International Research Journal of Biochemistry and Bioinformatics (ISSN-2250-9941) Vol. 1(7) pp. 168-177, August, 2011
Available online http://www.interesjournals.org/IRJBB
Copyright © 2011 International Research Journals

Full Length Research Paper

Bioinformatics techniques for designing antiviral agents against influenza

Yadish Bukhari, Sahar Fazal* and Hamid Rashid

Department of Bioinformatics, Mohammad Ali Jinnah University, Islamabad, Pakistan.

Accepted 18 July, 2011

Influenza virus infection is responsible for hundreds of thousands of deaths annually. Current vaccination strategies and antiviral drugs provide limited protection; therefore new strategies are needed. RNA interference is an effective means and is now widely used to knock-down gene expression in a sequence specific manner making it a powerful tool not only for studying gene function, but also for therapeutic purposes, including antiviral treatment, by targeting highly divergent pathogens including human immuno-deficiency virus (HIV), hepatitus C virus (HCV), influenza virus and SARS caronavirus, all of which pose enormous threats to global human health. In the present work, potential siRNA are being designed against segment HA and NA of influenza virus using a novel web-based online software system, siVirus, which provides functional, off-target minimized siRNAs targeting highly conserved regions of divergent viral sequences.

Keywords: influenza, antiviral agents, virus

INTRODUCTION

Influenza, commonly known as flu, is an infectious disease of birds and mammals caused by an RNA virus of the family *Orthomyxoviridae* (the influenza viruses). Also known as flu which is a contagious disease that is caused by the influenza virus. It attacks the respiratory tract in human (Carolyn *et al*., 2003). Acute viral infection of the respiratory tract, occurring in isolated cases, in epidemics, or in pandemics; is caused by serologically different strains of influenza viruses designated as A, B, and C, has a 3-day incubation period, and usually lasts for 3 to 10 days (George E *et al.,* 2001). There are three genera of influenza virus, identified by antigenic differences in their nucleoprotein and matrix protein:

- *Influenzavirus A* are the cause of all flu pandemics and are known to infect humans, other mammals and birds (avian influenza).
- *Influenzavirus B* are known to infect humans and seals.
- *Influenzavirus C* are known to infect humans and pigs.

The type A viruses are the most virulent human pathogens among the three influenza types and causes the most severe disease (Hay *et al.,* 2001). Influenza A viruses are further divided into subtypes based on the surface glycoproteins, hemagglutinin (HA) and neuraminidase (NA). At present, 16 HA subtype (H1–H16) and nine NA subtypes (N1–N9) have been recognized. Each virus has one HA and one NA antigen, apparently in any combination. (Suzuki., 2005). Influenza virus mutations are subtyped specific. siRNAs have a well-defined structure: a short (usually 21-nucleotide-long) double-strand of RNA (dsRNA) with 2-nucleotide 3' overhangs on either end. Each strand has a 5' phosphate group and a 3' hydroxyl (-OH) group. siRNAs can also be introduced into cells using nanoparticles and other methods to bring about the specific knockdown of a gene of interest. Essentially any gene of whose the sequence is known can thus be targeted based on sequence complementarily with an appropriately tailored siRNA. This method has made siRNAs an important tool for gene function and drug target validation studies in the post-genomic era (Qing G *et al.,* 2004). Bioinformatics approaches are significant to come up with intelligent solutions to existing health related issues e.g. designing molecules against virus infection. siRNA's can be used for gene silencing or for therapeutic purposes like vaccines (David B *et al.,* 2006).

*Corresponding author E-mail: sahar@jinnah.edu.pk

MATERIAL AND METHODS

Sequence data for gene segments 4 (hemagglutinin) and 6 (neuraminidase) of influenza A virus were downloaded from GenBank. To facilitate the computational analysis of these background datasets, those isolates with identical sequences were removed from the analysis. siVirus interface (http://**sivirus**.rnai.jp/.html) was used to obtain results. siVirus especially focuses on antiviral siRNA design and provides:

- Highly conserved target sites for designing antiviral siRNAs that would resist viral mutational escapes. siRNA sequences are selected based on their *degree of conservation*, defined as the proportion of viral sequences that are targeted by the corresponding siRNA, wfigureith complete matches (*i.e.*, 21/21 matches). All possible siRNA candidates targeting every other position of user-selected viral sequences are generated and their degrees of conservation are compared. Users can arbitrarily specify a set of viral sequences for the computation; for example, sequences can be selected from a specific geographic region(s) or a specific genotype(s) to design the best siRNAs tailored to specific user needs.
- Selection of Effective siRNAs based on guidelines of Ui-Tei *et al.*, Reynolds *et al.*, and Amarzguioui *et al.*

Those guidelines are:

1. A/U at the 5' end of the antisense strand
2. G/C at the 5' end of the sense strand
3. Atleast 5 A/U residues in the 5' terminal, one third of the antisense strand
4. The absence of any GC stretch of more than 9 nucleotide in length

If any of the above four guideline is not fulfilled than there is no gene silencing by a particular siRNA (Amarzguoui *et al.,* 2004).

Off-target minimized siRNAs. siVirus shows the number of off-target hits within two mismatches against human genes. Off-target hits are those which have dissimilarity to the target sequence. It is desirable to select a siRNA that has less off-target hits.

RESULTS

siRNA results with siVIRUS

A quite large number of possible siRNA candidates (i.e. substrings of length 21) targeting every other position of the Influenza virus A sequences were produced from the retrieved viral sequences. The analysis of these siRNA candidates revealed that highly conserved siRNAs constituted only 0.0% of the possible siRNAs if >90% conservation is expected. But the fraction increases to 100% in case of Hemmaglutinin segment of Influenza A virus even if the threshold of the conservation is relaxed to 80%. On the other hand, siRNAs predicted to be functional by one or more guidelines as mentioned in material and methods constituted in case of all the HA sequences present in siVirus interface 65.35%, in case of region specific HA sequences 51.16%, in case of all the NA sequences 84.08% and in case of region specific NA sequences 78.78% of the siRNAs (Figure 1, 2, 3, and 4). Taken together, siRNAs showing maximum conservation are given in their duplex form. In this condition, siRNAs can be designed for each full-length sequence of segment HA and NA of Influenza virus A. Graphs in figure 1, 2, 3 and 4 shows probability of siRNAs fulfilling at least one guideline as mentioned in material and methods.

