Size Unbiased Representative Enzymatically Generated RNAi (SULER) Library and Application for RNAi Therapeutic Screens

Tiejun Li,^{1-3,*} York Yuanyuan Zhu,^{1-3,*} Li Chen,^{3,*} Yuncheng Sun,^{1,2} Jian Yuan,^{1,2}
Michael Graham,⁴ and Peter French⁴

RNA interference (RNAi) libraries screens have become widely used for small RNA (sRNA) therapeutic targets development. However, conventional enzymatically libraries, typically prepared using the type 2 restriction enzyme *MmeI*, produce sRNAs between 18 and 20 bp, much shorter than the usual lengths of 19–23 bp. Here we develop a size unbiased representative enzymatically generated RNAi (SULER) library, which employs type 3 restriction modification enzyme *EcoP15I* to produce sRNAs ranging from 19 to 23 bp using a group of rationally designed linkers, which can completely mimic the length of sRNAs naturally generated by Dicer enzyme in living cells, and the screening results of SULER libraries showed high recombination rate and knockdown efficiency. SULER library provides a useful tool for RNAi therapeutics screening in a fast and simple way.

Introduction

IN RECENT YEARS, screening of random libraries targeting particular genes has become an attractive and useful method to define small RNA (sRNA) therapeutics, including small interfering RNA (siRNA), short hairpin RNA (shRNA) [1–3], and other modified siRNAs [4]. Small RNA library-based approaches have several advantages for unknown sequence discovery, including immediate clone availability, cost-effectiveness over chemically synthesized RNA, and the potential to define siRNAs and/or shRNA constructs for long-term gene silencing based on sequences determined from such libraries [5–7].

Small interfering RNAs are a class of 19- to 23-bp double-stranded RNAs (dsRNAs) endogenously generated or artificially designed that trigger gene silencing via RNA interference (RNAi) pathways. RNAi can be induced by exogenously introduced [1] or endogenous siRNAs [8–10], the latter potentially derived from endogenous dsRNA precursors arising from “natural” processes such as bidirectional transcription, transcription of inverted repeats, or hybridization of pseudogene transcripts to mRNAs [11–14]. Once dsRNAs are processed by Dicer, siRNA duplexes enter the RNA-induced silencing complex, where conserved components including Argonaute (Ago) proteins unwind and cleave the passenger

strand in the siRNA duplex, leaving the guide strand to pair with the target mRNA via perfect sequence complementarity, which in turn guides RNA cleavage [15].

MicroRNAs (miRNAs) are another class of endogenous sRNA (19–23 nt), which have been studied as regulators of gene expression in crucial biological processes, including cell development, differentiation, apoptosis, and proliferation [16,17]. MicroRNAs are post-transcriptional regulators that are processed from endogenously expressed transcripts [18,19] and have diverse functions in biology and disease.

RNA activation is a newly discovered mechanism of gene regulation triggered by small activating RNAs (saRNAs) that targets gene promoter regions instead of coding sequences and have been shown to activate the expression of endogenous genes [20–23], presenting a novel and natural tool for overexpressing functionally important genes for disease treatment.

Several strategies to enzymatically prepare shRNA libraries have been described, including EPRIL (enzymatic production of RNAi library) [24], REGS (restriction enzyme-generated siRNA) [25], SPEED (siRNA production by enzymatic engineering of DNA) [26], and others [27–30]. These are similar in principle but differ in details of their production. Typically, double-stranded (ds) complementary DNA (cDNA) is randomly cleaved into small fragments by

¹Small RNA Technology and Application Institute and ³Department of Pathological Anatomy, Nantong University, Nantong, China.

²Biomics Biotechnologies Co., Ltd., Nantong, China.

⁴Benitec Biopharma Ltd., Sydney, Australia.

*These authors contributed equally to this work.

DNaseI (or restriction enzymes in some instances) and subsequently ligated to an artificial adaptor that contains an *MmeI* restriction site, which specifies cleavage 18–20 bp from the recognition site [31]. The generated siRNAs only have a maximum size of 20 bp [32]. However, siRNAs generated by Dicer in mammalian cells typically range from 21 to 23 bp; such sequences are therefore underrepresented in conventional libraries. Moreover, conventional approaches use polymerase chain reaction (PCR) amplification of palindromic structures, which can lead to reduction in library complexity and potential loss of the best therapeutic molecules [28]. Size unbiased representative enzymatically generated RNAi (SURER) libraries avoid these issues.

Materials and Methods

Construction of SURER Library

Complementary DNA fragments corresponding to coding sequences (CDSs) of hepatitis B virus (HBV) DNA polymerase, baculoviral IAP repeat-containing 5 (BIRC5), and a conserved domain the microphthalmia-associated transcription factor (MITF) conserved domain, were prepared by Biomix Biotechnologies Co., Ltd. Sequences of linkers and primers used for preparing SURER libraries are listed in Table 1, libraries were prepared using the procedures that follow.

Step 1. *DNaseI* partial digestion. One hundred nanograms of the target cDNAs were partially digested to prepare 100–300 bp blunt-end fragments using 0.01–0.03 U *DNaseI*

(Roche) in a Mn^{2+} buffer at a final concentration of 10 mM Tris-HCl (pH 9.0), 2 mM $MgSO_4$, 10 mM KCl, 8 mM $(NH_4)_2SO_4$, and 1 mM $MnCl_2$. The reaction was performed on ice for 1 min followed by heat inactivation for 20 min at 70°C. The resulting products were analyzed on 1% agarose gel.

Step 2. Loop-1 linkers ligation. The partially digested products were ligated to individual loop-1 linkers (Sigma-Aldrich). These contained self-complementary sequences, which included recognition sequences for *EcoPI5I* (CTGCTG) and *FokI* (CATCC) (sequences are listed in Table 1). Ligations contained 2.5 μ L partially digested DNA from step 1, 0.5 μ L 10 $\times$ Ligation Reaction Buffer (NEB), 1 μ L (10 mM) loop-1 linker; 0.5 L 10 mM ATP (Sigma-Aldrich), and 0.5 μ L T4 DNA Ligase (NEB); and were incubated at 16°C overnight.

