BMC Bioinformatics

SpliceV: analysis and publication quality printing of linear and circular RNA splicing, expression and regulation

Nathan Ungerleider and Erik Flemington[*]

Abstract

Background: In eukaryotes, most genes code for multiple transcript isoforms that are generated through the complex and tightly regulated process of RNA splicing. Despite arising from identical precursor transcripts, alternatively spliced RNAs can have dramatically different functions. Transcriptome complexity is elevated further by the production of circular RNAs (circRNAs), another class of mature RNA that results from the splicing of a downstream splice donor to an upstream splice acceptor. While there has been a rapid expansion of circRNA catalogs in the last few years through the utilization of next generation sequencing approaches, our understanding of the mechanisms and regulation of circular RNA biogenesis, the impact that circRNA generation has on parental transcript processing, and the functions carried out by circular RNAs remains limited.

Results: Here, we present a visualization and analysis tool, SpliceV, that rapidly plots all relevant forward- and back-splice data, with exon and single nucleotide level coverage information from RNA-seq experiments in a publication quality format. SpliceV also integrates analysis features that assist investigations into splicing regulation and transcript functions through the display of predicted RNA binding protein sites and the configuration of repetitive elements along the primary transcript.

Conclusions: SpliceV is an easy-to-use splicing visualization tool, compatible with both Python 2.7 and 3+, and distributed under the GNU Public License. The source code is freely available for download at https://github.com/flemingtonlab/SpliceV and can be installed from PyPI using pip.

Keywords: circRNA, Alternative splicing, Isoform, Exon skipping, Intron retention

Background

The majority of mammalian genes code for multiple transcript isoforms that contribute substantially to the vast complexity of both the mammalian transcriptome and proteome (E. T. [25, 38]). Each mature isoform is generated through a dynamic series of tightly coordinated actions that begin to occur as the nascent transcript is being synthesized [3]. The growing precursor RNA is sequentially bound by a myriad of RNA binding proteins (RNABPs) and small nucleolar RNAs (snoRNAs; reviewed in Wahl et al [37]) as the exon-intron boundaries become defined through these specific ribonucleoprotein complex interactions. The assembled ribonucleoprotein complex, termed the spliceosome, facilitates intron excision and covalent ligation of flanking exons across the gene locus, ultimately generating a mature transcript isoform.

While each exon-intron boundary inherently contains a splice site, contiguous exons are not always spliced together. Retained introns (Y. [18]), skipped exons [20], and cryptic splice sites [14] commonly diversify the profile of fully processed transcript isoforms. Splice site proximity, defined by RNA secondary structure, is a major factor in splice site selection [28]. Intron length and the presence or absence of inverted repeats can impact the physical distance between splice donor and acceptor [33]. Branch point sequence motifs [43] and nucleotides adjacent to splice sites [5] fine tune the strength of snoRNA interactions. Further, variations in polypyrimidine tracts can preferentially attract one

* Correspondence: erik@tulane.edu
Department of Pathology, Tulane Cancer Center, Tulane University School of Medicine, New Orleans, LA 70112, USA

RNABP over another [32]. An additional layer of regulation is provided by the cellular abundance and availability of individual RNABPs and snoRNAs, allowing for tissue and context specificity of RNA processing and alternative isoform expression [29].

The same splicing reaction that generates mature mRNAs can also fuse a downstream splice donor to an upstream splice acceptor (much like tying the end of a string to the beginning), in effect circularizing the transcript [34]. These circular RNAs (circRNAs) are covalently closed transcripts that inherently lack 5′ or 3′ ends, thereby enabling them to escape exonuclease destruction. This class of RNA has recently been shown to be evolutionarily conserved ([30]; P. L. [39]), highly abundant in humans, and for some genes, is the most prevalent transcript isoform [31]. The 3′ to 5′ back-splicing reaction, required for circRNA biogenesis, correlates with the speed of precursor transcript elongation (Y. [42]), occurring more frequently at splice sites flanked by long introns and introns containing reverse complementary sequences [12]. To date, little is known regarding the function of the vast majority of circRNAs. Of the relatively few that have been characterized, some have been shown to serve as microRNA sponges [11, 22], as direct regulators of parental gene expression (Z. [19]), in signaling between cells [15] and even as templates for translation [16, 24]. Further evidence of their importance in the cell is accumulating and their functions and mechanisms of action are being found to be generally quite distinct from their cognate linear counterparts [6].