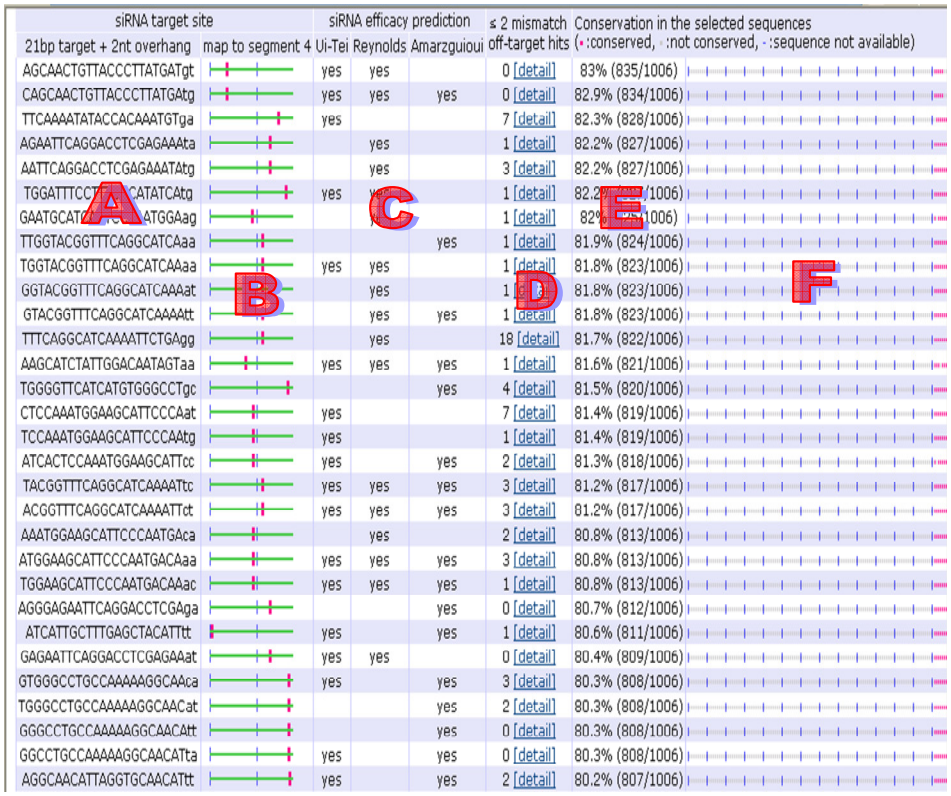
siRNA results with siDIRECT

siRNA duplex shows the double stranded sequence that is sense and anti-sense strands of siRNA. Oligo sequences means short sequences as here the length of sequence is 21bp. *Mismatch tolerance* describes the minimum number of mismatches between the siRNA sequence and any non-targeted sequence. An siRNA sequence of mismatch tolerance three does not match any off-target candidates with fewer than three mismatches. A higher mismatch tolerance of siRNA sequence indicates its high specificity in the presence of some mismatches.

DISCUSSION

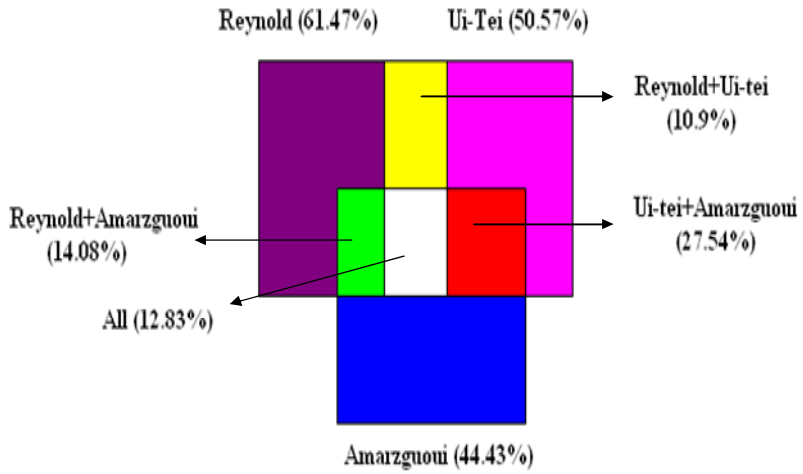
Influenza virus infection is a major public health problem, causing millions of cases of severe illness and as many as 500,000 deaths each year world wide. Factors contributing to the pathogenecity of influenza in human are not fully understood. Clearly the ability of the influenza virus to evolve rapidly and reassort contributes to the potential for epidemics and pandemics and may also contribute to the severity of disease.The restriction of influenza to the respiratory tract in human diferentiates the disease from that in bird, in which replication is normally restricted to the GI tract but pathogenic strains can cause systemic infections.The HA and NA of the avian viruses has a crucial role in pathogenecity (Taubenberger *et al.*, 2000).

