

Short report

Open Access

## Optimal design and validation of antiviral siRNA for targeting HIV-1

Yuki Naito\*<sup>1</sup>, Kyoko Nohtomi<sup>2</sup>, Toshinari Onogi<sup>2</sup>, Rie Uenishi<sup>2</sup>, Kumiko Ui-Tei<sup>1</sup>, Kaoru Saigo<sup>1</sup> and Yutaka Takebe\*<sup>2</sup>

Address: <sup>1</sup>Department of Biophysics and Biochemistry, Graduate School of Science, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, 113-0033, Japan and <sup>2</sup>Laboratory of Molecular Virology and Epidemiology, AIDS Research Center, National Institute of Infectious Diseases, 1-23-1 Toyama, Shinjuku-ku, Tokyo, 162-8640, Japan

Email: Yuki Naito\* - [y-naito@RNAi.jp](mailto:y-naito@RNAi.jp); Kyoko Nohtomi - [notomi@nih.go.jp](mailto:notomi@nih.go.jp); Toshinari Onogi - [onogit@nih.go.jp](mailto:onogit@nih.go.jp); Rie Uenishi - [uenishir@nih.go.jp](mailto:uenishir@nih.go.jp); Kumiko Ui-Tei - [ktei@biochem.s.u-tokyo.ac.jp](mailto:ktei@biochem.s.u-tokyo.ac.jp); Kaoru Saigo - [saigo@biochem.s.u-tokyo.ac.jp](mailto:saigo@biochem.s.u-tokyo.ac.jp); Yutaka Takebe\* - [takebe@nih.go.jp](mailto:takebe@nih.go.jp)

\* Corresponding authors

Published: 8 November 2007

Received: 6 August 2007

*Retrovirology* 2007, **4**:80 doi:10.1186/1742-4690-4-80

Accepted: 8 November 2007

This article is available from: <http://www.retrovirology.com/content/4/1/80>

© 2007 Naito et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

### Abstract

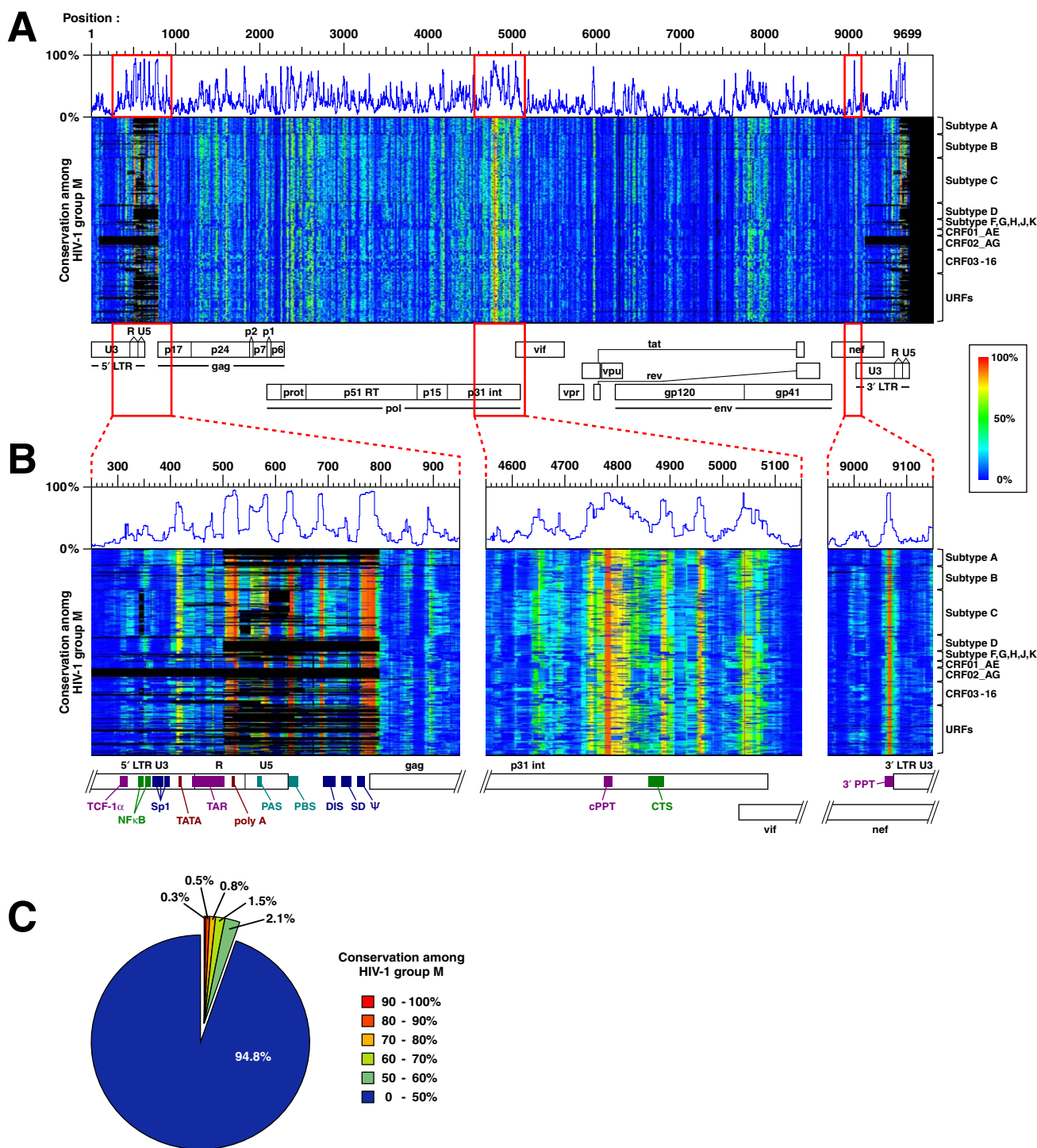
We propose rational designing of antiviral short-interfering RNA (siRNA) targeting highly divergent HIV-1. In this study, conserved regions within HIV-1 genomes were identified through an exhaustive computational analysis, and the functionality of siRNAs targeting the highest possible conserved regions was validated. We present several promising antiviral siRNA candidates that effectively inhibited multiple subtypes of HIV-1 by targeting the best conserved regions in pandemic HIV-1 group M strains.