Exploring the relationship between linear and circular RNA isoforms of a common parental gene can be facilitated by utilizing Next Generation Sequencing (NGS) technology. NGS based approaches have provided the framework to study the abundance of individual transcript isoforms at a large scale, allowing investigators to compare circular and linear isoform abundance. However, the majority of bioinformatic pipelines require prior knowledge of transcript structure. While useful for broad scale interpretations, these approaches fail to resolve the abundances of both linear and circular isoforms of each gene, the function of which can dramatically differ from one another. Between linear transcripts alone, alternatively spliced isoforms can code for proteins that are truncated [4], lack specific functional domains [26], have completely unique amino acid sequences [1], and in some cases, alter cell fate entirely [4, 10]. Here we present a visualization tool, SpliceV, that facilitates detailed exploration and visualization of transcript isoform expression in publication quality format. SpliceV facilitates within- and across-sample analyses and includes the display of predicted cis and trans regulatory factors to further assist in the biogenesis and function studies. Together, SpliceV should be a useful tool for a wide spectrum of the RNA biology research community.
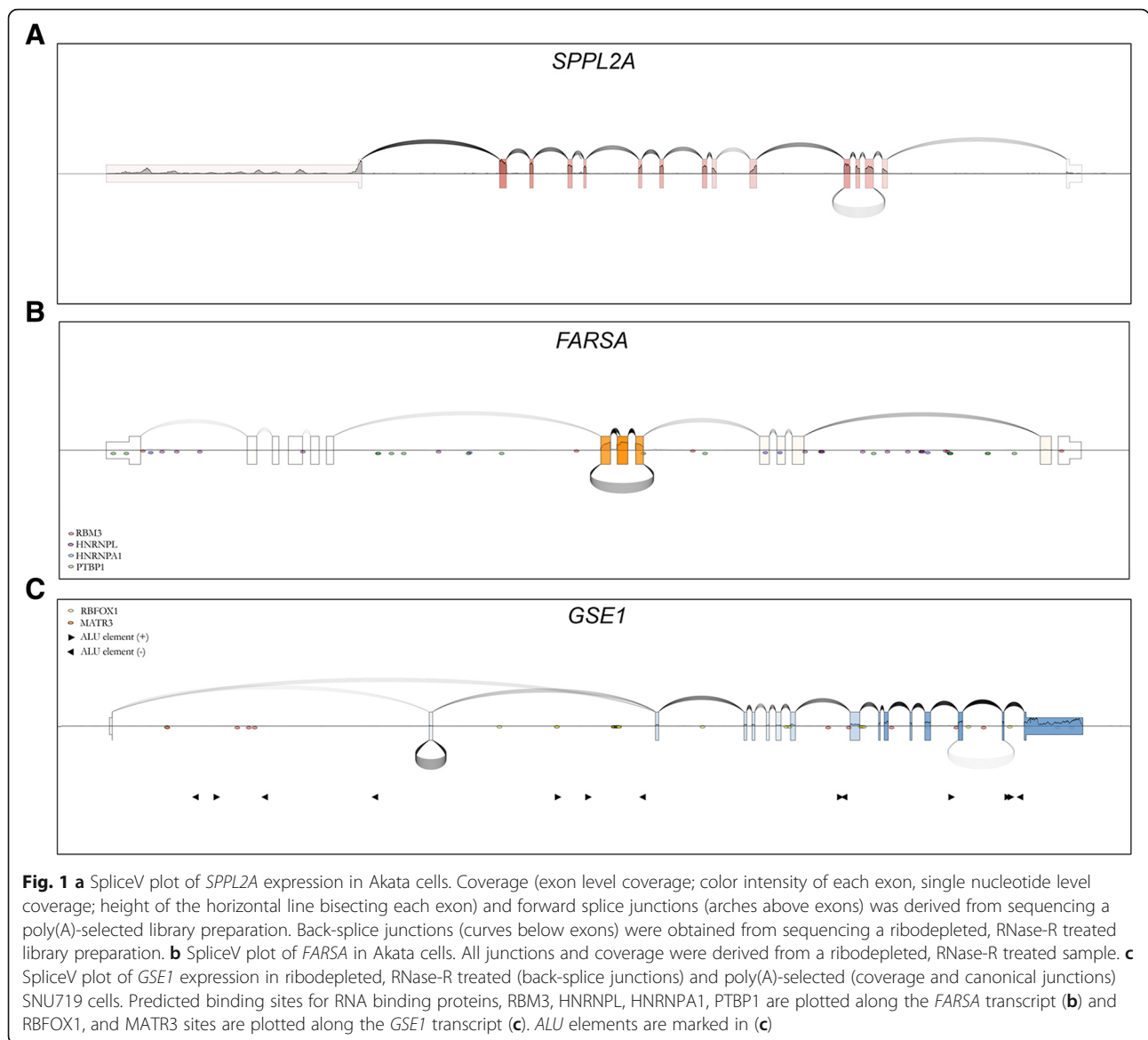
Implementation

