Chapter 11

Base-Editor-Mediated circRNA Knockout by Targeting Predominantly Back-Splice Sites

Xu-Kai Ma, Xiang Gao, Mei Cao, and Li Yang

Abstract

Back-splicing of eukaryotic exon(s) leads to the production of covalently closed circular RNAs (circRNAs). Generally, most circRNAs contain overlapping sequences to their cognate linear RNAs from the same gene loci, leading to difficulties in distinguishing them from each other. A recent study has shown that some circRNAs can be specifically depleted by using base editing systems to target their predominantly back-splice sites for circularization, suggesting an efficient approach for circRNA knockout (KO). Here, we describe the detailed protocol for applying base editors to disrupt back-splice sites of predominantly circularized exons for circRNA KO at the genomic DNA level.

Key words Circular RNA, circRNA, Back-splice site, Base editor, Knockout

1 Introduction

Covalently closed circular RNAs (circRNAs), which are produced from back-splicing of exons of RNA precursors, are widely expressed in mammalian cells [1–6]. Importantly, some relatively abundant circRNAs have been shown to be involved in different physiological and pathological conditions with distinct modes of actions [4, 6]. However, whether most other circRNAs can participate in various biological pathways remains largely unexplored. This is mainly due to the lack of efficient approaches for the gain-of-function (GOF) and loss-of-function (LOF) studies of circRNAs.

Given the fact of nearly complete sequence overlapping between circRNAs and their cognate linear RNAs [1, 4, 6], commonly used GOF and LOF approaches for the functional annotation of linear RNAs might be not directly applicable to circRNA studies [3]. For instance, when overexpressing sequences

Xu-Kai Ma and Xiang Gao contributed equally with all other contributors.

Christoph Dieterich and Marie-Laure Baudet (eds.), *Circular RNAs*, Methods in Molecular Biology, vol. 2765,
https://doi.org/10.1007/978-1-0716-3678-7_11,

containing circRNA-forming exon(s), unintended linear RNA precursor and mature linear transcript(s) are more likely to be co-expressed with desired circRNA products in cells [7, 8]. Although linear RNA by-products could be reduced to a minimal level with well-designed circRNA overexpression plasmids [9, 10], additional controls are required to differentiate the contribution of measured biological effects from circRNAs or their co-expressed linear RNAs. Of note, the overexpressed linear RNA by-products during in vitro circularized RNA preparation can be removed by enzymatic digestion and further purification [11]. However, how to efficiently get rid of linear RNA by-products from in-cell circRNA overexpression is yet solved.

Meanwhile, using shRNA/RNAi for specific circRNA LOF study usually encounters with the side effect of the simultaneous depletion of cognate linear RNAs as well, due to the complementarity of a half-RNAi sequence to cognate linear RNAs [3, 12–14]. Recently, with well-designed guide RNAs (gRNAs) targeting back-splicing junction (BSJ) site sequences, efficient circRNA depletion has been achieved by using CRISPR–RfxCas13d systems, and its side effect on the cognate linear RNAs was shown to be minimal. The high specificity of RfxCas13d-BSJ-gRNA is rendered by its longer (≥18 nucleotides, preferentially ≥22 nucleotides) spacer sequence requirement for RfxCas13d targeting [13], compared with relatively shorter sequence complementarity of RNAi (core sequence less than ten nucleotides) [13, 15, 16]. A detailed protocol for efficient circRNA knockdown (KD) by using RfxCas13d is provided by another chapter in this collection {Wu H and Chen LL, Functional circular RNA screening via RfxCas13d/BSJ-gRNA system, 2023}.

Similarly, circRNA knockout (KO) at the genomic DNA level is also challenging [3]. Indeed, only a very few circRNAs have been reported to be successfully knocked out, either by deleting paired intronic complementary sequences (ICSs) across circularized exons [17] or directly deleting the circularized exon [18]. However, these cases are more likely exceptions rather than general roles [3]. In the former case, as different pairs of ICSs can affect the same circRNA formation [19], the removal of a paired ICS is indirect and inadequate for efficient circRNA KO [3]. In the latter case, as the same exon(s) can be either back-spliced for circRNAs or canonically spliced for cognate linear RNA [19] in most gene loci, the removal of whole circularized exon(s) for circRNA KO could inevitably impair the processing of linear RNAs [3].

Other than subtracting full-length circularized exons, successful circRNA KO has been also achieved by specifically targeting back-splice site sequences [20] by base editors (BEs), which conjugate nucleobase deaminases with CRISPR/Cas proteins [21, 22]. Of note, since many splice sites are involved in both back-splicing and canonical splicing, targeting those sites can

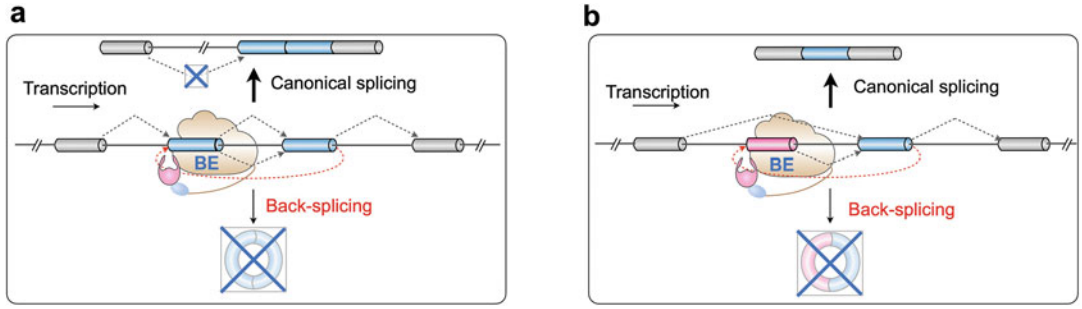

Fig. 1 Schematic of base editing and its outcomes at splice sites for linear RNA and/or circular RNA (circRNA). (**a**) Illustration of base editing at a splice site involved in both back-splice and canonical splice, by which the expression of both circRNA and linear RNA is affected. (**b**) Illustration of base editing at a back-splice site, by which only circRNA expression is repressed

