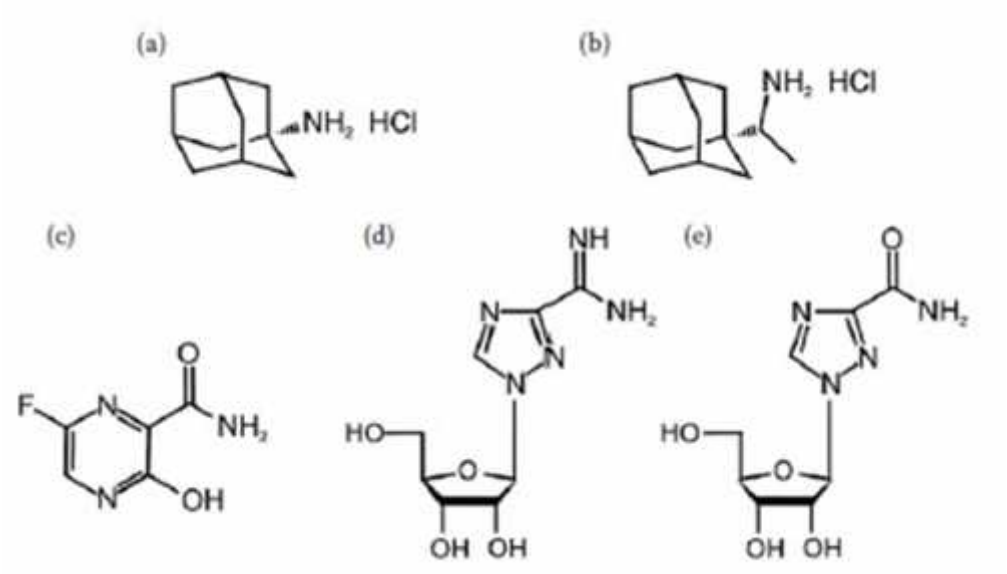


Figure: Chemical structures of the most important anti-influenza drugs;
(a) amantadine, (b) rimantadine, (c) favipiravir (T-705), (d) ribavirin, (e) viramidine.

Two basic classes of RNA polymerase inhibitors have been described based on different mechanisms of action. The first group is represented by nucleoside analogues for the blocking of viral RNA chain elongation (Tisdale *et al*. 1995). A typical member of this group is favipiravir (T-705, Figure), which is an inhibitor of influenza A, B and C strains, including variants resistant to amantadine or oseltamivir. This compound is currently in the stage of clinical testing (Furuta *et al*. 2005). Other nucleoside analogs with antiinfluenza activity include ribavirin (Virazole®) and its derivative viramidine, originaly licensed for treatment of hepatitis C infections (Figure). Their application is, however, sometimes connected with the development of haemolytic anaemia (Sidwell *et al*. 2005). The second class of antiviral molecules targeting the influenza polymerase is represented by compounds which block the endonuclease and cap-binding domains of the polymerase holoenzyme. These antivirals include cap analogues (Lv et al. 2011), short capped oligonucleotides (Tado *et al*. 2001), and small organic compounds, such as 4-substitued 2,4-phenylbutanoic acid (Hastings *et al*. 1996) and flutimide isolated from the fungus *Delitschia confertaspora* (Tomassini *et al*. 1996).

9.1.4 Inhibitors of neuraminidase

Neuraminidase, also referred to as sialidase, is an antigenic glycoprotein anchored in the surface envelope of the influenza virions, which hydrolytically cleaves the terminal sialic acid from the host cell receptors (Figure). Thus, it plays a crucial role in the release of viral progeny from the membranes of infected cells, prevents self-aggregation of virions and facilitates the movement of the infectious viral particles in the mucus of the respiratory epithelia (Matrosovich *et al*. 2004; Suzuki *et al*. 2005). Influenza neuraminidase has been established as a key drug target for the prophylaxis and treatment of influenza infections, predominantly for the following reasons: Firstly, the structure of the influenza neuraminidase active site is highly conserved between influenza A and B strains, making neuraminidase an attractive target for the development of broad-spectrum inhibitors (Yen *et al*. 2006). Secondly, resistance to neuraminidase inhibitors develops less commonly than to other anti-influenza drugs. Nevertheless, the intensive application of neuraminidase inhibitors for influenza treatment results in a permanently increasing number of drug-resistant strains (Garcia *et al*. 2009). Thirdly, in contrast to adamantanes, neuraminidase inhibitors are mostly well tolerated in patients under therapy (Cao *et al*. 2012). Finally, neuraminidase protein is a freely accessible target for antiviral molecules with an extracellular mode of action.

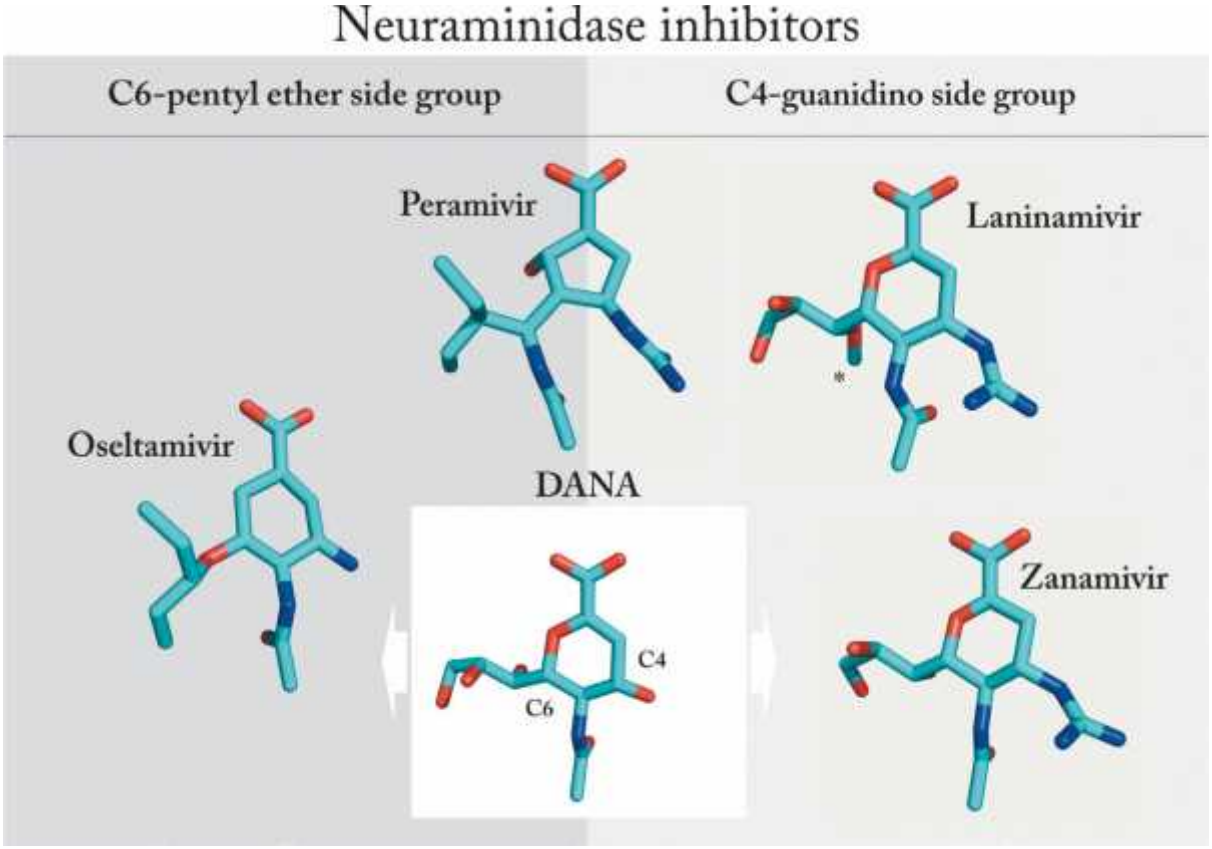


Figure: Chemical structure of DANA with neuraminidase inhibitors **(**Collins *et al.* 2008; McKimm 2013).

The development of neuraminidase inhibitors started in the middle 1970s, when the first structural analogues of sialic acid were described and denoted as DANA (2-deoxy-2,3-didehydro- N-acetyl neuraminic acid) and its trifluoroacetyl derivative FANA (Schulman and Palese 1975). At present, several licensed anti-influenza medications are available on the market, most notably the inhalant zanamivir with the trademark Releza®, and the orally administered oseltamivir (Tamiflu®) having excellent bioavailability and relatively long half-life *in vivo* (He *et al.* 1999; Greengard *et al.* 2000) (Figure). In response to the emergence of some oseltamivir-resistant influenza strains, peramivir and laninamivir have been recently developed (Bantia *et al.* 2006; Kubo *et al.* 2010). New-generation neuraminidase inhibitors are currently under investigation, e.g., multimeric forms of zanamivir (Watson *et al.* 2004), dual-targeted bifunctional antivirals (Liu *et al.* 2012), and several herbal remedies, such as flavonols, alkaloids and saponins (Jeong *et al.* 2009).

9.2 Development of neuraminidase inhibitor molecules

The neuraminidase inhibitors zanamivir and oseltamivir interfere with the release of progeny influenza virus from infected host cells, a process that prevents infection of new host cells and thereby halts the spread of infection in the respiratory tract (Fig.). Since replication of influenza virus in the respiratory tract reaches its peak between 24 and 72 hours after the onset of the illness, drugs such as the neuraminidase inhibitors that act at the stage of viral replication must be administered as early as possible. In contrast to the adamantanes, the neuraminidase inhibitors are associated with very little toxicity and are far less likely to promote the development of drug-resistant influenza. As a class, the neuraminidase inhibitors are effective against all neuraminidase subtypes and, therefore, against all strains of influenza, a key point in epidemic and pandemic preparedness and an important advantage over the adamantanes, which are effective only against sensitive strains of influenza A. These new drugs, if used properly, have great potential for diminishing the effects of influenza infection.

All influenza viruses bear two surface glycoproteins, a hemagglutinin and a neuraminidase, which are the antigens that define the particular strain of influenza. The variation of these molecules over time permits the virus to evade human immune responses and therefore necessitates the formulation of a new vaccine each year. The hemagglutinin is a sialic acid receptor–binding molecule and mediates entry of the virus into the target cell. The neuraminidase — the target molecule of the neuraminidase inhibitor compounds — cleaves the cellular-receptor sialic acid residues to which the newly formed particles are attached (Fig.). This cleavage releases the viruses, which can now invade new cells. Without neuraminidase, infection would be limited to one round of replication, rarely enough to cause disease. Neuraminidase may also facilitate viral invasion of the upper airways, possibly by cleaving the sialic acid moieties on the mucin that bathes the airway epithelial cells (Matrosovich *et al*. 2004).

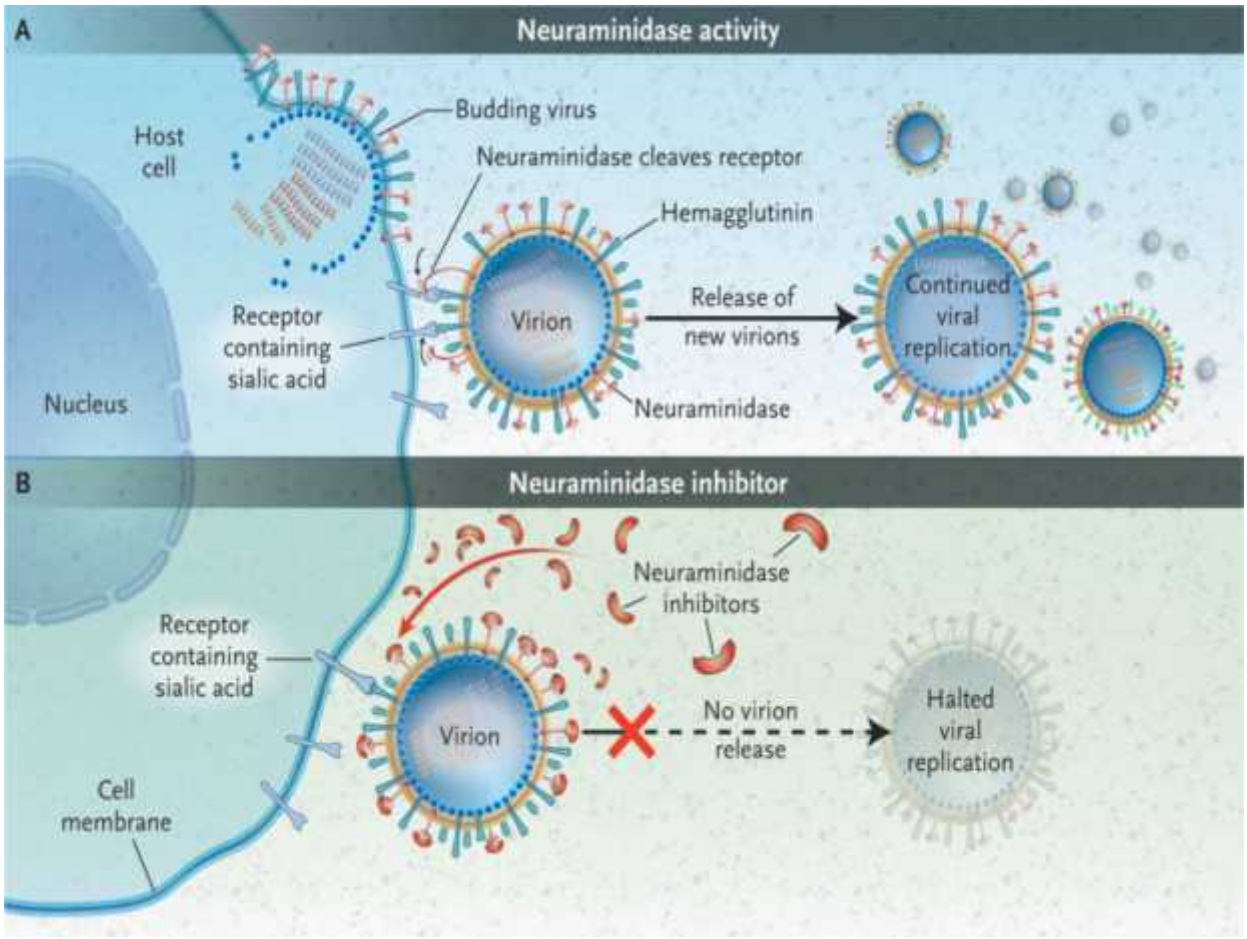


Figure: Mechanism of Action of Neuraminidase Inhibitors.

Panel A shows the action of neuraminidase in the continued replication of virions in influenza infection. The replication is blocked by neuraminidase inhibitors (Panel B), which prevent virions from being released from the surface of infected cells.

The ability of transition-state analogues of sialic acid to inhibit the influenza neuraminidase was first recognized in the 1970s, 3-5 but the design of highly effective inhibitors became feasible when analysis of the three-dimensional structure of influenza neuraminidase 6 disclosed the location and structure of the catalytic site. Potent inhibitors such as zanamivir closely mimic the natural substrate, fitting into the active site pocket and engaging the protein in the most energetically favorable interaction. 7-9 Zanamivir is administered by oral inhalation, which delivers the drug directly to the respiratory tract. Oseltamivir was developed through modifications to the sialic acid analogue framework (including the addition of a lipophilic side chain) that allow the drug to be used orally (Kim *et al*. 1997).

9.3 Resistance to the neuraminidase inhibitors

In the first few years after the approval of the NAIs it was generally believed that selection of antiviral resistance would go hand in hand with loss of virus pathogenicity and transmission (Carr *et al*. 2002; Ives *et al*. 2002). Therefore, antiviral resistance was thought not to be of major clinical significance. Indeed, viruses that were isolated early after the introduction of the NAIs were indeed attenuated (Carr *et al*. 2002; Ives *et al*. 2002). Unexpectedly, however, a dramatic increase of OS-resistant viruses was seen during the influenza season of 2007/2008 (Hauge *et al*. 2009). This virus did not seem to be attenuated in its ability to spread (Baz *et al*. 2010), in fact it had completely replaced the OS-sensitive variant by the end of 2008 (Besselaar *et al*. 2008; Meijer *et al*. 2009). In the discussion of this thesis a hypothesis is proposed, which may explain its emergence and spread.

9.3.1 Mechanisms of resistance to neuraminidase inhibitors

The active site of the NA is highly conserved among different influenza viruses (Russell *et al*. 2006). However, the combination of small differences in the active sites of N1, N2 and B NAs and those between the NAIs (Figure), lead to different options for development of resistance. Most amino acid changes causing NAI resistance are located in or are in close proximity to the active site (Van *et al*. 2011). All amino acids described in the introduction are numbered according to those in N2 NA.

9.3.1.1 N1 neuraminidase

In type 1 NA, the major amino acid change causing antiviral resistance is a histidine to tyrosine change at position 274 (H274Y). This change causes an increase of the 50% inhibitor constant (IC50) of about 150-fold for OSC and 80-fold for PER (Samson *et al*. 2013), but no increase in IC50 for ZA. This mutation is the major OS-resistance change in influenza A/H1N1 and A/H5N1 viruses (Le *et al*. 2005; Beigel *et al*. 2005; and De *et al*. 2005).

For binding of the pentoxyl group of OSC into the hydrophobic pocket reorientation is required of a glutamic acid residue at position 276 (E276) towards H274. In case of the H274Y change this reorientation is blocked by the larger tyrosine residue (Collins *et al*. 2008; McKimm *et al*. 2013). Since the C6-group of ZA is different it does not involve reorientation

of E276 upon binding (Collins *et al*. 2008; Russell *et al*. 2006 and Van *et al*. 2012). The I222R and S246N changes alter the hydrophobic pocket also and make binding of OSC to a H274Y mutant even more difficult (Collins *et al*. 2008; McKimm *et al*. 2013). A glutamine to lysine change at position 136 (Q136K) was previously shown to cause ZA (300-fold) and peramivir (70-fold) resistance in H1N1 viruses (Hurt *et al*. 2009). More recently it was shown that the presence of the Q136K mutation seemed to be an artefact of virus propagation in Madin Darby Canine Kidney (MDCK) cell cultures (Kaminski *et al*. 2012).

9.3.2 Mutations conferring drug resistance

Genotypic analysis of resistance to M2 blockers is relatively straightforward as only a few substitutions occurring at five codons within the M2 gene (codons 26, 27, 30, 31 and 34) have been linked to amantadine/rimantadine resistance **(**Abed *et al.* 2005**)**. Importantly, these M2 mutants are fully virulent and transmissible between humans **(**Hayden *et al.* 1989**)**. The NA mutations conferring resistance to NAIs vary according to the viral subtype/ type and the NAI20 (Table). In the N1 subtype, the most frequently encountered mutation is the H275Y that confers highly reduced inhibition to oseltamivir, moderate crossresistance to the investigational agent peramivir and susceptibility to zanamivir **(**Pizzorno *et al.* 2011**)**. Various amino acid changes at residue 223 (I? R/V) can also confer reduced inhibition to oseltamivir and/or to zanamivir (Pizzorno *et al.* 2012 and van *et al.* 2010**)**. In the N2 subtype, the most frequent mutations conferring highly reduced inhibition to oseltamivir are E119V and R292K, the latter being also associated with reduced inhibition to zanamivir (Whitley *et al.* 2001 and Baz *et al.* 2006**)**.

Reported changes associated with NAI resistance in B viruses include mainly R150K and D197N (Mishin *et al.* 2005 and Gubareva *et al.* 1998**)**. Of note, some NA mutations reported at codons 136 and 151 have a questionable clinical impact because they have been almost exclusively detected after cell passages and rarely in primary clinical samples **(**Hurt *et al.* 2009; Dapat *et al.* 2010 and Sheu *et al.* 2008**)**. Finally, very few zanamivir-resistant influenza viruses have been reported in clinical samples so far which may be explained by the modest use of this inhaled antiviral and also possibly by a higher genetic barrier for resistance due to a greater structural homology to the natural substrate, sialic acid **(**McKimm *et al.* 2012**)**.

Table: N1 neuraminidase mutations causing reduced sensitivity to NAIs

Influenza (sub)type	NA mutation a	Virus source/ NAI used for selection	Phenotype in NA inhibition assays b			References
			OSC	ZA	PER	
A(H1N1)	Q136K	In vitro (clinic?)/none	S	HRI	ND	(Kaminski *et al.* 2012; Hurt *et al.* 2009)
	H274Y	Clinic/OSC	HRI	S	HRI	(Van *et al*. 2008)
2009 A(H1N1)	E119V	Reverse genetics	RI	HRI	RI	(Pizzorno *et al*. 2011)
	E119G	Reverse genetics	S	HRI	RI	(Pizzorno *et al*. 2011)
	I222R	Clinic/OSC	RI	RI	ND	(Van *et al*. 2010)
	I222R/H274Y	Clinic/OSC	HRI	RI	HRI	(Nguyen *et al*. 2009)
	I222V/H274Y	Clinic/OSC	HRI	S	ND	(Centers *et al*. 2009)
	S246N/H274Y	Clinic/OSC	HRI	S	HRI	(Hurt *et al*. 2011)
	H274Y	Clinic/OSC	HRI	S	HRI	(Baz *et al.* 2009; Hurt *et al.* 2012)
	N294S	Reverse genetics	HRI	S	RI	(Pizzorno *et al.* 2011)
A(H5N1)	E119G	In vitro/ZA	S	HRI	RI/HRI	(Hurt *et al.* 2009)
	D198G	In vitro/ZA	RI	RI	S	(Hurt *et al.* 2009)
	H274Y	Clinic/OSC	HRI	S	HRI	(Le *et al.* 2009)
	N294S	Clinic/OSC	RI	S	S	(Le *et al.* 2009)

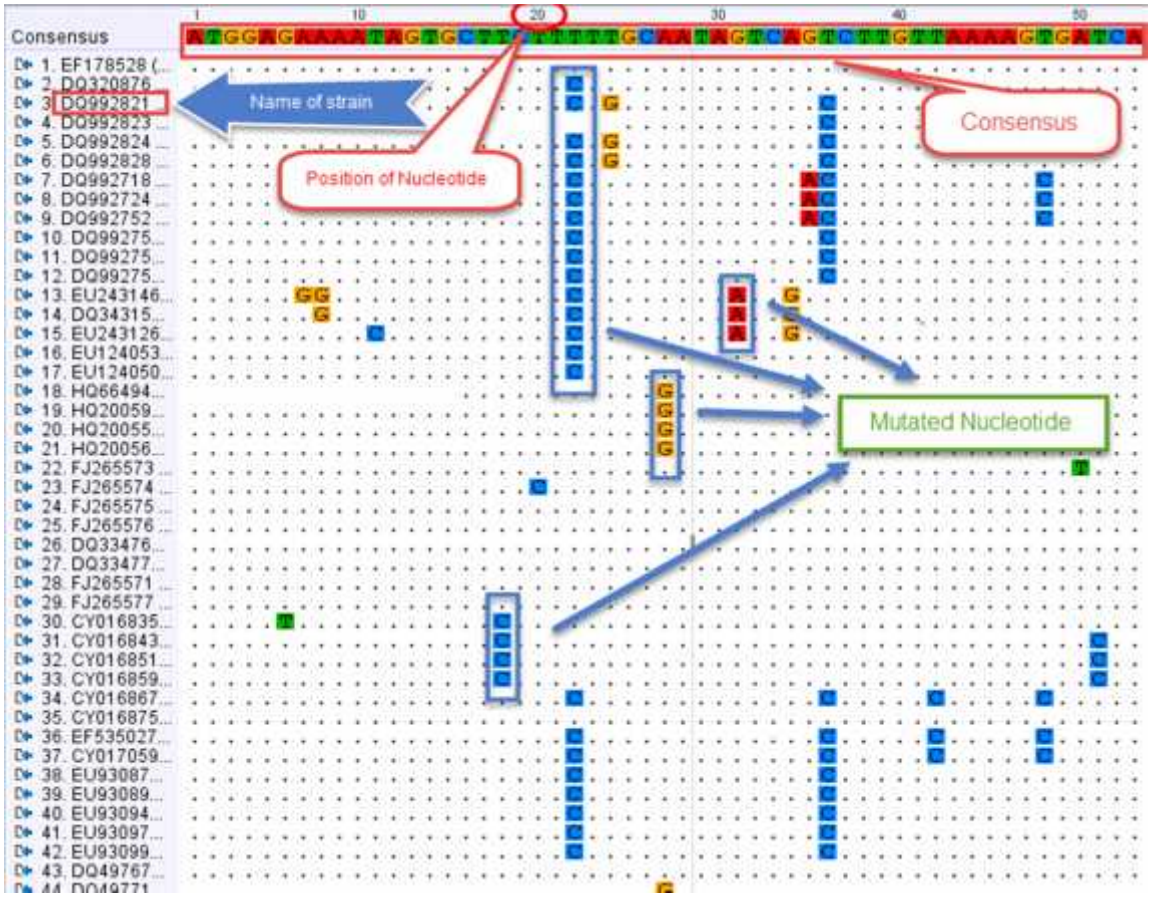
aNumbers indicate the position of the substituted residue in the NA amino acid sequence (N2 numbering). bS:<10-fold increase in IC50 over wild type); RI: 10–100-fold increase in IC50 over wild type); HRI: >100- fold increase in IC50 over wild type. ND: Not determined. A(H1N1): seasonal H1N1 viruses, A(H1N1) pdm09: Swine origin H1N1 viruses responsible for the 2009 pandemic. Table adapted from (Samson *et al.* 2013).

