Quantitative Biology 2020, 8(1): 31–42
https://doi.org/10.1007/s40484-020-0194-5

RESEARCH ARTICLE

Construction of a CRISPR-based paired-sgRNA library for chromosomal deletion of long non-coding RNAs

Minzhen Tao, Qiaochu Mu[#], Yurui Zhang[##,*], Zhen Xie[*]

MOE Key Laboratory of Bioinformatics and Bioinformatics Division, Center for Synthetic and System Biology, Department of Automation, Beijing National Research Center for Information Science and Technology, Tsinghua University, Beijing 100084, China
* Correspondence: yr.zhang@syngen.tech, zhenxie@tsinghua.edu.cn

Received May 28, 2019; Revised September 27, 2019; Accepted November 19, 2019

Background: **Derived from an adaptive bacterial immune system, the clustered regularly interspaced palindromic repeats (CRISPR)/CRISPR-associated 9 (Cas9) system has shown great potential in high-throughput functional genomic screening, especially for protein-coding genes. However, it is still challenging to apply the similar strategy to study non-coding genomic elements such as long non-coding RNAs (lncRNAs) or clusters of microRNAs, because short insertions or deletions may not be sufficient to generate loss-of-function phenotypes.**
Methods: **Here, we presented a systematic strategy for designing a CRISPR-based paired-sgRNA library for high-throughput screening in non-coding regions. Due to the abundance of lncRNAs and their diverse regulatory roles *in vivo*, we repurposed microarray datasets to select 600 highly expressed lncRNAs in non-small-cell lung cancer and designed two schemes for lncRNA deletion with ~20 paired-sgRNAs for each lncRNA. Through Golden-Gate assembly, we generated a pooled CRISPR-based library with a total of 12,878 sgRNA pairs.**
Results: **Over 80% of paired-sgRNAs were recovered from final pooled library with a relatively even distribution. Cleavage efficiency of sgRNA pairs was validated through experiments of transient transfection and viral infection. Moreover, randomly selected paired-sgRNAs showed that efficient deletion of genomic DNA could be achieved with a deletion size within the range of 500 to 3000 bp.**
Conclusions: **In summary, we have demonstrated a strategy to design and construct a pooled paired-sgRNA library to generate genomic deletion in the lncRNA regions, validated their deletion efficiency and explored the relationship of deletion efficiency with respect to deletion size. This method would be also suitable for investigation of other uncharacterized non-coding genomic regions in mammalian cells in an efficient and cost-effective manner.**

Keywords: CRISPR/Cas9; paired-sgRNA; chromosomal deletion; non-coding genomic elements; lncRNA

Author summary: The CRISPR/Cas9 system has shown great potential in functional genomic screening, especially for protein-coding genes. However, it is challenging to apply the similar strategy to study non-coding genomic elements, because short insertions or deletions may not be sufficient to generate loss-of-function phenotypes. In this paper, we proposed a strategy to design and construct a CRISPR-based paired-sgRNA library for chromosomal deletions of lncRNA loci in mammalian cells and confirm the cleavage efficiency through experiments. This approach demonstrates a simple and scalable tool for genome-wide functional study of non-coding elements in mammalian cells.

[#]Present address: Zhejiang Tmall Techonogy Co., Ltd, Yuhang District, Hangzhou 311121, China

[##]Present address: Beijing Syngentech Co., Ltd, Zhongguancun Life Science Park, Changping District, Beijing 102206, China

INTRODUCTION

Derived from an adaptive bacterial immune system, the clustered regularly interspaced palindromic repeats (CRISPR)/CRISPR-associated 9 (Cas9) system is a versatile and powerful tool for mammalian genome modification. The system requires two components when applied as an experimental tool: a single guide RNA (sgRNA) that recognizes a specific site in the genome through Watson-Crick base pairing, and a Cas9 nuclease that binds the sgRNA and generates a DNA double-strand break. The CRISPR/Cas9 system is easy to be manipulated by simply altering the short sequences of sgRNAs, and therefore it is utilized for wide applications such as genome editing [1–3], transcriptional regulation [4–8] and genomic function interrogation [9–18].

The CRISPR/Cas9 system can be widely applied in cell-based functional screening assays for hundreds of coding genes in parallel [9–14], in which pooled libraries of sgRNAs target the coding regions of genes. The sgRNA guided DNA cleavage often causes frameshift of coding genes by generating short insertions or deletions (indels), allowing the investigation of gene function on phenotypes such as cell growth. Although NHEJ mediated indels are reported to cause significant impact on the region of 3′ UTR of miRNA binding sites [19], indels may not be sufficient to produce loss-of-function phenotypes for large non-coding genomic elements, such as gene/miRNA clusters or lncRNAs, for which specific chromosomal deletions are more suitable. Although paired-sgRNAs have been successfully used to generate genomic deletions for lncRNAs [20–24], enhancers [25], microRNAs [26] and gene clusters [27], a systematic strategy for designing a CRIPSR-based paired-sgRNA library for high-throughput screening in non-coding regions has been scarcely described in details.

In this study, we presented designs for genomic deletions with paired-sgRNAs, provided methods to generate two types of pooled paired-sgRNA plasmids and validated the efficiency of pooled library. Due to the abundance of lncRNAs *in vivo* [28] and their diverse regulatory roles in chromosome silencing [29], chromatin modification remodeling [30], transcriptional regulation [22,23] and nuclear transport [31], we targeted highly expressed lncRNAs in non-small-cell lung cancer (NSCLC) for chromosomal deletion to demonstrate the utility of this approach.

RESULTS

Repurposing microarray data for lncRNA selection in NSCLC

Inspired by previous study in which cancer-type specific dysregulations of lncRNA have been found through integrative analyses [32], we searched for overexpressed lncRNAs in NSCLC. Expression profiles of lncRNA can be extracted from microarray data through the reannotation of probes, despite the fact that lncRNAs are not the intentional targets [33–37]. Therefore, we retrieved 27 microarray datasets performed on the H1650 cell line, a commonly used cell line in NSCLC research, from the Gene Expression Omnibus database (Table 1).

To repurpose these available array-based data, we designed a method to reannotate the probes from Affymetrix Human Genome Arrays, excluding low-quality and ambiguous probes and keeping lncRNAs transcripts with at least 3 matched probes (Fig. 1A). Then, we developed a pipeline to analyze and derive expressions of lncRNAs, including preprocessing raw microarray data with two different methods MAS5.0 and GCRMA, summarizing expression values of uniquely mapped lncRNAs according to the reannotation file. After listing lncRNAs in descending order according to their expression values, top 600 overlapped highly expressed lncRNAs were selected (Fig. 1B, Supplementary Table S1). In addition to the selected lncRNAs, we chose 18 well characterized lncRNAs associated with the occurrence and development of NSCLC as positive controls [38–57] (Supplementary Table S1).

Design of paired-sgRNAs for lncRNA deletion

We designed two schemes for lncRNA deletion to disrupt the function of selected lncRNAs: one targeting the promoter plus the lncRNA coding region (pg-type) and the other only targeting the coding region of the lncRNA (gg-type) (Fig. 1C). We considered the 300-bp upstream of each transcription start site (TSS) as the assumed promoter, because it covered the region of core promoter and proximal promoter, which contained a large number of transcription factor binding sites for basal transcription and transcription activation or repression activities respectively [58–61]. After the definition of targeted regions, we sorted for all possible sgRNAs with the proper protospacer adjacent motif (PAM) NGG through the online algorithm of CRISPR-ERA [62]. We set rules to filter sgRNAs only if (i) their sequences mapped to the intended loci with up to two mismatches, (ii) the sum of efficacy score (E-score) and specificity score (S-score) of each sgRNA was greater than 0 to ensure cutting efficiency, and (iii) the sgRNAs did not include the UUUU/TTTT polymer.