Step 3. Single primer PCR (PCR-1). Loop-1 ligation products were then amplified using single primer PCR with the 5'BH1 primer (Sigma-Aldrich). PCR-1 amplification reactions contained 0.5 μ L loop-1 linker ligation product, 2 μ L single primer 5'BH1 (20 μ M), 1 μ L dNTP (10 mM, Sangon), 5 μ L 10 $\times$ PCR buffer [200 mM Tris-HCl (pH 8.4), 500 mM KCl, 15 mM $MgCl_2$], 0.5 μ L Taq DNA polymerase (Clontech), and double-distilled water (ddH₂O) added to a final volume of 50 μ L. Reactions were preheated at 95°C for 1 min, followed by 28 amplification cycles of 95°C for 15 s and 68°C for 1 min. PCR products were analyzed on 1% agarose gel. PCR-1 products were purified by extracting with two volumes of phenol:chloroform:isoamyl alcohol (25:24:1) (Sangon) and precipitated with 1/10 volume of 3 mol/L NaOAc (Sangon),

TABLE 1. SEQUENCES OF LINKERS AND PRIMERS

Oligo name	Sequence (5'-3')
19-bp Loop-1 linker	pCTTTTTTCTGCTGCATCCCTGAACTGGGATCCGTTGGATGTGTACACATCCAA CGGATCCCAGTTCAGGGATGCAGCAGAAAAAAG
20-bp Loop-1 linker	pCTTTTTTCTGCTGGCATCCCTGAACTGGGATCCGTTGGATGTGTACACATCCAAAC GGATCCCAGTTCAGGGATGCCAGCAGAAAAAAG
21-bp Loop-1 linker	pCTTTTTTCTGCTGGGCATCCCTGAACTGGGATCCGTTGGATGTGTACACATCCAAAC GGATCCCAGTTCAGGGATGCCAGCAGAAAAAAG
22-bp Loop-1 linker	pCTTTTTTCTGCTGGGGCATCCCTGAACTGGGATCCGTTGGATGTGTACACATCCAAAC GGATCCCAGTTCAGGGATGCCCCAGCAGAAAAAAG
23-bp Loop-1 linker	pTTTTCTGCTGGGGGCATCCCTGAACTGGGATCCGTTGGATGTGTACACATCCAAAC GGATCCCAGTTCAGGGATGCCCCAGCAGAAAAA
Loop-2 linker	pCTTTTTGAGACCGACATCCCTCTGCAGACGATCCATCAGAGTCAGCTGACTCTGA TGGATCGTCTGCAGAGGGATGTCGGTCTCAAAAAG
5'BH1 primer	ACACATCCAACGGATCCCAGTTCAG
3'LG primer	GACTCTGATGGATCGTCTGCAGAG
5'U6 primer	AAGGTCGGGCAGGAAGAGGGC
3'H1 primer	TATTTGCATGTCGCTATGTGTTCT
NC sense	TTCTCCGAACGTGTCACGTTT
NC antisense	ACGTGACACGTTCCGAGAATT
HBV forward primer	TGTGGTTATCCTGCGTTAATG
HBV reverse primer	GCGTCAGCAAACACTTGG
MITF forward primer	CATCACCTTCAACAACAAC
MITF reverse primer	ATGCTCATACTGCTCCTC
BIRC5 forward primer	ACCGCATCTCTACATTCAAG
BIRC5 reverse primer	CAAGTCTGGCTCGTTCTC
GAPDH forward primer	GAAGGTGAAGGTCGGAGTC
GAPDH reverse primer	GAAGATGGTGATGGGATTTC

NC, negative control; p, phosphate; BIRC5, baculoviral IAP repeat-containing 5; GAPDH, glyceraldehyde 3-phosphate dehydrogenase; HBV, hepatitis B virus; MITF, microphthalmia-associated transcription factor.

1.5 μ L Glycogen (20 mg/mL, Sangon), and 2.5 fold volumes of ice-cold ethanol (Sangon); samples were kept at -20°C for 4 hrs, DNA pellets were collected by centrifugation and resuspended in 20 μ L ddH₂O and stored at -20°C .

Step 4. *EcoP15I* digestion. The PCR-1 products from the five (19-, 20-, 21-, 22-, and 23-bp) loop-1 linkers were mixed in equal quantities and then digested with *EcoP15I* (NEB). Ten microliters PCR-1 product mixtures were restricted in 100- μ L reactions containing 10 μ L 10 $\times$ NEBuffer-3 (NEB); 10 μ L ATP (10 mM, Sigma-Aldrich); 1 μ L bovine serum albumin (BSA) (10 mg/mL, NEB); 10 μ L *EcoP15I* (100 U, NEB); and 59 μ L ddH₂O overnight at 37°C . *EcoP15I* digestion products were purified by phenol-chloroform extraction and alcohol precipitation as described above. DNA pellets were collected by centrifugation and resuspended in 11 μ L ddH₂O.

Step 5. End-filling and loop-2 linker ligation. *EcoP15I*-generated sticky ends were end-filled with T4 DNA polymerase by adding 1.5 μ L T4 DNA polymerase (4.5 U), 1.5 μ L NEBuffer-2 (NEB), and 2 μ L dNTP (1 mM, Sangon) to a total 15- μ L reaction and incubated at 37°C for 15 min. Fragments of appropriate sizes (63-67 bp) were purified from 20% Tris-Borate-Ethylene Diamine Tetraacetic Acid (TBE) polyacrylamide gels using QIAEX II Gel Extraction Kit (Qiagen) and eluted with ddH₂O. The polyacrylamide gel electrophoresis (PAGE)-purified blunt-ended DNA fragments were ligated to a loop-2 linker containing a *FokI* recognition sequence (CATCC). Next, 2.5 μ L of DNA isolated by PAGE from step 4 was supplemented with 0.5 μ L 10 $\times$ ligation reaction buffer [500 mM Tris-HCl (pH 7.5, 25°C), 100 mM MgCl₂, 100 mM DTT, 25 μ g/mL BSA]; 1 μ L loop-2 linker (20 mM, Sigma-Aldrich); 0.5 μ L ATP (10 mM, Sigma-Aldrich); 0.5 μ L T4 DNA Ligase (NEB); and 0.5 μ L ddH₂O. Reactions were incubated at 16°C overnight.