generally repress both circular and linear RNA expression (Fig. 1a). Thus, it is highly suggested to carefully select predominantly back-splice sites for circRNA depletion by BEs [20]. In this chapter, we describe in detail how to use base editors to knock out circRNAs by disrupting back-splice sites of predominantly circularized exons (Fig. 1b). This technique is promising for efficient circRNA depletion at both individual and large-scale levels, which will facilitate our studies to address the roles of circRNAs in various biological processes.

2 Materials

2.1 Hardware and Software

1. 64-bit computer running Linux.
2. 8 GB of RAM (16 GB preferred).
3. CIRCexplorer2.

CIRCexplorer2 [19] and its updated versions are available at https://github.com/YangLab/CIRCexplorer2, and the detailed document is at https://circexplorer2.readthedocs.io/en/latest. This chapter is based on the version 2.4.0.

4. Python 2.7.

The python can be fetched from https://www.python.org. This chapter is based on version 2.7.18.

5. HISAT2.

HISAT2 [23] can be downloaded from http://daehwankimlab.github.io/hisat2/download/, and the manual of HISAT2 is at http://daehwankimlab.github.io/hisat2/manual/. This chapter is based on version 2.0.5.

6. TopHat2 and TopHat-Fusion.

The TopHat2 [24] can be downloaded from https://ccb.jhu.edu/software/tophat, and the information on TopHat2 is at https://ccb.jhu.edu/software/tophat/manual.shtml. This chapter is based on version 2.0.12.

7. Samtools.

The samtools [25] can be fetched from https://sourceforge.net/projects/samtools/files/samtools. This chapter is based on version 0.1.19.

8. BEDTools.

The bedtools [26] can be fetched from http://bedtools.readthedocs.io/en/latest. This chapter is based on version 2.26.0.

9. UCSC utilities.

UCSC utilities, including genePredToGtf (20180809), gtfToGenePred (20180809), bedGraphToBigWig (version 4), and bedToBigBed (version 2.7), are available at https://hgdownload.soe.ucsc.edu/admin/exe.

10. featureCounts.

The featureCounts [27] can be fetched from https://subread.sourceforge.net/featureCounts.html. This chapter is based on version 1.5.1.

11. BEable-GPS.

BEable-GPS [28] is accessible via the website http://yang-laboratory.com/BEable-GPS/ and can be downloaded from https://github.com/suduwoniu/BEable-GPS/BEable_Cas9_CtoT.py (version 1.0).

12. Bioedit.

Bioedit can be downloaded from https://bioedit.software.informer.com/7.2/. This chapter is based on version 7.2.

13. Trimmomatic.

The Trimmomatic [29] can be fetched from http://www.usadellab.org/cms/index.php?page=trimmomatic. This chapter is based on version 0.35.

14. Other python-related packages.

Python-related packages, including pysam (version 0.15.2), pybedtools (version 0.7.5), docopt (version 0.6.2), and scipy (version 1.2.1), are available at https://pypi.python.org.

15. Example codes.

The example codes shown in this chapter are also accessible at the following link: https://github.com/YangLab/BEable_circRNA_methods.

Of note, we suggest users to carefully read the associated documentation of each software before installation.

2.2 Reference Genome and Gene Annotation

1. hg38.fa.

"hg38.fa" contains human primary reference genome sequence (version GRCh38/hg38), which can be downloaded from http://hgdownload.soe.ucsc.edu/goldenPath/hg38/bigZips/hg38.chromFa.tar.gz.

2. hg38_gencode.gtf.

"hg38_gencode.gtf" is a General Transfer Format (GTF) file for gene annotation, which can be downloaded from ftp://ftp.ebi.ac.uk/pub/databases/gencode/Gencode_human/release_31/gencode.v31.annotation.gtf.gz.

This chapter is based on version 31.

2.3 RNA-Seq Datasets

To illustrate the profiling of circRNAs from deep sequencing datasets, previously published poly(A)+, poly(A)–, ribo–, and RNaseR-treated RNA-seq datasets of 293FT cells are used in this chapter (GEO: GSE172193 and GSE149691) [20].

2.4 Cell Lines and Agents for Evaluating Knockout Effect

293FT cell line: Thermo Fisher Scientific, catalog number R70007.

PGL3 vector: Addgene, Plasmid #48743.

PCMV-hA3A-eBE-Y130F: Addgene, Plasmid#113423.

BsaI-HF®v2: NEB, catalog number R3733L.

T4 DNA ligase: NEB, catalog number M0202L.

*Trans*1-T1 chemically competent cell: Transgen, catalog number CD501–02.

NucleoBond Xtra Midi kit: MACHEREY-NAGEL, catalog number 740410.50.

DMEM: Thermo Fisher Scientific, catalog number 11965118. Store at −4 °C.

