Evolutionary distances between proteins of the Influenza A Virus

Hatem Nassrat
B00393388
Bioinformatics: Winter 06/07
Dr. Christian Blouin

Table of Contents

Evolutionary distances between proteins of the Influenza A Virus

Introduction

Influenza, according to wikipedia, “is an infectious disease of birds and mammals caused by an RNA virus of the family Orthomyxoviridae” [1*]. The source of the infection (virus) would be an acute infection in similar animals (other humans) acting as the reservoir of the virus [5]. The virus is spread through aerial droplets inhaled into the respiratory tract or pharynx and has a short incubation period of 1-3 days, which may lead to epidemics [5]. Main victims of this disease are the young, elderly, or patients with chronic heart-lung problems [5].

The virion maybe rounded or long and filamentous. Each single stranded RNA is associated with one helical NP (Nucleoprotien), and is available in eight segments of the RNP (ribonucleoprotein). Where all eight are to be present for successful replication. The genome is enclosed inside a lipoprotein envelope. Matrix protein (MP1) is chemically bound to the RNP lining the inside of the envelope. The envelope has two types of protruding spikes. One is neuraminidase (NA), a box-shaped protein, which has enzymic properties. The second is a trimeric protein called HA (haemagglutinin). Which functions during attachment of the virus particle to a cell membrane, and combines with some receptors on a variety of cells such as red blood cells. The lipoprotein envelope makes the virion susceptible to heat, drying, detergents and solvents. [5]

Illustration 1: Influenza Virion [5]

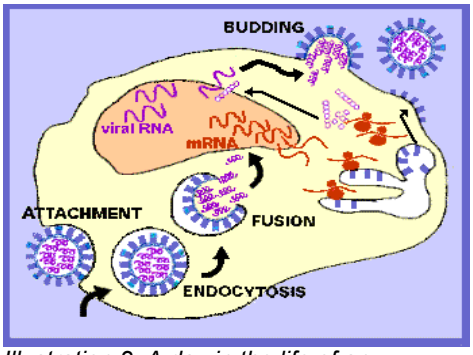

Illustration 2: A day in the life of an Influenza Virion [5]

This virus enters the cell, via a process know as endocytosis. The RNP is then released from the virus and transported into the nucleus. New viral protein is then created from the mRNA. The viral RNA is combined with new capsid protein and MP1 and are then transported to the cell surface where the HA and NA are added. The new virions then buds from the cell to attack other cells. The attacked cells stay alive initially allowing multiple virion to bud. [5]

Two different virions may attack one cell allowing for recombination, since influenza has eight segments to its genome, this allows for the development of new flu strains [5]. This virus has been evolving in nature for millennia, and in particular protein codings within the strains of this disease have been changing, bringing forth new more advanced versions of this virus. According to [5] avian and human strains recombining in pigs in the Far East may allow virulent human strains to evolve. The Influenza A category is classified due to its HA undergoing minor and occasionally major changes with some variation in NA [5]. “Because of the absence of RNA proofreading enzymes, the ... RNA transcriptase makes a single nucleotide insertion error roughly every 10 thousand nucleotides, which is the approximate length of the influenza v(iral) RNA”[2]. This project aims to compute the evolutionary distance between pairs of proteins within different influenza viri, through phylogeny analyses using likelihood ratios.

Every 10-15 years comes about an influenza virus strain with a new HA and occasionally a new NA as well. This antigenic shift in the virus usually causes a pandemic (major epidemic). Antigenic drifts (small variations) are experienced every 2-3 years, according to [5] this is due to selective pressure by antibodies in certain human populations pushing The virus strains adapt to survive. The chart bellow lists the main Influenza A isolated in previous years.

This study would be setup to analyze amino acid sequences from influenza. The sequences will be analyzed in groups, where sequences with the same type of HA and NA would be grouped. This will be used to measure evolutionary distances between pairs of proteins.

1874 --- (H3N8)	
1890 --- (H2N2)	.....Pandemic
1902 --- (H3N2)	
1918 --- (H1N1)	.....Pandemic
1933 --- (H1N1)	.....First strains isolated
1947 --- (H1N1)	.....Variation detected
1957 --- (H2N2)	....."Asian" Flu pandemic
1968 --- (H3N2)	....."Hong Kong" Flu pandemic
1976 --- (H1N1)	....."Swine" Flu, non-epidemic
1977 --- (H1N1) + (H3N2)	..."Russian" Flu epidemic

Chart 1: Different Influenza_A Strains [5]

Materials And Methods

A database table of Influenza A protein sequences was gathered from available gen banks. Each record of this table contains an amino acid sequence, its “GenInfo Identifier” (gi), the name of the protein that is sequenced, the HA type and NA type of the strain, and the strain identifier string (e.g. A/Bar-headed Goose/Qinghai/5/05(H5N1)). This data would be used in the analyses to judge the evolutionary distance between pairs of proteins.

