Title	Evolutionary Network Analysis Approach for Understanding Host Influenza A Virus Relationships
Sub Title	
Author	臼居, 優希
Publisher	慶應義塾大学湘南藤沢学会
Publication year	2013
Jtitle	生命と情報 No.20 (2013.) ,p.107- 115
JaLC DOI	
Abstract	Host cells have coexisted with viruses from ancient times, thereby both of host and virus have affected each other's evolution. Influenza A virus (IAV) undergoes remarkably rapid evolution to evade host defense systems, whereas host has developed various antiviral responses such as antibodies and RNA silencing systems. Recently, two research groups nave reportea that human microRNAs (miR-323, miR-491 and miR-654) bind to influenza A virus polymerase ß subunit (PB1) mRNA and miR-let-7c also binds to M1 mRNA, inhibiting the replication of IAV respectively. These observations suggested that host microRNAs (miRNAs) have a potential as an antiviral system against IAV, however, it remained unclear whether or not IAV has conserved miRNA-target sites during their dynamic evolution. Here, we investigated evolutionary relationships between host miRNAs and IAV by using IAV sequence data for about a hundred years. Firstly, large-scale evolutionary networks detected both of dynamic mutations by interspecies-transmission processes and gradual mutations after transmitting to new hosts. Secondly, conservation analysis based on Shannon's information theory proved that miRNA-target sites in PB1 and MI1 mRNAs were significantly conserved in human, swine and avian IAVs. Meanwhile, miR-323 and miR-491 were conserved in human and swine genome, and miR-let-7c was conserved in human, swine, and avian genome. Our results suggest that regulatory relationships of host miRNAs have been evolutionarily conserved in a broad host range, even though IAV mutate rapidly. We proposed following hypotheses : (1)host miRNAs targeted immutable regions of IAV mRNA, or (2) IAV might subvert host defense system to evade from host immune systems.
Notes	慶應義塾大学湘南藤沢キャンパス先端生命科学研究会 2013年度学生論文集 卒業論文ダイジェスト
Genre	Technical Report
URL	https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=KO92001004-00000020-0107

Evolutionary Network Analysis Approach for Understanding Host-Influenza A Virus Relationships

Department of Policy Management

臼居　優希

Abstract

Host cells have coexisted with viruses from ancient times, thereby both of host and virus have affected each other's evolution. Influenza A virus (IAV) undergoes remarkably rapid evolution to evade host defense systems, whereas host has developed various antiviral responses such as antibodies and RNA silencing systems. Recently, two research groups have reported that human microRNAs (miR-323, miR-491 and miR-654) bind to influenza A virus polymerase β subunit (PB1) mRNA and miR-let-7c also binds to M1 mRNA, inhibiting the replication of IAV respectively. These observations suggested that host microRNAs (miRNAs) have a potential as an antiviral system against IAV, however, it remained unclear whether or not IAV has conserved miRNA-target sites during their dynamic evolution. Here, we investigated evolutionary relationships between host miRNAs and IAV by using IAV sequence data for about a hundred years. Firstly, large-scale evolutionary networks detected both of dynamic mutations by interspecies-transmission processes and gradual mutations after transmitting to new hosts. Secondly, conservation analysis based on Shannon's information theory proved that miRNA-target sites in PB1 and M1 mRNAs were significantly conserved in human, swine and avian IAVs. Meanwhile, miR-323 and miR-491 were conserved in human and swine genome, and miR-let-7c was conserved in human, swine, and avian genome. Our results suggest that regulatory relationships of host miRNAs have been evolutionarily conserved in a broad host range, even though IAV mutate rapidly. We proposed following hypotheses: (1) host miRNAs targeted immutable regions of IAV mRNA, or (2) IAV might subvert host defense system to evade from host immune systems.