We ordered all potential sgRNAs according to their scores (the sum of its E-score and S-score), enumerated all possible paired-sgRNAs and multiplied their scores for the following evaluations. From the paired-sgRNAs with high multiplying scores, we separately chose approxi-

Table 1 Information of 27 microarray datasets of NSCLC from GEO

Number	Sample number	Project name
1	GSM99025	Gefitinib effect on various non-small cell lung cancer cell lines (HG-U133B)
2	GSM98979	Gefitinib effect on various non-small cell lung cancer cell lines (HG-U133A)
3	GSM1374703	Expression data from AstraZeneca internal cell lines
4	GSM1146860	Gene expression data from 56 lung cancer cell lines
5	GSM952570	Downstream targets of ID1 transcription factor in NSCLC
6	GSM952571	Downstream targets of ID1 transcription factor in NSCLC
7	GSM952572	Downstream targets of ID1 transcription factor in NSCLC
8	GSM952573	Downstream targets of ID1 transcription factor in NSCLC
9	GSM952574	Downstream targets of ID1 transcription factor in NSCLC
10	GSM902401	A novel five-gene signature predicts overall and recurrence-free survival in NSCLC
11	GSM902402	A novel five-gene signature predicts overall and recurrence-free survival in NSCLC
12	GSM785587	Gene expression patterns that predict sensitivity to epidermal growth factor receptor tyrosine kinase inhibitors in lung cancer cell lines and human lung tumors
13	GSM785588	Gene expression patterns that predict sensitivity to epidermal growth factor receptor tyrosine kinase inhibitors in lung cancer cell lines and human lung tumors
14	GSM785589	Gene expression patterns that predict sensitivity to epidermal growth factor receptor tyrosine kinase inhibitors in lung cancer cell lines and human lung tumors
15	GSM785590	Gene expression patterns that predict sensitivity to epidermal growth factor receptor tyrosine kinase inhibitors in lung cancer cell lines and human lung tumors
16	GSM785591	Gene expression patterns that predict sensitivity to epidermal growth factor receptor tyrosine kinase inhibitors in lung cancer cell lines and human lung tumors
17	GSM785592	Gene expression patterns that predict sensitivity to epidermal growth factor receptor tyrosine kinase inhibitors in lung cancer cell lines and human lung tumors
18	GSM785593	Gene expression patterns that predict sensitivity to epidermal growth factor receptor tyrosine kinase inhibitors in lung cancer cell lines and human lung tumors
19	GSM372755	Expression data from non small cell lung cancer cell lines
20	GSM254933	Paired MEK inhibited and control
21	GSM254934	Paired MEK inhibited and control
22	GSM274828	mRNA cancer cell line profiles
23	GSM274795	mRNA cancer cell line profiles
24	GSM254969	Anti-tumor activity of histone deacetylase inhibitors in non-small cell lung cancer cells
25	GSM206474	Death receptor O-glycosylation controls tumor-cell sensitivity to the proapoptotic ligand
26	GSM108819	Analysis of lung cancer cell lines
27	GSM108820	Analysis of lung cancer cell lines

mately 10 pairs for each lncRNA according to the two deletion schemes to meet the following requirements: (i) the cleavage offset of the pair was between 500 bp and 10 kb, (ii) the sgRNAs in the pair were not reused and (iii) the deletion region did not overlap with other genes. The offset distance above was chosen to optimize the deletion efficiency, because the deletion efficiency dropped rapidly when the deletion size was larger than 10 kb [63]. For the lncRNAs with very few paired-sgRNAs, we obtained enough pairs by reusing the sgRNAs with high targeting scores. Finally, we obtained 5253 and 5420 paired-sgRNAs for pg-type and gg-type lncRNA deletion, respectively (Supplementary Table S1).

For 18 well studied lncRNAs, we designed 10 paired-sgRNAs for each lncRNA following the scheme that only targeted the lncRNA coding region to serve as positive controls for further functional screening study. Additionally, 45 sgRNAs that did not target any loci in the human genome were designed. Performing a random combination of two sgRNAs above, 2025 paired-sgRNAs were generated as negative controls. In summary, the whole library contained a total of 12,878 sgRNA pairs.

Construction of a paired-sgRNA library

To facilitate further pooled functional screening, we cloned two distinct sgRNAs into a single lentiviral vector. In the construction of sgRNA pairs for selected and

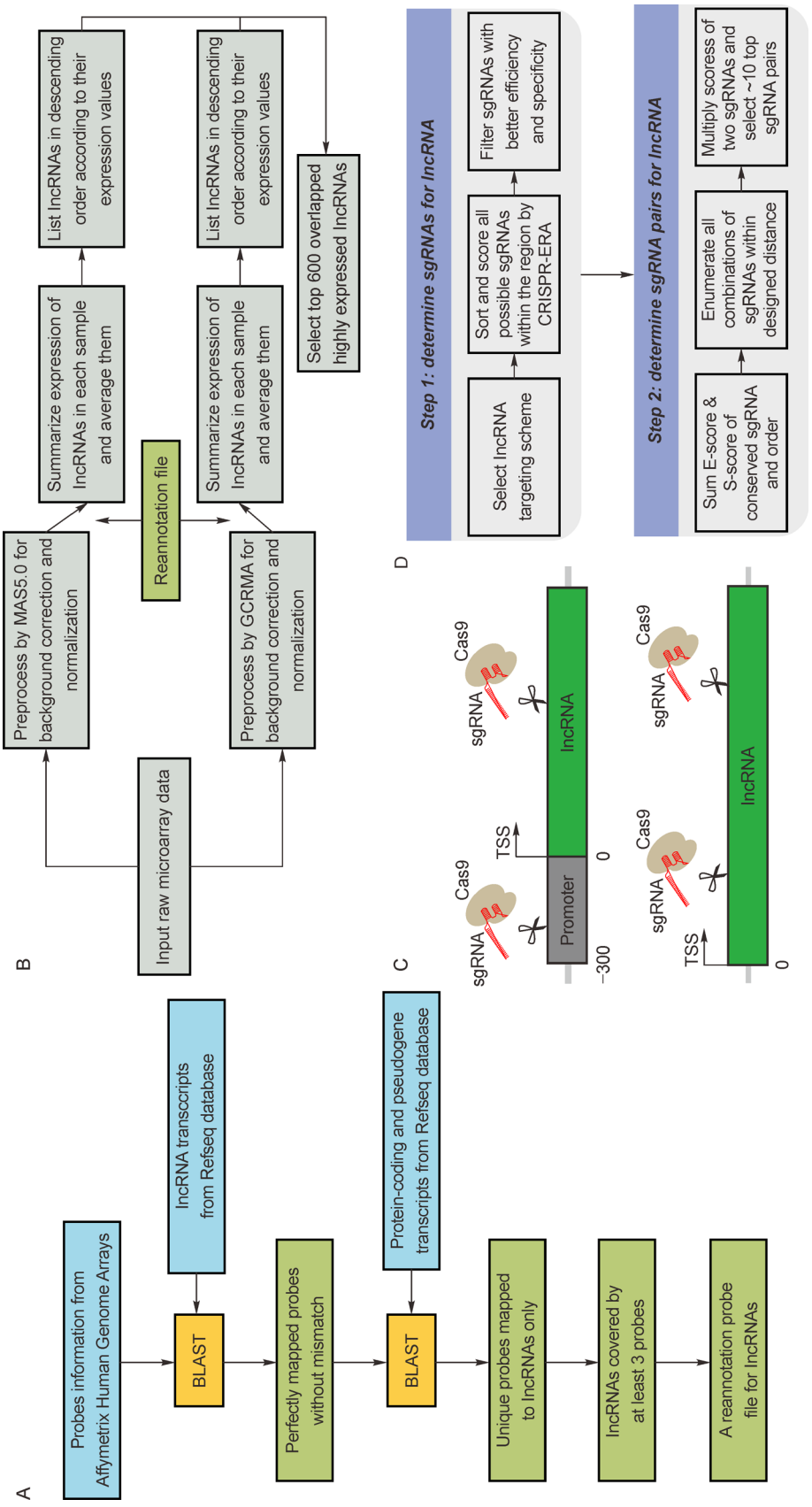

Figure 1. Workflow for repurposing microarray data and design of paired-sgRNAs for lncRNA deletion. (A) The method of reannotating Affymetrix microarray probes for lncRNAs. (B) Workflow for analyzing expression value and selection of lncRNAs. (C) Two schemes for lncRNA deletion. (D) Workflow for determining paired-sgRNAs for lncRNAs.