The analyses begins with grouping the protein sequences that came from viri of the same type (e.g. H1N1). Within this group amino acid sequences that code a certain protein would further sub grouped (e.g. grouping sequences coding the HA protein). The analyses to be performed is a paired analyses, to determine the evolution of some proteins with respect to others. To do so we would select sequences coding Protein 1, where within the same strain we have the coded sequence for Protein 2 also and vice versa. Thus generating two sets of sequences, one filled with codings of Protein 1, and the other of Protein 2.

These sequences are then to be aligned using Clustalw [2], then passed to Phyml [4] to generate a phylogeny tree for each of the two sets. The phylogeny is based on a gamma distribution with four rate categories. These two sets are expected to have the same topology, due to Darwin's theories of evolution. If the phylogeny estimators are correct then the separate trees should show similar results on which strain evolved next. To add to that, a third set is to be generated. The sequences from Protein1 are joined with their counterpart (coming from the same strain) sequences coding Protein2. This third set is then aligned and a phylogeny tree is built from the data. Similarly the tree for the concatenated sequence should have a similar topology.

The analyses would be based on the log likelihood of the generated trees. The following ratio is to be calculated, lk(Tree1) + lk(Tree2) - lk(Combined_Trees), where lk is the log likelihood. This ratio can be used to measure the evolutionary distance between Protein1 and 2. The closer this ratio is to Zero, the greater the assumption that proteins generating Tree1 and Tree2 have evolved at a similar pace. The greater the displacement of this ratio from zero, the greater the probability that one of the proteins has evolved faster than the other protein. The results of this experiment would allow identification of fast moving proteins within the Influenza A virus family.

Since the data in the database has been gathered in a general fashion, some viri types may have more strains than others. A correlation in the size of the ratio versus the size of the set that was analyzed could be problematic, as this leads to the assumption that larger sets produce may produce larger log likelihoods. This will be checked for to have confidence in the results produced.

A python program was written to do this calculation. The application generates phylogeny trees and likelihoods for each pair of protein for each type of virus. Following that the results are compiled to find likelihoods that are either too large or too small, thus indicating a difference in the evolutionary rate.

Results

The results show that most grouped sequences, produced a low log likelihood ratio, meaning that in most cases the pairs have evolved at a similar pace. However, some pairs had high scoring ratios. A sample of such pairs are to be displayed bellow.

H1N2 virus

The following protein pairs all had a high ratio score:

PA (Polymerase Acidic) vs. NP (Nucleoprotein)

This pair had the highest log likelihood ratio in the set, showing (according to the hypothesis) that one of these pairs has evolved at a higher rate than the other. Looking at the single likelihoods for the trees, shows that the NP within H1N2 seems to have evolved faster than the PA amino acid. Illustrated bellow are the phylogeny trees drawn using [1].

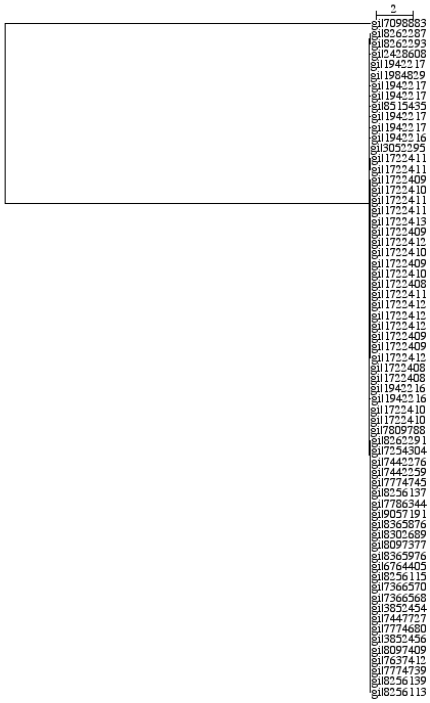

Illustration 4: PA (H1N2) Phylogeny Tree

Illustration 3: NP (H1N2) Phylogeny Tree

These two trees are of different topologies which would explain the difference in likelihood. The tree built using sequences that code the polymerase seem to be closer genetically (with one distant relative coming from the strain “A/swine/Bakum/1832/00(H1N2)”) than the sequences coding the NP sequences. This would also add to the hypothesis that the NP amino acid evolved differently in strains of the H1N2 Influenza A virus. Illustrated bellow is the phylogeny tree for the concatenated sequences (the sequences are labeled according to the label of their respective PA sequences).