Our software package is written in Python 3 but is backwards compatible with Python 2.7, relying only upon the third-party libraries, matplotlib [36], and pysam. Source code can be found at https://github.com/flemingtonlab/SpliceV and can be installed from PyPI using the Python package manager, pip. SpliceV is written with a GNU 3.0 public license, provided with anonymous download and installation. Full usage information can be found in Additional file 1.

SpliceV generates plots of coverage, splice junctions, and back-splice junctions with customizable parameters, depicting expression of both the linear and circular isoforms of a given gene. Standard formats (BAM, GTF, and BED) are accepted as input files. BAM files are sequentially accessed by our software (rather than in parallel). In practice, this means that SpliceV first determines the chromosomal coordinates that mark the beginning and end of the input gene. Next, it extracts reads that fall within that range from each BAM file (one BAM file at a time). As BAM files are indexed (either prior to running SpliceV, or automatically by SpliceV), this process never requires loading of the entire file into memory, and we have no reason to believe that a personal laptop computer would have difficulty running SpliceV on many BAM files at once. Because junction calling sensitivity can be improved using specialized software, canonical and back-splice junction information can be extracted directly from BAM files or input separately as BED-formatted files containing the coordinates and quantities of each junction. The user is provided the flexibility of normalizing expression of each exon across all samples or for exon normalization to be confined within each sample (this helps visualize alternative splicing, intron retention, and exon exclusion). As introns are generally much larger than exons, an option to reduce intron size by a user-defined amount is also provided. In an effort to guide interpretation of gene specific splicing patterns, predicted or empirically determined RNA binding protein binding sites can be added to the plots (Fig. 1b-c; a stepwise tutorial to reproduce these figures is outlined in Additional file 2) by supplying a list of coordinates or utilizing the consensus binding sequences determined by *Ray* et al [27]. Because inverted *ALU* repeat elements impact RNA secondary structure, we have also incorporated the option to add a track of *ALU* elements to the plot.

Results

Multiple computational pipelines have been developed to detect and quantify circRNAs from high throughput RNA sequencing data ([13, 22, 40]; X.-O. [8, 41]). As

Fig. 1 a SpliceV plot of *SPPL2A* expression in Akata cells. Coverage (exon level coverage; color intensity of each exon, single nucleotide level coverage; height of the horizontal line bisecting each exon) and forward splice junctions (arches above exons) was derived from sequencing a poly(A)-selected library preparation. Back-splice junctions (curves below exons) were obtained from sequencing a ribodepleted, RNase-R treated library preparation. **b** SpliceV plot of *FARSA* in Akata cells. All junctions and coverage were derived from a ribodepleted, RNase-R treated sample. **c** SpliceV plot of *GSE1* expression in ribodepleted, RNase-R treated (back-splice junctions) and poly(A)-selected (coverage and canonical junctions) SNU719 cells. Predicted binding sites for RNA binding proteins, RBM3, HNRNPL, HNRNPA1, PTBP1 are plotted along the *FARSA* transcript (**b**) and RBFOX1, and MATR3 sites are plotted along the *GSE1* transcript (**c**). *ALU* elements are marked in (**c**)