positive lncRNAs, the library preparation protocol comprised three cloning steps (Fig. 2A). First, the synthesized oligonucleotides, which included paired-sgRNA sequences separated by two *Esp*3I enzyme restriction sites, were inserted into donor vectors carrying the human U6 promoter. Then, the fragments containing a scaffold for the former sgRNA and promoter H1 were cloned into intermediate vectors by the Golden Gate cloning method, followed by excision of the chloramphenicol resistance marker by Cre/LoxP recombination. The pooled paired-gRNAs for these lncRNAs were divided into nine subpools, containing 1356, 1356, 1352, 1356, 1319, 1311, 1314, 1309 and 180 pairs, respectively.

A two-step cloning method was used to construct negative control paired-sgRNAs (Fig. 2B). The synthesized 24-bp targeting oligonucleotides were annealed and inserted into the donor vector by conventional ligation at

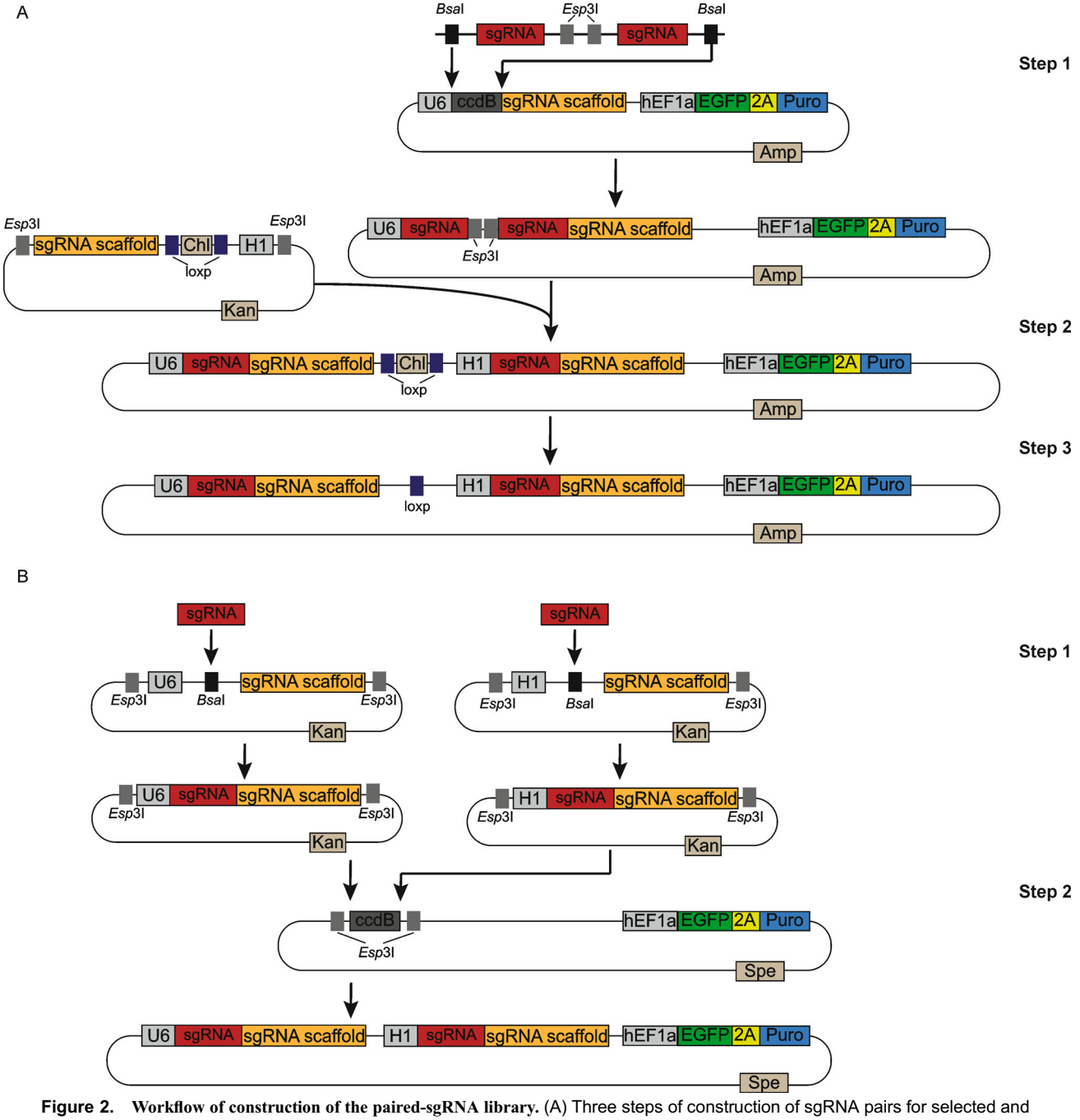

Figure 2. Workflow of construction of the paired-sgRNA library. (A) Three steps of construction of sgRNA pairs for selected and positive lncRNAs. (B) Two steps of construction of sgRNA pairs for negative controls.

the *Bsa*I site, and then tandem ligations of two sgRNAs were cloned to generate the final plasmid using the Golden Gate cloning method, serving as the tenth subpool.

Quality analyses of pooled paired-sgRNA library

We digested the pool plasmids from each subpool with *Xho*I enzyme. Three bands with the correct length were obtained in 10 subpools (Fig. 3A), which confirmed that the plasmids were correctly cloned. Further, we selected approximately 20 bacterial colonies from each subpool for Sanger sequencing, and approximately 80% of colonies mapped to the designed library (Table 2). Therefore, we mixed 10 subpools together equally according to their corresponding pairs to obtain the final pooled library. The 576-bp fragments harboring paired-sgRNA sequences were amplified by PCR from the final pooled library, and then sequenced using HiSeq 2500 with the paired-end 250-bp (PE250) mode. We found approximately 80% of paired-sgRNAs mapped to the designed library (Fig. 3B), with a relatively even distribution with 73.01% coverage of pairs within a 64-fold range (Fig. 3C).