Step 6. PCR-2. Ligation products were diluted with 45 μ L ddH₂O, and 0.5 μ L of these used as templates for PCR-2 amplification. Reactions contained 0.5 μ L diluted ligation products; 5 μ L 10 $\times$ PCR buffer [200 mM Tris-HCl (pH 8.4); 500 mM KCl; 15 mM MgCl₂]; 1 μ L 3'LG primer (10 μ M, Sigma-Aldrich); 1 μ L 5'BH1 primer (10 μ M, Sigma-Aldrich); 1 μ L dNTP (10 mM, Sangon); and 0.5 μ L Taq DNA polymerase (Clontech) in a final volume of 50 μ L. Reactions were preheated at 95°C for 1 min and amplified using 28 cycles of 95°C for 15 sec and 68°C for 1 min. After amplification, 10 μ L of PCR products were analyzed on a 20% TBE polyacrylamide gel, and the remaining fragments purified using phenol-chloroform extraction and ethanol precipitation as described above; DNA pellets were resuspended in 20 μ L ddH₂O.

Step 7. *FokI* digestion. PCR-2 products were digested with *FokI* enzyme; 50- μ L reactions contained 20 μ L purified PCR-2 products (from step 6), 2 μ L *FokI* (10 U, NEB), 5 μ L NEBuffer-3 (NEB), and 23 μ L ddH₂O, which was incubated at 37°C for 100 min. After digestion, 10 μ L of products were analyzed on a 20% TBE polyacrylamide gel and appropriate bands purified by PAGE as described above.

Step 8. Cloning to RNAi expression vector. *FokI*-digested fragments were ligated into linearized pRNAi-U6H1/Neo.

pRNAi-U6H1/Neo was digested with *BsmBI* and dephosphorylated with calf intestinal alkaline phosphatase according to manufacturer's (NEB) protocols. One hundred nanogram vectors were ligated to 2 μ L gel-purified *FokI* digested insert DNA from step 7 using T4 DNA ligase according to the manufacturer's protocol (NEB). Ligations were purified by ethanol precipitation and then electrotransformed into MegaX DH10BTM Cell (Invitrogen) using a MicroPulser (Bio-Rad) according to the manufacturer's protocols and transformed colonies selected with Kanamycin.

Preparation of pRNAi-U6H1/Neo

The pRNAi-U6H1/Neo vector was modified from pDsIPHER-GFP vector (MolecularA). An insert DNA fragment was synthesized (Biomics), this contained the human U6 and H1 promoter sequences, two inversely oriented *BsmBI* restriction sites, and an *SfiI* site: these sequences were flanked by *HindIII* or *MluI* sites and were cloned into *HindIII*/*MluI* digested pDsIPHER-GFP. The sequence of the fragment was: 5'-AAGCTTCTCGAGGAATTCAAGGTCGGGCAGGAAGAGGGCCTATTTCCCATGATTCCTTCATATTTGCATATACGATACAAGGCTGTTAGAGAGATAATTAGAATT AATTTGACTGTAAACACAAAGATATTAGTACAAAA TACGTGACGTAGAAAGTAATAATTTCTTGGGTAGTT TGCAGTTTTTAAAATTATGTTTTTAAAATGGACTATCA TATGCTTACCGTAACTTGAAAGTATTTTCGATTTCTTG GGTTTATATATCTTGTGGAAGGACGAAAAAAGAG ACGGCCGTGTTGGCCGTCTCTTTTTTGAGTGGTCTCA TACAGAACTTATAAGATTCCCAAATCCAAAGACATT TCACGTTTATGGTGATTTCCCAAGACACATAGCGCATGCAAATATCTCGAGACGCGT-3'.

Validation of cloning efficiency in SURER library

To analyze SURER libraries, bacterial colonies were randomly picked and inoculated into 400 μ L Lysogeny broth medium [1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1% (w/v) NaCl, pH 7.0] in 48-well plates. Cultures were incubated at 37°C with a continuous shaking for 2 h at 220 rpm. Cultures were screened with PCR by 5'U6 primer and 3'H1 primer (Invitrogen). Reactions contained 2 μ L cultured bacteria, 0.5 μ L 5'U6 primer (10 μ M), 0.5 μ L 3'H1 primer (10 μ M), 0.5 μ L dNTP (10 mM, Sangon), 3 μ L 10 $\times$ PCR buffer [200 mM Tris-HCl (pH 8.4), 500 mM KCl, 15 mM MgCl₂], 0.3 μ L Taq DNA polymerase (Clontech), and ddH₂O to a final volume of 30 μ L. Reactions were preheated at 95°C for 5 min and amplified for 25 cycles at 95°C for 15 s, 62°C for 30 s, and 72°C for 40 s, then held at 72°C for 7 min for final extension. Five microliters of each PCR product were analyzed on 1% agarose gel. *SfiI* digestion was used to define clones containing inserts; 4 μ L of the above PCR reactions were digested with *SfiI* according to the manufacturer's protocol (NEB). Products were analyzed on 1% agarose gel, and plasmids of validated clones prepared using Plasmid Mini Kit (Qiagen) and sequenced with 5'U6 or 3'H1 primer.

sRNA expression cassettes preparation

Short RNA expression cassettes for cell transfection were prepared using PCR with 5'U6 and 3'H1 primers. The PCR reaction contained 1 μ L plasmid of sRNA clone (10 ng), 1 μ L

5'U6 primer (10 μ M), 1 μ L 3'H1 primer (10 μ M), 1 μ L dNTP (10 mM, Sangon), 5 μ L 10 $\times$ PCR buffer [200 mM Tris-HCl (pH 8.4), 500 mM KCl, 15 mM MgCl₂], 0.5 μ L Taq DNA polymerase (Clontech), and ddH₂O to a final volume of 50 μ L. Reactions were preheated at 95°C for 5 min and amplified for 25 cycles of 95°C for 15 s, 62°C for 30 s, and 72°C for 40 s, then held at 72°C for 7 min for final extension. Two microliters of each PCR product was analyzed on 1% agarose gel.

