Methods 196 (2021) 17–22

Contents lists available at ScienceDirect

Methods

journal homepage: www.elsevier.com/locate/ymeth

Reconstruction of circular RNAs using Illumina and Nanopore RNA-seq datasets

Jinyang Zhang [a,b], Fangqing Zhao [a,b,c,d,*]

[a] Beijing Institutes of Life Science, Chinese Academy of Sciences, Beijing 100101, China
[b] University of Chinese Academy of Sciences, Beijing 100049, China
[c] Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming 650223, China
[d] Key Laboratory of Systems Biology, Hangzhou Institute for Advanced Study, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Hangzhou, China

ARTICLE INFO

Keywords:
Circular RNA
Transcript reconstruction
Transciptome

ABSTRACT

High-throughput RNA sequencing has enabled the extensive detection of circular RNAs (circRNAs) in eukaryotic organisms. However, most circRNAs are derived from exonic regions and possess sequences that are highly overlapped to their cognate linear mRNAs, which makes the reconstruction of the internal structure and full-length circular transcripts a challenging aspect in circRNA studies. To solve this problem, we provide a step-by-step protocol for the full-length reconstruction of circRNAs using CIRI-full and CIRI-long in Illumina and Nanopore RNA-seq libraries. By combining experimental and computational methods, we are able to effectively characterize the full-length landscape of circRNAs, which provide an important basis to explore the biogenesis and biological function of circRNAs.

1. Introduction

Circular RNAs (circRNAs) are a large class of RNA molecules with covalently closed structures. With the development of high-throughput sequencing technologies, studies have identified millions of circRNAs in eukaryotic organisms, and demonstrated the important roles of these circular transcripts in various biological processes and metabolic mechanisms [1–5]. The majority of reported circRNA act as competing endogenous RNAs like microRNA (miRNA) sponges [6,7] and RNA-binding protein (RBP) sponges [8–11], while other circRNAs are also suggested to participate in other processes like innate immune response [3] or DNA damage repairs [12]. Notably, the perspective that miRNA sponging is a common function of circRNAs has been challenged by recent studies [13,14]. Nevertheless, the functional roles of most circRNAs still largely depend on the existence of miRNA and RBP binding sites. Thus, reconstruction of the full-length circRNA sequences can provide essential information through characterization and mapping of these *cis*-acting regulatory sequences, which have become a crucial step in downstream functional analysis.

Due to their circular structure, circRNAs do not contain poly(A) tail structures like linear mRNA transcripts, and most RNA sequencing strategies using oligo-d(T) enrichment are unable to detect circRNAs. Moreover, most circRNAs are derived from exonic regions, which further increases the difficulty of distinguishing linear and circular reads in these overlapping regions [15]. Thus, the identification of circRNAs is currently based on the detection of reads spanning the back-spliced junction (BSJ) of circRNAs [16–22]. Most tools employ splice-aware aligners like TopHat/TopHat Fusion [23] or STAR [24] to align sequencing reads against the reference genome, or detect back-splicing events from split alignment results [20]. Furthermore, a number of specific experimental and computational tools have been developed to identify the internal structure of circRNAs (Table 1). Among these tools, CIRCexplorer2 [18], FcircSEC [25], and FUCHS [26] are able to detect the internal structure of circular isoforms, but do not provide a direct reconstruction of their full-length sequences. Other tools including circAST [27], circseq_cup [28], CIRI-full [29] and Cirit [30] are based on *de novo* reads assembly, which provides direct evidence of the full-length circRNA sequences. Notably, CIRI-full introduced reverse overlap (RO) feature for the full-length reconstruction of circRNAs and utilized bwa mem algorithm for accurate mapping of assembled sequences. Utilizing both *de novo* assembly and reference-based alignment strategies, CIRI-full can effectively reconstruct full-length sequences of circRNA transcripts regardless of their expression levels. Compared to other tools, CIRI-full provides user-friendly executable files in jar format which do not require an additional installation. However, the ability to reconstruct circRNAs is still limited by the sequencing length of Illumina RNA-

* Corresponding author at: Beijing Institutes of Life Science, Chinese Academy of Sciences, Beijing 100101, China.
E-mail address: zhfq@biols.ac.cn (F. Zhao).

https://doi.org/10.1016/j.ymeth.2021.03.017
Received 17 January 2021; Received in revised form 21 March 2021; Accepted 23 March 2021
Available online 27 March 2021

sequencing technology, where circRNAs larger than 300bps cannot be effectively reconstructed using popular paired-end 150 bp RNA-seq datasets.

In recent studies, long reads sequencing technologies including PacBio and Oxford Nanopore have been employed to detect the full-length circular sequence [31–34]. Specifically, isoCirc uses nuclease to remove 5′ overhang during reverse transcription of circRNAs and employs SplintR ligase to produce circular cDNA products. Then, rolling circle amplification (RCA) of the circular cDNAs generates long cDNA molecules for long-read sequencing, where each cDNA contains multiple copies of the template sequence. Afterwards, consensus sequences of each long-reads were generated, and aligned to reference genome for back-spliced junction detection. Meanwhile, CIRI-long implements additional poly(A) tailing before RNase R treatment to increase circRNA enrichment efficiency and uses fragment size selection to enrich long cDNAs that reverse transcribed multiple rounds on circRNAs directly. For data processing, CIRI-long utilizes a similar strategy for repetitive sequence identification and includes additional clustering and correction module to aggregate results from multiple samples. By combining circular reverse transcription and size selection strategy along with long reads sequencing technology, CIRI-long and isoCirc are able to effectively detect full-length circRNA transcripts and provide convincing evidence of their internal structure at the same time. In this study, we provide a detailed instruction to reconstruct full-length circRNA sequences using CIRI-full and CIRI-long from Illumina RNA-seq or Nanopore sequencing data, respectively (Fig. 1).