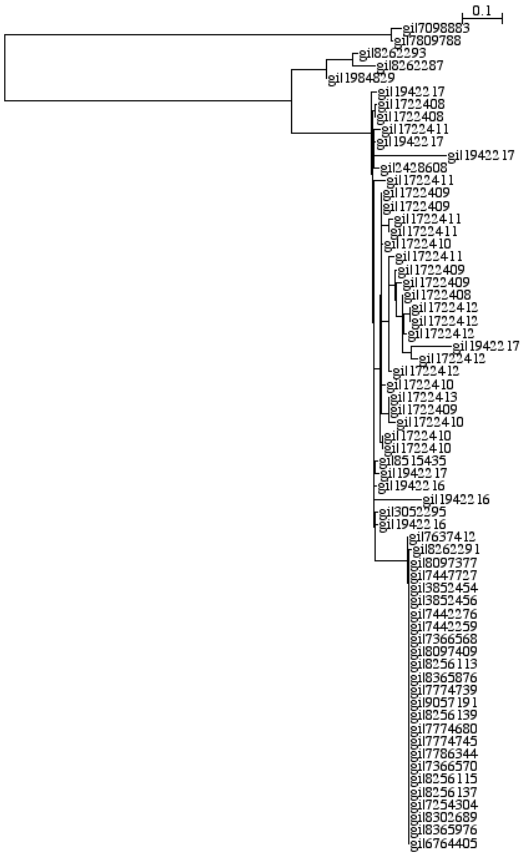

Illustration 5: PA+NP (H1N2) Phylogeny Tree

This tree has a different topology, the topological structure is induced by the distances between the sequences coding the NP amino acid, while keeping some of the evolutionary distance (within the tree) of the PA sequences. The log likelihoods of the previously illustrated trees is summarized in the ratio, (-4039.626319 + -2380.280615) - (-12499.408361) = 6079.501427, where the first likelihood is the PA tree likelihood.

Other high scoring pairs within the H1N2 virus include:

PB1 vs NA

Within the H1N2 virus, this amino acid pair also had a relatively high score, indicating a mismatch in the rate of evolution. The log likelihoods of their trees is, (-3147.911264 + -2641.114318) - -6507.893942 = 718.868360, showing differences in the two protein phylogenies. (Appendix A shows the three trees generated by Phyml)

PB1 vs NS2

This amino acid pair similarly produced a large ratio. The log likelihood summary of the trees is, -2956.597706 + -673.348382 - -4262.593613 = 632.647525, which shows that the likelihood of the NS2 tree is very small. From the overall topology of the trees we see a slightly faster evolutionary movement in the polymerase PB1. (Appendix A shows the Trees)

H6N5 Influenza A Virus

NP vs NA

This amino acid pair within the H6N5 virus strains seemed to be the most diverse, giving a likelihood ratio of: -1555.197691 + -1959.487073 - -5655.285922 = 2140.601158. The closeness of the log likelihoods of the first two trees versus the large likelihood of the third tree shows that the topologies of the first two trees maybe different. The evidence is pointing towards the hypothesis that the NP amino acid came from a different parent (source) in the H6N5 strains than the NA.

NP vs HA

The likelihood ratio given by, -1571.376155 + -2542.587186 - -4220.493956 = 106.53061, points to the hypothesis that the evolution of the HA in the H6N5 virus is faster than the evolution of the NP amino acid. The hypothesis is based on the knowledge of evolution of the HA amino acid as opposed to the evolution of the NA (described above).

H7N2 Influenza A Virus

NA vs HA

The ratio for this protein pair is -2487.483025 + -3027.464985 - -5971.337327 = 456.389317 (where the first number is the NA tree likelihood). The three trees seem to have a common distant relative from the strain “A/dk/Hong Kong/293/1978(H7N2)”. The NA in that strain has a slightly close relative from the strain “A/Chicken/New York/13142-5/? (H7N2)”. the question mark in the strain year is because of ambiguity, due to strains gathered in multiple years. The HA phylogeny tree seems to show faster evolution, when looking at the branching within the tree. On the other hand, the NA tree seems to be more distributed, showing a slow rate if evolution. The trees generated here can be seen in Appendix A.

H6N2 Influenza A Virus

NA vs HA

Log likelihood ratio for this pair is, -2105.657638 + -2941.473046 - -5370.264902 = 323.134218. In this situation the branch lengths within the HA tree are longer than those of the NA tree. Also there are a larger number of distinct levels in the HA tree and the way in which the levels deepen in that tree infer that the HA protein has been evolving faster within the H6N2 virus strains. Appendix A shows both trees along with the concatonated tree generated through Phyml[4].

General Results

The final log likelihood ratios can be seen for all sequences that had a large displacement from zero in Appendix B. Appendix C contains all the likelihood ratios that showed an equal evolutionary drift (due to having a likelihood ratio close to zero).

Conclusions

The validity of this type of analyses is questionable, due to the many variables involved in calculating phylogenies, and the variables that are yet to be discovered by phylogeny scientists. Keeping that in mind, we still are able to judge some evolutionary distances using the available methods. Pairs that tend to score high (likelihood ratio) tend to have a polymerase as a unit of the pair. From the results we can see, that a reason for this may be the stability of the polymerase sequence versus the diversity of its partners that force this high score in the likelihood estimation. Another observation directed towards the NS2 protein, show that in some Influenza A Virus types, move slower than the other proteins, even the polymerase. Also the hypothesis that HA tends to mutate at a faster rate than the other proteins seems to hold.

In conclusion, understanding the evolution of the Influenza A virus may help predict the nature of the new strains that we may face in the future. Doing so may allow scientists to prepare for the next epidemic and maybe able to prevent its spread. In general viri phylogeny analyses may lead to major discoveries that may save lives.

