Computational approaches and challenges in the analysis of circRNA data

Barry Digby [1]*[ID]**, Stephen P. Finn** [2][ID] **and Pilib ÓBroin** [1][ID]

1   School of Mathematical and Statistical Sciences, University of Galway, Galway, Ireland
2   Department of Histopathology and Morbid Anatomy, Trinity Translational Medicine Institute, Dublin, Ireland
*   Correspondence: b.digby1@universityofgalway.ie

Abstract: Circular RNAs (circRNA) are a class of non-coding RNA, forming a single-stranded covalently closed loop structure generated via back-splicing. Advancements in sequencing methods & technologies in conjunction with algorithmic developments of bioinformatics tools have enabled researchers to characterise the origin and function of circRNAs, with practical applications as a biomarker of diseases becoming increasingly relevant. Computational methods developed for circRNA analysis are predicated on detecting the chimeric back-splice junction of circRNAs whilst mitigating false-positive sequencing artefacts. In this review, we discuss in detail computational strategies developed for circRNA identification, highlighting a selection of tools strengths, weaknesses and blind-spots. In addition to circRNA identification tools, we describe methods for characterising circRNAs role within the cometing endogenous RNA (ceRNA) network, their interactions RNA-binding proteins and publicly available databases for rich circRNA annotation.

Keywords: circRNAs; bioinformatics tools; computational challenges; non-coding RNA; next generation sequencing

1. Introduction

Originally detected in plant viroids [1], yeast mitochondrial RNA [2] and hepatitis δ virus [3], circRNAs were believed to be discarded intermediates from intron lariat branching or by-products of aberrant splicing [4,5]. During the advent of next-generation sequencing (NGS) technologies, circRNAs went largely unstudied – their detection hindered by the design of microarray oligonucleotide probes and poly-A selection protocols in RNA-Sequencing (RNA-Seq) technologies preferentially selecting messenger RNA (mRNA) sequences. Recent advancements in bioinformatics methods coupled with a shift in the paradigm of RNA-Seq protocols have enabled the detection of circRNAs, with interest in the field rejuvenated when a landmark study by Salzman et al. (2012) identified RNA transcripts containing 'scrambled exons' characteristic of circRNAs in hyperdiploid acute lymphoblastic leukaemia diagnostic bone marrow samples [6]. Subsequent studies by Jeck et al. (2012) [7], Hansen et al. (2013) [8] and Memczak et al. (2013) [9] identified thousands of circRNAs in metazoans. Moreover, Hansen et al., (2013) [8] and Memczak et al. (2013) [9] demonstrated CDR1as and circSry competitively bind micro RNA (miRNA) sequences via complimentary sites within their mature spliced sequence, suggesting a regulatory role within the ceRNA network. These foundational works ushered in a plethora of novel research on circRNAs characterising their origin, biogenesis, structure and functions [10–13].

circRNAs exhibit stage and tissue-specific expression [14–16] and are enriched in exosomes [17]. Coupled with their elevated stability in contrast to other RNA molecules, circRNAs represent a promising biomarker for diseases; circ-ITCH acts as a tumor supressor in lung cancer by inhibiting the Wnt/β-Catenin pathway [18], circPVT1 acts as an oncogene in head and neck squamous cell carcinoma displaying overexpression in tumor samples harbouring TP53 mutations [19]. In addition to applications as a biomarker, circRNAs

can be constructed to target and sequester overexpressed oncogenic miRNAs. The synthetically generated circRNA *CM21D* was produced via t-RNA splicing and administered in experimental glioblastoma models, inhibiting miR-21-5p thus restoring the expression of tumor suppressor genes [20]. Given their diverse role within cells, it is imperative to accurately identify and annotate the functions of circRNAs using computational methods in conjunction with sequencing data.

Multiple bioinformatics tools exist for identifying circRNAs in RNA-Seq datasets via the detection of chimeric reads representative of circRNA back-splice junctions (BSJ). Once a circRNA has been identified, researchers are often interested in quantifying its expression, predicting interactions with other small RNAs and RBPs, examining the ratio of circRNA expression vs. its cognate parent gene and performing differential circRNA expression analyses. In this review, we discuss the current landscape of bioinformatics tools for circRNA analyses encompassing circRNA identification and annotation, circRNA quantification, circRNA functional prediction and the computational challenges presented therein. Whilst the remit of this review focuses on the computational analysis of circRNAs, we briefly detail the origin, biogenesis and structure of circRNAs as this information is necessary for understanding the algorithms employed by circRNA identification tools. For full reviews of circRNAs biogenesis, degradation, translation, cellular transport, downstream interactions and evolutionary conservation, we point readers to a selection of recent reviews [21–28]. Finally, we provide an overview of the currently available circRNA databases containing rich annotations for circRNAs derived from various tissue sources and cell lines using RNA-Seq datasets, predicted miRNA and RBP targets using CLIP-Seq and circRNAs associated with diseases.

2. circRNA biogenesis and structure

Canonical linear mRNA splicing involves processing precursor mRNA (pre-mRNA) to remove intronic sequences and the joining together of exon sequences to form a mature mRNA transcript. This process is mediated via the spliceosomal machinery composed of small nuclear ribonucleoproteins which recognise conserved 5' splice sites, 3' splice sites and a branch site within the intronic sequence. Spliceosomal machinery binds to the intron's upstream 5' splice donor site, pairing it with the downstream branch site forming a lariat loop structure. Following this, the downstream 3' splice acceptor site of the intron splice site is brought in close proximity to the 3' end of the exon where via a process of transesterification, the exons 3' hydroxyl group attacks the phosphodiester bond of the 3' intron splice acceptor site, covalently joining the exons producing a mature mRNA, releasing the intron lariat structure [29].

circRNA formation relies on canonical splice site signals and spliceosome machinery [30] however, in contrast to linear RNAs, circRNAs are formed by a process known as back-splicing in which a downstream 5' splice donor site is reversely joined to an upstream 3' splice acceptor site forming a covalently closed loop structure [7,11]. This circularization process is mediated by one of three factors: complimentary intron pairing, RNA binding proteins (RBP) and lariat formation (Figure1). 1) Intron pairing: RNA pairing across flanking introns in pre-mRNA can occur due to the complementarity of short non-coding *cis* regulatory elements such as repetitive Alu sequences or GU-rich sequences forming a hairpin structure that brings the downstream 5' splice acceptor site in close proximity to the upstream 3' acceptor site facilitating back-splicing and the formation of exonic circRNAs (EcircRNAs) and exon-intron circRNAs (EIcircRNAs) [31,32]. 2) RBP pairing: *trans*-factors in pre-mRNA bind to RBP such as Muscleblind (MBNL1) and Quaking (QKI) which in turn bridge flanking introns facilitating back-splicing and the generation of EcircRNAs and EIcircRNAs [13,33]. 3) Lariat formation: during canonical splicing, an exon-skipping event produces an intermediate lariat containing excised exons and introns. Subsequent internal lariat splicing produces one of a single EcircRNAs [34] or intronic circRNAs (ciRNAs) driven by base-pairing of splice-site motifs in exonic and intronic sequences, respectively. [35]. The three categories of circRNAs classes (EcircRNA, EIcircRNA and ciRNA) present
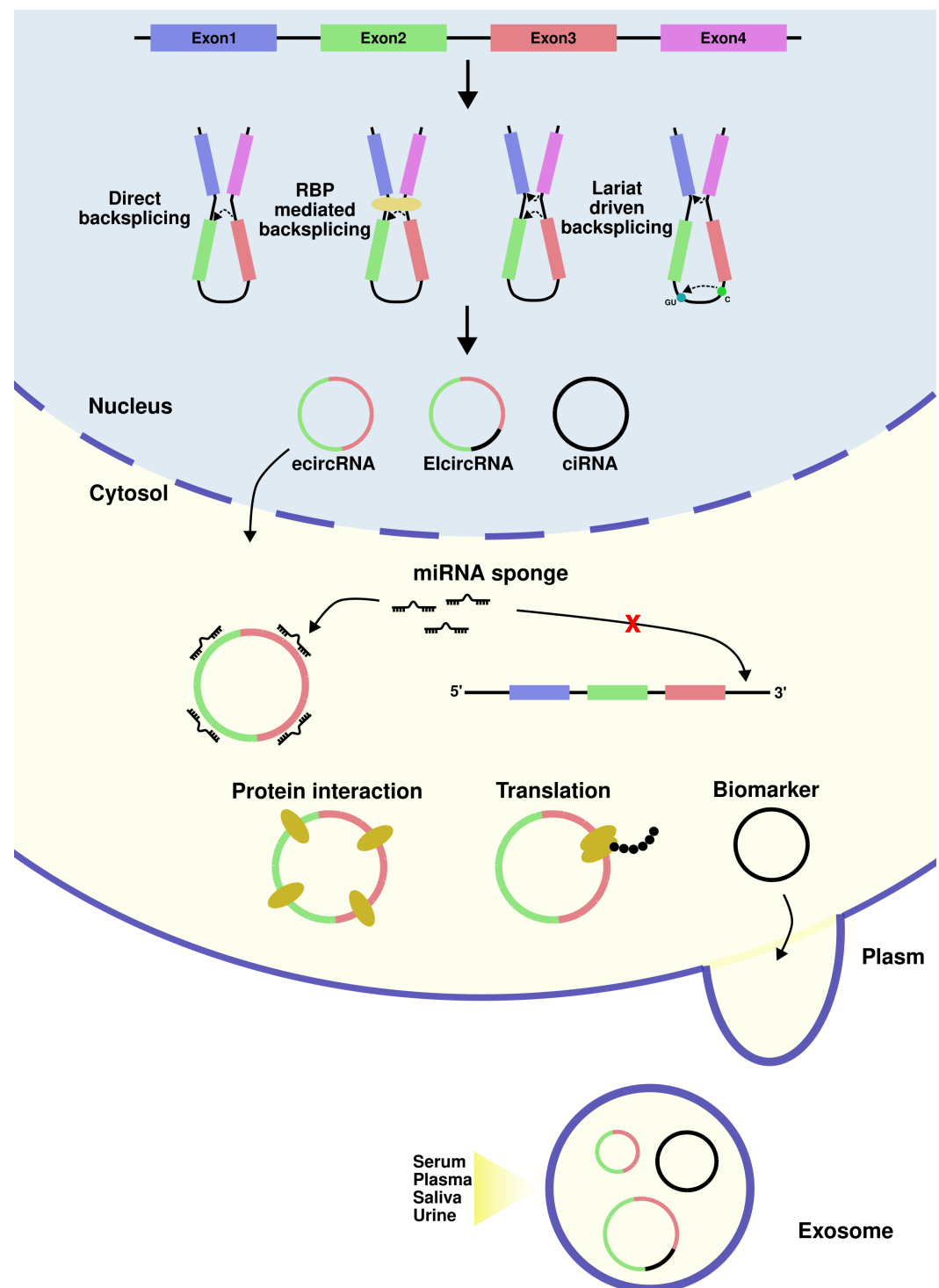
Figure 1. Biogenesis and fucntional mechanisms of circRNAs. **Nucleus**: circRNA biogenesis via direct back-splicing driven by complementary intronic sequences, RNA-binding proteins or lariat loops containing skipped exons or introns. **Cytosol**: Cap independent protein translation, RBP binding and miRNA sequestration via complementary miRNA response element sites. **Plasm**: exosomal circRNAs facilitating non-invasive biomarker discovery.

diverse internal structures in the final mature spliced sequence: 1) EcircRNAs containing all underlying exons after canonical splicing; 2) alternative EcircRNAs comprised of a subset of underlying exons as a by-product of alternative-splicing; 3) EIcircRNAs composed of both intronic and exonic sequences; and 4) ciRNAs containing intronic sequences.

circRNAs unique covalently closed loop structure lacking 5' and 3' tails confers resistance to RNase R degradation, granting them a much higher half-life than their linear mRNA counterparts [36,37]. This feature makes circRNAs an attractive biomarker in disease-based settings, with reports of circRNAs exhibiting differential expression in gastric, colorectal, ovarian and lung cancers, and enzalutamide resistant LNCaP cell lines [17,38,39]. Furthermore, circRNAs can be packaged and exported from the cell to bodily fluid via exosomes [40,41] facilitating the use of non-invasive liquid biopsies to monitor disease progression [42–47]. The mechanism of circRNA degradation and clearance remains an active area of research. Studies have found that miRNAs can facilitate the degradation of circRNAs via Argonaute 2 (Ago2)-mediated degradation supporting the hypothesis of circRNAs role within the ceRNA network [48]. Other works demonstrate RNase H1-mediated degradation of ciRNAs with high GC content [49] and the degradation of circRNAs containing m^6A modifications via endoribonucleolytic cleavage [50].

3. Principles and challenges for circRNA identification

3.1. Library preparation

circRNAs represent a mere 1-3% of the transcriptome pool in total RNA sequencing libraries [51], dictating novel strategies to enrich circRNA libraries prior to sequencing (Figure2). A typical eukaryotic RNA-Seq library preparation protocol involves the preferential selection of RNAs with polyA tails and/or the depletion of ribosomal RNAs (rRNAs). Due to circRNAs covalently closed loop structure, polyA selection in libraries will almost completely remove all circular transcripts in a sample. By contrast, circRNAs are retained in rRNA-depleted samples and are enriched in samples treated with ribonuclease R (RNase R) to deplete linear RNAs. Random priming is preferred to oligo(dT) priming, as the former method generates random oligonucleotide sequences for cDNA synthesis independent of polyA sequences, whilst the latter will generate libraries biased towards linear RNAs. One method, termed "RNase R treatment followed by polyadenylation and polyA$^+$ RNA depletion" (RPAD) has emerged as a leading candidate for circRNA library preparation yielding the highest number of circRNAs and the highest sensitivity in a benchmark study [52]. RPAD employs the sequential depletion of linear RNAs via RNase R treatment, polyadenylation of remaining linear RNAs and a final round of polyA$^+$ depletion using oligo(dT) beads followed by ribosomal RNA (rRNA) depletion to yield a high concentration circRNA library for sequencing [53,54]. In the absence of the RPAD method, rRNA depletion and/or RNase R$^+$ are sufficient for generating RNA-Seq datasets for circRNA detection and have been used in benchmark studies analysing the performance of circRNA identification tools [55].