2. Materials

2.1. Hardware and software

1. Computer or server running 64-bit Linux distribution
2. GB of RAM
3. CIRI2 (v2.0.6) and CIRI-AS(v1.0.1)
   CIRI2 and CIRI-AS can be downloaded from https://sourceforge.net/projects/ciri/files/CIRI2 and https://sourceforge.net/projects/ciri/files/CIRI-AS, respectively.
4. CIRI-Full (v2.0) and CIRI-vis (v1.4)
   CIRI-Full can be downloaded from https://sourceforge.net/projects/ciri/files/CIRI-full and https://sourceforge.net/projects/ciri/files/CIRI-vis, respectively.
5. CIRI-long (v0.6.0)
   CIRI-long can be obtained from https://github.com/bioinfobiols/CIRI-long, installation instruction is available at https://ciri-cookbook.readthedocs.io/en/latest/CIRI-long_1_-installation.html.
6. Perl (>5.8)
   The Perl programming language should be packaged within the Linux operating system. If not installed, the binary distribution of Perl can be simply downloaded from ActiveState Perl https://www.activestate.com/products/perl/downloads.
7. Python 3
   The Python programming language can be easily installed using Anaconda3 https://repo.anaconda.com/archive/Anaconda3-2020.11-Linux-x86_64.sh
8. Java (>1.6) Runtime Environment
   The Java SE Runtime Environment 8u271 can be fetched from https://www.oracle.com/java/technologies/javase-jre8-downloads.html.
9. Bwa (v0.7.17-r1188)
   The source code of Bwa (Burrows-Wheeler Aligner) can be downloaded from https://github.com/lh3/bwa.
10. Porechop
    Porechop can be downloaded from https://github.com/rrwick/Porechop.

2.2. Reference genome and annotation

1. Reference genome
   The nucleotide sequence of the mouse reference genome (Release M25, GRCm38.p6) can be obtained from ftp://ftp.ebi.ac.uk/pub/databases/gencode/Gencode_mouse/release_M25/GRCm38.primary_assembly.genome.fa.gz. For convenience, we renamed the reference genome as mm10_genome.fa in the following instructions.
2. Gene annotation
   The GENCODE gene annotation on the mouse genome can be downloaded from ftp://ftp.ebi.ac.uk/pub/databases/gencode/Gencode_mouse/release_M25/gencode.vM25.annotation.gtf.gz.

2.3. RNA-seq datasets

Public Illumina and Nanopore RNA-seq datasets from previous studies [29] can be obtained from the Genome Sequence Archive [35] (https://bigd.big.ac.cn/gsa/) under accession number CRR026043 and CRR194208.

3. Analysis of circRNAs

3.1. Library preparation for circRNA sequencing

To effectively reconstruct the full-length sequence of circRNAs, sequencing strategies with longer read length are preferred (e.g. paired-end 250 bp sequencing or Nanopore sequencing). Here, we used two different datasets from mouse brain samples to demonstrate the ability of circRNA full-length reconstruction. The libraries were constructed and sequenced using paired-end 250 bp (Illumina Hiseq 2500) and long read single-molecule sequencing (Nanopore MinION), respectively. Detailed information about library preparation and sequencing can be obtained from previously published articles [29,34]. In brief, total RNA was extracted from mouse brain samples, and ribosomal RNA was depleted using the RiboErase kit (Fig. 2a). Then, optional poly(A) tailing treatments were employed before RNase R treatment to increase the digestion efficiency of RNase R. The RNase R-treated RNA was then reverse transcribed using random hexamers and SMARTer primers, and

Table 1
Summary of the computational tools for reconstructing circRNA transcripts. The programming languages, aligners/assemblers and URL of tools are listed. The "Full-length assembly" column represents the ability to assemble full-length circRNAs from sequencing reads.

Sequencing Strategy	Tool	Programming Language	Aligner / Assembler	Full-length assembly	URL
Illumina	circAST	Python	Tophat	Yes	https://github.com/xiaofengsong/CircAST
	CIRCexplorer2	Python	TopHat2/TopHat-Fusion	No	https://github.com/YangLab/CIRCexplorer2
	circseq_cup	Python	TopHat / STAR / segemehl	Yes	https://github.com/bioinplant/circseq-cup
	CIRI-full	Java	BWA mem	Yes	https://sourceforge.net/projects/ciri/files/CIRI-full
	Cirit	Java	IDBA-tran or Velvet	Yes	http://www.bio-add.org/CIRIT
	FcircSEC	R	–	No	https://github.com/tofazzal4720/FcircSEC
	FUCHS	Python & R	STAR	No	https://github.com/dieterich-lab/FUCHS
Nanopore	CIRI-long	Python	Minimap2 & BWA	Yes	https://github.com/bioinfo-biols/CIRI-long
	isoCirc	Python & R	Minimap2	Yes	https://github.com/Xinglab/isoCirc