circRNAs lack a poly(A) tail, ribodepleted library preparations are essential for circRNA detection. RNA preparations can then be treated with the exonuclease, RNase R, which exclusively digests linear RNAs, to increase the depth of circRNA coverage. To demonstrate the utility of SpliceV, we used libraries prepared from poly(A) selected (enriched for polyadenylated linear RNAs) or ribodepleted-RNase R-treated RNA from the Burkitt's Lymphoma cell line, Akata, and the gastric carcinoma cell line, SNU719. Reads from each library were aligned using the STAR aligner v2.6.0a [7] to generate BAM and splice junction BED files. We further processed our alignments using find_circ [22] to interrogate the unmapped reads for back-splice junctions. Our first plot displays a prominent circular RNA formed via back-splicing from exon 5 to 3 of *SPPL2A* (Fig. 1a). For

this plot, back-splicing (under arches) derived from RNase R-seq data is plotted with forward splicing (over arches) and exon level (exon color intensity) and single nucleotide level (horizontal line graph) coverage from poly(A)-RNA-seq data from Akata cells to illustrate circRNA data in the context of linear poly(A) transcript expression. Exon level coverage display provides easy visualization of selective exon utilization: for example, using forward- and back-splicing and coverage data derived from RNase R-seq data (Akata cells) show enriched coverage of the circular RNA exons 6–8 of the *FARSA* gene (Fig. 1b). Nevertheless, the simultaneous display of single nucleotide level coverage includes additional information that can help provide more detailed clarity in interpretation. For example, while the last exon of *SPPL2A* (Fig. 1a) shows low exon level
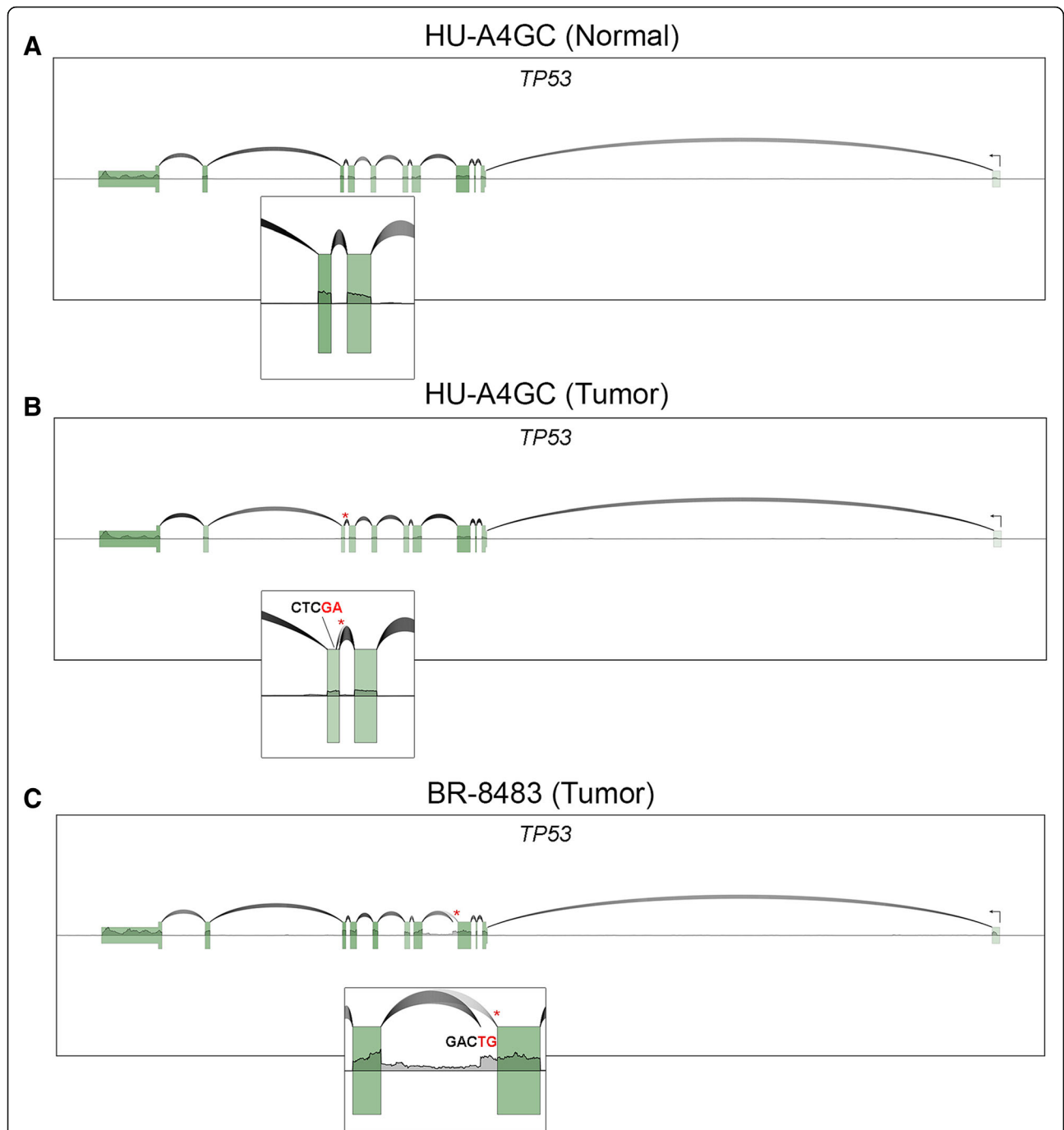
Fig. 2 a SpliceV plot of RNA coverage and splicing from normal stomach tissue (wild type TP53). **b** SpliceV plot of a gastric tumor with a 1 base (T) deletion at a splice acceptor (chr17:7673610, HG38 genome build), disrupting the splicing from exon 8–9 and causing the utilization of a novel cryptic downstream splice acceptor at position chr17:7673590; resulting in a frameshift deletion. **c** SpliceV plot of a gastric tumor with a G > A splice donor variant at chr17:7675993. Part of the intron is retained (chr17:7675884–7,675,993) and a novel intronic splice donor site is utilized, with the same upstream acceptor. This introduces a frameshifting insertion into the protein coding sequence. Asterisks indicate the SNV location and insets are enlarged representations of the transcript structure. Nucleotide sequences at the cryptic splice sites are labeled, with the junctions occurring between the red and black bases in each figure. These samples were initially provided by The Cancer Genome Atlas [2], with alignments obtained from the Genomic Data Commons [9]