3.2. Sequencing artefacts

Technical artefacts introduced during sequencing can lead to the generation of false positives during circRNA identification. Reverse transcriptase, an enzyme used to synthesise complementary cDNA strands can undergo a template-switching event when brought in close proximity to a different RNA template with a suitable region for priming [56]. The original, incomplete synthesized strand is carried to the newly 'switched' cDNA template where reverse transcriptase continues generating a chimeric molecule capable of mimicking alternative splicing and backsplicing. Alarmingly, these template switching events can account for up to 35-55% of the isoforms computationally detected for a gene [57]. Sequencing libraries that use adapter-ligation steps are at risk of generating chimeric sequences – albeit at a much lower level. Finally, incorrectly called bases at the beginning or end of exons in genes containing highly homologous sequences can generate false positive splice site signals (GT/AG, etc.) [58]. With respect to circRNAs, these sequencing errors can lead
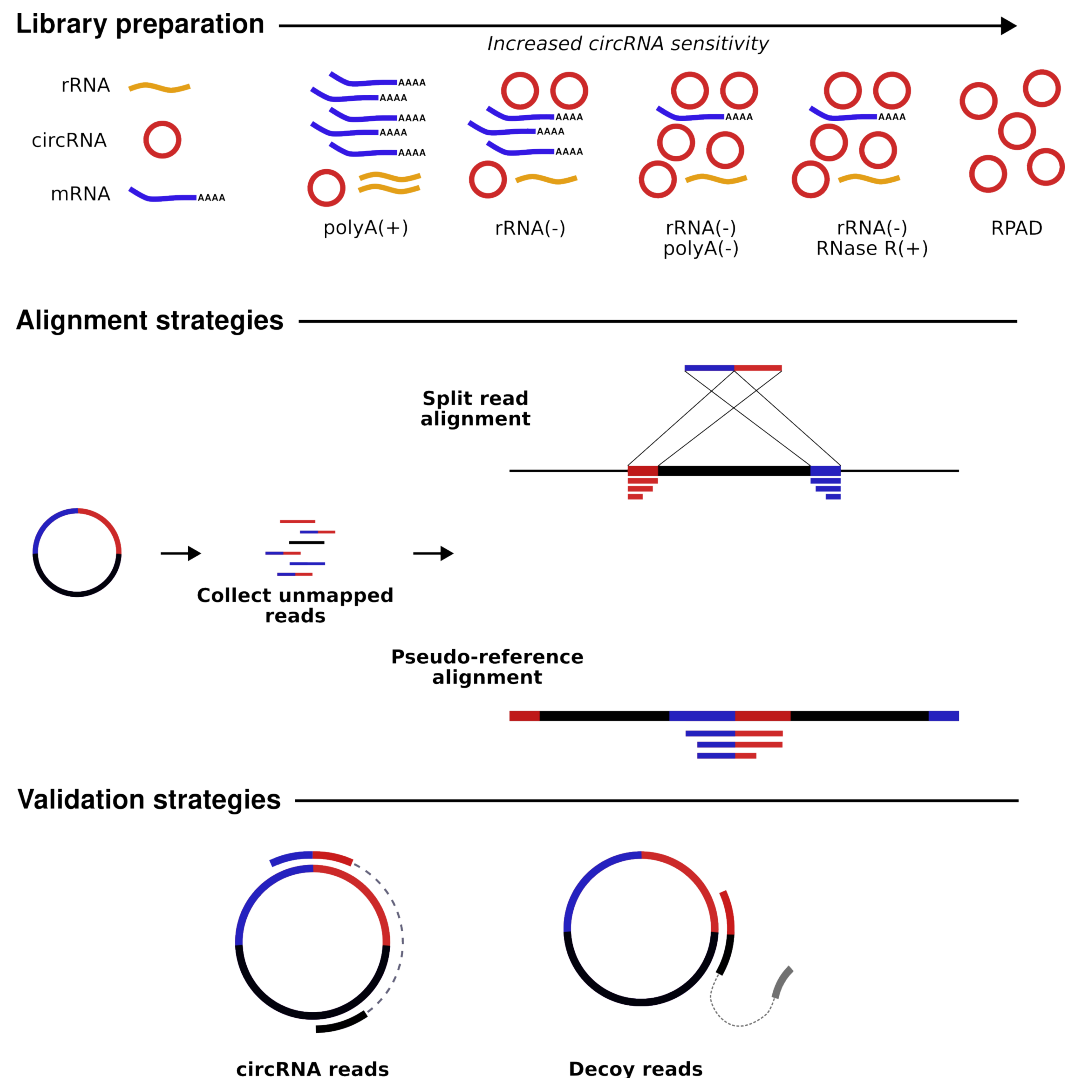
Figure 2. Advancements in biochemical and bioinformatics strategies for circRNA detection. **Library preparation**: Left to right, in order of increasing circRNA sensitivity; polyA(+): unsuitable for circRNA detection, preferentially selects mRNAs; rRNA(-): ribosomal RNA depletion yielding a library with circRNAs and mRNAs; rRNA(-) & polyA(-): ribosomal RNA depletion in conjunction with polyadenylation and depletion of polyA transcripts; rRNA(-) & RNase R(+): ribosomal RNA depletion in conjunction with RNase R ribonuclease treatment depleting mRNAs; RPAD: RNase R treatment followed by polyadenylation and polyA^{R}NA depletion, yielding a highly pure circRNA library. **Alignment strategies**: Unmapped reads to the reference genome can be used to perform split alignment to identify reads aligning in opposite directions, or to map reads to a BSJ pseudo-reference. **Validation strategies**: Paired-end sequencing greatly reduces false-positives, mates that map across the BSJ whose mates reside within the same transcript are indictive of circRNAs.

to sites that are mistakenly identified as backsplice sites when identifying circRNAs in samples. Due to the low levels of circRNA expression in cells when compared to other RNA transcripts, the presence of sequencing artefacts cannot be overcome by simply applying a read depth filter to quantification results. circRNA identification tools typically require paired-end data to overcome this source of error by requiring read 1 to map to the back-splice junction (BSJ) and its corresponding read 2 pair to map in the same transcript within a fixed distance [6].

3.3. BSJ-based circRNA identification

The main step of any circRNA analysis is the identification of circRNAs in RNA-Seq datasets. This analysis is predicated on the detection of the BSJ, i.e the scrambled exon junctions representing the joining of an upstream 5' donor site to the downstream 3' acceptor site to form a circular structure. The majority of circRNA identification tools can be classed as one of two sub-groups: segmented-based-approach, whereby anchors (fixed-length segments taken from the end of reads) are extracted from unmapped sequencing reads and re-mapped to the genome; or pseudo-reference based in which a custom database of manually curated BSJ sites are generated and used to map the sequencing reads (Figure 2). The former strategy allows for *de novo* circRNA identification whilst the latter is constrained to exons contained within the reference annotation file used for constructing the pseudo-reference database. As circRNA identification tools evolved, developers began to blend the two approaches to optimize the process of circRNA identification.

The first circRNA identification analysis performed by Salzman et al. (2012) [6] used the pseudo-reference based approach to identify circRNAs in ALL samples. Reads that mapped contiguously to RefSeq annotated genes using Bowtie were considered representative of linear transcripts and removed from the analysis. Subsequently, the RefSeq database was used to create a custom database of all intragenic exon-exon junctions against which reads that failed to align were mapped. An exon scrambling event was flagged if read 1 mapped to a non-canonical exon-exon junction as defined by the custom RefSeq database and read 2 mapped within the same transcript. The number of reads spanning the scrambled exon junction was used to estimate the relative abundance of candidate circRNAs. In contrast to the pseudo-reference based approach, the first tool created for the purpose of circRNA detection 'find_circ' [9] utilises the segmented-based approach. Firstly, paired-end reads are aligned to the genome to extract reads that do not contiguously align. A customised script then splits unmapped reads to obtain 20 nucleotide anchor sequences originating from the 5' and 3' ends of the reads. The anchors are re-aligned to the genome, with anchors mapping in reverse orientation extended to identify the breakpoint site in the anchor. The resulting BED file is filtered to arrive at a set of circRNAs: 1) splice sites must be flanked by GU/AG signals; 2) unambivalant breakpoints; 3) < 2 mismatches in the extension procedure; 4) breakpoint cannot reside more than 2 nucleotides inside the anchor; 5) > 2 reads must support the BSJ site and 6) splice sites must be within 100Kb.

UROBORUS [59] adopts a similar approach to 'find_circ', collecting and extracting 20bp anchors from reads that failed to map contiguously to the reference genome using TopHat. Anchor segments representative of a circRNA BSJ site mapped in reverse orientation within the same transcript with an overhang of >20bp at either end of both segments are termed 'balanced mapped junctions' (BMJ) whilst segments with an overhang <20bp in one read are termed 'unbalanced mapped junctions' (UMJ) and subject to different extension strategies. Both BMJ reads are separately extended outwards to the nearest splice-site to form paired-end segments, whilst for UMJ reads, the shorter mate is discarded and the single mapped seed is outwardly extended to the nearest splice-site (an extension distance no longer than the length of the single read - 3bp). Bowtie is then used to remap the paired-end and single extended segments; segments aligning to the reference genome in the opposite orientation with read support >2 as detected by the UROBORUS algorithm are representative of circular candidates. Post-transcriptional exon shuffling finder (PTESfinder) [60] combines the segment-based approach with the pseudo-reference based

approach to identify circular candidates using Bowtie. Briefly, 20bp segment anchors are extracted from the ends of input sequencing reads and mapped to the reference transcriptome. Anchors that map to the same gene but in an inverted orientation are identified and used to construct a pseudo-reference (termed 'PTES constructs') by concatenating the last 65bp of the underlying 5' exon and the first 65bp of the 3' exon. Reads are then aligned to the PTES construct, genomic and transcriptomic references in order to generate mapping scores for circular candidate filtering. Candidates are marked as circRNAs when they exhibit high mapping scores to the PTES construct and low scores to the genomic/transcriptome reference.

The concept of mapping reads to genomic, transcriptome and BSJ databases to filter circRNA candidates was further improved upon by KNIFE [61]. KNIFE maps reads to rRNA sequences, genomic, transcriptome and a customised BSJ reference database using Bowtie2, discarding circular candidates that map with high scores to one of the other databases listed above. Where paired-end RNA-Seq data is available, the candidate reads spanning the BSJ site are subset into circRNA and decoy reads based on mapping information available to mitigate false-positive BSJ reads generated by sequencing errors. For reads that fail to map to any of the databases, a *de novo* analysis is performed using Bowtie coupled with a segment-based approach whereby segments are used to construct a *de novo* index. The unmapped reads are then re-aligned to the *de novo* index using Bowtie2 with the same criteria for pseudo-reference based alignment. KNIFE is one of the first circRNA identification tools to employ a statistical framework by obtaining a posterior probability for each circRNA candidate to predict if it is a true positive using a logistic generalized linear model (GLM) based on the alignment features of read 1. In contrast to the circRNA identification tools discussed thus far which require extracting anchor sequences to identify putative BSJ sites using Bowtie/Bowtie2, BWA and STAR are capable of directly detecting breakpoint events and chimeric fusions during read alignment, respectively. circRNA identification tools utilizing BWA/STAR therefore circumventing the need to manually extract anchors for BSJ identification using customized scripts – streamlining the process of circRNA identification and reducing computational overheads.

CircRNA Identifier (CIRI) [62] is one such tool that utilises BWA-MEM sequence alignment mapping (SAM) information to identify reads in which two segments of the read align in chiastic order termed 'paired chiastic clipping' (PCC) signals. Subsequent filtering leveraging paired-end mapping (PEM) information, GT-AG splice signals and mapping rates to homologous sequences removes false positives to arrive at a set of high-confidence circRNAs. One shortcoming of CIRI is its handling of unbalanced junction reads. Unbalanced junction reads are segments of length <19bp which are ignored by BWA-MEM to prevent multi-mapping/erroneous mapping, therefore lacking the necessary alignment information in the SAM file for CIRI to detect PCC signals. CIRI uses a dynamic programming algorithm to re-map unbalanced junction reads to balanced junction reads originating from the same junctions detected in the first alignment step. This step is computationally expensive and leads to the generation of false positives, an area specifically addressed by the tools successor. CIRI2 [63] is more cautious when addressing unbalanced junction reads and balanced junction reads with low mapping quality by utilising a maximum likelihood estimation (MLE) based on multiple seed matching. The undetermined segment of a putative BSJ read is divided into n seeds of length m (e.g 50bp segment divided into five seeds of length 10bp) to determine if the segment belongs to a forward splice region or a back-splice region based on the matched-seed numbers derived from the two regions which follow a binomial distribution with distinct probabilities. Since its publication in 2018, CIRI2 has become one of the most popular circRNA identification tools and has since been subsumed by CIRIquant [64] which extends its functionality by creating a pseudo-reference based on circular candidates detected by CIRI2, against which candidate reads are re-aligned using HISAT2. In addition to improved alignment, CIRIquant performs RNase R correction, linear RNA quantification and automated differential expression analysis of circRNAs. ACFS [65] is another identification tool that uses BWA however, its approach to circRNA

identification is somewhat unorthodox. ACFS converts paired-end data to single-end data and collapses the reads prior to alignment, borrowing a common strategy used for miRNA alignment and quantification. After identifying candidate reads containing segments mapping in inverse orientation, ACFS uses maximum entropy models to predict the underlying BSJ sequence most likely to be generated by splicing. The advantage of this approach is non-canonical dinucleotide splice sites are considered. The authors also point to the tool's ability to detect fusion circRNAs generated by chromosomal translocation events. This raises the question as to how ACFS controls for sequencing artefacts which can mimic fusion events - particularly when the tool discards paired-end read information.

circRNA_finder [14] and CIRCexplorer [66] were the first tools to use the outputs from the STAR aligner to identify circRNAs. CIRCexplorer offers additional support for TopHat, however since the deprecation of the TopHat we focus on STAR herein. STAR is capable of directly detecting and writing chimeric reads to the output binary alignment map (BAM) file or separately to a junctions.out tab-separated text file when '–chimSegmentMin' is set to a positive integer. Both circRNA_finder and CIRCexplorer take advantage of the lightweight junctions.out file which contain in each line the genomic coordinates and CIGAR flags corresponding to each read segment that comprise the chimeric RNA molecule. circRNA_finder imposes filtering on the putative circRNAs, allowing at most 3 mismatches, uniquely mapped reads, a maximum distance between splice-donor sites of 100kb and that if one read spans the BSJ site, its mate should reside within the interval between the splice donor and acceptor site. Interestingly, CIRCexplorer does not impose such filtering strategies. It instead benefits from using an input reference gene annotation file to annotate putative circRNAs thereby constraining results to exon-exon boundaries contained within the reference file, reducing the rate of false-positives. CIRCexplorer was superseded by CIRCexplorer2 [67], adding a suite of new modules for circRNA identification including alignment using TopHat/TopHat-Fusion, annotation of circRNAs, *de novo* assembly of novel circRNAs, characterising alternative-splicing events within circRNAs and support for parsing BWA, MapSplice, STAR and Segemehl outputs. Similarly to CIRCexplorer, the deprecation of TopHat has resulted in CIRCexplorer2 largely becoming a tool for the downstream parsing and annotation of outputs from BWA, MapSplice, Segemehl and STAR. DCC [55] is yet another circRNA identification tool that harnesses the power of the STAR aligner. In its recommended workflow, paired-end mates are mapped using STAR and each individual mate is processed in the same manner, generating three output files per sample – joint mapping, mate1 and mate2 junctions.tab files. DCC also offers a junction ratio test using CircTest to formally test variation in expression between circRNAs and their parent gene. We have noted that the sensitivity of circRNA identification tools using STAR can be drastically increased by implementing STAR 2-Pass mode, in which the chimeric junctions detected in all samples during the first mapping stage can be collected and incorporated into the reference genome on the fly during the second pass mapping stage for a sample. This method comes at the cost of increased false positives [68] and as such we recommend users adopt an ensemble approach or set suitably strict filtering parameters on detected circRNAs when employing STAR 2-Pass mode with circRNA_finder, CIRCexplorer2 or DCC.