Validation of the library cleavage performance

To validate the chromosomal cleavage performance of the pooled library, we randomly selected 12 paired-sgRNAs from two targeting schemes for lncRNA deletion. The deletion size of selected pairs was from 500 to 3000 bp, which constituted the majority in our library. Using a paired-sgRNA which did not target any loci in the genome as the negative control, we transfected HEK293-FT cells with plasmids expressing functional positive paired-sgRNAs and Cas9 protein. After 5 days of puromycin selection, cells were harvested to extract genomic DNA for deletion detection. PCR results showed various deletion rates, with 1.3%–87.8% for pg-type targeting scheme (Fig. 4A) and 10.8%–78.6% for gg-type (Fig. 4B). We conducted the experiments three times and the summarized results suggested that there was no

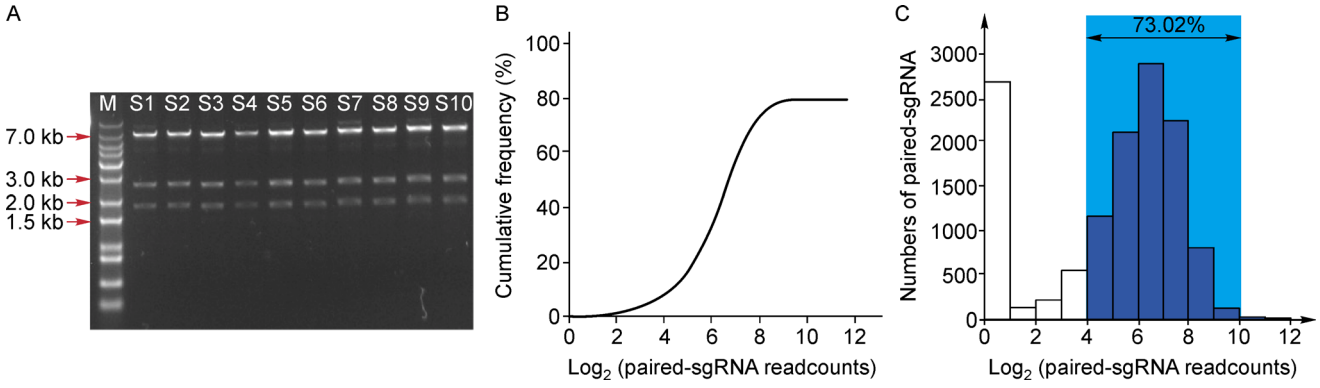

Figure 3. Quality analyses of pooled paired-sgRNA library. (A) Gel results of *Xho*I digestion for 10 subpools, in which three fragments (1721 bp/ 2374 bp/ 7342 bp) were obtained. (B) Cumulative distribution of log transformed paired-sgRNA counts in final pooled library. (C) Histogram of log transformed paired-sgRNA counts in final pooled library, and the shaded areas represented the maximal coverage of sgRNA pairs within a 64-fold range.

Table 2 Sanger sequencing results of plasmids from each subpool

Subpool	Number of bacterial colonies	Number of correct colonies	Correct rate
1	27	21	0.78
2	23	19	0.83
3	22	18	0.82
4	28	20	0.71
5	24	20	0.83
6	19	14	0.74
7	22	16	0.73
8	20	16	0.80
9	22	16	0.73
10	20	19	0.95
Total	227	179	0.79

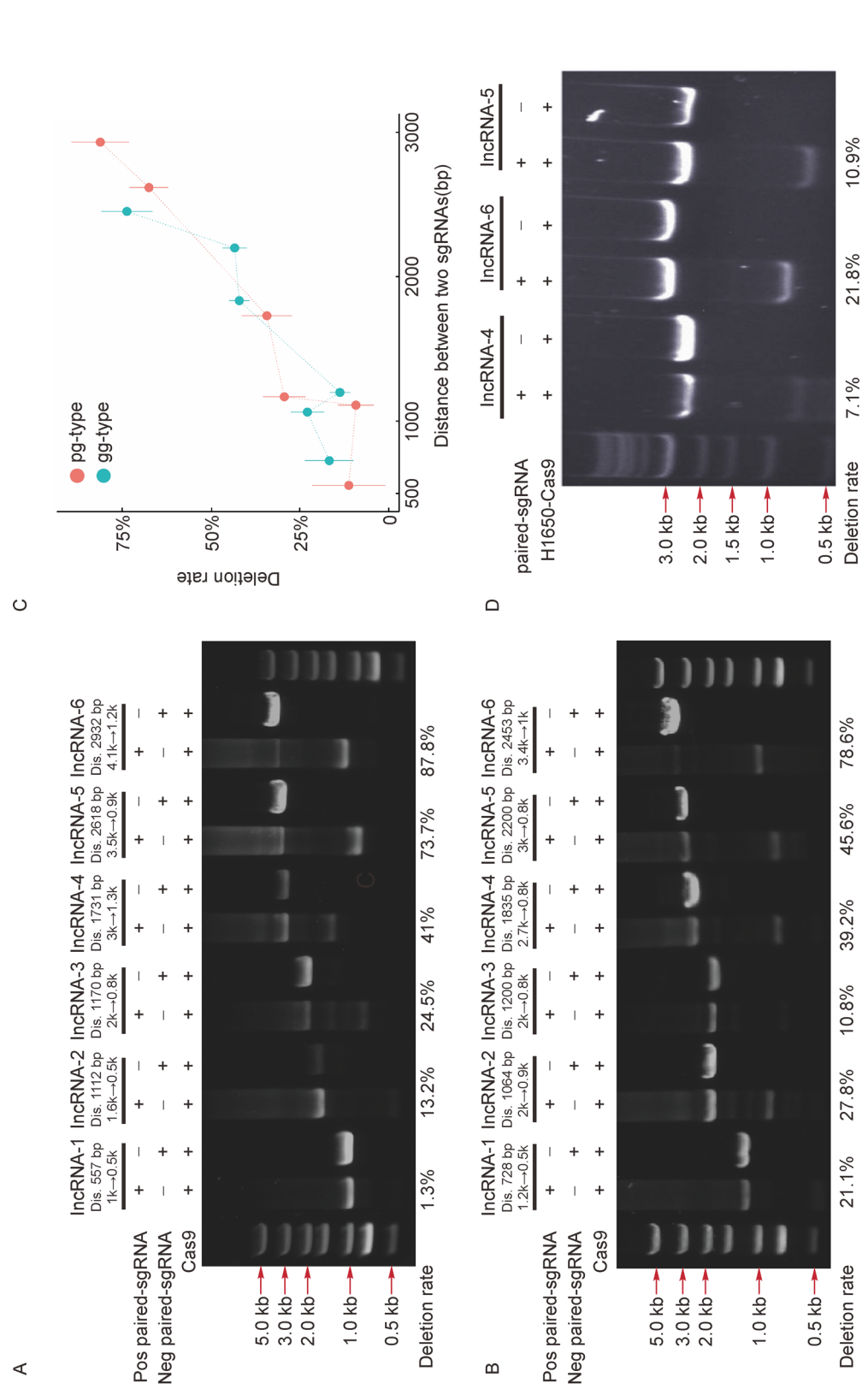

Figure 4. Validation of the library cleavage performance. (A–B) Gel results of amplicons from PCR for lncRNA deletion targeted by paired-sgRNAs of pg-type (A) and gg-type (B). Dis. represented the deletion size of each sgRNA pair, length before and after the arrow represented the size of non-deletion band and deletion band, and pos paired-sgRNA and neg paired-sgRNA represented the pair that could conduct lncRNA deletion and the pair that did not target any loci. (C) Summary of three replicates of transfection experiment for lncRNA deletion, represented by mean±s.d.. (D) Gel results of amplicons from PCR in the lentiviral infections for lncRNA deletion.

significant difference between two targeting schemes (Fig. 4C). Despite GC content, presence of polythymidine and specificity considered in CRISPR-ERA to provide sgRNAs with well performance, the cutting efficiency of paired-sgRNAs might be influenced by the deletion size and accessibility of sgRNAs to their target sites. Within the range of 500–3000 bp, the higher efficiency was observed when the deletion size was larger. Additionally, we noticed that there were weak bands in some lncRNA deletions located between the deletion band and non-deletion band (Fig. 4A, B), which we speculated could be large indels ranging from ~–200 to ~+500 bp [63].

