Detecting circular RNA from high-throughput sequence data with deBruijn graph.

Abstract:

- To detect circular RNA, we used a new method to classify circular RNA, which is CircDBG, which is based on the De Bruijn diagram.
- We mentioned the existence of an RNA when we compile the RNA with the reading, which CircDBG found to be fictitious based on high sequence data.
- We used the CircDBG approach, which is based on De Bruijn, since it was differentiated from other methods by its ability to minimise bias, reduce operating time, and alert it to the existence of any potential circular RNA, aswell as its ability to balance accuracy and sensitivity.

Introduction:

- Circular RNA has a circular structure and is a type of non-coding RNA. Exons are present in many circular RNAs,but they are not converted into proteins. Circular RNA plays an essential role in gene regulation. And plays an important role in some human diseases.
- There are two types of experimental methods currently that can be used to identify circular RNA . First, since circRNAs lack a poly(A) tail , they can be retained in rRNA-depleted libraries by using expected depletion profileto assess results. Second, circRNA has the potential to be enriched in RNase R was used to digest linear RNA in libraries, making it easier to detect lowly expressed circRNA.
- Multiple people can be sequenced at the same time using high-throughput sequencing technologies. Recently,bioinformatics tools for circRNA detection from RNA sequence reads have been created.
- Some of them require gene annotation Those methods could be divided into two categories: (a) readsmappingbased methods, such as CIRI/CIRI2 ,CIRCExplorer , Find-circ and CircRNAFinderand (b) k-mer-based methods,such as CircMarker ..Reads-mapping-based methods first map the RNA-seqreads onto a reference. For this purpose, CIRI uses BWA, while bowtie and Tophat (TopHat-fusion) are used by Find-circ and CIRCExplorerrespectively.
- methods have two major issues. First, reads-mapping based tools are often computationally inefficient because mapping all reads can be slow, yet we note that many RNA-seq reads are irrelevant to circRNA detection.Second, these tools may miss circRNA in some cases due to errors in reads mapping.
- we developed a k-mer-based tool called CircMarker , which uses an efficient k-mer table for circular RNA detection. Compared with the readsmapping-based method, CircMarker has two major advantages. First,CircMarker looks for the circRNArelated reads for detection and does not depend on any third party mapping tool. Thus CircMarker is much faster than reads-mapping-based methods, especially for small data. Second, since the minimum comparison unit for CircMarker is a k-mer rather than reads, it can toleratemore errors and find more circular RNAs.
- We present CircDBG, a new de Bruijn graph-based approach for detecting circular RNA. Distinctive CircDBG takes the de Bruijn graph and applies it in a new way. a specialised method for referring to circular RNA, which is the first circular RNA algorithm based on the de Bruijn graph the identification of Experiments focused on simulated and real-world data Using real-world data, we show that this new approach is efficient.

Realated work

This article was originally published in BMC Genomics, Volume 20. This is a good example. BMC Genomics, Volume 21 Supplement, has published this report. 1st of January, 2020: Selected papers from the 14th International Symposium on ISBRA-18 (International Symposium on Bioinformatics Research and Applications): genomics The whole storey is here. The supplement's contents can be found at https://bmcgenomics.biomedcentral.com/articles/supplements/volume-21-supplement-1
The publication costs are funded by IIS-1526415 and CCF-1718093 from US National Science Foundation. Funding agency has no role in design of the study and collection,analysis, and interpretation of data and in writing the manuscript

1. Hsiao K-Y, Sun HS, Tsai S-J. Circular rna–new member of noncoding rna with novel functions. Exp Biol Med. 2017;242(11):1136–41.
2. Greene J, Baird A-M, Brady L, Lim M, Gray SG, McDermott R, Finn SP.Circular rnas: biogenesis, function and role in human diseases. Front Mol Biosciences. 2017;4:38.
3. Boeckel J-N, Jaé N, Heumüller AW, Chen W, Boon RA, Stellos K, Zeiher AM, John D, Uchida S, Dimmeler S. Identification and characterization of hypoxia-regulated endothelial circular RNA. Circ Res. 2015;117(10):884–890.
4. Holdt LM, Kohlmaier A, Teupser D. Molecular roles and function of circular RNAs in eukaryotic cells. Cell Mol Life Sci. 2017;75(6):1071–1098.
5. Szabo L, Salzman J. Detecting circular rnas: bioinformatic and experimental challenges. Nat Rev Genet. 2016;17(11):679.
6. Salzman J, Gawad C, Wang PL, Lacayo N, Brown PO. Circular rnas are the predominant transcript isoform from hundreds of human genes in diverse cell types. PloS ONE. 2012;7(2):30733.
7. Gao Y, Wang J, Zhao F. Ciri: an efficient and unbiased algorithm for de novo circular rna identification. Genome Biol. 2015;16(1):4.
8. Gao Y, Zhang J, Zhao F. Circular RNA identification based on multiple seed matching. Brief Bioinforma. 2017;19(5):803–810.
9. Zhang X-O, Wang H-B, Zhang Y, Lu X, Chen L-L, Yang L. Complementary sequence-mediated exon circularization. Cell. 2014;159(1):134–47.
10. Memczak S, Jens M, Elefsinioti A, Torti F, Krueger J, Rybak A, Maier L,Mackowiak SD, Gregersen LH, Munschauer M, et al. Circular rnas are a large class of animal rnas with regulatory potency. Nature.2013;495(7441):333.
11. Westholm JO, Miura P, Olson S, Shenker S, Joseph B, Sanfilippo P, Celniker SE, Graveley BR, Lai EC. Genome-wide analysis of drosophila circular rnas reveals their structural and sequence properties and age-dependent neural accumulation. Cell Rep. 2014;9(5):1966–80.
12. Li X, Chu C, Pei J, Mandoiu I, Wu Y. Circmarker: a fast and accurate ˇalgorithm for circular rna detection. BMC Genomics. 2018;19(6):175
13. Li H, Durbin R. Fast and accurate short read alignment withburrows–wheeler transform. Bioinformatics. 2009;25(14):1754–60.
14. Langmead B, Salzberg SL. Fast gapped-read alignment with bowtie 2.Nat Methods. 2012;9(4):357.
15. Trapnell C, Pachter L, Salzberg SL. Tophat: discovering splice junctions with rna-seq. Bioinformatics. 2009;25(9):1105–11.
16. Chen X, Han P, Zhou T, Guo X, Song X, Li Y. circrnadb: a comprehensivedatabase for human circular rnas with protein-coding annotations. Sci Rep. 2016;6:34985.
17. Glažar P, Papavasileiou P, Rajewsky N. circbase: a database for circular rnas. RNA. 2014;20(11):1666–70.