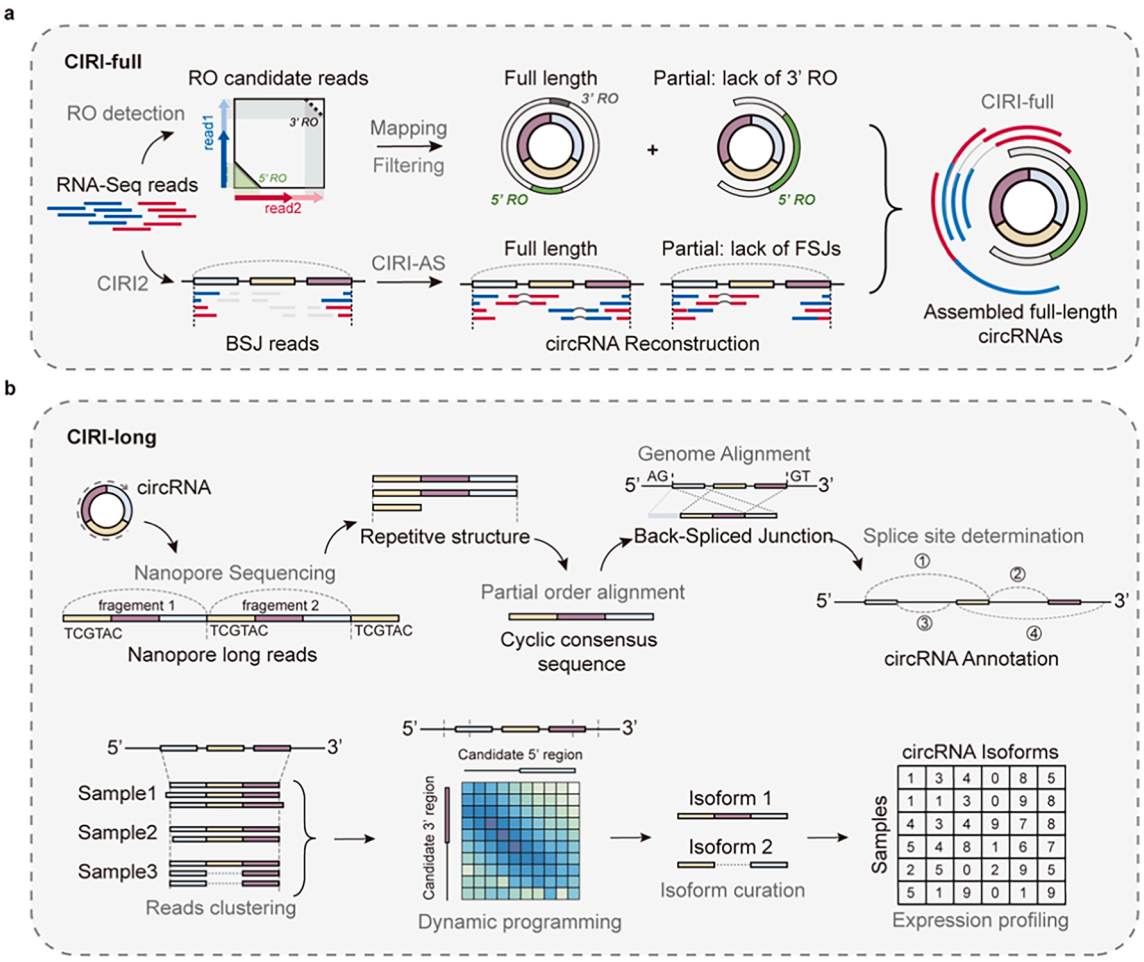

Fig. 1. The workflow of CIRI-full and CIRI-long. a CIRI-full detects the reverse overlap (RO) features in Illumina paired-end RNA-seq reads, and also employs CIRI2 and CIRI-AS pipelines to detect circular exons from back-spliced junction reads. b Top: repetitive structures in nanopore sequencing reads are detected, and cyclic consensus sequences (CCSs) for each read are constructed to reduce sequencing errors. Then, CCSs are aligned to the reference genome, and back-spliced junctions are detected from the alignment results. Bottom: candidate reads from multiple samples are further clustered based on their alignment position, and back-spliced junction coordinates are corrected using dynamic programming.

selected to proper size using magnetic beads. Finally, size-selected cDNAs were sequenced using Illumina Hiseq 2500 (paired-end 250 bp) or Nanopore MinION according to the manufacturer's instruction (Fig. 2b).

