BIOINFORMATION
Discovery at the interface of physical and biological sciences

open access www.bioinformation.net Hypothesis

Volume 7(3)

Palindromes drive the re-assortment in Influenza A

Abdullah Zubaer[1, 2*] & Simrika Thapa[1, 2]

[1]Swapnojaatra Bioresearch Laboratory, DataSoft Systems, Dhaka-1215, Bangladesh; [2]Shahjalal University of Science and Technology, Bangladesh; Abdullah Zubaer - Email: abdullah.zubaer@datasoft-bd.com; Phone: +880-2-9110169, Ext: 104; Fax: +880-29114747; *Corresponding Author

Received September 08, 2011; Accepted September 11, 2011; Published September 28, 2011

Abstract:
Different subtypes of Influenza A virus are associated with species specific, zoonotic or pandemic Influenza. The cause of its severity underlies in complicated evolution of its segmented RNA genome. Although genetic shift and genetic drift are well known in the evolution of this virus, we reported the significant role of unique RNA palindromes in its evolution. Our computational approach identified the existence of unique palindromes in each subtype of Influenza A virus with its absence in Influenza B relating the fact of virulence and vigorous genetic hitchhiking in Influenza A. The current study focused on the re-assortment event responsible for the emergence of pandemic-2009 H1N1 virus, which is associated with outgrow of new palindrome and in turn, changing its RNA structure. We hypothesize that the change in RNA structure due to the presence of palindrome facilitates the event of re-assortment in Influenza A. Thus the evolutionary process of Influenza A is much more complicated as previously known, and that has been demonstrated in this study.

Background:
Influenza A viruses are single-stranded RNA viruses of negative sense with an eight-segmented genome and belong to the family Orthomyxoviridae **[1]**. The hemagglutinin (HA) is an envelope glycoprotein in Influenza A that is responsible for the sialic acid binding as well as host recognition **[2]**. It is also a key antigen against which humoral immune responses are directed **[1]**, and hence the study on it is needed for understanding the biology, and the control and prevention of Influenza A viruses **[3]**. Influenza A viruses are sub-typed according to the reactivity of their surface antigens, haemagglutinin (HA) and neuraminidase (NA), combination of the HA and NA represents a subtype **[4]**. The emergence of new subtypes is driven by two mechanisms, i.e., re-assortment and mutation, termed also as genetic shift and genetic drift respectively **[5]**. Previous study showed that the re-assortment does the mixing of two genomes of Influenza A virus, during which HA and NA could come from different lineages. Besides, purifying selection helped them to adapt with host **[6]**. In some cases, positive selection has taken place that could be the cause of new subtypes **[7]**. Besides these concepts, we performed a computational approach to study the evolution of Influenza virus from a different point of view, by investigating the presence of palindromes in the gene of Influenza viruses, each from different subtypes. Although some previous statistical and experimental studies of palindromes in other classes of viral genomes, such as the double stranded DNA viruses, bacteriophages, retro viruses, etc., have been performed **[8-12]**, this is a novel approach to analyze the gene pattern and evolution of Influenza A. Prior study revealed that the hemagglutinin gene of pandemic-2009 H1N1 has directly come from the triple re-assortant H1N2 through re-assortment, and similarly that of triple re-assortant H1N2 came from the classical swine H1N1 through the same mechanism **[6]**. We have compared the palindrome as well as RNA structure of these segments. The proposition of this study would disclose a new dimension to detect the complicated evolution of Influenza A virus.

Methodology:
The nucleotide sequences of hemagglutinin (HA) of Influenza A and Influenza B have been retrieved from Influenza database at NCBI (http://www.ncbi.nlm.nih.gov/genomes/FLU/), as HA being the virulent factor for Influenza viruses **[2]**. HA sequences from each of the 12 highly virulent subtypes of Influenza A were downloaded. The classical swine H1N1 was included,

since it is involved in re-assortment, for the emergence of triple re-assortant H1N2 subtype from which the HA has been derived in the 2009-pandemic H1N1 **[6]**. To search for the presence of palindromic sequence in the HA gene, all retrieved sequences of 12 subtypes of Influenza A and 6 sequences of Influenza B were submitted to the program "Palindrome" from Mobyle Portal @ Pasteur **[13]** (a web-based site with repository of sequence and structural analysis). Palindromes of length more than or equal to 10 bases were only considered. Mismatch was not allowed in software parameter. The structural changes among the HA gene segments involved in re-assortment were identified using M-fold **[14]** program for the comparative folding pattern in the palindromic site, to analyze the influence of palindrome in the folding pattern of the RNA of the virus.