coverage, there is an evident drop in single nucleotide level coverage soon after the splice acceptor site, likely illustrating the utilization of an upstream poly(A) site (3′ UTR shortening [21]). Therefore, while exon level coverage provides illustrative qualities for some more macroscopic analyses (e.g. enriched exon coverage of circularized exons (Fig. 1b in RNase R-seq data)), single nucleotide coverage provides granularity when needed.

The need that initially inspired us to develop SpliceV was the lack of available software to plot back-splicing in the context of coverage and forward splicing (for example, see Fig. 1a). This is not only useful for simple presentation of circRNA splicing information, but can also aid interpretation. For example, the display of forward splicing and coverage from poly(A)-seq data in the context of back-splicing data from RNase R-seq data for the GSE1 gene provides evidence of circle formation of exon 2 which precludes its inclusion in the cognate linear GSE1 isoform (Fig. 1c). In this case, exon 2 exclusion introduces a frameshift, ablating the canonical function of this gene.

To add utility to SpliceV in transcript biogenesis and isoform function analyses, we also incorporated the display of RNA binding protein predictions (Fig. 1b) based on empirically determined binding motifs (*Ray* et al) and user supplied *ALU* element sites (Fig. 1c). These features can assist the user in assessing the mechanisms of forward splicing, back splicing, alternative splicing, intron retention, etc. Further, since loaded RNA binding proteins control transcript localization as well as activity, these features can help assist the user in investigating transcript function.

To further illustrate the utility of SpliceV in investigational efforts, we next used SpliceV to visualize isoform level expression in two Gastric Carcinomas and one normal gastric tissue sample from The Cancer Genome Atlas (TCGA) [2]. Whole Exome Sequencing variant calls revealed that each of the two tumor samples had unique splice site mutations in the critical tumor suppressor gene, TP53 [17]. The Genomic Data Commons pipeline [9] for gene expression quantification revealed a slight increase in TP53 RNA levels in the tumor samples. Because the mutations in both tumors occurred in intronic regions, the impact on protein output is not easily determined. Using SpliceV to visualize RNA-seq data (Fig. 2), however, revealed likely haplotypic ablation of the mutated splice acceptor (Fig. 2b) or donor (Fig. 2c) site in these two samples. This led to the utilization of cryptic splice sites that produced frameshifts in each of the resulting transcripts. Also evident in sample BR-8483, based on the single nucleotide coverage line graph, is extensive intron retention, likely causing the resulting intron retained transcript to be subjected to non-sense mediated RNA decay. In both of these cases, SpliceV was able to assist in determining the negative impact of these two mutations on TP53 function, findings that are otherwise opaque to the user.