Finally, there exist splice-aware aligners that are capable of directly handling unmapped reads for detecting circRNAs during the alignment step. Non-co-linear scan (NCLscan) [69] and segemehl [70] are two popular tools for this task however, as NCLscan uses the proprietary aligner Novoalign, it's use is dependent on an active Novocraft membership. For this reason segemehl is considered the more popular tool in academic circles and has been incorporated into CIRCexplorer2 and intergration-based tools.

3.4. Integration-based identification methods

A study by Hansen et al. (2016) [71] highlighted the discrepancies in results generated by the most popular circRNA identification tools at the time (circRNA_finder, CIRCexplorer, CIRI, find_circ and MapSplice). Strikingly, only 854 circRNAs were identified by all tools out

of the 5071 unique circRNAs detected, indicating that the choice of circRNA identification tool drastically impacts analyses. Furthermore, the use of RNase R$^+$ and RNase R$^-$ libraries from the same samples permitted the calculation of false positives returned by each tool. By analysing each paired combination of circRNA identification tools, the authors show that CIRCexplorer + CIRI had the highest rate of false positives (10.76%), whilst circRNA_finder + MapSplice formed the lowest false-positive rate amongst analysed pairs (8.3%). Perhaps the biggest takeaway from the study was the combination of all tools yielded a false positive rate of 6.56%, trading increased precision at the cost of reduced sensitivity. In 2018 Hansen [72] performed the same analysis again, this time using 11 circRNA identification tools (ACFS, CIRCexplorer, CIRCexplorer2, CIRI, CIRI2, DCC, find_circ, KNIFE, MapSplice and UROBORUS). Results echoed those from 2016, however Hansen made pertinent points surrounding tool selection for an ensemble approach: users should combine results from circRNA identification tools that utilise different aligners to avoid biases. One such example is circRNA_finder and DCC, which both use the STAR aligner. These two algorithms are thus less suited for pairing as the false positives generated are likely to be inherent to the aligner used. The analyses performed by Hansen et al. set new standards for best practices surrounding circRNA detection, ushering in a new class of circRNA identification pipelines termed 'integrated tools' in which the user can select one or multiple tools for circRNA identification analysis with an automated intersection of results based on user-defined parameters.

CircomPara, developed by Gaffo et al. (2017) [73] represents the first integration-based identification tool offering users the choice of CIRI, CIRCexplorer (STAR, BWA or Segemehl) and find_circ. Results are configurable by requiring detected circRNAs to have n reads spanning their BSJ site and/or circRNAs to be called by at least n tools. Requiring only input sequencing reads, a reference FASTA file and a reference annotation file, the workflow streamlines the process of circRNA identification for users by automatically generating the required genome indices, reformatting reference annotation files and executing scripts for the analysis. The authors have also made the considerable effort to create a docker container with all of the necessary software for the analysis included, circumventing the need to install from source. Gaffo et al. made substantial upgrades to CircomPara in 2022 by releasing CircomPara2 [74]. In addition to offering updated circRNA identification tools to the user (CIRI2, CIRCexplorer2 (BWA, Segemehl, STAR, TopHat), DCC and find_circ), the workflow includes an improved expression estimate step when consolidating results from multiple tools. In CircomPara, circRNA abundances from multiple methods were calculated using the median of library-normalized BSJ counts across tools. In CircomPara2, the authors identify for each method the number of unique reads spanning the BSJ site of a circRNA thereby preserving the information returned by each tool used. Similar to CircomPara, CircomPara2 is packaged in a docker container facilitating rapid execution for users.

Several other integration tools exist for circRNA identification [75–79], however they operate by using as input previously generated results from circRNA identification tools. In our opinion, this does not meet the standards set by CircomPara and CircomPara2 which produce results directly from raw sequencing reads. One such novel integration tool that meets these standards is nf-core/circrna, a workflow for the quantification, miRNA target prediction and differential expression analysis of circRNAs [80]. The workflow takes as input raw sequencing reads, a reference FASTA, reference gene annotation file and performs all of the preprocessing steps and execution scripts required for a circRNA analysis using circRNA_finder, CIRIquant, CIRCexplorer2 (STAR), DCC, find_circ, MapSplice and Segemehl. Similarly to CircomPara, the user can specify custom filtering parameters dictating the intersection strategies used on results. With support for 18 species, the workflow additionally performs automatic miRNA target prediction using miRanda and TargetScan, and automated differential expression analysis of circRNAs between phenotypes of interest provided in an optional metadata file. Developed using nextflow DSL2, the workflow requires Java version >8, the latest version of nextflow and a client to automatically down-

load software packages for each analysis step (Docker, Apptainer, Conda) facilitating rapid 'out-of-the-box' deployment using a single command.

3.5. Full circle reconstruction

The first iteration of circRNA detection tools discussed above are predicated on identifying circRNAs via the presence of BSJ reads in sequencing data. Whilst this is an effective method to detect and quantify circRNAs in RNA-Seq data, the underlying mature spliced sequence (i.e the internal structure) of circRNAs remained opaque. circRNAs are subject to internal splicing events and intron retention (EIcircRNAs), thus assuming that all of the underlying exons are retained within a circRNA will lead to false positives when predicting their targets based on sequence alignment against miRNA and RBP databases. To overcome this limitation and elucidate circRNA isoforms, coverage of paired-end RNA-Seq reads through the circRNA are utilised to characterise read densities amongst exons within the circular transcript. Full characterization of circRNAs (FUCHs) [81] was the first tool released capable of delineating circular isoforms, accepting as input results from circRNA_finder, CIRI2, CIRCexplorer2, and DCC in conjunction with a BAM file containing chimeric reads, linear reads and unmapped reads. The first step is to isolate circular reads from the BAM file, then identify splicing events within the circular transcript by detecting exon-skipping events in reads. The coordinates of the skipped exons are used to generate coverage statistics, assigning reads to one of two circular isoforms. The output files generated detail the circular candidate's genomic location coupled with read depth for each underlying exon. In this way, researchers can delineate the spliced transcript by removing exons with a read count of 0. The developers of CIRI2 have developed the software package CIRI-AS [82] for *de novo* detection of circRNA exons. Using the outputs from CIRI2 and a BAM file generated by BWA-MEM, the algorithm works by analysing local alignment positions of segments within candidate BSJ reads and its paired mate to identify forward spliced junctions representative of joined circular exons. For each circexon candidate, sequencing depth variation, BSJ read pair coverage and splice junctions from non-BSJ reads are taken into account. CIRI-AS can be performed without a reference GTF file, permitting flexible usage with non-reference organisms. In addition to detecting circexons, CIRI-AS can detect intronic or intergenic circRNA fragments (ICFs) when adequate sequencing depth is provided. CIRI-full [83] is yet another tool for full resolution of circRNAs internal structure developed by the same group. The main premise of CIRI-full revolves around the detection of reverse overlap (RO) reads. During reverse transcription, the circular structure of circRNAs can cause continuous circumnavigation of reverse transcriptase within the circRNA, producing read pairs that overlap in reverse orientation. Moreover, the presence of a 3'-RO overlap in both RO reads indicates the full circle has been transcribed facilitating full circRNA reconstruction. For RO reads that do not overlap due to insert size length, CIRI-full borrows information from CIRI2 (BSJ sites) and CIRI-AS (circexons) to produce a reconstructed circRNA. Next, a forward-splice graph (FSG) is constructed by assembling BSJ and RO reads within a detected circRNA BSJ site to model the read coverage of each path using Monte Carlo simulations, providing resolution of circRNA isoforms.

Machine learning circRNA identification

circRNA biogenesis can be attributed to hallmarks within the mature sequence and the flanking regions: homologous sequences in flanking introns [84], inverted repeats [85], the presence of ALU/tandem repeats [7] and the density of single nucleotide polymorphisms (SNP) [86]. These hallmarks coupled with evolutionary conservation and secondary structure have been identified as the top ranking features for discriminating circRNAs from other classes of lncRNAs using statistical based approaches (minimum redundancy maximum relevance (mRMR) method [87], Random Forest, LightGBM, XGBoost [88] and multiple kernel learning [89]). A limitation of these methods is that splice site and back-splice junctions are ignored, focusing instead on sequence context. Given the unique BSJ of circRNAs, it is key to understand the properties and relationships between splice sites that constitute

canonical linear splicing and a circular back-splicing event. circDeep [90] and DeepCirCode [91] analyzes the nucleotide sequences of two splice sites and models the flanking regions to predict circRNAs by leveraging sequences from circBase and circRNADb in conjunction with deep learning models. Junction encoders and deep interation (JEDI) among splice sites [92] is a tool that optimizes a deep learning model for circRNA prediction in the absence of annotated back-splice sites as training data (zero-shot learning). Unlike its predecessors, JEDI is not limited to interrogating only two splice sites. In this way, it can model the sequence context and flanking regions of all splice sites within an transcript, making it an effective tool for classifying circRNAs that are derived from genes which also produce linear transcripts.

Table 1: Bioinformatic tools for circRNA identfication, quantification, isoform detection, full circle reconstruction, target prediction and differential expression analysis.

Tool name	Description*	Installation**	Aligner***	Language	Ref
ACFS	Identification & quantification of circRNAs	Manual	BWA, BLAT	Perl	[65]
ACValidator	Assembly based circRNA detection	Manual, pip	BWA, Bowtie2	Python2	[93]
ANNOgesic	Archael/bacterial circRNA identification	Manual, pip3, Docker	segemehl	Python3	[94]
AutoCirc	Fast identification of circRNAs	Manual	Bowtie2	C++,Perl	[95]
BIQ	Identify circRNAs using k-mers spanning BSJ	Manual	k-mer	C++,Perl, JavaScript	[96]
CircAST	Full circle reconstruction, assemble & quantify circular isoforms	Manual	TopHat	Python2	[97]
CircDBG	De Bruijn graph detection of circRNAs	Manual	k-mer	C++	[98]
circDeep	circRNA ifdentification using deep learning	Manual	k-mer	Python3	[90]
CIRCexplorer	Identify, quantify & annotate circRNAs	Conda, pip, BioContainers	STAR, TopHat	Python2	[66]
CIRCexplorer2	Identify, quantify & annotate circRNAs with updated *De novo* module	Conda, pip, BioContainers	BWA, MapSplice, Segemehl, STAR, TopHat	Python2	[67]
CIRCexplorer3	Compare circRNA & linear expression	Manual	Hisat2, StringTie	Python3	[99]
circfull	Detect circRNA isoforms from nanopore reads	Manual	minimap2	Python3	[100]
circLGB-circMRT	Predicting circRNA regulatory interactions by machine learning	Manual	*NA*	Python2	[101]
circMeta	Downstream functional analysis of circRNAs	devtools	*NA*	R	[77]
CircMarker	Fast identification of circRNAs	Manual	k-mer	C++, JavaScript	[102]

Table 1: Bioinformatic tools for circRNA identification, quantification, isoform detection, full circle reconstruction, target prediction and differential expression analysis. (Continued)

CircomPara	Automated detection of circRNAs usign integrated approach	Manual, Docker	Multiple	Python2, R	[73]
CirComPara2	Automated detection of circRNAs using integrated approach	Manual, Docker	Multiple	Python2, R	[74]
CircPro	circRNA coding potential using RNA-Seq & Ribo-Seq reads	Manual	BWA	Perl	[103]
circRNA_finder	Identification of circRNAs	Conda, BioContainers	STAR	Perl	[14]
circRNAFisher	Identification of circRNAs using statistical framework	Manual	Bowtie2	Perl	[104]
circRNAprofiler	Downstream functional analysis of circRNAs	BiocManager	*NA*	R	[78]
circRNAwrap	Automated detection, abundance, isoform & DE analysis	Manual	Multiple	UNIX	[75]
circtools	Suite of tools to identify, quantify, visualise & perform DEA	Docker, pip3	STAR	Python3,R	[105]
CIRI	Identification & quantification of circRNAs	Manual	BWA	Perl	[62]
CIRI2	Identification & quantification of circRNAs	Manual	BWA	Perl	[63]
CIRI-AS	Alternative circRNA splicing	Manual	BWA	Perl	[82]
CIRI-full	Full circle reconstruction	Manual	BWA	JavaScript	[83]
CIRI-long	Identify circRNAs in Nanopore reads	Manual	BWA, minimap2	C++, Python3	[106]
CIRIquant	Identification, quantification, RNAase R correction, DEA of circRNAs	Manual	BWA, HISAT2	Perl	[64]
CIRI-viz	Visualising circRNA alignments & isoforms	Manual	*NA*	JavaScript	*NA*
CirRBP	Stacked generalization ensemble deep learning model to identify RBP binding sites	Web server	*NA*	*NA*	[107]
CirRNAPL	circRNA identification using machine learning	Web server	*NA*	*NA*	[108]
CRIP	Predict circRNA-RBP interactions using neural networks & stacked codon-encoding	Manual	*NA*	Python3	[109]
CYCLeR	Reconstruct & quantify circRNAs	Docker	BWA, STAR, kallisto	R	[110]
DCC	Identify & quantify circRNAs	Conda, BioContainers	STAR	Python2	[55]
DEBKS	Downstream differential circRNA analysis	Conda, pip	*NA*	Python3	[76]
DeepCirCode	circRNA ifdentification using deep learning	Manual	*NA*	Pyton2, R	[91]
exceRpt	extracellular circRNA profiling	Manual, Docker	Bowtie2, STAR	JavaScript, UNIX	[111]
FcircSEC	Full sequence reconstruction	CRAN, devtools	*NA*	R	[79]

Table 1: Bioinformatic tools for circRNA identfication, quantification, isoform detection, full circle reconstruction, target prediction and differential expression analysis. (Continued)

find_circ	Identification & quantification of circRNAs	Conda, BioContainers	Bowtie2	Python2	[9]
FUCHS	Alternative circRNA splicing	Manual	*NA*	Python2, R	[81]
hppRNA	Workflow for mRNA, lncRNA, circRNA identification & quantification	Manual	STAR	Perl	[112]
isoCirc	Full length circRNA isoform reconstruction	Manual, pip	Minimap2		[113]
JEDI	circular RNA prediction based on junction encoders and deep interaction among splice sites	Manual	*NA*	Python3	[92]
KNIFE	Statistical based circRNA identification	Manual	Bowtie2		[61]
MapSplice2	Splice-aware aligner	Conda, BioContainers	Bowtie	C++	[114]
miARma	circRNA quantification, miRNA targets & DEA	Manual	BWA	Perl	[115]
NCLcomparator	Screening of NCLscan results	Manual	*NA*	Python2,R, UNIX	[116]
NCLscan	Identification of non-co-linear transcripts	Manual	BWA, BLAT	Perl,Python2	[69]
nf-core/circrna	Autmated circRNA quantification, miRNA target prediction & DEA	Conda,Docker	Multiple	nextflow	[80]
Pcirc	Random forest plant circRNA identification	Manual	Bowtie2, TopHat2	Python3, R	[117]
PcircRNA_finder	Plant circRNA identification	Manual	*NA*	Perl, Python2	[118]
PRAPI	circRNA identification from PacBio Iso-Seq	Manual	GMAP	Python2	[119]
PredcircRNA	Classification of circRNAs using hybrid features	Manual	*NA*	Python2	[89]
PredcircRNATool	circRNA detection based on thermodynamic properties of flanking introns	Manual	*NA*	Python2	[120]
PTESfinder	Identify post-transcriptional exon shuffling events	Manual	Bowtie	JavaScript, UNIX	[60]
RAISE	Identification, quantification & internal structure	Manual	Bowtie2, BWA, HISAT2, STAR, StringTie	UNIX	[121]
ROP	Identify RNAs in unmapped reads	Manual	*NA*	Python2	[122]
segemehl	Short read mapper capable of detecting circRNAs	Conda, BioContainers	segemehl	C++	[70]
StackCirRNAPred	Classification of circRNAs using Random Forest, LightGBM & XGBoost	Manual	*NA*	*NA*	[88]
STARChip	Identify circRNAs from STAR junction files	Manual	*NA*	Perl, UNIX	[123]
Ularcirc	Rshiny visualisation of circRNAs	devtools	*NA*	R	[124]

423

Table 1: Bioinformatic tools for circRNA identfication, quantification, isoform detection, full circle reconstruction, target prediction and differential expression analysis. (Continued)

UROBORUS	circRNA identification		Manual	Bowtie	Perl	[59]

* *DEA* = Differential expression analysis

** *Manual*, requires one of source installation from GitHub, compilation using make, prerequisite software to be previosuly installed or a combination of all three. *BioContainers*, all Conda packages are automatically converted to container images hosted on BioContainers. Available via container clients such as singularity, docker etc.