Discussion:

Palindromes in nucleic acid consist of nucleotide sequences that read the same from the 5'-end to the 3'-end **[15]**. The presence of palindrome in a gene sequence decreases the sequence entropy **[16]**. Thus the presence of palindrome in a gene sequence is evolutionary important. Search for palindrome results in the presence of one or more palindromes in each subtype of Influenza A as listed in **Table 1 (see Supplementary material)**. However, reports lack palindrome in Influenza B. Palindrome of each subtype is unique and not similar to other subtypes. The frequent occurrence of palindrome in Influenza A and nullification of palindromic occurence in Influenza B signify that palindrome is specific to Influenza A, which is much more dangerous than Influenza B with regard to disease causing ability. This also correlates the palindromic occurrence with the event of expedited gene exchange in Influenza A termed as re-assortment, which is a special mechanism of producing new subtype, since Influenza B lacks both palindromes and the re-assortment event. Thus we hypothesize that palindromes play a significant role in the re-assortment of Influenza A virus. The HA gene of pandemic-2009 A/H1N1 virus directly came from triple re-assortant virus H1N2, and H1N2 came from classical swine H1N1 through re-assortment **[6]**. The lineages from where HA came through re-assortment in the 2009 pandemic influenza virus were compared, which disclosed the loss and gain of new palindrome during re-assortment. HA gene of pandemic-2009 H1N1 was found to have two palindromes with 10 and 13 arm-size, H1N2 with only a 10 arm-sized palindrome, which is different from H1N1, but the 13-arm palindrome of 2009 H1N1 was similar to the palindrome of classical swine H1N1 (as marked in bold in **Table 1, see supplementary material**). This indicates HA from H1N1 has been mutated in such a way that it had lost its own palindrome and gained a new during re-assortment. HA segment of pandemic-2009 H1N1 regained the palindrome of classical swine H1N1, although that palindrome disappeared when it was in triple re-assortant H1N2 subtype. Interestingly in the second event of re-assortment during 2009, HA gene came from H1N2 directly, but mutated in a mysterious way so that it gained a new palindrome to that of H1N2, whereas regained the palindrome of classical swine H1N1. This disclosure suggested that the change in palindrome is prerequisite for the event of re-assortment. In addition, the lack of palindrome does not allow Influenza B to re-assort. Thus the palindromes have significant influence on the re-assortment event driving the evolution of Influenza A virus.