Keywords : microRNA, Influenza A virus, PB1, M1, evolution, host-virus interaction

1. Introduction

Influenza A viruses (IAVs) undergo remarkably rapid evolution to evade host defense systems, whereas host has developed various antiviral responses such as antibodies and RNA silencing systems. MicroRNA (miRNA) is a class of small non-coding RNA molecules (~22 nucleotides length) which play important roles in the cell (Lewis et al., 2005; Rosalind et al., 1993). Recently, It has become clear that host miRNAs are strongly related to the battle of between host and pathogen in some virus infections, and also IAV infection (Xiao and Rajewsky, 2009). In 2010, Song *et al.* reported that three human miRNAs (miR-323, miR-491, and miR-654) bind to almost the same 3′ coding regions in the H1N1 IAV PB1 mRNA, inhibiting the replication of the H1N1 IAV through degradation of PB1 mRNA in MDCK cells (Song et al., 2010). Afterwards, Ma *et al.* elucidated that miR-let-7c directly targets the 3′ untranslated regions (3′ UTR) of IAV M1 complementary RNA (cRNA) and down-regulate M1 expression at both of cRNA and protein level. They also showed that miR-let-7c inhibit IAV replication in A549 cells (Ma et al., 2012). These researches have shed light on host-derived miRNAs that inhibit the IAV replication. However, there is a crucial fact that IAV evolve more rapidly than host immune systems (Taubenberger and Morens, 2013). Indeed, the IAV has caused pandemic outbreaks several times (Gao et al., 2013; Novel Swine-Origin Influenza et al., 2009). Therefore, it is important issue to clarify whether or not regulatory relationships between host and IAVs via miRNAs have been evolutionarily conserved during the rapid evolution of IAV.

The aim of this study is to clarify the evolutionary relationships between host and IAVs via host miRNAs during the IAV's dynamic evolution. Evolutionary biology is a powerful approach to understand the evolutionary relationships between hosts and pathogens, thus we constructed large-scale evolutionary networks of IAVs by using a large amount of IAV sequence data. Large-scale evolutionary networks could detected both of dynamic mutations by interspecies-transmission processes and gradual mutations after transmitting to new hosts. Accordingly, we performed a conservation analysis of the relationships between host miRNAs and IAV. Entropy analysis proved that miRNA-target sites in PB1 and M1 mRNAs were significantly conserved in human, swine and avian IAVs. Meanwhile, miR-323 and miR-491 were conserved in human and swine genome, and miR-let-7c was conserved in human, swine, and avian genome. Our results suggest that regulatory relationships of host miRNAs have been evolutionarily conserved in a broad host range, even though IAV mutate rapidly.

2. Materials & Methods

2.1 Evolutionary analysis of influenza A virus

2.1.1 Sequence data

319,011 influenza A virus genomes during 1902-2013 were retrieved from the NCBI Influenza Virus Resource on September 14, 2013 (Bao et al., 2008). Excluding laboratory strains, only strains which have complete eight segments and isolated year information were selected. 20,072 complete strains (9,320 human, 1,837 swine and 8,915 avian IAVs) were used for this analysis.

2.1.2 Spectral Clustering for constructing evolutionary networks

We calculated the sequence similarity scores (Basic Local Alignment Search Tool [BLAST] bit scores) for all the influenza A virus PB1 proteins from 1902 to 2013 based on a round-robin BLASTP (BLAST 2.2.25) analysis with a cut off at E value ≤ 1e-5. The sequence similarity scores are defined as S'_{bit} (x, y) and indicate the bit score between the "database" sequence x and the "query" sequence y. The bit score were normalized according to the following equation:

$$sim(x,y) = \frac{max\,(bit\ score\,(x,y)\,,bit\ score\,(y,x))}{max\,(bit\ score\,(x,x)\,,bit\ score\,(y,y))} \cdots\cdots (1)$$

with $0 \leq \mathrm{Sim}(x, y) \leq 1$, where Sim ($x$, y) represent the normalized sequence similarity between two sequences x and y. Each Sim (x, y) value was then calculated against all pairs of PB1 proteins, and a weighted-undirected graph clustering algorithm using SCPS 0.9.5, and network graph was visualized from the clustering results by using Cytoscape 2.8.2 (Nepusz et al., 2010; Smoot et al., 2011).

2.2 Conservation analysis of target site by microRNA

2.2.1 Entropy analysis of influenza A virus mRNAs and conservation analysis of host miRNAs

In order to measure the variability of influenza A virus nucleotides, we used Shannon entropy methodology. Entropy $H(x)$ was calculated according to the following equation:

$$H = -\sum_{i=a,t,g,c} P_i \log P_i \cdots\cdots (2)$$

, where p (i) represent the probability of genome position (i). The entropy value increases with n (x), the total number of variants observed at position x, it is also sensitive to the relative frequency of the variants, such that it decreases when one variant is clearly dominant. Only sequences that contain a valid amino acid at position x were used for the entropy computation, and alignment gaps were ignored. Host miRNA sequences were obtained from miRBase, and were aligned by using ClustalW (Kozomara and Griffiths-Jones, 2011).