Chapter 10

Data Analysis & Result

10.1 Data Analysis

We have collected random strains of different year (2005, 2006 and 2007) from the search result. We will analyze a lot of strain from different year. We will make nucleotide sequence, amino acid sequence and tree from those strains. Phylogenetic tree helps to find out the diverse strain hence it helps to find out the evolution. Phylogenetic tree cannot be made without nucleotide or amino acid sequences. Diverse strain shows mutation in different nucleotide and amino acid position. From literature search we will try to find the virology of these diverse strains. In case of unavailability to find out the virology of the identified diverse strain we will find out the reported mutant position of different strain. These mutant positions can infect more avian species, reduced sensitivity to anti-viral drugs and even they may infect to human also. We will also correlate the reported data with our experimental data.



Here is one example of nucleotide sequence.

Figure: Overview of Nucleotide sequence view in geneious.

Left panel shows the name of strains. This is the selected list of strains to be analyzed. Top row is the consensus. Nucleotide dissimilarities are showing in different colors. Dotted lines are the similarities. We can also find the mutant position of different nucleotides.

We can also make the amino acid alignment by translating the nucleotide sequence. Throughout the analysis we will use amino acid sequence to find out the mutation of different species as amino acid shows the significant way to distinguish the small differences of different species.

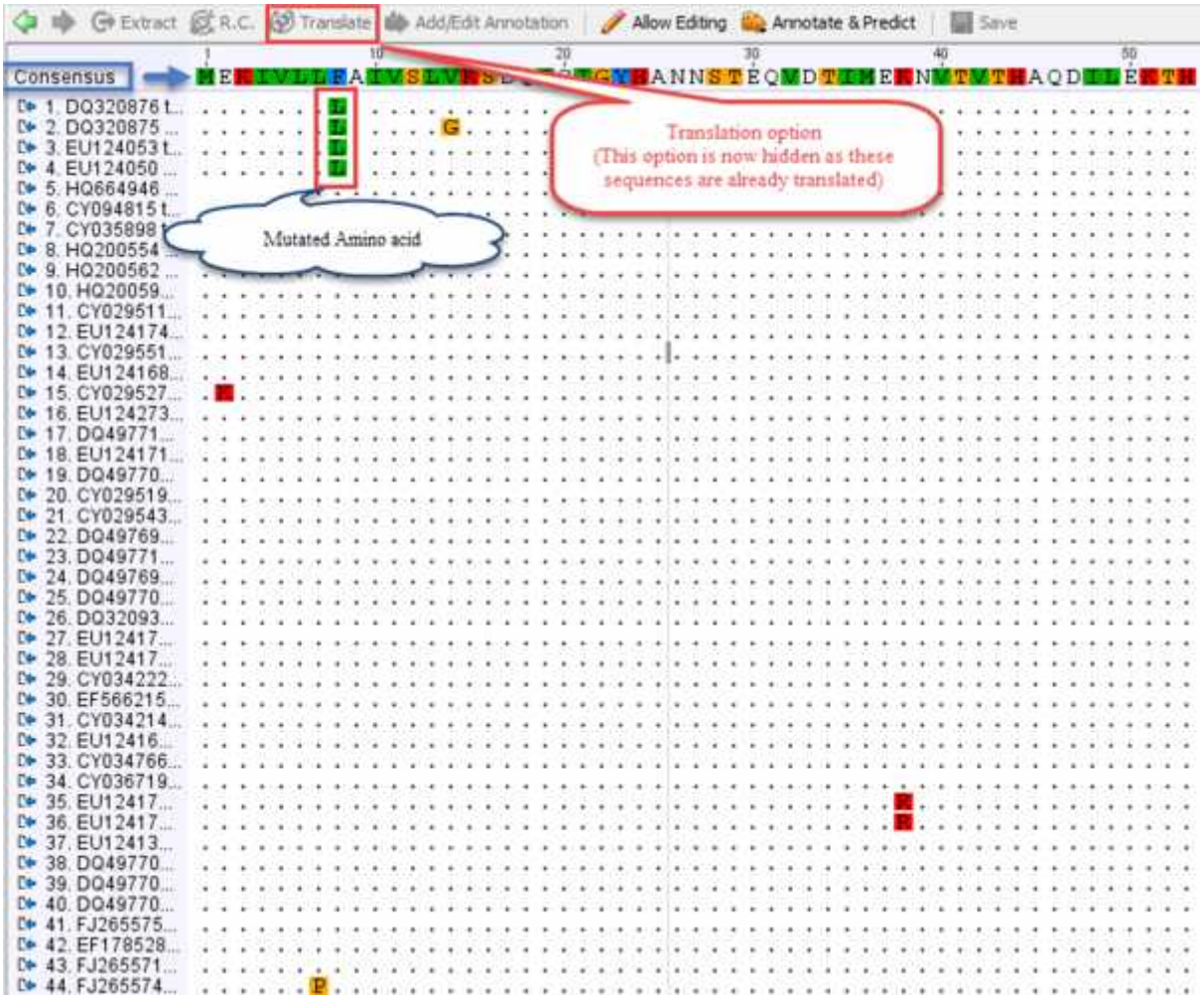


Figure: Overview of Amino Acid sequence view.

It is very easy to find out the mutation point from amino acid sequence as there is less number of amino acid sequence then nucleotide sequence. Phylogenetic tree can be made from both nucleotide sequence and amino acid sequence alignment. We will make phylogenetic tree here with the nucleotide sequence.

In each year we have taken 150+ random strains to analyze. To ease our analysis we have shorten the description of downloaded strain in document table. These lists of abbreviations are as follows:

Abbreviations: Fujian – FJ, Vietnam – VN, Guangxi – GX, Jiangxi – JX, Hunan – HN, Astrakhan – AKH, Yunnan – YN, Guiyang – GY, Xinjiang – XJ, Cambodia – KH, Egypt – EGP, Nigeria – NG, Bavaria – BVR, Qinghai – QH, Sweden – SE, Laos – LAO, Kuwait – KW, Chicken – CK, Duck – DK, Goose – GS, Swan – SN, Bar-Headed Goose – BHGS, Muscovy Duck – MDK, Mallard – Mall, Wild Duck – WDK, Swan – SN, Mute Swan – MSN, Cygnus olor – CGO, Tufted Duck – TDK, Goldeneye Duck – GDK, Herring Gull – HGL.

In this analysis all trees are made with Neighbor joining method. The numbers which will show above the branch node represents the neighbor-joining bootstrap value generated from 1,000 replicates. Data collected form NCBI using BLAST search inside geneious. Scale bar represents the nucleotide substitutions per site.

10.2 Result

In 2005 we have selected total 169 strains of Neuraminidase (NA) strain of H5N1.

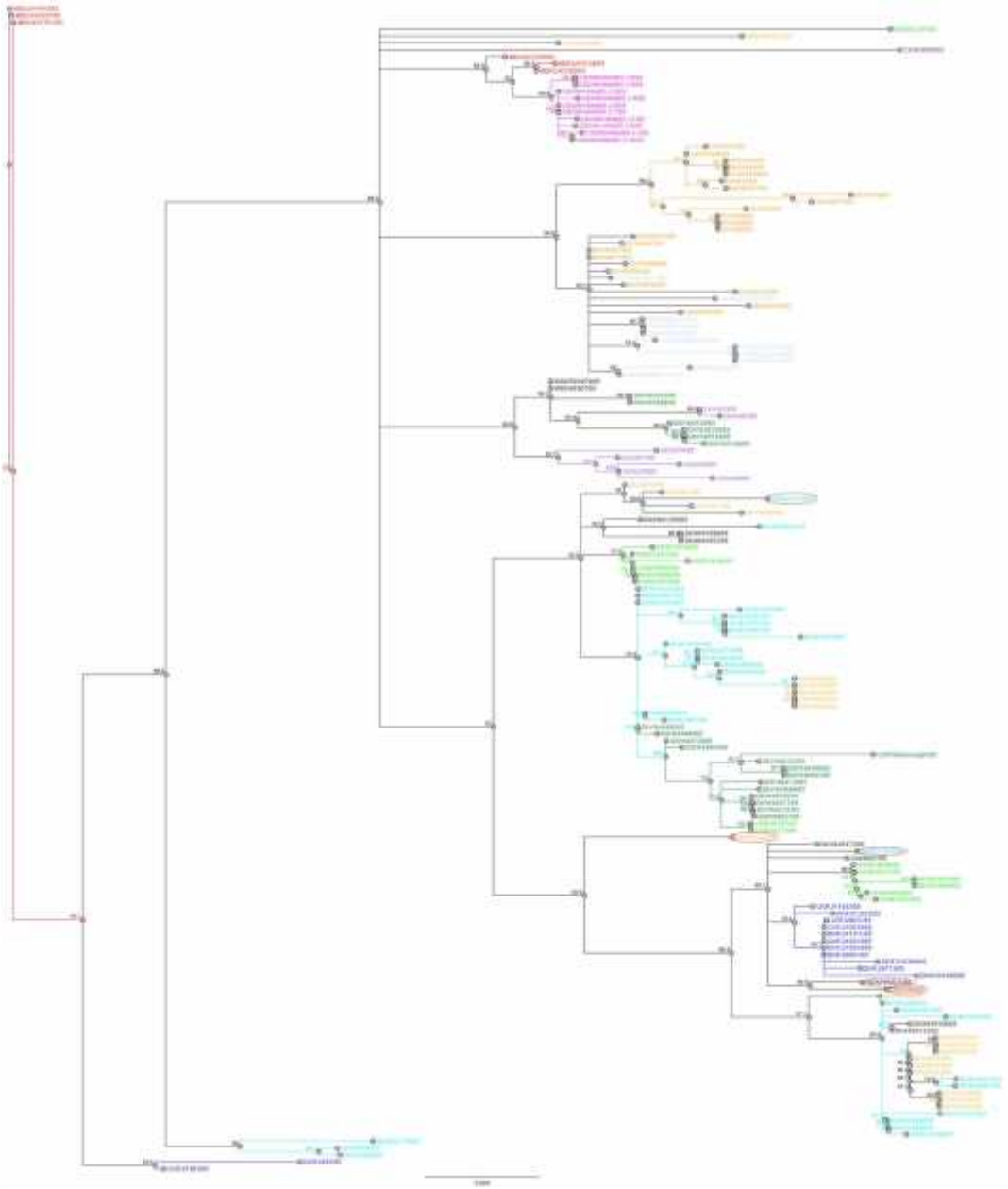


Figure: Phylogenetic tree of the year 2005.

List of diverse strains from our analysis of 2005 are given below.

Accession number	Strain name
EU930918	DK/VN/205/05
EF124227	GS/GX/5414/05
HM172200	DK/JX/80/05
EU329179	WDK/HN/021/05
EF124280	GS/GX/3316/05

In 2006 we have selected total 148 strains of Neuraminidase (NA) strain of H5N1.

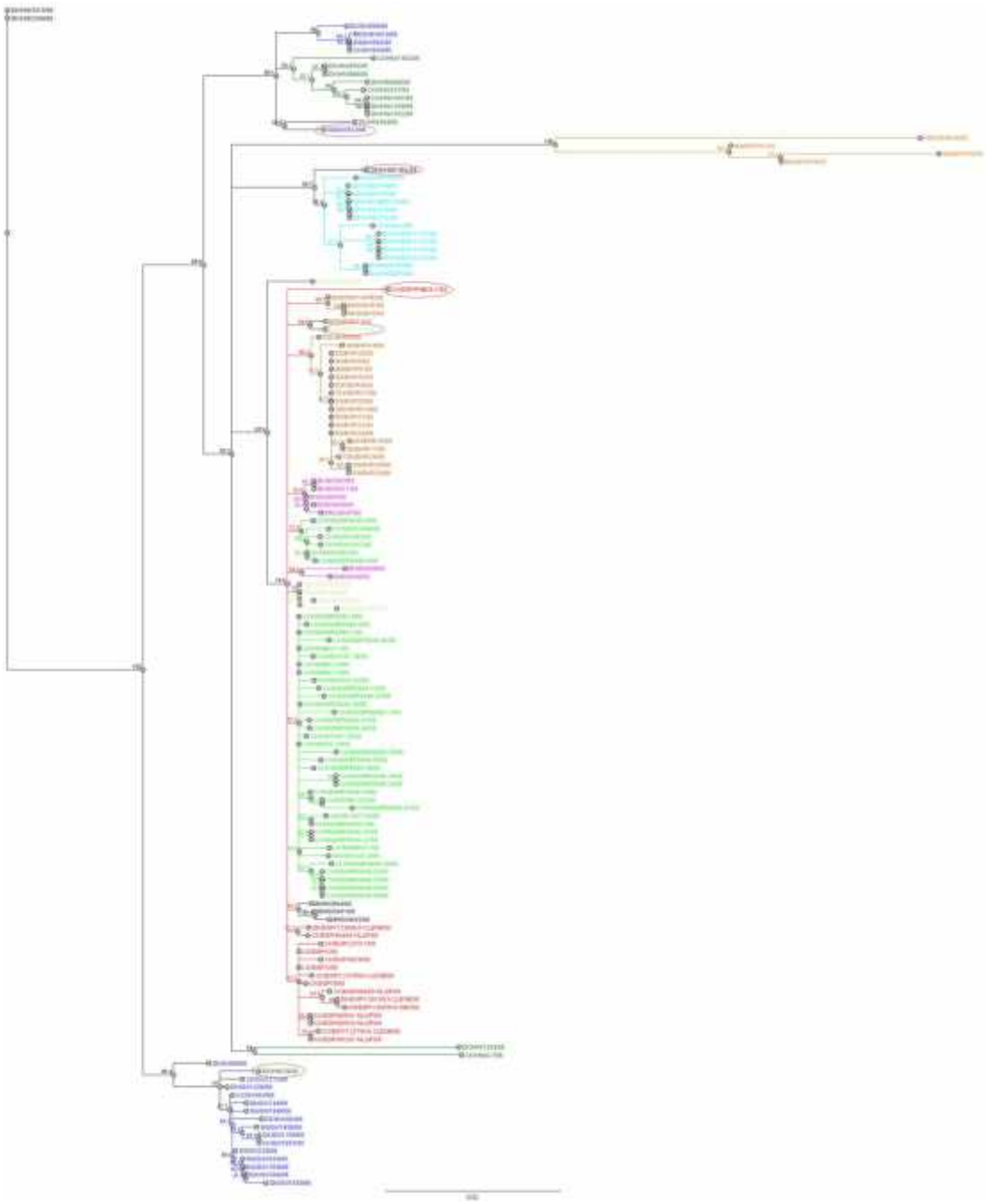


Figure: Phylogenetic tree of the year 2006.

List of diverse strains from our analysis of 2006 are given below.

Accession number	Strain name
HQ200451	DK/KH/D14AL/06
AB601142	CK/EGP/RIMD6-1/06
EU889100	HGL/SE/V1116/06
HM172177	DK/HN/29/06
EF124273	GS/GX/532/06

In 2007 we have selected total 158 strains of Neuraminidase (NA) strain of H5N1.

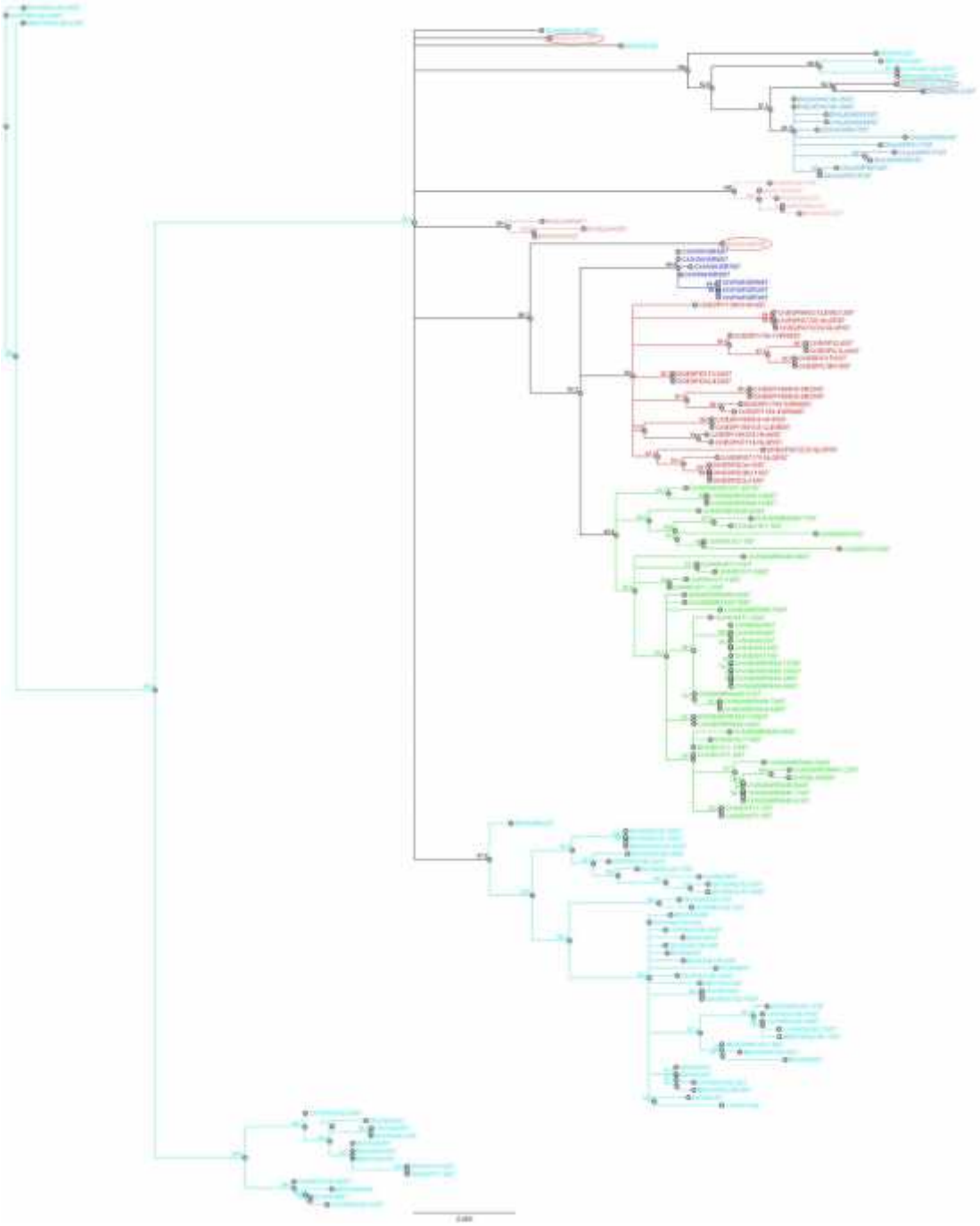


Figure: Phylogenetic tree of the year 2007.

List of diverse strains from our analysis of 2007 are given below.

Accession number	Strain name
FJ602852	BHGS/QH/13/07
FJ461727	BHGS/QH/F/07
CY030455	DK/VN/NCVD-31/07

From three years we have got some strains which seems to diverse from our analysis. We did our literature search but we did not get any information about these diverse strains. So it seems to us that these strains are not responsible for anti-viral resistance. Neighbor joining method and bootstrap value shows that these diverse strains are showing Neuraminidase resistance to influenza virus.

From the phylogenetic tree of 2005 we came to see that all VN like viruses formed a cluster. Among them GS/GX/3316/05 shows 59.8 bootstrap value with DK/VN/12/05. Number of bootstrap value is showing the similarities between these two strains. Strain GS/GX/3316/05 may be migrated or may be mutated which can show antigenic drift.GS/GX/5414/05 shows 65.7 bootstrap values with DK/HN/5472/05. Number of bootstrap value is showing the similarities between these two strains. Strain GS/GX/5414/05 may be migrated or may be mutated which can show antigenic drift. DK/JX/80/05 and WDK/HN/021/05 viruses are closely related to each other as they both are detached from their region and showing the same bootstrap value 69.8. One VN virus (DK/VN/205/05) forming a clade with HB like viruses. Genetic character of this virus is different from other VN like viruses as this is detached from the cluster of VN like viruses showing the antigenic drift.

Phylogenetic tree analysis of 2006 shows that there are few diverse strains which may show genetic drift. One virus from Egypt (CK/EGP/RIMD6-1/06) form clade with the QG like viruses. HGL/SE/V1116/06 shows 94.6 bootstrap values with MSN/BVR/12/06. Number of bootstrap value is showing the similarities between these two strains. Strain HGL/SE/V1116/06 may be migrated or may be mutated which can show antigenic drift. Among them GS/GX/532/06 shows 50.8 bootstrap values with DK/HN/344/06. Number of bootstrap value is showing the similarities between these two strains. Strain GS/GX/532/06 may be migrated or may be mutated which can show antigenic drift. One HN like virus (DK/HN/29/06) forming a clade with Guangxi (GX) like viruses. One KH virus (DK/KH/D14AL/06) forming a clade. Genetic character of this virus is different from other KH like viruses as this is detached from the cluster of KH like viruses showing the antigenic drift.

Phylogenetic tree analysis of 2007 shows that there are few diverse strains which may show genetic drift.Among them DK/VN/NCVD-31/07 shows 92.4 bootstrap values with DK/LAO/PO127/07. Number of bootstrap value is showing the similarities between these two strains. Strain DK/VN/NCVD-31/07 may be migrated or may be mutated which can

show antigenic drift. TWO QH virus (BHGS/QH/13/07 and BHGS/QH/F/07) forming a separate clade. Genetic character of this virus is different from other QH like viruses as these are detached from the cluster of QH like viruses showing the antigenic drift.

Chapter 11

Discussion & Conclusion

11.1 Discussion

Oseltamivir and zanamivir are neuraminidase inhibitors (NAIs) that are commonly used against type A and type B influenza infections. Oseltamivir is administered orally, whereas zanamivir is inhaled. Another drug of this type is peramivir, which is currently being studied. NAIs competitively bind to the highly conserved NA active site by mimicking sialic acid (N-acetylneuraminic acid), which is the natural substrate of NA. This inhibits the enzyme's key function by destroying neuraminic acid-containing receptors, which prevents the release of progeny virions from infected cells and any possible dissemination to neighbouring cells (Neal Nathanson *et al*. 2007). The highly conserved NA enzyme active site is comprised of catalytic amino acid residues that directly interact with the substrate (R118, D151, R152, R224, E276, R292, R371, and Y406) and framework (E119, R156, W178, S179, D198N, I222, E227, H274, E277, N294, and E425) residues that support the catalytic residues (Ferraris *et al*. 2008; & Le *et al*. 2010).

Since the first human cases in 1997 in Hong Kong (Claas *et al*. 1998; & Subbarao *et al*. 1998) sporadic human infection with highly pathogenic avian influenza A(H5N1) virus has caused illness in more than 600 persons in 15 countries in Asia, the Middle East, Europe, and Africa, with an overall mortality rate of approximately 60% (WHO, 2009). Because the seasonal influenza vaccine does not elicit effective immunity against H5N1 influenza viruses, we must rely on antiviral drugs to combat these deadly viruses; thus, the acquisition of NAI resistance by H5N1 influenza viruses is a serious public health concern. Oseltamivir-resistant H5N1 viruses with the H274Y NA mutation have been isolated from three patients during drug treatment or prophylaxis (de Jong *et al*. 2005; Le QM, *et al*. 2005) and those with the N294S NA mutation, from two patients in Egypt (Earhart *et al*. 2005). In addition, the highly pathogenic A/ Hanoi / 30408 / 2005(H5N1) influenza virus that was isolated from a patient treated with oseltamivir had a mixed oseltamivir-sensitive and oseltamivir-resistant population (Le QM, *et al*. 2005). Ten resistant clones that were randomly picked from plaques of the virus in MDCK cells possessed either a H274Y or N294S NA mutation (Le QM, *et al*. 2005).

The virulence of H5N1 viruses carrying either a H274Y or N294S NA mutation was addressed using cell culture, mouse, and ferret models. In a mouse model, recombinant A/ Vietnam / 1203 / 2004-like (H5N1) influenza viruses possessing either the H274Y or N294S

NA substitution exhibited lethality similar to that of the wild-type virus (Yen *et al*. 2007). For the less-virulent A/ Hanoi / 30408 / 2005(H5N1) oseltamivir resistant clone, the N294S NA substitution attenuated the virus in mice, although the degree of attenuation was lower than that caused by the H274Y NA substitution (Kiso *et al*. 2011). Le and collegues (Le QM *et al*. 2005) reported that the oseltamivir-resistant A/ Hanoi / 30408 / 2005 (H5N1) influenza virus with the H274Y NA mutation was attenuated compared to the wildtype virus, as reflected by less-efficient replication in the ferret upper respiratory tract. However, the A/ Hanoi / 30408 / 2005(H5N1) influenza virus does not cause severe infection in ferrets inoculated with 2 · 105 PFU per animal (Le QM *et al*. 2005). This mild infectivity is in contrast to the virulence of human A/ Vietnam / 1203 / 2004 (H5N1) virus in ferrets. Inoculation with A/ Vietnam / 1203 / 2004(H5N1) virus with a dose as low as 10 EID50 resulted in severe disease, and the deaths of two of three animals inoculated; higher doses (102 or 103 EID50) caused high fever, substantial weight loss, anorexia, and extreme lethargy, and werelethal to all animals (Govorkova *et al*. 2007). Compared to the wild-type virus, a recombinant A/ Vietnam / 1203 / 2004-like (H5N1) virus with the H274Y NA mutation replicated to similar titers in the upper respiratory tract of ferrets and caused similar disease signs; none of the animals survived when infected with either oseltamivir-resistant or -sensitive viruses (Govorkova *et al*. 2010).