Fetal bovine serum: Thermo Fisher Scientific, catalog number 10099141. Store at −20 °C.

Puromycin: Invivogen, catalog number ant-pr-1.

Trypsin–EDTA (0.25%), phenol red: Thermo Fisher Scientific, catalog number 25200072.

Lipofectamine 3000 transfection reagent: Thermo Fisher Scientific, catalog number L3000015.

Opti-MEM® I reduced-serum medium: Thermo Fisher Scientific, catalog number 11058021.

DPBS: Thermo Fisher Scientific, catalog number 14190136.

TRIzol® reagent: Thermo Fisher Scientific, catalog number 15596018.

TURBO DNA-free™ kit: Thermo Fisher Scientific, catalog number AM1907.

Ribonuclease R (RNase R): Lucigen, catalog number RNR07250.

Phenol–chloroform–isoamyl alcohol (25:24:1, v/v): Life Technologies, catalog number 15593–031.

LiCl, Sigma-Aldrich, catalog number L9650–500 G.

Glycogen, RNA grade: Thermo SCIENTIFIC, catalog number R0551.

75% ethanol (v/v).

SuperScript™ III Reverse Transcriptase: Invitrogen, catalog number 18080.

Random hexamers: TaKaRa, catalog number RR037A.

dNTP mixture: TaKaRa, catalog number T4030.

Recombinant RNasin® Ribonuclease Inhibitor: Promega, catalog number N2511.

Premix PrimeSTAR HS: Takara, catalog number R40A.

StarPrep gel extraction kit: GenStar, catalog number D205-01.

TIANamp genomic DNA kit: TIANGEN, catalog number DP304.

THUNDERBIRD SYBR qPCR mix: Toyobo, catalog number QPS-201.

3 Methods

3.1 Genome-Wide Profiling of circRNAs

First, the quality control tool, FastQC, is used to evaluate the quality of four RNA-seq datasets, such as per base sequence quality and overrepresented sequences to indicate whether RNA-seq datasets are qualified for subsequent analyses. For the 293FT ribo– RNA-seq dataset, Trimmomatic was used for the removal of adaptor sequences and low-quality bases at both ends of reads. The other three datasets are used directly in downstream analyses.

Here all four RNA-seq datasets passed this evaluation.

Command lines:

fastqc 293FT_pp.fq.gz 293FT_pm.fq.gz 293FT_rm.R1.fq.gz 293FT_rm.R2.fq.gz 293FT_rr.fq.gz

trimmomatic-0.35.jar PE -threads 3 293FT_rm.R1.fq.gz 293FT_rm.R2.fq.gz 293FT_rm.R1.trimmed.fq.gz 293FT_rm.

R1.unpaired.fq.gz 293FT_rm.R2.trimmed.fq.gz 293FT_rm.R2.unpaired.fq.gz TruSeq3-PE-2.fa:2:30:10:8:true LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:30

mv 293FT_rm.R1.trimmed.fq.gz 293FT_rm.R1.fq.gz

mv 293FT_rm.R2.trimmed.fq.gz 293FT_rm.R2.fq.gz

Then we create indexes of the human reference genome (hg38.fa).

Command lines:

bowtie-build hg38.fa hg38

bowtie2-build hg38.fa hg38

hisat2-build hg38.fa hg38

Then, reads are aligned to the human reference genome (hg38) by the aligner, HISAT2 (Fig. 2a). The input files contain the HISAT2 genome index file (hg38), the GENCODE splice site annotation file (hg38_gencode_sp.txt), and RNA-seq dataset files (293FT_pp.fq.gz, 293FT_pm.fq.gz, 293FT_rm.R1.fq.gz, 293FT_rm.R2.fq.gz, and 293FT_rr.fq.gz). The output file is a HISAT2 aligned result file in the sam format (293FT_pp.sam, 293FT_pm.sam, 293FT_rm.sam, and 293FT_rr.sam). The sam format file is converted to a bam format file (293FT_pp_sorted.bam, 293FT_pm_sorted.bam, 293FT_rm_sorted.bam, and 293FT_rr_sorted.bam), and HISAT2-unmapped fragments are extracted in a fastq format file (293FT_pm_unmapped.fq, 293FT_rm_unmapped.fq, and 293FT_rr_unmapped.fq). Command lines that use poly(A)– 293FT RNA-seq dataset as an example for the analysis are listed below:

hisat2 --no-softclip --score-min L,-16,0 --mp 7,7 --rfg 0,7 --rdg 0,7 --dta -k 1 --max-seeds 20 -p 20 -x hg38 --known-splicesite-infile hg38_gencode_sp.txt -U 293FT_pm.fq.gz -S 293FT_pm.sam &>! 293FT_pm_hisat2.log

samtools view -Sb -f 4 293FT_pm.sam > 293FT_pm_unmapped.bam

bamToFastq -i 293FT_pm_unmapped.bam -fq 293FT_pm_unmapped.fq

samtools view -Sb -F 4 293FT_pm.sam > 293FT_pm_mapped.bam

samtools sort 293FT_pm_mapped.bam 293FT_pm_sorted

samtools index 293FT_pm_sorted.bam

To retrieve fragments aligned to BSJ sites, the HISAT2-unmapped fragments (293FT_pm_unmapped.fq, 293FT_rm_unmapped.fq, and 293FT_rr_unmapped.fq) were re-aligned by TopHat-Fusion (Fig. 2a). Except HISAT2-unmapped fragment file, input files also include bowtie1 index file for reference genome (hg38). The output is a folder containing TopHat2 mapping results