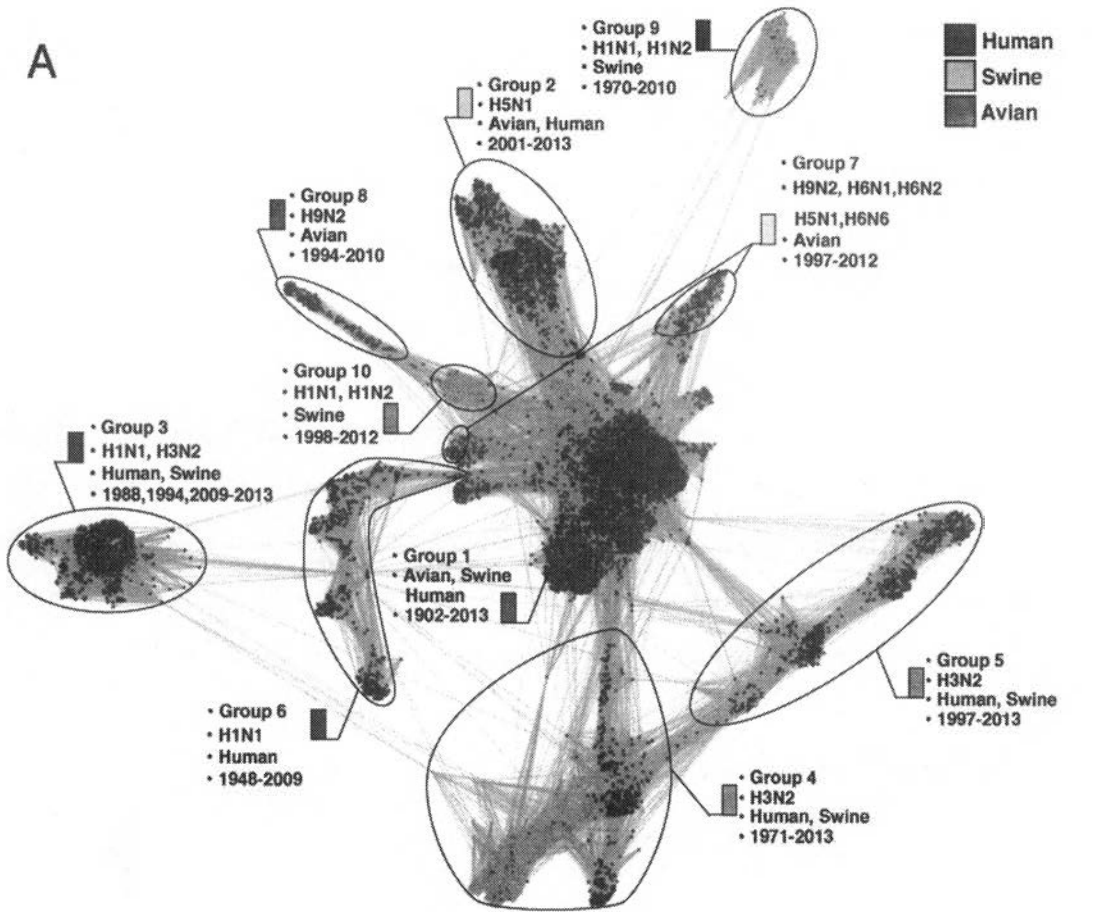
3. Results and Discussions

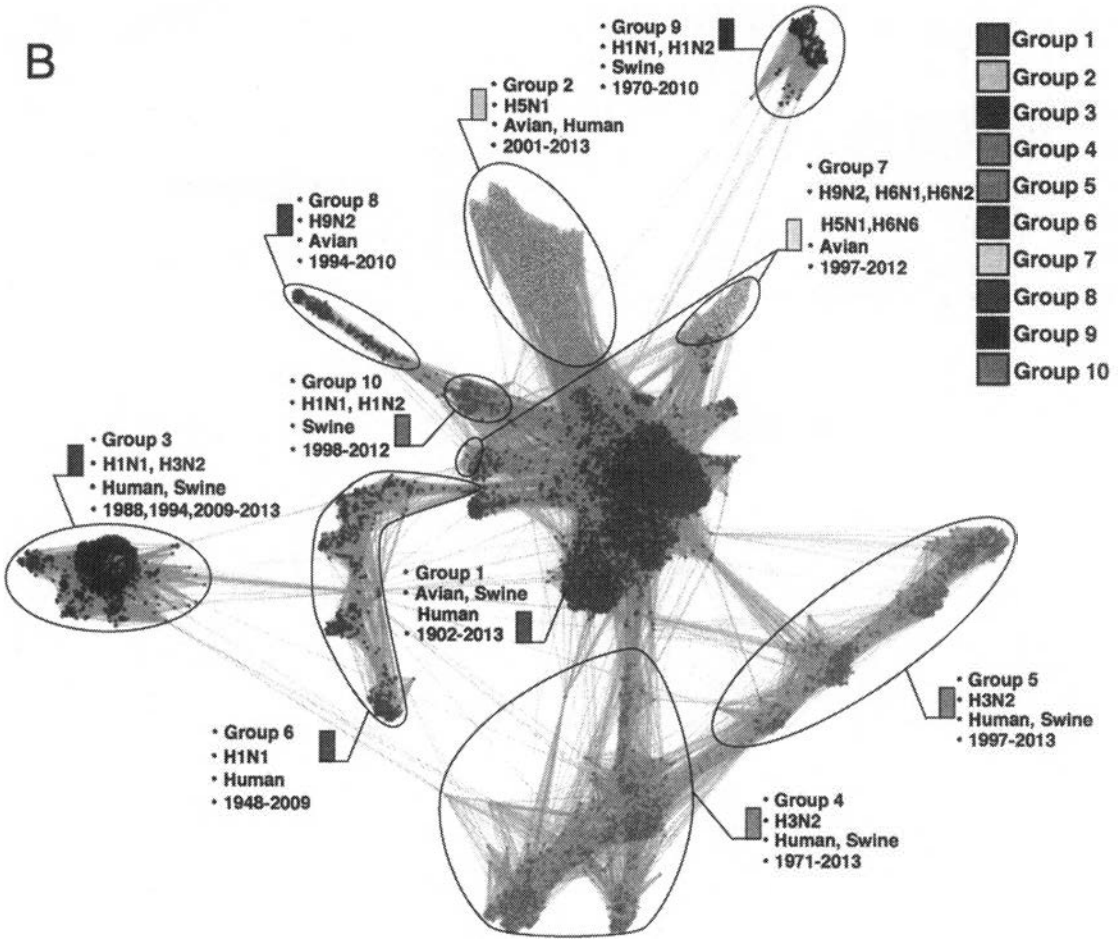
3.1 Evolutionary network of influenza A virus PB1 proteins during 1902-2013

In this study, a total of 20,072 complete strains of avian, swine and human IAVs were examined, and we constructed large-scale evolutionary networks of IAVs by spectral clustering methodology. These large-scale evolutionary networks are effective to understand the location of a specific IAV in the entire IAV evolution. Only evolutionary network of PB1 proteins is shown in the thesis digest **(Fig 1)**. PB1 is more highly conserved protein than external HA and NA protein. Therefore, the evolutionary network of IAV PB1 proteins is not divided. This network is classified into viral categories by the difference of host species (human, swine and avian) **(Fig 1 A)**. Especially, the most of avian IAVs located in the center of the network, whereas the groups of swine and human influenza viruses located peripherally and they extended in straight line individually, deriving from avian influenza viral gene pool. These straight evolutions are due to the high selective pressure when viruses adapt to mammalian hosts. Also, swine IAVs are more settled in contrast with human IAVs, which are due to the limited transmission caused by swine IAVs' inability to move to farther areas.

In Figure 1B, the network of PB1 proteins was classified into 10 groups, which implies that there are various different characters of PB1 proteins. It was reported that 2009 pandemic influenza PB1 proteins were derived from human H3N2 viruses via triple reassortment swine H1N2 and H1N1 viruses. In this network, Group 3 (2009 pandemic influenza) is connected with Group 4 (Triple reassortment swine flu), but they also have connections with Group 6 (Spanish flu and its progeny viruses). It was known that European avian-like swine (H1N1) occurred in reassortment with avian H9N2 viruses (Bi et al., 2010), and there are also detected the interactions between Group 10 (European avian-like swine) and Group 7 (avian H9N2 viruses) in this network.