References

1. NJplot.

2. Gibson T, Thompson J, and Higgins D. ClustalW. 1.83: .

3. Huelsenbeck JP, Crandall KA. Phylogeny estimation and hypothesis testing using maximum likelihood. Annu Rev Ecol Syst 1997; 28: 437-466.

4. Jean-Francois Dufayard, Stephane Guindon, and Vincent Lefort. PhyML.

5. Keen A. Influenza. 2000. Retrieved 04/04, 2007, from http://web.uct.ac.za/depts/mmi/jmoodie/influen2.html

6. Perriere G, Gouy M. WWW-query: An on-line retrieval system for biological sequence banks. Biochimie 1996; 78: 364-369.

Appendix A

(H1N2) PB1 vs NA

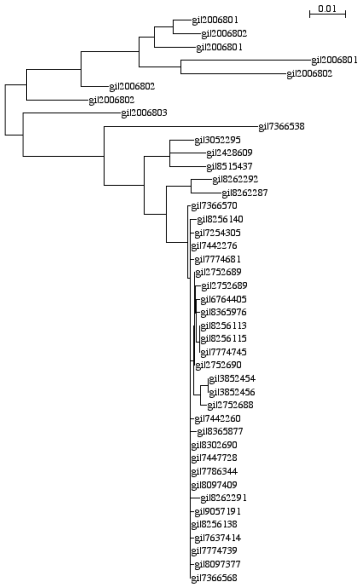

Illustration 6: PB1 (H1N2) Tree

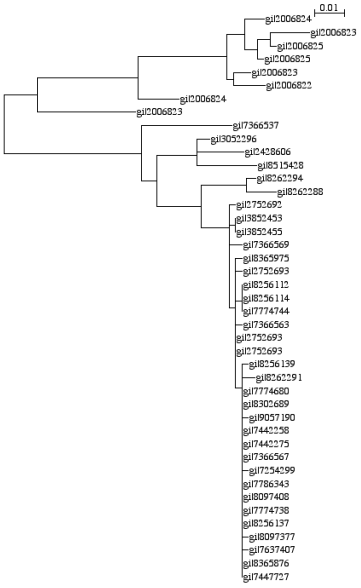

Illustration 7: (H1N2) NA Tree

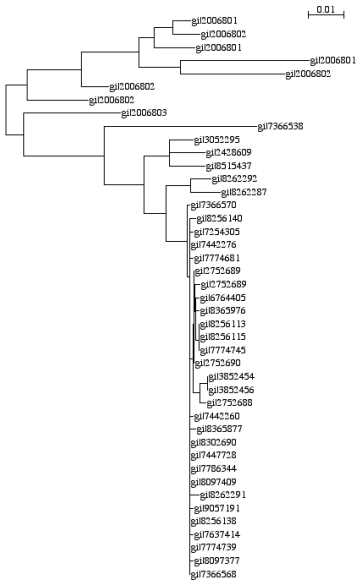

Illustration 8: PB1+NA (H1N2) Tree

(H1N2) PB1 vs NS2

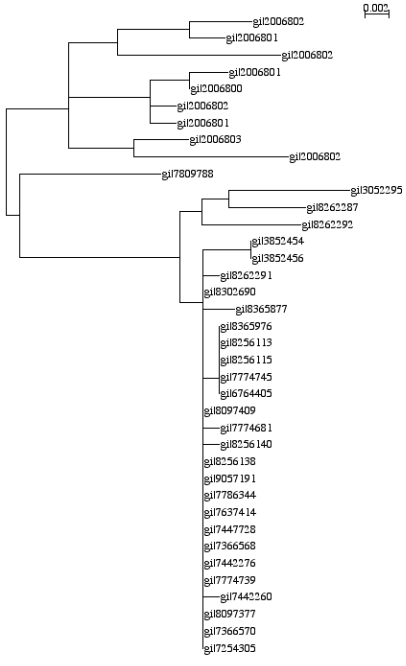

Illustration 9: PB1 (H1N2)

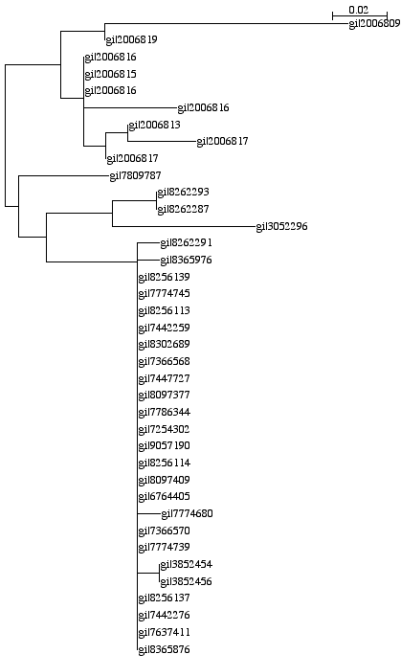

Illustration 11: NS2 (H1N2)