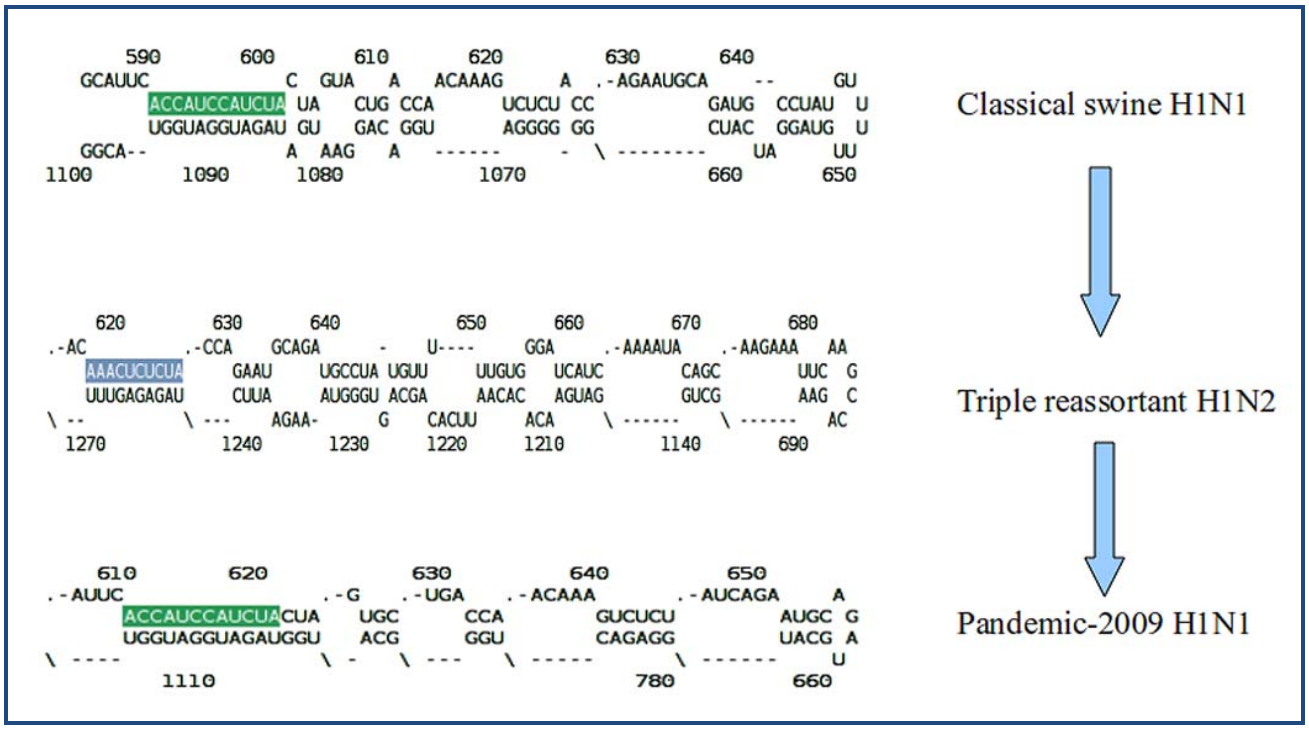

Figure 1: The change in the RNA folding due to the loss and gain of the palindrome in HA gene during the re-assortment event for the emergence of virulent pandemic virus. Classical swine A/H1N1 covers 590- 1100 of palindrome while folding, H1N2 folds at 620- 1270 position, and A/H1N1/2009 palindrome folds at 610- 1110. The folding position and folding pattern are different from each other.

Determination of folding pattern by Mfold revealed the influence of palindrome in RNA folding. Change in palindrome causes change in RNA structure by changing the fold length and pattern as shown in **Figure 1**. It was reported that there is no positive selection in HA of H1N1 rather purifying selection act on it **[6]**. Thus the mutations in HA have less effect on amino acid level, but on its own RNA gene structure. The change in gene structure, with the aid of palindrome, influences the molecular mechanism of replication, expression and other molecular signaling **[17-19]**. Influence of palindrome in RNA structure of Influenza gene could determine the rate of replication and gene expression, as well as viral packaging through which the re-assortment takes place.

Conclusion:

Influenza A viruses are evolving very frequently causing risks to human and several animals. The insight into their evolutionary mechanisms is necessary to prevent their disease causing ability. Here, the novel approach to find the relation between the palindrome in gene and the re-assortment process reveals the evolutionary mechanisms. This study concludes that palindromic change thus drives the event of genetic shift or re-assortment in Influenza A virus.

Acknowledgement:

We are thankful to Shakhinur Islam Mondal and Mohammad Lashkar for their valuable suggestions and supervision in this study.

