Available online at www.sciencedirect.com
ScienceDirect
Biochemical and Biophysical Research Communications 358 (2007) 272–276
BBRC
www.elsevier.com/locate/ybbrc

A randomized lentivirus shRNA library construction

Qiang Guo, Yirong Kong, Ling Fu, Ting Yu, Junjie Xu *, Wei Chen *

State Key Laboratory of Pathogen and Biosecurity, Beijing Institute of Microbiology and Epidemiology, 20 Dongdajie, Fengtai, Beijing 100071, China

Received 12 April 2007
Available online 27 April 2007

Abstract

Vector based shRNA (short hairpin RNA) expression library has been widely used to screen functional genes. For two main methods that have been used to generate short hairpin RNA libraries, chemical synthesis is too expensive to be widely used and the low efficiency of enzymatic approach makes it difficult to construct. We have developed a protocol to construct a new kind of shRNA library called randomized shRNA library. Within three steps chemically synthesized randomized 19-mers DNA were efficiently converted to double-stranded DNA fragments containing shRNA templates. This kind of shRNA library permits simple and economic construction, providing another choice for whole-genome phenotypic screening of genes.

Keywords: shRNA; RNA interference; shRNA library; Lentivirus

RNAi (RNA interference), a gene suppression phenomenon triggered by double stranded RNA [1], is a powerful tool in determining gene-function relationships. Large-scale phenotypic or pathway-driven screens of shRNA libraries has been used to disclose genes related to a certain function and identify novel genes that may be targets for therapy in cancer and other diseases.

There have been two main methods to generate genome-wide vector-based shRNA libraries. The first is to chemically synthesize oligonucleotides [2,3]. Typically, shRNA sequences are chosen by a certain design algorithms, and then the coding sequences are synthesized in the form of double-stranded DNA molecules containing shRNA templates and are cloned into a Pol III-driven expression vector. Libraries constructed with this method targeting more than 10,000 different human genes have been successfully used for screening [4,5]. But this method is expensive and time-consuming because it requires chemically synthesizing thousands of oligonucleotides, followed by cloning and sequence validating. The second method is to generate shRNA libraries from pool of double-stranded cDNAs [6–9]. After several steps of cut, ligation and amplification, cDNAs are converted to double-stranded DNA molecules containing shRNA templates, and then cloned into a Pol III-driven expression vector. Construction shRNA library from cDNA is economic; however, the multiple-step process makes it a technical difficulty because of the low overall efficiency.

Herein, we report the development of a method to construct a new type shRNA library that allows the simple and economic construction and genomewide screens. Within three steps chemically synthesized randomized 19-mers DNA were converted to double-stranded DNA fragments containing shRNA templates. After restriction endonuclease cutting, the fragments coding for shRNA were cloned into a lentivirus shRNA expression vector with human U6 promoter. This kind of shRNA library permits simple and economic construction, providing another choice for whole-genome phenotypic screening of genes.

Materials and methods

Oligonucleotides

Pre-hairpin: 5′-nnnnnnnnnnnnnnnnnnnttggatccttcaaagagaggatccaa-3′.
Y-linker: LY5: 5′-gccgaggaccgcggtccgactcgagtttatatatcttgtggaaaggatagtccacgt-3′; LY3: 5′-cgtggactctcctttccacaagattcggacctagaggctaaggcctcccatcgc-3′.

* Corresponding authors. Fax: +86 10 63815273.
E-mail addresses: xujj02@yahoo.com.cn (J. Xu), chenwei0226@yahoo.com.cn (W. Chen).

0006-291X/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.bbrc.2007.04.123

PCR primers: Forward: 5′-gccgaggaccgcggtccgactcgagtttatat-3′; Reverse: 5′-gcgatgggaggccttagcctctaggtccga-3′.

The procedure of randomized shRNA Library construction

Pre-hairpin fill-in. The chemically synthesized pre-hairpin oligonucleotide was annealed after phosphorylation with T4 polynucleotide kinase (New England Biolabs) according to the manufacturer's instructions. The annealing reaction was done by incubating dissolved oligonucleotide at 90 °C for 4 min, and then at 70 °C for 10 min, and slowly cooling the annealed oligos to 10 °C within 30 min. The annealed oligonucleotide was treated with Ex Taq DNA polymerase (TaKaRa) at 60 °C for 2 h.