As a negative control (NC), 100 μ M NC sense and anti-sense strand (Table 1, Biomix) were annealed by using 95°C heating for 10 min and cooled to room temperature. Two microliters of annealed oligos were cloned into linearized, phosphorylated pRNAi-U6H1/Neo as described above. Reactions were purified and transformed into *Escherichia coli* (*E. coli*), plasmids were sequenced as described above, and an appropriate clone was selected.

Cell culture and transfection

Human HepG2 2.2.15, A431, and HepG2 cell lines (Biomix) were used to assay silencing. Cells were cultured in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum (Gibco) and incubated at 37°C in 5% CO₂. One hundred thousand cells per well were seeded 24 h before transfection into 96-well plates at 70%–80% confluence. Transfection of 100 ng sRNA expression cassettes and 50 nM chemical synthesized sRNAs were conducted with Lipofectamine® 2000 (Invitrogen) according to the manufacturer's recommendations.

Real-time quantitative PCR analysis

The levels of mRNAs encoding HBV DNA polymerase, MIF, and BIRC5 were determined by real-time quantitative PCR (RT-qPCR). At 48 h post-transfection, mRNAs were extracted using TurboCapture mRNA Kit (Qiagen) using the manufacturer's protocol and used as templates for RT-qPCR reactions. The primers used to detect HBV, MIF, and BIRC5 as shown in Table 1. Target mRNA levels were normalized to glyceraldehyde phosphate dehydrogenase (GAPDH) mRNA levels. Triplicate RT-qPCR reactions were analyzed for each mRNA sample. The 25- μ L RT-qPCR reactions contained 4 μ L template RNA (50 ng), 12.5 μ L of 2 $\times$ SensiMix One-Step (Quantance), 1 μ L forward and reverse primers (10 μ M, Biomix), 0.5 μ L 50 $\times$ SYBR Green I (Quantance), and RNase-free water to 25 μ L. Reverse transcription was at 42°C for 30 min; reactions were then preheated at 95°C for 7 min, followed by 45 amplification cycles, which involved denaturation at 95°C for 20 sec, annealing at 60°C for 30 sec, and extension at 72°C for 30 sec.

Three siRNAs—c-2, c-38, and c-42—were synthesized; their sequences were based on the sequences of clones 2, 38, and 42 from the BIRC5 SURER library. These were transfected into HepG2 cell line and mRNAs extracted for the RT-qPCR method described above; the NC sRNA (sense: UUCUCCGAACGUGUCACGUGdTdT; antisense: ACGUGA CACGUUCGAGAAAdTdT) was used negative control.

Western blot analysis

HepG2 cells were plated at a concentration of 1×10^6 cells per well in 6-well plates and grown for 24 h until they reached 70%–80% confluence then transfected with sRNAs c-2, c-38,

and c-42. Cells were harvested 48 h after transfection lysed in ice-cold cell M-PER Mammalian Protein Extraction Reagent (Pierce), and proteins separated on SDS-PAGE and transferred to polyvinylidene fluoride membranes (Millipore). To detect BIRC5, filters were probed with a polyclonal rabbit anti-human BIRC5 (Abcam) followed by goat anti-rabbit immunoglobulin G–horseradish peroxidase (IgG-HRP) (Santa Cruz). β -actin levels were used for controls and were determined using mouse anti-human β -actin (Santa Cruz) followed by goat anti-mouse IgG-HRP (Santa Cruz). Proteins were visualized by chemiluminescence with ECL (Promega). The relative amount of proteins on the blots was determined by Image J software (National Institutes of Health).

Cell proliferation assay

HepG2 cell proliferation was measured by Cell Counting Kit-8 (CCK-8) detection kit (Dojindo). Cells were plated at a concentration of 5×10^3 cells per well in 96-well plates. At 24 h, 48 h, 72 h, and 96 h after siRNA transfection, CCK-8 solution was applied followed by 2 h incubation at 37°C. Absorbance values of all wells were then determined at 490 nm in Microplate Reader (Bio-Rad).

Statistical analysis

All experiments were performed independently three times, the results were shown as mean $\pm$ standard deviation, and statistical analyses were performed using SPSS19.0 software. The differences between groups were compared using Student's t-test to assess statistical significance. All *P* values were based on a two-sided statistical analysis, and *P* < 0.05 was considered to indicate statistical significance.

Results

SURER library preparation

SURER library construction is shown diagrammatically in Fig. 1, a detailed protocol is described in "Materials and Methods," and primers used to construct libraries are listed in Table 1. First, double-stranded DNAs (dsDNAs) were randomly and partially digested to produce 100- to 300-bp fragments by *DNAseI* in the presence of Mn²⁺ ions [33] (Fig. 1, Step 1). The fragments were then ligated to individual loop-1 linkers (Fig. 2A), which contained the recognition sequences of *EcoP15I* and *FokI* adjacent to a universal PCR anchor sequence (Fig. 1, Step 2). As shown in detail in Fig. 2A, five different loop-1 linkers were used to generate independent pools of 19-, 20-, 21-, 22-, and 23-bp cDNA inserts. *EcoP15I* cleaves 25–27 bp outside of its recognition site, thus adjusting the number of the poly A/T sequence between the *EcoP15I* recognition site in loop-1 linker(s) and cDNA fragments, specifying the size of restriction fragments (e.g. 19–23 bp) generated by this protocol. Single loop-1 linkers can be used if a distinct length of sRNA is required.