### Findings

RNA interference (RNAi) 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 applications including antiviral treatments [1,2]. The replication of a wide range of viruses can be successfully inhibited using RNAi with both short interfering RNA (siRNA) and siRNA expression vectors [3,4]. However, for RNA viruses such as HIV-1, designing functional siRNAs that target viral sequences is problematic because of their extraordinarily high genetic diversity. We analyzed 495 entries of near full-length HIV-1 group M sequences available in the Los Alamos HIV Sequence Database, and selected the highest-possible conserved target sites for designing optimal antiviral siRNAs. It is known that RNAi-resistant viral mutants emerge rapidly when targeting viral sequences due to their high mutation rate [5-7]. Since highly conserved sequences are likely to contain structurally or functionally constrained

elements, our approach is anticipated to resist viral mutational escape.

First, we performed a detailed analysis on the HIV-1 genome to identify highly conserved targets by using 495 near full-length genome sequences of HIV-1 group M (listed in Additional file 1). Every possible 21-mer was generated from all of the HIV-1 group M sequences, and their conservations among the 495 HIV-1 sequences were exhaustively determined using siVirus engine [8]. We defined 'conservation' as the percentage of sequence entries out of the 495 HIV-1 sequences that showed perfect identity (*i.e.*, 21/21 matches) with the cognate 21-mer. Since many of the HIV-1 sequence entries lack 5' untranslated region (5' UTR), the 3' LTR sequence was used to compensate for the lack of 5' LTR sequences in order to avoid underestimating conservation in such regions. For the regions that cannot be compensated for in this way (depicted in Figure 1A and 1B left panel, colored



**Figure 1**  
**Conservations of siRNA target sequences among HIV-1 group M.** (A) A total of 4,417,157 siRNA targets were generated from the 495 HIV-1 sequences, and their conservations within the HIV-1 genomes are represented using a color density plot. The line plot above the color chart represents the highest value in each position. (B) A detailed view of the three conserved regions; 5' LTR, the cPPT/CTS in the integrase gene, and 3' PPT. 'Position' indicates the 5'-most position of each 21-mer. The landmarks of the HIV-1 genome are adjusted to align at the center of the siRNAs by shifting 10 bp to the left. (C) Pie chart indicating the percentage of the 4,417,157 siRNA target sites at each conservation level.

black), conservation was calculated by considering only the HIV-1 sequences that contain the corresponding regions. The result revealed that HIV-1 genomes are not conserved for consecutive 21 bp for the most part, resulting in the poor conservation of many of the 21-mers over the HIV-1 sequences (Figure 1A, colored blue). As shown in Figure 1C, only 5.2% of the possible 21-mers are >50% conserved. Furthermore, highly (>70%) conserved 21-mers constitute only 1.6% of all 21-mers. It is of note that many of the published anti-HIV-1 siRNA sequences do not fall into this 'highly conserved' category (Additional file 2 and [9]). From these results, we anticipate that most of the possible siRNAs are not suitable for the efficient targeting of HIV-1.