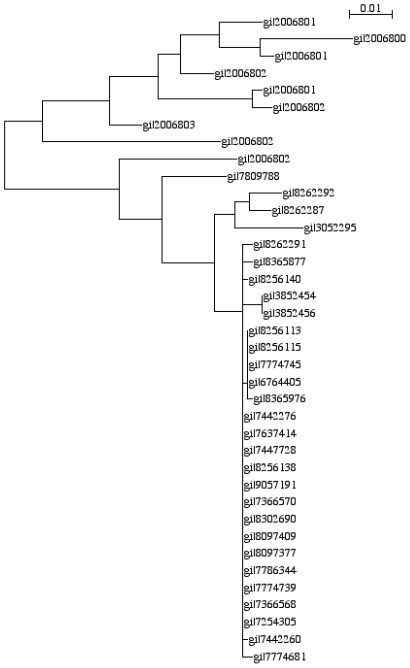

Illustration 10: PB1+NS2 (H1N2)

(H7N2) HA vs NA

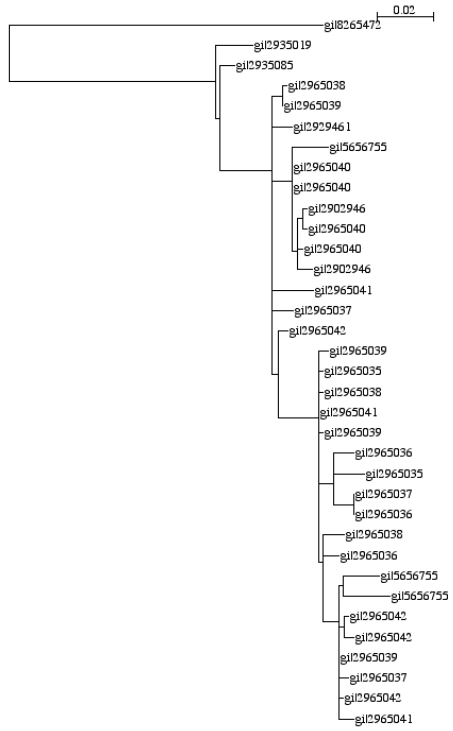

Illustration 12: HA (H7N2)

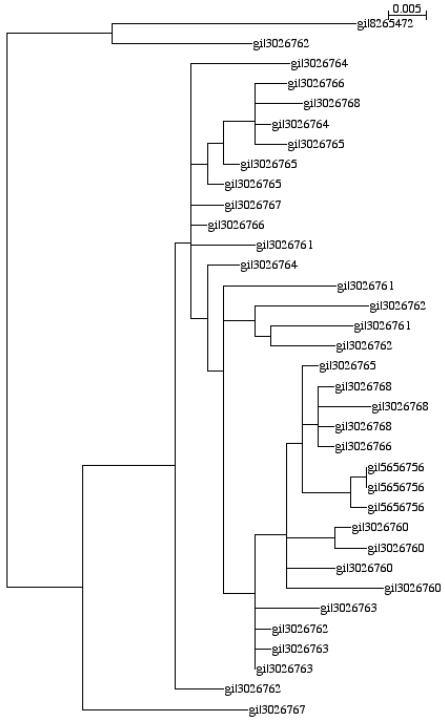

Illustration 13: NA (H7N2)

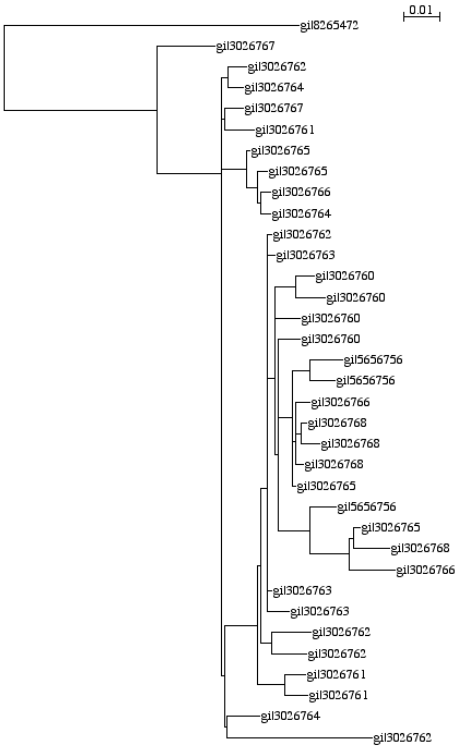

Illustration 14: HA+NA (H7N2)

(H6N2) HA vs NA

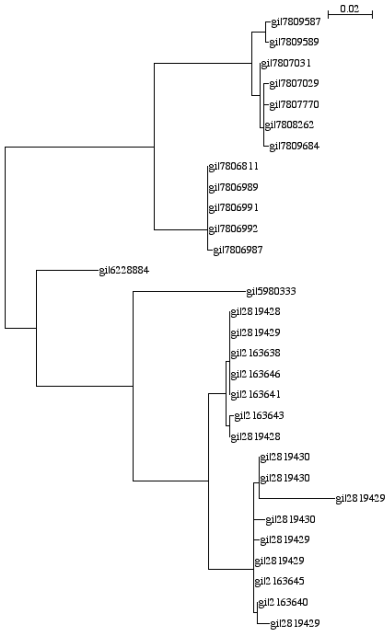

Illustration 16: HA (H6N2)