Efficient cleavage of dsDNAs by *EcoP15I* requires the presence of two inversely oriented restriction sites with head-to-head orientations [34]. After blunt-end ligation of loop-1 linker(s) to dsDNAs, a single primer PCR (PCR-1) was performed (Fig. 1, Step 3) for each linker ligation; these were then pooled for subsequent steps.

EcoP15I digestion was carried out after PCR-1 (Fig. 1, Step 4) to generate 19- to 23-bp fragments derived from the

FIG. 1. General scheme for SURER (size unbiased representative enzymatically generated RNA interference [RNAi]) library construction. *Step 1:* Partial *DNaseI* digestion of gene-specific complementary DNAs (cDNAs) to produce 100- to 300-bp fragments. *Step 2:* Three kinds of molecules (-a, -b, and -c) with different linker orientations are generated after loop-1 linker ligation. Only molecule-a can be digested by *EcoP15I* since the enzyme requires the presence of two inversely oriented recognition sites for efficient cleavage. The loop-1 linker contains the recognition sites of *EcoP15I* (to generate fragments of appropriate size) and *FokI* (for cloning into an expression vector). *Step 3:* Single primer polymerase chain reaction (PCR) amplification (PCR-1). *Step 4:* *EcoP15I* digestion, which generates 19- to 23-bp small RNA (sRNA) encoding fragments from PCR-1 amplified fragments. *Step 5:* End-filling by T4 DNA polymerase followed by loop-2 linker ligation. *Step 6:* Double primers PCR amplification (PCR-2) to amplify appropriate fragments. *Step 7:* *FokI* digestion to generate sRNA fragments. *Step 8:* Cloning into linearized pRNAi-U6H1/Neo.

original cDNA. The loop-1 linkers contained sites for both *EcoP15I* and a type-2 restriction site, *FokI*, adjacent to a PCR anchor sequence (Fig. 2A). *EcoP15I* cleaves 25–27 bp outside of their recognition sequences, thus adjusting the number of the poly “A/T” sequence between the *EcoP15I* recognition site in loop-1 linker(s) and cDNA fragments, specifying the size of restriction fragments (e.g. 19–23 bp) generated by this protocol.

EcoP15I cleavage leaves two base-pair 5' overhangs; these were then end-filled using T4 DNA polymerase in the presence of dNTPs [35] (Fig. 1, Step 5). After gel purification, the blunt-ended DNAs were ligated to the loop-2 linker (Fig. 2B). These ligation products were then used as templates for the second PCR amplification (PCR-2; Fig. 1, Step 6). The loop-2 linker creates a PCR anchor and specifies *FokI* restriction sites for subsequent cloning of amplified fragments. *FokI* cleavage of PCR-2 amplified products generates 4-nt 5' overhangs of amplified fragments, in this instance 5' with AAAA overhangs at each end (Fig. 1, Step 7).

The *FokI* digested products were then cloned into the RNAi expression vector pRNAi-U6H1/Neo (Fig. 1, Step 8). A map of this is shown in Fig. 3, it contains two opposing polymerase III (pol III) promoters (U6 and H1) separated by a

polylinker. The polylinker contains two *BsmBI* sites, cleavage of which generated cloning vectors with TTTT 3' overhangs, compatible with the *FokI* digested PCR-2 products (Fig. 1, Step 7). These poly T and poly A tracts act as pol III transcriptional terminators. The polylinker also contains a *SfiI* restriction site, loss of which can be conveniently used to screen colonies for recombinant molecules.

Ligations were transformed into *E. coli*, to generate pools of recombinants containing fragments of cDNAs expected to range from 19 to 23 bp. Resultant colonies were screened with PCR by 5'U6 and 3'H1 primers (Table 1) and clones lacking the diagnostic *SfiI* restriction site were sequenced to characterize libraries.

SURER library from cDNA encoding DNA polymerase of Hepatitis B virus

A full-length cDNA of HBV DNA polymerase (NCBI Accession No. U95551) CDS was synthesized as two fragments of 1625 bp and 874 bp (Fig. 4A). DNA was randomly and partially digested to generate blunt-ended fragments using *DNaseI*. The preferable length for library construction was about 100–300 bp (Fig. 4B).

FIG. 4. Gel electrophoretic analysis showing steps in preparation of SURER library from HBV DNA polymerase cDNA. (A) The full-length cDNA of HBV DNA polymerase was synthesized as two fragments of 1625 bp (1) and 874 bp (2). (B) Partial digestion of 1625 bp (1) and 874 bp (2) target DNAs with *DNaseI* generated fragments ranging from 100 to 300 bp. (C) Single primer PCR amplification (PCR-1) to generate appropriate substrates for library preparation. Substrates for these reactions were ligations using the 19-bp loop-1 linker (1); 20-bp loop-1 linker (2); 21-bp loop-1 linker (3); 22-bp loop-1 linker (4); or 23-bp loop-1 linker (5). (D) *EcoP15I* cleavage of pooled PCR-1 fragments. (E) PCR-2 amplification using 5'BH1 and 3'LG primers. (F) *FokI* digestion of PCR-2 amplified products. Diagrammatic representations of products are shown to the right. M, 1kb plus DNA ladder (Invitrogen); LMW, low molecular weight DNA ladder (NEB).

$\geq 50\%$, 16 of which resulted in $\geq 70\%$ knockdown efficiency (Fig. 6A). Sixty-three sRNA clones from the MITF SURER library were screened for activity; 32 sRNAs were effective in silencing MITF by $\geq 50\%$, 14 of these resulted in $\geq 70\%$ knockdown efficiency (Fig. 6B). Twenty-six sRNA clones from the BIRC5 SURER library were screened for activity; 13 sRNAs were effective in silencing BIRC5 by $\geq 50\%$, 4 of these resulted in $\geq 70\%$ knockdown efficiency (Fig. 6C).