The influence of multiple genes on the fitness of viruses carrying the H274 NA mutation cannot be excluded. In our study, we focused on additional NA mutations, and sequence analysis of individual NA clones (Govorkova *et al*. 2007; & Mitamura *et al*. 2004) was done to identify potential host-dependent and compensatory NA mutations. We found that the NA mutations E119A and N294S, which confer cross-resistance to oseltamivir and zanamivir (Ferraris *et al*. 2008; & McKimm *et al*. 2005) can emerge spontaneously H5N1 influenza virus in ferrets. Further, we observed that mutations at NA catalytic (R292K) and framework (I222L and N294S) sites and in close proximity to the NA enzyme active site (V116I, I117T/V, Q136H, K150N, and A250T) emerged without drug pressure in both pairs of H5N1 viruses. Compensatory mutations in NA or other genes may mitigate any fitness cost imposed by resistance mutations. Our study identified 6 potential compensatory NA changes (D103V, F132S, I254V, E276A, H296L, and F466S) that may affect the fitness of viruses with the H274Y NA mutation. We suggest that NA mutations at residues I254 (I254V) and E276 (E276A) are of importance. Interestingly, we observed differences in predominance of I254V and E276A NA mutations in different genetic backgrounds: the I254V mutation was

identified in the A/Vietnam/1203/04 (H5N1)-like genetic background and E276A in the A/Turkey/15/06 (H5N1)-like genetic background. Moreover, the I254V NA mutation was identified only when ferrets were inoculated with the mixtures of VN-WT and VN-H274Y viruses, not in ferrets inoculated with VN-H274Y virus. None of the potential compensatory NA mutations was identified in the original inoculum used to infect ferrets. The H274Y NA mutation causes a large shift in the position of the side chain of the neighboring E276 residue (Collins *et al*. 2008), which must form a salt bridge with R224 to accommodate the large hydrophobic pentyl ether group of oseltamivir (Smith *et al*. 2002). The E276A substitution results in a shorter, nonpolar side chain at that residue and thus can affect interactions in this region. Residue I254 is located near the NA active site, and although it does not alter polarity, it results in a shorter side chain and thus may indirectly affect the residues in the NA active site.

One important conclusion from these studies is that a particular NAI resistance-associated marker can cause different effects on fitness in different H5N1 virus genetic and virulence backgrounds. Deficiency in NA function caused by an NAI resistance mutation may not be deleterious for highly pathogenic H5N1 viruses because of the extremely efficient replication of these viruses.

The molecular mechanisms that enable avian influenza viruses to cross the species barrier and transmit efficiently in humans are incompletely understood. Some experiments have been done to identify the transmission pattern and it shows poorer transmission from infected to susceptible animals [Alexander *et al.* 1978, Westbury *et al.* 1979 & Ito *et al.* 2001]. Migration process can influence transmission of viruses. Migratory birds can carry pathogens from country to country thereby playing a role distributing influenza viruses. Avian influenza A consists of two major glycoproteins which are Hemagglutinin (HA) and Neuraminidase (NA) [Gambotto *et al.* 2008]. HA glycoproteins are more prone to attach to the cell surface sialic acid receptors. There is a difference between host surface receptors on the target cell which is believed to be the possible restrictive factor of avian influenza.

Human infections are periodic. In some cases these viruses are accompanied by high mortality. As a result they are the major concern about the potential H5N1 as an endemic virus.

Although human infections are sporadic, they are accompanied by high mortality, raising major concerns about the potential of H5N1 as a pandemic virus [Ungchusak *et al.* 2005]. Fortunately, H5N1 viruses have not yet naturally acquired the ability to stably transmit between humans [Wang *et al.* 2008 & Imai *et al.* 2012]. One factor that limits transmission of avian viruses in humans is the receptor specificity of the hemagglutinin (HA) [Rott *et al.* 1992]. Avian viruses, like H5N1, preferentially bind to α 2, 3 sialosides (avian-type receptors), whereas human viruses prefer α 2,6sialosides (human-type receptors that are found in the human respiratory tract).

Before functioning as a virus it needs post translational cleavage by host proteases [Tong *et al.* 2013]. HA followed by NA are important antigenic determinant from which neutralizing antibodies are directed. There are several subtypes of HA and NA. 18 different HA subtypes (H1 to H18) and 11 different NA subtypes (N1 to N11) are found [Shinya *et al.* 2006].

In humans, the SA α 2,6 Gal receptor is expressed mainly in the upper airway, while the SA α 2,3 Gal receptor is expressed in alveoli and the terminal bronchiole [Gambaryan *et al.* 2006]. A virus with good affinity to both SA α 2,3 Gal and SA α 2,6 Gal receptors may be a very dangerous one, which could both infect efficiently via its binding to SAα2,6Gal in the upper airway and cause severe infection in the lung via its binding to SAα2,3Gal.

Data used in this study are obtained inside using nucleotide BLAST search from publicly available database of National Centre for Biotechnology Information (NCBI). Multiple sequence alignments, editing, assembly of strains were performed in windows platform with the Geneious program version 7.1.3 (trial).

In this study we will analyze some avian neuraminidase (H5N1) of different years. Analysis includes building nucleotide sequence and translating them into amino acid sequence. Then we will study amino acid positions with respect to some reported mutation to see the genetic pattern. After analyzing we will try to find out whether there are any similarities between avian and human or not. There are some reported avian H5N1 strains that affect human which are A/Goose/Hong Kong/739.2/2002 [Yohei *et al.* 2011], A/duck/Egypt/D1Br12/2007 [Nipa *et al.* 2009], A/Duck/Singapore/3/97 [De *et al.* 2014], A/egret/Egypt/1162/2006 [Maines *et al.* 2011]. All of these strain show preferential binding to Sia α (2, 6) Gal receptor that can infect a human. Few specific positions of amino acids are responsible for this binding.The

NA of VN1203 in a PR8 backgrand (designated N158D/N224K/Q226L or N158D/N224K/Q226L/T318I, respectively) preferentially bind to Siaa2,6Gal and attach to human tracheal epithelia [Chen *et al.* 2012 & Le *et al.* 2005]. In our avian neuraminidase (H5N1) analysis we did not find the exact location where reported mutations are occurred. But we found similar amino acid near the reported mutated position. We have analyzed around (before and after the mutation point) twenty positions with respect to the reported mutation point.

In this study we will analyze some avian neuraminidase (H5N1) of different years. Analysis includes building nucleotide sequence and translating them into amino acid sequence. Then we will study amino acid positions with respect to some reported mutation to see the genetic pattern.

Twenty amino acids, their symbols and short forms are given below

Nonpolar (Hydrophobic) amino acids			**Polar (Hydrophilic) amino acids**		
Amino acid (AA)	**Short form**	**Symbol**	**Amino acid (AA)**	**Short form**	**Symbol**
Glycine	Gly	G	Serine	Ser	S
Alanine	Ala	A	Threonine	Thr	T
Valine	Val	V	Cysteine	Cys	C
Leucine	Leu	L	Tyrosine	Try	Y
Isoleucine	Ile	I	Asparagine	Asn	N
Methionine	Met	M	Glutamine	Gln	Q
Phenylalanine	Phe	F			
Tryptophan	Trp	W			
Proline	Pro	P			

Electrically charged (Negative) amino acids			**Electrically charged (Positive) amino acids**		
Amino acid (AA)	**Short form**	**Symbol**	**Amino acid (AA)**	**Short form**	**Symbol**
Aspartic acid	Asp	D	Lysine	Lys	K
Glutamic acid	Glu	E	Arginine	Arg	R
			Histidine	His	H

Table: Summary of some reported position which can be responsible for antigenic shift from avian to human.

Reported neuraminidase amino acid position with mutation that bind α 2, 6 receptor can affect human.

Amino acid mutation position with reference	Short form	Amino acid in Geneious in reported position
Asn-158-Asp (Maines *et al.* 2011)	N158D	T
Asn-224-Lys (Chen *et al.* 2012)	N224K	W
Gln-226-Leu (Maines *et al.* 2011)	Q226L	N
Thr-318-Ile (Chen *et al.* 2012)	T318I	E

Table: Correlation of Reported neuraminidase amino acid position with mutation with our experimental strain, and their mutation pattern (exact and around twenty positions): Original Amino Acid analysis.

Reported AA with mutation point	AA in Geneious in reported position	Exact AA which matches with the reported data with respect to specific mutation point (before 20 positions)	Exact AA which matches with the reported data with respect to specific mutation point (after 20 positions)
N158D	T158S(1) Polar to Polar T158A(2) Polar to Nonpolar	N142 N147S(1) Polar to Polar	N176
N224K	W224	N205S(7) Polar to Polar N205D(1) Polar to Negative N214	N226S(1) Polar to Polar N227Y(1) Polar to Polar N227S(1) Polar to Polar N240
Q226L	N226S(1) Polar to Polar		Q232P(1) Polar to Nonpolar
T318I	E318K(4) Electrically charged Negative to Positive		T239A(1) Polar to Nonpolar T239K(1) Polar to Positive

In case of N158D (Asn-158-Asp) reported position, we found amino acid Threonine (Thr) in our software (Geneious). We found two mutations here, Threonine (Thr) to Serine (Ser) in one strain and from Threonine (Thr) to Alanine (Ala) in two strains. We found amino acid Asparagine (Asn) in two positions (142 & 147) which is located within twenty positions before 158. We found one mutation here from Asparagine (Asn) to Serine (Ser) in one strain. We have also found Asparagine (Asn) in one position (176) which is located within twenty positions after 158. We found no mutation here.We found T158S and T158A mutation which indicates that polarity is changed from polar to polar and from polar to non-polar as Threonine (Thr), Serine (Ser) is Polar and Alanine (Ala) is non-polar.

In case of N224K (Asn-224-Lys) reported position, we found amino acid Tryptophan (Trp) in our software (Geneious). We found amino acid Asparagine (Asn) in two positions (205 & 214) which is located within twenty positions before 224. We found two mutations here from Asparagine (Asn) to Serine (Ser) in seven strains and from Asparagine (Asn) to Aspartic acid (Asp) in one strain. We have also found Asparagine (Asn) in three positions (226, 227 and 240) which is located within twenty positions after 224. We found three mutations here from Asparagine (Asn) to Serine (Ser) in one strain, Asparagine (Asn) to Tyrosine (Try) in one strain and from Asparagine (Asn) to Serine (Ser) in one strain.We found N205S and N205D mutation which indicates that polarity is changed from polar to polar and from polar to negative as Threonine (Thr), Serine (Ser) is Polar and Aspartic acid (Asp) is negative.We also found N227Y mutation which indicates that polarity is changed from polar to polar as Threonine (Thr), Serine (Ser) is Polar and Tyrosine is polar.

In case of Q226L (Gln-226-Leu) reported position, we found amino acid Threonine (Thr) in our software (Geneious). We found one mutations here, Threonine (Thr) to Serine (Ser) in one strain. We found amino acid Glutamine (Gln) in one position (232) which is located within twenty positions after 226. In this position there is one mutation here fron Glutamine (Gln) to Proline (Pro) in one strain. We found N226S and Q232P mutation which indicates that polarity is changed from polar to polar and from polar to non-polar as Threonine (Thr), Serine (Ser) is Polar and Proline (Pro) is non-polar.

In case of T318I (Thr-318-Ile) reported position, we found amino acid Glutamic acid (Glu) in our software (Geneious). We found one mutation here form Glutamiic (Glu) to Lysine (Lys) in four strain. We found amino acid Threonine (Thr) in position (239) which is located

within twenty positions after 318. We found two mutations here from Threonine (Thr) to Alanine (Ala) in one strain and from Threonine (Thr) to Lysine (Lys) in one strain. In T239K, T239A and E318K mutations which indicate that polarity is changed from polar to polar, polar to positive and negative to polar as Threonine (Thr) is Polar, Alanine (Ala) is non-polar, Lysine (Lys) is positive and Glutamic acid (Glu) is negative.

Table: Correlation of Reported neuraminidase amino acid position with mutation with our experimental strain, and their mutation pattern (exact and around twenty positions): Mutated Amino Acid analysis.

Reported AA with mutation point	**AA in Geneious in reported position**	**Exact AA which matches with the reported data with respect to specific mutation point (before 20 positions)**	**Exact AA which matches with the reported data with respect to specific mutation point (after 20 positions)**
N158D	T158S(1) Polar to Polar T158A(2) Polar to Nonpolar	D143 D152	
N224K	W224	K212	
Q226L	N226S(1) Polar to Polar	L211V(1) Nonpolar to Nonpolar	L229
T318I	E318K(4) Electrically charged Negative to Positive		I321T(1) Nonpolar to Polar I321M(3) Nonpolar to Nonpolar I321V(2) Nonpolar to Nonpolar I324

In case of N158D (Asn-158-Asp) reported position, we found amino acid Threonine (Thr) in our software (Geneious). We found two mutations here, Threonine (Thr) to Serine (Ser) in one strain and from Threonine (Thr) to Alanine (Ala) in two strains. We found amino acid Asparagine (Asn) in two positions (143 & 152) which is located within twenty positions before 158. There is no mutation here. We found T158S and T158A mutation which indicates

that polarity is changed from polar to polar and from polar to non-polar as Threonine (Thr), Serine (Ser) is Polar and Alanine (Ala) is non-polar.

In case of N224K (Asn-224-Lys) reported position, we found amino acid Tryptophan (Trp) in our software (Geneious). We found amino acid Lysine (Lys) in one positions (212) which is located within twenty positions before 224. There is no mutation here.

In case of Q226L (Gln-226-Leu) reported position, we found amino acid Threonine (Thr) in our software (Geneious). We found one mutations here, Threonine (Thr) to Serine (Ser) in one strain. We found amino acid Leucine (Leu) in one position (211) which is located within twenty positions before 226. . In this position there is one mutation here fronLeucine (Leu) to Valine (Val) in one strain. We found amino acid Leucine (Leu) in one position (229) which is located within twenty positions after 226. There is no mutation hereWe found N226S and L211V mutation which indicates that polarity is changed from polar to polar and from non-polar to non-polar as Threonine (Thr), Serine (Ser) is Polar and Proline (Pro), Valine (Val) is non-polar.

In case of T318I (Thr-318-Ile) reported position, we found amino acid Glutamic acid (Glu) in our software (Geneious). We found one mutation here form Glutamiic (Glu) to Lysine (Lys) in four strains. We found amino acid Isoleucine (Ile) in two positions (321 and 324) which is located within twenty positions after 318. We found three mutations here from Isoleucine (Ile) to Threonine (Thr) in one strain, from Isoleucine (Ile) to Methionine (Met) in three strains and from Isoleucine (Ile) to Valine (Val) in two strains. In I321T, I321M, I321V and E318K mutations which indicate that polarity is changed from non-polar to non-polar and non-polar to polar as Isoleucine (Ile), Methionine (Met) and Valine (Val) is nonpolar and Threonine (Thr) is polar.

In this study we will analyze some avian neuraminidase (H5N1) of different years. Analysis includes building nucleotide sequence and translating them into amino acid sequence. Then we will study amino acid positions with respect to some reported mutation to see the genetic pattern.

We found some avian amino acid position H274Y, H275Y, N294S, N295S, D198G , D199G , V116A, H252Y, S246N and I222T are specific reported position which can be responsible

for N1 (H1N1) neuraminidase mutations causing reduced sensitivity to NAIs [Le QM *et al*. 2005, Pizzorno *et al*. 2011, Hurt *et al*. 2009 & Nguyen *et al*. 2012]. In our avian neuraminidase (H5N1) analysis we did not find the exact location where reported mutations are occurred. But we found similar amino acid near the reported mutated position. We have analyzed around (before and after the mutation point) twenty positions with respect to the reported mutation point.

Table: Summary of some reported position which can be responsible for N1 (H5N1) neuraminidase mutations causing reduced sensitivity to NAIs.

Amino acid mutation position with reference	Short form	Amino acid in Geneious in reported position	NAIs
His-274-Try (Le QM *et al*. 2005)	H274Y	V	Oseltamivir Peramivir
His-275-Try (Pizzorno *et al*. 2011)	H275Y	E	Oseltamivir Peramivir
Asn-294-Ser (Le QM *et al*. 2005)	N294S	E	Oseltamivir
Asn-295-Ser (Pizzorno *et al*. 2011)	N295S	I	Oseltamivir
Asp-198-Gly (Hurt *et al*. 2009)	D198G	I	Oseltamivir Zanamivir
Asp-199-Gly (Pizzorno *et al*. 2011)	D199G	G	Oseltamivir Zanamivir
Val -116- Ala (Nguyen *et al*. 2012)	V116A	V	Oseltamivir Zanamivir
His -252- Try (Nguyen *et al*. 2012)	H252Y	S	Oseltamivir
Ser -246- Asn (Nguyen *et al*. 2012)	S246N	V	Oseltamivir
Ile -222- Thr (Nguyen *et al*. 2012)	I222T	K	Oseltamivir

Table: Analysis of our study data with respect to reported mutation point causing reduced sensitivity to NAIs of avian H5N1.

Correlation of Reported NAIs resistance specific avian amino acid position with our experimental strain, and their mutation pattern (exact and around twenty positions): Original Amino Acid analysis.

Reported AA with mutation point	AA in Geneious in reported position	Exact AA which matches with the reported data with respect to specific mutation point (before 20 positions)	Exact AA which matches with the reported data with respect to specific mutation point (after 20 positions)
H274Y	V274I (2) Nonpolar to Nonpolar		H282
H275Y	E275		H282
N294S	E294G Electrically charged Negative to Nonpolar	N280D(1) Polar to Electrically charged Negative N280Y(1) Polar to Polar N280T(1) Polar to Polar N280K(1) Polar to Electrically Charged Positive	N302 N307 N314
N295S	I295V(7) Nonpolar to Nonpolar	N280D(1) Nonpolar to Electrically charged Negative N280Y(1) Polar to Polar N280T(1) Polar to Polar N280K(1) Polar to Electrically Charged Positive	N302 N307 N314
D198G	I198	D191	D204
D199G	G199	D191	D204 D219

V116A	V116I(2) Nonpolar to Nonpolar	V99 I Nonpolar to Nonpolar V114	
H252Y	S252N(4) Polar to Polar S 252 G(4) Polar to Nonpolar		
S246N	V246L(3) V246I(6) Nonpolar to Nonpolar (L, I)	S 234 S 242 F(1) Polar to Nonpolar	S252G(4) Polar to Nonpolar S252N(4) Polar to Polar S257
I222T	K222Q(1) K222R(3) K222N(1) Electrically charged Positive to Polar (Q, N)and to Positive (R)	I216V(1) I217 I221V(1) Nonpolar to Nonpolar	I228V(1) I228T(2) I228L(1) Nonpolar to Nonpolar (V, L) and to Polar (T)

In case of H274Y (His-274-Try) reported position, we found amino acid Valine (V) in our software (Geneious). We found one mutation here, Valine (V) to Isoleucine (I) in two strains. We found amino acid Histidine (H) in one positions (282) which are located after twenty positions 274. We found V 274 I (2) mutation in two strains which indicates that there is polarity changed from non-polar to non-polar as Valine (V) is non-polar and Isoleucine (I) is non-polar.

In case of H275Y (His-274-Try) reported position, we found amino acid Glutamic acid (E) in our software (Geneious). There are no mutations here. We found amino acid Histidine (H) in one positions (282) which are located after twenty positions 274. There are no mutations here.

In case of N294S (Asn-294-Ser) reported position, we found amino acid Glutamic acid (E) in our software (Geneious). We found one mutation here, Glutamic acid (E) to Glycine (G) in one strain. We found Asparagine (Asn) in one position (280) which is located within twenty positions before 294. In 280 positions there are four mutations here from Asparagine (N) to Aspartic acid (D) in one strain, from Asparagine (N) to Tyrosine (Y) in one strain, from Asparagine (N) to Threonine (T) in one strain and from Asparagine (N) to Lysine (K) in one strain. We found amino acid Asparagine (N) in three positions (302, 307 and 314) which are located after twenty positions 294.There are no mutation here. In N280D, N280Y, N280T

and N280K mutations which are indicate that there is polarity changed from polar to electrically charged negative, to positive and to polar as Asparagine (N) is polar, Aspartic acid (D) is electrically charged negative, Tyrosine (Y) is polar, Threonine (T) is polar and Lysine (K) is electrically charged positive. In E294G mutation which is indicates that there is polarity change from electrically charged negative to nonpolar as Glutamic acid (E) is negative and Glycine (G) is nonpolar.

In case of N295S (Asn-295-Ser) reported position, we found amino acid Isoleucine (I) in our software (Geneious). We found one mutation here, Isoleucine (I) to Valine (V) in seven strains. We found Asparagine (Asn) in one position (280) which is located within twenty positions before 295. In 280 positions there are four mutations here from Asparagine (N) to Aspartic acid (D) in one strain, from Asparagine (N) to Tyrosine (Y) in one strain, from Asparagine (N) to Threonine (T) in one strain and from Asparagine (N) to Lysine (K) in one strain. We found amino acid Asparagine (N) in three positions (302, 307 and 314) which are located after twenty positions 295.There are no mutation here. We found I295V mutation which indicates that there is polarity changed from non-polar to non-polar as Isoleucine (I) is non-polar and Valine (V) is non-polar. In N280D, N280Y, N280T and N280K mutations which are indicate that there is polarity changed from polar to electrically charged negative, to positive and to polar as Asparagine (N) is polar, Aspartic acid (D) is electrically charged negative, Tyrosine (Y) is polar, Threonine (T) is polar and Lysine (K) is electrically charged positive.

In case of D198G (Asp-198-Gly) reported position, we found amino acid Isoleucine (I) in our software (Geneious). There are no mutations here. We found amino acid Aspartic acid (D) in one positions (191) which are located before twenty positions 198. We found amino acid Aspartic acid (D) in one positions (204) which are located after twenty positions 198. There are no mutations here.

In case of D199G (Asp-199-Gly) reported position, we found amino acid Glycine (G) in our software (Geneious). There are no mutations here. We found amino acid Glycine (G) in one positions (191) which is located within twenty positions before 199. There are no mutations here. We found amino acid Aspartic acid (D) in two positions (204 and 219) which are located after twenty positions 199. There are no mutations here.

In case of V116A (Val -116- Ala) reported position, we found amino acid Valine (V) in our software (Geneious). There is one mutation here from Valine (V) to Isoleucine (I) in two

strains. We found amino acid Valine (V) in two positions (99 and 114) which is located within twenty positions before 116 and one mutation here from Valine (V) to Isoleucine (I) in one strain. We found V116I mutation which is indicates that there is polarity changed from non-polar to non-polar as Valine (V) is non-polar and Isoleucine (I) is non-polar.

In case of H252Y (His -252- Try) reported position, we found amino acid Serine (S) in our software (Geneious). There are two mutations here from Serine (S) to Asparagine (N) in four strains and from Serine (S) to Glycine (G) in four strains. We found S252N mutation which is indicates that there is polarity changed from polar to polar as Serine (S) is polar and Asparagine (N) is polar. We found S252G mutation which is indicates that there is polarity changed from Polar to non-polar as Serine (S) is polar and Glycine (G) is non-polar.

In case of S246N (Ser -246- Asn) reported position, we found Valine (V) in our software (Geneious). There are two mutations here from Valine (V) to Leucine (L) in three strains and from Valine (V) to Isoleucine (I) in six strains. We found V246L and V246I mutation which is indicates that there is polarity changed from non-polar to non-polar as Valine (V) is non-polar, Leucine (L) is non-polar and Isoleucine (I) is non-polar. We found amino acid Serine (S) in two positions (234 and 242) which is located within twenty positions before 246. In 234 positions there is no mutation here. In 242 positions there is one mutation from Serine (S) to Phenylalanine (F) in one strain.

We found S242F mutation which is indicates that there is polarity changed from polar to non-polar as Serine (S) is polar and Phenylalanine (F) is non-polar. We found amino acid Serine (S) in two positions (252 and 257) which is located within twenty positions after 246. In 252 positions there are two mutations here from Serine (S) to Glycine (G) in four strains and from Serine (S) to Asparagine (N) in four strains. In 257 positions there is no mutation here. We found S252G mutation which is indicates that there is polarity changed from polar to non-polar as Serine (S) is polar and Glycine (G) is non-polar. We found S252N mutation which is indicates that there is polarity changed from Polar to Polar as Serine (S) is polar and Asparagine (N) is polar.

In case of I222T (Ile -222- Thr) reported position, we found Lysine (K) in our software (Geneious). There are three mutations here from Lysine (K) to Glutamine (Q) in one strain, from Lysine (K) to Arginine (R) in three strains and from Lysine (K) to Asparagine (N) in one strain. We found K222Q, K222R and K222N mutation which is indicates that there is

changed electrically charged from positive (K) to positive (R) and from positive (K) to polar (Q and N).