However, our analysis has identified several distinct regions that are highly conserved in the HIV-1 genome (Figure 1B). Such regions include the regulatory domains responsible for the viral gene expression, such as the TATA sequence and polyadenylation signal (AAUAAA). In addition, several regions essential for the regulation of viral replication were also highly conserved, including the primer activation signal (PAS)[10], primer binding site (PBS), packaging signal ( $\Psi$ ), central polypurine tract (cPPT), central termination sequence (CTS), and 3' polypurine tract (3' PPT). All of these highly conserved sequences are constrained at the nucleotide sequence level or by their RNA secondary structure in order to execute their functions. In contrast, regions constrained by amino acid sequences were not necessarily conserved at the nucleotide sequence level due to the wobbling of the third base in the codon (data not shown). siRNAs targeting the highly conserved regions are expected to overwhelm the high level of sequence diversity of the HIV-1 genome, and also to reduce the chances of viral mutational escapes.

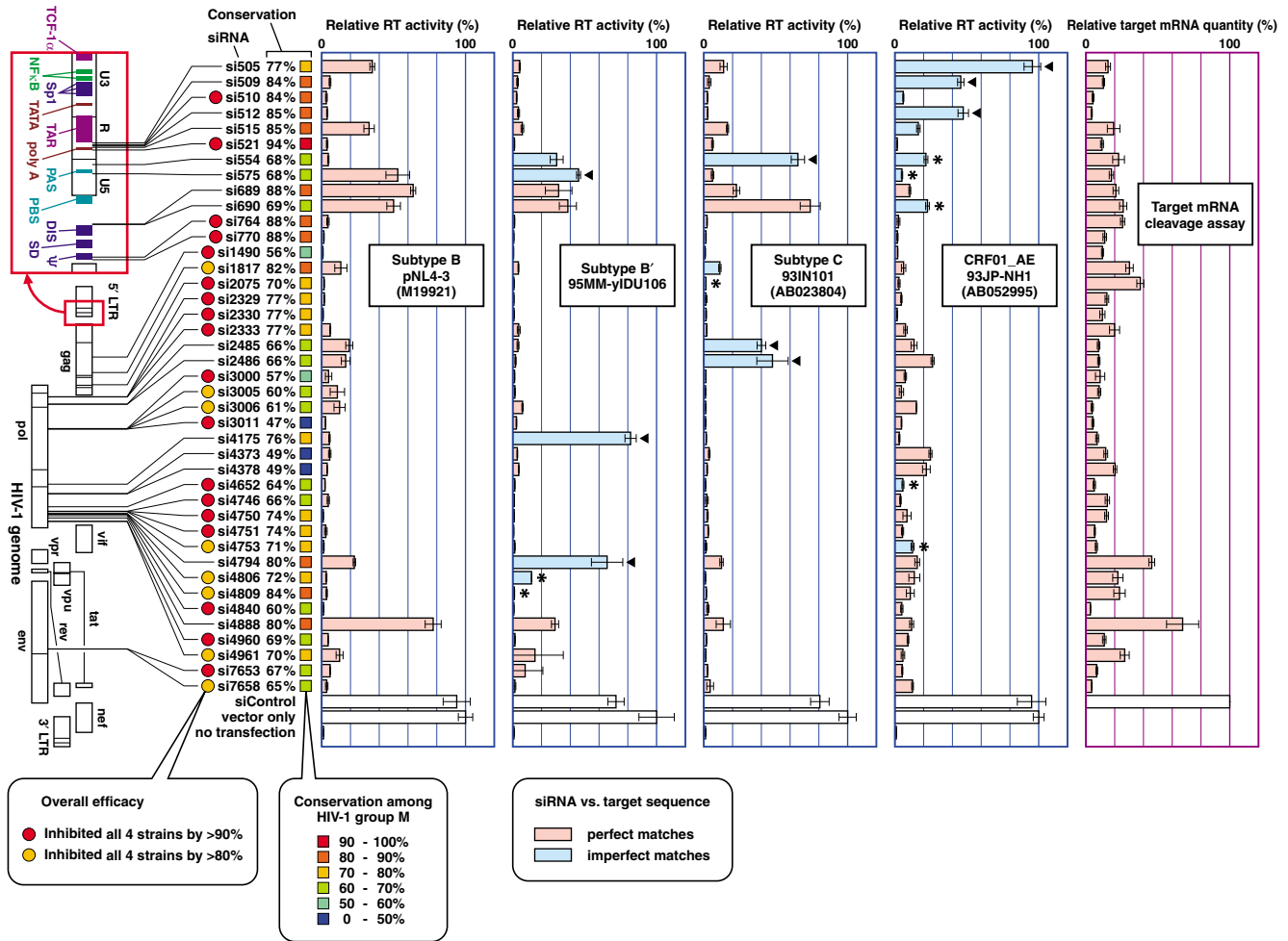
Total of 216 highly conserved (>70%) siRNA targets identified in this study are listed in Additional file 3. In mammalian RNAi, the efficacy of each siRNA varies markedly depending on its sequence. According to our guidelines for the selection of effective siRNAs [11,12], 31 out of 216 siRNAs were predicted to be functional. Similarly, 30 and 44 siRNAs are functional according to the algorithms reported by Reynolds *et al.* [13], and Amarzguioui *et al.* [14], respectively (Additional file 3). This suggests that only a limited fraction of 21-mers is best suited for use as functional antiviral siRNAs.