In start the scourge of Influenza has probably affected mankind for several millenia and is feared because of its dramatic impact during pandemics (Micheal *et al.,* 2001). Over the past few years novel class of antiviral agents, the neuraminadase inhibitors has been found to be safe and effective in the prevention and treatment of influenza. But they too provide partial protection. Previously available agents, the M2 inhibitors amantadine and rimantadine, could only be used to treat



siRNA target site		siRNA efficacy prediction			≤ 2 mismatch	Conservation in the selected sequences
21bp target + 2nt overhang	map to segment 4	Ui-Tei	Reynolds	Amarzguioui	off-target hits	(•:conserved, ◦:not conserved, -:sequence not available)
AGCAACTGTTACCCTTATGATgt		yes	yes		0 [detail]	83% (835/1006)
CAGCAACTGTTACCCTTATGAtg		yes	yes	yes	0 [detail]	82.9% (834/1006)
TTCAAAATATACCACAAATGTga		yes			7 [detail]	82.3% (828/1006)
AGAATTCAGGACCTCGAGAAAta			yes		1 [detail]	82.2% (827/1006)
AATTCAGGACCTCGAGAAATAtg			yes		3 [detail]	82.2% (827/1006)
TGGATTTCCT[illegible]CATATCAtg		yes	[illegible]		1 [detail]	82.2% [illegible]/1006)
GAATGCAT[illegible]ATGGAag			[illegible]		1 [detail]	82% [illegible]1006)
TTGGTACGGTTTCAGGCATCAaa				yes	1 [detail]	81.9% (824/1006)
TGGTACGGTTTCAGGCATCAAaa		yes	yes		1 [detail]	81.8% (823/1006)
GGTACGGTTTCAGGCATCAAAat			yes		1 [detail]	81.8% (823/1006)
GTACGGTTTCAGGCATCAAAAtt			yes	yes	1 [detail]	81.8% (823/1006)
TTTCAGGCATCAAAATTCTGAgg			yes		18 [detail]	81.7% (822/1006)
AAGCATCTATTGGACAATAGTaa		yes	yes	yes	1 [detail]	81.6% (821/1006)
TGGGGTTCATCATGTGGGCCTgc				yes	4 [detail]	81.5% (820/1006)
CTCCAAATGGAAGCATTCCCAat		yes			7 [detail]	81.4% (819/1006)
TCCAAATGGAAGCATTCCCAAtg		yes			1 [detail]	81.4% (819/1006)
ATCACTCCAAATGGAAGCATTcc		yes		yes	2 [detail]	81.3% (818/1006)
TACGGTTTCAGGCATCAAAATtc		yes	yes	yes	3 [detail]	81.2% (817/1006)
ACGGTTTCAGGCATCAAAATTct		yes	yes	yes	3 [detail]	81.2% (817/1006)
AAATGGAAGCATTCCCAATGAca			yes		2 [detail]	80.8% (813/1006)
ATGGAAGCATTCCCAATGACAaa		yes	yes	yes	3 [detail]	80.8% (813/1006)
TGGAAGCATTCCCAATGACAAac		yes	yes	yes	1 [detail]	80.8% (813/1006)
AGGGAGAATTCAGGACCTCGAga				yes	0 [detail]	80.7% (812/1006)
ATCATTGCTTTGAGCTACATTtt		yes		yes	1 [detail]	80.6% (811/1006)
GAGAATTCAGGACCTCGAGAAat		yes	yes		0 [detail]	80.4% (809/1006)
GTGGGCCTGCCAAAAAGGCAAca		yes		yes	3 [detail]	80.3% (808/1006)
TGGGCCTGCCAAAAAGGCAACat				yes	2 [detail]	80.3% (808/1006)
GGGCCTGCCAAAAAGGCAACAtt				yes	0 [detail]	80.3% (808/1006)
GGCCTGCCAAAAAGGCAACATta		yes		yes	0 [detail]	80.3% (808/1006)
AGGCAACATTAGGTGCAACATtt		yes		yes	2 [detail]	80.2% (807/1006)

Figure 1: siRNAs for Hemagglutinin
A = siRNA sequence.
B = siRNA position.
C = siRNA efficacy predictions.
D = Off-target search results.
E = Degrees of conservation.
F = Conservation in the user-selected sequences.

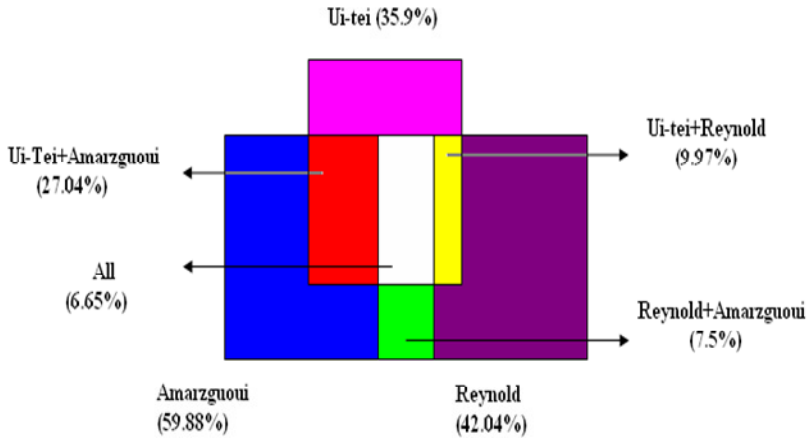


Graph 1: Probability of siRNA for Hemagglutinin

- Amarzguoui
- Ui-Tei
- Reynold
- Reynold + Amarzguoui
- Ui-Tei + Amarzguoui
- Reynold + Ui-Tei