We found amino acid Isoleucine (I) in three positions (216, 217 and 221) which is located within twenty positions before 222. In 216 positions there is one mutation here from Isoleucine (I) to Valine (V) in one strain and in 221 positions there is one mutation here from Isoleucine (I) to Valine (V) in one strain. In 217 positions there is no mutation here. In I216V and I221V mutations which are indicate that there is polarity changed from non-polar to non-polar as Isoleucine (I) is non-polar and Valine (V) is non-polar.

We found amino acid Isoleucine (I) in one positions (228) which are located within twenty positions after 222. In 228 positions there are three mutations here from Isoleucine (I) to Valine (V) in one strain, from Isoleucine (I) to Threonine (T) in two strains and from Isoleucine (I) to Leucine (L) in one strain. In I228V, I228T and I228L mutations which are indicate that there is polarity changed from non-polar to non-polar and non-polar to polar as Isoleucine (I) is non-polar, Valine (V) is non-polar, Leucine (L) is non-polar and Threonine (T) is polar.

Table: Correlation of Reported NAIs resistance specific avian amino acid position with our experimental strain, and their mutation pattern (exact and around twenty positions): Mutated Amino Acid analysis.

Reported AA with mutation point	AA in Geneious in reported position	Exact AA which matches with the reported data with respect to specific mutation point (before 20 positions)	Exact AA which matches with the reported data with respect to specific mutation point (after 20 positions)
H274Y	V274I(2) Nonpolar to Nonpolar	Y258	Y281F(2) Polar to Nonpolar Y283 Y289N(1) Polar to Polar
H275Y	E275	Y258	Y281F(2) Polar to Nonpolar Y283 Y289N(1) Polar to Polar
	E294G	S287	S306

N294S	Electrically charged Negative to Nonpolar		S312
N295S	I295V(7) Nonpolar to Nonpolar	S287	S306 S312
D198G	I198	G192	G199 G202 G206A(1) Nonpolar to Nonpolar G206E(2) Nonpolar to Electrically Charged Negative G215
D199G	G199	G 192	G202 G206A(1) Nonpolar to Nonpolar G206E(2) Nonpolar to Electrically Charged Negative G215
V116A	V11I(2) Nonpolar to Nonpolar	A98	
H252Y	S252N(4) Polar to Polar S252G(4) Polar to Nonpolar		Y258
S246N	V246L(3) V246I(6) Nonpolar to Nonpolar (L, I)	N240	N253D(1) Polar to Electrically charged Negative N253H(1) Polar to Electrically charged Positive N253S(2) Polar to Polar
I222T	K222Q(1) K222R(3) K222N(1) Electrically charged Positive to Polar (Q, N)and to Positive (R)	T220 T218	T231

In case of H274Y (His-274-Try) reported position, we found amino acid Valine (V) in our software (Geneious). We found one mutation here, Valine (V) to Isoleucine (I) in two strains. We found amino acid Tyrosine (Y) in one position (258) which is located before twenty positions 274. There are no mutations here. We found amino acid Tyrosine (Y) in three positions (281, 283 and 289) which are located after twenty positions after 274. There are two mutations here from Tyrosine (Y) to Phenylalanine (F) in two strains and from Tyrosine (Y) to Asparagine (N) in one strain. We found V274I mutation in two strains which indicates that there is polarity changed from non-polar to non-polar as Valine (V) is non-polar and Isoleucine (I) is non-polar. We found Y281F and Y289N mutations which indicate that there is polarity changed from polar to non-polar and to polar as Tyrosine (Y) is polar, Phenylalanine (F) is non-polar and Asparagine (N) is polar.

In case of H275Y (His-274-Try) reported position, we found amino acid Glutamic acid (E) in our software (Geneious). We found amino acid Tyrosine (Y) in one position (258) which is located before twenty positions 275. There are no mutations here. We found amino acid Tyrosine (Y) in three positions (281, 283 and 289) which are located after twenty positions after 275. There are two mutations here from Tyrosine (Y) to Phenylalanine (F) in two strains and from Tyrosine (Y) to Asparagine (N) in one strain. We found Y281F and Y289N mutations which indicate that there is polarity changed from polar to non-polar and to polar as Tyrosine (Y) is polar, Phenylalanine (F) is non-polar and Asparagine (N) is polar.

In case of N294S (Asn-294-Ser) reported position, we found amino acid Glutamic acid (E) in our software (Geneious). We found one mutation here, Glutamic acid (E) to Glycine (G) in one strain. We found amino acid Serine (S) in one position (287) which is located within twenty positions before 294. We found amino acid Serine (S) in two positions (306 and 312) which is located within twenty positions after 294. There are no mutations here. In E294G mutation which is indicates that there is polarity change from electrically charged negative to nonpolar as Glutamic acid (E) is negative and Glycine (G) is nonpolar.

In case of N295S (Asn-295-Ser) reported position, we found amino acid Isoleucine (I) in our software (Geneious). We found one mutation here, Isoleucine (I) to Valine (V) in seven strains. We found amino acid Serine (S) in one position (287) which is located within twenty positions before 295. We found amino acid Serine (S) in two positions (306 and 312) which is located within twenty positions after 295. There are no mutations here. We found I295V

mutation which indicates that there is polarity changed from non-polar to non-polar as Isoleucine (I) is non-polar and Valine (V) is non-polar.

In case of D198G (Asp-198-Gly) reported position, we found amino acid Isoleucine (I) in our software (Geneious). There are no mutations here. We found amino acid Glycine (G) in one position (192) which is located within twenty positions before 198. We found amino acid Glycine (G) in four positions (199, 202, 206 and 215) which is located within twenty positions after 198. We found two mutations here from Glycine (G) to Alanine (A) in one strain and from Glycine (G) to Glutamic acid (E) in two strains. In G206A and G206E mutations which indicates that there is polarity changed from nonpolar to nonpolar and nonpolar to electrically charge negative as Glycine (G) and Alanine (A) is nonpolar and Glutamic acid (E) is electrically charged negative.

In case of D199G (Asp-198-Gly) reported position, we found amino acid Isoleucine (I) in our software (Geneious). There are no mutations here. We found amino acid Glycine (G) in one position (192) which is located within twenty positions before 199. We found amino acid Glycine (G) in three positions (202, 206 and 215) which is located within twenty positions after 199. We found two mutations here from Glycine (G) to Alanine (A) in one strain and from Glycine (G) to Glutamic acid (E) in two strains. In G206A and G206E mutations which indicates that there is polarity changed from nonpolar to nonpolar and nonpolar to electrically charge negative as Glycine (G) and Alanine (A) is nonpolar and Glutamic acid (E) is electrically charged negative.

In case of V116A (Val -116- Ala) reported position, we found amino acid Valine (V) in our software (Geneious). There is one mutation here from Valine (V) to Isoleucine (I) in two strains. We found amino acid Alanine (A) in one position (98) which is located within twenty positions before 116. There are no mutations here. We found V116I mutation which is indicates that there is polarity changed from non-polar to non-polar as Valine (V) is non-polar and Isoleucine (I) is non-polar.

In case of H252Y (His -252- Try) reported position, we found amino acid Serine (S) in our software (Geneious). There are two mutations here from Serine (S) to Asparagine (N) in four strains and from Serine (S) to Glycine (G) in four strains. We found S252N mutation which is indicates that there is polarity changed from polar to polar as Serine (S) is polar and Asparagine (N) is polar. We found amino acid Tyrosine (Y) in one position (258) which is located within twenty positions after 252. There are no mutations here. We found S252G

mutation which is indicates that there is polarity changed from Polar to non-polar as Serine (S) is polar and Glycine (G) is non-polar.

In case of S246N (Ser -246- Asn) reported position, we found Valine (V) in our software (Geneious). There are two mutations here from Valine (V) to Leucine (L) in three strains and from Valine (V) to Isoleucine (I) in six strains. We found amino acid Asparagine (N) in one position (240) which is located within twenty positions before 246. There are no mutations here. We found amino acid Asparagine (N) in one position (253) which is located within twenty positions after 246. There are three mutations here from Asparagine (N) to Aspartic acid (D) in one strain, from Asparagine (N) to Histidine (H) in one strain and fron Asparagine (N) to Serine (S) in two strains. We found N253D, N253H and N253S mutation which is indicates that there is polarity changed from polar to electrically charged negative, to positive and to polar as Asparagine (N) is polar, Aspartic acid (D) is electrically charged negative, Histidine (H) is electrically charged positive and Serine (S) is polar. We found V246L and V246I mutation which is indicates that there is polarity changed from non-polar to non-polar as Valine (V) is non-polar, Leucine (L) is non-polar and Isoleucine (I) is non-polar.

In case of I222T (Ile -222- Thr) reported position, we found Lysine (K) in our software (Geneious). There are three mutations here from Lysine (K) to Glutamine (Q) in one strain, from Lysine (K) to Arginine (R) in three strains and from Lysine (K) to Asparagine (N) in one strain. We found amino acid Threonine (T) in two positions (220 and 218) which is located within twenty positions before 222. We found amino acid Threonine (T) in one positions (231) which is located within twenty positions after 222.There are no mutations here. We found K222Q, K222R and K222N mutation which is indicates that there is changed electrically charged from positive (K) to positive (R) and from positive (K) to polar (Q and N).

11.2 Conclusion

Neuraminidase inhibitors (NAIs) are first-line agents for the treatment and prevention of influenza virus infections. Resistance to the NAIs can be both drug and virus type or subtype specific. The effects of NAI resistance NA mutations on the fitness and transmissibility of influenza viruses may vary depending on several factors: location of the mutation (catalytic or framework residue), NA type／subtype, virus genetic background, existence of permissive secondary NA mutations, degree of NA functional loss, and an appropriate functional NA–HA balance. In addition, differences in the host's immune response and genetic background can contribute to such variation. No single measure can easily describe the extent of fitness of influenza viruses carrying drug resistance mutations.

The most frequently reported change conferring oseltamivir resistance in that viral context is the H274Y neuraminidase mutation (N1 numbering). Recent studies have shown that, in the presence of the appropriate permissive mutations, the H274Y variant can retain virulence and transmissibility in some viral backgrounds. Most oseltamivir-resistant influenza A virus infections can be managed with the use of inhaled or intravenous zanamivir, another NAI. New NAI compounds and non-neuraminidase agents as well as combination therapies are currently in clinical evaluation for the treatment for severe influenza infections. In our Geneious we found some similar amino acids near the reported mutated positions that may cause reduce sensitivity to NAIs. In our article we analyzed phylogenetic tree and found some diverse strains those may be migrated or may be mutated which can show antigenic drift.

At present, it is clear that mutation conferring resistance to the currently approved antiviral drugs is a growing problem, so it is very important to continue research in various areas that would allow better resolution of the problem: First, the knowledge of mutation's type that allows virus avoids the action of drugs, will allow an understanding as to how these mutations arise and how we must avoid them. Also, this knowledge, will allow a better design of new drugs.

Part-Two (2)

Analysis of Neuraminidase Inhibitors Resistance to H1N1 of Influenza Virus A from 2010 and 2013

Chapter 1

Introduction

1.1 Virology

Influenza is an enveloped negative-strand RNA virus with single stranded segmented RNA genome consisting of eight segments (Lamb & Choppin 1983). The eight segments of influenza A viral RNA (vRNA) encode ten recognized gene products which are PB1, PB2, PA polymerases, HA, NP, NA, Ml and M2 proteins, and NS1 and NS2 proteins (Robert *et al.* 1992).

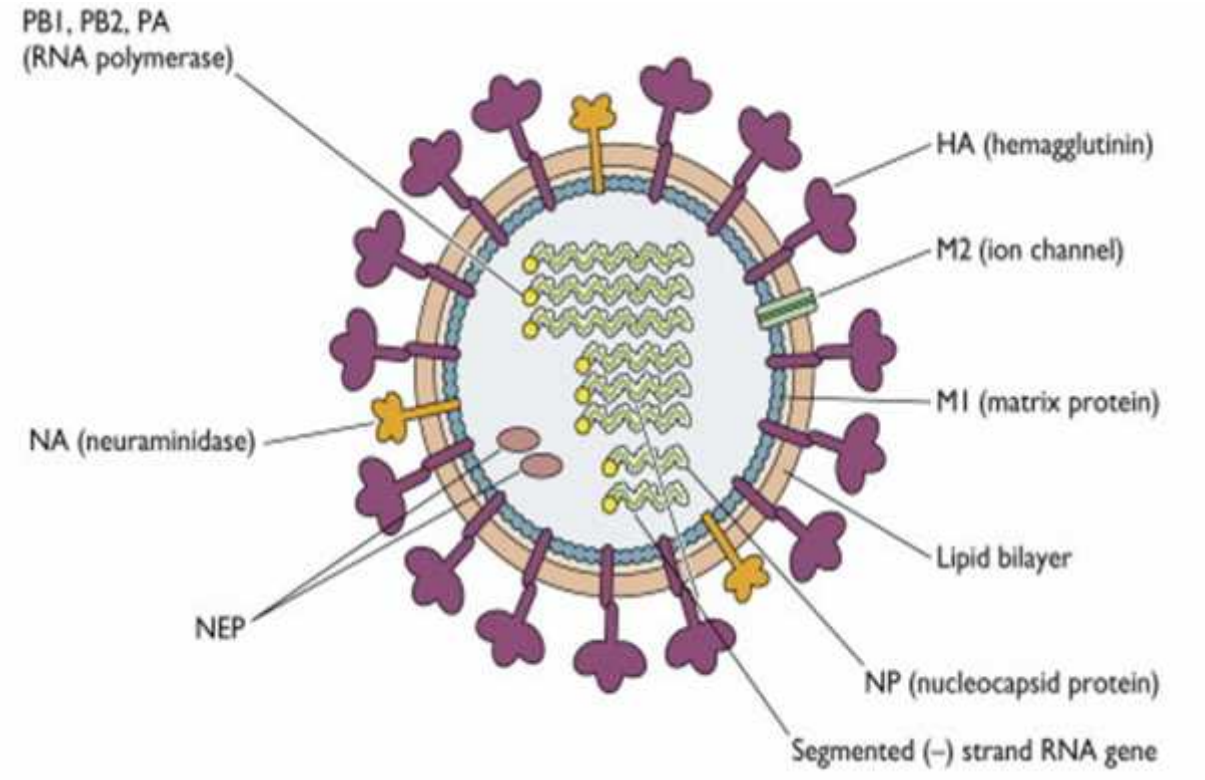


Figure: Structure of the Influenza Virus (Ludwig *et al.* 2003).

1.1.1 PB2 Polymerase

PB2 polymerase is encoded by RNA segment 1. PB2 form the trimeric vRNA dependent RNA polymerase (Engelhardt & Fodor 2006). The RNA polymerase is responsible for the transcription and replication of the viral RNA genome in the nucleus of infected cells. The PB2 protein specifically plays a role in generating 5′-capped RNA fragments from cellular pre-mRNA molecules that are used as primers for viral transcription (Engelhardt & Fodor 2006; Guilligay *et al.* 2008). Endonucleolytic cleavage of these cap structures from host mRNAs is also at least in part a function of PB2. The role of PB2 in the other virus-directed RNA synthetic processes, i.e., synthesis of full-length template cRNA and new negative-sense viral RNA (vRNA), is not known since these processes do not require host cap priming. Newly synthesized PB2 proteins migrate to the nucleus of infected cells (Robert *et al.* 1992).

1.1.2 PB1 Polymerase

PB1 polymerase is encoded by RNA segment 2. PB1 is the best characterized functionally. Biochemical and structural analyses recognize PB1 as responsible for RNA chain elongation. PB1 contains amino acid motifs common to all RNA-dependent RNA polymerases and RNA-dependent DNA polymerases (Muller *et al.* 1994). PB1 proteins localize in the nucleus of infected cells.

1.1.3 PA polymerase

PA polymerase is encoded by RNA segment 3. The PA subunit induces a generalized proteolytic process when expressed from cloned cDNA (Sanz-Ezquerro *et al.* 1995). It also localizes in the infected cell nucleus and is a member of the RNA-dependent RNA polymerase complex along with PB1 and PB2, but its role in viral RNA synthesis is unknown. There is evidence for possible roles as a protein kinase or as a helix-unwinding protein (Robert *et al.* 1992).

Avian influenza A consists of two major glycoproteins which are Hemagglutinin (HA) and Neuraminidase (NA) (Ito *et al.* 2001). There are several subtypes of HA and NA. 18 different HA subtypes (H1 to H18) and 11 different NA subtypes (N1 to N11) are found (Tong *et al.* 2013).

1.1.4 Hemagglutinin (HA)

HA is encoded by RNA segment 4. HA glycoproteins are more prone to attach to the cell surface sialic acid receptors. HA represents a crucial component of current vaccines.

HA is a trimer-forming glycoprotein (Wilson *et al.* 1981) performing two crucial functions:

- Binding of the virus to the surface receptors of the host cell through the sialic acid and
- Releasing of the viral genome into the cytoplasm during the process of membrane fusion (Huang *et al.* 1981; Lenard & Miller 1981).

HA is synthesized in endoplasmic reticulum as a single protein. To be activated, i.e., render the virus infective and fusogenic, its cleavage is necessary by the specific host cell proteases into two chains (subunits) HA1 a HA2 (Lazarowitz *et al.* 1973). There is a difference

between host surface receptors on the target cell which is believed to be the possible restrictive factor of avian influenza. HA gene of avian cell binds to Sia2-3Galactose-containing receptor which is different from human Sia2-6Galactose containing receptor (Matrosovich *et al.* 1999). Before functioning as a virus it needs post translational cleavage by host proteases (Rott 1992). HA followed by NA are important antigenic determinant from which neutralizing antibodies are directed.

1.1.5 Nucleoprotein (NP)

NP is encoded by RNA segment 5. NP binds RNA with high affinity in a sequence-independent manner (Baudin *et al.* 1994; Yamanaka *et al.* 1990), most likely through the positively charged cleft identified in the atomic structure between the head and body domains (Ye *et al.* 2006; Ng *et al.* 2008). In addition, the structure revealed the presence of a tail loop

by which NP molecules self-associate into oligomeric structures, which has been shown by mutagenesis to be required for RNP (Ribonucleoprotein) activity (Chan *et al.* 2010). NP has also been shown biochemically to interact with polymerase subunits PB1 and PB2 (Biswas *et al.* 1998), but information on the interaction domains is limited (Poole *et al.* 2004). It is generally thought that during replication the 5′ terminus of the emerging transcript is bound sequence-specifically and co-transcriptionally by “free” polymerase which then serves as a nucleation step for the sequence-independent sequential encapsidation of the transcript by NP (Ye *et al.* 2006; Tao & Ye 2010; Portela & Digard 2002).

1.1.6 Neuraminidase (NA)

NA is encoded by RNA segment 6. NA is the second major antigenic determinant for neutralizing antibodies. NA is glycosylated and possesses an amino-terminal hydrophobic sequence which functions both as signal for transport to the cell membrane and as transmembrane domain; it is not cleaved away. By catalyzing the cleavage of glycosidic linkages to sialic acid on host cell and virion surfaces, this glycoprotein prevents aggregation of virions thus facilitating the release of progeny virus from infected cells. Inhibition of this important function represents the most effective antiviral treatment strategy to date (Menno & Tran 2006). Antiviral drugs like Oseltamivir and Zanamivir (NA inhibitors) inhibit this important function are the key to the antiviral treatment.

1.1.7 M1 Protein

M1 protein is encoded by RNA segment 7. The matrix protein Ml is the most highly conserved of all the influenza A virus proteins (McCauley & Mahy 1983) and is the smallest (252 amino acids) and most abundant of the structural proteins. The close association of MI both with the lipid envelope and with the ribonucleoprotein particles, comprising nucleoprotein (NP), polymerases and RNA, in budding and mature virions (Patterson *et al.* 1988) points to a key role in viral assembly. M1 can bind RNA directly in the absence of nucleoprotein (NP) and suggest that properties previously thought to be mediated by NP:M1 interactions may be attributable to the RNA-binding activity of M1 (Larrissa & George 1989). In the infected cell it is present in both cytoplasm and nucleus.

1.1.8 NS protein

The NS1 protein is an interferon antagonist59,60 that blocks the activation of transcription factors and IFN-b-stimulated gene products, and binds to double-stranded RNA (dsRNA) to prevent the dsRNA-dependent activation of 29-59 oligo(A) synthetase, and the subsequent activation of RNase L, an important player in the innate immune response. Recently obtained structural data are expected to help in the identification of domains that are critical for the biological functions of NS1.

The NS1 proteins of H5N1 viruses confer resistance to the antiviral effects of interferon and are associated with high levels of pro-inflammatory cytokines27,28,30,65,66; the resulting cytokine imbalance probably contributes to the high mortality of H5N1 virus infections in humans. Several amino acids in NS1 have now been shown to affect virulence
65,67,68. The S-OIVs possess the low-pathogenic-type amino acid at these positions. However, available data suggest that these amino acid changes affect virulence in a strain-specific manner, whereas a multibasic HA cleavage sequence and PB2-Lys 627 seem to be universal determinants of viral pathogenicity.

1.2 Seasonal H1N1

Early results demonstrated that mutations in the NA which conferred reduced NAI sensitivity also impacted on the function of the NA such that resistant viruses were compromised in their fitness and unlikely to be transmitted (Herlocher *et al.* 2004 & Herlocher *et al.* 2002). However, in late 2007, several seasonal H1N1 viruses with an H274Y mutation were isolated in Norway. There was a minimal use of oseltamivir in Norway, and none of the patients had a

history of drug exposure. Subsequent testing revealed 183 of 272 isolates (67%) bore this mutation. This virus was clearly fit, and transmissible and within weeks resistant viruses were detected in North America, Europe, and Asia (Dharan *et al.* 2009; Matsuzaki *et al.* 2010 & Meijer *et al.* 2007-2008). The resistant viruses continued to spread to the southern hemisphere, ultimately displacing the sensitive virus. More than 90% of H1N1 isolates were resistant by 2008–2009, with IC50s in the fluorescent assay generally in the 500–1000 nM range. It appears that permissive mutations had evolved that enabled the NA to tolerate the H274Y mutation, maintaining fitness of the enzyme (Bloom *et al.* 2010; Abed *et al.* 2011; Rameix-Welti *et al.* 2008 & Collins *et al.* 2009). The substitutions include R193G, R221Q, V233M, and D343N (R194G, R222Q, V234M, and D344N in N1 numbering). While H274Y is the primary mutation seen in N1 viruses, a seasonal H1N1 virus with an I222V mutation, conferring reduced susceptibility to oseltamivir, was also detected in surveillance of community isolates from untreated patients (Monto *et al.* 2008).

1.3 Pandemic influenza A(H1N1)pdm09

The sudden emergence and spread of the swine-derived influenza virus from Mexico led to the displacement of the oseltamivir-resistant seasonal H1N1 virus by the new A(H1N1)pdm09 virus. However, given the awareness of oseltamivir resistance, much closer monitoring has been carried out, by both phenotypic testing by enzyme assay and sequencing. Viruses with the H274Y mutation have been detected in patients after treatment or prophylaxis with oseltamivir or peramivir (Inoue *et al.* 2010; Valinotto *et al.* 2009 & Renaud *et al.* 2010). and also in untreated patients. Resistance has been detected as early as 48 hours post-treatment (Inoue *et al.* 2010). One of the earliest reports of resistance was in summer campers in the USA undergoing oseltamivir prophylaxis. Resistant virus was isolated from the second patient several days after contact with the first. In addition to the H274Y mutation, both viruses had a common I222V (I223V in N1 numbering) mutation, suggesting possible human–human spread (CDC 2009). Another case was reported after prophylaxis of a family contact of an infected patient (Baz *et al.* 2009). As the dose of oseltamivir for prophylaxis is only half that for therapy (75 mg once daily), administering subtherapeutic doses when virus replication has already begun could increase the selection of resistant A(H1N1)pdm09 virus. There are numerous reports of the emergence of resistant viruses among immunocompromised patients undergoing oseltamivir treatment or prophylaxis, which is not unexpected due to the longer periods of therapy (Harvala *et al.* 2009; Tramontana *et al.* 2009;

Campanini *et al.* 2009; Memoli *et al.* 2010 & Chan *et al.* 2010). There are also reports of transmission of resistant viruses in hospitalized settings among immunocompromised patients (Chen *et al.* 2011; Moore *et al.* 2011). Of more concern is the isolation and transmission of viruses in the community among patients with no history of drug use. A traveler from the USA, identified as feverish upon entry into Hong Kong, was one of the first patients from which resistant virus was isolated with no known exposure to oseltamivir. The nasopharyngeal specimen contained a mixture of wild-type (47%) and H274Y-resistant (53%) viruses. After culture in MDCK cells, the sample contained 98% H274Y virus. Hence, the mutation did not compromise replication in vitro. H274Y-resistant virus was also detected in 7 of 10 untreated students traveling on a train in Vietnam.59 More recently, a community cluster of 29 patients infected with a H274Y virus has been identified in Australia, 28 of whom had no known drug exposure (Hurt *et al.* 2011), thus demonstrating the fitness and transmissibility of the H274Y A(H1N1)pdm09 virus. There are various reports of the assessment of fitness of resistant A(H1N1)pdm09 viruses in vitro and in animal models. Some report no compromise in fitness (Hamelin *et al.* 2010 & Seibert *et al.* 2010), while others demonstrate reduced fitness (Brookes *et al.* 2011 & Duan *et al.* 2010). In addition to the early report of an I222V (I223V in N1 numbering) mutation I222R mutations have been reported. I222R variants emerged in two immunocompromised patients, one treated sequentially with oseltamivir than zanamivir and a second treated with oseltamivir, 0ne virus had both I222R and H274Y mutations (Nguyen *et al.* 2010 & Van *et al.* 2010).