Y-Linker and hairpin ligation. The product of fill-in was extracted with phenol/chloroform (SonGon ShangHai), precipitated with ethanol, resuspended in 10 μl of water. Six microliters of Y-Linker (3 μg) and 1.5 μl of T4 ligase (450 U) were added, and the mixture incubated at 16 °C for 2 h. We separated the reaction mixture on a 10% Tris–borate–EDTA PAGE gel and gel-purified the desired band with Qiaex II beads (Qiagen).

PCR amplification. First, 100 μl reaction mix was prepared (2.5 U polymerase La Taq DNA polymerase (TaKaRa), 2 μl ligation product, 10 μl 10× buffer, 8 μl dNTP (2.5 mM), 0.4 μM primers), then divided into ten tubes, each containing 10 μl reaction mix. PCR was carried out with the following cycling parameters: 94 °C for 5 min, 94 °C 30 s, 75 °C 1 min, for 20, 22, 24, 26, 28, 30, 32, 34, 36, or 38 cycles. The PCR products were compared by 1.5% agarose gel electrophoresis. The cycle number yielding the greatest quantity of the 170 bp product was chosen for amplifying in large quantity. After PCR, an additional PCR cycle was performed with the addition of fresh PCR reaction mix (buffer, dNTPs, primers and polymerase).

Cloning. The 170 bp PCR product was digested with XhoI and MlyI at 37 °C for 5 h. The 120 bp insert was separated on a 1% agarose gel, purified, and cloned into the expression vector pLENTI-hU6m (see below). We ligated 100 ng XhoI–MlyI fragment to 1 μg of pLENTI-hU6m and transformed it into Sthl4 (Invitrogen) bacterial competent cells. We scraped the resulting bacterial colonies and isolated the shRNA constructs.

Removing excess loop sequence: We digested the plasmids with BamHI and self-ligated them to produce the final shRNA constructs.

Cell culture and cotransfection

CHO cells were seeded 24 h before transfection in F12 plus 10% FCS in 96-well plates at 3×10^4/cell. Transfection was conducted with Lipofectamine 2000 (Invitrogen) according to the manufacturer's instructions. The amounts of plasmid used in cotransfection were as follows: 50 ng pEGFP-N1 plasmid, 150 ng shRNA expression plasmid. Samples were analyzed at 24 h after transfection.

Construction of shRNA expression vectors

The viral backbone of our lentivirial expression constructs is pLENTIV5/Dest (Invitrogen). It was modified by inserting a human U6 promoter driving a shRNA between ClaI and MluI. hU6-EGFP has an hU6 promoter driving expression of a shRNA targeting EGFP (5′-ggcacaagctggagtacaacta-3′). hU6-IL-1β has an hU6 promoter driving expression of a shRNA targeting IL-1β (5′-gcaggcagtatcactcattgt-3′). hU6m-EGFP has a modified hU6 promoter driving expression of a shRNA targeting EGFP (5′-ggcacaagctggagtacaacta-3′) with the last 38 nucleotides changed to 5′-ctcgagtttatatatcttgtggaaaggatagtccacg-3′. hU6m-EGFPlm has a modified loop changed to 5′-ttggatccaa-3′.