siRNA target site		siRNA efficacy prediction			≤ 2 mismatch	Conservation in the selected sequences
21bp target + 2nt overhang	map to segment 4	Ui-Tei	Reynolds	Amarzguioui	off-target hits	(•:conserved, ·:not conserved, -:sequence not available)
CAGCAACTGTTACCCTTATGAtg		yes	yes	yes	0 [detail]	72.2% (86/119)
AGCAACTGTTACCCTTATGATgt		yes	yes		0 [detail]	72.2% (86/119)
TGGTACGGTTTCAGGCATCAAaa		yes	yes		1 [detail]	72.2% (86/119)
GGTACGGTTTCAGGCATCAAAat			yes		1 [detail]	72.2% (86/119)
GTACGGTTTCAGGCATCAAAAtt			yes	yes	1 [detail]	72.2% (86/119)
TACGGTTTCAGGCATCAAAATtc		yes	yes	yes	3 [detail]	72.2% (86/119)
ACGGTTTCAGGCATCAAAATTct		yes	yes	yes	3 [detail]	72.2% (86/119)
TTTCAGGCATCAAAATTCTGAgg			yes		18 [detail]	72.2% (86/119)
ATCATTGCTTTGAGCTACATTtt		yes		yes	1 [detail]	71.4% (85/119)
TTCTCAGAAGTAGAAGGGAGAat				yes	7 [detail]	71.4% (85/119)
CTCAGAAGTAGAAGGGAGAATtc		yes		yes	4 [detail]	71.4% (85/119)
TGGATTTCCTTTGCCATATCAtg		yes	yes		1 [detail]	71.4% (85/119)
TATGACCATGATGTATACAGAga			yes		4 [detail]	71.4% (85/119)
TGACCATGATGTATACAGAGAtg			yes		2 [detail]	71.4% (85/119)
AGGGAGAATTCAGGACCTCGAga				yes	0 [detail]	69.7% (83/119)
AGAATTCAGGACCTCGAGAAAta			yes		1 [detail]	69.7% (83/119)
AATTCAGGACCTCGAGAAATAtg			yes		3 [detail]	69.7% (83/119)
GGCAACATTAGGTGCAACATTtg		yes	yes	yes	1 [detail]	69.7% (83/119)
GAATGCATCACTCCAAATGGAag			yes		1 [detail]	69.7% (83/119)
GAGAATTCAGGACCTCGAGAAat		yes	yes		0 [detail]	68.9% (82/119)
AGGCAACATTAGGTGCAACATTt		yes		yes	2 [detail]	68.9% (82/119)
TTAGGTGCAACATTTGCATTTga				yes	2 [detail]	68.9% (82/119)
TCTGAATGCATCACTCCAAATgg			yes		3 [detail]	68.9% (82/119)
CTCCAAATGGAAGCATTCCCAat		yes			7 [detail]	68% (81/119)
TCCAAATGGAAGCATTCCCAAtg		yes			1 [detail]	68% (81/119)
AAATGGAAGCATTCCCAATGAca			yes		2 [detail]	68% (81/119)
TTCAAAATATACCACAAATGTga		yes			7 [detail]	68% (81/119)
AGGTGCAACATTTGCATTTGAgt		yes	yes	yes	2 [detail]	68% (81/119)
ATGGAAGCATTCCCAATGACAaa		yes	yes	yes	3 [detail]	67.2% (80/119)
TGGAAGCATTCCCAATGACAAac		yes	yes	yes	1 [detail]	67.2% (80/119)

Figure 2: siRNAs for Asian and Russian Hemagglutinin



Graph 2: Probability of siRNA for Asian and Russian Hemagglutinin
Graph 2: Probability of siRNA for Asian and Russian Hemagglutinin
siRNA results with siDIRECT for Hemagglutinin
1. siRNA Information for 83% conservation:
siRNA duplex
5'-CAACUGUUACCCUUAUGAUGU sense strand
UCGUUGACAAUGGGAAUACUA-5' antisense strand
Oligo sequences
>Sense_strand
CAACUGUUACCCUUAUGAUGU
>Antisense_strand
AUCAUAAGGGUAACAGUUGCU
Mismatch tolerance
Both-strand mismatch tolerance: 3Plus-strand mismatch tolerance: 3
2. siRNA Information for 82.9% conservation:
siRNA duplex
5'-GCAACUGUUACCCUUAUGAUG sense strand
GUCGUUGACAAUGGGAAUACU-5' antisense strand

Oligo sequences
>Sense_strand
GCAACUGUUACCCUUAUGAUG
>Antisense_strand
UCAUAAGGGUAACAGUUGCUG
Mismatch tolerance
Both- strandmismatch tolerance:
Plus-strand mismatch tolerance: 3
3. siRNA Information for 82.3% conservation:
siRNA duplex

```
5'-CAAAAUAUACCACAAAUGUGA      sense strand
   : : : : : : : : : : : : : : : : : : : :
AAGUUUUAUAUGGUGUUUACA-5'      antisense strand
```

Oligo sequences
>Sense_strand
CAAAAUAUACCACAAAUGUGA
>Antisense_strand
ACAUUUGUGGUAUAUUUUGAA
Mismatch tolerance
Both-strand mismatch tolerance: 2 Plus-strand mismatch tolerance: 2
4. siRNA Information for 82.2% conservation:
i. siRNA duplex

```
5'-AAUUCAGGACCUCGAGAAAUA      sense strand
   : : : : : : : : : : : : : : : : : : : :
UCUUAAGUCCUGGAGCUCUUU-5'      antisense strand
```

Oligo sequences
>Sense_strand
AAUUCAGGACCUCGAGAAAUA
>Antisense_strand
UUUCUCGAGGUCCUGAAUUCU
Mismatch tolerance
Both-strand mismatch tolerance: 3 Plus-strand mismatch tolerance: 3
ii. siRNA duplex

```
5'-UUCAGGACCUCGAGAAAUAUG      sense strand
   : : : : : : : : : : : : : : : : : : : :
UUAAGUCCUGGAGCUCUUUAU-5'      antisense strand
```

Oligo sequences
>Sense_strand
UUCAGGACCUCGAGAAAUAUG
>antisense_strand
UAUUUCUCGAGGUCCUGAAUU
Mismatch tolerance
Both-strand mismatch tolerance: 2 Plus-strand mismatch tolerance: 2
iii. siRNA duplex

```
5'-GAUUUCCUUUGCCAUAUCAUG      sense strand
   : : : : : : : : : : : : : : : : : : : :
ACCUAAAGGAAACGGUAUAGU-5'      antisense strand
```