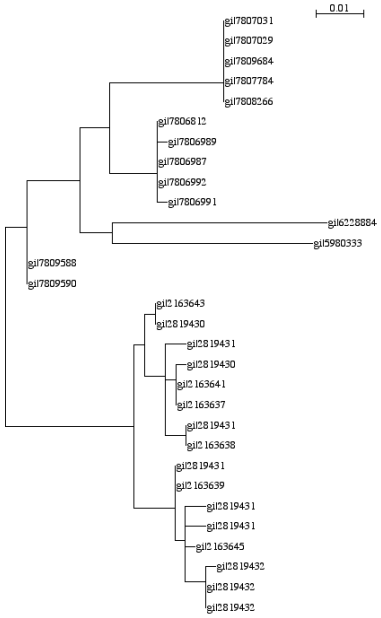

Illustration 15: NA (H6N2)

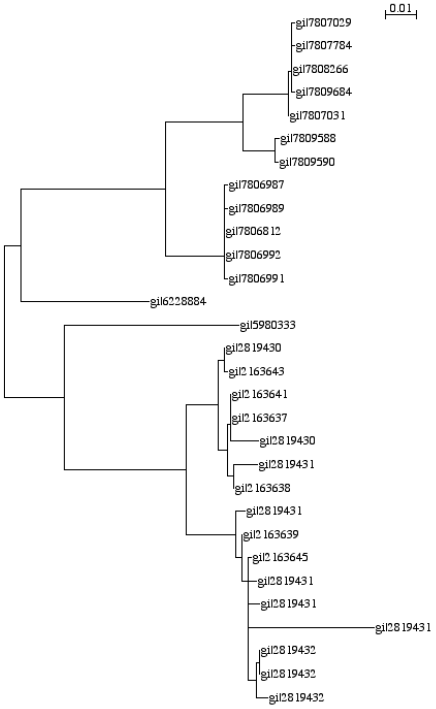

Illustration 17: HA+NA (H6N2)

Appendix B

PANP,H1N2 -> 6079.501427
PANP,H6N1 -> 2222.582660
PB1NP,H6N1 -> 2170.935292
NPNA,H6N5 -> 2140.601158
M2M1,H6N2 -> 1685.199689
PB2NP,H3N8 -> 1635.920114
PB1HA,H7N1 -> 774.016892
PB2NA,H6N1 -> 745.061392
PB1NA,H1N2 -> 718.868360
PB2PB1,H2N2 -> 708.086385
PB2PB1,H4N2 -> 697.356806
PB1PA,H3N8 -> 690.243694
NPNA,H3N8 -> 686.656592
PB1NA,H3N8 -> 666.979068
PB2NP,H2N2 -> 662.029955
PB1NS2,H1N2 -> 632.647525
PB1NP,H7N1 -> 594.173817
PB2HA,H4N2 -> 590.659304
PB2PB1,H6N1 -> 521.145566
NAHA,H7N2 -> 456.389317
PB2NS1,H3N8 -> 409.335536
PB1M1,H7N1 -> 397.931941
PB2NA,H3N8 -> 370.875420
PB1NS1,H3N8 -> 357.414384
PANS1,H3N8 -> 341.239134
NPNS1,H3N8 -> 337.739756
PB1-F2PB1,H1N1 -> 329.878243
NAHA,H6N2 -> 323.134218
M2HA,H1N2 -> 315.031797
M2M1,H6N6 -> 287.521246
M2NA,H1N2 -> 277.174188
PB2PB1-F2,H1N1 -> 262.968528
PB2NP,H6N1 -> 235.574483
NPHA,H6N2 -> 224.474455
PB1NP,H1N2 -> 215.651532
M2M1,H6N4 -> 214.063671
NS1NA,H7N3 -> 212.455858
PB1-F2PA,H1N1 -> 212.139725
PB1-F2NP,H1N1 -> 211.994332
PB1M1,H1N2 -> 211.156029
PB1-F2NS1,H1N1 -> 198.036979
PB1M2,H7N1 -> 180.810330
M2M1,H9N7 -> 170.796313
NS1NA,H6N1 -> 161.530924
PB2HA,H3N8 -> 155.714414
PB1-F2M2,H1N1 -> 153.325944
NS2NA,H1N2 -> 152.594646
NPM1,H1N2 -> 148.109534
PANS1,H7N3 -> 142.627629
NPNS1,H6N8 -> 138.879604
NS1NA,H11N9 -> 138.175389
PB2HA,H6N2 -> 136.573807
NS1HA,H7N1 -> 133.205090
NS2HA,H1N2 -> 132.157867
PB1-F2NS2,H1N1 -> 131.406251
NS1NA,H5N3 -> 130.040234
PANS1,H12N5 -> 129.560917
NS1HA,H2N3 -> 128.824466
PB1-F2M1,H1N1 -> 128.780773
PB1NS1,H6N8 -> 125.400460
PB1NS1,H10N7 -> 124.092479
PB2NS1,H10N7 -> 123.594030
PB2NS1,H11N9 -> 122.465424
NPNS1,H5N2 -> 122.152254
NPM2,H1N2 -> 121.468562
NS1HA,H10N7 -> 120.328882