Conclusions

Here we present a new tool, SpliceV, that facilitates investigations into transcript biogenesis, isoform function and the generation of publication quality figures for the RNA biologist. SpliceV is fast (taking full advantage of the random access nature of BAM files), customizable (allowing users to control plotting aesthetics), and can filter data and make cross-sample comparisons. It is modular in structure, allowing for the inclusion of new features in future package releases. SpliceV should provide value to the toolkit of investigators studying RNA biology and function and should speed the time frame from data acquisition, data analysis to publication of results.

Availability and requirements

Project name: SpliceV

Project home page: https://github.com/flemingtonlab/SpliceV

Operating system: Platform independent

Programming language: Python

Other requirements: Python 2.7 or Python 3.0+

License: GNU Public License

Any restrictions to use by non-academics: License needed

Additional files

Additional file 1: Command line parameters for SpliceV. (PDF 251 kb)

Additional file 2: An example pipeline to generate the required files for SpliceV analysis. (PDF 148 kb)

References

1. Amara SG, Jonas V, Rosenfeld MG, Ong ES, Evans RM. Alternative RNA processing in calcitonin gene expression generates MRNAs encoding different polypeptide products. Nature. 1982;298(5871):240–4. https://doi.org/10.1038/298240a0.
2. Bass AJ, Thorsson V, Shmulevich I, Reynolds SM, Miller M, Bernard B, Hinoue T, et al. Comprehensive molecular characterization of gastric adenocarcinoma. Nature. 2014;513(7517):202–9. https://doi.org/10.1038/nature13480.
3. Bentley DL. Coupling MRNA processing with transcription in time and space. Nat Rev Genet. 2014;15(3):163–75 https://doi.org/10.1038/nrg3662.
4. Boise LH, González-García M, Postema CE, Ding L, Lindsten T, Turka LA, Mao X, Nuñez G, Thompson CB. Bcl-x, a Bcl-2-related gene that functions as a dominant regulator of apoptotic cell death. Cell. 1993;74(4):597–608. https://doi.org/10.1016/0092-8674(93)90508-N.
5. Cartegni L, Wang J, Zhu Z, Zhang MQ, Krainer AR. ESEfinder: a web resource to identify Exonic splicing enhancers. Nucleic Acids Res. 2003;31(13):3568–71. https://doi.org/10.1093/nar/gkg616.
6. Chen S, Huang V, Xu X, Livingstone J, Soares F, Jeon J, Zeng Y, et al. Widespread and functional RNA circularization in localized prostate Cancer. Cell. 2019;176(4):831–843.e22 https://doi.org/10.1016/j.cell.2019.01.025.
7. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. STAR: ultrafast universal RNA-Seq aligner. Bioinformatics. 2013;29(1):15–21. https://doi.org/10.1093/bioinformatics/bts635.
8. Gao Y, Wang J, Zhao F. CIRI: an efficient and unbiased algorithm for de novo circular RNA identification. Genome Biol. 2015;16(1):4. https://doi.org/10.1186/s13059-014-0571-3.
9. Grossman RL, Heath AP, Ferretti V, Varmus HE, Lowy DR, Kibbe WA, Staudt LM. Toward a shared vision for Cancer genomic data. N Engl J Med. 2016; 375(12):1109–12 https://doi.org/10.1056/NEJMp1607591.
10. Hammes A, Guo J-K, Lutsch G, Leheste J-R, Landrock D, Ziegler U, Gubler M-C, Schedl A. Two splice variants of the Wilms' tumor 1 gene have distinct functions during sex determination and nephron formation. Cell. 2001; 106(3):319–29 https://doi.org/10.1016/S0092-8674(01)00453-6.
11. Hansen TB, Jensen TI, Clausen BH, Bramsen JB, Finsen B, Damgaard CK, Kjems J. Natural RNA circles function as efficient MicroRNA sponges. Nature. 2013;495(7441):384–8 https://doi.org/10.1038/nature11993.
12. Ivanov A, Memczak S, Wyler E, Francesca Torti HT, Porath MR, Orejuela MP, et al. Analysis of intron sequences reveals hallmarks of circular RNA biogenesis in animals. Cell Rep. 2015;10(2):170–7. https://doi.org/10.1016/J.CELREP.2014.12.019.
13. Jeck WR, Sorrentino JA, Wang K, Slevin MK, Burd CE, Liu J, Marzluff WF, Sharpless NE. Circular RNAs are abundant, conserved, and associated with ALU repeats. RNA (New York, NY). 2013;19(2):141–57. https://doi.org/10.1261/rna.035667.112.
14. Krainer AR, Conway GC, Kozak D. The essential pre-MRNA splicing factor SF2 influences 5' splice site selection by activating proximal sites. Cell. 1990; 62(1):35–42 https://doi.org/10.1016/0092-8674(90)90237-9.
15. Lasda, Erika, and Roy Parker. 2016. "Circular RNAs co-precipitate with extracellular vesicles: a possible mechanism for CircRNA clearance." edited by Pierre Busson. PLoS One 11 (2): e0148407. https://doi.org/10.1371/journal.pone.0148407.
16. Legnini I, Di Timoteo G, Rossi F, Morlando M, Briganti F, Sthandier O, Fatica A, et al. Circ-ZNF609 is a circular RNA that can be translated and functions