Primers 5′-ggtatcgataaggtcgggcaggaaga-3′ and 5′-acaacgcgtccaaaaataaaatggattagcggtgtttcgtcctttccacaagat-3′ was used to amplify human U6 promoter from genomic DNA of HEK293 cells. The PCR fragment was digested with ClaI and MluI, and then inserted into pLENTIV5/Dest. The resulting plasmid was called pLENTI-hU6. Primers 5′-ggtatcgataag gtcgggcaggaaga-3′ and 5′-acaacgcgtccaaaaataaaatgcgtggactatcctttccacaa gatatataaactcgagaaatcgaaatactttcaagttac-3′ was used to amplify modified human U6 promoter. The PCR fragment was digested with ClaI and MluI, and then inserted into pLENTIV5/Dest. The resulting plasmid was called pLENTI-hU6m. An XcmI site was introduced upstream of MluI site. The PolIII transcription termination signal was contained in the XcmI site. Primers 5′-ggtatcgataagg tcgggcaggaaga-3′ and 5′-tgtacgcgtaa aaaaggcacaagctggagtacaactatctcttgaatagttgtactccagcttgtgcctggaaaggacgaa acacc-3′ were used to construct hU6-EGFP. Primers 5′-ggtatcgataaggt cgggcaggaaga-3′ and 5′-tgtacgcgtaaaaaaggcacaagctggagtacaactatctcttgaa tagttgtactccagcttgtgccgtggactatcctttccacaag-3′ were used to construct hU6m-EGFP. Primers 5′-ggtatcgataaggtcgggcaggaaga-3′ and 5′-tgtacgcg taaaaaaggcacaagctggagtacaactattggatccaatagttgtactccagcttgtgccgtggactatc ctttccacaag-3′ were used to construct hU6m-EGFPlm. PCR product was digested with ClaI and MluI and cloned into the pLENTI4V5/Dest digested with ClaI and MluI.

Results and discussion

Effect of modifications of hU6 promoter sequence and hairpin Loop on RNA interference

The shRNA expression vector pLENTI-hU6m contains a modified human U6 promoter sequence into which XhoI and MlyI was introduced. To test whether our modifications alter the activity of the promoter, a shRNA targeting EGFP with a modified human U6 promoter sequence was made. It was tested for RNA interference efficiency on CHO cells. After cotransfected with the plasmid pEGFP-N1, the fluorescent was examined 24 h after transfection. Fig. 1D shows that this shRNA with a modified human U6 promoter sequence could effectively suppress EGFP expression at a molar ratio of 3:1, and there is no obvious difference in suppression efficiency compared with positive control (Fig. 1C). Therefore, the modification doesn't affect the promoter activity.

The "loop" we created in the shRNA is a palindrome loop. We tested whether this change could effect the processing of shRNA to a functional siRNA. As we show in Fig. 1E, the shRNA target EGFP with this palindrome loop also suppressed the expression of EGFP efficiency.

On the basis of these tests, we constructed our shRNA library with the expression vector pLENTI-hU6m.

Generation of randomized shRNA Library

Fig. 2 outlines our method for the generating of a pool of double-stranded DNA comprising a large and diverse population of 19-bp inverted repeats from a synthesized randomized single strand DNA. The main steps are described below:

1. Chemically synthesizing a single strand DNA containing a randomized 19 nt sequence. After annealing, this synthesized sequence can form a secondary structure (we call it pre-hairpin) shown in Fig. 2. It contains three parts: (a) a loop, (b) a 6-bp complementary region, (c) a randomized 19-bp single strand. The single strand region was filled-in with a Taq DNA polyrease which