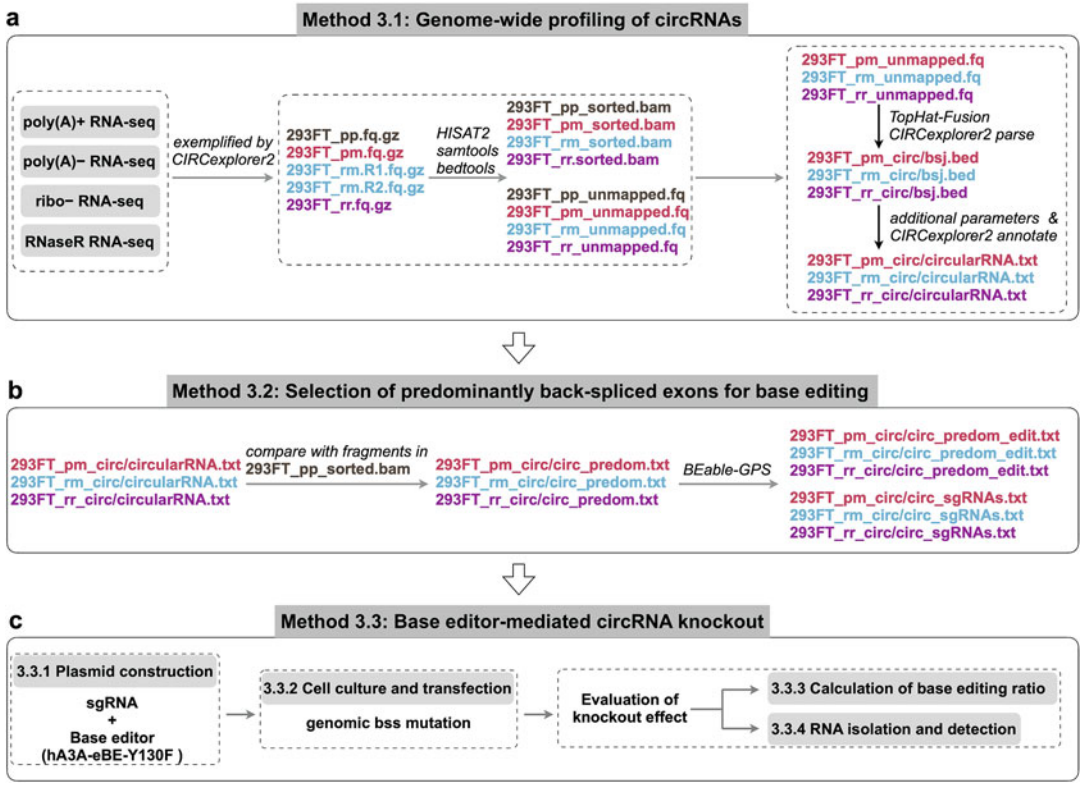

Fig. 2 A diagram of circular RNA (circRNA) annotation, sgRNA design, and evaluation of circRNA knockout by base editors. (**a**) Profiling of circRNAs from poly(A)–, ribo–, or RNaseR 293FT RNA-seq datasets with the CIRCexplorer2 pipeline. See text for details. (**b**) Designing of sgRNAs for mutating splice sites of predominantly circularized exons. See text for details. (**c**) Workflow of base editing back-splice sites of predominantly circularized exons and evaluating corresponding circRNA knockout effect. See text for details

(293FT_pm_fusion, 293FT_rm_fusion, and 293FT_rr_fusion). Right after that, we use CIRCexplorer2 to get fragments mapped to BSJ sites. The input file is the TopHat-Fusion mapping result (accepted_hits.bam in 293FT_pm_fusion, 293FT_rm_fusion, and 293FT_rr_fusion). The output is a new folder (293FT_pm_circ, 293FT_rm_circ, and 293FT_rr_circ) containing fragments aligned to BSJ sites (bsj.bed).

Command lines that use poly(A)– 293FT RNA-seq dataset as an example for the analysis are listed below:

```
tophat2 -o 293FT_pm_fusion -p 20 --fusion-search --keep-fasta-order
    --bowtie1 --no-coverage-search hg38 293FT_pm_unmapped.fq
    &> 293FT_pm_fusion.log

mkdir 293FT_pm_circ

CIRCexplorer2 parse -f -t TopHat-Fusion 293FT_pm_fusion/accep-
    ted_hits.bam -b 293FT_pm_circ/bsj.bed &> ! 293FT_pm_parse.
    log
```

To obtain high-confidence circRNAs with predominantly circularized exons, additional parameters are used to identify high-confidence BSJ site annotation (Fig. 2a) (additional parameters: mapped fragments ≥3, containing GU/AG splice site motif with 3-nt offset, length between two splice sites ≤30,000 nt). These annotations are merged with GENCODE gene annotation (hg38_gencode_merged.txt) for identifying circRNAs (circularRNA.txt in 293FT_pm_circ, 293FT_rm_circ, and 293FT_rr_circ). Then circRNAs are annotated with gene annotation file hg38_gencode_merged.txt. Except for the gene annotation file, input files also include the reference genome file (hg38.fa), and the "bsj.bed" file with the aforementioned BSJ information. The output file is "circularRNA.txt" in an extended BED12 format (in the folders 293FT_pm_circ, 293FT_rm_circ, and 293FT_rr_circ), including circRNA genome position, exons, and many other information (*see* **Note 1**).