3.2. Data collection

Public RNA-seq data can be obtained from GSA using wget command. The GSA run CRR026043 contains paired-end 250 bp data for mouse brain samples, and CRR194208 contains a nanopore sequencing library for full-length circRNA reconstruction. Reads are renamed using the '-O' option to human-readable file names.

```
$ wget -O mouse_brain_pe250_1.fastq.gz \
ftp://download.big.ac.cn/gsa3/CRA000348/CRR026043/
CRR026043_f1.fq.gz
$ wget -O mouse_brain_pe250_2.fastq.gz \
ftp://download.big.ac.cn/gsa3/CRA000348/CRR026043/
CRR026043_r2.fq.gz
$wget -O ont_basecalled_reads.fastq.gz \
ftp://download.big.ac.cn/gsa/CRA003317/CRR194208/
CRR194208.fq.gz
```

3.3. Full-length assembly of circRNAs in Illumina sequencing data

First, indexes of the reference genome (mm10_genome.fa) are created to perform BWA alignment of sequencing reads. For large genomes like the mouse genome (~2.7G) in this protocol, the option "-a bwtsw" is recommended. Built index is stored in files with the following suffix (mm10_genome.amb / mm10_genome.ann / mm10_genome.bwt / mm10_genome.pac / mm10_genome.sa)

Command line:
$ bwa index -a bwtsw mm10_genome.fa

Then, the BWA mem algorithm is used to align paired-end sequencing reads to the reference genome. Option "-T 19" is used to keep all alignment results longer than 19 bp. The alignment results are stored in the SAM format (*reads.sam*). If running with a multi-core processor, users can use "-t" option to specific number of multiple processes, which can largely reduce the running step of this step.

Command line:
$ bwa mem -T 19 -t 8 mm10_genome.fa mouse_brain_pe250_1.fastq.gz mouse_brain_pe250_2.fastq.gz>./read.sam

Afterwards, CIRI2 is used to predict circRNA back-spliced junctions from alignment SAM. The results are stored in the text file specified by

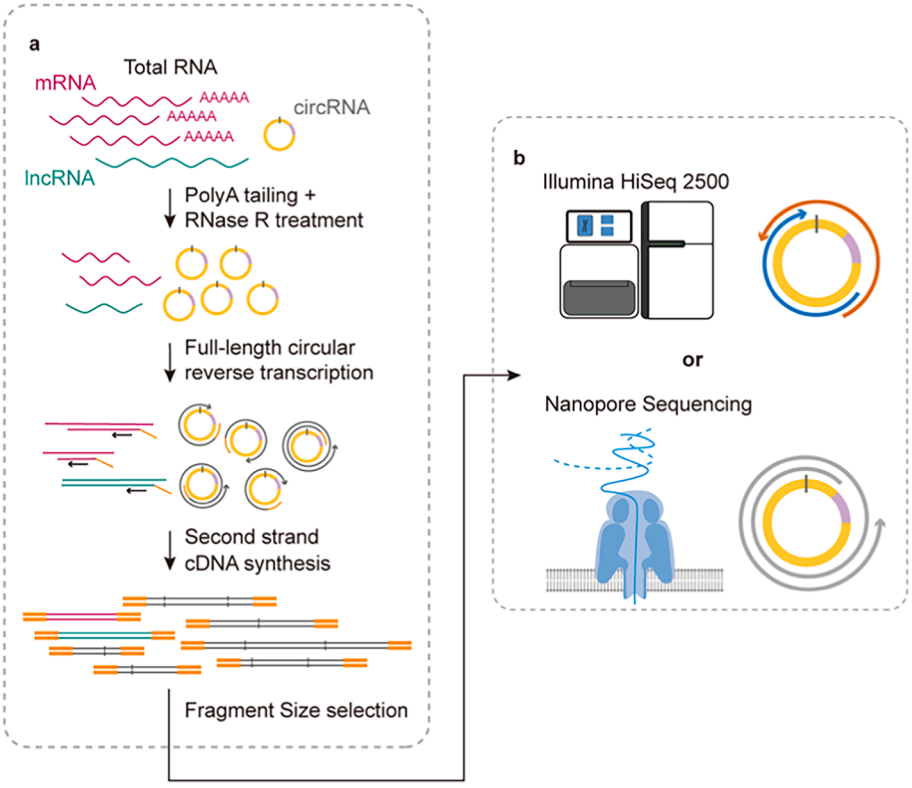

Fig. 2. Schematic view of the experimental protocol. a Total RNA is extracted and subjected to RNase R treatment to degrade linear RNAs. Optional poly(A) tailing can be employed to increase the effect of RNase R digestion. The remaining RNAs are reverse transcribed using random SMARTer primers, and selected to proper library size with magnetic beads. b The size-selected cDNA libraries are further sequenced using the Illumina HiSeq 2500 (paired-end 250 bp) and Nanopore MinION platform.

option "-O", which includes the circRNA coordinates (chr, circRNA_start, circRNA_end, circRNA_type, gene_id and strand) and the number of supporting reads (#junction reads). This step requires a relatively large amount of RAM resources, thus the thread number (specified by option "-T") should be no larger than 16 for servers with less than 128 GB of RAM.

Command line:
$ perl CIRI2.pl -I reads.sam -O prefix.ciri -F mm10_genome.fa -A gencode.vM25.annotation.gtf -T 8

CIRI-AS [36] is used to detect circular exons (cirexons) from the alignment position of circular reads. The option "-D yes" is required to keep intermediate results (*prefix_coverage.list / prefix_jav.list / prefix_splice.list*) for downstream analysis. The detailed information of cirexons and their alternative splicing events are stored in the output files (*prefix.list & prefix_AS.list*).

