Virologica Sinica (2020) 35:359–361
https://doi.org/10.1007/s12250-020-00221-6

www.virosin.org
www.springer.com/12250

LETTER

Computational Identification of Small Interfering RNA Targets in SARS-CoV-2

Wei Chen[1,2] · Pengmian Feng[1] · Kewei Liu[2] · Meng Wu[2] · Hao Lin[3]

Received: 6 March 2020 / Accepted: 3 April 2020 / Published online: 15 April 2020

Dear Editor,

At the end of 2019, a new virus, called Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) was reported (Benvenuto *et al.* 2020; Zhu *et al.* 2020). The sequences of SARS-CoV-2 reported by different research groups demonstrated that it is a positive strand RNA virus. The sequence of SARS-CoV-2 is approximately 30 kb long, and could encodes spike, envelope, membrane, nucleocapsid proteins, etc. (Phan 2020). These proteins are responsible for replicating the viral genome as well as generating nested transcripts that are used in the synthesis of the viral proteins.

As of April 3 2020, there was more than 1 million cases of SARS-COV-2 reported to World Health Organization with 50,000 deaths globally. However, there have been no effective measures to prevent or treat the severe complications caused by SARS-COV-2.

RNA interference (RNAi) is a native and specific post-transcriptional gene silencing mechanism (Bobbin and Rossi 2016). The progress initiated by double-stranded RNA (dsRNA) to manipulate gene expression (RNAi) has been proved highly effective, at least 10 times more effective than either using sense or antisense RNAs alone (Chalk and Sonnhammer 2002). The RNAi triggered by dsRNA is a phenomenon of homology-dependent gene silencing and may play certain roles in affecting the process of virus expression and proliferation. Recently, several reports have demonstrated the use of RNAi in blocking virus infection and replication in animal cells (Ge *et al.* 2003), suggesting that the small interfering RNA (siRNA, 21–25 nt long) plays an important role in RNAi-related gene silencing pathways (Elbashir *et al.* 2001). Progress has been made in anti-HIV and anti-HCV drug design by applying the method of RNA interference (Wilson *et al.* 2003). The effectiveness of siRNA for inhibiting SARS coronavirus genes expression was also demonstrated by Shi *et al.* (2005). Besides silencing the targeted genes, the siRNAs can also inhibit the replication of the virus. For example, it has been demonstrated that, by targeting the Leader sequence of SARS-CoV, the siRNA demonstrate a strong inhibitory effect on SARS-CoV replication (Li *et al.* 2005). More recently, a CRISPR/Cas13d system was proposed for the treatment of SARS-COV-2 (Nguyen *et al.* 2020). These results indicate that both RNAi and CRISPR/Cas technology might become potential therapeutic approaches for treating viral diseases.

Accordingly, as complementary to the CRISPR/Cas13d system, we proposed an RNAi based strategy that might interfere the gene expression and block the replication of SARS-COV-2. The main idea of this strategy is to search for siRNA targets in the virus genome, which will be recognized and cleaved by the RNA-induced silencing complex (RISC).

In this work, we performed theoretical predictions of the potential siRNA targets in the virus genome. We firstly collected the representative SARS-COV-2 genome (MN908947, https://www.ncbi.nlm.nih.gov/nuccore/MN908947) and the mutation information of the SARS-COV-2 genomes from the 2019nCoVR database (Zhao *et al.* 2020), which is available at https://bigd.big.ac.cn/ncov/. The 2019nCoVR database not only integrates genomic and proteomic sequences of SARS-COV-2 from different resources, but also provides a series of scientific

✉ Wei Chen
greatchen@ncst.edu.cn

✉ Hao Lin
hlin@uestc.edu.cn

[1] Innovative Institute of Chinese Medicine and Pharmacy, Chengdu University of Traditional Chinese Medicine, Chengdu 611137, China

[2] Center for Genomics and Computational Biology, School of Life Sciences, North China University of Science and Technology, Tangshan 063210, China

[3] Key Laboratory for Neuro-Information of Ministry of Education, School of Life Science and Technology, Center for Informational Biology, University of Electronic Science and Technology of China, Chengdu 610054, China

Table 1 siRNA target sequence in plus strand of coronavirus (MN908947).