1.4 Mode of transmission

Investigations of all the most recently confirmed cases, in China, Turkey and Indonesia have identified that the most likely source of exposure is close contact with infected or dead poultry or their faeces. According to World Health Organization (WHO), almost all human sufferers caught the disease directly from birds and only tiny number of infected person has infected another. Other sources are flesh (even refrigerated/frozen carcasses), eggs from infected birds which can harbor the virus both outside and within the cell. Indirect transmission is also possible through contaminated clothing, footwear, vehicles and equipments. Contaminated feed, water, insects, rodents, cats, dogs can also act as vectors and transmit the disease (Bridges *et al*. 2002; Mounts *et al*. 1999).

1.5 Route of transmission

Route of transmission is mainly respiratory tract. Conjunctiva may be affected directly. The incubation period for H5N1 avian influenza ranges from 2 to 8 days and possibly as long as 17 days. WHO recommends that an incubation period of 7 days be used for field investigations and monitoring of patients' contacts (Bridges *et al*. 2003; Salgado *et al*. 2002; and van *et al*. 2006).

1.6 Antigenic Drift (Gradual, Small Changes)

Antigenic drift refers to small, gradual changes that occur through point mutations in the two genes that contain the genetic material to produce the main surface proteins, hemagglutinin (HA) and neuraminidase (NA). These point mutations occur unpredictably and result in minor changes to these surface proteins. Antigenic drift produces new virus strains that may not be recognized by antibodies to earlier influenza strains. This process works as follows: a person infected with a particular influenza virus strain develops antibody against that strain. As newer virus strains appear, the antibodies against the older strains might not recognize the "newer" virus and infection with a new strain can occur. (Centers for disease control and prevention, Department for disease control and prevention Avian Influenza (Bird Flu) November 18, 2005).

1.7 Antigenic Shift (Major, abrupt change)

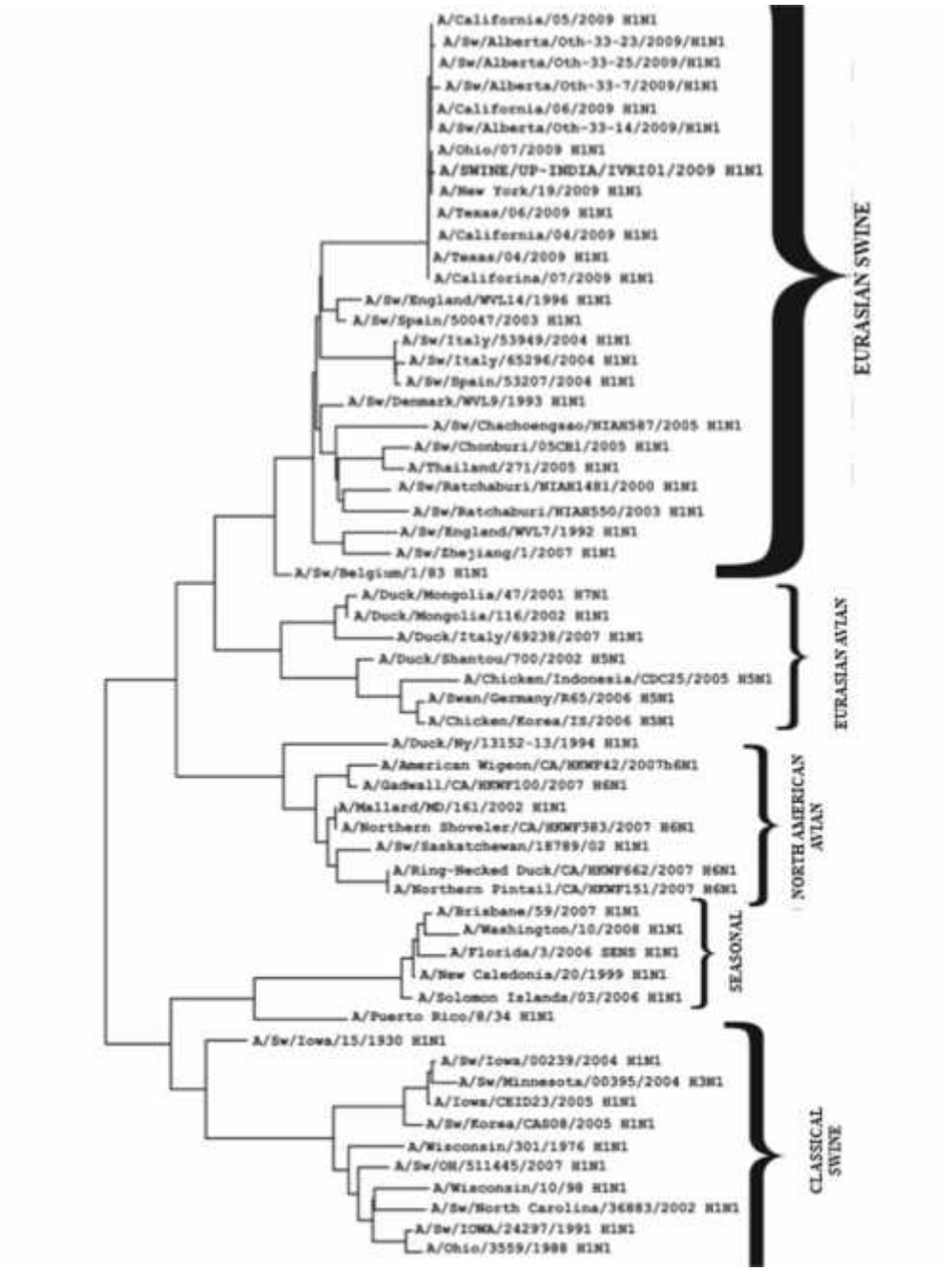
The event of their replacement on a global scale is commonly referred to as an antigenic shift in the virus population (Paul *et al.* 2013). It refers to an abrupt, major change to produce a novel influenza A virus subtype in humans that was not currently circulating among people. Antigenic shift can occur either through direct animal (poultry) to human transmission or through mixing of human influenza A and animal influenza A virus genes to create a new human influenza A subtype virus through a process called genetic reassortment. Antigenic shift results in a new human influenza A subtype. A global influenza pandemic (worldwide spread) may occur if three conditions are met: (Centers for disease control and prevention, Department for disease control and prevention Avian Influenza (Bird Flu). November 18, 2005)

- A new subtype of influenza A virus is introduced into the human population.
- The virus causes serious illness in humans.
- The virus can spread easily from person to person in a sustained manner.

Chapter 2

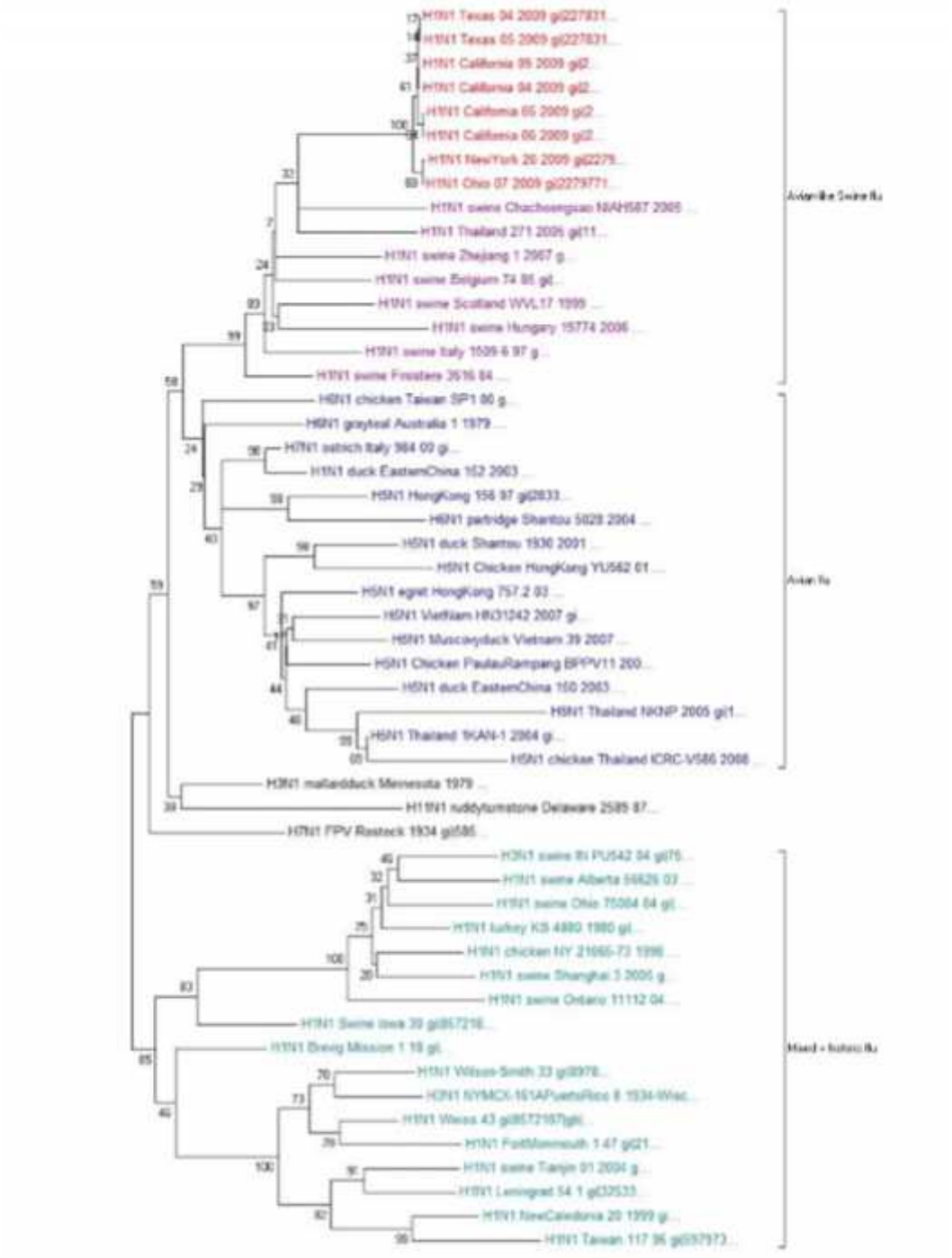
Some Reported Phylogenetic Tree of NA Protein and their Interpretation

2.1 Phylogentic tree based on nucleotide sequence of neuraminidase (NA) gene segment of influenza A viruses human, swine and avian origin. A/Sw/UP-India/IVRI01/2009 H1N1 is highlighted with bold letters (Neumann *et al.* 2009).



Phylogenetic tree of NA gene

2.2 Phylogenetic tree of neuraminidase protein sequences of the N1 subtype family.



Phylogenetic tree of NA gene

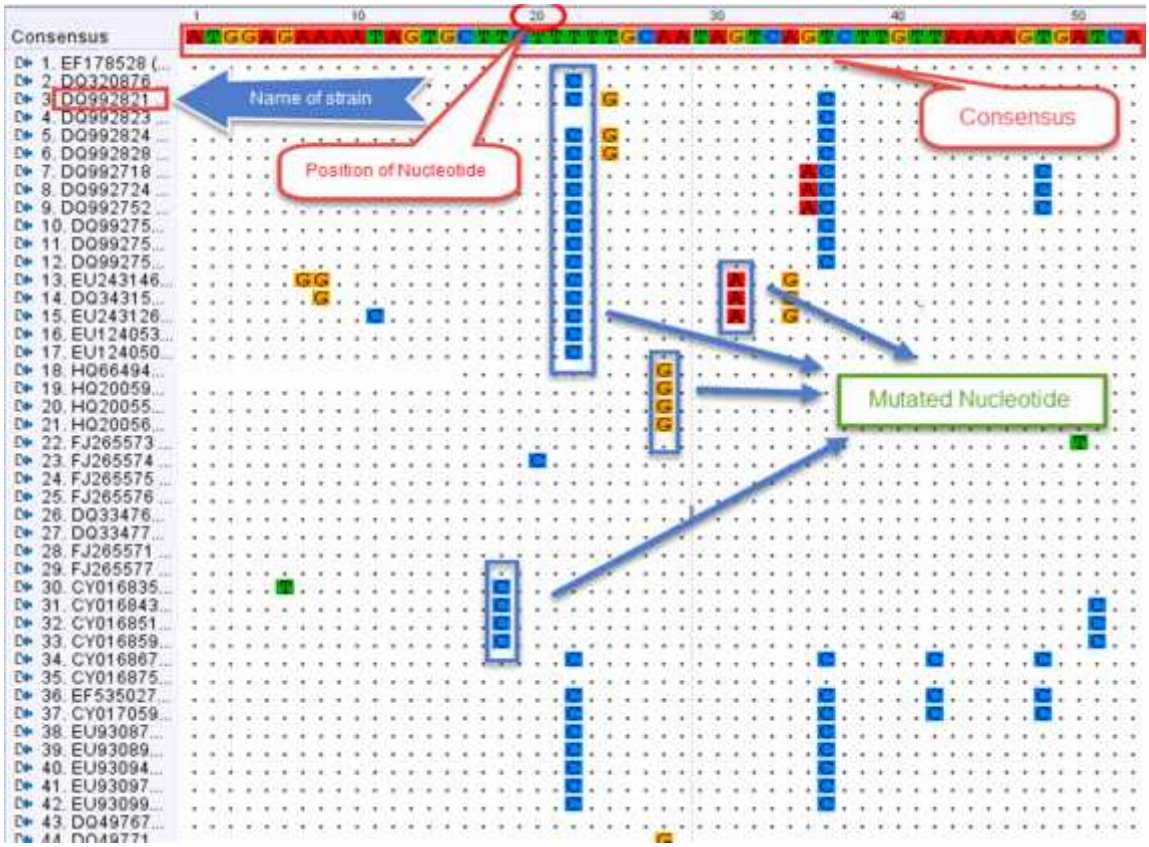
Panorama genotypic diversity of H1N1 subtype swine influenza viruses in mainland China Taken together, on the basis of the phylogenetic analyses of all eight gene segments of H1N1 SIVs isolated from pigs in mainland China from 1992 to 2011, the viruses in the present study could be divided into 10 genotypes, which were shown. The genetic composition of the genotypes of CS, EA, 2009H, RSH, and ESH series contained the same segments origin respectively. Of these genotypes CS-H3N2H, EA-H9N2A, EA-TR, and 2009H-EA and TR-EA included the distinct gene sources. The genotypes of 2009H, ESH, RSH, CS-H3N2H and 2009H-EA were directly related to human influenza viruses, including strains harbored genes derived from human influenza viruses. The eight gene segments of genotype RSH, ESH, and 2009H derived from recent human influenza viruses, early human influenza viruses, and 2009 human pandemic influenza viruses respect- ively. Genotype CS-H3N2, such as A/swine/Shandong/ 1123/2008/H1N1, its NS gene originated from human H3N2 influenza viruses, the remaining segments from the classic swine influenza viruses. NS gene of EA-TR genotype viruses belonged to the North American triple reassortant influenza viruses, with other genes derived from the Eurasian avian-like influenza viruses. Genotype EA-H9N2A represented that virus recombinant PB1 gene from avian influenza viruses, remaining genes origi- nated from the Eurasian avian-like influenza viruses. The 2009 H-EA genotype strains appeared to evolve from the Eurasian avian-like influenza viruses by recombinant HA, NA gene, and the internal genes from the 2009 human influenza viruses. HA, M, and NA gene of TR- EA genotype strains belonged to the Eurasian avian-like influenza viruses, the other genes derived from the North American triple reassortant viruses. These findings showed that novel genotypes of H1N1 SIVs were existed in mainland China, and further revealed the diversity and complexity of H1N1 SIVs in mainland China (Liu *et al.* 2009).

Chapter 3

Data Analysis & Result

3.1 Data Analysis

We have collected random strains of different year (2010 and 2013) from the search result. We will analyze a lot of strain from different year. We will make nucleotide sequence, amino acid sequence and tree from those strains. Phylogenetic tree helps to find out the diverse strain hence it helps to find out the evolution. Phylogenetic tree cannot be made without nucleotide or amino acid sequences. Diverse strain shows mutation in different nucleotide and amino acid position. From literature search we will try to find the virology of these diverse strains. In case of unavailability to find out the virology of the identified diverse strain we will find out the reported mutant position of different strain. These mutant positions can infect more avian species, reduced sensitivity to anti-viral drugs and even they may infect to human also. We will also correlate the reported data with our experimental data.



Here is one example of nucleotide sequence.

Figure: Overview of Nucleotide sequence view in geneious.

Left panel shows the name of strains. This is the selected list of strains to be analyzed. Top row is the consensus. Nucleotide dissimilarities are showing in different colors. Dotted lines are the similarities. We can also find the mutant position of different nucleotides.

We can also make the amino acid alignment by translating the nucleotide sequence. Throughout the analysis we will use amino acid sequence to find out the mutation of different species as amino acid shows the significant way to distinguish the small differences of different species.

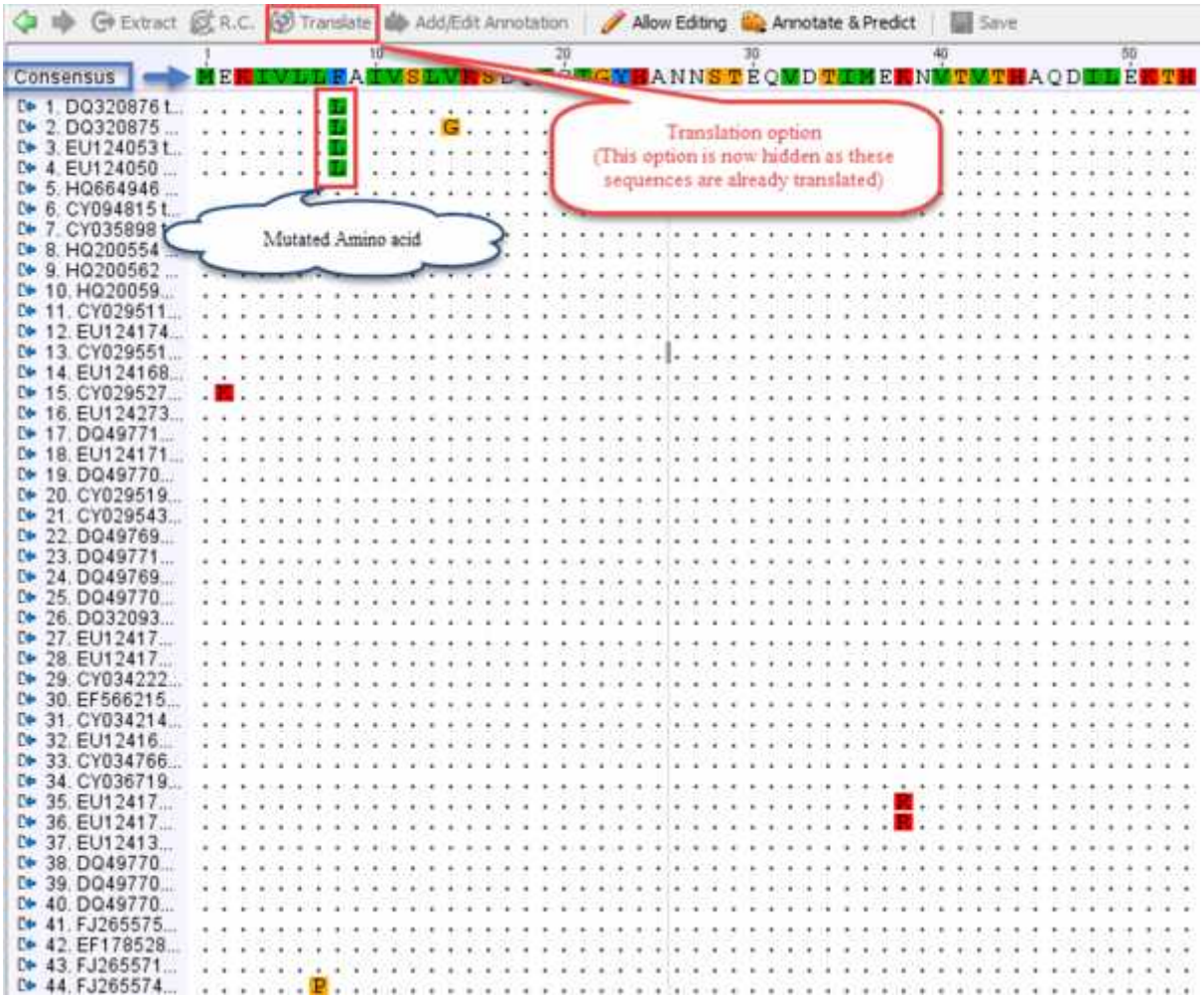


Figure: Overview of Amino Acid sequence view.

It is very easy to find out the mutation point from amino acid sequence as there is less number of amino acid sequence then nucleotide sequence. Phylogenetic tree can be made

from both nucleotide sequence and amino acid sequence alignment. We will make phylogenetic tree here with the nucleotide sequence.

In each year we have taken 150+ random strains to analyze. To ease our analysis we have shorten the description of downloaded strain in document table. These lists of abbreviations are as follows:

Abbreviations: Singapore-SP, Swine-SN, Taiwan-TN, Helsinki-HS.

In this analysis all trees are made with Neighbor joining method. The numbers which will show above the branch node represents the neighbor-joining bootstrap value generated from 1,000 replicates. Data collected form NCBI using BLAST search inside geneious. Scale bar represents the nucleotide substitutions per site.

3.2 Result

In 2010 and 2013 we have selected total 38 strains of Neuraminidase (NA) strain of H1N1.

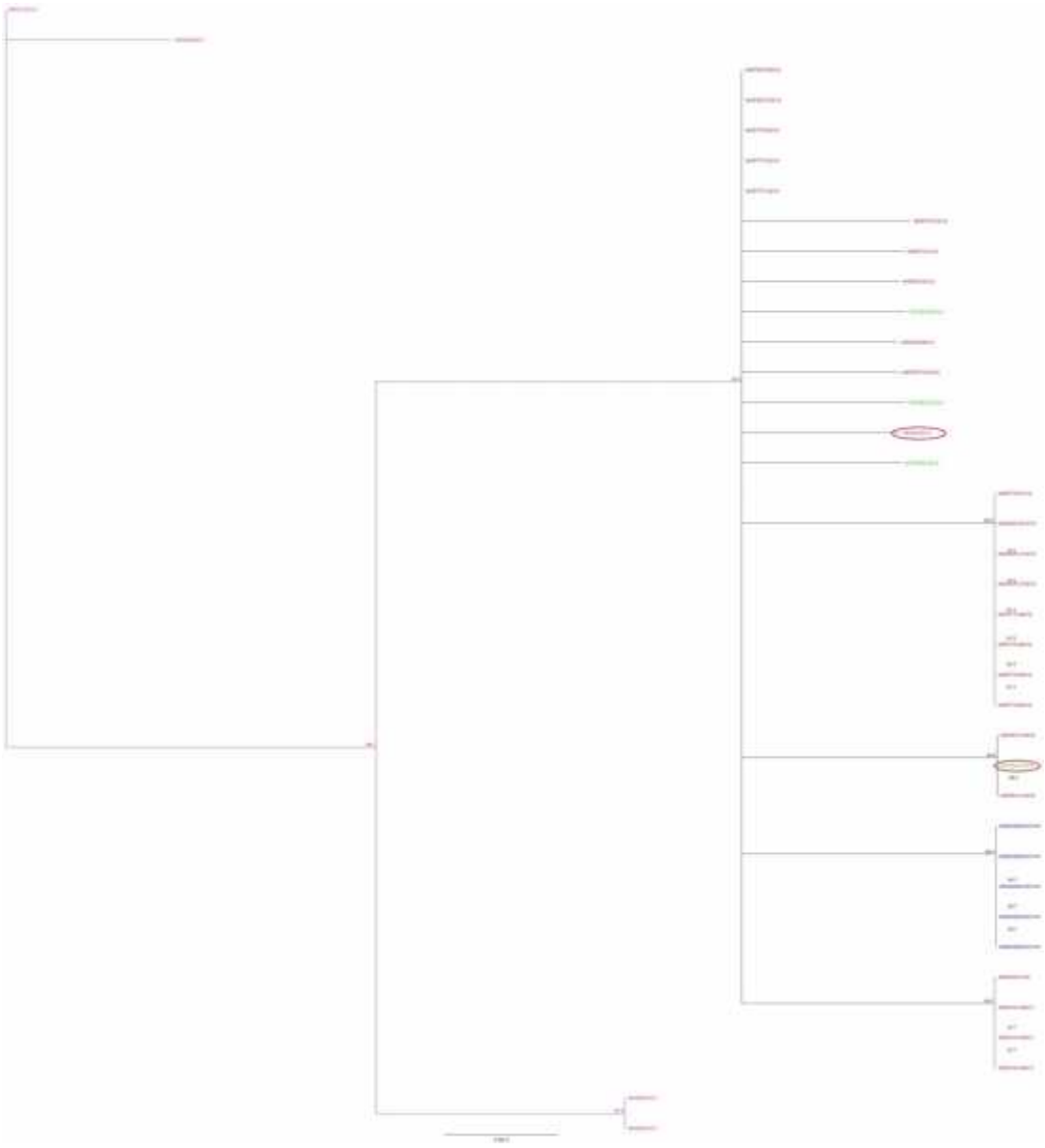


Figure: Phylogenetic tree of the year 2010 and 2013.

