Bioinformatics, 36(9), 2020, 2934–2935
doi: 10.1093/bioinformatics/btaa033
Advance Access Publication Date: 17 January 2020
Applications Note

Gene expression

Visualization of circular RNAs and their internal splicing events from transcriptomic data

Yi Zheng[1] and Fangqing Zhao [1,2,*]

[1]Beijing Institutes of Life Science, Chinese Academy of Sciences, Beijing 100101, China and [2]Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming 650223, China

*To whom correspondence should be addressed.

Associate Editor: Anthony Mathelier

Received on December 1, 2019; revised on January 9, 2020; editorial decision on January 10, 2020; accepted on January 13, 2020

Abstract

Summary: Circular RNAs (circRNAs) are proved to have unique compositions and splicing events distinct from canonical mRNAs. However, there is no visualization tool designed for the exploration of complex splicing patterns in circRNA transcriptomes. Here, we present CIRI-vis, a Java command-line tool for quantifying and visualizing circRNAs by integrating the alignments and junctions of circular transcripts. CIRI-vis can be applied to visualize the internal structure and isoform abundance of circRNAs and perform circRNA transcriptome comparison across multiple samples.

Availability and implementation: https://sourceforge.net/projects/ciri/files/CIRI-vis.

Contact: zhfq@biols.ac.cn

Supplementary information: Supplementary data are available at *Bioinformatics* online.

1 Introduction

Recent studies have demonstrated the prevalence of a vast number of circular RNAs (circRNAs) in eukaryotic organisms (Gao and Zhao, 2018; Kristensen *et al.*, 2019). We have developed a number of bioinformatic tools to identify, reconstruct, quantify and annotate circRNAs from transcriptomic data (Gao *et al.*, 2015, 2016, 2018; Ji *et al.*, 2019; Zhang *et al.*, 2020; Zheng *et al.*, 2019), which have greatly expanded our knowledge of circRNAs on a genome-wide scale. Previous studies found that circRNAs contain multiple types of alternative splicing events and more than a half of circRNAs may have more than one splicing isoform (Gao *et al.*, 2016; Ji *et al.*, 2019). Considering that there is a significant overlap of genomic localization of circRNAs with linear mRNAs, interpretation and exploration of circular transcript isoforms remains a great challenge in circRNA studies.

Data visualization is indispensable to the exploration of complex splicing patterns in transcriptomes. Currently, there are a number of bioinformatic tools developed for the visualization of canonical mRNA transcripts but very few for circRNAs. Recently, CircView was designed for circRNA visualization, which can display multiple information, including exon composition, genomic location and miRNA and RBP binding sites (Feng *et al.*, 2018). However, it still lacks the ability to explore the mapping details of RNA-seq reads on circRNAs and is unable to visualize their internal structure and splicing events. Here, we developed CIRI-vis, a novel visualization approach that integrates multiple functions including circRNA transcript reconstruction, quantification and visualization.

2 Materials and methods

2.1 Retrieval of the alignment information of circRNAs

In order to retrieve the corresponding alignment information of circRNAs, CIRI-AS (Gao *et al.*, 2016) or CIRI-full (Zheng *et al.*, 2019) can be used to process the RNA-seq dataset. Briefly, BWA-MEM is firstly employed to align raw RNA-seq data against the reference genome with option '-T 19' (only output alignment with score not lower than 19). Next, CIRI-AS is applied to output circRNA's detailed information using option '-D yes', which provides two essential data for circRNA visualization and quantification: forward splice events and library length distribution. Finally, CIRI-full is used to reconstruct and quantify circRNAs by integrating both back-splice junction (BSJ) and reverse overlap (RO) of identified circRNAs.

2.2 Splicing isoform reconstruction and visualization in circRNAs

CIRI-vis first builds a forward splice graph for each circRNA using BSJ reads that are clustered with the same BSJ location. The node and the edge represent the circular exon (cirexon) and its associated forward splice junction, respectively. To exhaustively decompose the splice graph into paths, a modified depth-first search (DFS) method is applied. Considering that some cirexons and splice events may not be detected in certain circRNAs due to poor coverage or short library insert size, this DFS method is modified to search across BSJ and solve non-redundant path with or without breakpoint. For unconstructed isoforms, CIRI-vis will mark out their

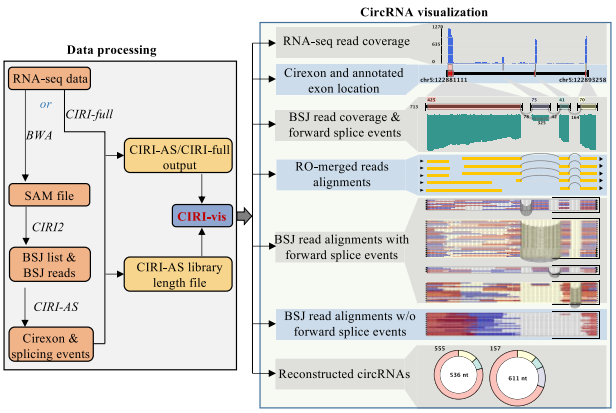

Fig. 1. The workflow of CIRI-vis including circRNA data processing and visualization

undefined regions with dashed lines. After the reconstruction of circRNA isoforms, CIRI-vis loads the file generated by CIRI-AS ('–D yes'), which contains the mapping distances of read pairs that are aligned to the same exon, representing the library insert size. Then, read pairs are simulated for each path following the inferred insert size distribution. Using these simulated read pairs on different paths, a Monte Carlo model of the expected coverage distribution on cirexons and expected isoform number on splice sites are constructed (Zheng *et al.*, 2019). After obtaining all the circular isoforms and their expression abundance, CIRI-vis can visualize the corresponding alignment and splice information in an intuitive way (Fig. 1).