Target 5′–3′	Position	Region	Length	AR (RAR)	Number of mutation strain
AAUAGUUUAAAAAUUACAG**A**AGA	6509–6531	Orf1ab	23	20 (20)	1
UCCUUCUUUAGAAACUAUACA	7168–7188	Orf1ab	21	18 (12.6)	0
UGGUUUCACUACUUUCUGUUU	11,997–12,017	Orf1ab	21	15 (10.5)	0
UUCACUACUUUCUGUUUUGCU	12,001–12,021	Orf1ab	21	15 (10.5)	0
AUGUCAUCCCUACUAUAACUCAAA	15,041–15,064	Orf1ab	24	18 (18)	0
UUAAAAUAUAAUGAAAAUGGA	22,391–22,411	S	21	18 (12.6)	0
CUUGAAGCCCCUUUUCUCUAUCUUU	25,693–25,717	Orf3a	25	18 (12.6)	0
CAACUAUAAAUUAAACACA**G**A	27,128–27,148	M	21	19 (19)	2
UUGAAUACACCAAAAGAUCACAUU	28,688–28,711	N	24	18 (18)	0

The bold and underlined characters indicate the SNP found in different strains.

services, such as variation visualization, variation annotations, AI diagnosis, etc.

Next, we folded the SARS-COV-2 genome (MN908947) in a window of 3000 nucleotides with the step of 1500 nucleotides by using RNAstructure (version 4.5) program (Bellaousov *et al.* 2013). Only those 21–25 nt long non-base-paired regions can be served as the potential targets of siRNA (Huang *et al.* 2008), which is called free segments. The long non-base-paired region containing one or several short stems (total length of stems 1–3 base pairs), called quasi-free segments (Ji and Luo 2004), was also considered in the present work.

A given RNA sequence segment may have different configurations of secondary structure with lower free energy. The total frequency of a segment occurring in non-base-paired region of different folds (20 folds are selected for each segment) is called appearance rate (AR). If each quasi-free case is multiplied by a reduced factor in numeration, namely, by 0.9 for 1 base pair, 0.8 for 2 base pair, and 0.7 for 3 base pairs (base pairs may be continuous in structure or disconnected) then the total number of folds is called reduced appearance rate (RAR) (Ji and Luo 2004).

To guarantee the safety of the designed drug, we further performed alignment of the free and quasi-free segments with human genome (hg 38) by using BLAST and deleted the matching ones in siRNA target candidates.

Finally, we obtained nine potential siRNA targets in the SARS-COV-2 genome (MN908947). The information about their position and region in the virus genome, length, AR and RAR was provided in Table 1.

In addition, we also analyzed the mutations of the target sequences by comparing all the 143 high quality strains in the 2019nCoVR database (as of March 15, 2020). SNP were found in two of the nine target sequences (indicated by bold character in Table 1). For the potential target 'AAUAGUUUAAAAAUUACAGAAGA', only one SNP was found in the strain BetaCoV/Wuhan/HBCDC-HB-05/2020, which is a coding_sequence_variant that changes the coding sequence. For 'CAACUAUAAAUUAAACACAGA', the SNP was found in the strain BetaCoV/Singapore/6/2020 and BetaCoV/Singapore/2/2020, respectively, which is a missense_variant that changes G to A resulting in a different amino acid sequence. These results indicate that the selected targets are conserved among the existing SARS-COV-2 genomes.

Although there are still some challenges that needed to be overcome for the clinic applications of siRNA, progresses have been made to solve the fundamental problems, such as off-target effects and effective delivery. For example, the position-specific chemical modification of siRNAs could can significantly reduce off targeting; safe and effective *in vivo* delivery systems have also been developed, such as nanoparticles, cationic lipids, antibodies, cholesterol, aptamers delivery strategies. Therefore, we hope that the above results would be useful in drug design and treatments against SARS-COV-2.