References:

[1] Peiris JS *et al. J Clin Virol.* 2009 **45**: 169 [PMID: 19540800]
[2] Skehel JJ & Wiley DC. *Annu Rev Biochem.* 2000 **69**: 531 [PMID: 10966468]
[3] Zhang Y *et al. Biochem Biophys Res Commun.* 2009 **387**: 405 [PMID: 19615337]
[4] Hay AJ *et al. Philos Trans R Soc Lond B Biol Sci.* 2001 **356**: 1861 [PMID: 11779385]
[5] Holmes EC *et al. PloS Biol.* 2005 **3**: e300 [PMID: 16026181]
[6] Mondal SI *et al. J Microbiol Biotechnol.* 2010 **20**: 1500 [PMID: 21124053]
[7] Li W *et al. Virol J.* 2011 **8**: 183 [PMID: 21507270]
[8] Cain D *et al. J Virol.* 2001 **75**: 3731 [PMID: 11264362]
[9] Dirac AM *et al. J Gen Virol.* 2002 **83**: 2533 [PMID: 12237437]
[10] Hill MK *et al. J Virol.* 2003 **77**: 8329 [PMID: 12857902]
[11] Leung MY *et al. J Comput Biol.* 2005 **12**: 331 [PMID: 15857246]
[12] Rocha EP *et al. Genome Res.* 2001 **11**: 946 [PMID: 11381024]
[13] Rice P *et al. Trends Genet.* 2004 **16**: 276 [PMID: 10827456]
[14] Zuker M. *Nucleic Acids Res.* 2003 **31**: 3406 [PMID: 12824337]
[15] Giel-Pietraszuk M *et al. J Protein Chem.* 2003 **22**: 109 [PMID: 12760415]
[16] Rentzeperis D *et al. J Phys Chem B.* 2002 **106**: 9945
[17] Usdin K *et al. Genome Res.* 2008 **18**:1011 [PMID: 18593815]
[18] Verstrepen KJ *et al. Nat Genet.* 2005 **37**: 986 [PMID: 16086015]
[19] Andrade MA *et al. J Struct Biol.* 2001 **134**: 117 [PMID: 11551174]

Edited by P Kangueane
Citation: Zubaer & Thapa. Bioinformation 7(3): 115-119 (2011)

BIOINFORMATION

open access

Supplementary material:

Table 1: Palindromes in HA gene of Influenza A virus with their position and length. The palindromes in each subtype are unique except for 2009-pandemic H1N1 and classical swine H1N1 as shown in bold.

Subtype	Accession	Palindrome	Length
H1N1 (2009-pandemic)	CY083910	610 **accatccatcta**c 622 1116 **tggtaggtagat**g 1104	13
		618 tctactagtg 627 775 agatgatcac 766	10
H1N1 (classical swine)	EU139827	590 **accatccatcta** 601 1096 **tggtaggtagat** 1085	12
H1N2	AY233393	617 aaactctcta 626 1271 tttgagagat 1262	10
H2N2	CY021813	1089 atggcaagga 1098 1616 taccgttcct 1607	10
H3N2	CY026035	253 aatagatgct 262 755 ttatctacga 746	10
		358 tgtgccggat 367 406 acacggccta 397	10
H4N2	CY005955	788 gtcttcaaca 797 1306 cagaagttgt 1297	10
		1293 agtatgttga 1302 1369 tcatacaact 1360	10
H5N1	HM172104	545 ataccaacca 554 1101 tatggttggt 1092	10
		1002 aaatagtcctct 1013 1044 tttatcaggaga 1033	12
H7N2	AY240877	574 catcactctg 583 814 gtagtgagac 805	10
H7N3	CY015006	525 ggcatttcccca 536 964 ccgtaaaggggt 953	12
		595 catcactctgga 606 859 gtagtgagacct 848	12
H7N7	AY338459	188 ttcccaggat 197 347 aagggtccta 338	10

		449	caaccagtgc	458	10
			||||||||||		
		1454	gttggtcacg	1445	
H10N2	CY076269	358	atggaaagtgg	368	11
			|||||||||||		
		635	taccttcacc	625	
		414	caattcagct	423	10
			||||||||||		
		1314	gttaagtcga	1305	
H11N6	CY014679	313	acattgtggaa	323	11
			|||||||||||		
		961	tgtaacacctt	951	
H12N1	CY006006	1224	acaagcaatt	1233	10
			||||||||||		
		1344	tgttcgttaa	1335	
Influenza B	EU605942	Absent			
Influenza B	EU605944	Absent			
Influenza B	EU605945	Absent			
Influenza B	EU605943	Absent			
Influenza B	EU605941	Absent			
Influenza B	EU605946	Absent			