*** *NA* refers to downstream tools that consume previously generated circRNA identififcation tool outputs as input, or classification tools that leverage experimentally validated interactions for prediction tasks.

4. Differential expression analysis

Once the circRNA transcriptome has been characterised in samples, it is often the goal of researchers to perform differential expression analysis (DEA) between phenotypes of interest using the generated circRNA count matrix. DEA can be performed Manually using popular tools such as DESeq2 [125], EdgeR [126] and limma-voom [127]. Both DESeq2 and EdgeR fit a negative binomial distribution to the counts matrix and use generalized linear models to perform statistical tests, whereas limma-voom computes observational weights for a linear model using mean-variance relationship between samples on the logarithmic scale. A common filtering step prior to DEA is to require ≥ 2 reads spanning the BSJ site of quantified circRNAs. Whilst this will result in a count matrix with higher confidence circRNAs, there remains the problem of multiple zero values present in columns (samples) in which the high confidence circRNAs were not detected resulting in a sparse matrix. In our experience, providing a sparse matrix to the DESEq2/EdgeR/limma-voom packages will result in an error when calculating the library size factors for normalization. To remedy this, we suggest applying a pseudocount to the sparse matrix prior to performing DEA.

A major factor of DEA that has only recently been considered is the increasingly popular use of multiple quantification tools to generate the final count matrix [72,74,80]. This comes with the upside of increasing the recall rate of the quantification analysis by overlapping the calls of multiple quantification tools, however the number of called reads spanning the BSJ site for a circRNA are likely divergent across the quantification tools employed [71,72]. This presents the issue of which function to apply when consolidating reads from multiple tools; should researchers average circRNA expression across multiple tools? Perhaps they may be inclined to take the read count value returned for a given circRNA by a tool. Regardless of the function applied, there will at the very least be a loss of information and at worst, a significant overestimation of a circRNAs expression by selecting highly expressed outliers. To overcome this issue, Buratin et al. (2022) [128] perform DEA by modelling the effect of the phenotype of interest whilst simultaneously modelling the variance of circRNA reads between different quantification tools as a random effect using generalized linear mixed models e.g: circRNA counts $\sim$ phenotype group + (1|quantification tool 1) + (1|quantification tool 2) etc. In this manner, one can obtain robust differentially expressed circRNAs estimates without discarding any of the information obtained from mutliple quantification tools. We recommend users adopt this approach when using a consensus based approach to calling circRNAs, a method that has been shown to increase the performance of the quantification step [72].

There has also been considerable efforts made to automate the process of differential expression analysis of circRNAs for researchers. CIRIquant, nf-core/circrna and CircTest offer automated differential expression analysis of circRNAs using edgeR, DESeq2 and a custom beta binomial distribution coupled with an ANOVA test, respectively. The main drawback of using automated differential expression analysis pipelines are the constraints placed on the complexity of the model design; these methods are only capable of analyzing the predictor variable whilst controlling for the effect of covariates, unable to facilitate

more complex designs with additive, interactive or nested effects. For complex designs, we recommend users perform differential expression analysis manually.

5. circRNA interactions

ceRNA networks

circRNAs can function as miRNA sponges when they enter the cytoplasm [8,9], forming a role within the ceRNA network by competitively binding miRNAs and sequestering the degradation of its mRNA target. The predicted interactions of circRNA-miRNAs and miRNA-mRNAs targets are used to create a tri-partite ceRNA network representing the interplay between circRNA-miRNA-mRNA in cells. Researchers can achieve this by using existing databases, performing manual predictions using sequence alignment tools against databases, or a combination of both. Several publicly available databases exist containing predicted circRNA-miRNA interactions in downloadable files such as CircBase [129] and CSCD [130]. Additionally, starBase [131] offers an API function to submit requests for predicted circRNA-miRNA targets. Once the circRNA-miRNA pairs have been generated, the miRNAs can be used as inputs for deriving miRNA-mRNA interactions. Given miRNAs have been studied since the early 1990s (compared to circRNAs relatively new revival in 2012), there exist multiple databases for predicting miRNA-mRNA pairs. miRBase [132], miRTaRBase [133], miRNet [134] and TargetScan [135] represent a selection of the available databases for this task.

To predict circRNA-miRNA and miRNA-mRNA targets manually, users can avail of the multiple sequence alignment tools miRanda [136] and TargetScan [135]. The full mature spliced sequence of each circRNA can be scanned for miRNA response element (MRE) sites by passing the sequence in FASTA format to each tool. TargetScan offers the advantage of reporting each miRNA match as a 6-mer, 7-mer or 8-mer, detailing the number of matching nucleotides in the circRNA MRE site and the miRNA seed region. To reduce the number of false positives in the analysis, users can adopt three strategies: 1) remove 6-mers sites that are considered poorly conserved in comparison to 7-mer and 8-mers; 2) overlap results between miRanda and TargetScan; or 3) overlap predicted MRE sites with AGO2 binding sites. These filtering steps can be applied to circRNA-miRNA and miRNA-mRNA predictions alike. Finally, in the event expression data between phenotypes is available for circRNAs, miRNAs and mRNAs, users may wish to apply filtering to conform to the ceRNA hypothesis by selecting circRNA-miRNA-mRNAs subgraphs in which the circRNA exhibits up-regulation, its target miRNA is down-regulated and the target mRNA of the down-regulated miRNA is up-regulated. The inverse filtering expression can be applied to generate a ceRNA network modelling up- and down-regulated circRNAs. Tripartite networks can be visualised and analysed using Cytoscape [137] and its plugins for network analysis. The main challenge in performing manual circRNA-miRNA predictions is providing an accurate mature spliced sequence to each tool, details of which are discussed in section 3.5.

circRNA-RBP prediction

Whilst circRNA-miRNA binding is the most studied functionality of circRNAs, there is increasing evidence to suggest circRNAs interact with RBPs at multiple iterations of their life cycle. Quaking (QKI), FUS, HNRNPL, RBM20 and Muscleblind are all RBPs which bind to specific intronic motifs, promoting the formation of circRNAs [12,138–141], whilst ADAR1 and DHX9 have been shown to destabilize inverted Alu repeats, supressing back-splicing [84,142]. *CircPABPN1* has been shown to modulate the transcription of its cognate mRNA *PABPN1* by competitively binding and reducing the availability of HuR, a translational activator protein [143]. Additionally, *circFoxo3* binds p21 and CDK2 RBPs, forming ternary complexes inhibiting cyclin E/CDK2 complex formation, arresting cells in G1 phase [144].

circRNA-RBP interactions can be characterised using cross-linking and immunoprecipitation (CLIP-seq) datasets however, the assay suffers from limitations. Firstly, CLIP-Seq

reads are produced via enzymatic degredation, producing single-end reads of length <50bp. These short, single-end reads are unsuitable for traditional circRNA identification tools developed for RNA-Seq data which suffer from poor mapping estimates when using short reads and in the absence of paired-end reads, will generate high rates of false-positives. To accurately identify circRNAs in CLIP-Seq data, researchers can use Clirc, a computational tool capable of detecting BSJ sites bound to RBPs [145]. Clirc collapses reads to remove PCR duplicates prior to constructing a psuedo-reference based on publicly available human, mouse and drosophila circRNAs and circRNAs detected in ENCODE datasets using CIRI2. Reads that contiguously align to the reference genome are discarded, whilst reads mapping to the pseudo-reference are indicative of BSJ sites in circRNAs. The authors concede that Clirc is constrained to detecting circRNAs in the pseudo-reference and cannot detect circRNAs *de novo*. Additionally, Clirc can only detect RBPs that directly bind to the BSJ site as distinguishing RBPs binding to 'linear' sequences in circRNAs/mRNAs remains intractable.

Databases such as circInteractome [146] and starBase [131] host results of circRNA-RBP interactions validated using CLIP-Seq experimental data. Due to the costs associated with CLIP-seq, there have been several computational methods developed to predict circRNA-RBP interactions by analysing motif sequences. CircRNAs Interact with Proteins (CRIP) is tool that represents circRNA-RBP interactions as a binary classification problem. The authors have developed a novel sequence encoding scheme whereby RNA triplets are represented as pseudo-amino acids, one-hot encoded and passed to a convolutional neural network and a bidirectional long- and short-term memory network to exploit sequence information of 37 RBPs and the corresponding 32,216 circRNAs they bind [109]. Source code and training data are freely available, allowing users to leverage the information provided by circinteractome to predict circRNA-RBP interactions using their own circRNA sequence data. We concede that this requires a high degree of computational competency to run, in which case users may find CirRBP [107] a more suitable alternative. CirRBP utilizes a stacked ensemble deep learning model to predict RBP binding sites within a user supplied circRNA sequence, sourcing circRNA-RBP binding information from circinteractome, starBase and CSCD2. The authors have packaged the underlying algorithm and models used for CirRBP as a publicly available webserver (http://www.bioinformatics.team) greatly reducing the computational barrier to entry for researchers to perform circRNA-RBP predictions.

6. circRNA databases

Multiple circRNA databases currently exist providing users with circRNA annotations, predicted interactions, mature spliced sequence and expression estimates across cell lines. Typically, these databases are derived from a selection of published ribosomal depleted RNA-Seq datasets [6,7,9,12,16,58,84,147] and processed using a circRNA identification pipeline. It is worth noting that there is no universal 'gold standard' pipeline for circRNA identification, thus each database will vary in their results. For example, circBase [129] and CIRCpedia [148] use the find_circ identification tool, whilst CSCD2 [149] employs CIRI, CIRCexplorer2, circRNA_finder and find_circ to produce its database, allowing users to identify which circRNAs have been called by multiple tools. Other databases such as circRNADb [150] host circRNA annotations collated from published literature, removing biases inherent to specific pipelines. With respect to the fucntional interactions of circRNAs, the starBase [131] and TRCirc [151] databases contain RNA-RNA interactions and RNA-protein interactions using CLIP-Seq and CHIP-Seq data, respectively. Researchers can also search disease specific circRNAs backed by experimental findings in published literature via Circ2Disease [152].

One of the key challenges facing researchers when using circRNA databases is the lack of a standardised naming format for circRNAs. Chen et al. (2023) [153] use the example of the functional *FAM20A* circRNA to depict the wild discrepancies between nomenclature: HSA_CIRCpedia_64725 in CIRCpedia, hsa-FAM120A_0006 in circAtlas,

hsa_circFAM120A_007 in circBank, and hsa_circ_0001875 in circBase. With 'hsa' being the only common string between the four identifiers, Chen et al. proposed a novel naming convention for circRNAs. They provide examples for classic exonic circRNAs, EIcircRNAs and ciRNAs: 1) *circCOX5A(2,3)* an exonic circRNA derived from the *COX5A* gene that uses exons 2 and 3; 2) *circCAMSAP1(2,RI,3)* an exonic-intronic circRNA derived from the *CAMSAP1* gene that uses exons 2 and 3, with a retained intron (RI) between exons 2 and 3; 3) *ciANKRD52(2)* an intronic circRNA derived from the *ANKRD52* gene that retains the second intron. Whilst we agree that the standardisation of circRNA identifiers is a useful endeavour, we believe it may take several years to adopt. The main hurdle we observe is that the most popular circRNA identification tools that are currently in use by researchers are solely based on identifying BSJ reads in RNA-Seq data. These tools cannot fully resolve the internal structure of circRNAs, thus researchers are not able to accurately annotate internal exon/intron usage required to conform to the proposed naming convention. With this in mind, an end-to-end pipeline for full circle reconstruction capable of outputting the updated naming convention proposed by Chen et al. would be of tremendous benefit to the research community. Secondly, CircBase, CSCD2 and CIRCPedia remain hugely popular databases for circRNAs. It would take a large effort to re-analyze and update the circRNA identifiers present in these databases. However, as the data used to produce these databases is publicly available, this is not an impossible task – particularly if an end-to-end annotation pipeline was developed. Finally, the research community must agree on the reference used for reporting circRNAs (e.g ENSEMBL [154], RefSeq [155]) and make clear which reference annotation files were used for circRNA annotation in order to make results reproducible.