Acknowledgements The authors would like to express gratitude to the Editor and anonymous reviewers for their constructive comments. This work was supported by the National Nature Scientific Foundation of China (31771471, 61772119) and the Natural Science Foundation for Distinguished Young Scholar of Hebei Province (No. C2017209244).

Compliance with Ethical Standards

Conflict of interest The authors declare that they have no conflict of interest.

Animal and Human Rights Statement This article does not contain any studies with human or animal subjects performed by any of the authors.

References

Bellaousov S, Reuter JS, Seetin MG, Mathews DH (2013) Rnastructure: web servers for RNA secondary structure prediction and analysis. Nucleic Acids Res 41:W471–W474

Benvenuto D, Giovanetti M, Ciccozzi A, Spoto S, Angeletti S, Ciccozzi M (2020) The 2019-new coronavirus epidemic: evidence for virus evolution. J Med Virol 92:455–459

Bobbin ML, Rossi JJ (2016) Rna interference (rnai)-based therapeutics: delivering on the promise? Annu Rev Pharmacol Toxicol 56:103–122

Chalk AM, Sonnhammer EL (2002) Computational antisense oligo prediction with a neural network model. Bioinformatics 18:1567–1575

Elbashir SM, Harborth J, Lendeckel W, Yalcin A, Weber K, Tuschl T (2001) Duplexes of 21-nucleotide rnas mediate rna interference in cultured mammalian cells. Nature 411:494–498

Ge Q, McManus MT, Nguyen T, Shen CH, Sharp PA, Eisen HN, Chen J (2003) Rna interference of influenza virus production by directly targeting mrna for degradation and indirectly inhibiting all viral RNA transcription. Proc Natl Acad Sci USA 100:2718–2723

Huang C, Li M, Chen C, Yao Q (2008) Small interfering rna therapy in cancer: mechanism, potential targets, and clinical applications. Expert Opin Ther Targets 12:637–645

Ji FM, Luo LF (2004) Prediction for target sites of small interfering RNA duplexes in sars coronavirus. Genome Biol 5:P6

Li T, Zhang Y, Fu L, Yu C, Li X, Li Y, Zhang X, Rong Z, Wang Y, Ning H, Liang R, Chen W, Babiuk LA, Chang Z (2005) Sirna targeting the leader sequence of sars-cov inhibits virus replication. Gene Ther 12:751–761

Nguyen TM, Zhang Y, Pandolfi PP (2020) Virus against virus: a potential treatment for 2019-ncov (sars-cov-2) and other RNA viruses. Cell Res 30:189–190

Phan T (2020) Novel coronavirus: from discovery to clinical diagnostics. Infect Genet Evol 79:104211

Shi Y, Yang DH, Xiong J, Jia J, Huang B, Jin YX (2005) Inhibition of genes expression of sars coronavirus by synthetic small interfering rnas. Cell Res 15:193–200

Wilson JA, Jayasena S, Khvorova A, Sabatinos S, Rodrigue-Gervais IG, Arya S, Sarangi F, Harris-Brandts M, Beaulieu S, Richardson CD (2003) Rna interference blocks gene expression and RNA synthesis from hepatitis c replicons propagated in human liver cells. Proc Natl Acad Sci USA 100:2783–2788

Zhao WM, Song SH, Chen ML, Zou D, Ma LN, Ma YK, Li RJ, Hao LL, Li CP, Tian DM, Tang BX, Wang YQ, Zhu JW, Chen HX, Zhang Z, Xue YB, Bao YM (2020) The 2019 novel coronavirus resource. Yi Chuan 42:212–221

Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, Zhao X, Huang B, Shi W, Lu R, Niu P, Zhan F, Ma X, Wang D, Xu W, Wu G, Gao GF, Tan W, China Novel Coronavirus I, Research T (2020) A novel coronavirus from patients with pneumonia in china, 2019. N Engl J Med 382:727–733