2.3 Implementation

CIRI-vis is a command-line tool for circRNA internal structure visualization, which is implemented in Java programming language. It can be run on the operating systems compatible with Java 1.8 (or greater) including Mac OS, Linux and Windows. CIRI-vis utilizes Graphics2D in JAVA to display the layout of BSJ read pairs and their associated forward splice events in a Scalable Vector Graphics (SVG) format. With the help of the BATIK package, SVG files can be converted into various formats, including PDF. CIRI-vis also allows users to visually compare the same circRNA locus across multiple samples in the same image. Summary statistics can be exported to a tab-delimited text file, which includes the circRNA location, strand, expression level, isoform abundance and cirexon composition.

3 Results

3.1 Visualization of circRNAs with complex alternative splicing events

CIRI-vis provides as much detailed information as possible for users to inspect circRNA isoforms and their internal forward splice events (Fig. 1). For paired-end RNA-seq datasets, alignments of BSJ reads and RO-merged reads (Zheng *et al.*, 2019) are shown in red/blue lines and orange lines, respectively. Curved black lines represent the forward splice junctions present within circRNAs. To better understand the internal structure of circRNAs, we applied CIRI-vis to a RNA-seq dataset of human whole-brain sample (SRA Project ID: PRJNA475651), which contains a large number of circRNAs with multiple types of alternative splicing events. As shown in Supplementary Figure S1, circRNA chr13:33091994|33101669 on gene N4BP2L2 serves as an example to visualize the complex internal structure of circRNA as it is highly expressed and contains at least six alternative splicing events. Based on the split mapping of BSJ read pairs, seven cirexons can be detected within this BSJ. Specifically, five of them that have been annotated in the Gene transfer format file showed a higher sequencing coverage than the other two unannotated cirexons. Although the presence of multiple alternative splice events greatly increases the complexity of forward splice graph, RO-merged reads provide strong evidence for the splice structures of the two isoforms (length 438 and 388 nt, respectively), which represent the top two most abundant isoforms in this circRNA locus, with 1620 and 501 BSJ read counts, respectively.

3.2 Comparison of circRNA internal structures among multiple samples

We further applied CIRI-vis to compare the internal splicing changes between paired samples. Firstly, we used CIRI-full to detect and reconstruct circRNAs from the RNA-seq datasets of hepatocellular carcinoma tumor tissues and adjacent normal tissues (Yang *et al.*, 2017) and selected one circRNA for visualization (Supplementary Fig. S2A). This circRNA is located in the gene NRIP1 on chromosome 21, which consists of four cirexons and four alternative splicing isoforms. Using CIRI-vis, we can determine and compare the splice junction frequency (Supplementary Fig. S2B), cirexon coverage (Supplementary Fig. S2C) and expression levels of circular isoforms (Supplementary Fig. S2D) between tumor and normal tissues, which clearly show the discrepancy between paired samples. Besides the output of SVG or PDF-formatted graph files, all the detailed information related to circRNA's sequence, expression and splicing pattern can be stored in text files for further analysis. For users who need to analyze more than three samples, CIRI-vis can integrate circRNA visualization of multiple samples into one single graph. As shown in Supplementary Figure S3, the relative abundance of each isoform and their internal splicing junctions in all samples can be quantified, visualized and compared.

4 Conclusion

CIRI-vis is the first tool to integrate the following functions: circRNA transcript reconstruction, isoform quantification and customized visualization. Relying on these integrated functions, CIRI-vis can greatly improve our understanding of the complexity of circRNA's structure and splicing pattern, which will undoubtedly contribute to the circRNA community.

Funding

This work was supported by the grants from National Key Research & Development Program [2018YFC0910400]; and National Natural Science Foundation of China [31722031, 91640117, 91940306, 31671364 and 91531306].

Conflict of Interest: none declared.

References

Feng,J. *et al.* (2018) CircView: a visualization and exploration tool for circular RNAs. *Brief. Bioinform.*, **19**, 1310–1316.
Gao,Y. *et al.* (2015) CIRI: an efficient and unbiased algorithm for de novo circular RNA identification. *Genome Biol.*, **16**, 4.
Gao,Y. *et al.* (2016) Comprehensive identification of internal structure and alternative splicing events in circular RNAs. *Nat. Commun.*, **7**, 12060.
Gao,Y. *et al.* (2018) Circular RNA identification based on multiple seed matching. *Brief. Bioinform.*, **19**, 803–810.
Gao,Y. and Zhao,F. (2018) Computational strategies for exploring circular RNAs. *Trends Genet.*, **34**, 389–400.
Ji,P. *et al.* (2019) Expanded expression landscape and prioritization of circular RNAs in mammals. *Cell Rep.*, **26**, 3444–3460.
Kristensen,L.S. *et al.* (2019) The biogenesis, biology and characterization of circular RNAs. *Nat. Rev. Genet.*, **20**, 675–691.
Yang,Y. *et al.* (2017) Recurrently deregulated lncRNAs in hepatocellular carcinoma. *Nat. Commun.*, **8**, 14421.
Zhang,J. *et al.* (2020) Accurate quantification of circular RNAs identifies extensive circular isoform switching events. *Nat. Commun.*, **11**, 90.
Zheng,Y. *et al.* (2019) Reconstruction of full-length circular RNAs enables isoform-level quantification. *Genome Med.*, **11**, 2.