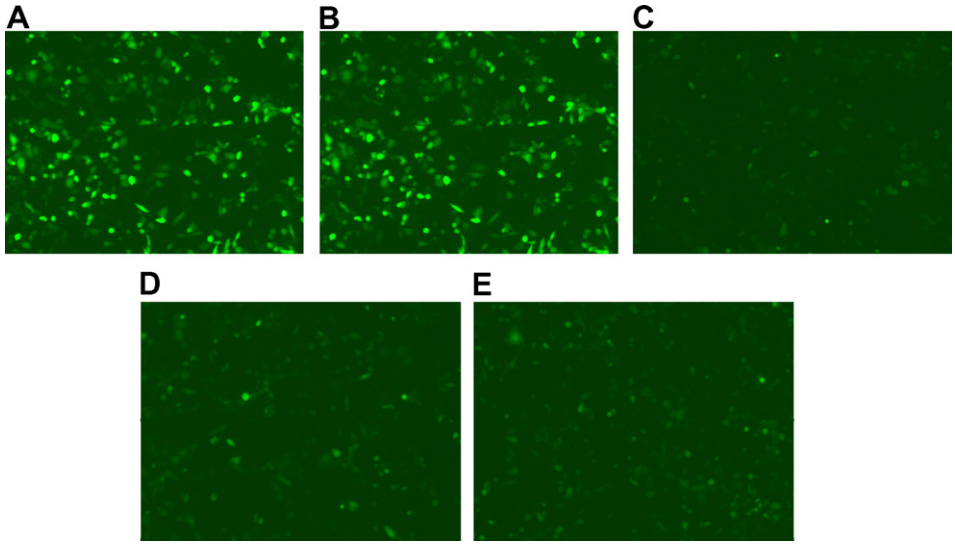

Fig. 1. Suppression of EGFP expression by cotransfection of shRNA constructs and pEGFP-N1. (A) pEGFP-N1 + pLENTI-hU6; (B) pEGFP-N1 + pLENTI-hU6-shRNA-IL-β; (C) pEGFP-N1 + pLENTI-hU6-shRNA-EGFP; (D) pEGFP-N1 + pLENTI-hU6m-shRNA-EGFP; (E) pEGFP-N1 + pLENTI-hU6m-shRNA-EGFPlm.

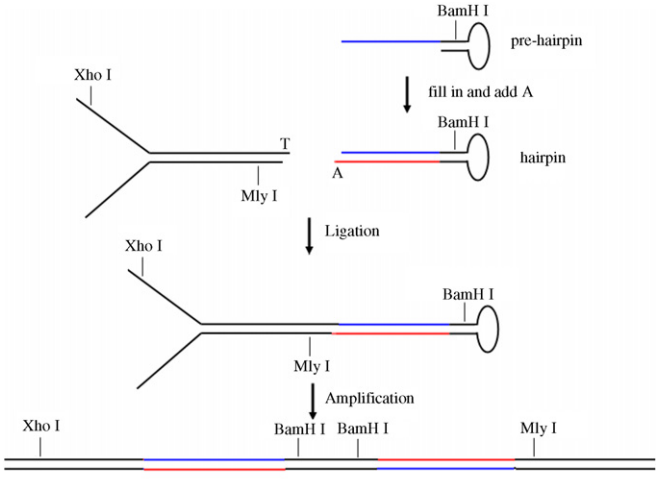

Fig. 2. Generating of a pool of double-stranded DNA comprising a large and diverse population of 19-bp inverted repeats from a synthesized randomized single strand DNA. See text for details. A single C-T mismatch was designed in MlyI site. After amplification, only one flanking sequence contains a MlyI site.

can add a single A at the 3′ end to form a randomized hairpin structure (Fig. 3A). This A prevents self-ligation in the next ligation reaction.

2. Generation of a DNA hairpin with noncomplementaty end. We designed a "Y" oligonucleotide with three features. (1) A single 3′ T overhang to avoid self-ligation. (2) Long noncomplementary arms designed for anchoring PCR primers with high melting temperatures. (3) A single basepair mismatch within the 18-bp stem region, resulting in the MlyI site on only one arm of the double-stranded PCR products. (4) A XhoI site within one noncomplementary arms, resulting in the XhoI site on only one arm of the double-stranded PCR products. We ligated this Y molecular to randomized hairpin (the ligation product was named YIU, Fig. 3B). Excess Y oligonucleotide was added to increase ligation efficiency.
3. PCR amplification of YIU. The ligated YIU molecular can be PCR-amplified with high melting temperatures. We designed two primers corresponding to the two non-complementary arms. The annealing and extension temperature during PCR cycling was set at 75 °C to destabilize intramolecular hairpins. In the late cycles of PCR amplification, product reannealing competes with primer annealing that causes the formation of byproduct, an X shaped product [9]. To minimize the product reannealing, an additional PCR cycle was performed with the addition of fresh PCR reaction mix (buffer, dNTPs, primers and polymerase) (Fig. 3C).
4. Cloning. After purification, the PCR product was digested with XhoI and MlyI and the desired 120-bp fragment was separated by Agarose gels (Fig. 3D). We first digested vector with XcmI, then with T4 DNA polymerase to get blunt-end, and subsequently cut with XhoI. After purification, fragment and vector were ligated. The result product still included excess 3′ loop. We have designed a BamHI site to remove the excess sequence. After digestion, the religation yielded expression-ready vectors. We transformed it into Stbl4 bacterial competent cells.
5. Library validation. We isolated and sequenced 20 independent constructs. Of these, 19 constructs contained inserts with the appropriate structures and all were unique (Table 1), one clone had no insert and one clone had insert containing TTTTT. Average G + C content is 43.8%. These results show that our methods can be used to generate a high-complexity library (1.2×10^7) with 95% of the clones encoding shRNA inserts of the appropriate size.