List of diverse strains from our analysis of 2010 and 2013 are given below.

Accession number	Strain name
JN381405	A/TN/65229/10
KF559896	A/HS/636/13

From two years we have got some strains which seems to diverse from our analysis. We did our literature search but we did not get any information about these diverse strains. So it seems to us that these strains are not responsible for antigenic shift. Neighbor joining method and bootstrap value shows that this diverse strain is showing antigenic drift which may transfer to the other avian in the same country or other different country as well through migratory process.

From the phylogenetic tree of 2010 and 2013 we came to see that all TN like viruses formed a cluster. Among them A/TN/65229/10 shows 64.5 bootstrap value with A/SP/KK140/10. Number of bootstrap value is showing the similarities between these two strains. Strain A/TN/65229/10 may be migrated or may be mutated which can show antigenic drift. A/HS/636/13 shows 86.5 bootstrap values with A/TN/65337/10 and others strains. Number of bootstrap value is showing the similarities between others strains. Strain A/HS/636/13 may be migrated or may be mutated which can show antigenic drift.

Chapter 4

Discussion & Conclusion

4.1 Discussion

The molecular mechanisms that enable avian influenza viruses to cross the species barrier and transmit efficiently in humans are incompletely understood. Some experiments have been done to identify the transmission pattern and it shows poorer transmission from infected to susceptible animals [Alexander *et al.* 1978, Westbury *et al.* 1979 & Goto *et al.* 2001]. Migration process can influence transmission of viruses. Migratory birds can carry pathogens from country to country thereby playing a role distributing influenza viruses. Avian influenza A consists of two major glycoproteins which are Hemagglutinin (HA) and Neuraminidase (NA) [Gambotto *et al.* 2008]. HA glycoproteins are more prone to attach to the cell surface sialic acid receptors. There is a difference between host surface receptors on the target cell which is believed to be the possible restrictive factor of avian influenza.

Human infections are periodic. In some cases these viruses are accompanied by high mortality. As a result they are the major concern about the potential H1N1 as an endemic virus.

The pandemic H1N1 influenza viruses also possess the N1 NA from avian sources [Cohen *et al.* 2009, Garten *et al.* 2009]. Unlike previously circulating viruses, the pandemic H1N1 viruses contained a complex of influenza virus genes of Eurasian and North American swine influenza virus origin [Smith *et al.* 2009] which were previously derived from reassorted genes of human, swine, and avian lineages. This novel reassortment of HA and NA genes resulted in a virus that was effectively transmitted in humans. In type 1 NA, the major amino acid change causing antiviral resistance is a histidine to tyrosine change at position 274 (H274Y) (Table 1). This change causes an increase of the inhibitor of OSC and PeR [Samson *et al.* 2013]. This mutation is the major OS-resistance change in influenza A/H1N1 viruses [Le *et al.* 2005].

Before functioning as a virus it needs post translational cleavage by host proteases [Tong *et al.* 2013]. HA followed by NA are important antigenic determinant from which neutralizing antibodies are directed. There are several subtypes of HA and NA. 18 different HA subtypes (H1 to H18) and 11 different NA subtypes (N1 to N11) are found [**Shinya** *et al.* 2006].

Data used in this study are obtained inside using nucleotide BLAST search from publicly available database of National Centre for Biotechnology Information (NCBI). Multiple sequence alignments, editing, assembly of strains were performed in windows platform with the Geneious program version 7.1.3 (trial).

In this study we will analyze some swine neuraminidase (H1N1) of different years. Analysis includes building nucleotide sequence and translating them into amino acid sequence. Then we will study amino acid positions with respect to some reported mutation to see the genetic pattern.

We found some avian amino acid position N295S, N294S, H275Y, H274Y, S247N, S246N, E119V and I222R are specific reported position which can be responsible for N1 (H1N1) neuraminidase mutations causing reduced sensitivity to NAIs [Pizzorno *et al*. 2011, Le QM *et al*. 2005 & Nguyen *et al*. 2012]. In our swine H1N1 analysis we did not find the exact location where reported mutations are occurred. But we found similar amino acid near the reported mutated position. We have analyzed around (before and after the mutation point) twenty positions with respect to the reported mutation point.

Twenty amino acids, their symbols and short forms are given below

Nonpolar (Hydrophobic) amino acids			**Polar (Hydrophilic) amino acids**		
Amino acid (AA)	**Short form**	**Symbol**	**Amino acid (AA)**	**Short form**	**Symbol**
Glycine	Gly	G	Serine	Ser	S
Alanine	Ala	A	Threonine	Thr	T
Valine	Val	V	Cysteine	Cys	C
Leucine	Leu	L	Tyrosine	Try	Y
Isoleucine	Ile	I	Asparagine	Asn	N
Methionine	Met	M	Glutamine	Gln	Q
Phenylalanine	Phe	F			
Tryptophan	Trp	W			
Proline	Pro	P			

Electrically charged (Negative) amino acids			Electrically charged (Positive) amino acids		
Amino acid (AA)	**Short form**	**Symbol**	**Amino acid (AA)**	**Short form**	**Symbol**
Aspartic acid	Asp	D	Lysine	Lys	K
Glutamic acid	Glu	E	Arginine	Arg	R
			Histidine	His	H

Table: Summary of some reported position which can be responsible for N1 (H1N1) neuraminidase mutations causing reduced sensitivity to NAIs.

Amino acid mutation position with reference	Short form	Amino acid in Geneious in reported position	NAIs
Asn-295-Ser (Pizzorno *et al*. 2011)	N295S	V	Oseltamivir Peramivir
Asn-294-Ser (Le QM *et al*. 2005)	N294S	C	Oseltamivir Peramivir
His-275-Try (Pizzorno *et al*. 2011)	H275Y	A	Oseltamivir Peramivir
His-274-Try (Le QM *et al*. 2005)	H274Y	N	Oseltamivir Peramivir
Ser-246-Asn (Pizzorno *et al*. 2011)	S247N	T	Oseltamivir Peramivir
Ser-246-Asn (Nguyen *et al*. 2012)	S246N	M	Oseltamivir Peramivir

Glu-119-Val (Pizzorno *et al*. 2011)	E119V	R	Oseltamivir Peramivir
Ile-222-Arg (Nguyen *et al*. 2012)	I222R	S	Oseltamivir Zanamivir

Analysis of our study data with respect to reported mutation point causing reduced sensitivity to NAIs of swine H1N1.

Table: Correlation of Reported NAIs resistance specific avian amino acid position with our experimental strain, and their mutation pattern (exact and around twenty positions): Original Amino Acid analysis.

Reported AA with mutation point	AA in Geneious in reported position	Exact AA which matches with the reported data with respect to specific mutation point (before 20 positions)	Exact AA which matches with the reported data with respect to specific mutation point (after 20 positions)
N295S	V295	N277	N299 N304 N311 N313
N294S	C294	N277	N299 N304 N311 N313
H275Y	A275		H279Y(1) Electrically charged Positive to Polar

H274Y	N274		H279Y(1) Electrically charged Positive to Polar
S247N	T247	S241 S233	S251N(1) Polar to Polar S256
S246N	M246	S241 S233	S251N(1) Polar to Polar S256
E119V	R119		E120K(1) Electrically charged Negative to Positive E129
I222R	S222	I220 I216 I215	I227T(1) Nonpolar to Polar

In case of N295S (Asn-295-Ser) reported position, we found Valine (V) in our software (Geneious). There is no mutation here. We found amino acid Asparagine (N) in one positions (277) which is located within twenty positions before 295. There is no mutation here. We found amino acid Asparagine (N) in four positions (299, 304, 311 and 313) which is located within twenty positions after 295. There are no mutations here.

In case of N294S (Asn-294-Ser) reported position, we found Cysteine (C) in our software (Geneious). There is no mutation here. We found amino acid Asparagine (N) in one positions (277) which is located within twenty positions before 294. There is no mutation here. We found amino acid Asparagine (N) in four positions (299, 304, 311 and 313) which is located within twenty positions after 294. There are no mutations here.

In case of H275Y (His-275-Try) reported position, we found Alanine (A) in our software (Geneious). There is no mutation here. We found amino acid Histidine (H) in one positions (279) which is located within twenty positions after 275. There is one mutation here from Histidine (H) to Tyrosine (Y) in one strain. We found H279Y mutation which is indicate that there is polarity changed from electrically charged Positive to polar as Histidine (H) is electrically charged Positive and Tyrosine (Y) is polar.

In case of H274Y (His-274-Try) reported position, we found Asparagine (N) in our software (Geneious). There is no mutation here. We found amino acid Histidine (H) in one positions (279) which is located within twenty positions after 274. There is one mutation here from Histidine (H) to Tyrosine (Y) in one strain. We found H279Y mutation which is indicate that there is polarity changed from electrically charged Positive to polar as Histidine (H) is electrically charged Positive and Tyrosine (Y) is polar.

In case of S247N (Ser-247-Asn) reported position, we found Threonine (T) in our software (Geneious). There is no mutation here. We found amino acid Serine (S) in two positions (233 and 241) which is located within twenty positions before 247. There are no mutations here. We found amino acid Serine (S) in two positions (251 and 256) which is located within twenty positions after 247.There is one mutation here from Serine (S) to Asparagine (N) in one strain. We found S251N mutation which is indicates that there is polarity changed from polar to polar as Serine (S) is polar and Asparagine (N) is polar.

In case of S246N (Ser-246-Asn) reported position, we found Methionine (M) in our software (Geneious). There is no mutation here. We found amino acid Serine (S) in two positions (233 and 241) which is located within twenty positions before 246. There are no mutations here. We found amino acid Serine (S) in two positions (251 and 256) which is located within twenty positions after 246.There is one mutation here from Serine (S) to Asparagine (N) in one strain. We found S251N mutation which is indicates that there is polarity changed from polar to polar as Serine (S) is polar and Asparagine (N) is polar.

In case of E119V (Glu-119-Val) reported position, we found Arginine (R) in our software (Geneious). There is no mutation here. We found amino acid Glutamic acid (E) in two positions (120 and 129) which is located within twenty positions after 119. There is one mutation here from Glutamic acid (E) to Lysine (K) in one strain. We found E120K mutation

which is indicates that there is polarity changed from electrically charged negative to Positive as Glutamic acid (E) is electrically charged negative and Lysine (K) is electrically charged Positive.

In case of I222R (Ile-222-Arg) reported position, we found Serine (S) in our software (Geneious). There is no mutation here. We found amino acid Isoleucine (I) in three positions (225, 216 and 220) which is located within twenty positions before 222. There are no mutations here. We found amino acid Isoleucine (I) in one position (227) which is located within twenty positions after 222. There is one mutation here from Isoleucine (I) to Threonine (T) in one strain. We found I227T mutation which is indicates that there is polarity changed from non-polar to polar as Isoleucine (I) is non-polar and Threonine (T) is polar.

Table: Correlation of Reported NAIs resistance specific avian amino acid position with our experimental strain, and their mutation pattern (exact and around twenty positions): Mutated Amino Acid analysis.

Reported AA with mutation point	AA in Geneious in reported position	Exact AA which matches with the reported data with respect to specific mutation point (before 20 positions)	Exact AA which matches with the reported data with respect to specific mutation point (after 20 positions)
N295S	V295	S290N(2) Polar to Polar S290G(1) Polar to Nonpolar S289 S284	S303 S309
N294S	C294	S290N(2) Polar to Polar S290G(1) Polar to Nonpolar S289 S284	S303 S309
H275Y	A275	Y257	Y278 Y280 Y286

H274Y	N274	Y257	Y278 Y280 Y286
S247N	T247	N239	
S246N	M246	N239 N226S(3) Polar to Polar	
E119V	R119	V117I(3) Nonpolar to Nonpolar V115	
I222R	S222		R224 R229

In case of N295S (Asn-295-Ser) reported position, we found Valine (V) in our software (Geneious). There is no mutation here. We found amino acid Serine (S) in three positions (284, 289 and 290) which is located within twenty positions before 295. There are two mutations here from Serine (S) to Asparagine (N) in two strains and from Serine (S) to Glycine (G) in one strain. We found S290N and S290G mutations which are indicate that there is polarity changed from polar to polar and from polar to non-polar as Serine (S), Asparagine (N) is polar and Glycine (G) is non-polar. We found amino acid Serine (S) in two positions (303 and 309) which is located within twenty positions after 295. There are no mutations here.

In case of N294S (Asn-294-Ser) reported position, we found Cysteine (C) in our software (Geneious). There is no mutation here. We found amino acid Serine (S) in three positions (284, 289 and 290) which is located within twenty positions before 294. There are two mutations here from Serine (S) to Asparagine (N) in two strains and from Serine (S) to Glycine (G) in one strain. We found S290N and S290G mutations which are indicate that there is polarity changed from polar to polar and from polar to non-polar as Serine (S), Asparagine (N) is polar and Glycine (G) is non-polar. We found amino acid Serine (S) in two positions (303 and 309) which is located within twenty positions after 294. There are no mutations here.

In case of H275Y (His-275-Try) reported position, we found Alanine (A) in our software (Geneious). There is no mutation here. We found amino acid Tyrosine (Y) in one positions (257) which is located within twenty positions before 275. There is no mutation here. We found amino acid Tyrosine (Y) in three positions (278, 280 and 286) which are located within twenty positions after 275. There are no mutations here.

In case of H274Y (His-274-Try) reported position, we found Alanine (A) in our software (Geneious). There is no mutation here. We found amino acid Tyrosine (Y) in one positions (257) which is located within twenty positions before 274. There is no mutation here. We found amino acid Tyrosine (Y) in three positions (278, 280 and 286) which are located within twenty positions after 274. There are no mutations here.

In case of S247N (Ser-247-Asn) reported position, we found Threonine (T) in our software (Geneious). There is no mutation here. We found amino acid Asparagine (N) in one position (239) which is located within twenty positions before 247. There are no mutations here.

In case of S246N (Ser-246-Asn) reported position, we found Threonine (T) in our software (Geneious). There is no mutation here. We found amino acid Asparagine (N) in one position (239) which is located within twenty positions before 246. There are no mutations here.

In case of E119V (Glu-119-Val) reported position, we found Arginine (R) in our software (Geneious). There is no mutation here. We found amino acid Valine (V) in two positions (115 and 117) which is located within twenty positions before 119. There is one mutation here from Valine (V) to Isoleucine (I) in three strains. We found V117I mutation which is indicates that there is polarity changed from non-polar to polar as Valine (V) is non-polar and Isoleucine (I) is polar.

In case of I222R (Ile-222-Arg) reported position, we found Serine (S) in our software (Geneious). There is no mutation here. We found amino acid Arginine (R) in two positions (224 and 229) which is located within twenty positions after 222. There are no mutations here.

4.2 Conclusion

Resistance to the NAIs can be both drug and virus type or subtype specific. A summary of the effects of mutations referred to in this review is presented in Table 1. In the N1 subtype, the most frequently encountered mutation is the H275Y that confers highly reduced inhibition to oseltamivir, moderate cross resistance to the investigational agent peramivir and susceptibility to zanamivir. Various amino acid changes at residue 222 (I ?R) or 223 (I ?R/V) can also confer reduced inhibition to oseltamivir and/or to zanamivir. In our Geneious we found some similar amino acids near the reported mutated positions that may cause reduce sensitivity to NAIs.

The global outbreak of the human pandemic influenza A virus (H1N1) has caused a significant alertness in the general public [Dawood *et al.* 2009]. Analysis of the genetic material allows a prediction of the new influenza virus characteristics including its capacity to spread the infection, degree of virulence, and drug resistance. In our article we analyzed phylogenetic tree and found some diverse strains those may be migrated or may be mutated which can show antigenic drift.

At present, it is clear that mutation conferring resistance to the currently approved antiviral drugs is a growing problem, so it is very important to continue research in various areas that would allow better resolution of the problem: First, the knowledge of mutation's type that allows virus avoids the action of drugs, will allow an understanding as to how these mutations arise and how we must avoid them. Also, this knowledge, will allow a better design of new drugs.

Chapter 12

References

Part One (1)

12.1 References

Aggarwal S, Dewhurst S, Takimoto T, Kim B (2011): Biochemical impact of the host adaptation-associated PB2 E627K mutation on the temperature-dependent RNA synthesis kinetics of influenza A virus polymerase complex. Journal of Biological Chemistry 286, 34504–34513.

Abed Y, Goyette N, Boivin G. Generation and characterization of recombinant influenza A (H1N1) viruses harboring amantadine resistance mutations. Antimicrob Agents Chemother 2005; 49:556– 559.

Abed Y, Baz M, Boivin G. Impact of neuraminidase mutations conferring influenza resistance to neuraminidase inhibitors in the N1 and N2 genetic backgrounds. Antivir Ther 2006; 11:971–976.

Baldauf, S. L. (2003) Phylogeny for the faint of heart: a tutorial. *Trends in Genetics.* 19 (6). p.345-351.

Bender C, Hall H, Huang J, Klimov A, Cox N, Hay A, et al. Characterization of the surface proteins of influenza A (H5N1) viruses isolated from humans in 1997–1998. Virology. 1999; 254:115–23.

Baum, L.G.; Paulson, J.C. The N2 neuraminidase of human influenza virus has acquired a substrate specificity complementary to the hemagglutinin receptor specificity. *Virology* 1991, *180*, 10–15.

Bantia S, Arnold CS, Parker CD, Upshaw R, Chand P (2006): Anti-influenza virus activity of peramivir in mice with single intramuscular injection. Antiviral Research 69, 39–45.

Baudin, F., Bach, C., Cusack, S., & Ruigrok, R. W. (1994) Structure of influenza virus RNP. I. Influenza virus nucleoprotein melts secondary structure in panhandle RNA and exposes the bases to the solvent. *Embo J.* 13. p.3158–65.

Besselaar TG, Naidoo D, Buys A, Gregory V, McAnerney J, et al. (2008) Widespread oseltamivir resistance in influenza A viruses (H1N1), South Africa. Emerg Infect Dis 14: 1809- 1810.

Beigel JH, Farrar J, Han AM, Hayden FG, Hyer R, et al. (2005) Avian influenza A (H5N1) infection in humans. N Engl J Med 353: 1374-1385.

Bridges, C.B., W. Lim, J. Hu Primmer, L. Sims, K. Fukuda, K.H. Mak, T. Rowe, W.W. Thompson, L. Conn, X. Lu, N.J. Cox and J.M. Katz, 2002. Risk of influenza A (H5N1) infection among poultry workers, Hong Kong, 1997-1998. J. Infect. Dis., 185: 1005-1010.

Belshe, R.B., 2005. Avian influenza A (H5N1) infection in humans. New England Journal of Medicine, 353: 1374-1385.

Birney, E., Thompson, J. D. & Gibson, T. J. (1996) Pair Wise and Search Wise: finding the optimal alignment in a simultaneous comparison of a protein profile against all DNA translation frames. *Nucleic Acids Research*. 24. p.2730–2739.

Biswas, S. K., Boutz, P. L. & Nayak, D. P. (1998) Influenza virus nucleoprotein interacts with influenza virus polymerase proteins. *J Virol.* 72. p.5493–501.

Bridges, C.B., M.J. Kuehnert, C.B. Hall, 2003. Transmission of influenza: implications for control in health care settings. Clin. Infect. Dis., 37: 1094-1101.

Beigel J, Bray M (2008): Current and future antiviral therapy of severe seasonal and avian influenza. Antiviral Research 78, 91–102.

Bright RA, Medina MJ, Xu XY, Perez-Oronoz G, Wallis TR, Davis XHM, Povinelli L, Cox NJ, Klimov AI (2005): Incidence of adamantane resistance among influenza A (H3N2) viruses isolated worldwide from 1994 to 2005: a cause for concern. Lancet 366, 1175–1181.

Barr IG, Deng YM, Iannello P, Hurt AC, Komadina N (2008): Adamantane resistance in influenza A(H1) viruses increased in 2007 in South East Asia but decreased in Australia and some other countries. Antiviral Research 80, 200–205.

Baz M, Abed Y, Papenburg J, Bouhy X, Hamelin ME, et al. (2009) Emergence of oseltamivirresistant pandemic H1N1 virus during prophylaxis. N Engl J Med 361: 2296-2297.

Baz M, Abed Y, McDonald J, Boivin G. Characterization of multidrugresistant influenza A/H3N2 viruses shed during 1 year by an immunocompromised child. Clin Infect Dis 2006; 43:1555–1561.

Boivin S, Cusack S, Ruigrok RWH, Hart DJ (2010): Influenza A virus polymerase: structural insights into replication and host adaptation mechanisms. Journal of Biological Chemistry 285, 28411–28417.

Boivin, S.; Hart, D.J. Interaction of the influenza A virus polymerase PB2 C-terminal region with importin alpha isoforms provides insights into host adaptation and polymerase assembly. *J. Biol. Chem.* 2011, *286*, 10439–10448.

Cao B, Wang DY, Yu XM, We LQ, Pu ZH, Gao Y, Wang J, Dong JP, Li XL, Xu Q, Hu K, Chen BY, Yu YS, Song SF, Shu YL, Wang C (2012): An uncontrolled openlabel, multicenter study to monitor the antiviral activity and safety of inhaled zanamivir (as Rotadisk via Diskhaler device) among Chinese adolescents and adults with influenza-like illness. Chinese Medical Journal 125, 3002–3007.

Carr J, Ives J, Kelly L, Lambkin R, Oxford J, et al. (2002) Influenza virus carrying neuraminidase with reduced sensitivity to oseltamivir carboxylate has altered properties in vitro and is compromised for infectivity and replicative ability in vivo. Antiviral Res 54: 79-88.

Claas, C.J., A.D.M. Osterhaus, R. Beek, J. 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, 351: 472-477.

Claas EC, Osterhaus AD, van Beek R et al. Human influenza A H5N1 virus related to a highly pathogenic avian influenza virus. Lancet
1998; 351:472–477.

Chen, L. M. et al. In vitro evolution of H5N1 avian influenza virus toward humantype-receptor specificity. Virology 422, 105–113 (2012).

Chan, W.-H., Ng, A. K.-L., Robb, N. C., Lam, M. K.-H., Chan, P. K.-S., Au, S. W.-N., Wang, J.-H., Fodor, E. & Shaw, P.-C. (2010) Functional analysis of the influenza virus H5N1 nucleoprotein tail loop reveals amino acids that are crucial for oligomerization and ribonucleoprotein activities. *J Virol*. 84. p.7337–45.

Chen G-W, Chang S-C, Mok C-K, Lo Y-L, Kung Y-N, Huang J-H, et al. Genomic signatures of human versus avian infl uenza A viruses. Emerg Infect Dis. 2006;12:1353–60.

Chen H, Smith GJD, Li KS et al. Establishment of multiple sublineages of H5N1 influenza virus in Asia: implications for pandemic control. Proc Natl Acad Sci USA 2006;103:2845–2850.

Chutinimitkul, S.; van Riel, D.; Munster, V.J.; van den Brand, J.M.; Rimmelzwaan, G.F.; Kuiken, T.; Osterhaus, A.D.; Fouchier, R.A.; de Wit, E. *In vitro* assessment of attachment pattern and replication efficiency of H5N1 influenza A viruses with altered receptor specificity. *J. Virol.* 2010, *84*, 6825–6833.

Conenello GM, Zamarin D, Perrone LA, Tumpey T, Palese P. A single mutation in the PB1-F2 of H5N1 (HK/97) and 1918 infl uenza A viruses contributes to increased virulence. PLoS Pathog. 2007;3:1414–21. DOI: 10.1371/journal.ppat.0030141.

Charles, D. M. & Robert, R. S. (1957) A quantitative approach to a problem in classification. *Evolution.* 11. p.130-162.

Collins, P. J., L. F. Haire, Y. P. Lin, J. Liu, R. J. Russell, P. A. Walker, J. J. Skehel, S. R. Martin, A. J. Hay, and S. J. Gamblin. 2008. Crystal structures of oseltamivir-resistant influenza virus neuraminidase mutants. Nature 453: 1258–1261.

Centers for Disease C, Prevention (2009) Oseltamivir-resistant 2009 pandemic influenza A (H1N1) virus infection in two summer campers receiving prophylaxis--North Carolina, 2009. MMWR Morb Mortal Wkly Rep 58: 969-972.

Collins PJ, Haire LF, Lin YP, Liu J, Russell RJ, et al. (2008) Crystal structures of oseltamivirresistant influenza virus neuraminidase mutants. Nature 453: 1258-1261.

Davies WL, Hoffmann CE, Paulshock M, Wood TR, Haff RF, Grunert RR, Watts JC, Hermann EC, Neumayer EM, Mcgahen JW (1964): Antiviral activity of 1-adamantanamine (amantadine). Science 144, 862–863.

Dapat C, Suzuki Y, Saito R et al. Rare influenza A (H3N2) variants with reduced sensitivity to antiviral drugs. Emerg Infect Dis 2010; 16:493–496.

De Jong MD, Tran TT, Truong HK, Vo MH, Smith GJ, et al. (2005) Oseltamivir resistance during treatment of influenza A (H5N1) infection. N Engl J Med 353: 2667-2672.