Figure 1: Evolutionary network of influenza A virus PB1 proteins during 1902-2013. Total 20,072 amino acid sequences of influenza A virus PB1 protein were classified based on the sequence similarity. The dot symbol represents each PB1 protein of influenza A virus strain, and the line indicated the sequence similarity. (A) The network is colored by host species. (B) The network is colored by clusters, and is classified into 10 clusters: (1) consisting of avian flu during 1902-2013, (2) consisting of highly pathogenic H5N1 avian influenza viruses (HPAIVs) between 2001 and 2013, (3) mainly consisting of 2009 pandemic influenza, (4) consisting of human H3N2 between 1971 and 2011 (Hong Kong flu and its progeny viruses) and swine H3N2 between 1977 and 2013 (Triple reassortment swine flu), (5) consisting of human H3N2 between 1997 and 2012 (progeny viruses of Hong Kong flu), (6) consisting of human H1N1 between 1948 and 2009 (Spanish flu and its progeny viruses), (7) consisting of avian H9N2, H6N1, H6N2, H5N1, H6N6 during 1997-2012, (8) consisting of avian H9N2 between 1994 and 2010, (9) consisting of swine H1N1, H1N2 between 1970 and 2010 (Classical swine flu), (10) consisting of swine H1N1 and H1N2 between 1998 and 2012 (European avian-like swine flu).


A
• Group 9
• H1N1, H1N2
• Swine
• 1970-2010
Human
Swine
Avian
• Group 2
• H5N1
• Avian, Human
• 2001-2013
• Group 7
• H9N2, H6N1,H6N2
H5N1,H6N6
• Avian
• 1997-2012
• Group 8
• H9N2
• Avian
• 1994-2010
• Group 10
• H1N1, H1N2
• Swine
• 1998-2012
• Group 3
• H1N1, H3N2
• Human, Swine
• 1988,1994,2009-2013
• Group 1
• Avian, Swine
Human
• 1902-2013
• Group 5
• H3N2
• Human, Swine
• 1997-2013
• Group 6
• H1N1
• Human
• 1948-2009
• Group 4
• H3N2
• Human, Swine
• 1971-2013


B
• Group 9
• H1N1, H1N2
• Swine
• 1970-2010
Group 1
Group 2
Group 3
Group 4
Group 5
Group 6
Group 7
Group 8
Group 9
Group 10
• Group 2
• H5N1
• Avian, Human
• 2001-2013
• Group 7
• H9N2, H6N1,H6N2
H5N1,H6N6
• Avian
• 1997-2012
• Group 8
• H9N2
• Avian
• 1994-2010
• Group 10
• H1N1, H1N2
• Swine
• 1998-2012
• Group 3
• H1N1, H3N2
• Human, Swine
• 1988,1994,2009-2013
• Group 1
• Avian, Swine
Human
• 1902-2013
• Group 5
• H3N2
• Human, Swine
• 1997-2013
• Group 6
• H1N1
• Human
• 1948-2009
• Group 4
• H3N2
• Human, Swine
• 1971-2013

3.2 miRNA-target sites were conserved in influenza A viruses

To measure the conservation of miRNA-target sites by host miRNAs, we performed entropy calculation analysis of PB1 mRNAs based on Shannon's entropy. The Shannon's entropy values at each nucleotide position of avian, swine and human IAVs are shown in Figure 2. As a result of entropy calculation analysis, miRNA-target sites were highly conserved in human IAVs, and also conserved among swine and avian IAVs. Conservation analysis of host miRNAs revealed that two of three host miRNAs (miR-323, miR-491) are conserved among human and swine. However, avian species (*Gallus gallus, Taeniopygia guttata*) do not have three miRNAs (miR-323, miR-491 and miR-654). It might be caused by the shortage of the registration number in avian miRNAs. Indeed, 2,578 *Homo sapiens* miRNAs, 326 *Sus Scrofa* miRNA, 996 *Gallus gallus* and 334 *Taeniopygia guttata* are registered in miRBase (Version. 20).