PANS1,H2N3 -> 120.061780
PB1NS1,H11N9 -> 119.855873
PB2NS1,H12N5 -> 116.894733
PB2NS1,H2N3 -> 116.481351
NS1HA,H6N1 -> 115.291595
M1NA,H1N2 -> 114.432662
PB1NS1,H3N5 -> 114.171436
PANS1,H4N8 -> 113.988637
M2NS1,H1N2 -> 113.963473
PB1NS1,H2N3 -> 113.495938
PB1NS1,H4N8 -> 113.276610
PANS1,H10N7 -> 113.149667
M1NA,H7N2 -> 113.030952
PB1-F2NS1,H3N8 -> 112.858438
PANS1,H11N9 -> 112.034853
PB2NS1,H4N8 -> 112.009498
NS1HA,H6N8 -> 110.954459
PB1NS1,H11N2 -> 110.450564
PB1-F2NS1,H11N9 -> 110.330604
PB2NS1,H6N8 -> 110.029620
NS1NA,H6N8 -> 109.698853
PB1NS1,H12N5 -> 109.460551
PB1NS1,H4N2 -> 109.024317
PB2PA,H3N8 -> 108.803402
PB1-F2PB1,H3N8 -> 107.219984
NPHA,H6N5 -> 106.530615
PB1-F2PA,H3N8 -> 106.344040
PANS1,H6N8 -> 106.246748
NPNS1,H2N3 -> 105.752757
NS1NA,H2N3 -> 105.605663
NS1NA,H10N7 -> 105.500281
PB1NS1,H3N1 -> 105.394888
PB2PA,H7N7 -> 105.088541
PANS1,H8N4 -> 104.050426
NPHA,H6N8 -> 103.504683
PANS1,H7N1 -> 103.042805
PB2NS1,H3N5 -> 102.918397
NPNS1,H11N9 -> 102.040641
NPNS1,H10N7 -> 101.563180
M2NA,H7N2 -> 100.991958
PB2NS1,H3N1 -> 100.568714
PANP,H2N3 -> -425.803466
PB2HA,H6N1 -> -907.251896
PB2PA,H6N1 -> -1404.558521
PB1PA,H6N1 -> -1986.533627

Appendix C

M2HA,H4N2 -> 0.982077
PB2NA,H15N9 -> 0.971286
PB1NA,H15N9 -> 0.961941
PB2HA,H13N2 -> 0.941715
NS2NS1,H6N6 -> 0.939621
PB2PA,H2N1 -> 0.938899
PB1M1,H4N2 -> 0.923900
PANA,H15N9 -> 0.911469
PB1NP,H6N3 -> 0.903222
NPNS2,H9N5 -> 0.886513
PB1-F2NS2,H9N5 -> 0.864287
M2NS2,H3N3 -> 0.851132
PB1-F2NP,H9N5 -> 0.838157
NPNS1,H6N6 -> 0.831960
PANS1,H9N6 -> 0.829868
M2NS2,H9N6 -> 0.827947
NPM1,H10N1 -> 0.826918
NPM1,H13N2 -> 0.826431
NPM1,H9N5 -> 0.826242
NPHA,H13N2 -> 0.803395
NAHA,H6N6 -> 0.793131
M2HA,H8N4 -> 0.791608
NAHA,H15N9 -> 0.792105
M2HA,H9N5 -> 0.778246
PB2PB1,H10N1 -> 0.778001
M1NS2,H11N1 -> 0.776214
M2HA,H13N2 -> 0.773501
PANP,H4N1 -> 0.768378
PB2PB1,H4N4 -> 0.749959
M1NS1,H9N5 -> 0.748920
NPNA,H15N9 -> 0.728424
PB1-F2M2,H9N5 -> 0.720906
PB2PB1,H6N3 -> 0.712055
M2HA,H9N6 -> 0.710448
PB1NS1,H6N6 -> 0.703000
PB2PB1,H5N3 -> 0.702150
NPNS2,H7N7 -> 0.689070
PB1-F2NS2,H6N3 -> 0.683549
PB1PA,H5N3 -> 0.659786
PB1NP,H6N6 -> 0.650269
PB1-F2M2,H6N3 -> 0.630550
PAM1,H9N5 -> 0.625872
M2HA,H6N6 -> 0.619510
M2NA,H13N2 -> 0.618319
PANP,H13N2 -> 0.613288
PAM2,H3N3 -> 0.607943
NS2HA,H2N8 -> 0.604406
M2NS2,H11N3 -> 0.589175
NPNS2,H9N6 -> 0.582972
PB1M1,H4N9 -> 0.583423
PB2PA,H4N1 -> 0.572316
PB1M2,H4N9 -> 0.531054
PAM2,H10N1 -> 0.529880
PANP,H10N7 -> 0.523062
PANP,H6N3 -> 0.493637
NS2NA,H13N6 -> 0.491397
NS2NS1,H9N6 -> 0.485815
PB1M2,H13N2 -> 0.471386
PB2M2,H4N8 -> 0.459076
M1NA,H15N9 -> 0.436447
NS2NS1,H2N8 -> 0.424708
NS1NA,H15N9 -> 0.398258
PB2PA,H5N3 -> 0.394335
NS1HA,H2N8 -> 0.378707
PANS2,H9N5 -> 0.353104
M2NS1,H9N5 -> 0.340652