De Jong MD, Van Cam B, Qui PT, Hien VM, Thanh TT, et al. (2005) Brief report: Fatal avian influenza A (H5N1) in a child presenting with diarrhea followed by coma. New England Journal of Medicine 352: 686-691.

De Jong M, Thanh TT, Khanh TH et al. Oseltamivir resistance during treatment of influenza A (H5N1) infection. N Engl J Med 2005; 353:2667–2672.

Deng, T.; Engelhardt, O.G.; Thomas, B.; Akoulitchev, A.V.; Brownlee, G.G.; Fodor, E. Role of ran binding protein 5 in nuclear import and assembly of the influenza virus RNA polymerase complex. *J. Virol.* 2006, *80*, 11911–11919.

Dorigo, M. & Caro, G. Di. (1999) The Ant Colony Optimization Meta-Heuristic, in New Ideas in Optimization. Corne, D., Dorigo, M., and Glover, F., Eds. London: McGraw-Hill. p.11-32.

Durbin, R., Eddy, S., Krogh, A., & Mitchison, G. (1999) Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids. *Reprint ed. Cambridge: Cambridge*

University Press. Earhart KC, Elsayed NM, Saad MD et al. Oseltamivir resistance mutation N294S in human influenza A(H5N1) virus in Egypt. J Infect Public Health 2009; 2:74–80.

Egan, A. N. & Crandall, K. A. (2006) Theory of Phylogenetic Estimation," in Evolutionary Genetics: Concepts and Case Studies. 1 ed Oxford University Press, USA. p. 426-436.

Engelhardt, O. G. & Fodor, E. (2006) Functional association between viral and cellular transcription during influenza virus infection. *Rev. Med. Virol.* 16. p.329–345.

Feng F, Miura N, Isoda N, Sakoda Y, Okamatsu M, Kida H, Nishimura SI (2010): Novel trivalent anti-influenza reagent. Bioorganic and Medicinal Chemistry Letters 20, 3772–3776.

Ferraris, O. & Lina, B. (2008). Mutations of neuraminidase implicated in neuraminidase inhibitors resistance. *Journal Clinical Virology*. Vol. 41, No. 1, (January 2008), pp.13–19.

Furuta Y, Takahashi K, Kuno-Maekawa M, Sangawa H, Uehara S, Kozaki K, Nomura N, Egawa H, Shiraki K (2005): Mechanism of action of T-705 against influenza virus. Antimicrobial Agents and Chemotherapy 49, 981–986.

Felsenstein, J. (1985) Confidence intervals on phylogenies: An approach using the bootstrap. *Evolution.* 39. p.783–791.

Fodor, E.; Smith, M. The PA subunit is required for efficient nuclear accumulation of the PB1 subunit of the influenza A virus RNA polymerase complex. *J. Virol.* 2004, *78*, 9144–9153.

Gabriel, G.; Dauber, B.; Wolff, T.; Planz, O.; Klenk, H.D.; Stech, J. The viral polymerase mediates adaptation of an avian influenza virus to a mammalian host. *Proc. Natl. Acad. Sci. USA* 2005, *102*, 18590–18595.

Gambaryan, A.; Webster, R.; Matrosovich, M. Differences between influenza virus receptors on target cells of duck and chicken. *Arch. Virol.* 2002, *147,* 1197–1208.

Guan Y, Poon LL, Cheung CY, Ellis TM, Lim W, Lipatov AS, et al. H5N1 influenza: a protean pandemic threat. Proc Natl Acad Sci U S A. 2004; 101:8156–61.

Guilligay, D., Tarendeau, F., Resa-Infante, P., Coloma, R., Crepin, T., Sehr, P., Lewis, J., Ruigrok, R. W., Ortin, J., Hart, D. J. & Cusack, S. (2008) The structural basis for cap binding by influenza virus polymerase subunit PB2. *Nat. Struct. Mol. Biol.* 15. p.500–506.

Gamblin, S.J., L.F. Haire and R.J. Russell, 2004. The structure and receptor binding properties of the 1918 influenza hemagglutinin. Science, 303: 1838-42.

Galvao MGA, Santos MARC, da Cunha AJLA (2012): Amantadine and rimantadine for influenza A in children and the elderly.Cochrane Database of Systematic Reviews, vol.23.

Garcia J, Sovero M, Torres AL, Gomez J, Douce R, Barrantes M, Sanchez F, Jimenez M, Comach G, de Rivera I, Agudo R, Kochel T (2009): Antiviral resistance in influenza viruses circulating in Central and South America based on the detection of established genetic markers. Influenza and Other Respiratory Viruses 3, 69–74.

Gubareva LV, Matrosovich MN, Brenner MK, Bethell RC, Webster RG. Evidence for zanamivir resistance in an immunocompromised child infected with influenza B virus. J Infect Dis 1998; 178:1257– 1262.

Govorkova EA, Ilyushina NA, Boltz DA, Douglas A, Yilmaz N, Webster RG. Efficacy of oseltamivir therapy in ferrets inoculated with different clades of H5N1 influenza virus. Antimicrob Agents Chemother 2007; 51:1414–1424.

Galvao MGA, Santos MARC, da Cunha AJLA (2012): Amantadine and rimantadine for influenza A in children and the elderly. Cochrane Database of Systematic Reviews,vol.23.

Greengard O, Poltoratskaia N, Leikina E, Zimmerberg J, Moscona A (2000): The anti-influenza virus ahent 4-GU-DANA (Zanamivir) inhibits cell fusion mediated by human parainfluenza virus and influenza virus HA. Journal of Virology 74, 11108–11114.

Hartwell, L. H., Hood, L., Goldberg, M. L., Reynolds, A. E., Silver, L. M. & Veres, R. C. (2008) *Genetics: From Genes to Genomes, 3rd Ed.* McGraw-Hill: New York.

Hennig, W. (1966) "Phylogenetic Systematics." University of Illinois Press, Urbana.

Hall, B. G. (2007) Phylogenetic Trees Made Easy: A How-to Manual, 3 ed Sinauer Associates, Inc. Pevsner, J. (2003) Molecular Phylogeny and Evolution. 1 ed John Wiley and Sons, Inc. p. 368-388.

Hastings JC, Selnick H, Wolanski B, Tomassini JE (1996): Anti-influenza virus activities of 4-substituted 2,4-dioxobutanoic acid inhibitors. Antimicrobial Agents and Chemotherapy 40, 1304–1307.

He G, Massarella J, Ward P (1999): Clinical pharmacokinetics of the prodrug oseltamivir and its active metabolite Ro 64-0802. Clinical Pharmacokinetics 37, 471–484.

Hillis, D. M., & Bull, J. J. (1993) An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. *Syst. Biol.* 42. p.182–192.

Hatta, M.; Hatta, Y.; Kim, J.H.; Watanabe, S.; Shinya, K.; Nguyen, T.; Lien, P.S.; Le, Q.M.; Kawaoka, Y. Growth of H5N1 influenza A viruses in the upper respiratory tracts of mice. *PLoS Pathog.* 2007, *3*, 1374–1379.

Hauge SH, Dudman S, Borgen K, Lackenby A, Hungnes O (2009) Oseltamivir-resistant influenza viruses A (H1N1), Norway, 2007-08. Emerg Infect Dis 15: 155-162.

Hurt AC, Holien JK, Parker M, Kelso A, Barr IG (2009) Zanamivir-resistant influenza viruses with a novel neuraminidase mutation. J Virol 83: 10366-10373.

Hurt AC, Lee RT, Leang SK, Cui L, Deng YM, et al. (2011) Increased detection in Australia and Singapore of a novel influenza A(H1N1)2009 variant with reduced oseltamivir and zanamivir sensitivity due to a S247N neuraminidase mutation. Euro Surveill 16.

Huang, R. T. C., Rott, R. & Klenk, H. D. (1981) Influenza viruses cause hemolysis and fusion of cells. *Virology. 110 (1).* p.243-247.

Hurt AC, Hardie K, Wilson NJ, Deng YM, Osbourn M, et al. (2012) Characteristics of a Widespread Community Cluster of H275Y Oseltamivir-Resistant A(H1N1)pdm09 Influenza in Australia. J Infect Dis 206: 148-157.

Hurt AC, Holien JK, Barr IG (2009) In vitro generation of neuraminidase inhibitor resistance in A(H5N1) influenza viruses. Antimicrob Agents Chemother 53: 4433-4440.

Hurt AC, Holien JK, Parker M, Kelso A, Barr IG (2009) Zanamivir-resistant influenza viruses with a novel neuraminidase mutation. J Virol 83: 10366-10373.

Hayden FG, Belshe RB, Clover RD, Hay AJ, Oakes MG, Soo W. Emergence and apparent transmission of rimantadine-resistant influenza A virus in families. N Engl J Med 1989; 321:1696–1702.

Hurt AC, Holien JK, Parker M, Kelso A, Barr IG. Zanamivir-resistant influenza viruses with a novel neuraminidase mutation. J Virol 2009; 83:10366–10373.

Hale, B.G.; Randall, R.E.; Ortin, J.; Jackson, D. The multifunctional NS1 protein of influenza A viruses. *J. Gen. Virol.* 2008, *89*, 2359–2376.

Ito, T., Goto, H., Yamamoto, E., Tanaka, H., Takeuchi, M., Kuwayama, M., Kawaoka, Y. & Otsuki, K (2001) Generation of a highly pathogenic avian influenza A virus from an avirulent field isolate by passaging in chickens. *J Virol.* 75. p.4439-4443.

Ives JA, Carr JA, Mendel DB, Tai CY, Lambkin R, et al. (2002) The H274Y mutation in the influenza A/H1N1 neuraminidase active site following oseltamivir phosphate treatment leave virus severely compromised both in vitro and in vivo. Antiviral Res 55: 307-317.

Jeffrey, R. & Eric, C. R. (2007) Review of Phylogenetic Tree Construction. Bioinformatics Review. *University of Louisville Bioinformatics Laboratory Technical Report Series.*
Efron, B. (1979) Bootstrapping methods: Another look at the jackknife. *Ann. Stat.* 7. p. 1–26.

Jeong HJ, Ryu YB, Park SJ, Kim JH, Kwon HJ, Kim JH, Park KH, Rho MC, Lee WS (2009): Neuraminidase inhibitory activities of flavonols isolated from Rhodiola rosea roots and their in vitro anti-influenza viral activities. Bioorganic and Medicinal Chemistry 17, 6816–6823.

Kiso M, Ozawa M, Le MT et al. Effect of an asparagine-to-serine mutation at position 294 in neuraminidase on the pathogenicity of highly pathogenic H5N1 influenza A virus. J Virol 2011; 85:4667–4672.

Kim CU, Lew W, Williams MA, et al. Influenza neuraminidase inhibitors possessing a novel hydrophobic interaction in the enzyme active site: design, synthesis, and structural analysis of carbocyclic sialic acid analogues with potent anti-influenza activity. J Am Chem Soc 1997;119:681-90.

Kobasa, D.; Kodihalli, S.; Luo, M.; Castrucci, M.R.; Donatelli, I.; Suzuki, Y.; Suzuki, T.; Kawaoka, Y. Amino acid residues contributing to the substrate specificity of the influenza A virus neuraminidase. *J. Virol.* 1999, *73*, 6743–6751.

Kubo S, Tomozawa T, Kakuta M, Tokumitsu A, Yamashita M (2010): Laninamivir prodrug CS-8958, a long-acting neuraminidase inhibitor, shows superior anti-influenza virus activity after a single administration. Antimicrobial Agents and Chemotherapy 54, 1256–1264.

Kaminski MM, Ohnemus A, Staeheli P, Rubbenstroth D (2012) Pandemic 2009 H1N1 influenza A virus carrying a Q136K mutation in the neuraminidase gene is resistant to zanamivir but exhibits reduced fitness in the guinea pig transmission model. J Virol.

Kiryu, H., Tabei, Y., Kin, T. & Asai, K. (2007) Murlet: a practical multiple alignment tool for structural RNA sequences. *Bioinformatics*. 23. p.1588–1589.

Le QM, Wertheim HF, Tran ND, van Doorn HR, Nguyen TH, et al. A community cluster of oseltamivir-resistant cases of 2009 H1N1 influenza. N Engl J Med 362: 86-87.

Le QM, Kiso M, Someya K, Sakai YT, Nguyen TH, et al. (2005) Avian flu - Isolation of drugresistant H5N1 virus. Nature 437: 1108-1108.

Ludwig S, Planz O, Pleschka S, Wolff T (2003): Influenzavirus- induced signaling cascades: targets for antiviral therapy? Trends in Molecular Medicine 9, 46–52.

Le, L., Lee, EH., Hardy, DJ., Truong, TN. & Schulten K. (2010) Molecular dynamics simulations suggest that electrostatic funnel directs binding of Tamiflu to influenza N1 neuraminidases. PLoS Computational Biology. Vol. 6, No. 9, (September 2010), pp. e1000939.

Lenard, J. & Miller, D. K. (1981) P^H-dependent hemolysis by influenza, Semliki Forest virus, and Sendai virus. *Virology*. *110.* p.479-482.

Lazarowitz, S. G., Golberg, A. R. & Choppin, P. W. (1973) Proteolytic cleavage by plasmin of the HA polypeptide of influenza virus: Host cell activation of serum plasminogen. *Virology*. *56.* p.172-180.

Le QM, Kiso M, Someya K et al. Avian flu: isolation of drug-resistant H5N1 virus. Nature 2005; 437:1108.

Leipe, D. D., Wainright, P. O., Gunderson, J. H., Porter, D., Patterson, D. J., Valois, F., Himmerich, S., & Sogin, M. L. (1994) The Stramenopiles from a molecular perspective: 16S-like rRNA sequences from *Labyrinthuloides minutum* and *Cafeteria roenbergensis. Phycologia.* 33. p.369–377.

Lee KK, Pessi A, Gui L, Santoprete A, Talekar A, Moscona A, Porotto M (2011): Capturing a fusion intermediate of influenza hemagglutinin with a cholesterol- conjugated peptide, a new antiviral strategy for influenza virus. Journal of Biological Chemistry 286, 42141–42149.

Lv HM, Guo XL, Gu RX, Wei DQ (2011): Free energy calculations and binding analysis of two potential antiinfluenza drugs with polymerase basic protein-2 (PB2). Protein and Peptide Letters 18, 1002–1009.

Larson AM, Wang HM, Cao Y, Jiang TJ, Chen JZ, Klibanov A (2012): Conjugating drug candidates to polymeric chains does not necessarily enhance anti-influenza activity. Journal of Pharmaceutical Sciences 101, 3896–3905.

Tado M, Abe T, Hatta T, Ishikawa M, Nakada S, Yokota T, Takaku H (2001): Inhibitory effect of modified 5'-capped short RNA fragments on influenza virus RNA polymerase gene expression. Antiviral Chemistry and Chemotherapy 12, 353–358.

Lamb, R. A. & Choppin, P. W. (1983) The gene structure and replication of influenza virus. *Ann Rev Biochem*. 52. p.467-506.

László (1999) *Protein Evolution.* Blackwell Science, Ltd: Malden, MA. Laura E.: http://education-portal.com/academy/lesson/cladograms-and-phylogenic-trees-evolution-classifiations.html#lesson.

Luo GX, Torri A, Harte WE, Danetz S, Cianci C, Tiley L, Day S, Mullaney D, Yu KL, Ouellet C, Dextraze P, Meanwell N, Colonno R, Krystal M (1997): Molecular mechanism underlying the action of a novel fusion inhibitor of influenza A virus. Journal of Virology 71, 4062–4070.

Labadie, K.; Dos Santos Afonso, E.; Rameix-Welti, M.A.; van der Werf, S.; Naffakh, N. Hostrange determinants on the PB2 protein of influenza A viruses control the interaction between the viral polymerase and nucleoprotein in human cells. *Virology* 2007, *362*, 271–282.

Larrissa, W. & George, G. B. (1989) RNA-binding properties of influenza A virus matrix protein Ml. *Nucleic Acids Research.* 17(21). p.8569 – 8580.

Lv HM, Guo XL, Gu RX, Wei DQ (2011): Free energy calculations and binding analysis of two potential antiinfluenza drugs with polymerase basic protein-2 (PB2). Protein and Peptide Letters 18, 1002–1009.

Le QM, Kiso M, Someya K, Sakai YT, Nguyen TH, et al. (2005) Avian flu: isolation of drugresistant H5N1 virus. Nature 437: 1108.

Liu KC, Fang JM, Jan JT, Cheng TJR, Wang SY, Yang ST, Cheng YSE, Wong CH (2012): Enhanced anti-influenza agents conjugated with anti-inflammatory activity. Journal of Medicinal Chemistry 55, 8493–8501.

Mitamura, Y. Kiso, M., K. Sakai-Tagawa, K. Shiraishi, C. Kawakami, K. Kimura, F. G. Hayden, N. Sugaya, and Y. Kawaoka. 2004. Resistant influenza A viruses in children treated with oseltamivir: descriptive study. Lancet 364:759–765.

McKimm-Breschkin, J. L. 2005. Management of influenza virus infections with neuraminidase inhibitors: detection, incidence, and implications of drug resistance. Treat. Respir. Med. 4(2):107–116.

Meijer A, Lackenby A, Hungnes O, Lina B, van-der-Werf S, et al. (2009) Oseltamivir-resistant influenza virus A (H1N1), Europe, 2007-08 Season. Emerg Infect Dis 15: 552-560.

Matrosovich MN, Matrosovich TY, Gray T, Roberts NA, Klenk HD (2004): Neuraminidase is important for the initiation of influenza virus infection in human airway epithelium. Journal of Virology 78, 12665–12667.

Matrosovich MN, Gambaryan AS, Teneberg S, Piskarev VE, Yamnikova SS, Lvov DK, et al. Avian infl uenza A viruses differ from human viruses by recognition of sialyloligosaccharides and gangliosides and by a higher conservation of the HA receptor-binding site. Virology. 1997;233:224–34. DOI: 10.1006/viro.1997.8580.

Matrosovich M, Zhou N, Kawaoka Y, Webster R. The surface glycoproteins of H5 influenza viruses isolated from humans, chickens, and wild aquatic birds have distinguishable properties. J Virol. 1999;73:1146–55.

Muller, R., Poch, O., Delarue, M., Bishop, D. H. & Bouloy, M. (1994) Rift Valley fever virus L segment: correction of the sequence and possible functional role of newly identified regions conserved in RNA-dependent polymerases. *J. Gen. Virol.* 75. p.1345–1352.

Matrosovich, M., Zhou, N., Kawaoka, Y. & Webster, R. (1999) The Surface Glycoproteins of H5 Influenza Viruses Isolated from Humans, Chickens, and Wild Aquatic Birds Have Distinguishable Properties. *Journal of Virology.* 73 (2). p.1146–1155.

Menno, D. D. J. & Tran, T. H. (2006) Review - Avian influenza A (H5N1). *Journal of Clinical Virology.* 35. p.2–13.

McCauley, J. W. & Mahy, B. W. (1983) Structure and function of the influenza virus genome. *Biochem. J.* 211. p.281-294.

Matrosovich, M.N.; Matrosovich, T.Y.; Gray, T.; Roberts, N.A.; Klenk, H.D. Human and avian influenza viruses target different cell types in cultures of human airway epithelium. *Proc. Natl. Acad. Sci. USA* 2004, *101*, 4620–4624.

Matrosovich, M.; Tuzikov, A.; Bovin, N.; Gambaryan, A.; Klimov, A.; Castrucci, M.R.; Donatelli, I.; Kawaoka, Y. Early alterations of the receptor-binding properties of H1, H2, and H3 avian influenza virus hemagglutinins after their introduction into mammals. *J. Virol.*2000, *74*, 8502–8512.

Matrosovich, M.N.; Krauss, S.; Webster, R.G. H9N2 influenza A viruses from poultry in Asia have human virus-like receptor specificity. *Virology* 2001, *281*, 156–162.

Melen, K.; Fagerlund, R.; Franke, J.; Kohler, M.; Kinnunen, L.; Julkunen, I. Importin alpha nuclear localization signal binding sites for STAT1, STAT2, and influenza A virus nucleoprotein. *J. Biol. Chem.* 2003, *278*, 28193–28200.

Manz, B.; Brunotte, L.; Reuther, P.; Schwemmle, M. Adaptive mutations in NEP compensate for defective H5N1 RNA replication in cultured human cells. *Nat. Commun.* 2012, *3*, 802.

Malakhov MP, Aschenbrenner LM, Smee DF, Wandersee MK, Sidwell RW, Gubareva LV, Mishin VP, Hayden FG, Kim DH, Ing A, Campbell ER, Yu M, Fang F (2006): Sialidase fusion protein as a novel broad-spectrum inhibitor of influenza virus infection. Antimicrobial Agents and Chemotherapy 50, 1470–1479.

Moss RB, Hansen C, Sanders RL, Hawley S, Li TJ, Steigbigel RT (2012): A phase II study of DAS181, a novel host directed antiviral for the treatment of influenza infection. Journal of Infectious Diseases 206, 1844–1851.

Mounts, A.W., H. Kwong, H.S. Izurieta, Y. Ho, T. Au, M. Lee, C. Buxton Bridges, S.W. Williams, K.H. Mak, J.M. Katz, W.W. Thompson, N.J. Cox and K. Fukuda, 1999. Case-control study of risk factors for avian influenza A (H5N1) disease, Hong Kong, 1997. J. Infect. Dis., 180: 505-508.

McKimm-Breschkin JL, Rootes C, Mohr PG, Barrett S, Streltsov VA. In vitro passaging of a pandemic H1N1/09 virus selects for viruses with neuraminidase mutations conferring high-level resistance to oseltamivir and peramivir, but not to zanamivir. J Antimicrob Chemother 2012; 67:1874–1883.

Miotto O, Heiny AT, Tan TW, August JT, Brusic V (2008): Identification of human-to-human transmissibility factors in PB2 proteins of influenza A by large-scale mutual information analysis. Bmc Bioinformatics 9 (Suppl.1),S18.

Maines, T. R. et al. Effect of receptor binding domainmutations on receptor binding andtransmissibility ofavianinfluenzaH5N1viruses. Virology413, 139–147(2011).

Mukaigawa, J.; Nayak, D.P. Two signals mediate nuclear localization of influenza virus (A/WSN/33) polymerase basic protein 2. *J. Virol.* 1991, *65*, 245–253.

Mishin VP, Hayden FG, Gubareva LV. Susceptibilities of antiviralresistant influenza viruses to novel neuraminidase inhibitors. Antimicrob Agents Chemother 2005; 49:4515–4520.

Matrosovich MN, Matrosovich TY, Gray T, Roberts NA, Klenk HD. Neuraminidase is important for the initiation of influenza virus infection in human airway epithelium. J Virol 2004;78:12665-7.

Nguyen HT, Fry AM, Loveless PA, Klimov AI, Gubareva LV (2010) Recovery of a multidrugresistant strain of pandemic influenza A 2009 (H1N1) virus carrying a dual H275Y/I223R mutation from a child after prolonged treatment with oseltamivir. Clin Infect Dis 51: 983- 984.

Ng, A. K.-L., Zhang, H., Tan, K., Li, Z., Liu, J.-H,.Chan, P. K.-S., Li, S.-M., Chan, W.-Y., Au, S. W.-N., Joachiniak, A., Walz, T., Wang, J.-H. & Shaw, P.-C. (2008) Structure of the influenza virus A H5N1 nucleoprotein: implications for RNA binding, oligomerization, and vaccine design. *FASEB J.* 22 (10). p.3638–47.

Neal Nathanson (ed). (2007). *Viral pathogenesis and immunity.* ISBN 13:978-0-12-369464-5 Academic Press. London UK.

Narla SN, Sun XL (2012): Immobilized sialyloligo-macroligand nd its protein binding specificity. Biomacromolecules 13, 1675–1682.

Nguyen HT, Fry AM, Loveless PA, Klimov AI, Gubareva LV (2010) Recovery of a multidrugresistant strain of pandemic influenza A 2009 (H1N1) virus carrying a dual H275Y/I223R mutation from a child after prolonged treatment with oseltamivir. Clin Infect Dis 51: 983- 984.

Naffakh, N.; Massin, P.; Escriou, N.; Crescenzo-Chaigne, B.; van der Werf, S. Genetic analysis of the compatibility between polymerase proteins from human and avian strains of influenza A viruses. *J. Gen. Virol.* 2000, *81*, 1283–1291.

Ozawa, M.; Fujii, K.; Muramoto, Y.; Yamada, S.; Yamayoshi, S.; Takada, A.; Goto, H.; Horimoto, T.; Kawaoka, Y. Contributions of two nuclear localization signals of influenza A virus nucleoprotein to viral replication. *J. Virol.* 2007, *81*, 30–41.

Pizzorno A, Bouhy X, Abed Y, Boivin G. Generation and characterization of recombinant pandemic influenza A(H1N1) viruses resistant to neuraminidase inhibitors. J Infect Dis 2011; 203:25–31.

Pizzorno A, Abed Y, Bouhy X et al. Impact of mutations at residue I223 of the neuraminidase protein on the resistance profile, replication level, and virulence of the 2009 pandemic influenza virus. Antimicrob Agents Chemother 2012; 56:1208–1214.

Pizzorno A, Bouhy X, Abed Y, Boivin G (2011) Generation and characterization of recombinant pandemic influenza A(H1N1) viruses resistant to neuraminidase inhibitors. J Infect Dis 203: 25-31.