Table 2: Online databases for circRNAs

Database	Data available	Organisms[1]	Reference
AtCircDB	A. thaliana circRNAs, miRNA targets	ath	[156]
circAtlas	circRNA sequences, conservation score, miRNA & RBP targets, isoforms, expression in tissues, junction ratio, reported diseases	hsa, mml, mmu, rno, ocu, clf, fca, ssc, oar, gga	[157]
circBase	circRNA sequences, circRNA expression in cell lines/tissues	hsa, mmu, cel, lch, lme	[129]
CircBank[*]	miRNA targets, m^6A modifications, conservation, mutations and coding potential	hg19	[158]
CircInteractome	miRNA & RBP targets, primer design, siRNA sites	hsa	[146]
CircNet	miRNA & RBP targets, ceRNA networks construction, coding potential	hsa	[159]
CircFunBase	circRNA predicted function, miRNA & RBP targets, visual representation of ceRNA network	ath, osa, tae, sly, gsp, hvu, ade, hsa, mml, rno, mmu, gga, ssc, bta, dme, ocu	[160]
CIRCpedia	circRNA exprression in cells and tissues	hsa, mmu, cel, dme, dre, rno	[148]
CSCD2	circRNAs in cancer, target miRNAs & RBP, coding potential	hsa	[149]
circRNADb	exonic circRNAs, coding potential	hsa	[150]
CircRiC[*]	circRNAs in cancer	hsa	[161]
Circ2Disease	circRNAs associated with diseases	hsa	[152]
CircR2Disease	circRNAs associated with diseases	hsa	[162]

Table 2: Online databases for circRNAs (Continued)

Circ2Traits[*]	circRNA-miRNA disease associations	hsa	[163]
circRNADisease[*]	circRNAs associated with diseases	hsa	[164]
CropCircDB	circRNAs in maize & rice	osa, zma	[165]
DeepBase	circRNA, lncRNA, miRNAs in tissues and cancers	hsa	[166]
exoRBase	Atlas of mRNAs, lncRNAs,& circRNAs in extracellular vesicles	hsa	[167]
MiOncoCirc	Compendium of circRNA datasets in cancer	hsa	[168]
NeuroCirc	circRNA expression in brain regions, circQTLs	hsa	[169]
PlantCircNet[*]	ceRNA regulatory networks	ath, gma, hvu, osa, sly, tae, zma, bdi	[170]
PlantCircBase	circRNAs in plants, ceRNA regulatory network	ath, gma, hvu, osa, sly, tae, zma, gar, ghi, gra, ptr, stu, csi, nbe, pbe, osi	[171]
starBase	RNA-RNA interactions based on CLIP-Seq data	23 species	[131]
TRCirc	Transcriptional regulation of circRNAs using CHIP-Seq data	hsa	[151]

[1] Species abbreviations: ade, *Actinidia deliciosa*; ath, *Arabidopsis thaliana*; bdi, *Brachypodium distachyon*; bta, *Bos taurus*; cel, *Caenorhabditis elegans*; csi, *Camellia sinensis*; dme, *Drosophila melanogaster*; dre, *Danio rerio*; gar, *Gossypium arboreum*; gga, *Gallus gallus*; ghi, *Gossypium hirsutum*; gma, *Glycine max*; gra, *Gossypium raimondii*; gsp, *Gossypium spp.*; hsa, *Homo sapiens*; hvu, *Hordeum vulgare*; lch, *Latimeria chalumnae*; lme, *Latimeria menadoensis*; mml, *Macaca mulatta*; mmu, *Mus musculus*; nbe, *Nicotiana benthamiana*; ocu, *Oryctolagus cuniculus*; osa, *Oryza sativa*; osi, *Oryza sativa ssp. indica*; pbe, *Pyrus betulifolia*; ptr, *Poncirus trifoliata*; rno, *Rattus norvegicus*; sly, *Solanum lycopersicum*; ssc, *Sus scrofa*; stu, *Solanum tuberosum*; sly, *Solanum lycopersicum*; tae, *Triticum aestivum*; zma, *Zea mays*

[*] URL not accessible at time of drafting review

7. Concluding remarks

circRNAs are a class of non-coding RNAs which are gaining recognition for their roles in cellular processes as transcriptional regulators. Despite circRNAs representing an increasingly popular area of research, it is still in its infancy with respect to computational approaches. This is perhaps most apparent in a subset of widely utilised circRNA identification tools that are entirely predicated on detecting the BSJ of circRNAs in sequencing data—mimicking microarray technologies which rely on short oligonucleotide probes to detect and quantify mRNA transcripts. Whilst these tools are useful for detecting and quantifying circRNAs, they are unable to fully resolve the full circRNA sequence or delineate circRNA isoforms generated from the same BSJ. In this regard, circRNA tools are severely lacking in comparison to RNA-Seq tools. Efforts have been made to address these limitations, however uptake from the research community has been slow. Tools for full circle reconstruction are difficult to use and require specific inputs generated from circRNA identification tools. Their outputs are typically a modified iteration of BED12 format, with circRNA IDs reported as 'chr:start|stop' and its underlying exon sequence reported in separate columns. In our opinion, it is of utmost importance to standardise the nomenclature of circRNAs using the suggestions by Chen et al. [153]. Once a standardised format has been agreed upon, the research community can begin developing circRNA identification tools that can report circRNAs accurately with standardized naming conventions. This will feed circRNA databases, and may eventually lead to the creation of a reference circRNA GTF/GFF file which will greatly enhance quantification tools. We estimate that once these steps have been completed, there is no reason why the circRNA research community cannot

develop rapid circRNA identification/quantification tools that mirror those in the RNA-Seq space such as Kallisto [172] and Salmon [173].

Abbreviations

The following abbreviations are used in this manuscript:

circRNA	Circular RNA
ncRNA	Non-coding RNA
mRNA	Messenger RNA
miRNA	Micro RNA
lncRNA	Long non-coding RNA
tRNA	Transfer RNA
pre-mRNA	Precursor mRNA
ciRNA	Intron circRNA
ecircRNA	Exonic circRNA
EIcircRNA	Exon-Intron circRNA
RNA-Seq	RNA-Sequencing
RBP	RNA-binding protein
RNase R	Ribonuclease R
rRNA	Ribosomal RNAs
ceRNA	Competing endogenous RNA
CLIP-seq	Crosslinking immunoprecipitation-high-throughput-sequencing
MRE	MicroRNA response element site
NGS	Next generation sequencing
ALL	Acute lymphoblastic leukaemia
MBLN1	Musclebind
QKI	Quaking
LNCaP	Lymph Node Carcinoma of the Prostate
Ago2	Argonaute 2
m^6A	N6-Methyladenosine
oligo(DT)	Short sequence of deoxy-thymidine nucleotides
cDNA	complementary DNA
BSJ	Back-splice junction
BMJ	Balanced mapped junctions
UMJ	Unbalanced mapped junctions
GLM	Generalized linear model
MLE	Maximum likelihood estimation
DEA	Differential expression analysis
RO	Reverse overlap
ALU	Arthrobacter luteus
SNP	Single nucleotide polymorphism
ANOVA	Analysis of variance
FUS	Fused in sarcoma
HNRNPL	heterogeneous nuclear ribonucleoprotein L
RBM20	RNA binding motif protein 20
ADAR1	Adenosine deaminases acting on RNA 1
DHX9	DExH-box helicase 9
CDK2	Cyclin-dependent kinase/cell division kinase 2

References

1. Sanger, H.L.; Klotz, G.; Riesner, D.; Gross, H.J.; Kleinschmidt, A.K. Viroids are single-stranded covalently closed circular RNA molecules existing as highly base-paired rod-like structures. *Proceedings of the National Academy of Sciences* **1976**, *73*, 3852–3856, [https://www.pnas.org/content/73/11/3852.full.pdf]. https://doi.org/10.1073/pnas.73.11.3852.

2. Arnberg, A.C.; Van Ommen, G.J.B.; Grivell, L.A.; Van Bruggen, E.F.J.; Borst, P. Some yeast mitochondrial RNAs are circular. *Cell* **1980**, *19*, 313–319. https://doi.org/10.1016/0092-8674(80)90505-X.

3. Kos, A.; Dijkema, R.; Arnberg, A.C.; van der Meide, P.H.; Schellekens, H. The hepatitis delta (δ) virus possesses a circular RNA. *Nature* **1986**, *323*, 558–560. https://doi.org/10.1038/323558a0.

4. Cocquerelle, C.; Mascrez, B.; Hétuin, D.; Bailleul, B. Mis-splicing yields circular RNA molecules. *FASEB J.* **1993**, *7*, 155–160. https://doi.org/10.1096/fasebj.7.1.7678559.

5. Qian, L.; Vu, M.N.; Carter, M.; Wilkinson, M.F. A spliced intron accumulates as a lariat in the nucleus of T cells. *Nucleic Acids Research* **1992**, *20*, 5345–5350, [https://academic.oup.com/nar/article-pdf/20/20/5345/7073334/20-20-5345.pdf]. https://doi.org/10.1093/nar/20.20.5345.

6. Salzman, J.; Gawad, C.; Wang, P.L.; Lacayo, N.; Brown, P.O. Circular RNAs Are the Predominant Transcript Isoform from Hundreds of Human Genes in Diverse Cell Types. *PLOS ONE* **2012**, *7*, 1–12. https://doi.org/10.1371/journal.pone.0030733.

7. Jeck, W.R.; Sorrentino, J.A.; Wang, K.; Slevin, M.K.; Burd, C.E.; Liu, J.; Marzluff, W.F.; Sharpless, N.E. Circular RNAs are abundant, conserved, and associated with ALU repeats. *RNA* **2012**, *19*, 141–157. https://doi.org/10.1261/rna.035667.112.

8. Hansen, T.B.; Jensen, T.I.; Clausen, B.H.; Bramsen, J.B.; Finsen, B.; Damgaard, C.K.; Kjems, J. Natural RNA circles function as efficient microRNA sponges. *Nature* **2013**, *495*, 384–388, [23446346]. https://doi.org/10.1038/nature11993.

9. Memczak, S.; Jens, M.; Elefsinioti, A.; Torti, F.; Krueger, J.; Rybak, A.; Maier, L.; Mackowiak, S.D.; Gregersen, L.H.; Munschauer, M.; et al. Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature* **2013**, *495*, 333–338, [23446348]. https://doi.org/10.1038/nature11928.

10. Danan, M.; Schwartz, S.; Edelheit, S.; Sorek, R. Transcriptome-wide discovery of circular RNAs in Archaea. *Nucleic Acids Res.* **2012**, *40*, 3131–3142, [22140119]. https://doi.org/10.1093/nar/gkr1009.

11. Jeck, W.R.; Sharpless, N.E. Detecting and characterizing circular RNAs. *Nat. Biotechnol.* **2014**, *32*, 453–461. https://doi.org/10.1038/nbt.2890.

12. Ashwal-Fluss, R.; Meyer, M.; Pamudurti, N.R.; Ivanov, A.; Bartok, O.; Hanan, M.; Evantal, N.; Memczak, S.; Rajewsky, N.; Kadener, S. circRNA biogenesis competes with pre-mRNA splicing. *Mol. Cell* **2014**, *56*, 55–66, [25242144]. https://doi.org/10.1016/j.molcel.2014.08.019.

13. Chen, L.L.; Yang, L. Regulation of circRNA biogenesis. *RNA Biol.* **2015**, *12*, 381–388. https://doi.org/10.1080/15476286.2015.1020271.

14. Westholm, J.O.; Miura, P.; Olson, S.; Shenker, S.; Joseph, B.; Sanfilippo, P.; Celniker, S.E.; Graveley, B.R.; Lai, E.C. Genomewide analysis of Drosophila circular RNAs reveals their structural and sequence properties and age-dependent neural accumulation. *Cell Rep.* **2014**, *9*, 1966. https://doi.org/10.1016/j.celrep.2014.10.062.

15. Gruner, H.; Cortés-López, M.; Cooper, D.A.; Bauer, M.; Miura, P. CircRNA accumulation in the aging mouse brain. *Sci. Rep.* **2016**, *6*. https://doi.org/10.1038/srep38907.

16. Rybak-Wolf, A.; Stottmeister, C.; Glažar, P.; Jens, M.; Pino, N.; Giusti, S.; Hanan, M.; Behm, M.; Bartok, O.; Ashwal-Fluss, R.; et al. Circular RNAs in the Mammalian Brain Are Highly Abundant, Conserved, and Dynamically Expressed. *Mol. Cell* **2015**, *58*, 870–885, [25921068]. https://doi.org/10.1016/j.molcel.2015.03.027.

17. Li, P.; Chen, S.; Chen, H.; Mo, X.; Li, T.; Shao, Y.; Xiao, B.; Guo, J. Using circular RNA as a novel type of biomarker in the screening of gastric cancer. *Clin. Chim. Acta.* **2015**, *444*, 132–6, [25689795]. https://doi.org/10.1016/j.cca.2015.02.018.

18. Wan, L.; Zhang, L.; Fan, K.; Cheng, Z.X.; Sun, Q.C.; Wang, J.J. Circular RNA-ITCH Suppresses Lung Cancer Proliferation via Inhibiting the Wnt/β-Catenin Pathway. *Biomed Res. Int.* **2016**, *2016:1579490.*, Biomed, [27642589]. https://doi.org/10.1155/2016/1579490.

19. Verduci, L.; Ferraiuolo, M.; Sacconi, A.; Ganci, F.; Vitale, J.; Colombo, T.; Paci, P.; Strano, S.; Macino, G.; Rajewsky, N.; et al. The oncogenic role of circPVT1 in head and neck squamous cell carcinoma is mediated through the mutant p53/YAP/TEAD transcription-competent complex. *Genome Biol.* **2017**, *18*, 1–24. https://doi.org/10.1186/s13059-017-1368-y.

20. Bayat, H.; Pourgholami, M.H.; Rahmani, S.; Pournajaf, S.; Mowla, S.J. Synthetic miR-21 decoy circularized by tRNA splicing mechanism inhibited tumorigenesis in glioblastoma in vitro and in vivo models. *Mol. Ther. Nucleic Acids* **2023**, *32*, 432–444. https://doi.org/10.1016/j.omtn.2023.04.001.

21. Greene, J.; Baird, A.M.; Brady, L.; Lim, M.; Gray, S.G.; McDermott, R.; Finn, S.P. Circular RNAs: Biogenesis, Function and Role in Human Diseases. *Front. Mol. Biosci.* **2017**, *4*, 265463. https://doi.org/10.3389/fmolb.2017.00038.

22. Kristensen, L.S.; Andersen, M.S.; Stagsted, L.V.W.; Ebbesen, K.K.; Hansen, T.B.; Kjems, J. The biogenesis, biology and characterization of circular RNAs. *Nat. Rev. Genet.* **2019**, *20*, 675–691. https://doi.org/10.1038/s41576-019-0158-7.

23. Li, X.; Yang, L.; Chen, L.L. The Biogenesis, Functions, and Challenges of Circular RNAs. *Mol. Cell* **2018**, *71*, 428–442, [30057200]. https://doi.org/10.1016/j.molcel.2018.06.034.

24. Ren, L.; Jiang, Q.; Mo, L.; Tan, L.; Dong, Q.; Meng, L.; Yang, N.; Li, G. Mechanisms of circular RNA degradation. *Commun. Biol.* **2022**, *5*, 1–6. https://doi.org/10.1038/s42003-022-04262-3.

25. Yang, Q.; Li, F.; He, A.T.; Yang, B.B. Circular RNAs: Expression, localization, and therapeutic potentials. *Mol. Ther.* **2021**, *29*, 1683–1702. https://doi.org/10.1016/j.ymthe.2021.01.018.

26. Huang, A.; Zheng, H.; Wu, Z.; Chen, M.; Huang, Y. Circular RNA-protein interactions: functions, mechanisms, and identification. *Theranostics* **2020**, *10*, 3503. https://doi.org/10.7150/thno.42174.

27. Panda, A.C. Circular RNAs Act as miRNA Sponges. *Adv. Exp. Med. Biol.* **2018**, *1087:67-79.*, Adv, [30259358]. https://doi.org/10.1007/978-981-13-1426-1_6.

28. Santos-Rodriguez, G.; Voineagu, I.; Weatheritt, R.J. Evolutionary dynamics of circular RNAs in primates. *eLife* **2021**. https://doi.org/10.7554/eLife.69148.

29. RNA Splicing | Learn Science at Scitable, 2019. [Online; accessed 31. Aug. 2023].

30. 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*, 103–111. https://doi.org/10.1016/j.celrep.2014.12.002.

31. Wilusz, J.E. Repetitive elements regulate circular RNA biogenesis. *Mobile Genetic Elements* **2015**, *5*, 39. https://doi.org/10.1080/2159256X.2015.1045682.

32. Zhang, X.O.; Wang, H.B.; Zhang, Y.; Lu, X.; Chen, L.L.; Yang, L. Complementary Sequence-Mediated Exon Circularization. *Cell* **2014**, *159*, 134–147. https://doi.org/10.1016/j.cell.2014.09.001.