Command line:
$ perl CIRI_AS.pl -S reads.sam -C prefix.ciri -F mm10_genome.fa -A gencode.vM25.annotation.gtf -O prefix -D yes

Then, CIRI-full is used for *de novo* prediction of full-length circular reads with reverse overlap (RO) pattern. RO feature is the overlapped region in amplified circular transcripts, where the 5′- or 3′- ends of paired reads can be reversely overlapped with each other (Fig. 1a). Reads with RO features will be merged and aligned to the reference genome for back-spliced junction detection. In the RO2 step, the length of sequencing reads need to be specified using option "-l", and circRNAs detected from the RO reads are stored in output prefix_ro2_info.list.

Command line:
$ java -jar CIRI_full.jar RO1 −1 mouse_brain_pe250_1.fastq.gz −2 mouse_brain_pe250_2.fastq.gz -o prefix
$ bwa mem -T 19 -t 8 mm10_genome.fa ./prefix_ro1.fq>prefix_ro1.sam
$ java -jar CIRI_full.jar RO2 -r mm10_genome.fa -s ./prefix_ro1.sam -l 250 -o prefix

Finally, CIRI-full can merge CIRI-AS results and RO results from the previous step. In this step, CIRI-full will summarize the internal structure of circRNAs reported by CIRI-AS and CIRI-full, and output the full-length and partially estimated structure of circRNA. The input files *prefix.ciri*, *prefix_jav.list*, and *prefix_ro2_info.list* are generated using CIRI2, CIRI-AS and CIRI-full that have been described in the previous steps. The output file *prefix_merge_circRNA_detail.anno* contains information of circRNAs, including BSJ (circRNA id), Chr, Start, End, GTF-annotated_exon, Cirexon, Coverage, BSJ_reads_information, RO_reads_information (full-length reads), Original_gene, and strand.

Command line:
$ java -jar CIRI-full.jar Merge -a gencode.vM25.annotation.gtf -c ./prefix.ciri -as ./prefix_jav.list -ro ./prefix_ro2_info.list -o prefix -r mm10_genome.fa

Moreover, CIRI-full provides a command line wrapper for one-step analysis of full-leng circRNAs reconstruction including all steps described above. The final outputs of CIRI-full are stored under directory *output_dir/CIRI-full_output/prefix_merge_circRNA_detail.anno*

Command line:
$ mkdir CIRI-full_output
$ java -jar CIRI-full.jar Pipeline −1 mouse_brain_pe250_1.fastq.gz −2 mouse_brain_pe250_2.fastq.gz -a gencode.vM20.annotation.gtf -r mm10_genome.fa -d output_dir -o prefix -t 8

3.4. Visualization of circRNA structure

CIRI-vis [37] is a tool that provides visualization of circular isoform structure along with alignment results and coverage of circular reads. Moreover, it can estimate the abundance of circRNA isoforms derived

from the same back-spliced junction. CIRI-vis is largely dependent on the result of CIRI-full, and has been packed in the latest version of CIRI-full.

Command line:
$ java -jar ~/bioinfo_tools/CIRI-vis-1.1.jar -i ./prefix_merge_circRNA_detail.anno -l ./prefix_library_length.list -r mm10_genome.fa -d ./CIRI-vis_out -rank 10

CIRI-vis will generate one pdf for each circRNA. For example, Fig. 3 shows a circRNA derived from chr1:53865041–53891260 junction. Circles in the bottom indicate two fully reconstructed (left and right) and one partially recovered isoforms (middle) detected within this back-spliced junction. The option "-rank N" can be used to limit the display of top X% expressed circRNAs, which can largely avoid generating too many meaningless visualization results for lowly-expressed circRNAs. The output CIRI-vis_out/stout.list contains 14 columns, including Image_ID (The name of the pdf file), Circle_ID (circRNA id), Chr, Start, end, total_exp (total number of back-spliced junction reads), isoform_number, isoform_exp, isoform_length, isoform_state (whether the isoform is fully reconstructed), strand, gene_id, isoform_cirexon (0–0 represent partially reconstructed breakpoint)