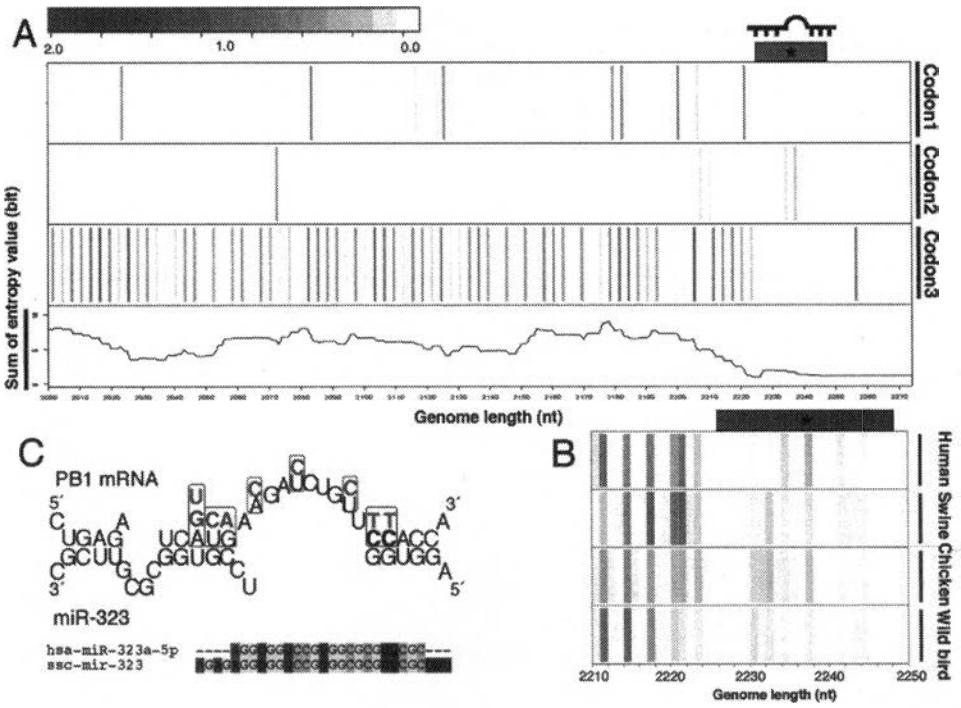


Figure 2: Conservation of miRNA-target site on influenza A virus PB1 mRNAs. (A) Entropy values were calculated based on Shannon entropy theory with 7,350 nucleotide sequences of human influenza A virus PB1 mRNAs. The x-axis represents the nucleotide position of influenza A virus genomes (nt). The red and white heat map indicated the entropy (bit) values which measured the variability of viral genomes at single position. The red rectangle represents the lower conserved sites and white ones indicated the highly conserved sites. The chart is divided into codon and sum of entropy values at the 30 nucleotides length. Host miRNA-targeting sites were enclosed in the star bar. (B) Entropy valued were calculated using 1,491 swine, 1,629 chicken, 5,993 wild birds-infecting influenza A virus PB1 mRNAs. (C) Binding pattern of host miR-323 and viral PB1 mRNA were predicted by using RNAhybrid, and Conservation of host miRNAs between human and swine.

3.3 Biological meanings of host microRNAs's regulation against influenza A viruses

These results raised up the a new question that why the regulatory relationships between host miRNAs and IAVs have been conserved during the extremely dynamic evolution of IAVs. Currently, we propose two contrasting hypotheses to this question; (1) host has targeted the immutable regions of IAVs. (2) IAVs utilize host miRNA's regulation in order to modulate the amount of viral replication.

IAVs are segmented viruses that possess eight-segmented genome in each viral particle. It was suggested that IAVs ensure correct packaging of the eight segments by viral RNA-RNA interactions (Gavazzi et al., 2013; Noda et al., 2012). The 3′ coding regions on PB1 mRNAs are part of packaging signal sites, and miRNA-target site also cover these packaging signal sites (Marsh et al., 2008). However, it is endurable for IAVs to put sequence substitutions in miRNA-target sites if genome-packaging mechanisms occur by RNA-RNA interactions. Because RNA-RNA interactions require only the conservation of second RNA structures, not sequence mutations. Thus, the reason of conservativeness of miRNA-target sites requires additional explanation.