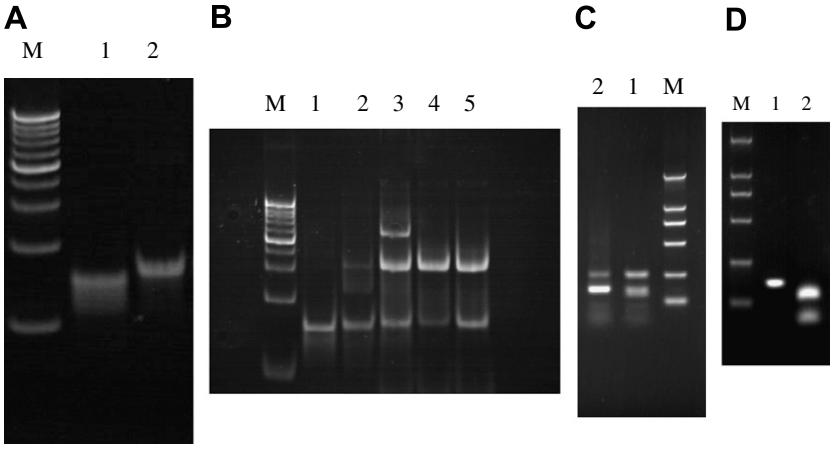

Fig. 3. Generation of randomized shRNA Library. (A) Pre-hairpin fill-in. M, 20 bp ladder; 1, pre-hairpin; 2, hairpin with a 3′ A overhang. (B) Y-Linker and hairpin ligation. M, 20 bp ladder; 1, hairpin with a 3′ A overhang; 2, hairpin self-ligation; 3, Y-Linker and hairpin ligation, the resulting products migrate at 120 bp region; 4, Y-Linker self-ligation; 5, Y-Linker, migrate at 60 bp region. (C) Conversion of the X molecular into double-stranded DNA by a PCR+1 protocol. M, DL-2000 (TaKaRa); 1, PCR amplification of YIU with 30 cycles results in an extra band seen in the 250 bp region. 2, When the PCR products are diluted with fresh PCR reaction mix and subjected to one additional PCR cycle, the 250 bp band disappeared. (D) Restriction digestion of double-stranded DNA of YIU. M, DL-2000 (TaKaRa); 1, the YIU double-stranded DNA generated by PCR; 2, the YIU double-stranded DNA generated by PCR is digested with XhoI and MlyI.

The size of a random library would be one of important factor to successful screen. A library that cover all genes, but has limited size would facilitate the screening. The theoretical complexity of randomly 19-mers is 2.75×10^{11}. It is too large to library maintenance and screening applications. According to Du et al. report, at positions 1, 2, 18, 19 of the target site, siRNA has the high levels of mismatch tolerance, and significant level of tolerance at positions 3, 4, 12–17 was found [10,11]. This observation was corroborated by the fact that some of the genes that share 15- to 16-nt identities with the siRNA can be efficiently modulated by the siRNA [12]. Taking this aspect into consideration, the effective complexity of a fully randomized 19-mer shRNA library could be only $\sim 1 \times 10^8$, The library we constructed contains about 1.2×10^7 independent clones from 1 μg of starting plasmid vector, which means that for any mRNA of around 2 kb from any organism, there will be 24–240 shRNA against it in our library.

Compared with generation of shRNA library by enzymatic engineering of cDNA, construction of a randomized shRNA library is more rapid and convenient. It could be used to screen all possible genes in different cell types and different species where shRNA is applicable.