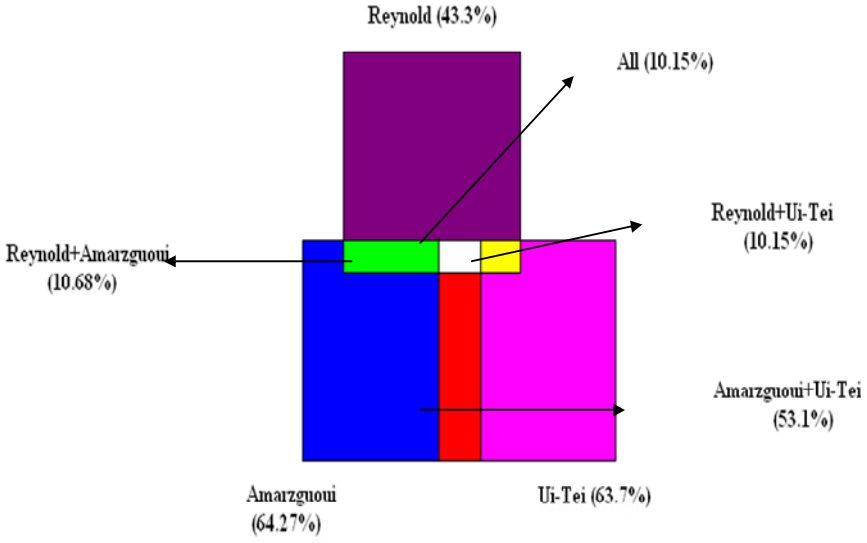
Oligo sequences
>Sense_strand
GAUUUCCUUUGCCAUAUCAUG
>antisense_strand
UGAUAUGGCAAAGGAAAUCCA
Mismatch tolerance
Both-strand mismatch tolerance:3 Plus-strand mismatch tolerance: 3
5. siRNA Information for 82% conservation:
siRNA duplex

```
5'-AUGCAUCACUCCAAAUGGAAG      sense strand
   : : : : : : : : : : : : : : : : : : : :
CUUACGUAGUGAGGUUUACCU-5'      antisense strand
```

Oligo sequences
>Sense_strand
AUGCAUCACUCCAAAUGGAAG
>antisense_strand
UCCAUUUGGAGUGAUGCAUUC
Mismatch tolerance
Both-strand mismatch tolerance: 3 Plus-strand mismatch tolerance: 3

siRNA target site		siRNA efficacy prediction			≤ 2 mismatch	Conservation in the selected sequences
21bp target + 2nt overhang	map to segment 6	Ui-Tei	Reynolds	Amarzguioui	off-target hits	(+:conserved, -:not conserved, -:sequence not available)
GTGGACCTCAAACAGTATTGTtg		yes		yes	0 [detail]	85.9% (1123/1307)
TGGACCTCAAACAGTATTGTTgt		yes		yes	1 [detail]	85.8% (1122/1307)
GACCTCAAACAGTATTGTTGTgt		yes		yes	2 [detail]	85.2% (1114/1307)
TAGCATGGTCCAGCTCAAGTTgt		yes	yes	yes	2 [detail]	85% (1112/1307)
ATGGTCCAGCTCAAGTTGTCAcg		yes		yes	4 [detail]	85% (1112/1307)
CTCAAACAGTATTGTTGTGTTtt		yes		yes	2 [detail]	84.9% (1110/1307)
TCCAGCTCAAGTTGTCACGATgg		yes		yes	1 [detail]	84.3% (1103/1307)
AACAGTATTGTTGTGTTTTGTgg		yes		yes	9 [detail]	84.3% (1103/1307)
TCAAACAGTATTGTTGTGTTTtg			yes		6 [detail]	84.3% (1102/1307)
CTCAAGTTGTCACGATGGAAAag		yes	yes	yes	0 [detail]	84.1% (1100/1307)
AGCTCAAGTTGTCACGATGGAaa			yes		2 [detail]	84% (1099/1307)
TCAAGTTGTCACGATGGAAAAgc			yes		0 [detail]	84% (1099/1307)
TGCAACTGCTAGCTTCATTTAtg		yes	yes	yes	0 [detail]	83.3% (1090/1307)
AATGCAACTGCTAGCTTCATTta			yes		3 [detail]	83.3% (1089/1307)
ATGCAACTGCTAGCTTCATTTat		yes	yes	yes	0 [detail]	83.3% (1089/1307)
AGTTGTCACGATGGAAAAGCAtg			yes		2 [detail]	82.6% (1080/1307)
ATGACGTGTGGATGGGAAGAAca		yes		yes	2 [detail]	82.3% (1076/1307)
GAACCTTATGTGTCATGCGATcc				yes	0 [detail]	82.1% (1074/1307)
TGCTTTTATGTGGAGTTGATAag		yes		yes	1 [detail]	82.1% (1074/1307)
ATGTGGAGTTGATAAGGGGAAgg				yes	1 [detail]	82.1% (1074/1307)
GTGCTTTTATGTGGAGTTGATaa		yes	yes	yes	2 [detail]	82% (1073/1307)
AATCCAAATCAAAAGATAATAac				yes	5 [detail]	81.5% (1066/1307)
ATCCAAATCAAAAGATAATAAca		yes	yes	yes	6 [detail]	81.4% (1064/1307)
GAAAAGCATGGCTGCATGTTgt			yes		6 [detail]	81.3% (1063/1307)
TGGAAAAGCATGGCTGCATGTtt		yes			1 [detail]	81.2% (1062/1307)
ATTACAGGATTTGCACCTTTTtc			yes		2 [detail]	81.1% (1061/1307)
TTACAGGATTTGCACCTTTTTct			yes	yes	1 [detail]	81.1% (1060/1307)
ACGTGTGGATGGGAAGAACGAtc		yes			0 [detail]	80.8% (1057/1307)
AACATAACAGAGATAGTGTATTt		yes		yes	4 [detail]	80.7% (1056/1307)
ATTTGCACCTTTTTCTAAGGAca			yes		3 [detail]	80.2% (1049/1307)