Command line that uses poly(A)– 293FT RNA-seq dataset as an example for the analysis is listed below:

CIRCexplorer2 annotate -r hg38_gencode_merged.txt *-g hg38.fa -b* 293FT_pm_circ*/bsj.bed -o* 293FT_pm_circ*/circularRNA.txt &>!* 293FT_pm_*annotate.log*

3.2 Selection of Predominantly Circularized Exons for Base Editing

To search predominantly circularized exons for base editing, HISAT2-mapped fragments from 293FT poly(A)+ RNA-seq are included in the analyses (Fig. 2b). We use featureCounts to calculate the number of HISAT2-mapped fragments of 293FT poly(A)+ RNA-seq at back-splice sites. If the number of HISAT2-mapped fragments ≤3 (*see* **Note 2**), we define the exon as a predominantly circularized exon. Those circRNAs with predominantly circularized exons are listed in file circ_predom.txt (in the folders 293FT_pm_circ, 293FT_rm_circ, and 293FT_rr_circ).

To design specific sgRNAs for targeted BSJs, flanking regions of their back-splice sites are retrieved to find nearby PAM motifs that could fit the targeted bases at back-splice sites into the editing windows of available BEs, such as hA3A-eBE-Y130F, by previously reported BEable-GPS method [28] (Fig. 2b). Command lines that use 3′ back-splice sites of circRNAs identified from poly(A)– 293FT RNA-seq dataset as an example for the analysis are listed below:

python extract_bss.py 293FT_pm_circ/circ_predom.txt

python extract_sequences.py hg38.fa 293FT_pm_circ/3bss.bed 293FT_pm_circ/3bss.seq

python extract_sequences.py hg38.fa 293FT_pm_circ/3bss2th.bed 293FT_pm_circ/3bss2th.seq

python BEable_Cas9_CtoT.py 293FT_pm_circ/3bss.seq NGG 20 3 8 > 293FT_pm_circ/3bss_hA3A_Y130F.txt

python BEable_Cas9_CtoT.py 293FT_pm_circ/3bss2th.seq NGG 20 3 8 > 293FT_pm_circ/3bss2th_hA3A_Y130F.txt

By combining editing information of 5′ and 3′ back-splice sites (3bss_hA3A_Y130F.txt, 3bss2th_hA3A_Y130F.txt, 5bss_hA3A_Y130F.txt, and 5bss2th_hA3A_Y130F.txt), BE-editable circRNAs and their corresponding sgRNAs for predominantly circularized exons are obtained (circ_predom_edit.txt and circ_sgRNAs.txt in 293FT_pm_circ, 293FT_rm_circ, and 293FT_rr_circ).

3.3 Evaluation of Knockout Effect

Here, we select one circRNA, *circRALY(NE,2)* with predominantly circularized novel exon (NE and *see* **Note 3**), as an example to illustrate the design and the whole experiment setup for BE-mediated circRNA KO (Figs. 2c and 3). Of note, this circRNA, previously listed as *circRALY-nov* in Gao et al. [20], has now been renamed as *circRALY(NE,2)* according to the most recently published nomenclature for circRNA naming [1].

3.3.1 Plasmid Construction

Construction of the sgRNA plasmid was achieved by inserting the sgRNA sequence into the BsaI sites of the PGL3 vector (Fig. 2c).

1. Linearize the PGL3 vector by BsaI restriction enzyme and recover DNA with the StarPrep Gel Extraction Kit.
2. Anneal the sgRNA oligos by heating at 95 °C at a thermocycler for 5 min, and then gradually cool down to room temperature at a speed of −1 °C per 30 s.
3. Mix 100 ng of the linearized PGL3 vector with 5 ng of annealed oligos, and ligate the oligos to the vector by T4 DNA ligase.
4. Transform the ligation mixture into *Trans*1-T1 chemically competent cells following the standard molecular cloning protocol.
5. Check the positive clones by Sanger sequencing and prepare sgRNA plasmid by NucleoBond Xtra Midi kit.

3.3.2 Cell Culture and Transfection

1. 293FT cells (Thermo) were maintained in DMEM supplemented with 10% fetal bovine serum and 0.1% penicillin/streptomycin at 37 °C in a 5% CO_2 cell culture incubator.
2. The day before transfection, seed 293FT cells in a 12-well plate at a density of 2×10^5 cells per well.
3. Carry out co-transfection of 1.7 μg BE expressing vector and 1.1 μg sgRNA expressing vector with Lipofectamine™ 3000 Reagent according to the manufacturer's protocol.
4. 24 h after transfection, add puromycin into the culture medium at a final concentration of 1 μg/mL to select transfected cells.

5. After 3 days, collect cells into a 1.5 mL tube by trypsinization and centrifugation. Wash the cell pellet twice with DPBS and then store it at −80 °C for further analyses.

3.3.3 Calculation of Base Editing Ratio

1. Extract genomic DNA from collected cells with the TIANamp Genomic DNA Kit according to the manufacturer's protocol.
2. Amplify the genomic DNA fragments flanking target sites with primers spanning the base editing site (Fig. 3a, and *see* **Note 4**), and separate the products with 1.5% agarose gel electrophoresis (Fig. 3b).
3. Cut and recover the target bands with the StarPrep Gel Extraction Kit.