33. Singh, D.; Kesharwani, P.; Alhakamy, N.A.; Siddique, H.R. Accentuating CircRNA-miRNA-Transcription Factors Axis: A Conundrum in Cancer Research. *Front. Pharmacol.* **2022**, *12*, 784801. https://doi.org/10.3389/fphar.2021.784801.

34. Zaphiropoulos, P.G. Circular RNAs from transcripts of the rat cytochrome P450 2C24 gene: correlation with exon skipping. *Proceedings of the National Academy of Sciences* **1996**, *93*, 6536–6541, [https://www.pnas.org/content/93/13/6536.full.pdf]. https://doi.org/10.1073/pnas.93.13.6536.

35. Zhang, Y.; Zhang, X.O.; Chen, T.; Xiang, J.F.; Yin, Q.F.; Xing, Y.H.; Zhu, S.; Yang, L.; Chen, L.L. Circular Intronic Long Noncoding RNAs. *Molecular Cell* **2013**, *51*, 792–806. https://doi.org/https://doi.org/10.1016/j.molcel.2013.08.017.

36. Suzuki, H.; Zuo, Y.; Wang, J.; Zhang, M.Q.; Malhotra, A.; Mayeda, A. Characterization of RNase R-digested cellular RNA source that consists of lariat and circular RNAs from pre-mRNA splicing. *Nucleic Acids Res.* **2006**, *34*, e63., [16682442]. https://doi.org/10.1093/nar/gkl151.

37. Enuka, Y.; Lauriola, M.; Feldman, M.E.; Sas-Chen, A.; Ulitsky, I.; Yarden, Y. Circular RNAs are long-lived and display only minimal early alterations in response to a growth factor. *Nucleic Acids Res.* **2016**, *44*, 1370–1383, [26657629]. https://doi.org/10.1093/nar/gkv1367.

38. Bachmayr-Heyda, A.; Reiner, A.T.; Auer, K.; Sukhbaatar, N.; Aust, S.; Bachleitner-Hofmann, T.; Mesteri, I.; Grunt, T.W.; Zeillinger, R.; Pils, D. Correlation of circular RNA abundance with proliferation – exemplified with colorectal and ovarian cancer, idiopathic lung fibrosis and normal human tissues - Scientific Reports. *Sci. Rep.* **2015**, *5*, 1–10. https://doi.org/10.1038/srep08057.

39. Lim, M.C.J.; Baird, A.M.; Greene, J.; McNevin, C.; Ronan, K.; Podlesniy, P.; Sheils, O.; Gray, S.G.; McDermott, R.S.; Finn, S.P. hsa_circ_0001275 Is One of a Number of circRNAs Dysregulated in Enzalutamide Resistant Prostate Cancer and Confers Enzalutamide Resistance In Vitro. *Cancers* **2021**, *13*. https://doi.org/10.3390/cancers13246383.

40. Li, Y.; Zheng, Q.; Bao, C.; Li, S.; Guo, W.; Zhao, J.; Chen, D.; Gu, J.; He, X.; Huang, S. Circular RNA is enriched and stable in exosomes: a promising biomarker for cancer diagnosis. *Cell Res.* **2015**, *25*, 981–984. https://doi.org/10.1038/cr.2015.82.

41. Shi, X.; Wang, B.; Feng, X.; Xu, Y.; Lu, K.; Sun, M. circRNAs and Exosomes: A Mysterious Frontier for Human Cancer. *Mol. Ther. Nucleic Acids* **2020**, *19:384-392.*, [31887549]. https://doi.org/10.1016/j.omtn.2019.11.023.

42. Wang, J.; Zhang, Q.; Zhou, S.; Xu, H.; Wang, D.; Feng, J.; Zhao, J.; Zhong, S. Circular RNA expression in exosomes derived from breast cancer cells and patients. *Epigenomics* **2019**. https://doi.org/10.2217/epi-2018-0111.

43. Wu, X.; Shi, M.; Lian, Y.; Zhang, H. Exosomal circRNAs as promising liquid biopsy biomarkers for glioma. *Front. Immunol.* **2023**, *14*, 1039084. https://doi.org/10.3389/fimmu.2023.1039084.

44. Pan, B.; Qin, J.; Liu, X.; He, B.; Wang, X.; Pan, Y.; Sun, H.; Xu, T.; Xu, M.; Chen, X.; et al. Identification of Serum Exosomal hsa-circ-0004771 as a Novel Diagnostic Biomarker of Colorectal Cancer. *Front. Genet.* **2019**, *10*, 486216. https://doi.org/10.3389/fgene.2019.01096.

45. Li, T.; Sun, X.; Chen, L. Exosome circ_0044516 promotes prostate cancer cell proliferation and metastasis as a potential biomarker. *J. Cell. Biochem.* **2020**, *121*, 2118–2126. https://doi.org/10.1002/jcb.28239.

46. Chen, W.; Quan, Y.; Fan, S.; Wang, H.; Liang, J.; Huang, L.; Chen, L.; Liu, Q.; He, P.; Ye, Y. Exosome-transmitted circular RNA hsa_circ_0051443 suppresses hepatocellular carcinoma progression. *Cancer Lett.* **2020**, *475*, 119–128. https://doi.org/10.1016/j.canlet.2020.01.022.

47. Louis, C.; Desoteux, M.; Coulouarn, C. Exosomal circRNAs: new players in the field of cholangiocarcinoma. *Clin. Sci.* **2019**, *133*, 2239–2244. https://doi.org/10.1042/CS20190940.

48. Hansen, T.B.; Wiklund, E.D.; Bramsen, J.B.; Villadsen, S.B.; Statham, A.L.; Clark, S.J.; Kjems, J. miRNA-dependent gene silencing involving Ago2-mediated cleavage of a circular antisense RNA. *EMBO J.* **2011**, *30*, 4414–4422, [21964070]. https://doi.org/10.1038/emboj.2011.359.

49. Li, X.; Zhang, J.L.; Lei, Y.N.; Liu, X.Q.; Xue, W.; Zhang, Y.; Nan, F.; Gao, X.; Zhang, J.; Wei, J.; et al. Linking circular intronic RNA degradation and function in transcription by RNase H1. *Sci. China Life Sci.* **2021**, *64*, 1795–1809. https://doi.org/10.1007/s11427-021-1993-6.

50. Park, O.H.; Ha, H.; Lee, Y.; Boo, S.H.; Kwon, D.H.; Song, H.K.; Kim, Y.K. Endoribonucleolytic Cleavage of m6A-Containing RNAs by RNase P/MRP Complex. *Mol. Cell* **2019**, *74*, 494–5078, [30930054]. https://doi.org/10.1016/j.molcel.2019.02.034.

51. Guria, A.; Sharma, P.; Srikakulam, N.; Baby, A.; Natesan, S.; Pandi, G. Cost-Effective Transcriptome-Wide Profiling of Circular RNAs by the Improved-tdMDA-NGS Method. *Front. Mol. Biosci.* **2022**, *9*, 886366. https://doi.org/10.3389/fmolb.2022.886366.

52. Shi, H.; Zhou, Y.; Jia, E.; Liu, Z.; Pan, M.; Bai, Y.; Zhao, X.; Ge, Q. Comparative analysis of circular RNA enrichment methods. *RNA Biol.* **2022**, *19*, 55–67. https://doi.org/10.1080/15476286.2021.2012632.

53. Panda, A.C.; De, S.; Grammatikakis, I.; Munk, R.; Yang, X.; Piao, Y.; Dudekula, D.B.; Abdelmohsen, K.; Gorospe, M. High-purity circular RNA isolation method (RPAD) reveals vast collection of intronic circRNAs. *Nucleic Acids Res.* **2017**, *45*, e116. https://doi.org/10.1093/nar/gkx297.

54. Pandey, P.R.; Rout, P.K.; Das, A.; Gorospe, M.; Panda, A.C. RPAD (RNase R treatment, polyadenylation, and poly(A)+ RNA depletion) method to isolate highly pure circular RNA. *Methods* **2019**, *155*, 41–48, [30391514]. https://doi.org/10.1016/j.ymeth.2018.10.022.

55. Cheng, J.; Metge, F.; Dieterich, C. Specific identification and quantification of circular RNAs from sequencing data. *Bioinformatics* **2016**, *32*, 1094–1096. https://doi.org/10.1093/bioinformatics/btv656.

56. Cocquet, J.; Chong, A.; Zhang, G.; Veitia, R.A. Reverse transcriptase template switching and false alternative transcripts. *Genomics* **2006**, *88*, 127–131. https://doi.org/10.1016/j.ygeno.2005.12.013.

57. Roy, C.K.; Olson, S.; Graveley, B.R.; Zamore, P.D.; Moore, M.J. Assessing long-distance RNA sequence connectivity via RNA-templated DNA–DNA ligation. *eLife* **2015**, *4*. https://doi.org/10.7554/eLife.03700.

58. Salzman, J. RNA Isoform Discovery Through Goodness of Fit Diagnostics. In *Statistical Analysis of Next Generation Sequencing Data*; Springer: Cham, Switzerland, 2014; pp. 261–276. https://doi.org/10.1007/978-3-319-07212-8_13.

59. Song, X.; Zhang, N.; Han, P.; Moon, B.S.; Lai, R.K.; Wang, K.; Lu, W. Circular RNA profile in gliomas revealed by identification tool UROBORUS. *Nucleic Acids Res.* **2016**, *44*, e87. https://doi.org/10.1093/nar/gkw075.

60. Izuogu, O.G.; Alhasan, A.A.; Alafghani, H.M.; Santibanez-Koref, M.; Elliott, D.J.; Jackson, M.S. PTESFinder: a computational method to identify post-transcriptional exon shuffling (PTES) events. *BMC Bioinf.* **2016**, *17*, 1–11. https://doi.org/10.1186/s12859-016-0881-4.

61. Szabo, L.; Morey, R.; Palpant, N.J.; Wang, P.L.; Afari, N.; Jiang, C.; Parast, M.M.; Murry, C.E.; Laurent, L.C.; Salzman, J. Statistically based splicing detection reveals neural enrichment and tissue-specific induction of circular RNA during human fetal development. *Genome Biol.* **2015**, *16*, 1–26. https://doi.org/10.1186/s13059-015-0690-5.

62. Gao, Y.; Wang, J.; Zhao, F. CIRI: an efficient and unbiased algorithm for de novo circular RNA identification. *Genome Biol.* **2015**, *16*, 1–16. https://doi.org/10.1186/s13059-014-0571-3.

63. Gao, Y.; Wang, J.; Zheng, Y.; Zhang, J.; Chen, S.; Zhao, F. Comprehensive identification of internal structure and alternative splicing events in circular RNAs. *Nat. Commun.* **2016**, *7*, 1–13. https://doi.org/10.1038/ncomms12060.

64. Zhang, J.; Chen, S.; Yang, J.; Zhao, F. Accurate quantification of circular RNAs identifies extensive circular isoform switching events. *Nat. Commun.* **2020**, *11*, 1–14. https://doi.org/10.1038/s41467-019-13840-9.

65. You, X.; Conrad, T.O.F. Acfs: accurate circRNA identification and quantification from RNA-Seq data. *Sci. Rep.* **2016**, *6*, 1–11. https://doi.org/10.1038/srep38820.

66. Zhang, X.O.; Wang, H.B.; Zhang, Y.; Lu, X.; Chen, L.L.; Yang, L. Complementary sequence-mediated exon circularization. *Cell* **2014**, *159*, 134–147, [25242744]. https://doi.org/10.1016/j.cell.2014.09.001.

67. Zhang, X.O.; Dong, R.; Zhang, Y.; Zhang, J.L.; Luo, Z.; Zhang, J.; Chen, L.L.; Yang, L. Diverse alternative back-splicing and alternative splicing landscape of circular RNAs. *Genome Res.* **2016**, *26*, 1277–1287. https://doi.org/10.1101/gr.202895.115.

68. Engström, P.G.; Steijger, T.; Sipos, B.; Grant, G.R.; Kahles, A.; Rätsch, G.; Goldman, N.; Hubbard, T.J.; Harrow, J.; Guigó, R.; et al. Systematic evaluation of spliced alignment programs for RNA-seq data. *Nat. Methods* **2013**, *10*, 1185–1191. https://doi.org/10.1038/nmeth.2722.

69. Chuang, T.J.; Wu, C.S.; Chen, C.Y.; Hung, L.Y.; Chiang, T.W.; Yang, M.Y. NCLscan: accurate identification of non-co-linear transcripts (fusion, trans-splicing and circular RNA) with a good balance between sensitivity and precision. *Nucleic Acids Res.* **2016**, *44*, e29. https://doi.org/10.1093/nar/gkv1013.

70. Hoffmann, S.; Otto, C.; Kurtz, S.; Sharma, C.M.; Khaitovich, P.; Vogel, J.; Stadler, P.F.; Hackermüller, J. Fast mapping of short sequences with mismatches, insertions and deletions using index structures. *PLoS Comput. Biol.* **2009**, *5*, e1000502., [19750212]. https://doi.org/10.1371/journal.pcbi.1000502.

71. Hansen, T.B.; Venø, M.T.; Damgaard, C.K.; Kjems, J. Comparison of circular RNA prediction tools. *Nucleic Acids Res.* **2016**, *44*, e58. https://doi.org/10.1093/nar/gkv1458.

72. Hansen, T.B. Improved circRNA Identification by Combining Prediction Algorithms. *Front. Cell Dev. Biol.* **2018**, *6*. https://doi.org/10.3389/fcell.2018.00020.

73. Gaffo, E.; Bonizzato, A.; Te Kronnie, G.; Bortoluzzi, S. CirComPara: A Multi-Method Comparative Bioinformatics Pipeline to Detect and Study circRNAs from RNA-seq Data. *Non-Coding RNA* **2017**, *3*. https://doi.org/10.3390/ncrna3010008.

74. Gaffo, E.; Buratin, A.; Dal Molin, A.; Bortoluzzi, S. Sensitive, reliable and robust circRNA detection from RNA-seq with CirComPara2. *Briefings Bioinf.* **2022**, *23*. https://doi.org/10.1093/bib/bbab418.

75. Li, L.; Bu, D.; Zhao, Y. CircRNAwrap – a flexible pipeline for circRNA identification, transcript prediction, and abundance estimation. *FEBS Lett.* **2019**, *593*, 1179–1189. https://doi.org/10.1002/1873-3468.13423.

76. Liu, Z.; Ding, H.; She, J.; Chen, C.; Zhang, W.; Yang, E. DEBKS: A Tool to Detect Differentially Expressed Circular RNA. *Genomics Proteomics Bioinformatics* **2021**. https://doi.org/10.1016/j.gpb.2021.01.003.

77. Chen, L.; Wang, F.; Bruggeman, E.C.; Li, C.; Yao, B. circMeta: a unified computational framework for genomic feature annotation and differential expression analysis of circular RNAs. *Bioinformatics* **2020**, *36*, 539–545. https://doi.org/10.1093/bioinformatics/btz606.

78. Aufiero, S.; Reckman, Y.J.; Tijsen, A.J.; Pinto, Y.M.; Creemers, E.E. circRNAprofiler: an R-based computational framework for the downstream analysis of circular RNAs. *BMC Bioinf.* **2020**, *21*, 1–9. https://doi.org/10.1186/s12859-020-3500-3.