To confirm the activity of BIRC clones, three siRNAs were chemically synthesized based on the sequences of clones 2, 38, and 42 (c-2, c-38, c-42) and assayed for knockdown of the endogenous BIRC5 gene in HepG2 cells, and the three designed [37] and synthetic siRNAs (s-1, s-2, s-3) used for comparison. Levels of BIRC5 mRNA and protein were determined in siRNA treated cells 48 h after transfection using RT-qPCR and western blots. The result of RT-qPCR showed that the knockdown efficiencies of designed siRNAs were all lower than 50%, not stronger than screened siRNAs ($> 80\%$) (Fig. 7A). Furthermore, BIRC5 protein was strongly reduced in HepG2 cells treated with c-2, c-38, and c-42 compared with negative control (c-NC) treated and untransfected cells

(Fig. 7B). Silencing of BIRC5 was expected to inhibit cell proliferation in siRNA transfected cells. CCK-8 assays in cells transfected with c-2, c-38, and c-42 demonstrated reduced signals at 48 h, 72 h, and 96 h post-transfection compared with c-NC treated and untransfected cells (Fig. 7C), indicating reduced proliferation.

Discussion

The SURER library protocol has several advantages over alternative methods. Importantly the lengths of sRNAs expressed from SURER libraries are representatively distributed from 19 to 23 bp, as a consequence of using different loop-1 linkers. In addition, the SURER cloning strategy does not result in incorporation of additional sequences, such as linkers, into clones. Moreover, the PCR amplification steps incorporated into the library construction process do not seem to appreciably bias clone representation (Fig. 5) and also ensures adequate materials are available for the various steps in the protocol.

The representative sRNA size distribution from 19 to 23 bp in all SURER libraries was achieved using five rationally

FIG. 5. Site and size distribution profiles of clones randomly selected from HBV, MITF, and BIRC5 SURER libraries. The *x*-axis represents cDNA sequences; sequences of sRNA-expressing clones are aligned above these. The *y*-axis denotes the size of sRNA clones (≤ 19 , 20, 21, 22 and ≥ 23). (A) 604 clones from HBV DNA polymerase SURER library. (B) 120 clones from MITF conserved domain SURER library. (C) 59 clones from BIRC5 SURER library.

designed loop-1 linkers (Fig. 2A) where the size of the resulting cDNA fragments is controlled by the number of A nucleotide (nt) between cDNA fragments and the *EcoPI5I* site sites in Loop-1 linkers used for library construction (Fig. 1). To meet our internal RNAi therapeutic screening purpose, steps 2 and 3 were performed with each individual loop-1 primers and analyzed on 1% agarose gel (Fig. 4C); samples were then pooled for subsequent manipulations. The clones we analyzed showed the expected size distribution, similar to the natural length (19–23 bp) of sRNAs generated by Dicer processing in living cells [38,39]. Moreover, clones of the different size classes seemed to be randomly distributed across the cDNA sequences used to prepare the SURER libraries (Fig. 5). The five different sizes cover almost all sites of a target gene which is determined by the initial *DNaseI* partial digestion (Fig. 1) as previously described [24,28,40,41]. However, we did find some sizes were outside of this range <5% clones gave sizes >23 bp and <19 bp, respectively. We suspect *EcoPI5I*, a type-3 restriction/modification enzyme, does not always cleave precisely at 25–27 bp from its recognition site. This phenomenon has also been observed when using *MmeI* (18–20 bp) for library preparation enzymatically [26].

FokI restriction sites were incorporated into both the loop-1 and loop-2 linkers to generate TTTT cohesive ends, which permits cloning into *BsmBI*-digested pRNAi-U6H1/Neo, these sequences also act as pol III transcriptional terminators. The loop-1 linkers contain additional bases (Ns in Fig. 2A) which maintained appropriate spacing between the *FokI* sites and the A/T regions to ensure an appropriate AAAA overhang is generated. The cloning efficiency for SURER libraries reached $\sim 3 \times 10^5$ colony-forming units (~ 2.5 μ g final cDNA was used

for vector ligation), with a $>90\%$ recombination rate (data not shown). Given a gene is 3 kb in length, a gene-specific library with a $\sim 3 \times 10^5$ independent clones is large enough to cover each sRNA site for a gene targeted. In this study, the three SURER libraries were successfully constructed from the templates of full-length cDNA (HBV DNA polymerase and BIRC5) or a domain of MITF gene. The length of the starting templates ranged from 429 (BIRC5) to 2499 bp (HBV DNA polymerase). As shown in Fig. 5, the distribution of sRNAs in these libraries is quite representative, particularly for the HBV polymerase, where larger numbers of clones were analyzed, where more than 80% of sRNA clones were non-overlapping recombinants (data not shown).

The SURER protocol involves two PCR steps. The first (Fig. 1, Step 3) is a single primer PCR, which is used to specifically amplify substrates containing opposing *BsmBI* sites present in the loop-1 primers (Fig. 1, Step 2). The second (Fig. 1, Step 6) is a conventional PCR used to specifically amplify molecules containing both loop-1 and loop-2 primers, which are required for cloning into the expression cassette. The three SURER libraries described here show excellent coverage of the target genes, indicating these steps had not significantly biased representation. The protocol does not involve amplification of palindromes another potential source of bias [28].

Constructs specifying potent knockdown for three targets (HBV, MITF, and BIRC5) were defined by screening SURER libraries. For HBV, 19.8% (100/507) of HBV SURER library clones gave $\geq 50\%$ gene knockdown, 3.2% of clones showed $\geq 70\%$ gene knockdown (Fig. 8), none of these have been previously described. We have developed siRNAs and ddRNAi constructs based on these sequences, which also

FIG. 6. Gene silencing efficiency of SURER clones determined using real-time quantitative PCR (RT-qPCR), normalized to GAPDH. **(A)** Relative levels of HBV polymerase mRNA in HepG2 2.2.15 cells treated with different sRNA expression cassettes selected from SURER library. **(B)** Relative mRNA levels of MITF in A431 cells treated with different sRNA expression cassettes. **(C)** Relative mRNA levels of BIRC5 in HepG2 cells treated with different sRNA expression cassettes.