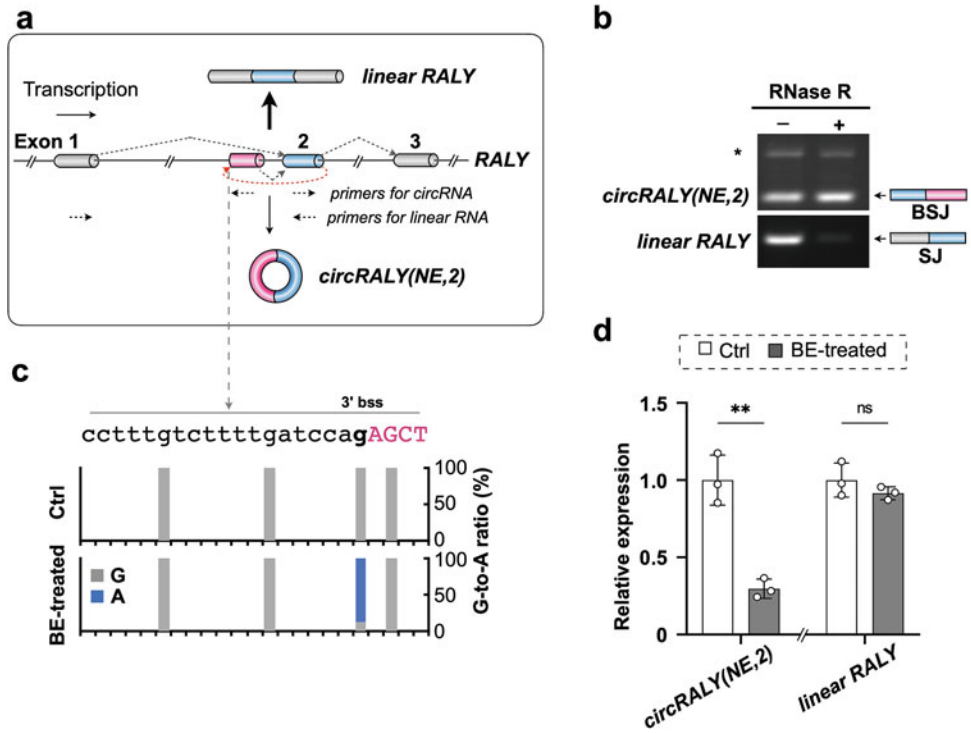

Fig. 3 Evaluation of circular RNA (circRNA) knockout effect by base editing back-splice sites. (**a**) Schematic of partial *RALY* gene organization (according to the ENST00000375114.7 transcript from GENCODE v31). A novel exon (NE, pink bar), is predominantly circularized for the production of *circRALY(NE,2)*. Of note, this specific *NE* is rarely spliced for the expression of cognate linear *RALY* RNA. (**b**) Validation of *circRALY(NE,2)* with divergent primers [20]. Note that the *circRALY(NE,2)* remains stable after RNase R treatment, while the linear *RALY* RNA is largely degraded. * indicates amplified products with divergent primers for *circRALY(NE,2)*, possibly due to a rolling PCR. (**c**) Calculation of base editing efficiency at the 3′ back-splice site of the specific predominantly circularized NE [20]. Context sequences of targeted 3′ (b)ss were shown by a, t, c, and g for intron or by A, T, C, and G for exon. (**d**) Expression of *circRALY(NE,2)* and linear *RALY* RNA after base editing. Note that *circRALY(NE,2)* is specifically repressed, while its linear *RALY* RNA is barely affected by RT-qPCR [20]

4. Analyze the base editing ratio by DNA sequencing (Fig. 2c). To calculate the C-to-T (/G-to-A) editing by CBEs or A-to-G (/T-to-C) editing by ABEs at each target site, heights of A, T, C, and G signals of Sanger sequencing were retrieved by Bioedit and processed by the following equation: Editing ratio = $[C_{height}/(C_{height} + T_{height})$ or $A_{height}/(A_{height} + G_{height})]$ (Fig. 3c and *see* **Note 5**).

3.3.4 RNA Isolation and Detection

1. Add 1 mL TRIzol to collected cells and isolate total RNAs (*see* **Note 6**) according to the manufacturer's protocol. The RNA concentration was determined by the OD_{260} using Nanodrop (Thermo).
2. Treat 10 μg total RNA with 1 μL DNase I at 37 °C for 30 min to remove genomic DNA contamination, and extract RNA with phenol–chloroform–isoamyl alcohol (*see* **Note 7**).
3. (Optional) Enrich circRNAs by RNase R digestion:
   (a) Dilute 4 μg total RNA to a total volume of 34 μL with RNase-free water in a 1.5 mL tube.
   (b) Denature the RNA at 68 °C for 5 min, and place the tube on ice immediately for 2 min.
   (c) Split the denatured RNA into two aliquots in two new 1.5 mL tubes, one for RNase R digestion and the other one for control.
   (d) For RNase R digestion, add 2 μL 10 RNase R Reaction buffer and 20 U RNase R. For control treatment, add 2 μL 10 RNase R Reaction buffer and 1 μL RNase-free water to a final volume of 20 μL. Mix thoroughly, and quickly spin the tubes for a few seconds.
   (e) Incubate at 37 °C for 30 min in a ThermoMixer with a thermoblock.
   (f) Inactive the enzyme at 70 °C for 5 min.
   (g) Add 80 μL RNase-free water to the reaction to a total volume of 100 μL and proceed with phenol/chloroform extraction, followed by ethanol precipitation and air drying of the RNA pellet.
   (h) Dissolve the RNA in 20 μL RNase-free water.
4. Synthesize first-strand cDNA:
   (a) Mix 1 μg total RNA, 1 μL random hexamers (100 μM), and 1 μL dNTP Mix(10 mM each) with RNase-free water to 13 μL. For the RNase R digestion assay, add an equal volume of RNase R enriched RNA or non-treated RNA.
   (b) Heat the mixture at 65 °C for 5 min, and immediately chill on ice for 2 min.