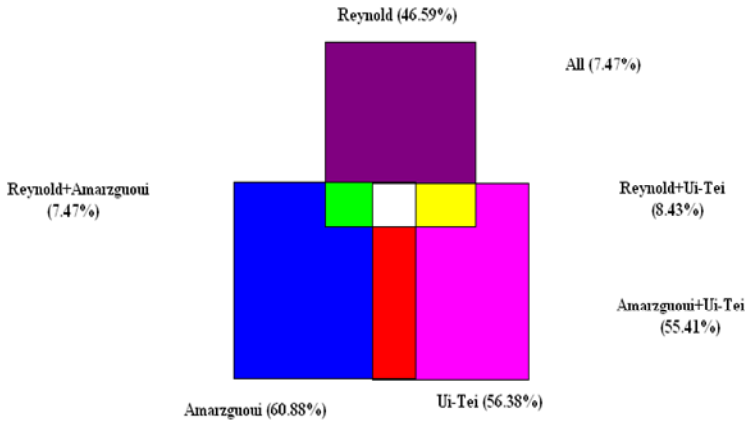
Figure 3: siRNAs for Neuraminidase



Graph 3. Probability of siRna for Neuraminidase

siRNA target site		siRNA efficacy prediction			≤ 2 mismatch	Conservation in the selected sequences
21bp target + 2nt overhang	map to segment 6	Ui-Tei	Reynolds	Amarzguioui	off-target hits	(*:conserved, -:not conserved, .:sequence not available)
AATCCAAATCAAAAGATAATAac				yes	5 [detail]	75.3% (122/162)
ATGAATCCAAATCAAAAGATAat		yes	yes	yes	8 [detail]	74% (120/162)
GAATCCAAATCAAAAGATAATaa				yes	6 [detail]	74% (120/162)
ATCCAAATCAAAAGATAATAAca		yes	yes	yes	6 [detail]	74% (120/162)
TCAAACAGTATTGTTGTTTTtg			yes		6 [detail]	74% (120/162)
AACAGTATTGTTGTGTTTTGTgg		yes		yes	9 [detail]	74% (120/162)
ATGGTCCAGCTCAAGTTGTCAcg		yes		yes	4 [detail]	73.4% (119/162)
TCCAGCTCAAGTTGTCACGATgg		yes		yes	1 [detail]	73.4% (119/162)
GTGGACCTCAAACAGTATTGTtg		yes		yes	0 [detail]	73.4% (119/162)
TGGACCTCAAACAGTATTGTTgt		yes		yes	1 [detail]	73.4% (119/162)
GACCTCAAACAGTATTGTTGTgt		yes		yes	2 [detail]	73.4% (119/162)
CTCAAACAGTATTGTTGTGTTtt		yes		yes	2 [detail]	73.4% (119/162)
TAGCATGGTCCAGCTCAAGTTgt		yes	yes	yes	2 [detail]	72.8% (118/162)
GGAGTGAAAGGCTGGGCCTTTga				yes	2 [detail]	72.8% (118/162)
CTCAAGTTGTCACGATGGAAAag		yes	yes	yes	0 [detail]	72.2% (117/162)
TCAAGTTGTCACGATGGAAAAgc			yes		0 [detail]	72.2% (117/162)
AGTTGTCACGATGGAAAAGCAtg			yes		2 [detail]	72.2% (117/162)
AGTGAAAGGCTGGGCCTTTGAca			yes		4 [detail]	72.2% (117/162)
ATGAATGAGTTGGGTGTTCCAtt		yes	yes		1 [detail]	71.6% (116/162)
AATGAGTTGGGTGTTCCATTTca			yes		2 [detail]	71.6% (116/162)
TGAGTTGGGTGTTCCATTTCAtt			yes		1 [detail]	71.6% (116/162)
AGCTCAAGTTGTCACGATGGAaa			yes		2 [detail]	71.6% (116/162)
TGCAACTGCTAGCTTCATTTAtg		yes	yes	yes	0 [detail]	71.6% (116/162)
AACATAACAGAGATAGTGTATtt		yes		yes	4 [detail]	70.9% (115/162)
ATTTGCACCTTTTTCTAAGGAca			yes		3 [detail]	70.9% (115/162)
AATGCAACTGCTAGCTTCATTta			yes		3 [detail]	70.9% (115/162)
ATGCAACTGCTAGCTTCATTTat		yes	yes	yes	0 [detail]	70.9% (115/162)
TTGCACCTTTTTCTAAGGACAat		yes	yes	yes	3 [detail]	70.3% (114/162)
TGCACCTTTTTCTAAGGACAAtc		yes		yes	2 [detail]	70.3% (114/162)

Figure 4: siRNAs for Asian and Russian Neuraminidase



Graph 4: Probability of siRNA for Asian and Russian Neuraminidase

- **siRNA results with siDIRECT for Neuraminidase**

1. siRNA Information for 85.9% conservation:

siRNA duplex

```
5'-GGACCUCAAACAGUAUUGUUG      sense strand
   : : : : : : : : : : : : : : : : : : : :
CACCUGGAGUUUGUCAUAACA-5'      antisense strand
```

Oligo sequences

>Sense_strand
GGACCUCAAACAGUAUUGUUG
>Antisense_strand
ACAAUACUGUUUGAGGUCCAC

Mismatch tolerance

Both-strand mismatch tolerance: 3 Plus-strand mismatch tolerance: 3

2. siRNA Information for 85.8% conservation:
siRNA duplex

```
5'-GACCUCAAACAGUAUUGUUGU      sense strand
   : : : : : : : : : : : : : : : : : : : : :
ACCUGGAGUUUGUCAUAACAA-5'      antisense strand
```