17. in Myogenesis. Mol Cell. 2017;66(1):22–37.e9. https://doi.org/10.1016/j.molcel.2017.02.017.
17. Levine AJ, Momand J, Finlay CA. The P53 tumour suppressor gene. Nature. 1991;351(6326):453–6 https://doi.org/10.1038/351453a0.
18. Li Y, Bor Y-c, Misawa Y, Xue Y, Rekosh D, Hammarskjöld M-L. An intron with a constitutive transport element is retained in a tap messenger RNA. Nature. 2006;443(7108):234–7 https://doi.org/10.1038/nature05107.
19. Li Z, Huang C, Bao C, Chen L, Lin M, Wang X, Zhong G, et al. Exon-intron circular RNAs regulate transcription in the nucleus. Nat Struct Mol Biol. 2015; 22(3):256–64 https://doi.org/10.1038/nsmb.2959.
20. Marchionni MA, Goodearl ADJ, Chen MS, Bermingham-McDonogh O, Kirk C, Hendricks M, Danehy F, et al. Glial growth factors are alternatively spliced ErbB2 ligands expressed in the nervous system. Nature. 1993;362(6418):312–8 https://doi.org/10.1038/362312a0.
21. Mayr C, Bartel DP. Widespread shortening of 3′UTRs by alternative cleavage and polyadenylation activates oncogenes in Cancer cells. Cell. 2009;138(4): 673–84 https://doi.org/10.1016/J.CELL.2009.06.016.
22. Memczak S, Jens M, Elefsinioti A, Torti F, Krueger J, Rybak A, Maier L, et al. Circular RNAs are a large class of animal RNAs with regulatory potency. Nature. 2013;495(7441):333–8 https://doi.org/10.1038/nature11928.
23. O'Grady T, Cao S, Strong MJ, Concha M, Wang X, Splinter S, BonDurant MA, et al. Global bidirectional transcription of the Epstein-Barr virus genome during reactivation. J Virol. 2014;88(3):1604–16. https://doi.org/10.1128/JVI.02989-13.
24. Pamudurti NR, Bartok O, Jens M, Ashwal-Fluss R, Stottmeister C, Ruhe L, Hanan M, et al. Translation of CircRNAs. Mol Cell. 2017;66(1):9–21.e7. https://doi.org/10.1016/j.molcel.2017.02.021.
25. Pan Q, Shai O, Lee LJ, Frey BJ, Blencowe BJ. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. Nat Genet. 2008;40(12):1413–5 https://doi.org/10.1038/ng.259.
26. Rauscher F, Morris J, Tournay O, Cook D, Curran T, Hastie ND. Binding of the Wilms' tumor locus zinc finger protein to the EGR-1 consensus sequence. Science. 1990;250(4985):1259–62 https://doi.org/10.1126/science.2244209.
27. Ray D, Hilal Kazan KB, Cook MT, Weirauch HS, Najafabadi XL, Gueroussov S, et al. A compendium of RNA-binding motifs for decoding gene regulation. Nature. 2013;499(7457):172–7 https://doi.org/10.1038/nature12311.
28. Reed R, Maniatis T. A role for exon sequences and splice-site proximity in splice-site selection. Cell. 1986;46(5):681–90. https://doi.org/10.1016/0092-8674(86)90343-0.
29. Rosenfeld MG, Mermod J-J, Amara SG, Swanson LW, Sawchenko PE, Rivier J, Vale WW, Evans RM. Production of a novel neuropeptide encoded by the calcitonin gene via tissue-specific RNA processing. Nature. 1983;304(5922): 129–35 https://doi.org/10.1038/304129a0.