With more stable cleavage performance than pairs of pg-type, we used 3 well performed sgRNA pairs of gg-type targeting lncRNA-4, lncRNA-5 and lncRNA-6 to produce lentiviral particles to infect H1650-Cas9 cells (H1650 cells that were engineered to stably express Cas9). Similarly after puromycin selection and genomic DNA extraction, we conducted PCR for deletion detection. These three pairs showed deletion rate between 7% –21% (Fig. 4D), in which the pair for lncRNA-6 was the highest and the other two were close as they did in transient transfections.

DISCUSSION

Since known protein-coding sequences compose a minority of the whole human genome [64], functional identification of non-coding genomic regions such as lncRNAs is needed. We here described a method to design and establish a CRISPR-based paired-sgRNA library for genomic deletion. We detailed the selection of overexpressed lncRNAs in NSCLC, criteria of paired-sgRNA design and methods to construct the final pooled library, whose cleavage efficiency was confirmed by several selected plasmids.

There are several limitations to our strategy that could be further optimized. First, although we used two different promoters to express sgRNAs to prevent recombination, not all designed pairs were recovered in our deep sequencing results (Fig. 3B, C). This could be improved with careful design of sequence constructs and primers for lower recombination rates and higher sequencing quality [65]. As shown in Fig. 4 that the cleavage efficiency of paired-sgRNAs varied along with the deletion size, the library could be optimized to achieve better cleavage performance for functional screening. Finally, due to various locations of lncRNA in the genome [66], it is necessary to exclude genomic regions that overlap with other functional elements to avoid misleading results.

In conclusion, we have demonstrated a strategy to design and construct pooled paired-sgRNAs to generate genomic deletion in the lncRNA region, explored the relationship of deletion efficiency with respect to deletion size and observed the better performance with the larger deletion distance within the range of 500 to 3000 bp.This method would be also suitable for investigation of other uncharacterized long non-coding genomic regions in mammalian cells in an efficient and cost-effective manner.

MATERIALS AND METHODS

Construction of paired-sgRNA subpools for lncRNA deletion

Chip synthesis of oligonucleotide sequences

The oligonucleotides containing each paired-sgRNA sequences, separated by a short spacer harboring two *Esp*3I sites, were designed and synthesized on a chip. At the 5′ and 3′ ends of each oligonucleotide were short sequences including a *Bsa*I site and a fragment used to distinguish different subpools. The schematic construct was fragment1-*Bsa*I-sgRNA1-*Esp*3I-*Esp*3I-sgRNA2-*Bsa*I-fragment2.

Step one of construction

For the construction of each subpool, primers targeting the *Bsa*I site and flanking fragments of oligonucleotides were used for amplification to generate 127-bp dsDNA molecules. Primers for nine subpools were as follows:

Subpool-1: TGGTGATAGGTAAGGATGGC; CGGCTCAGTATTGCGATTAC

Subpool-2: TCGACACCACTATACACCAC; GGCCCGTGAGAGTATAAAGA

Subpool-3: CATGTAGTGCAGCCATTCTC; GGGCACAGCAATCAAAAGTA

Subpool-4: TCTAGGTTTCGGCTTCATGT; GGTGCATGGGAGGAACTATA

Subpool-5: ATACTGCTGGGCTGGATATC; TCCTGAGAGAATACGGATGC

Subpool-6: ACCCAAAGAACTCGATTCCT; GCTAAATGGAGTGAGGAGGT

Subpool-7: AGTCTTAGGCTTGGAGTGTC; GTAGGCTGAGTAGTGATCCC

Subpool-8: GCTCTCCGCTATCAGTAACA; GACGAAGTTCACTAGACCCA

Subpool-9: GCCTATCCTCTAGTTCTGCC; TCGAGTTAGATTGTCACCCC

The amplicons were purified from a gel and ligated to the vector of pDonor 1 (U6-*Bsa*I-ccdB-*Bsa*I-sgRNA_scaffold-hEF1a-EGFP-2A-Puro) using the Golden Gate cloning method. To ensure no loss of representation, 10 parallel transformations were performed and plated onto 15-cm plates with ampicillin selection. The intermediate

vectors were named after pDonor 2, laying out as U6-sgRNA1-*Esp*3I-*Esp*3I-sgRNA2-sgRNA_scaffold-hEF1a-EGFP-2A-Puro.

Step two of construction

The gel-purified fragments generated by *Esp*3I digestion of pDonor 3 (*Esp*3I- sgRNA_scaffold-loxp-Chl-loxp-H1-*Esp*3I) were ligated to pDonor 2 using the Golden Gate cloning method. To ensure no loss of representation, 10 parallel transformations were performed and plated onto 15-cm plates with chloramphenicol and ampicillin selection. The intermediate vectors were named after pDonor 4, with the following configuration: U6-sgRNA1-sgRNA_scaffold-loxp-Chl-loxp-H1-sgRNA2-sgRNA_scaffold-hEF1a-EGFP-2A-Puro.

Step three of construction

The pDonor 4 plasmids were transformed into competent cells expressing cyclization recombination, and plated onto 15-cm plates with ampicillin selection. To ensure no loss of representation, 10 parallel transformations were performed. In this step, the cassette of chloramphenicol was deleted, generating the final constructs (U6-sgRNA1-sgRNA_scaffold-loxp-H1-sgRNA2-sgRNA_scaffold-hEF1a-EGFP-2A-Puro).

Construction of paired-sgRNA subpools for negative controls

Step one of construction

We purchased 45 pairs of self-complementary oligonucleotides synthesized in 96-well plates. Each pair of oligonucleotides was mixed and annealed. These products were mixed in the same ratio, and then ligated into pDonor 5 (*Esp*3I-U6-*Bsa*I-sgRNA_scaffold-*Esp*3I) and pDonor 6 (*Esp*3I-H1-*Bsa*I-sgRNA_scaffold-*Esp*3I) digested by *Bsa*I, respectively. To ensure no loss of representation, 2 parallel transformations were performed and plated onto 15-cm plates with kanamycin selection. The products were intermediate vectors named after pDonor 7 (*Esp*3I-U6-sgRNA-sgRNA_scaffold-*Esp*3I) and pDonor 8 (*Esp*3I-H1-sgRNA-sgRNA_scaffold-*Esp*3I).

Step two of construction

Intermediate vectors of pDonor 7 and pDonor 8 were mixed equally and then digested with *Esp*3I to insert into pDonor 9 (*Esp*3I-ccdB-*Esp*3I-hEF1a-EGFP-2A-puro) through the Golden Gate cloning method, generating paired-sgRNAs in tandem. To ensure no loss of representation, two parallel transformations were performed and plated onto 15-cm plates with spectinomycin selection. Finally, we obtained the randomly paired-sgRNA subpool, in which the structure of plasmids was U6-sgRNA1-sgRNA_scaffold-H1-sgRNA2-sgRNA_scaffold-hEF1a-EGFP-2A-Puro.

Transfection, lentivirus production and cell infection

HEK293-FT cells were seeded onto four 6-well plates with 1 mL at a density of 1×10^6 cell/mL the day before cell transfections. Apart from 1 μg of Cas9 plasmid, each well was also transfected with 1 μg of functional paired-sgRNA plasmid or negative paired-sgRNA plasmid with Lipofectamine 3000 according to the manufacturer's protocol. Puromycin (3 μg/mL) was added 24 hours after transfection and maintained for 5 days to harvest genomic DNAs.