Oligo sequences
>Sense_strand
GACCUCAAACAGUAUUGUUGU
>Antisense_strand
AACAAUACUGUUUGAGGUCCA
Mismatch tolerance
Both-strand mismatch tolerance: 3 Plus-strand mismatch tolerance: 3
2. siRNA Information for 85.2% conservation:
siRNA duplex

```
5'-CCUCAAACAGUAUUGUUGUGU      sense strand
   : : : : : : : : : : : : : : : : : : : :
CUGGAGUUUGUCAUAACAACA-5'      antisense strand
```

Oligo sequences
>Sense_strand
CCUCAAACAGUAUUGUUGUGU
>antisense_strand
ACAACAAUACUGUUUGAGGUC
Mismatch tolerance
Both-strand mismatch tolerance: 2 Plus-strand mismatch tolerance: 3
4. siRNA Information for 85% conservation:
i. siRNA duplex

```
5'-GCAUGGUCCAGCUCAAGUUGU      sense strand
   : : : : : : : : : : : : : : : : : : : : :
AUCGUACCAGGUCGAGUUCAA-5'      antisense strand
```

Oligo sequences
>Sense_strand
GCAUGGUCCAGCUCAAGUUGU
>Antisense_strand
AACUUGAGCUGGACCAUGCUA
Mismatch tolerance
Both-strand mismatch tolerance: 2 Plus-strand mismatch tolerance: 2
ii. siRNA duplex

```
5'-GGUCCAGCUCAAGUUGUCACG      sense strand
   : : : : : : : : : : : : : : : : : : : : :
UACCAGGUCGAGUUCAACAGU-5'      antisense strand
```

Oligo sequences
>Sense_strand
GGUCCAGCUCAAGUUGUCACG
>Antisense_strand
UGACAACUUGAGCUGGACCAU
Mismatch tolerance
Both-strand mismatch tolerance: 2 Plus-strand mismatch tolerance: 2
5. siRNA Information for 74.6% conservation (Asia and Russia):
siRNA duplex

```
5'-UCCAAAUCAAAAGAUAAUAAC      sense strand
   : : : : : : : : : : : : : : : : : : : : :
UUAGGUUUAGUUUUCUAUUAU-5'      antisense strand
```

Oligo sequences
>Sense_strand
UCCAAAUCAAAAGAUAAUAAC
>Antisense_strand
UAUUAUCUUUUGAUUUGGAUU
Mismatch tolerance
Both-strand mismatch tolerance: 2 Plus-strand mismatch tolerance: 2

Influenza A infection and resistance develops rapidly (Wright *et al.*, 2002).

Inactivated vaccines are 60-80% effective against matched influenza strains (Hannon, 2002). Vaccination coverage is a problem world wide. Moreover, this strategy provides no protection against unexpected strains, outbreaks such as H5 and H7 avian influenza outbreaks in Honk Kong in 1997 and the Netherlands and Southeast Asia in 2003—2004, over pandemics.

RNA interference is an effective means or suppressing virus replication. Small interfering RNAs (siRNAs), mediators of RNAi, are short (21-26nt), double-stranded RNA duplexes that inhibit gene expression by inducing sequence-specific degradation of homologous mRNA (Gitlin *et al.,* 2003). A number of studies demonstrated inhibition of replication of RNA viruses by RNAi (Silva *et al.,* 2002) (jacque *et al.,* 2002), including HIV (Lee *et al,.* 2003), Polio virus (kapadia *et al,.* 2003), Hepatitus C virus (Randal *et al,.* 2003), West Nile virus, and influenza virus (Mcmanus *et al,.* 2003).

siRNA analysis through siVirus software has been done before for HIV in 495 near full-length HIV-1

sequences were analyzed. The analysis of those siRNA candidates revealed that highly conserved siRNAs constituted only 0.3% of the possible siRNAs for >90% conservation. The fraction was still as small as 0.8% even the threshold of the conservation was relaxed to 80%. On the other hand, siRNAs predicted to be functional by one or more guidelines constituted 35.5% 0f the 4417157 siRNAs. Taken together, siRNAs that were >80% conserved, and satisfied at least one guideline constituted only 0.2% of the siRNAs. In our results 0% highly conserved siRNAs were obtained while using >90% conservation parameter. However on decreasing the percentage of conservation close to 80% than siRNAs constituted 100% conservation in case of neuraminidase. We reported a screen of 90522 siRNA candidates for targeting 1006 sequences of the HA and 13576 siRNA candidates for targeting 1307 sequences of the NA segment of influenza genome.

Our results showing typical output from siVirus for designing anti-Influenza Virus A siRNAs. The sequences from Influenza Virus A subtypes A and B, which are the most prevalent Influenza Virus A genotypes circulating in Asia, were selected. The results were sorted by the degree of conservation, and filtered to display siRNAs that satisfy at least one efficacy guideline. It is desirable to select an siRNA that has less off target hits as they predict the dissimilarity with the target sequence. With 0 or 1 off-target hits we got 17 siRNA in case of Hemagglutinin and 13 siRNAs in case of Neuraminidase.

To test the validity of siVirus, siRNAs satisfying the guideline by Ui-Tei *et al.* showed better results i.e. the maximum conserved regions of Influenza Virus A segments HA and NA using siVirus. Thus guidelines of Ui-tei *et al* predicted efficient siRNA against influenza A virus.

CONCLUSION

The efforts to develop antiviral agents directly from viral genome sequences has led to short interfering RNA (siRNA), the mediator of a sequence-selective inhibition known as RNA interference (RNAi). The current *in silico* study resulted in finding various siRNA agents which can initially be tested in animal model and the results may have implications in developing prophylactic as well as therapeutic treatment against influenza infection. The outcome of this study supports the growing expectation that siRNA can fulfill the need for moving rapidly from gene sequence to targeted therapeutic agents for many previously intractable disease targets. siRNA can be effective when given to animals before or in established influenza infections and this intervention would be useful during influenza outbreaks, Further development of this technology may provide an effective strategy for controlling influenza and other viral diseases.