(c) Spin the tube briefly. Add 4 μL 5 First-Strand Buffer, 1 μL 0.1 M DTT, 1 μL Rnasin, and 1 μL SuperScript™ III RT. Gently mix well and incubate at 25 °C for 5 min.

(d) Incubate at 50 °C for 60 min.

(e) Inactivate the reaction by heating at 70 °C for 15 min.

(f) Store the cDNA at −20 °C until use.

5. Analyze circRNA expression with PCR:

(a) Design primers for circRNA or cognate linear RNA detection. Use divergent primers across the back-splicing junction site or convergent primers spanning at least two exons flanking the predominant exons to evaluate circular and linear RNA splice/expression, respectively.

(b) PCR amplification with appropriate cDNA templates and corresponding primers.

(c) Run the PCR products on a 1.5% agarose gel.

(d) Recovery of the target bands for sequencing.

6. Perform the real-time PCR using SYBR Green™ Real-time PCR Master Mix according to the manufacturer's protocol (Fig. 3d and *see* **Note 8**).

(a) Prepare each sample as follows: 5 μL SYBR Green Mix, 0.4 μL forward primer (10 μM), 0.4 μL reverse primer (10 μM), 1 μL cDNA, and 3.2 μL RNase-free water.

(b) Then run the qPCR reaction on a QuantStudio™ 6 Flex Real-Time PCR System.

4 Discussion

A large number of circRNAs have been profiled from different cell lines/tissues across eukaryotes, mainly by computational analysis of massive deep sequencing datasets [2–6]. Generally, circRNAs are co-expressed with their cognate linear RNAs from the same gene loci. That is why commonly used approaches that work well for linear RNAs may not discriminate circRNAs from their cognate linear RNAs. Possibly due to these technical obstacles, biological potentials of most circRNAs are still not clearly addressed [4, 6].

To solve this problem, new approaches with careful design have been applied to achieve efficient effects specifically on circRNAs but not linear ones, including using the state-of-the-art BEs for circRNA LOF studies [20]. Compared with the whole exon deletion by CRISPR-Cas9 for circRNA KO, BE-mediated circRNA depletion only requires the change of a few bases near back-splice sites.

However, as many circularized exons can be also co-linearly spliced for the production of linear RNAs, BE-mediated circRNA KO is inevitably challenged by the common limitation of simultaneous effects on cognate circular and linear RNAs with overlapping sequences (Fig. 1a). Thus, we suggest to screen/choose exons predominantly included in circRNAs but not in cognate linear RNAs for BE-targeting (Fig. 1b). With the development of applying other methods to achieve efficient and specific circRNA KD, such as by RfxCas13d targeting BSJ sites, the application of BEs for circRNA KO can be complementarily used for functional exploration of circRNAs.

5 Notes

1. Detailed information can be retrieved from https://circexplorer2.readthedocs.io/en/latest/modules/annotate/.
2. Different numbers of HISAT2-mapped fragments can be used as the threshold for filtering, which is also dependent on sequence depth. Here, we used at least three HISAT2-mapped fragments (for 293FT RNase R sample has ~38 M total reads) for the selection of high-confidence circRNA prediction.
3. The label with NE, indicating novel exon, is not ideal, and it should be viewed as a temporary placeholder. Further efforts to annotate the transcriptome—especially by incorporating circRNAs into databases such as GENCODE and RefSeq—should allow more unambiguous names to be given in the future.
4. Here, we generally design DNA fragment size at about 400 bp for amplification and sequencing.
5. To evaluate base editing efficiency at the target site, nearby bases across the spacer region should also be calculated, to achieve a convincing knockout effect.
6. For RNA experiments, use nuclease-free reagents, tubes, and tips.
7. The inactivator can be used to inactivate the DNase to perform an immediate reverse transcription reaction.
8. If there is more than one circular or linear transcript produced at the target gene locus, expressions of all these circRNAs and linear RNAs need to be detected.

Acknowledgments

This work was supported by the National Science Foundation of China (NSFC, grant number 31925011) and the Ministry of Science and Technology of China (MoST, grant numbers 2021YFA1300503 and 2019YFA0802804).