For the functional validation, 23 siRNAs from Additional file 3, and 18 additional siRNAs targeting moderately-conserved regions were selected based on the following criteria: (I) predicted to be functional by the algorithm of Ui-Tei *et al.* [11,12], and (II) the sequence has perfect identity with pNL4-3 (GenBank [M19921](#)). The 41 siRNA sequences selected and their target sites are detailed in

Additional file 4. We first tested the efficacy of each siRNA using target mRNA cleavage assay (Additional file 5 and [15]). Briefly, a vector expressing reporter mRNA that contains the siRNA target site was cotransfected into HeLa cells with the corresponding siRNA, and the mRNA cleavage activity of the siRNA was evaluated by measuring the quantity of surviving mRNA using real-time RT-PCR. This assay allows us to directly monitor the sequence-dependent potency of siRNA itself, without being affected by the differences in target gene expression level or target secondary structures. The result showed that 39 out of the 41 siRNAs gave >60% silencing at 5 nM (Figure 2, rightmost panel). si4794 and si4888 were not functional, probably due to the long consecutive Gs in si4794 and internal palindromes (AAAAUUUU) in si4888 [11,13].

Next, siRNAs were evaluated for their antiviral efficacy against three evolutionary-distant groups of HIV-1: subtypes B and B' (Thailand variant of subtype B [16]); subtype C; and CRF01\_AE. Each siRNA was cotransfected into HeLa cells at 5 nM with one of the four infectious molecular clones: pNL4-3 (subtype B); 95MM-yIDU106 (subtype B'); 93IN101 (subtype C); or 93JP-NH1 (CRF01\_AE). Culture supernatants were collected 48 h after transfection and the viral reverse transcriptase activity was measured (Additional file 5 and [17]). The results show that 26 of the 41 siRNAs effectively inhibited viral replication of all four strains by >80% (Figure 2, marked with red or orange circles). Of the remaining 15 siRNAs, 13 of them (except si4794/4888) were shown to be functional in the target mRNA cleavage assay, and 12 of them (except si690/4794/4888) inhibited the replication of at least one viral strain by >80%, indicating that the designed siRNAs have the potential to induce RNAi. In several viral strains, nucleotide substitutions in their target sites essentially abolished the inhibition of viral replication (Figure 2, blue bars with arrowheads). However, mismatches near the ends of the target sites (see Additional file 6) did not necessarily abolish the siRNA efficacy (Figure 2, blue bars with asterisks). si689 and si690 did not inhibit viral replication even though these siRNAs perfectly matched to their target sites (confirmed by DNA sequencing of the infectious molecular clones). This is probably due to the stable secondary structure at the si689-690 target sites in both BMH (branched multiple hairpin) conformation and LDI (long distance interaction) conformation of the HIV-1 leader RNA [18] (see Additional file 4). It should be noted that the efficacy of si575 differed when targeting pNL4-3 and 93IN101. One possible explanation for this is the secondary structure differences among HIV-1 subtypes, which may alter the accessibility of the si575 target site.