It is commonly known that the mutation rate of IAVs is extremely higher than that of host. According to Darwinian evolution theory, populations that have resistance to unfavorable environments have been selected. Therefore, IAVs that are able to evade from host-miRNAs's regulations should be selected if host miRNAs inhibit IAV propagation. However, the regulatory relationships between host miRNA and IAVs have been evolutionary conserved for over one hundred years and beyond the host species. Considering these results, it is supportive that host miRNA's gene regulation is profitable for IAVs. Viruses cannot replicate and metabolite anymore without host cells, and there are always active of the trade-off relation of viral pathogenicity and its propagation rates during viral infection (Kerr and Best, 1998). Therefore, IAVs might have a strategy to make the most of the progeny, keeping an adequate number of host alive.

4. Conclusion

Elucidating of the revolutionary relationships between host and viruses by large-scale evolutionary network enable us to understand evolutionary history of IAVs clearly and also detect both of the dynamic and gradual mutations. This evolutionary approach revealed that regulatory relationships between host and IAVs via miRNA has been evolutionarily conserved for about a one hundred years in stead of extremely high mutation rates of IAVs, Furthermore, the evolutionary conservation relationships of host miRNAs conserved over the host species. Importantly, this study will provide us novel biological meanings of host miRNA's regulations against viruses and hints to deal with theses severe pathogens in the future.

Acknowledgements

I was truly lucky that I met Professor Masaru Tomita who taught me how interesting and exciting science can be, and this led me to the Institute for Advanced for Biosciences (IAB) when I was a freshman in college. I would like to show my appreciation for this wonderful opportunity he gave me and this study would not be possible without financial support of Yamagata Prefecture, Tsuruoka City and Keio University. At IAB, I was able to meet a lot of amazing people. I wish to express my sincere appreciation to my supervisor Professor Akio Kanai for his passionate guidance, enthusiasm, patience, and immense knowledge during these four years. He not only provided extensive guidance, but he also gave me considerable encouragement and treated me drinks & sweeties and joking. He is absolutely my best role model for a scientist, mentor, and teacher. Importantly, I believe this study could not be done without significant contributions from my two special advisors, Kahori T. Ikeda and Motomu Matsui. I had countless discussions with them and received their warming help. Kahori T. Ikeda was always there to help me anytime, anywhere. She is an independent and thoughtful woman, I always respect her. Motomu Matsui has been teaching me scientific thinking and computer programming skills since I was junior in college. I thought I wanted to be a scientist like him who is able to inspire other young scientists. I would like to thanks to Dr. Hitomi Sano, Dr. Kazuharu Arakawa, Dr. Yasuhiro Naito, Dr. Kenji Nakahigashi for providing exciting lectures and giving important comments to my research. On another note, I am grateful to my present and former RNA group members, Dr. Kosuke Fujishima, Dr. Yuka W. Iwasaki, Dr. Yoshiki Ikeda, Kiyofumi Hamashima, Yuka Hirose, Shinnosuke Murakami, Junnosuke Imai, Gakuto Makino, Asaki Kobayashi, Asako Sato and Emiko Noro for their constructive discussions. Shinnosuke Murakami was my first advisor, and he generously taught me the way to research. Dr. Yuka W. Iwasaki gave me opportunity to start this study. Dr. Kosuke Fujishima inspired me to decide to live as a scientist. And, it was greatest memory to attend 2013 RNA society conference at Davos, Switzerland. When I was sophomore, I was able to really enjoy Tsuruoka life thanks to biocamp members (Emmy Umeda, Hiroaki Abe, Guo Jing, Junnosuke Imai, Keita Kamezaki, Gembu Maryu, Hitomi Matsui, Mayuko Tsuruoka, Reina Fukuzawa, Sayaka Saito, Seiko Nakatsuka, Satoru Shirokaze, Sou Ishiguro and Yuka Sato). I thank my labmates (Jun Yamakubo, Riko Aoki, and Tamami Toki) for sharing memorable time with me. I was also able to make great memories with TTCK students (Arisa Matsuura, Chiharu Ishii, Hitoshi Iuchi, Kentaro Hayashi, Dr Nobuaki Kono, Satoshi Tamaki, Yuri Yamamoto, Yuriko Hasebe and Megumi Uetaki). I especially thank my great friend, Midori Kawasaki. She always listened to my worries and encouraged me. I also appreciate the support from my SFC secretary members (Ayumi Mikami, Ayami Mizukami, Miwa Hiramoto) and TTCK staff members (Akiko Shiozawa, Aya Saitoh, Katsunori Komatsu, Kazune Sugahara, Maki Oike, Nozomi Nakamura and Yoko Tsuchiya). Finally, I would like to express my sincere gratitude to my family for the immeasurable amount of support and love.