FIG. 7. Knockdown of BIRC5 with small interfering RNAs (siRNAs) defined by SURER screen. **(A)** Relative levels of BIRC5 mRNA in HepG2 cells treated with siRNAs: c-2, c-38, and c-43 were determined using RT-qPCR and normalized to levels of GAPDH mRNA, the three designed and synthetic siRNAs (s-1, s-2, s-3) used for comparison. **(B)** Relative protein levels of BIRC5 in cells treated with the indicated siRNAs were determined by western blots and normalized to levels of β -actin. **(C)** Cell proliferation in BIRC5 siRNA-treated cells was analyzed using CCK-8 assays. Growth curve of cells is shown for each treatment at 0, 24, 48, 72, and 96 hours. Values were given as mean $\pm$ SD of three separate experiments with triplicate wells per condition. * $P < 0.05$.

showed strong silencing activity (data not shown). For MITF, 50.8% (32/63) of clones gave $\geq 50\%$ gene knockdown, and 22.2% of clones showed $\geq 70\%$ gene knockdown (Fig. 8). For BIRC5, three clones (2, 38, 42) showed 80%–90% knockdown of mRNA (Fig. 6C). siRNAs were synthesized based on the sequences of these and potent knockdown confirmed at the level of mRNA, protein, and biological activity (Fig. 7). Only a small region (429 bp) of the BIRC5 gene screened, which was too short to be designed and found siRNAs with high knockdown efficiency using existing computer algorithms [37,42–43]. Screening SURER libraries overcomes the shortcomings of computer design algorithms.

In summary, we provide an innovative library construction method for RNAi screens. The SURER library construction protocol can be readily expanded to other kind of small RNA libraries such as miRNA and saRNA libraries where RNAs in the range of 19 to 23 bp show biological activity. Sequences derived from all or parts of open reading frames (ORFs), 5' or 3'UTRs or promoter regions of target sequences can be readily screened. Incorporation of the pRNAi-U6H1/Neo expression cassette into new vectors might further increase the potential of the SURER strategy. The SURER library opens new ways for RNAi therapeutic screens in a fast and simple manner.

FIG. 8. Knockdown efficiency in three different SURER libraries. **(A)** 19.8% of siRNA clones silenced HBV polymerase by $\geq 50\%$, 3.2% of which resulted in $\geq 70\%$ knockdown efficiency. **(B)** 50.8% of siRNAs were effective in silencing MITF by $\geq 50\%$, and 22.2% of these resulted in $\geq 70\%$ knockdown efficiency from the MITF SURER library. **(C)** 50% of siRNAs were effective in silencing BIRC5 by $\geq 50\%$, and 23.1% of these resulted in $\geq 70\%$ knockdown efficiency from the BIRC5 SURER library.

Acknowledgments

This work was supported National Science and Technology Major Project from National Natural Science Foundation of China (2011ZX09401-012); National High Technology Research and Development Program of China (863 Program, 2012AA022501). The HBV work was financially supported by Benitec Biopharma Ltd.

Author Disclosure Statement

No competing financial interests exist.