79. Hossain, Md.T.; Peng, Y.; Feng, S.; Wei, Y. FcircSEC: An R Package for Full Length circRNA Sequence Extraction and Classification. *Int. J. Genomics* **2020**, *2020*, 9084901. https://doi.org/10.1155/2020/9084901.

80. Digby, B.; Finn, S.P.; Ó Broin, P. nf-core/circrna: a portable workflow for the quantification, miRNA target prediction and differential expression analysis of circular RNAs. *BMC Bioinf.* **2023**, *24*, 1–15. https://doi.org/10.1186/s12859-022-05125-8.

81. Metge, F.; Czaja-Hasse, L.F.; Reinhardt, R.; Dieterich, C. FUCHS—towards full circular RNA characterization using RNAseq. *PeerJ* **2017**, *5*, e2934. https://doi.org/10.7717/peerj.2934.

82. Gao, Y.; Wang, J.; Zheng, Y.; Zhang, J.; Chen, S.; Zhao, F. Comprehensive identification of internal structure and alternative splicing events in circular RNAs. *Nat. Commun.* **2016**, *7*, 1–13. https://doi.org/10.1038/ncomms12060.

83. Gao, Y.; Wang, J.; Zheng, Y.; Zhang, J.; Chen, S.; Zhao, F. Comprehensive identification of internal structure and alternative splicing events in circular RNAs. *Nat. Commun.* **2016**, *7*, 1–13. https://doi.org/10.1038/ncomms12060.

84. Ivanov, A.; Memczak, S.; Wyler, E.; Torti, F.; Porath, H.T.; Orejuela, M.R.; Piechotta, M.; Levanon, E.Y.; Landthaler, M.; Dieterich, C.; et al. Analysis of intron sequences reveals hallmarks of circular RNA biogenesis in animals. *Cell Rep.* **2015**, *10*, 170–177, [25558066]. https://doi.org/10.1016/j.celrep.2014.12.019.

85. Dubin, R.A.; Kazmi, M.A.; Ostrer, H. Inverted repeats are necessary for circularization of the mouse testis Sry transcript. *Gene* **1995**, *167*, 245–248, [8566785]. https://doi.org/10.1016/0378-1119(95)00639-7.

86. Thomas, L.F.; Sætrom, P. Circular RNAs are depleted of polymorphisms at microRNA binding sites. *Bioinformatics* **2014**, *30*, 2243–2246, [24764460]. https://doi.org/10.1093/bioinformatics/btu257.

87. Chen, L.; Zhang, Y.H.; Huang, G.; Pan, X.; Wang, S.; Huang, T.; Cai, Y.D. Discriminating cirRNAs from other lncRNAs using a hierarchical extreme learning machine (H-ELM) algorithm with feature selection. *Mol. Genet. Genomics* **2018**, *293*, 137–149. https://doi.org/10.1007/s00438-017-1372-7.

88. Wang, X.; Liu, Y.; Li, J.; Wang, G. StackCirRNAPred: computational classification of long circRNA from other lncRNA based on stacking strategy. *BMC Bioinf.* **2022**, *23*, 1–15. https://doi.org/10.1186/s12859-022-05118-7.

89. Pan, X.; Xiong, K. PredcircRNA: computational classification of circular RNA from other long non-coding RNA using hybrid features. *Mol. Biosyst.* **2015**, *11*, 2219–2226, [26028480]. https://doi.org/10.1039/c5mb00214a.

90. Chaabane, M.; Williams, R.M.; Stephens, A.T.; Park, J.W. circDeep: deep learning approach for circular RNA classification from other long non-coding RNA. *Bioinformatics* **2020**, *36*, 73–80. https://doi.org/10.1093/bioinformatics/btz537.

91. Wang, J.; Wang, L. Deep learning of the back-splicing code for circular RNA formation. *Bioinformatics* **2019**, *35*, 5235–5242, [31077303]. https://doi.org/10.1093/bioinformatics/btz382.

92. Jiang, J.Y.; Ju, C.J.T.; Hao, J.; Chen, M.; Wang, W. JEDI: circular RNA prediction based on junction encoders and deep interaction among splice sites. *Bioinformatics* **2021**, *37*, i289–i298. https://doi.org/10.1093/bioinformatics/btab288.

93. Sekar, S.; Geiger, P.; Adkins, J.; Tassone, E.; Serrano, G.; Beach, T.G.; Liang, W.S. ACValidator: A novel assembly-based approach for in silico verification of circular RNAs. *Biol. Methods Protoc.* **2020**, *5*. https://doi.org/10.1093/biomethods/bpaa010.

94. Yu, S.H.; Vogel, J.; Förstner, K.U. ANNOgesic: a Swiss army knife for the RNA-seq based annotation of bacterial/archaeal genomes. *GigaScience* **2018**, *7*, giy096. https://doi.org/10.1093/gigascience/giy096.

95. Zhou, C.; Molinie, B.; Daneshvar, K.; Pondick, J.V.; Wang, J.; Van Wittenberghe, N.O.; Xing, Y.; Giallourakis, C.C.; Mullen, A.C. Genome-wide maps of m6A circRNAs identify widespread and cell-type-specific methylation patterns that are distinct from mRNAs. *Cell Rep.* **2017**, *20*, 2262. https://doi.org/10.1016/j.celrep.2017.08.027.

96. Menzel, P.; Meyer, I.M. BIQ: A method for searching circular RNAs in transcriptome databases by indexing backsplice junctions. *bioRxiv* **2019**, p. 556993, [556993]. https://doi.org/10.1101/556993.

97. Wu, J.; Li, Y.; Wang, C.; Cui, Y.; Xu, T.; Wang, C.; Wang, X.; Sha, J.; Jiang, B.; Wang, K.; et al. CircAST: Full-length Assembly and Quantification of Alternatively Spliced Isoforms in Circular RNAs. *Genomics Proteomics Bioinformatics* **2019**, *17*, 522–534. https://doi.org/10.1016/j.gpb.2019.03.004.

98. Li, X.; Wu, Y. Detecting circular RNA from high-throughput sequence data with de Bruijn graph. *BMC Genomics* **2020**, *21*. https://doi.org/10.1186/s12864-019-6154-7.

99. Ma, X.K.; Wang, M.R.; Liu, C.X.; Dong, R.; Carmichael, G.G.; Chen, L.L.; Yang, L. CIRCexplorer3: A CLEAR Pipeline for Direct Comparison of Circular and Linear RNA Expression. *Genomics Proteomics Bioinformatics* **2019**, *17*, 511–521. https://doi.org/10.1016/j.gpb.2019.11.004.

100. Liu, Z.; Tao, C.; Li, S.; Du, M.; Bai, Y.; Hu, X.; Li, Y.; Chen, J.; Yang, E. circFL-seq reveals full-length circular RNAs with rolling circular reverse transcription and nanopore sequencing. *eLife* **2021**. https://doi.org/10.7554/eLife.69457.

101. Zhang, G.; Deng, Y.; Liu, Q.; Ye, B.; Dai, Z.; Chen, Y.; Dai, X. Identifying Circular RNA and Predicting Its Regulatory Interactions by Machine Learning. *Front. Genet.* **2020**, *11*. https://doi.org/10.3389/fgene.2020.00655.

102. Li, X.; Chu, C.; Pei, J.; Măndoiu, I.; Wu, Y. CircMarker: a fast and accurate algorithm for circular RNA detection. *BMC Genomics* **2018**, *19*, 79–87. https://doi.org/10.1186/s12864-018-4926-0.

103. Meng, X.; Chen, Q.; Zhang, P.; Chen, M. CircPro: an integrated tool for the identification of circRNAs with protein-coding potential. *Bioinformatics* **2017**, *33*, 3314–3316. https://doi.org/10.1093/bioinformatics/btx446.

104. Jia, G.y.; Wang, D.l.; Xue, M.z.; Liu, Y.w.; Pei, Y.c.; Yang, Y.q.; Xu, J.m.; Liang, Y.c.; Wang, P. CircRNAFisher: a systematic computational approach for de novo circular RNA identification. *Acta Pharmacol. Sin.* **2019**, *40*, 55. https://doi.org/10.1038/s41401-018-0063-1.

105. Jakobi, T.; Uvarovskii, A.; Dieterich, C. circtools-a one-stop software solution for circular RNA research. *Bioinformatics* **2019**, *35*, 2326–2328, [30462173]. https://doi.org/10.1093/bioinformatics/bty948.

106. Zhang, J.; Hou, L.; Zuo, Z.; Ji, P.; Zhang, X.; Xue, Y.; Zhao, F. Comprehensive profiling of circular RNAs with nanopore sequencing and CIRI-long. *Nat. Biotechnol.* **2021**, *39*, 836–845. https://doi.org/10.1038/s41587-021-00842-6.

107. Wang, Z.; Lei, X. A web server for identifying circRNA-RBP variable-length binding sites based on stacked generalization ensemble deep learning network. *Methods* **2022**, *205:179-190.*, [35810958]. https://doi.org/10.1016/j.ymeth.2022.06.014.

108. Niu, M.; Zhang, J.; Li, Y.; Wang, C.; Liu, Z.; Ding, H.; Zou, Q.; Ma, Q. CirRNAPL: A web server for the identification of circRNA based on extreme learning machine. *Comput. Struct. Biotechnol. J.* **2020**, *18*, 834–842. https://doi.org/10.1016/j.csbj.2020.03.028.

109. Zhang, K.; Pan, X.; Yang, Y.; Shen, H.B. CRIP: predicting circRNA–RBP-binding sites using a codon-based encoding and hybrid deep neural networks. *Cold Spring Harbor Lab* **2019**. https://doi.org/10.1261/rna.070565.119.

110. Stefanov, S.R.; Meyer, I.M. CYCLeR—a novel tool for the full isoform assembly and quantification of circRNAs. *Nucleic Acids Res.* **2023**, *51*, e10. https://doi.org/10.1093/nar/gkac1100.

111. Rozowsky, J.; Kitchen, R.R.; Park, J.J.; Galeev, T.R.; Diao, J.; Warrell, J.; Thistlethwaite, W.; Subramanian, S.L.; Milosavljevic, A.; Gerstein, M. exceRpt: A Comprehensive Analytic Platform for Extracellular RNA Profiling. *Cell Syst.* **2019**, *8*, 352–3573, [30956140]. https://doi.org/10.1016/j.cels.2019.03.004.

112. Wang, D. hppRNA—a Snakemake-based handy parameter-free pipeline for RNA-Seq analysis of numerous samples. *Briefings Bioinf.* **2018**, *19*, 622–626. https://doi.org/10.1093/bib/bbw143.

113. Xin, R.; Gao, Y.; Gao, Y.; Wang, R.; Kadash-Edmondson, K.E.; Liu, B.; Wang, Y.; Lin, L.; Xing, Y. isoCirc catalogs full-length circular RNA isoforms in human transcriptomes. *Nat. Commun.* **2021**, *12*, 1–11. https://doi.org/10.1038/s41467-020-20459-8.

114. Wang, K.; Singh, D.; Zeng, Z.; Coleman, S.J.; Huang, Y.; Savich, G.L.; He, X.; Mieczkowski, P.; Grimm, S.A.; Perou, C.M.; et al. MapSplice: Accurate mapping of RNA-seq reads for splice junction discovery. *Nucleic Acids Res.* **2010**, *38*, e178. https://doi.org/10.1093/nar/gkq622.

115. Andrés-León, E.; Núñez-Torres, R.; Rojas, A.M. miARma-Seq: a comprehensive tool for miRNA, mRNA and circRNA analysis. *Sci. Rep.* **2016**, *6*, 1–8. https://doi.org/10.1038/srep25749.

116. Chen, C.Y.; Chuang, T.J. NCLcomparator: systematically post-screening non-co-linear transcripts (circular, trans-spliced, or fusion RNAs) identified from various detectors. *BMC Bioinf.* **2019**, *20*, 1–11. https://doi.org/10.1186/s12859-018-2589-0.

117. Yin, S.; Tian, X.; Zhang, J.; Sun, P.; Li, G. PCirc: random forest-based plant circRNA identification software. *BMC Bioinf.* **2021**, *22*, 1–14. https://doi.org/10.1186/s12859-020-03944-1.

118. Chen, L.; Yu, Y.; Zhang, X.; Liu, C.; Ye, C.; Fan, L. PcircRNA_finder: a software for circRNA prediction in plants. *Bioinformatics* **2016**, *32*, 3528–3529. https://doi.org/10.1093/bioinformatics/btw496.

119. Gao, Y.; Wang, H.; Zhang, H.; Wang, Y.; Chen, J.; Gu, L. PRAPI: post-transcriptional regulation analysis pipeline for Iso-Seq. *Bioinformatics* **2018**, *34*, 1580–1582. https://doi.org/10.1093/bioinformatics/btx830.

120. Liu, Z.; Han, J.; Lv, H.; Liu, J.; Liu, R. Computational identification of circular RNAs based on conformational and thermodynamic properties in the flanking introns. *Comput. Biol. Chem.* **2016**, *61*, 221–225. https://doi.org/10.1016/j.compbiolchem.2016.02.003.

121. Li, L.; Zheng, Y.C.; Kayani, M.U.R.; Xu, W.; Wang, G.Q.; Sun, P.; Ao, N.; Zhang, L.N.; Gu, Z.Q.; Wu, L.C.; et al. Comprehensive analysis of circRNA expression profiles in humans by RAISE. *Int. J. Oncol.* **2017**, *51*, 1625–1638. https://doi.org/10.3892/ijo.2017.4162.

122. Mangul, S.; Yang, H.T.; Strauli, N.; Gruhl, F.; Porath, H.T.; Hsieh, K.; Chen, L.; Daley, T.; Christenson, S.; Wesolowska-Andersen, A.; et al. ROP: dumpster diving in RNA-sequencing to find the source of 1 trillion reads across diverse adult human tissues. *Genome Biol.* **2018**, *19*, 1–12. https://doi.org/10.1186/s13059-018-1403-7.

123. Akers, N.K.; Schadt, E.E.; Losic, B. STAR Chimeric Post for rapid detection of circular RNA and fusion transcripts. *Bioinformatics* **2018**, *34*, 2364–2370. https://doi.org/10.1093/bioinformatics/bty091.

124. Humphreys, D.T.; Fossat, N.; Demuth, M.; Tam, P.P.L.; Ho, J.W.K. Ularcirc: visualization and enhanced analysis of circular RNAs via back and canonical forward splicing. *Nucleic Acids Res.* **2019**, *47*, e123. https://doi.org/10.1093/nar/gkz718.

125. Love, M.I.; Huber, W.; Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **2014**, *15*, 1–21. https://doi.org/10.1186/s13059-014-0550-8.

126. Robinson, M.D.; McCarthy, D.J.; Smyth, G.K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **2010**, *26*, 139. https://doi.org/10.1093/bioinformatics/btp616.

127. Law, C.W.; Chen, Y.; Shi, W.; Smyth, G.K. voom: precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol.* **2014**, *15*, 1–17. https://doi.org/10.1186/gb-2014-15-2-r29.

128. Buratin, A.; Romualdi, C.; Bortoluzzi, S.; Gaffo, E. Detecting differentially expressed circular RNAs from multiple quantification methods using a generalized linear mixed model. *Comput. Struct. Biotechnol. J.* **2022**, *20*, 2495. https://doi.org/10.1016/j.csbj.2022.05.026.