30. Rybak-Wolf A, Stottmeister C, Glažar P, Jens M, Pino N, Giusti S, Hanan M, et al. Circular RNAs in the mammalian brain are highly abundant, conserved, and dynamically expressed. Mol Cell. 2015;58(5):870–85. https://doi.org/10.1016/j.molcel.2015.03.027.
31. Salzman, Julia, Charles Gawad, Peter Lincoln Wang, Norman Lacayo, and Patrick O Brown. 2012. "Circular RNAs are the predominant transcript isoform from hundreds of human genes in diverse cell types." edited by Thomas Preiss. PLoS One 7 (2): e30733. https://doi.org/10.1371/journal.pone.0030733.
32. Singh R, Valcárcel J, Green MR. Distinct binding specificities and functions of higher eukaryotic Polypyrimidine tract-binding proteins. Science (New York, NY). 1995;268(5214):1173–6 https://doi.org/10.1126/SCIENCE.7761834.
33. Solnick D. Alternative splicing caused by RNA secondary structure. Cell. 1985;43(3):667–76 https://doi.org/10.1016/0092-8674(85)90239-9.
34. Starke S, Jost I, Rossbach O, Schneider T, Schreiner S, Hung L-H, Bindereif A. Exon circularization requires canonical splice signals. Cell Rep. 2015;10(1): 103–11 https://doi.org/10.1016/J.CELREP.2014.12.002.
35. Ungerleider, Nathan, Monica Concha, Zhen Lin, Claire Roberts, Xia Wang, Subing Cao, Melody Baddoo, et al. 2018. "The Epstein Barr virus CircRNAome." edited by Bryan R. Cullen. PLoS Pathog 14 (8): e1007206. https://doi.org/10.1371/journal.ppat.1007206.
36. van der Walt S, Chris Colbert S, Varoquaux G. The NumPy Array: a structure for efficient numerical computation. Computing in Science & Engineering. 2011;13(2):22–30 https://doi.org/10.1109/MCSE.2011.37.
37. Wahl MC, Will CL, Lührmann R. The spliceosome: design principles of a dynamic RNP machine. Cell. 2009;136(4):701–18 https://doi.org/10.1016/J.CELL.2009.02.009.
38. Wang ET, Sandberg R, Luo S, Khrebtukova I, Zhang L, Mayr C, Kingsmore SF, Schroth GP, Burge CB. Alternative isoform regulation in human tissue

transcriptomes. Nature. 2008;456(7221):470–6. https://doi.org/10.1038/nature07509.

39. Wang Peter L, Yun Bao, Muh-Ching Yee, Steven P. Barrett, Gregory J. Hogan, Mari N. Olsen, José R. Dinneny, Patrick O. Brown, and Julia Salzman. 2014. "Circular RNA is expressed across the eukaryotic tree of life." edited by Thomas Preiss. PLoS One 9 (3): e90859. https://doi.org/10.1371/journal.pone.0090859.

40. 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. https://doi.org/10.1016/J.CELREP.2014.10.062.

41. 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. https://doi.org/10.1016/J.CELL.2014.09.001.

42. Zhang Y, Xue W, Li X, Zhang J, Chen S, Zhang J-L, Yang L, Chen L-L. The biogenesis of nascent circular RNAs. Cell Rep. 2016;15(3):611–24. https://doi.org/10.1016/j.celrep.2016.03.058.

43. Zhuang Y, Weiner AM. A Compensatory Base change in human U2 SnRNA can suppress a branch site mutation. Genes Dev. 1989;3(10):1545–52. https://doi.org/10.1101/GAD.3.10.1545.