References

1. Chen LL, Bindereif A, Bozzoni I, Chang HY, Matera AG, Gorospe M, Hansen TB, Kjems J, Ma XK, Pek JW, Rajewsky N, Salzman J, Wilusz JE, Yang L, Zhao F (2023) A guide to naming eukaryotic circular RNAs. Nat Cell Biol 25(1):1–5
2. Kristensen LS, Andersen MS, Stagsted LVW, Ebbesen KK, Hansen TB, Kjems J (2019) The biogenesis, biology and characterization of circular RNAs. Nat Rev Genet 20(11):675–691
3. Li X, Yang L, Chen LL (2018) The biogenesis, functions, and challenges of circular RNAs. Mol Cell 71(3):428–442
4. Liu CX, Chen LL (2022) Circular RNAs: characterization, cellular roles, and applications. Cell 185(13):2390
5. Wilusz JE (2018) A 360 degrees view of circular RNAs: from biogenesis to functions, Wiley Interdiscip rev. RNA 9:e1478
6. Yang L, Wilusz JE, Chen LL (2022) Biogenesis and regulatory roles of circular RNAs. Annu Rev Cell Dev Biol 38:263–289
7. Li X, Liu CX, Xue W, Zhang Y, Jiang S, Yin QF, Wei J, Yao RW, Yang L, Chen LL (2017) Coordinated circRNA biogenesis and function with NF90/NF110 in viral infection. Mol Cell 67(2):214–227 e7
8. Liu CX, Li X, Nan F, Jiang S, Gao X, Guo SK, Xue W, Cui Y, Dong K, Ding H, Qu B, Zhou Z, Shen N, Yang L, Chen LL (2019) Structure and degradation of circular RNAs regulate PKR activation in innate immunity. Cell 177(4):865–880 e21
9. Kramer MC, Liang D, Tatomer DC, Gold B, March ZM, Cherry S, Wilusz JE (2015) Combinatorial control of drosophila circular RNA expression by intronic repeats, hnRNPs, and SR proteins. Genes Dev 29(20):2168–2182
10. Pamudurti NR, Bartok O, Jens M, Ashwal-Fluss R, Stottmeister C, Ruhe L, Hanan M, Wyler E, Perez-Hernandez D, Ramberger E, Shenzis S, Samson M, Dittmar G, Landthaler M, Chekulaeva M, Rajewsky N, Kadener S, Translation of CircRNAs (2017) Mol Cell 66(1):9–21 e7
11. Liu CX, Guo SK, Nan F, Xu YF, Yang L, Chen LL (2022) RNA circles with minimized immunogenicity as potent PKR inhibitors. Mol Cell 82(2):420–434 e6
12. Ai Y, Liang D, Wilusz JE (2022) CRISPR/Cas13 effectors have differing extents of off-target effects that limit their utility in eukaryotic cells. Nucleic Acids Res 50(11):e65
13. Li S, Li X, Xue W, Zhang L, Yang LZ, Cao SM, Lei YN, Liu CX, Guo SK, Shan L, Wu M, Tao X, Zhang JL, Gao X, Zhang J, Wei J, Li J, Yang L, Chen LL (2021) Screening for functional circular RNAs using the CRISPR-Cas13 system. Nat Methods 18(1):51–59
14. Zhang Y, Nguyen TM, Zhang XO, Wang L, Phan T, Clohessy JG, Pandolfi PP (2021) Optimized RNA-targeting CRISPR/Cas13d technology outperforms shRNA in identifying functional circRNAs. Genome Biol 22(1):41
15. Anderson EM, Birmingham A, Baskerville S, Reynolds A, Maksimova E, Leake D, Fedorov Y, Karpilow J, Khvorova A (2008) Experimental validation of the importance of seed complement frequency to siRNA specificity. RNA 14(5):853–861
16. Wolters NM, MacKeigan JP (2008) From sequence to function: using RNAi to elucidate mechanisms of human disease. Cell Death Differ 15(5):809–819
17. Zhang Y, Xue W, Li X, Zhang J, Chen S, Zhang JL, Yang L, Chen LL (2016) The biogenesis of nascent circular RNAs. Cell Rep 15(3): 611–624
18. Piwecka M, Glazar P, Hernandez-Miranda LR, Memczak S, Wolf SA, Rybak-Wolf A, Filipchyk A, Klironomos F, Cerda Jara CA, Fenske P, Trimbuch T, Zywitza V, Plass M, Schreyer L, Ayoub S, Kocks C, Kuhn R, Rosenmund C, Birchmeier C, Rajewsky N (2017) Loss of a mammalian circular RNA locus causes miRNA deregulation and affects brain function. Science 357(6357):eaam8526
19. Zhang XO, Dong R, Zhang Y, Zhang JL, Luo Z, Zhang J, Chen LL, Yang L (2016) Diverse alternative back-splicing and

alternative splicing landscape of circular RNAs. Genome Res 26(9):1277–1287

20. Gao X, Ma XK, Li X, Li GW, Liu CX, Zhang J, Wang Y, Wei J, Chen J, Chen LL, Yang L (2022) Knockout of circRNAs by base editing back-splice sites of circularized exons. Genome Biol 23(1):16
21. Chen J, Yang B, Yang L (2019) To BE or not to BE, that is the question. Nat Biotechnol 37(5): 520–522
22. Yang L, Chen J (2020) A tale of two moieties: rapidly evolving CRISPR/Cas-based genome editing. Trends Biochem Sci 45(10):874–888
23. Kim D, Langmead B, Salzberg SL (2015) HISAT: a fast spliced aligner with low memory requirements. Nat Methods 12(4):357–360
24. Kim D, Salzberg SL (2011) TopHat-fusion: an algorithm for discovery of novel fusion transcripts. Genome Biol 12(8):R72
25. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, S. (2009) Genome project data processing, the sequence alignment/map format and SAMtools. Bioinformatics 25(16): 2078–2079
26. Quinlan AR, Hall IM (2010) BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26(6):841–842
27. Liao Y, Smyth GK, Shi W (2014) featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 30(7):923–930
28. Wang Y, Gao R, Wu J, Xiong YC, Wei J, Zhang S, Yang B, Chen J, Yang L (2019) Comparison of cytosine base editors and development of the BEable-GPS database for targeting pathogenic SNVs. Genome Biol 20(1):218
29. Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30(15): 2114–2120