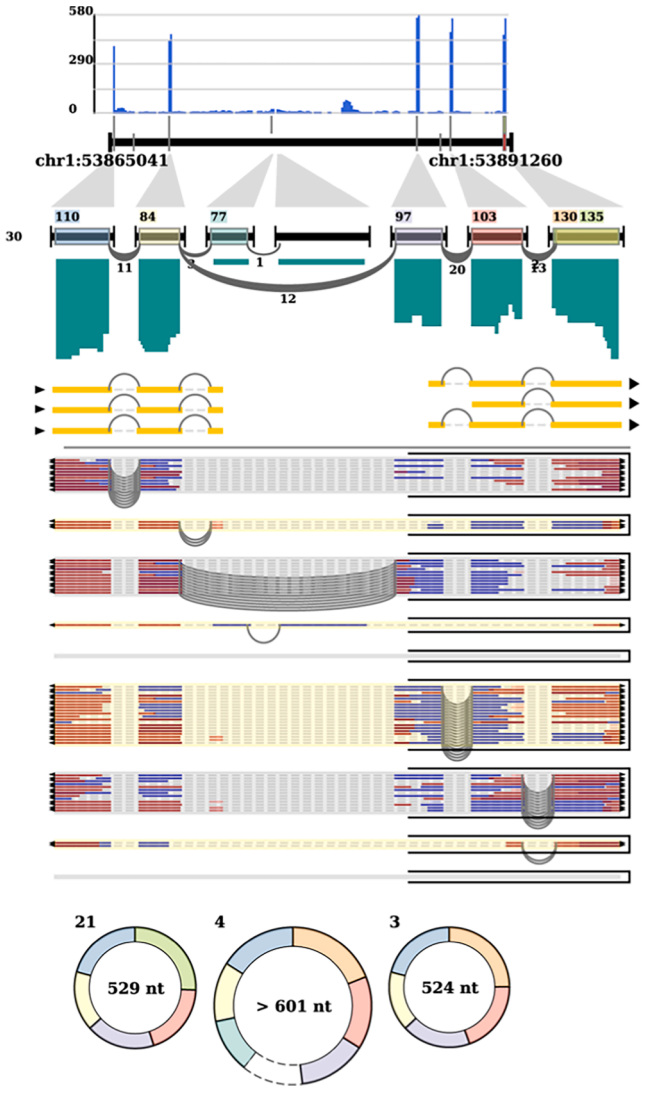

Fig. 3. Visualization of circRNA internal structures. The internal structures of three isoforms derived from the same back-spliced junction loci are shown. Tracks from top to bottom: coverage of genomic regions, length of circular exons (cirexons), coverage of cirexons, alignment pattern of BSJ reads, structure and expression levels of circular isoforms.

3.5. Detection of full-length isoforms using Nanopore sequencing data

Firstly, sequencing adapters and sample barcodes need to be removed from the nanopore sequencing data. During the library preparation step, additional SMARTer primers are subject to the library to enable amplification of non-polyA transcripts. Thus, SMARTer sequencing primers need to be trimmed from the end of reads prior to circRNA detection.

The users need to edit the adapter database of Porechop (porechop/adapters.py), adding sequences of SMARTer primers.

Adapter('SMARTer Barcode', start_sequence=('SMARTer_start', 'AAGCAGTGGTATCAACGCAGAGTAC'), end_sequence=('SMARTer_-end', 'GTACTCTGCGTTGATACCACTGCTT')),

Then, Porechop can be installed following instructions from https://github.com/rrwick/Porechop#install-from-source, and both SMARTer and Nanopore sequencing primers can be trimmed using Porechop.

Command line:
$ porechop -i ont_basecalled_reads.fastq.gz -o ont_trimmed_reads.fastq.gz –threads 8

Then, CIRI-long is used to predict circRNAs from long reads sequencing data. CIRI-long requires the same reference genome and gene annotation files as described above. The option "–canonical" can be specified to detect circRNAs with canonical "AG-GT" splice signals only.

Command line:
$ mkdir CIRI-long_output
$ CIRI-long call -i ont_trimmed_reads.fq.gz -o CIRI-long_output -p prefix -t 8 -r mm10_genome.fa -a gencode.vM25.annotation.gtf –canonical

The "call" module of CIRI-long will split raw sequencing reads into k-mers and detect repetitive circular sequences from long sequencing reads. The boundaries of repetitive fragments are determined by the occurrence of identical k-mers, and partial order alignment is performed to generate circular consensus sequences (CCS) for each circular candidate. Detected CCS are stored in *CIRI-long_output/prefix.cand_circ.fa*. Then, users can aggregate results from multiple samples using the "collapse" module, where candidate circRNAs are clustered according to the alignment position, and back-spliced junction positions are corrected using the flanking splicing signals. The input file is a tab-separated text file listing sample IDs and the absolute path of output files generated by the "call" module.

Command line:
$ printf "CIRI-long_output\tprefix.cand_circ.fa\n" > input.txt
$ CIRI-long collapse -i input.txt -o CIRI-long_output -p collapse_prefix -t 8 -r mm10_genome.fa -a gencode.vM25.annotation.gtf

By default, CIRI-long uses the canonical GT/AG splicing signal as well as splice sites annotated in the GTF file. However, users can also provide additional splice signal in the *CIRI/align.py* in the following format, the values represent the priority of each splicing signal:

```
SPLICE_SIGNAL = {
  ('GT', 'AG'): 0, # U2-type
  ('GC', 'AG'): 1, # U2-type
  ('AT', 'AC'): 2, # U12-type
  ('GT', 'AC'): 2, # U12-type
  ('AT', 'AG'): 2, # U12-type
}
```

Finally, the output file "*CIRI-long_output/collapse_prefix.info*" is

shown in standard GTF format, which contains additional information including circ_id, circ_type, circ_len, isoform, gene_id, gene_name, gene_type of each circRNA in the 9th columns. Besides, "*CIRI-long_output/collapse_prefix.expression*" provides a tab-separated matrix of supporting reads number of circRNAs in each sample, and can be subjected to differential expression analysis software like DESeq2 or edgeR for downstream analysis.

4. Conclusion

In this study, we described experimental and computational methods to reconstruct the full-length structure of circRNAs using both Illumina and Nanopore sequencing libraries. CIRI-full can detect full-length circles using reverse overlap features in short-read RNA-seq datasets, while CIRI-long is able to acquire direct evidence of full-length circRNA sequences using long reads sequencing technology. Both methods are able to effectively characterize the internal structure of circRNAs, which provide an important basis to explore the biogenesis and biological function of these circular transcripts.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by grants from the the Beijing Natural Science Foundation (JQ18020), the National Natural Science Foundation of China (32025009, 91940306) and the National Key R&D Program (2018YFC0910400).