HEK293-FT cells were seeded onto three 10-cm plates with 1 mL at a density of 3×10^6 cell/mL the day before cell transfections for lentivirus production. Each plate was transfected with 2.5 μg of the paired-sgRNA plasmid, 5 μg of pCMV-dR8.2-dvpr and 2.5 μg of pCMV-VSV-G with Lipofectamine 3000 according to the manufacturer's protocol. Lentiviruses were harvested and filtered 48 hours later. H1650-Cas9 cells were seeded onto 6-well plate with 1 mL at a density of 1×10^6 cell/mL the day before cell infection. On the next day, cells were infected with lentiviruses along with 8 μg/mL polybrene. Puromycin (1 μg/mL) was added two days after infection and maintained for 7 days to harvest genomic DNAs.

Validation of the library cleavage performance

Twelve plasmids were selected from the library to validate the cleavage efficiency of the paired-sgRNAs. The 6 transcript IDs of lncRNA targeted by paired-sgRNAs of pg-type were ENST00000442823, ENST00000438436, ENST00000425412, ENST00000445681, ENST00000500612, ENST00000520348, and 6 transcript IDs of lncRNA targeted by paired-sgRNAs of gg-type were ENST00000317114, ENST00000448587, ENST00000439321, ENST00000428008, ENST00000458653, ENST00000415590. After harvesting genomic DNAs from cultured cells both in transfection and infection experiments, PCR were performed to validate deletion of lncRNAs. Primers for PCR were designed at the upstream and downstream of the target loci, and the sequences of 12 sets primers were as follows:

pg-type:

LncRNA-1: 5′-CTGGAGCATAGTAAGTGCTG-3′, 5′-GTGAGGTAGGCTTTATGGC-3′;

LncRNA-2: 5′-CATAGAACATCCCGAACCC-3′, 5′-GTTCTTCGATTTCACAGAGG-3′;
LncRNA-3: 5′-CTGGTAGATCAGACGTCAC-3′, 5′-CCTTAGAGGCTTTCTCCGC-3′;
LncRNA-4: 5′-CATGTCTACTGATCGGAATG-3′, 5′-GCTCTCCTTAAACTCTGTGC-3′;
LncRNA-5: 5′-CATGACCCTATGTCAGGAG-3′, 5′-GCCTTGAACTCCTGGAATG-3′;
LncRNA-6: 5′-CTCGAACTCCTGACATCGG-3′, 5′-CAGCTGTCAGCCTCAATGAG-3′;
gg-type:
LncRNA-1: 5′-CACGGATGTAACCACAGCAC-3′, 5′-ACGCCTGCTTTCCAGATCC-3′;
LncRNA-2: 5′-ATCCGGATGCCTCGTCTTG-3′, 5′-TGTGGCTGTGGGACCTTAG-3′;
LncRNA-3: 5′-GGTTAGGCCCCTTGGAAG-3′, 5′-GTGGTTGAGAAGTGGAGCAC-3′;
LncRNA-4: 5′-CGGTTTGGTGCGTGTGAAGC-3′, 5′-CCCAACTTGGAAATGGGTC-3′;
LncRNA-5: 5′-GTTTCCGTTCCCCGCAGAC-3′, 5′-ACAGGCCAATGTCAGTCC-3′;
LncRNA-6: 5′-CTCACCACAGTGGGAAGTAC-3′, 5′-GCCTTGTTCAAAACTGGGC-3′.

SUPPLEMENTARY MATERIALS

The supplementary materials can be found online with this article at https://doi.org/10.1007/s40484-020-0194-5.

ACKNOWLEDGEMENTS

The research was supported by the National Natural Science Foundation of China (31471255 to Z.X.). The authors would thank members of Xie lab for helpful discussions and support as well as Syngentech for technical support.

COMPLIANCE WITH ETHICS GUIDELINES

The authors Minzhen Tao, Qiaochu Mu, Yurui Zhang and Zhen Xie declare that they have no conflict of interests.

All procedures performed in this study were in accordance with the ethical standards of the institution or practice at which the studies were conducted, and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

REFERENCES

1. Cho, S. W., Kim, S., Kim, J. M. and Kim, J. S. (2013) Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. Nat. Biotechnol., 31, 230–232
2. Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P. D., Wu, X., Jiang, W., Marraffini, L. A., *et al.* (2013) Multiplex genome engineering using CRISPR/Cas systems. Science, 339, 819–823
3. Mali, P., Yang, L., Esvelt, K. M., Aach, J., Guell, M., DiCarlo, J. E., Norville, J. E. and Church, G. M. (2013) RNA-guided human genome engineering via Cas9. Science, 339, 823–826
4. Gilbert, L. A., Larson, M. H., Morsut, L., Liu, Z., Brar, G. A., Torres, S. E., Stern-Ginossar, N., Brandman, O., Whitehead, E. H., Doudna, J. A., *et al.* (2013) CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. Cell, 154, 442–451
5. Maeder, M. L., Linder, S. J., Cascio, V. M., Fu, Y., Ho, Q. H. and Joung, J. K. (2013) CRISPR RNA-guided activation of endogenous human genes. Nat. Methods, 10, 977–979
6. Mali, P., Aach, J., Stranges, P. B., Esvelt, K. M., Moosburner, M., Kosuri, S., Yang, L. and Church, G. M. (2013) CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. Nat. Biotechnol., 31, 833–838
7. Perez-Pinera, P., Kocak, D. D., Vockley, C. M., Adler, A. F., Kabadi, A. M., Polstein, L. R., Thakore, P. I., Glass, K. A., Ousterout, D. G., Leong, K. W., *et al.* (2013) RNA-guided gene activation by CRISPR-Cas9-based transcription factors. Nat. Methods, 10, 973–976
8. Qi, L. S., Larson, M. H., Gilbert, L. A., Doudna, J. A., Weissman, J. S., Arkin, A. P. and Lim, W. A. (2013) Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell, 152, 1173–1183
9. Chen, S., Sanjana, N. E., Zheng, K., Shalem, O., Lee, K., Shi, X., Scott, D. A., Song, J., Pan, J. Q., Weissleder, R., *et al.* (2015) Genome-wide CRISPR screen in a mouse model of tumor growth and metastasis. Cell, 160, 1246–1260
10. Koike-Yusa, H., Li, Y., Tan, E. P., Velasco-Herrera, M. C. and Yusa, K. (2014) Genome-wide recessive genetic screening in mammalian cells with a lentiviral CRISPR-guide RNA library. Nat. Biotechnol., 32, 267–273
11. Parnas, O., Jovanovic, M., Eisenhaure, T. M., Herbst, R. H., Dixit, A., Ye, C. J., Przybylski, D., Platt, R. J., Tirosh, I., Sanjana, N. E., *et al.* (2015) A genome-wide CRISPR screen in primary immune cells to dissect regulatory networks. Cell, 162, 675–686
12. Shalem, O., Sanjana, N. E., Hartenian, E., Shi, X., Scott, D. A., Mikkelsen, T., Heckl, D., Ebert, B. L., Root, D. E., Doench, J. G., *et al.* (2014) Genome-scale CRISPR-Cas9 knockout screening in human cells. Science, 343, 84–87
13. Wang, T., Wei, J. J., Sabatini, D. M. and Lander, E. S. (2014) Genetic screens in human cells using the CRISPR-Cas9 system. Science, 343, 80–84
14. Zhou, Y., Zhu, S., Cai, C., Yuan, P., Li, C., Huang, Y. and Wei, W. (2014) High-throughput screening of a CRISPR/Cas9 library for functional genomics in human cells. Nature, 509, 487–491
15. Wong, A. S. L. and Cho, G. G., Cui, C.H., Pregernig, G., Milani, P., Adam, M., Perli, S.D., Kazer, S.W., Gaillard, A., Hermann, M., *et al.* (2016). Multiplexed barcoded CRISPR-Cas9 screening enabled by CombiGEM. Proc. Natl. Acad. Sci.USA,113, 2544–2549
16. Shen, J. P., Zhao, D., Sasik, R., Luebeck, J., Birmingham, A., Bojorquez-Gomez, A., Licon, K., Klepper, K., Pekin, D., Beckett, A. N., *et al.* (2017) Combinatorial CRISPR-Cas9 screens for *de novo* mapping of genetic interactions. Nat. Methods, 14, 573–576
17. Han, K., Jeng, E. E., Hess, G. T., Morgens, D. W., Li, A. and