Reference

Bao, Y., Bolotov, P., Dernovoy, D., Kiryutin, B., Zaslavsky, L., Tatusova, T., Ostell, J., and Lipman, D. (2008). The influenza virus resource at the National Center for Biotechnology Information. Journal of virology *82*, 596-601.

Bi, Y., Fu, G., Chen, J., Peng, J., Sun, Y., Wang, J., Pu, J., Zhang, Y., Gao, H., Ma, G., *et al.* (2010). Novel swine influenza virus reassortants in pigs, China. Emerging infectious diseases *16*, 1162-1164.

Gao, R., Cao, B., Hu, Y., Feng, Z., Wang, D., Hu, W., Chen, J., Jie, Z., Qiu, H., Xu, K., *et al.* (2013). Human infection with a novel avian-origin influenza A (H7N9) virus. The New England journal of medicine *368*, 1888-1897.

Gavazzi, C., Isel, C., Fournier, E., Moules, V., Cavalier, A., Thomas, D., Lina, B., and Marquet, R. (2013). An in vitro network of intermolecular interactions between viral RNA segments of an avian H5N2 influenza A virus: comparison with a human H3N2 virus. Nucleic acids research *41*, 1241-1254.

Kerr, P., and Best, S. (1998). Myxoma virus in rabbits. Revue scientifique et technique (International Office of Epizootics) *17*, 256-268.

Kozomara, A., and Griffiths-Jones, S. (2011). miRBase: integrating microRNA annotation and deep-sequencing data. Nucleic acids research *39*, 7.

Lewis, B., Burge, C., and Bartel, D. (2005). Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell *120*, 15-20.

Ma, Y.-J., Yang, J., Fan, X.-L., Zhao, H.-B., Hu, W., Li, Z.-P., Yu, G.-C., Ding, X.-R., Wang, J.-Z., Bo, X.-C., *et al.* (2012). Cellular microRNA let-7c inhibits M1 protein expression of the H1N1 influenza A virus in infected human lung epithelial cells. Journal of cellular and molecular medicine *16*, 2539-2546.

Marsh, G., Rabadán, R., Levine, A., and Palese, P. (2008). Highly conserved regions of influenza a virus polymerase gene segments are critical for efficient viral RNA packaging. Journal of virology *82*, 2295-2304.

Nepusz, T., Sasidharan, R., and Paccanaro, A. (2010). SCPS: a fast implementation of a spectral method for detecting protein families on a genome-wide scale. BMC bioinformatics *11*, 120.

Noda, T., Sugita, Y., Aoyama, K., Hirase, A., Kawakami, E., Miyazawa, A., Sagara, H., and Kawaoka, Y. (2012). Three-dimensional analysis of ribonucleoprotein complexes in influenza A virus. Nature communications *3*, 639.

Novel Swine-Origin Influenza, A.V.I.T., Dawood, F., Jain, S., Finelli, L., Shaw, M., Lindstrom, S., Garten, R., Gubareva, L., Xu, X., Bridges, C., *et al.* (2009). Emergence of a novel swine-origin influenza A (H1N1) virus in humans. The New England journal of medicine *360*, 2605-2615.

Rosalind, C.L., Rhonda, L.F., and Victor, A. (1993). The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell *75*.

Smoot, M., Ono, K., Ruscheinski, J., Wang, P.-L., and Ideker, T. (2011). Cytoscape 2.8: new features for data integration and network visualization. Bioinformatics (Oxford, England) *27*, 431-432.

Song, L., Liu, H., Gao, S., Jiang, W., and Huang, W. (2010). Cellular microRNAs inhibit replication of the H1N1 influenza A virus in infected cells. Journal of virology *84*, 8849-8860.

Taubenberger, J., and Morens, D. (2013). Influenza viruses: breaking all the rules. mBio *4*.

Xiao, C., and Rajewsky, K. (2009). MicroRNA control in the immune system: basic principles. Cell *136*, 26-36.