Table 1
Sequence isolated from the randomized shRNA library

	Sequence	GC%
1	ccaggtgcttacccacaag	57.9
2	atttcaa*ttttt*gtcattc	21.1
3	cttaggctgggctgctacg	63.2
4	tgacaggaaatcgtcaggc	52.6
5	ctatccttggttaatatgt	31.6
6	aagctgaaatactccggaa	42.1
7	ctgacttgaaagccgaaac	47.4
8	ttagcaagatttggcgata	36.8
9	gcatcacttaatgtaccta	36.8
10	agctgtgcccgtaaatgac	52.6
11	gtattttctgcaagcacat	36.8
12	atcttcctgttcaaatgta	31.6
13	no insert	—
14	gaacaacaacaagtggcaa	42.1
15	acttctccacattgccttc	42.1
16	ctgggcgactttgaattgc	52.6
17	gtgcttggagctacacagc	57.9
18	atctaagtatattgttctc	26.3
19	agctgctcacgatggtgaa	52.6
20	tcgggtgctagttttcatc	47.4

References

[1] A. Fire, S. Xu, M.K. Montgomery, S.A. Kostas, S.E. Driver, C.C. Mello, Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*, Nature 391 (1998) 806–811.
[2] J. Downward, Use of RNA interference libraries to investigate oncogenic signaling in mammalian cells, Oncogene 23 (51) (2004) 8376–8383.
[3] C. Sachse, C.J. Echeverri, Oncology studies using siRNA libraries: the dawn of RNAi-based genomics, Oncogene 23 (51) (2004) 8384–8391.
[4] K. Berns, E.M. Hijmans, J. Mullenders, T.R. Brummelkamp, A. Velds, M. Heimerikx, R.M. Kerkhoven, M. Madiredjo, W. Nijkamp, B. Weigelt, R. Agami, W. Ge, G. Cavet, P.S. Linsley, R.L. Beijersbergen, R.A. Bernards, A large-scale RNAi screen in human cells identifies new components of the p53 pathway, Nature 428 (6981) (2004) 431–437.
[5] P.J. Paddison, J.M. Silva, D.S. Conklin, M. Schlabach, M. Li, S. Aruleba, V. Balija, A. O'Shaughnessy, L. Gnoj, K. Scobie, K. Chang, T. Westbrook, M. Cleary, R. Sachidanandam, W.R. McCombie, S.J. Elledge, G.J. Hannon, A resource for large-scale RNA-interference-based screens in mammals, Nature 428 (6981) (2004) 427–431.
[6] D. Shirane, K. Sugao, S. Namiki, M. Tanabe, M. Iino, K. Hirose, Enzymatic production of RNAi libraries from cDNAs, Nat. Genet. 36 (2) (2004) 190–196.
[7] G. Sen, T.S. Wehrman, J.W. Myers, H.M. Blau, Restriction enzyme generated siRNA (REGS) vectors and libraries, Nat. Genet. 36 (2) (2004) 183–189.
[8] B. Luo, A.D. Heard, H.F. Lodish, Small interfering RNA production by enzymatic engineering of DNA (SPEED), Proc. Natl. Acad. Sci. USA 101 (15) (2004) 5494–5499.

[9] C. Du, B. Ge, Z. Liu, K. Fu, W.C. Chan, T.W. McKeithan, PCR-based generation of shRNA libraries from cDNAs, BMC Biotechnol. 21 (2006).
[10] Q. Du, H. Thonberg, J. Wang, C. Wahlestedt, Z. Liang, A systematic analysis of the silencing effects of an active siRNA at all single-nucleotide mismatched target sites, Nucleic Acids Res. 33 (5) (2005) 1671–1677.
[11] T. Holen, S.E. Moe, J.G. Sorbo, T.J. Meza, O.P. Ottersen, A. Klungland, Tolerated wobble mutations in siRNAs decrease specificity, but can enhance activity in vivo, Nucleic Acids Res. 33 (15) (2005) 4704–4710.
[12] A.L. Jackson, S.R. Bartz, J. Schelter, S.V. Kobayashi, J. Burchard, M. Mao, B. Li, G. Cavet, P.S. Linsley, Expression profiling reveals off-target gene regulation by RNAi, Nat. Biotechnol. 21 (6) (2003) 635–637.