129. Glažar, P.; Papavasileiou, P.; Rajewsky, N. circBase: a database for circular RNAs. *RNA* **2014**. https://doi.org/10.1261/rna.043687.113.

130. Xia, S.; Feng, J.; Chen, K.; Ma, Y.; Gong, J.; Cai, F.; Jin, Y.; Gao, Y.; Xia, L.; Chang, H.; et al. CSCD: a database for cancer-specific circular RNAs. *Nucleic Acids Res.* **2018**, *46*, 925–929, [29036403]. https://doi.org/10.1093/nar/gkx863.

131. Li, J.H.; Liu, S.; Zhou, H.; Qu, L.H.; Yang, J.H. starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein–RNA interaction networks from large-scale CLIP-Seq data. *Nucleic Acids Res.* **2014**, *42*, D92–D97. https://doi.org/10.1093/nar/gkt1248.

132. Kozomara, A.; Birgaoanu, M.; Griffiths-Jones, S. miRBase: from microRNA sequences to function. *Nucleic Acids Res.* **2019**, *47*, D155–D162. https://doi.org/10.1093/nar/gky1141.

133. Hsu, S.D.; Lin, F.M.; Wu, W.Y.; Liang, C.; Huang, W.C.; Chan, W.L.; Tsai, W.T.; Chen, G.Z.; Lee, C.J.; Chiu, C.M.; et al. miRTarBase: a database curates experimentally validated microRNA-target interactions. *Nucleic Acids Res.* **2011**, *39*, issue, [21071411]. https://doi.org/10.1093/nar/gkq1107.

134. Chang, L.; Zhou, G.; Soufan, O.; Xia, J. miRNet 2.0: network-based visual analytics for miRNA functional analysis and systems biology. *Nucleic Acids Res.* **2020**, *48*, W244–W251. https://doi.org/10.1093/nar/gkaa467.

135. Agarwal, V.; Bell, G.W.; Nam, J.W.; Bartel, D.P. Predicting effective microRNA target sites in mammalian mRNAs. *eLife* **2015**. https://doi.org/10.7554/eLife.05005.

136. Enright, A.J.; John, B.; Gaul, U.; Tuschl, T.; Sander, C.; Marks, D.S. MicroRNA targets in Drosophila. *Genome Biol.* **2003**, *5*, 1–14. https://doi.org/10.1186/gb-2003-5-1-r1.

137. Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N.S.; Wang, J.T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res.* **2003**, *13*, 2498. https://doi.org/10.1101/gr.1239303.

138. Conn, S.J.; Pillman, K.A.; Toubia, J.; Conn, V.M.; Salmanidis, M.; Phillips, C.A.; Roslan, S.; Schreiber, A.W.; Gregory, P.A.; Goodall, G.J. The RNA Binding Protein Quaking Regulates Formation of circRNAs. *Cell* **2015**, *160*, 1125–1134. https://doi.org/10.1016/j.cell.2015.02.014.

139. Errichelli, L.; Dini Modigliani, S.; Laneve, P.; Colantoni, A.; Legnini, I.; Capauto, D.; Rosa, A.; De Santis, R.; Scarfò, R.; Peruzzi, G.; et al. FUS affects circular RNA expression in murine embryonic stem cell-derived motor neurons. *Nat. Commun.* **2017**, *8*, 1–11. https://doi.org/10.1038/ncomms14741.

140. Fei, T.; Chen, Y.; Xiao, T.; Li, W.; Cato, L.; Zhang, P.; Cotter, M.B.; Bowden, M.; Lis, R.T.; Zhao, S.G.; et al. Genome-wide CRISPR screen identifies HNRNPL as a prostate cancer dependency regulating RNA splicing. *Proc. Natl. Acad. Sci. U.S.A.* **2017**, *114*, E5207–E5215. https://doi.org/10.1073/pnas.1617467114.

141. Khan, M.A.F.; Reckman, Y.J.; Aufiero, S.; van den Hoogenhof, M.M.G.; van der Made, I.; Beqqali, A.; Koolbergen, D.R.; Rasmussen, T.B.; van der Velden, J.; Creemers, E.E.; et al. RBM20 Regulates Circular RNA Production From the Titin Gene. *Circ. Res.* **2016**.

142. Aktaş, T.; Avşar Ilık, İ.; Maticzka, D.; Bhardwaj, V.; Pessoa Rodrigues, C.; Mittler, G.; Manke, T.; Backofen, R.; Akhtar, A. DHX9 suppresses RNA processing defects originating from the Alu invasion of the human genome. *Nature* **2017**, *544*, 115–119. https://doi.org/10.1038/nature21715.

143. Abdelmohsen, K.; Panda, A.C.; Munk, R.; Grammatikakis, I.; Dudekula, D.B.; De, S.; Kim, J.; Noh, J.H.; Kim, K.M.; Martindale, J.L.; et al. Identification of HuR target circular RNAs uncovers suppression of PABPN1 translation by CircPABPN1. *RNA Biol.* **2017**, pp. 361–369.

144. Du, W.W.; Yang, W.; Liu, E.; Yang, Z.; Dhaliwal, P.; Yang, B.B. Foxo3 circular RNA retards cell cycle progression via forming ternary complexes with p21 and CDK2. *Nucleic Acids Res.* **2016**, *44*, 2846–2858. https://doi.org/10.1093/nar/gkw027.

145. Zhang, M.; Wang, T.; Xiao, G.; Xie, Y. Large-Scale Profiling of RBP-circRNA Interactions from Public CLIP-Seq Datasets. *Genes* **2020**, *11*, 54. https://doi.org/10.3390/genes11010054.

146. Dudekula, D.B.; Panda, A.C.; Grammatikakis, I.; De, S.; Abdelmohsen, K.; Gorospe, M. CircInteractome: A web tool for exploring circular RNAs and their interacting proteins and microRNAs. *RNA Biol.* **2016**, *13*, 34–42, [26669964]. https://doi.org/10.1080/15476286.2015.1128065.

147. Maass, P.G.; Glažar, P.; Memczak, S.; Dittmar, G.; Hollfinger, I.; Schreyer, L.; Sauer, A.V.; Toka, O.; Aiuti, A.; Luft, F.C.; et al. A map of human circular RNAs in clinically relevant tissues. *J. Mol. Med.* **2017**, *95*, 1179–1189. https://doi.org/10.1007/s00109-017-1582-9.

148. Dong, R.; Ma, X.K.; Li, G.W.; Yang, L. CIRCpedia v2: An Updated Database for Comprehensive Circular RNA Annotation and Expression Comparison. *Genomics Proteomics Bioinformatics* **2018**, *16*, 226–233, [30172046]. https://doi.org/10.1016/j.gpb.2018.08.001.

149. Feng, J.; Chen, W.; Dong, X.; Wang, J.; Mei, X.; Deng, J.; Yang, S.; Zhuo, C.; Huang, X.; Shao, L.; et al. CSCD2: an integrated interactional database of cancer-specific circular RNAs. *Nucleic Acids Res.* **2022**, *50*, D1179–D1183. https://doi.org/10.1093/nar/gkab830.

150. Chen, X.; Han, P.; Zhou, T.; Guo, X.; Song, X.; Li, Y. circRNADb: A comprehensive database for human circular RNAs with protein-coding annotations. *Sci. Rep.* **2016**, *6*, 1–6. https://doi.org/10.1038/srep34985.

151. Tang, Z.; Li, X.; Zhao, J.; Qian, F.; Feng, C.; Li, Y.; Zhang, J.; Jiang, Y.; Yang, Y.; Wang, Q.; et al. TRCirc: a resource for transcriptional regulation information of circRNAs. *Briefings Bioinf.* **2019**, *20*, 2327–2333, [30184150]. https://doi.org/10.1093/bib/bby083.

152. Yao, D.; Zhang, L.; Zheng, M.; Sun, X.; Lu, Y.; Liu, P. Circ2Disease: a manually curated database of experimentally validated circRNAs in human disease. *Sci. Rep.* **2018**, *8*, 1–6. https://doi.org/10.1038/s41598-018-29360-3.

153. Chen, L.L.; Bindereif, A.; Bozzoni, I.; Chang, H.Y.; Matera, A.G.; Gorospe, M.; Hansen, T.B.; Kjems, J.; Ma, X.K.; Pek, J.W.; et al. A guide to naming eukaryotic circular RNAs. *Nat. Cell Biol.* **2023**, *25*, 1–5. https://doi.org/10.1038/s41556-022-01066-9.

154. Martin, F.J.; Amode, M.R.; Aneja, A.; Austine-Orimoloye, O.; Azov, A.G.; Barnes, I.; Becker, A.; Bennett, R.; Berry, A.; Bhai, J.; et al. Ensembl 2023. *Nucleic Acids Res.* **2023**, *51*, D933–D941. https://doi.org/10.1093/nar/gkac958.

155. O'Leary, N.A.; Wright, M.W.; Brister, J.R.; Ciufo, S.; Haddad, D.; McVeigh, R.; Rajput, B.; Robbertse, B.; Smith-White, B.; Ako-Adjei, D.; et al. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res.* **2016**, *44*, 733–745, [26553804]. https://doi.org/10.1093/nar/gkv1189.

156. Ye, J.; Wang, L.; Li, S.; Zhang, Q.; Zhang, Q.; Tang, W.; Wang, K.; Song, K.; Sablok, G.; Sun, X.; et al. AtCircDB: a tissue-specific database for Arabidopsis circular RNAs. *Briefings Bioinf.* **2019**, *20*, 58–65, [28968841]. https://doi.org/10.1093/bib/bbx089.

157. Wu, W.; Ji, P.; Zhao, F. CircAtlas: an integrated resource of one million highly accurate circular RNAs from 1070 vertebrate transcriptomes. *Genome Biol.* **2020**, *21*, 1–14. https://doi.org/10.1186/s13059-020-02018-y.

158. Liu, M.; Wang, Q.; Shen, J.; Yang, B.B.; Ding, X. Circbank: a comprehensive database for circRNA with standard nomenclature. *RNA Biol.* **2019**, *16*, 899. https://doi.org/10.1080/15476286.2019.1600395.

159. Chen, Y.; Yao, L.; Tang, Y.; Jhong, J.H.; Wan, J.; Chang, J.; Cui, S.; Luo, Y.; Cai, X.; Li, W.; et al. CircNet 2.0: an updated database for exploring circular RNA regulatory networks in cancers. *Nucleic Acids Res.* **2022**, *50*, 93–101, [34850139]. https://doi.org/10.1093/nar/gkab1036.

160. Meng, X.; Hu, D.; Zhang, P.; Chen, Q.; Chen, M. CircFunBase: a database for functional circular RNAs. *Database (Oxford).* **2019**, *2019:baz003.*, [30715276]. https://doi.org/10.1093/database/baz003.

161. Ruan, H.; Xiang, Y.; Ko, J.; Li, S.; Jing, Y.; Zhu, X.; Ye, Y.; Zhang, Z.; Mills, T.; Feng, J.; et al. Comprehensive characterization of circular RNAs in ~ 1000 human cancer cell lines. *Genome Med.* **2019**, *11*, 1–14. https://doi.org/10.1186/s13073-019-0663-5.

162. Fan, C.; Lei, X.; Tie, J.; Zhang, Y.; Wu, F.X.; Pan, Y. CircR2Disease v2.0: An Updated Web Server for Experimentally Validated circRNA–disease Associations and Its Application. *Genomics Proteomics Bioinformatics* **2022**, *20*, 435–445. https://doi.org/10.1016/j.gpb.2021.10.002.

163. Ghosal, S.; Das, S.; Sen, R.; Basak, P.; Chakrabarti, J. Circ2Traits: a comprehensive database for circular RNA potentially associated with disease and traits. *Front. Genet.* **2013**, *4*, 68415. https://doi.org/10.3389/fgene.2013.00283.

164. Sun, Z.Y.; Yang, C.L.; Huang, L.J.; Mo, Z.C.; Zhang, K.N.; Fan, W.H.; Wang, K.Y.; Wu, F.; Wang, J.G.; Meng, F.L.; et al. circRNADisease v2.0: an updated resource for high-quality experimentally supported circRNA-disease associations. *Nucleic Acids Res.* **2023**, p. gkad949. https://doi.org/10.1093/nar/gkad949.

165. Wang, K.; Wang, C.; Guo, B.; Song, K.; Shi, C.; Jiang, X.; Wang, K.; Tan, Y.; Wang, L.; Wang, L.; et al. CropCircDB: a comprehensive circular RNA resource for crops in response to abiotic stress. *Database* **2019**, *2019*, baz053. https://doi.org/10.1093/database/baz053.

166. Zheng, L.L.; Li, J.H.; Wu, J.; Sun, W.J.; Liu, S.; Wang, Z.L.; Zhou, H.; Yang, J.H.; Qu, L.H. deepBase v2.0: identification, expression, evolution and function of small RNAs, LncRNAs and circular RNAs from deep-sequencing data. *Nucleic Acids Res.* **2016**, *44*, 196–202, [26590255]. https://doi.org/10.1093/nar/gkv1273.

167. Lai, H.; Li, Y.; Zhang, H.; Hu, J.; Liao, J.; Su, Y.; Li, Q.; Chen, B.; Li, C.; Wang, Z.; et al. exoRBase 2.0: an atlas of mRNA, lncRNA and circRNA in extracellular vesicles from human biofluids. *Nucleic Acids Res.* **2022**, *50*, 118–128, [34918744]. https://doi.org/10.1093/nar/gkab1085.

168. Vo, J.N.; Cieslik, M.; Zhang, Y.; Shukla, S.; Xiao, L.; Zhang, Y.; Wu, Y.M.; Dhanasekaran, S.M.; Engelke, C.G.; Cao, X.; et al. The Landscape of Circular RNA in Cancer. *Cell* **2019**, *176*, 869–88113, [30735636]. https://doi.org/10.1016/j.cell.2018.12.021.

169. Walsh, K.; Gokool, A.; Alinejad-Rokny, H.; Voineagu, I. NeuroCirc: an integrative resource of circular RNA expression in the human brain. *Bioinformatics* **2021**, *37*, 3664–3666, [34028497]. https://doi.org/10.1093/bioinformatics/btab230.

170. Zhang, P.; Meng, X.; Chen, H.; Liu, Y.; Xue, J.; Zhou, Y.; Chen, M. PlantCircNet: a database for plant circRNA–miRNA–mRNA regulatory networks. *Database: The Journal of Biological Databases and Curation* **2017**, *2017*. https://doi.org/10.1093/database/bax089.

171. Chu, Q.; Zhang, X.; Zhu, X.; Liu, C.; Mao, L.; Ye, C.; Zhu, Q.H.; Fan, L. PlantcircBase: A Database for Plant Circular RNAs. *Mol. Plant* **2017**, *10*, 1126–1128, [28315753]. https://doi.org/10.1016/j.molp.2017.03.003.

172. Bray, N.L.; Pimentel, H.; Melsted, P.; Pachter, L. Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol.* **2016**, *34*, 525–527. https://doi.org/10.1038/nbt.3519.

173. Patro, R.; Duggal, G.; Love, M.I.; Irizarry, R.A.; Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* **2017**, *14*, 417–419. https://doi.org/10.1038/nmeth.4197.