The approach described here enabled us to select highly effective siRNAs against divergent HIV-1 strains at a high



**Figure 2**  
**Validation of 41 siRNAs.** The antiviral efficacy of each siRNA was tested against four HIV-1 infectious molecular clones: pNL4-2 (subtype B); 95MM-yIDU106 (subtype B'); 93IN101 (subtype C); or 93JP-NH1 (CRF01\_AE). The potency of each siRNA was tested using the target mRNA cleavage assay (rightmost panel). The ability of each siRNA to cleave its target was evaluated by the target mRNA cleavage assay.

rate. The highly effective siRNAs (>90% inhibition) with maximal conservation (>70%) identified in our study include si521 (poly A site; 94% conservation), si764/770 ( $\Psi$ ; 88%), si510 (TAR/poly A; 84%), si2075 (ribosomal slip site; 70%), si2329/2330/2333 (protease region; 77%), and si4750/4751/4753 (integrase region; 71–74%). These sites are found mostly in the 5' LTR, protease, and integrase regions (Figure 2). However, the extraordinarily high genetic diversity of HIV-1 obviously prevents us from designing a single siRNA that can nullify all HIV-1 strains currently circulating worldwide (Additional file 7). One possible approach is to combine multiple siRNAs targeting different conserved regions [19,20]. The siRNAs selected and validated in this study have the potential to target >99% of HIV-1 strains by combining only two siRNAs (Additional file 7), and also considered

to resist viral mutational escape. Our approach is expected to be highly applicable to therapeutic intervention for other pathogens of public health importance, including HCV, influenza virus, and SARS coronavirus, that are known to show high genetic diversity.

**Competing interests**

The author(s) declare that they have no competing interests.

**Authors' contributions**

YN performed the computational analyses and the target mRNA cleavage assays, participated in the design of the study, and drafted the manuscript. KN and TO performed the viral replication assays. RU analyzed the data. KU-T participated in the target mRNA cleavage assays, and was

involved in critically revising the manuscript. KS and YT supervised the entire study and wrote the manuscript.

## Additional material

### Additional file 1

The list of 495 near full-length genome sequences of HIV-1 group M. Click here for file  
[http://www.biomedcentral.com/content/supplementary/1742-4690-4-80-S1.pdf]

### Additional file 2

The list of published siRNA/shRNAs targeting HIV-1. Click here for file  
[http://www.biomedcentral.com/content/supplementary/1742-4690-4-80-S2.pdf]

### Additional file 3

The list of highly conserved siRNA targets identified in this study. Click here for file  
[http://www.biomedcentral.com/content/supplementary/1742-4690-4-80-S3.pdf]

### Additional file 4

The siRNA sequences and their target sites. The sequences of 41 siRNAs and their target sites are shown. The siRNA numbers indicate the nucleotide position in HXB2 (GenBank [K03455](#)). The conservation level of each siRNA in HIV-1 group M sequence is depicted in color chart at the rightmost column. BMH (branched multiple hairpin) and LDI (long distance interaction) conformations of the HIV-1 leader RNA and siRNAs targeting them are shown. Click here for file  
[http://www.biomedcentral.com/content/supplementary/1742-4690-4-80-S4.pdf]

### Additional file 5

Supplementary materials and methods. Click here for file  
[http://www.biomedcentral.com/content/supplementary/1742-4690-4-80-S5.pdf]

### Additional file 6

Target sites of the 41 siRNAs used in this study. Sequence alignment of the target site from the four HIV-1 infectious molecular clones: pNL4-2 (subtype B); 95MM-γIDU106 (subtype B'); 93IN101 (subtype C); or 93JP-NH1 (CRF01\_AE). Click here for file  
[http://www.biomedcentral.com/content/supplementary/1742-4690-4-80-S6.pdf]

### Additional file 7

Coverage of HIV-1 group M by single siRNA or two siRNAs. (A) Coverage of HIV-1 group M by 41 siRNAs used in this study. (B) Coverage of HIV-1 group M by combining two siRNAs from above. Coverage was calculated by considering only the HIV-1 sequences which contain the corresponding regions. Click here for file  
[http://www.biomedcentral.com/content/supplementary/1742-4690-4-80-S7.pdf]

## Acknowledgements

This study was supported in part by grants from the Ministry of Education, Culture, Sports, Science and Technology of Japan (to YN, KU-T, KS, and YT), the Ministry of Health, Labour and Welfare of Japan (to YT), and the Japan Health Sciences Foundation (to YT).