Bassik, M. C. (2017) Synergistic drug combinations for cancer identified in a CRISPR screen for pairwise genetic interactions. Nat. Biotechnol., 35, 463–474

18. Du, D., Roguev, A., Gordon, D. E., Chen, M., Chen, S.-H., Shales, M., Shen, J. P., Ideker, T., Mali, P., Qi, L. S., *et al.* (2017) Genetic interaction mapping in mammalian cells using CRISPR interference. Nat. Methods, 14, 577–580
19. Li, Y., Nowak, C. M., Withers, D., Pertsemlidis, A. and Bleris, L. (2018) CRISPR-based editing reveals edge-specific effects in biological networks. CRISPR J, 1, 286–293
20. Han, J., Zhang, J., Chen, L., Shen, B., Zhou, J., Hu, B., Du, Y., Tate, P. H., Huang, X. and Zhang, W. (2014) Efficient *in vivo* deletion of a large imprinted lncRNA by CRISPR/Cas9. RNA Biol., 11, 829–835
21. Ho, T. T., Zhou, N., Huang, J., Koirala, P., Xu, M., Fung, R., Wu, F. and Mo, Y. Y. (2015) Targeting non-coding RNAs with the CRISPR/Cas9 system in human cell lines. Nucleic Acids Res., 43, e17
22. Yin, Y., Yan, P., Lu, J., Song, G., Zhu, Y., Li, Z., Zhao, Y., Shen, B., Huang, X., Zhu, H., *et al.* (2015) Opposing roles for the lncRNA haunt and its genomic locus in regulating *HOXA* gene activation during embryonic stem cell differentiation. Cell Stem Cell, 16, 504–516
23. Paralkar, V. R., Taborda, C. C., Huang, P., Yao, Y., Kossenkov, A. V., Prasad, R., Luan, J., Davies, J. O. J., Hughes, J. R., Hardison, R. C., *et al.* (2016) Unlinking an lncRNA from its associated *cis* element. Mol. Cell, 62, 104–110
24. Covarrubias, S., Robinson, E. K., Shapleigh, B., Vollmers, A., Katzman, S., Hanley, N., Fong, N., McManus, M. T. and Carpenter, S. (2017) CRISPR/Cas-based screening of long non-coding RNAs (lncRNAs) in macrophages with an NF-κB reporter. J. Biol. Chem., 292, 20911–20920
25. Li, Y., Rivera, C. M., Ishii, H., Jin, F., Selvaraj, S., Lee, A. Y., Dixon, J. R. and Ren, B. (2014) CRISPR reveals a distal super-enhancer required for *Sox2* expression in mouse embryonic stem cells. PLoS One, 9, e114485
26. Xiao, A., Wang, Z., Hu, Y., Wu, Y., Luo, Z., Yang, Z., Zu, Y., Li, W., Huang, P., Tong, X., *et al.* (2013) Chromosomal deletions and inversions mediated by TALENs and CRISPR/Cas in zebrafish. Nucleic Acids Res., 41, e141
27. Essletzbichler, P., Konopka, T., Santoro, F., Chen, D., Gapp, B. V., Kralovics, R., Brummelkamp, T. R., Nijman, S. M. B. and Bürckstümmer, T. (2014) Megabase-scale deletion using CRISPR/Cas9 to generate a fully haploid human cell line. Genome Res., 24, 2059–2065
28. Rinn, J. L. and Chang, H. Y. (2012) Genome regulation by long noncoding RNAs. Annu. Rev. Biochem., 81, 145–166
29. Wutz, A., Rasmussen, T. P. and Jaenisch, R. (2002) Chromosomal silencing and localization are mediated by different domains of Xist RNA. Nat. Genet., 30, 167–174
30. Gupta, R. A., Shah, N., Wang, K. C., Kim, J., Horlings, H. M., Wong, D. J., Tsai, M. C., Hung, T., Argani, P., Rinn, J. L., *et al.* (2010) Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. Nature, 464, 1071–1076
31. Wilusz, J. E., Sunwoo, H. and Spector, D. L. (2009) Long noncoding RNAs: functional surprises from the RNA world. Genes Dev., 23, 1494–1504
32. Yan, X., Hu, Z., Feng, Y., Hu, X., Yuan, J., Zhao, S. D., Zhang, Y., Yang, L., Shan, W., He, Q., *et al.* (2015) Comprehensive genomic characterization of long non-coding RNAs across human cancers. Cancer Cell, 28, 529–540
33. Liao, Q., Liu, C., Yuan, X., Kang, S., Miao, R., Xiao, H., Zhao, G., Luo, H., Bu, D., Zhao, H., *et al.* (2011) Large-scale prediction of long non-coding RNA functions in a coding-non-coding gene co-expression network. Nucleic Acids Res., 39, 3864–3878
34. Mercer, T. R., Dinger, M. E., Sunkin, S. M., Mehler, M. F. and Mattick, J. S. (2008) Specific expression of long noncoding RNAs in the mouse brain. Proc. Natl. Acad. Sci. USA, 105, 716–721
35. Michelhaugh, S. K., Lipovich, L., Blythe, J., Jia, H., Kapatos, G. and Bannon, M. J. (2011) Mining Affymetrix microarray data for long non-coding RNAs: altered expression in the nucleus accumbens of heroin abusers. J. Neurochem., 116, 459–466
36. Gellert, P., Ponomareva, Y., Braun, T. and Uchida, S. (2013) Noncoder: a web interface for exon array-based detection of long non-coding RNAs. Nucleic Acids Res., 41, e20
37. Du, Z., Fei, T., Verhaak, R. G. W., Su, Z., Zhang, Y., Brown, M., Chen, Y. and Liu, X. S. (2013) Integrative genomic analyses reveal clinically relevant long noncoding RNAs in human cancer. Nat. Struct. Mol. Biol., 20, 908–913
38. Liu, X. H., Liu, Z. L., Sun, M., Liu, J., Wang, Z. X. and De, W. (2013) The long non-coding RNA HOTAIR indicates a poor prognosis and promotes metastasis in non-small cell lung cancer. BMC Cancer, 13, 464
39. Lu, K. H., Li, W., Liu, X. H., Sun, M., Zhang, M. L., Wu, W. Q., Xie, W. P. and Hou, Y. Y. (2013) Long non-coding RNA MEG3 inhibits NSCLC cells proliferation and induces apoptosis by affecting p53 expression. BMC Cancer, 13, 461
40. Wang, P., Chen, D., Ma, H. and Li, Y. (2017) LncRNA MEG3 enhances cisplatin sensitivity in non-small cell lung cancer by regulating miR-21-5p/SOX7 axis. OncoTargets Ther., 10, 5137–5149
41. Qiu, M., Xu, Y., Yang, X., Wang, J., Hu, J., Xu, L. and Yin, R. (2014) CCAT2 is a lung adenocarcinoma-specific long non-coding RNA and promotes invasion of non-small cell lung cancer. Tumour Biol., 35, 5375–5380
42. Shi, X., Sun, M., Liu, H., Yao, Y., Kong, R., Chen, F. and Song, Y. (2015) A critical role for the long non-coding RNA GAS5 in proliferation and apoptosis in non-small-cell lung cancer. Mol. Carcinog., 54, E1–E12
43. Tang, Q., Ni, Z., Cheng, Z., Xu, J., Yu, H. and Yin, P. (2015) Three circulating long non-coding RNAs act as biomarkers for predicting NSCLC. Cell. Physiol. Biochem., 37, 1002–1009
44. Nie, W., Ge, H. J., Yang, X. Q., Sun, X., Huang, H., Tao, X., Chen, W. S. and Li, B. (2016) LncRNA-UCA1 exerts oncogenic functions in non-small cell lung cancer by targeting miR-193a-3p. Cancer Lett., 371, 99–106
45. Li, P., Li, J., Yang, R., Zhang, F., Wang, H., Chu, H., Lu, Y., Dun, S., Wang, Y., Zang, W., *et al.* (2015) Study on expression of

lncRNA RGMB-AS1 and repulsive guidance molecule b in non-small cell lung cancer. Diagn. Pathol., 10, 63