References

1. Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE and Mello CC. (1998). Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 391:806–811.
2. Hannon GJ. (2002). RNA interference. *Nature* 418:244–251.
3. Martinez J, Patkaniowska A, Urlaub H, Lührmann R and Tuschl T. (2002). Single-stranded antisense siRNAs guide target RNA cleavage in RNAi. *Cell* 110:563–574.
4. Li T, Wu M, Zhu YY, Chen J and Chen L. (2014). Development of RNA interference-based therapeutics and application of multi-target small interfering RNAs. *Nucleic Acid Ther* 24:302–312.
5. Buchholz F, Kittler R, Slabicki M and Theis M. (2006). Enzymatically prepared RNAi libraries. *Nat Methods* 3:696–700.
6. Strezoska Ž, Licon A, Haimes J, Spayd KJ, Patel KM, Sullivan K, Jastrzebski K, Simpson KJ, Leake D, van Brabant Smith A and Vermeulen A. (2012). Optimized PCR conditions and increased shRNA fold representation improve reproducibility of pooled shRNA screens. *PLoS One* 7:e42341.
7. Tan H, Fan J, Bao J, Dy JG and Zhou X. (2012). A computational model for compressed sensing RNAi cellular screening. *BMC Bioinformatics* 13:337.
8. Kawasaki H, Suyama E, Iyo M and Taira K. (2003). siRNAs generated by recombinant human Dicer induce specific and significant but target site-independent gene silencing in human cells. *Nucleic Acids Res* 31:981–987.
9. Watanabe T, Totoki Y, Toyoda A, Kaneda M, Kuramochi-Miyagawa S, Obata Y, Chiba H, Kohara Y, Kono T, et al. (2008). Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. *Nature* 453:539–543.
10. Yang D, Goga A and Bishop JM. (2004). RNA interference (RNAi) with RNase III prepared siRNAs. *Methods Mol Biol* 252:471–482.
11. Allen E, Xie Z, Gustafson AM and Carrington JC. (2005). microRNA-directed phasing during trans-acting siRNA biogenesis in plants. *Cell* 121:207–221.
12. Corrêa RL, Steiner FA, Berezikov E and Ketting RF. (2010). MicroRNA-directed siRNA biogenesis in *Caenorhabditis elegans*. *PLoS Genet* 6:e1000903.
13. Dalakouras A, Moser M, Zwiebel M, Krczal G, Hell R and Wassenegger M. (2009). A hairpin RNA construct residing in an intron efficiently triggered RNA-directed DNA methylation in tobacco. *Plant J* 60:840–851.
14. Qi X, Bao FS and Xie Z. (2009). Small RNA deep sequencing reveals role for *Arabidopsis thaliana* RNA-dependent RNA polymerases in viral siRNA biogenesis. *PLoS One* 4:e4971.
15. Okamura K and Lai EC. (2008). Endogenous small interfering RNAs in animals. *Nat Rev Mol Cell Biol* 9:673–678.
16. Harfe BD. (2005). MicroRNAs in vertebrate development. *Curr Opin Genet Dev* 15:410–415.
17. Soifer HS, Rossi JJ and Sætrom P. (2007). MicroRNAs in disease and potential therapeutic applications. *Mol Ther* 15:2070–2079.
18. Bartel DP. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 116:281–297.
19. Tomari Y and Zamore PD. (2005). Perspective: machines for RNAi. *Genes Dev* 19:517–529.
20. Huang V, Qin Y, Wang J, Wang X, Place RF, Lin G, Lue TF and Li LC. (2010). RNAi is conserved in mammalian cells. *PLoS One* 5:e8848.
21. Janowski BA, Younger ST, Hardy DB, Ram R, Huffman KE and Corey DR. (2007). Activating gene expression in mammalian cells with promoter-targeted duplex RNAs. *Nat Chem Biol* 3:166–173.
22. Li LC, Okino ST, Zhao H, Pookot D, Place RF, Urakami S, Enokida H and Dahiya R. (2006). Small dsRNAs induce transcriptional activation in human cells. *Proc Natl Acad Sci (USA)* 103:17337–17342.
23. Portnoy V, Huang V, Place RF and Li LC. (2011). Small RNA and transcriptional upregulation. *Wiley Interdiscip Rev RNA* 2:748–760.
24. Shirane D, Sugao K, Namiki S, Tanabe M, Iino M and Hirose K. (2004). Enzymatic production of RNAi libraries from cDNAs. *Nat Genet* 36:190–196.
25. Shirane D, Sugao K, Namiki S, Tanabe M, Iino M and Hirose K. (2004). Restriction enzyme-generated siRNA (REGS) vectors and libraries. *Nat Genet* 36:183–189.
26. Luo B, Heard AD and Lodish HF. (2005). Small interfering RNA production by enzymatic engineering of DNA (SPEED). *Proc Natl Acad Sci (USA)* 101:5494–5499.
27. Pongratz C, Yazdanpanah B, Kashkar H, Lehmann MJ, Kräusslich HG and Krönke M. (2010). Selection of potent non-toxic inhibitory sequences from a randomized HIV-1 specific lentiviral short hairpin RNA library. *PLoS One* 5:e13172.
28. Seyhan AA, Vlassov AV, Ilves H, Egry L, Kaspar RL, Kazakov SA and Johnston BH. (2005). Complete, gene-specific siRNA libraries: production and expression in mammalian cells. *RNA* 11:837–846.
29. Shtutman M, Maliyekkel A, Shao Y, Carmack CS, Baig M, Warholc N, Cole K, Broude EV, Harkins TT, Ding Y, Roninson IB. (2010). Function-based gene identification using enzymatically generated normalized shRNA library and massive parallel sequencing. *Proc Natl Acad Sci (USA)* 107:7377–7382.
30. Xu J, Zeng JQ, Wan G, Hu GB, Yan H and Ma LX. (2009). Construction of siRNA/miRNA expression vectors based on a one-step PCR process. *BMC Biotechnol* 9:53.
31. Du C, Ge B, Liu Z, Fu K, Chan WC and McKeithan TW. (2006). PCR-based generation of shRNA libraries from cDNAs. *BMC Biotechnol* 6:28.
32. Tuscholski J, Skowron PM and Podhajski AJ. (1995). MmeI, a class-IIS restriction endonuclease: purification and characterization. *Gene* 157:87–92.
33. Campbell VW and Jackson DA. (1980). The effect of divalent cations on the mode of action of DNase I. The initial reaction products produced from covalently closed circular DNA. *J Biol Chem* 255:3726–3735.
34. Möncke-Buchner E, Mackeldanz P, Krüger DH and Reuter M. (2004). Overexpression and affinity chromatography

- purification of the Type III restriction endonuclease Eco-P15I for use in transcriptome analysis. *J Biotechnol* 114:99–106.
35. Tabor S, Struhl K, Scharf SJ and Gelfand DH. (2001). DNA-dependent DNA polymerases. *Curr Protoc Mol Biol* Chapter 3:Unit3.5.
 36. Livak KJ and Schmittgen TD. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-delta delta C(T)) method. *Methods* 25:402–408.
 37. Filhol O, Ciais D, Lajaunie C, Charbonnier P, Foveau N, Vert JP and Vandenbrouck Y. (2012). DSIR: assessing the design of highly potent siRNA by testing a set of cancer-relevant target genes. *PLoS One* 7:e48057.
 38. MacRae IJ, Ma E, Zhou M, Robinson CV and Doudna JA. (2008). In vitro reconstitution of the human RISC-loading complex. *Proc Natl Acad Sci (USA)* 105:512–517.
 39. Noland CL, Ma E and Doudna JA. (2011). siRNA repositioning for guide strand selection by human dicer complexes. *Mol Cell* 43:110–121.
 40. Singer O, Yanai A and Verma IM. (2004). Silence of the genes. *Proc Natl Acad Sci (USA)* 101:5313–5314.
 41. Tomimoto K, Yamakawa M and Tanaka H. (2012). Construction of a long hairpin RNA expression library using Cre recombinase. *J Biotechnol* 160:129–139.
 42. Gong W, Ren Y, Zhou H, Wang Y, Kang S and Li T. (2008). siDRM: an effective and generally applicable online siRNA design tool. *Bioinformatics* 24:2405–2406.
 43. Naito Y, Yoshimura J, Morishita S and Ui-Tei K. (2009). siDirect 2.0: updated software for designing functional siRNA with reduced seed-dependent off-target effect. *BMC Bioinformatics* 10:392.

Address correspondence to:

York Yuanyuan Zhu, PhD

Small RNA Technology and Application Institute

Nantong University

76 Changxing Road

E & T Development Area

Nantong, 226016

China

E-mail: yzhu@biomics.cn

Received for publication October 14, 2014; accepted after revision November 6, 2014.