References

[1] P. Ji, W. Wu, S. Chen, Y.i. Zheng, L. Zhou, J. Zhang, H. Cheng, J. Yan, S. Zhang, P. Yang, F. Zhao, Expanded Expression Landscape and Prioritization of Circular RNAs in Mammals, Cell Rep. 26 (12) (2019) 3444–3460.e5.
[2] L.S. Kristensen M.S. Andersen L.V.W. Stagsted K.K. Ebbesen T.B. Hansen J. Kjems The biogenesis, biology and characterization of circular RNAs 20 11 2019 675 691.
[3] C.-X. Liu, X. Li, F. Nan, S. Jiang, X. Gao, S.-K. Guo, W. Xue, Y. Cui, K. Dong, H. Ding, B. Qu, Z. Zhou, N. Shen, L. Yang, L.-L. Chen, Structure and Degradation of Circular RNAs Regulate PKR Activation in Innate Immunity, Cell 177 (4) (2019) 865–880.e21.
[4] Q. Zhao, J. Liu, H. Deng, R. Ma, J.-Y. Liao, H. Liang, J. Hu, J. Li, Z. Guo, J. Cai, X. Xu, Z. Gao, S. Su, Targeting Mitochondria-Located circRNA SCAR Alleviates NASH via Reducing mROS Output, Cell 183 (1) (2020) 76–93.e22.
[5] W. Wu, P. Ji, F. Zhao, CircAtlas: an integrated resource of one million highly accurate circular RNAs from 1070 vertebrate transcriptomes, Genome Biol. 21 (1) (2020) 101.
[6] T.B. Hansen, T.I. Jensen, B.H. Clausen, J.B. Bramsen, B. Finsen, C.K. Damgaard, J. Kjems, Natural RNA circles function as efficient microRNA sponges, Nature 495 (7441) (2013) 384–388.
[7] L.-L. Chen The expanding regulatory mechanisms and cellular functions of circular RNAs 21 8 2020 475 490.
[8] R. Ashwal-Fluss, M. Meyer, N. Pamudurti, A. Ivanov, O. Bartok, M. Hanan, N. Evantal, S. Memczak, N. Rajewsky, S. Kadener, circRNA biogenesis competes with pre-mRNA splicing, Mol. Cell 56 (1) (2014) 55–66.
[9] K. Abdelmohsen, A.C. Panda, R. Munk, I. Grammatikakis, D.B. Dudekula, S. De, J. Kim, J.H. Noh, K.M. Kim, J.L. Martindale, M. Gorospe, Identification of HuR target circular RNAs uncovers suppression of PABPN1 translation by CircPABPN1, RNA Biol. 14 (3) (2017) 361–369.
[10] L.M. Holdt, A. Stahringer, K. Sass, G. Pichler, N.A. Kulak, W. Wilfert, A. Kohlmaier, A. Herbst, B.H. Northoff, A. Nicolaou, G. Gabel, F. Beutner, M. Scholz, J. Thiery, K. Musunuru, K. Krohn, M. Mann, D. Teupser, Circular non-coding RNA ANRIL modulates ribosomal RNA maturation and atherosclerosis in humans, Nat. Commun. 7 (1) (2016) 12429.
[11] W.W. Du, W. Yang, E. Liu, Z. Yang, P. Dhaliwal, B.B. Yang, Foxo3 circular RNA retards cell cycle progression via forming ternary complexes with p21 and CDK2, Nucleic Acids Res. 44 (6) (2016) 2846–2858.
[12] X. Xu, J. Zhang, Y. Tian, Y. Gao, X. Dong, W. Chen, X. Yuan, W. Yin, J. Xu, K. Chen, C. He, L. Wei, CircRNA inhibits DNA damage repair by interacting with host gene, Mol. Cancer 19 (1) (2020) 128.
[13] L.S. Kristensen, K.K. Ebbesen, M. Sokol, T. Jakobsen, U. Korsgaard, A.C. Eriksen, T. B. Hansen, J. Kjems, H. Hager, Spatial expression analyses of the putative oncogene ciRS-7 in cancer reshape the microRNA sponge theory, Nat. Commun. 11 (1) (2020) 4551.
[14] J.U. Guo, V. Agarwal, H. Guo, D.P. Bartel, Expanded identification and characterization of mammalian circular RNAs, Genome Biol. 15 (7) (2014) 409.
[15] J. Zhang, S. Chen, J. Yang, F. Zhao, Accurate quantification of circular RNAs identifies extensive circular isoform switching events, Nat. Commun. 11 (1) (2020) 90.
[16] S. Memczak, M. Jens, A. Elefsinioti, F. Torti, J. Krueger, A. Rybak, L. Maier, S. D. Mackowiak, L.H. Gregersen, M. Munschauer, A. Loewer, U. Ziebold, M. Landthaler, C. Kocks, F. le Noble, N. Rajewsky, Circular RNAs are a large class of animal RNAs with regulatory potency, Nature 495 (7441) (2013) 333–338.
[17] L. Szabo, R. Morey, N.J. Palpant, P.L. Wang, N. Afari, C. Jiang, M.M. Parast, C. E. Murry, L.C. Laurent, J. Salzman, Statistically based splicing detection reveals neural enrichment and tissue-specific induction of circular RNA during human fetal development, Genome Biol. 16 (2015) 126.