46. Cui, Y., Zhang, F., Zhu, C., Geng, L., Tian, T. and Liu, H. (2017) Upregulated lncRNA SNHG1 contributes to progression of non-small cell lung cancer through inhibition of miR-101-3p and activation of Wnt/β-catenin signaling pathway. Oncotarget, 8, 17785–17794
47. Nie, F. Q., Sun, M., Yang, J. S., Xie, M., Xu, T. P., Xia, R., Liu, Y. W., Liu, X. H., Zhang, E. B., Lu, K. H., *et al.* (2015) Long noncoding RNA ANRIL promotes non-small cell lung cancer cell proliferation and inhibits apoptosis by silencing KLF2 and P21 expression. Mol. Cancer Ther., 14, 268–277
48. Sun, C., Li, S., Zhang, F., Xi, Y., Wang, L., Bi, Y. and Li, D. (2016) Long non-coding RNA NEAT1 promotes non-small cell lung cancer progression through regulation of miR-377-3p-E2F3 pathway. Oncotarget, 7, 51784–51814
49. Han, L., Zhang, E. B., Yin, D. D., Kong, R., Xu, T. P., Chen, W. M., Xia, R., Shu, Y. Q. and De, W. (2015) Low expression of long noncoding RNA PANDAR predicts a poor prognosis of non-small cell lung cancer and affects cell apoptosis by regulating Bcl-2. Cell Death Dis., 6, e1665
50. Luo, J., Tang, L., Zhang, J., Ni, J., Zhang, H. P., Zhang, L., Xu, J. F. and Zheng, D. (2014) Long non-coding RNA CARLo-5 is a negative prognostic factor and exhibits tumor pro-oncogenic activity in non-small cell lung cancer. Tumour Biol., 35, 11541–11549
51. Weber, D. G., Johnen, G., Casjens, S., Bryk, O., Pesch, B., Jöckel, K. H., Kollmeier, J. and Brüning, T. (2013) Evaluation of long noncoding RNA MALAT1 as a candidate blood-based biomarker for the diagnosis of non-small cell lung cancer. BMC Res. Notes, 6, 518
52. Zhang, E. B., Yin, D. D., Sun, M., Kong, R., Liu, X. H., You, L. H., Han, L., Xia, R., Wang, K. M., Yang, J. S., *et al.* (2014) P53-regulated long non-coding RNA TUG1 affects cell proliferation in human non-small cell lung cancer, partly through epigenetically regulating HOXB7 expression. Cell Death Dis., 5, e1243
53. Lin, P. C., Huang, H. D., Chang, C. C., Chang, Y. S., Yen, J. C., Lee, C. C., Chang, W. H., Liu, T. C. and Chang, J. G. (2016) Long noncoding RNA TUG1 is downregulated in non-small cell lung cancer and can regulate CELF1 on binding to PRC2. BMC Cancer, 16, 583
54. Fu, X., Li, H., Liu, C., Hu, B., Li, T. and Wang, Y. (2016) Long noncoding RNA AK126698 inhibits proliferation and migration of non-small cell lung cancer cells by targeting Frizzled-8 and suppressing Wnt/β-catenin signaling pathway. OncoTargets Ther., 9, 3815–3827
55. Yang, Y., Li, H., Hou, S., Hu, B., Liu, J. and Wang, J. (2013) The noncoding RNA expression profile and the effect of lncRNA AK126698 on cisplatin resistance in non-small-cell lung cancer cell. PLoS One, 8, e65309
56. Xie, X., Liu, H. T., Mei, J., Ding, F. B., Xiao, H. B., Hu, F. Q., Hu, R. and Wang, M. S. (2014) LncRNA HMlincRNA717 is down-regulated in non-small cell lung cancer and associated with poor prognosis. Int. J. Clin. Exp. Pathol., 7, 8881–8886
57. Nie, F. Q., Zhu, Q., Xu, T. P., Zou, Y. F., Xie, M., Sun, M., Xia, R. and Lu, K. H. (2014) Long non-coding RNA MVIH indicates a poor prognosis for non-small cell lung cancer and promotes cell proliferation and invasion. Tumour Biol., 35, 7587–7594
58. Solovyev, V. V., Shahmuradov, I. A. and Salamov, A. (2010) Identification of Promoter Regions and Regulatory Sites. In: Computational Biology of transcritption Factor Binding (*Methods in Molecular Biology*), Ladunga, I. (ed.). New York: Springer Science + Business Media, Humana Press, Chapter 5, 57–83
59. Tabach, Y., Brosh, R., Buganim, Y., Reiner, A., Zuk, O., Yitzhaky, A., Koudritsky, M., Rotter, V. and Domany, E. (2007) Wide-scale analysis of human functional transcription factor binding reveals a strong bias towards the transcription start site. PLoS One, 2, e807
60. Koudritsky, M. and Domany, E. (2008) Positional distribution of human transcription factor binding sites. Nucleic Acids Res., 36, 6795–6805
61. Nguyen, Q. H., Tellam, R. L., Naval-sanchez, M., Porto-neto, L. R., Barendse, W., Reverter, A., Haves, B., Kijas, J., Dalrymple, B. P. (2018) Mammalian genomic regulatory regions predicted by utilizing human genomics, transcriptomics and epigenetics data. GigaSceience, 7, 1–17.
62. Liu, H., Wei, Z., Dominguez, A., Li, Y., Wang, X. and Qi, L. S. (2015) CRISPR-ERA: a comprehensive designer tool for CRISPR genome editing, (gene) repression, and activation. Bioinformatics, 31, 3676–3678
63. Canver, M. C., Bauer, D. E., Dass, A., Yien, Y. Y., Chung, J., Masuda, T., Maeda, T., Paw, B. H. and Orkin, S. H. (2014) Characterization of genomic deletion efficiency mediated by CRISPR/Cas9 in mammalian cells. J. Biol. Chem., 289, 21312–21324
64. Hangauer, M. J., Vaughn, I. W. and McManus, M. T. (2013) Pervasive transcription of the human genome produces thousands of previously unidentified long intergenic noncoding RNAs. PLoS Genet., 9, e1003569
65. Zhu, S., Li, W., Liu, J., Chen, C. H., Liao, Q., Xu, P., Xu, H., Xiao, T., Cao, Z., Peng, J., *et al.* (2016) Genome-scale deletion screening of human long non-coding RNAs using a paired-guide RNA CRISPR-Cas9 library. Nat. Biotechnol., 34, 1279–1286
66. Goyal, A., Myacheva, K., Groß, M., Klingenberg, M., Duran Arqué, B. and Diederichs, S. (2017) Challenges of CRISPR/Cas9 applications for long non-coding RNA genes. Nucleic Acids Res., 45, e12