## References

1. Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC: **Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans***. *Nature* 1998, **391**:806-811.
2. Hannon GJ, Rossi JJ: **Unlocking the potential of the human genome with RNA interference**. *Nature* 2004, **431**:371-378.
3. Nielsen MH, Pedersen FS, Kjems J: **Molecular strategies to inhibit HIV-1 replication**. *Retrovirology* 2005, **2**:10.
4. Leonard JN, Schaffer DV: **Antiviral RNAi therapy: emerging approaches for hitting a moving target**. *Gene Ther* 2006, **13**:532-540.
5. Boden D, Pusch O, Lee F, Tucker L, Ramratnam B: **Human immunodeficiency virus type I escape from RNA interference**. *J Virol* 2003, **77**:11531-11535.
6. Das AT, Brummelkamp TR, Westerhout EM, Vink M, Madiredjo M, Bernards R, Berkhout B: **Human immunodeficiency virus type I escapes from RNA interference-mediated inhibition**. *J Virol* 2004, **78**:2601-2605.
7. Westerhout EM, Ooms M, Vink M, Das AT, Berkhout B: **HIV-1 can escape from RNA interference by evolving an alternative structure in its RNA genome**. *Nucleic Acids Res* 2005, **33**:796-804.
8. Naito Y, Ui-Tei K, Nishikawa T, Takebe Y, Saigo K: **siVirus: web-based antiviral siRNA design software for highly divergent viral sequences**. *Nucleic Acids Res* 2006, **34**:W448-W450.
9. ter Brake O, Berkhout B: **A novel approach for inhibition of HIV-1 by RNA interference: counteracting viral escape with a second generation of siRNAs**. *J RNAi Gene Silencing* 2005, **1**:56-65.
10. Beerens N, Groot F, Berkhout B: **Initiation of HIV-1 reverse transcription is regulated by a primer activation signal**. *J Biol Chem* 2001, **276**:31247-31256.
11. Ui-Tei K, Naito Y, Takahashi F, Haraguchi T, Ohki-Hamazaki H, Juni A, Ueda R, Saigo K: **Guidelines for the selection of highly effective siRNA sequences for mammalian and chick RNA interference**. *Nucleic Acids Res* 2004, **32**:936-948.
12. Naito Y, Yamada T, Ui-Tei K, Morishita S, Saigo K: **siDirect: highly effective, target-specific siRNA design software for mammalian RNA interference**. *Nucleic Acids Res* 2004, **32**:W124-W129.
13. Reynolds A, Leake D, Boese Q, Scaringe S, Marshall WS, Khvorov A: **Rational siRNA design for RNA interference**. *Nat Biotechnol* 2004, **22**:326-330.
14. Amarzguioui M, Prydz H: **An algorithm for selection of functional siRNA sequences**. *Biochem Biophys Res Commun* 2004, **316**:1050-1058.
15. Ui-Tei K, Naito Y, Saigo K: **Guidelines for the selection of effective short-interfering RNA sequences for functional genomics**. *Methods Mol Biol* 2007, **361**:201-216.
16. Kalish ML, Baldwin A, Raktham S, Wasi C, Luo CC, Schochetman G, Mastro TD, Young N, Vanichseni S, Rübsamen-Waigmann H, von Briesen H, Mullins JI, Delwart E, Herring B, Esparza J, Heyward WL, Osmanov S: **The evolving molecular epidemiology of HIV-1 envelope subtypes in injecting drug users in Bangkok, Thailand: implications for HIV vaccine trials**. *AIDS* 1995, **9**:851-857.
17. Willey RL, Smith DH, Lasky LA, Theodore TS, Earl PL, Moss B, Capon DJ, Martin MA: **In vitro mutagenesis identifies a region within the envelope gene of the human immunodeficiency virus that is critical for infectivity**. *J Virol* 1988, **62**:139-147.
18. Huthoff H, Berkhout B: **Two alternating structures of the HIV-1 leader RNA**. *RNA* 2001, **7**:143-157.
19. Nishitsuji H, Kohara M, Kannagi M, Masuda T: **Effective suppression of human immunodeficiency virus type I through a combination of short- or long-hairpin RNAs targeting essential sequences for retroviral integration**. *J Virol* 2006, **80**:7658-7666.
20. ter Brake O, Konstantinova P, Ceylan M, Berkhout B: **Silencing of HIV-1 with RNA interference: a multiple shRNA approach**. *Mol Ther* 2006, **14**:883-892.