Pizzorno A, Bouhy X, Abed Y, Boivin G (2011) Generation and characterization of recombinant pandemic influenza A(H1N1) viruses resistant to neuraminidase inhibitors. J Infect Dis 203: 25-31.

Paul, S. W., Michi, S., Mario, R. & Sunetra, G. (2013) The antigenic evolution of influenza: drift or thrift? *Phil Trans R Soc B.* 368: 20120200.

Philippe, L., Marco, S. & Anne-Mieke, V. (2009) The Phylogenetic Handbook. A Practical Approach to Phylogenetic Analysis and Hypothesis Testing. Second Edition. Cambridge University Press.

Patterson, S., Gross, J. & Oxford, J. S. (1988) The Intracellular Distribution of Influenza Virus Matrix Protein and Nucleoprotein in Infected Cells and Their Relationship to Hemagglutinin in the Plasma Membrane. *J. Gen. Virol.* 69. p.1859-1872.

Portela, A. & Digard, P. (2002) The influenza virus nucleoprotein: a multifunctional RNA-binding protein pivotal to virus replication. *J Gen Virol.* 83. p.723–34.

Poole, E., Elton, D., Medcalf, L. & Digard, P. (2004)Functional domains of the influenza A virus PB2 protein: identification of NP- and PB1-binding sites. *Virology.* 321. p.120–33.

Robert, G., Webster, william J., Bean, Owen, T., Gorman, Thomas, M., Chambers, T. & Yoshihiro, K. (1992) Evolution and Ecology of Influenza A Viruses. *Microbiological Reviews.* 56(1). p.152-179.

Rott, R. (1992) The pathogenic determinant of influenza virus. *Veterinary Microbiology.* 33. p.303-310.

Russell RJ, Haire LF, Stevens DJ, Collins PJ, Lin YP, et al. (2006) The structure of H5N1 avian influenza neuraminidase suggests new opportunities for drug design. Nature 443: 45-49.

Rosenberg MR, Casarotto MG (2010): Coexistence of two adamantane binding sites in the influenza A M2 ion channel. Proceedings of the National Academy of Sciences of the United States of America 107, 13866–13871.

Smith, B. J., J. L. Kimm-Breshkin, M. McDonald, R. T. Fernley, J. N. Varghese, and P. M. Colman. 2002. Structural studies of the resistance of influenza virus neuraminidase to inhibitors. J. Med. Chem. 45:2207–2212.

Sneath, P. H. A. & Sokal, R. R. (1973) Numerical Taxonomy. The Principles and. Practice of Numerical Classification.

Sandra, L. B. (2003) Phylogeny for the faint of heart: a tutorial. *TRENDS in Genetics.* 19.

Sanz-Ezquerro, J. J., Luna, S. de la, Ortín, J. & Nieto, A. (1995) Individual expression of influenza virus PA protein induces degradation of coexpressed proteins. *J. Virol.* 69. p.2420–2426.

Subbarao EK, London W, Murphy BR. A single amino acid in the PB2 gene of infl uenza A virus is a determinant of host range. J Virol. 1993;67:1761–4.

Stephen, A. S., Joseph, W., Brown, Cody, E. & Hinchliff (2013) Analyzing and Synthesizing Phylogenies Using Tree Alignment Graphs. *PLOS Computational Biology.* 9(9). e1003223.

Saitou, N. & Nei, M. (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4 (4). p.406-425.

Samson M, Pizzorno A, Abed Y, Boivin G (2013) Influenza virus resistance to neuraminidase inhibitors. Antiviral Res 98: 174-185.

Schweiger, B., Zadow, I. & Heckler, R. (2002) Antigenic drift and variability of influenza viruses.*Med MicrobiolImmunol.* 191(3–4). p.133–138.

Stray, S. J. & Pittman, L. B. (2012) Subtype- and antigenic site-specific differences in biophysical influences on evolution of influenza virus hemagglutinin. *Virol J.* 9. p.91.

Staschke KA, Hatch SD, Tang JC, Hornback WJ, Munroe JE, Colacino JM, Muesing MA (1998): Inhibition of influenza virus hemagglutinin-mediated membrane fusion by a compound related to podocarpic acid. Virology 248, 264–274.

Schnell JR, Chou JJ (2008): Structure and mechanism of the M2 proton channel of influenza A virus. Nature 451, 591–595.

Sanz-Ezquerro, J. J., Luna, S. de la, Ortín, J. & Nieto, A. (1995) Individual expression of influenza virus PA protein induces degradation of coexpressed proteins. *J. Virol.* 69. p.2420–2426.

Subbarao, E.K.; London, W.; Murphy, B.R. A single amino acid in the PB2 gene of influenza A virus is a determinant of host range. *J. Virol.* 1993, *67*, 1761–1764.

Steel, J.; Lowen, A.C.; Mubareka, S.; Palese, P. Transmission of influenza virus in a mammalian host is increased by PB2 amino acids 627K or 627E/701N. *PLoS Pathog.* 2009, *5*, e1000252.

Spesock, A.; Malur, M.; Hossain, M.J.; Chen, L.M.; Njaa, B.L.; Davis, C.T.; Lipatov, A.S.; York, I.A.; Krug, R.M.; Donis, R.O. The virulence of 1997 H5N1 influenza viruses in the mouse model is increased by correcting a defect in their NS1 proteins. *J. Virol.* 2011, *85*, 7048–7058.

Salgado, C.D., B.M. Farr, K.K. Hall, F.G. Hayden, 2002. Influenza in the acute hospital setting. Lancet Infect. Dis., 2: 145-155.

Skehel JJ, Wiley DC (2000): Receptor binding and membrane fusion in virus entry: The influenza hemagglutinin. Annual Review of Biochemistry 69, 531–569.

Sigal GB, Mammen M, Dahmann G, Whitesides GM (1996): Polyacrylamides bearing pendant alpha-sialoside groups strongly inhibit agglutination of erythrocytes by influenza virus: The strong inhibition reflects enhanced binding through cooperative polyvalent interactions. Journal of the American Chemical Society 118, 3789–3800.

Sidwell RW, Bailey KW, Wong MH, Barnard DL, Smee DF (2005): In vitro and in vivo influenza virus-inhibitory effects of viramidine. Antiviral Research 68, 10–17.

Suzuki T, Takahashi T, Guo CT, Hidari KIPJ, Miyamoto D, Goto H, Kawaoka Y, Suzuki Y (2005): Sialidase activity of influenza A virus in an endocytic pathway enhances viral replication. Journal of Virology 79, 11705–11715.

Sheu TG, Deyde VM, Okomo-Adhiambo M et al. Surveillance for neuraminidase inhibitor resistance among human influenza A and B viruses circulating worldwide from 2004 to 2008. Antimicrob Agents Chemother 2008; 52:3284–3292.

Subbarao K, Klimov A, Katz J et al. Characterization of an avian influenza A (H5N1) virus isolated from a child with a fatal respiratory illness. Science 1998; 279:393–396.

Shaw, M., L. Cooper, X. Xu, W. Thompson, S. Krauss, Y. Guan, N. Zhou, A. Klimov, N. Cox, R. Webster, W. Lim, K. Shortridge and K. Subbarao, 2002. Molecular changes associated with the transmission of avian influenza A H5N1 and H9N2 viruses to humans. J. Med. Virol., 66: 107-114.

Tisdale M, Ellis M, Klumpp K, Court S, Ford M (1995): Inhibition of influenza-virus transcription by 2'-deoxy- 2'-fluoroguanosine. Antimicrobial Agents and Chemotherapy 39, 2454–2458.

Twu, K.Y.; Kuo, R.L.; Marklund, J.; Krug, R.M. The H5N1 influenza virus NS genes selected after 1998 enhance virus replication in mammalian cells. *J. Virol.* 2007, *81*, 8112–8121.

Tarendeau, F.; Boudet, J.; Guilligay, D.; Mas, P.J.; Bougault, C.M.; Boulo, S.; Baudin, F.; Ruigrok, R.W.; Daigle, N.; Ellenberg, J.; *et al.* Structure and nuclear import function of the C-terminal domain of influenza virus polymerase PB2 subunit. *Nat. Struct. Mol. Biol.* 2007, *14*, 229–233.

Taubenberger, J.K., A.H. Reid and R.M. Lourens, 2005. Characterization of the 1918 influenza virus polymerase genes. Nature, 437: 889-93.

Tomassini JE, Davies ME, Hastings JC, Lingham R, Mojena M, Raghoobar SL, Singh SB, Tkacz JS, Goetz MA (1996): A novel antiviral agent which inhibits the endonuclease of influenza viruses. Antimicrobial Agents and Chemotherapy 40, 1189–1193.

Tao, Y. J. & Ye, Q. (2010) RNA virus replication complexes. *PLoS Pathogens.* 6. e1000943.

Tong, S., Zhu, X., Li, Y., Shi, M., Zhang, J., Bourgeois, M., Yang, H., Chen, X., Recuenco, S., Gomez, J., Chen, L. M., Johnson, A., Tao, Y., Dreyfus, C., Yu, W., McBride, R., Carney, P. J., Gilbert, A. T., Chang, J., Guo, Z., Davis, C. T., Paulson, J. C., Stevens, J., Rupprecht, C. E., Holmes, E. C., Wilson, I. A. & Donis, R. O. (2013) New World Bats Harbor Diverse Influenza A Viruses. *PLoS Pathogens*. 9 (10). e1003657.

Van der Vries E, Stelma FF, Boucher CA. Emergence of a multidrugresistant pandemic influenza A (H1N1) virus. N Engl J Med 2010; 363:1381–1382.

Van der Vries E, Schutten M, Boucher CA (2011) The potential for multidrug-resistant influenza. Current Opinion Infectious Diseases 24: 599-604.

Van der Vries E, Collins PJ, Vachieri SG, Xiong X, Liu J, et al. (2012) H1N1 2009 pandemic influenza virus: resistance of the I223R neuraminidase mutant explained by kinetic and structural analysis. PLoS pathogens 8: e1002914.

Van der Vries E, van den Berg B, Schutten M (2008) Fatal oseltamivir-resistant influenza virus infection. New England Journal of Medicine 359: 1074-1076.

Van der Vries E, Stelma FF, Boucher CA (2010) Emergence of a multidrug-resistant pandemic influenza A (H1N1) virus. New England Journal of Medicine 363: 1381-1382.

Wu, W.W.; Sun, Y.H.; Pante, N. Nuclear import of influenza A viral ribonucleoprotein complexes is mediated by two nuclear localization sequences on viral nucleoprotein. *Virol. J.* 2007, *4*, 49.

Wu, W.W.; Weaver, L.L.; Pante, N. Ultrastructural analysis of the nuclear localization sequences on influenza A ribonucleoprotein complexes. *J. Mol. Biol.* 2007, *374*, 910–916.

Wu, W.W.; Pante, N. The directionality of the nuclear transport of the influenza A genome is driven by selective exposure of nuclear localization sequences on nucleoprotein. *Virol. J.* 2009, *6*, 68.

Wagner, R.; Matrosovich, M.; Klenk, H.D. Functional balance between haemagglutinin and neuraminidase in influenza virus infections. *Rev. Med. Virol.* 2002, *12*, 159–166.

Whitley RJ, Hayden FG, Reisinger KS et al. Oral oseltamivir treatment of influenza in children. Pediatr Infect Dis J 2001; 20:127–133.

World Health Organization. Cumulative number of confirmed human cases of avian influenza A ⁄ (H5N1) reported to WHO, 2009. Available at http://www.who.int/csr/disease/avian_influenza/country/ cases_index.html (Accessed 10 August 2012).

Wong, S.S.Y. and K.Y. Yuen, 2006. Avian Influenza Virus Infections in Humans. Chest, 129: 156-168.

Watson KG, Cameron R, Fenton RJ, Gower D, Hamilton S, Jin B, Krippner GY, Luttick A, McConnell D, MacDonald SJF, Mason AM, Nguyen V, Tucker SP, Wu WY (2004): Highly potent and long-acting trimeric and tetrameric inhibitors of influenza virus neuraminidase. Bioorganic and Medicinal Chemistry Letters 14, 1589–1592.

Wilson, I. A., Skehel, J. J. & Wiley, D. C. (1981) Structure of the hemagglutinin membrane glycoprotein of influenza virus at 3 Å resolution. *Nature*. *289*. p.366-373.

Yen HL, Ilyushina NA, Salomon R, Hoffmann E, Webster RG, Govorkova EA. Neuraminidase inhibitor-resistant recombinant A ⁄ Vietnam ⁄ 1203 ⁄ 04 (H5N1) influenza viruses retain their replication efficiency and pathogenicity in vitro and in vivo. J Virol 2007; 81:12418–12426.

Yamanaka, K., Ishihama, A. & Nagata, K. (1990) Reconstitution of influenza virus RNA-nucleoprotein complexes structurally resembling native viral ribonucleoprotein cores. *J Biol Chem.* 265. p.11151–5.

Yen HL, Hoffmann E, Taylor G, Scholtissek C, Monto AS, Webster RG, Govorkova EA (2006): Importance of neuraminidase active-site residues to the neuraminidase inhibitor resistance of influenza viruses. Journal of Virology 80, 8787–8795.

Ye, Q., Krug, R. M. & Tao, Y. J. (2006) The mechanism by which influenza A virus nucleoprotein forms oligomers and binds RNA. *Nature.* 444. p.1078–82.

Zambon, M. C. (1999) Epidemiology and pathogenesis of influenza. *J Antimicrob Chemother*. 44 Suppl. B: 3-9.

Zuckerkandl, E. & Pauling, L. (1965) Evolutionary divergence and convergence in proteins. In Evolving Genes and Proteins (Bryson, V. and Vogel, H.J., eds) Academic Press. p. 97–166.

Part Two (2)

5.1 References

Abed Y, Pizzorno A, Bouhy X, Boivin G. Role of Permissive Neuraminidase Mutations in Influenza A/Brisbane/59/2007-like (H1N1) Viruses. PLoS Pathog 2011; 7:e1002431.

Baz M, Abed Y, Papenburg J et al. Emergence of oseltamivir-resistant pandemic H1N1 virus during prophylaxis. N Engl J Med 2009; 361:2296–2297.

Bloom JD, Gong LI, Baltimore D. Permissive secondary mutations enable the evolution of influenza oseltamivir resistance. Science 2010; 328:1272–1275.

Brookes DW, Miah S, Lackenby A, Hartgroves L, Barclay WS. Pandemic H1N1 2009 influenza virus with the H275Y oseltamivir resistance neuraminidase mutation shows a small compromise in enzyme activity and viral fitness. J Antimicrob Chemother 2011; 66:466–470.

Bridges, C.B., M.J. Kuehnert, C.B. Hall, 2003. Transmission of influenza: implications for control in health care settings. Clin. Infect. Dis., 37: 1094-1101.

Bridges, C.B., W. Lim, J. Hu Primmer, L. Sims, K. Fukuda, K.H. Mak, T. Rowe, W.W. Thompson, L. Conn, X. Lu, N.J. Cox and J.M. Katz, 2002. Risk of influenza A (H5N1) infection among poultry workers, Hong Kong, 1997-1998. J. Infect. Dis., 185: 1005-1010.

Campanini G, Piralla A, Rovida F et al. First case in Italy of acquired resistance to oseltamivir in an immunocompromised patient with influenza A/H1N1v infection. J Clin Virol 2010; 48:220–222.

CDC. Oseltamivir-resistant 2009 pandemic influenza A (H1N1) virus infection in two summer campers receiving prophylaxis– North Carolina, 2009. MMWR Morb Mortal Wkly Rep 2009; 58:969–972.

Chan PA, Connell NT, Gabonay AM et al. Oseltamivir-resistant 2009-2010 pandemic influenza A (H1N1) in an immunocompromised patient. Clin Microbiol Infect 2010; 16:1576–1578.

Chen LF, Dailey NJ, Rao AK et al. Cluster of oseltamivir-resistant 2009 pandemic influenza A (H1N1) virus infections on a hospital ward among immunocompromised patients–North Carolina, 2009. J Infect Dis 2011; 203:838–846.

Collins PJ, Haire LF, Lin YP et al. Structural basis for oseltamivir resistance of influenza viruses. Vaccine 2009; 27:6317–6323.

Dharan NJ, Gubareva LV, Meyer JJ et al. Infections with oseltamivir-resistant influenza A(H1N1) virus in the United States. JAMA 2009; 301:1034–1041.

Duan S, Boltz DA, Seiler P et al. Oseltamivir-resistant pandemic H1N1/2009 influenza virus possesses lower transmissibility and fitness in ferrets. PLoS Pathog 2010; 6:e1001022.

Engelhardt, O. G. & Fodor, E. (2006) Functional association between viral and cellular transcription during influenza virus infection. *Rev. Med. Virol.* 16. p.329–345.

Guilligay, D., Tarendeau, F., Resa-Infante, P., Coloma, R., Crepin, T., Sehr, P., Lewis, J., Ruigrok, R. W., Ortin, J., Hart, D. J. & Cusack, S. (2008) The structural basis for cap binding by influenza virus polymerase subunit PB2. *Nat. Struct. Mol. Biol.* 15. p.500–506.

Harvala H, Gunson R, Simmonds P et al. The emergence of oseltamivir-resistant pandemic influenza A(H1N1) 2009 virus amongst hospitalised immunocompromised patients in Scotland, NovemberDecember, 2009. Euro Surveill 2010; 15:pii:19536.

Hamelin ME, Baz M, Abed Y et al. Oseltamivir-resistant pandemic A/H1N1 virus is as virulent as its wild-type counterpart in mice and ferrets. PLoS Pathog 2010; 6:e1001015.

Herlocher ML, Truscon R, Elias S et al. Influenza viruses resistant to the antiviral drug oseltamivir: transmission studies in ferrets. J Infect Dis 2004; 190:1627–1630.

Herlocher ML, Carr J, Ives J et al. Influenza virus carrying an R292K mutation in the neuraminidase gene is not transmitted in ferrets. Antiviral Res 2002; 54:99–111.

Hurt AC, Hardie K, Wilson NJ et al. Community Transmission of Oseltamivir-Resistant A(H1N1)pdm09 Influenza. N Engl J Med 2011; 365:2541–2542.

Huang, R. T. C., Rott, R. & Klenk, H. D. (1981) Influenza viruses cause hemolysis and fusion of cells. *Virology. 110 (1).* p.243-247.

Inoue M, Barkham T, Leo YS et al. Emergence of oseltamivir-resistant pandemic (H1N1) 2009 virus within 48 hours. Emerg Infect Dis 2010; 16:1633–1636.

Ito, T., Goto, H., Yamamoto, E., Tanaka, H., Takeuchi, M., Kuwayama, M., Kawaoka, Y. & Otsuki, K (2001) Generation of a highly pathogenic avian influenza A virus from an avirulent field isolate by passaging in chickens. *J Virol.* 75. p.4439-4443.

Lamb, R. A. & Choppin, P. W. (1983) The gene structure and replication of influenza virus. *Ann Rev Biochem*. 52. p.467-506.

Ludwig S, Planz O, Pleschka S, Wolff T (2003): Influenzavirus- induced signaling cascades: targets for antiviral therapy? Trends in Molecular Medicine 9, 46–52.

Liu JH, Bi YH, Qin K, Fu GH, Yang J, Peng JS, Ma GP, Liu QF, Pu J, Tian FL: Emergence of European avian influenza virus-like H1N1 swine influenza A viruses in China. J Clin Microbiol 2009, 47: 2643–2646.

Lenard, J. & Miller, D. K. (1981) P^H-dependent hemolysis by influenza, Semliki Forest virus, and Sendai virus. *Virology*. *110.* p.479-482.

Lazarowitz, S. G., Golberg, A. R. & Choppin, P. W. (1973) Proteolytic cleavage by plasmin of the HA polypeptide of influenza virus: Host cell activation of serum plasminogen. *Virology*. *56.* p.172-180.

Matsuzaki Y, Mizuta K, Aoki Y et al. A two-year survey of the oseltamivir-resistant influenza A(H1N1) virus in Yamagata, Japan and the clinical effectiveness of oseltamivir and zanamivir. Virol J 2010; 7:53.

Meijer A, Lackenby A, Hungnes O et al. Oseltamivir-resistant influenza virus A (H1N1), Europe, 2007-08 season. Emerg Infect Dis 2009; 15:552–560.

Monto AS, McKimm-Breschkin JL, Macken C et al. Detection of influenza viruses resistant to neuraminidase inhibitors in global surveillance during the first 3 years of their use. Antimicrob Agents Chemother 2006; 50:2395–2402.

Memoli MJ, Hrabal RJ, Hassantoufighi A, Eichelberger MC, Taubenberger JK. Rapid selection of oseltamivir- and peramivir-resistant pandemic H1N1 virus during therapy in 2 immunocompromised hosts. Clin Infect Dis 2010; 50:1252–1255.

Moore C, Galiano M, Lackenby A et al. Evidence of person-to-person transmission of oseltamivir-resistant pandemic influenza A(H1N1) 2009 virus in a hematology unit. J Infect Dis 2011; 203:18–24.

Nguyen HT, Fry AM, Loveless PA, Klimov AI, Gubareva LV. Recovery of a multidrug-resistant strain of pandemic influenza A 2009 (H1N1) virus carrying a dual H275Y/I223R mutation from a child after prolonged treatment with oseltamivir. Clin Infect Dis 2010; 51:983–984.

Muller, R., Poch, O., Delarue, M., Bishop, D. H. & Bouloy, M. (1994) Rift Valley fever virus L segment: correction of the sequence and possible functional role of newly identified regions conserved in RNA-dependent polymerases. *J. Gen. Virol.* 75. p.1345–1352.

Mounts, A.W., H. Kwong, H.S. Izurieta, Y. Ho, T. Au, M. Lee, C. Buxton Bridges, S.W. Williams, K.H. Mak, J.M. Katz, W.W. Thompson, N.J. Cox and K. Fukuda, 1999. Case-control study of risk factors for avian influenza A (H5N1) disease, Hong Kong, 1997. J. Infect. Dis., 180: 505-508.

Neumann G. Noda T, Kawaoka Y. Emergence and pandemic 18.Potential of swine-origin H1N1 influenza virus. Nature 2009; 459: 931-9.

Paul, S. W., Michi, S., Mario, R. & Sunetra, G. (2013) the antigenic evolution of influenza: drift or thrift? *Phil Trans R Soc B.* 368: 20120200.

Rameix-Welti MA, Munier S, Le Gal S et al. Neuraminidase of 2007-2008 influenza A(H1N1) viruses shows increased affinity for sialic acids due to the D344N substitution. Antivir Ther 2011; 16:597–603.

Renaud C, Pergam SA, Polyak C et al. Early emergence of an H275Y mutation in a hematopoietic cell transplant recipient treated with intravenous peramivir. Transpl Infect Dis 2010; 12:513– 517.

Robert, G., Webster, william J., Bean, Owen, T., Gorman, Thomas, M., Chambers, T. & Yoshihiro, K. (1992) Evolution and Ecology of Influenza A Viruses. *Microbiological Reviews.* 56(1). p.152-179.

Sanz-Ezquerro, J. J., Luna, S. de la, Ortín, J. & Nieto, A. (1995) Individual expression of influenza virus PA protein induces degradation of coexpressed proteins. *J. Virol.* 69. p.2420–2426.

Salgado, C.D., B.M. Farr, K.K. Hall, F.G. Hayden, 2002. Influenza in the acute hospital setting. Lancet Infect. Dis., 2: 145-155. Van Riel, D., V.J. Munster, E. de Wit, G.F. Rimmelzwaan, R.A.M. Fouchier, A.D.M.E. Osterhaus and T. Kuiken, 2006. H5N1 virus attachment to lower respiratory tract. Science, pp: 312: 399.

Seibert CW, Kaminski M, Philipp J et al. Oseltamivir-resistant variants of the 2009 pandemic H1N1 influenza A virus are not attenuated in the guinea pig and ferret transmission models. J Virol 2010; 84:11219–11226.

Tong, S., Zhu, X., Li, Y., Shi, M., Zhang, J., Bourgeois, M., Yang, H., Chen, X., Recuenco, S., Gomez, J., Chen, L. M., Johnson, A., Tao, Y., Dreyfus, C., Yu, W., McBride, R., Carney, P. J., Gilbert, A. T., Chang, J., Guo, Z., Davis, C. T., Paulson, J. C., Stevens, J., Rupprecht, C. E., Holmes, E. C., Wilson, I. A. & Donis, R. O. (2013) New World Bats Harbor Diverse Influenza A Viruses. *PLoS Pathogens*. 9 (10). e1003657.

Tramontana AR, George B, Hurt AC et al. Oseltamivir resistance in adult oncology and hematology patients infected with pandemic (H1N1) 2009 virus, Australia. Emerg Infect Dis 2010; 16:1068– 1075.

Valinotto LE, Diez RA, Barrero PR et al. Emergence of intratreatment resistance to oseltamivir in pandemic influenza A H1N1 2009 virus. Antivir Ther 2010; 15:923–927.

Van der Vries E, Stelma FF, Boucher CA. Emergence of a multidrug-resistant pandemic influenza A (H1N1) virus. N Engl J Med 2010; 363:1381–1382.

Wilson, I. A., Skehel, J. J. & Wiley, D. C. (1981) Structure of the hemagglutinin membrane glycoprotein of influenza virus at 3 Å resolution. *Nature*. *289.* p.366-373.