REFERENCES

Alexander DJ (2006). “Report on avian influenza in the eastern hemisphere during 1997–2002.” *Avian Dis.* 47:792–797.

Amarzguoui M, Prydz H (2004). “An alorthm for selection of functional siRNA sequences.” *Biochem. Biophys. Res. Commun* 316: 1050-1058.

Carolyn B, Mathew J, Caroline B (2003). “Transmission of Influenza: Implictions for control in health care settings.” *Clinical. Infection. Dis.* 37: 1094-1101.

Cros J, Palese P (2003). "Trafficking of viral genomic RNA into and out of the nucleus: influenza, Thogoto and Borna dis. viruses.” *Virus Res.* 95 (1-2): 3-12.

David B, Muthiah M, Victor ., Dinah W (2006). “RNAi therapeutics: a potential new class of pharmaceutical drugs.” *Nature chemical. Biol.* 2: 711-719.

Gitlin L, Andino R (2003). *J. Virol.* 77: 7159-7165.

Ge Q, McManus MS, Nguyen T, Shen CH, Sharp PA, Eisen HN, Chen JZ (2003*).* *Proc Natl. Acad. Sci. USA* 100: 2718-2723.

George E, Mitchell W, Kevan L (2001). “Neutrophil survival is markedly reduced by incubation with influenza virus” *J. Leukocyte Biol.* 69: 50-56.

Gitlin L, Karelsky S, Andino R (2002). *Nature* 418: 430-434.

Hannon GJ (2002). *Nature* 418: 244-251.

Hay A, Gregory V, Douglas A, Lin Y (2001). "The evolution of human influenza viruses." *Philos. Trans R. Soc. Lond. B. Biol. Sci.* 356 (1416): 1861-70.

Hilleman M (2002). "Realities and enigmas of human viral influenza: pathogenesis, epidemiology and control.” *Vac.* 20 (25-26): 3068-87.

Jacque JM, Triques K, Stevenson M (2002). *Nature* 418 : 435-438.

Kumar S, Tamura K, Nei M (2004). “MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment.” *Brief. Bioinform.* 5:150–163.

Kash J, Goodman A, Korth M, Katze M (2006). "Hijacking of the host-cell response and translational control during influenza virus infection.” *Virus. Res.* 119 (1) :111-20.

Knobler S, Mack A, Mahmoud A, Lemon S (2005). “The story of influenza.” *The National Academies Press* pp 60–61.

Kapadia SB, Brideau-Andersoen A, Chisari FV (2003). *Proc Natl. Acad. Sci USA* 100: 2014-2018.

Lee MT, Coburn GA, McClure MO, Cullen BR (2003). *J. Virol* 77: 11964-11972.

Martin P, (2006). "2,500-year evolution of the term epidemic". *Emerg. Infect. Dis.* 12(6).

Micheal GI, Freerick GH (2001).”Therapeutic Options for the management of Influenza”. *Curent Opinion in pharmocol.* 1:482-490.

Monto A (2006). “The Fight Against Seasonal Influenza”.

Moscona A (2005). "Neuraminidase inhibitors for influenza". *N. Engl. J. Med.* 353 (13): 1363–73.

McCown M, Diamond M, Pekosz A (2003). *Virol.* 313: 514-524.

Nayak D, Hui E, Barman S (2004). "Assembly and budding of influenza virus.” *Virus Res* 106 (2): 147-65. Drake J (1993). "Rates of spontaneous mutation among RNA viruses". *Proc Natl Acad Sci U S A* 90 (9): 4171-5.

Nobusawa E, Sato K (2006). "Comparison of the mutation rates of human influenza A and B viruses.” *J. Virol.* 80 (7): 3675-8.

Potter CW, (2006). "A History of Influenza". *J. Appl. Microbial.* 91 (4): 572–579.

Qing G, Lily F, Ailin B, Tam N, Herman N, Jianzhu C, (2004). “Inhibition of Influenza virus production in virus-infected mice by RNA interference”’ *PNAS* 23: 8676-8681.

Reynolds A, Leake D, Boese Q, Scaringe S, Marshall WS, Kvorova A (2004) “Rational siRNA design for RNA interference.” *Nat. Biotechnol.* 22: 326-330.

Randall G, Grakoui A, Rice CM (2003). *Proc Natl. Acad. Sci USA* 100: 235-240.

Silva JM, Hammond SM, Hannon GJ (2002). *Trends Mol. Med.* 8: 505-508.

Sims LD, Domenech J, Benigno C, Kahn S, Kamata A, Lubroth **J**,

Martin V, Roeder P(2006). “Origin and evolution of highly pathogenic

H5N1 avian influenza in Asia." *Vet. Rec.* 157 : 159–164.
Suzuki Y (2005). "Sialobiology of influenza: molecular mechanism of host range variation of influenza viruses. " *Biol Pharm Bull* 28 (3): 399-408.
Taubenberger J (2006). "1918 Influenza: the mother of all pandemics." *Emerg. Infect. Dis.* 12 (1): 15–22.
Ui-Tei K, Naito Y, Takahashi F, Haraguchi T, Ohki-Hamazaki H, Juni A, Ueda R, Saigo K (2004). "Guidelines for the selection of highly effective siRNA sequences for mammalian and chick RNA interference." *Nucleic Acid Res.* 32: 936-948.
Wilson J, von Itzstein M (2003). "Recent strategies in the search for new anti-influenza therapies." *Curr. Drug Targets* 4(5): 389-408.
Wagner R, Matrosovich M, Klenk H (2002). "Functional balance between haemagglutinin and neuraminidase in influenza virus infections." *Rev. Med. Virol.* 12 (3): 159-66.
Wright PF, Webster RG, Knipe DM, Howley PM, Griffin DE, Martin MA, Lamb RA, Roizman B, Stratus SE (2002). *Virol.* pp 1533-1579.