NPM2,H9N5 -> 0.341588
PB1M2,H3N5 -> 0.319883
PB2PA,H6N3 -> 0.309560
M2NS2,H4N9 -> 0.301454
NS1HA,H6N6 -> 0.297242
PB1PA,H4N1 -> 0.257791
PAM2,H9N5 -> 0.251119
PAHA,H6N3 -> 0.234597
NS2NA,H15N9 -> 0.229982
M2NA,H9N6 -> 0.215758
M2M1,H6N9 -> 0.206910
M2NS1,H3N3 -> 0.207028
PB1NP,H4N1 -> 0.194543
NS2NS1,H13N2 -> 0.185495
M2NA,H15N9 -> 0.181902
PB1-F2NA,H15N9 -> 0.172988
NPM2,H9N6 -> 0.161280
M2NS1,H9N6 -> 0.137475
PB2PA,H9N6 -> 0.103083
PB2PA,H2N5 -> 0.099394
PB1HA,H6N6 -> 0.094744
PB1PA,H9N6 -> 0.090776
NS2HA,H8N4 -> 0.088364
PB1PA,H6N3 -> 0.074169
PB2PB1,H4N1 -> 0.063596
PB2NS2,H9N5 -> 0.060477
PB1HA,H6N3 -> 0.055726
PB2NP,H4N1 -> 0.048698
PB2NP,H6N3 -> 0.048332
NPHA,H6N6 -> 0.029514
PANS2,H9N6 -> 0.017941
PB1NS2,H9N6 -> 0.000647
PB2NS2,H9N6 -> 0.001259
M1HA,H15N9 -> 0.000000
M1NS1,H15N9 -> 0.000000
M1NS2,H15N9 -> -0.000000
M1NS2,H9N5 -> 0.000000
M2HA,H15N9 -> 0.000000
M2M1,H15N9 -> 0.000000
M2NS1,H15N9 -> 0.000000
M2NS2,H15N9 -> 0.000000
M2NS2,H9N5 -> -0.000000
NPHA,H15N9 -> 0.000000
NPM1,H15N9 -> -0.000000
NPM2,H15N9 -> 0.000001
NPNS1,H15N9 -> 0.000000
NPNS1,H9N6 -> 0.000064
NPNS2,H15N9 -> 0.000000
NS1HA,H15N9 -> 0.000000
NS2HA,H15N9 -> 0.000000
NS2NS1,H15N9 -> 0.000000
PAHA,H15N9 -> 0.000000
PAM1,H15N9 -> 0.000000
PAM2,H15N9 -> 0.000000
PANP,H15N9 -> 0.000000
PANS1,H15N9 -> -0.000001
PANS2,H15N9 -> 0.000000
PB1-F2HA,H15N9 -> -0.000000
PB1-F2M1,H15N9 -> 0.000000
PB1-F2M2,H15N9 -> 0.000000
PB1-F2NP,H15N9 -> 0.000000
PB1-F2NS1,H15N9 -> -0.000001
PB1-F2NS2,H15N9 -> 0.000000
PB1-F2PA,H15N9 -> 0.000000
PB1-F2PB1,H15N9 -> 0.000000
PB1HA,H15N9 -> 0.000000
PB1M1,H15N9 -> 0.000000
PB1M1,H9N5 -> 0.000000
PB1M2,H15N9 -> 0.000001
PB1M2,H9N5 -> 0.000000

PB1NP,H15N9 -> 0.000001
PB1NS1,H15N9 -> 0.000000
PB1NS2,H15N9 -> 0.000000
PB1NS2,H9N5 -> 0.000000
PB1PA,H15N9 -> 0.000000
PB2HA,H15N9 -> 0.000000
PB2M1,H15N9 -> 0.000000
PB2M1,H9N5 -> 0.000000
PB2M2,H15N9 -> 0.000000
PB2M2,H9N5 -> 0.000000
PB2NP,H15N9 -> 0.000000
PB2NS1,H15N9 -> -0.000001
PB2NS2,H15N9 -> 0.000000
PB2PA,H15N9 -> 0.000000
PB2PB1,H15N9 -> 0.000000
PB2PB1,H9N5 -> 0.000000
PB2PB1,H9N6 -> 0.000361
PB2PB1-F2,H15N9 -> 0.000000