[18] X.-O. Zhang, R. Dong, Y. Zhang, J.-L. Zhang, Z. Luo, J. Zhang, L.-L. Chen, L.i. Yang, Diverse alternative back-splicing and alternative splicing landscape of circular RNAs, Genome Res. 26 (9) (2016) 1277–1287.
[19] J. Cheng, F. Metge, C. Dieterich, Specific identification and quantification of circular RNAs from sequencing data, Bioinformatics 32 (7) (2016) 1094–1096.
[20] Y. Gao, J. Zhang, F. Zhao, Circular RNA identification based on multiple seed matching, Brief. Bioinform. 19 (5) (2018) 803–810.
[21] Y. Gao, J. Wang, F. Zhao, CIRI: an efficient and unbiased algorithm for de novo circular RNA identification, Genome Biol. 16 (2015) 4.
[22] Y. Gao, F. Zhao, Computational Strategies for Exploring Circular RNAs, Trends Genet. 34 (5) (2018) 389–400.
[23] D. Kim, S.L. Salzberg, TopHat-Fusion: an algorithm for discovery of novel fusion transcripts, Genome Biol. 12 (8) (2011) R72, https://doi.org/10.1186/gb-2011-12-8-r72.
[24] A. Dobin, C.A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski, S. Jha, P. Batut, M. Chaisson, T.R. Gingeras, STAR: ultrafast universal RNA-seq aligner, Bioinformatics 29 (1) (2013) 15–21.
[25] M.T. Hossain, Y. Peng, S. Feng, Y. Wei, FcircSEC: An R Package for Full Length circRNA Sequence Extraction and Classification, Int. J. Genomics Proteomics 2020 (2020) 9084901.
[26] F. Metge, L.F. Czaja-Hasse, R. Reinhardt, C. Dieterich, FUCHS-towards full circular RNA characterization using RNAseq, PeerJ 5 (2017) e2934.
[27] J. Wu, Y. Li, C. Wang, Y. Cui, T. Xu, C. Wang, X. Wang, J. Sha, B. Jiang, K. Wang, Z. Hu, X. Guo, X. Song, CircAST: Full-length Assembly and Quantification of Alternatively Spliced Isoforms in Circular RNAs, Genomics Proteomics Bioinformatics 17 (5) (2019) 522–534.
[28] C.-Y. Ye, X. Zhang, Q. Chu, C. Liu, Y. Yu, W. Jiang, Q.-H. Zhu, L. Fan, L. Guo, Full-length sequence assembly reveals circular RNAs with diverse non-GT/AG splicing signals in rice, RNA Biol. 14 (8) (2017) 1055–1063.
[29] Y. Zheng, P. Ji, S. Chen, L. Hou, F. Zhao, Reconstruction of full-length circular RNAs enables isoform-level quantification, Genome Med. 11 (1) (2019) 2.
[30] Y. Qin, T. Xu, W. Lin, Q. Jia, Q. He, K. Liu, J. Du, L. Chen, X. Yang, F. Du, M. Li, M. Chen, T. Tao, Z.-L. Ji, Reference-free and de novo Identification of Circular RNAs, Cold Spring Harbor Laboratory, 2020.
[31] X. You, I. Vlatkovic, A. Babic, T. Will, I. Epstein, G. Tushev, G. Akbalik, M. Wang, C. Glock, C. Quedenau, X.i. Wang, J. Hou, H. Liu, W. Sun, S. Sambandan, T. Chen, E.M. Schuman, W. Chen, Neural circular RNAs are derived from synaptic genes and regulated by development and plasticity, Nat. Neurosci. 18 (4) (2015) 603–610.
[32] K. Rahimi, M.T. Venø, D.M. Dupont, J. Kjems, Nanopore sequencing of full-length circRNAs in human and mouse brains reveals circRNA-specific exon usage and intron retention, bioRxiv (2019).
[33] R. Xin, Y. Gao, Y. Gao, R. Wang, K.E. Kadash-Edmondson, B. Liu, Y. Wang, L. Lin, Y. Xing, isoCirc catalogs full-length circular RNA isoforms in human transcriptomes, Nat. Commun. 12 (1) (2021) 266.
[34] J. Zhang, L. Hou, Z. Zuo, P. Ji, X. Zhang, Y. Xue, F. Zhao, Comprehensive profiling of circular RNAs with nanopore sequencing and CIRI-long, Nat. Biotechnol. (2021).
[35] Y. Wang, F. Song, J. Zhu, S. Zhang, Y. Yang, T. Chen, B. Tang, L. Dong, N. Ding, Q. Zhang, Z. Bai, X. Dong, H. Chen, M. Sun, S. Zhai, Y. Sun, L. Yu, L.i. Lan, J. Xiao, X. Fang, H. Lei, Z. Zhang, W. Zhao, others, GSA: genome sequence archive, Genomics Proteomics Bioinformatics 15 (1) (2017) 14–18.
[36] Y. Gao, J. Wang, Y. Zheng, J. Zhang, S. Chen, F. Zhao, Comprehensive identification of internal structure and alternative splicing events in circular RNAs, Nat. Commun. 7 (2016) 12060.
[37] Y. Zheng, F. Zhao, Visualization of circular RNAs and their internal splicing events from transcriptomic data, Bioinformatics 36(9) (2020) 2934-2935.