

MINI-REVIEW ARTICLE

A Brief Review of circRNA Biogenesis, Detection, and Function

Ying Liang¹, Niannian Liu¹, Le Yang¹, Jianjun Tang¹, Yinglong Wang^{1,*} and Meng Mei²¹College of Computer and Information Engineering, Jiangxi Agricultural University, Nanchang, China; ²Department of Basic teaching, Jiangxi Business School, Nanchang, China

ARTICLE HISTORY

Received: September 25, 2020
Revised: December 06, 2020
Accepted: December 22, 2020DOI:
10.2174/1389202922666210331130722

Abstract: Circular RNA (circRNA) is a non-coding molecule produced through alternative splicing of one or more exons of a gene in the presence of an RNA-induced silencing complex (RISC). Its formation depends on complementary intron sequences on both sides of the circularized sequence. CircRNA functions as a sponge for miRNA, playing the role of the transcriptional regulator or potential biomarker. It has an impact on fetal growth and on synaptic facilitation in the brain. In this review, we illustrate biogenesis mechanisms, characteristics, and functions of circRNAs. We also summarize methods using sequence feature and RNA next-generation sequencing data for circRNA prediction. Finally, we discuss the state of the research on circRNA in diseases, which will bring new contributions to future disease treatments.

Keywords: CircRNA, alternative splicing, miRNA sponge, complementary intron, biomarker, disease.

1. INTRODUCTION

Circular RNA (circRNA) is a non-coding RNA (ncRNA) molecule widely found in eukaryotes. CircRNA has no 5' cap and 3' end poly (A) tail and forms a ring structure by a covalent bond. CircRNA expression is stable because it is not easily degraded by RNA exonucleases [1]. CircRNA has a wide range of origins and tissue specificities and plays a variety of roles in the growth and development of organisms. CircRNA molecules are rich in microRNA (miRNA) binding sites and thus act as miRNA sponges in cells. Such binding hinders the inhibitory effects of miRNAs on their target genes and increases the expression level of these genes [2]. CircRNA was first discovered while studying the potato spindle tuber disease (PSTD) that is mediated by viruses with no protein coats, made of a single-stranded genome that forms a closed ring RNA molecule [3]. Later, using electron microscopy, researchers observed the existence of circRNA in the cytoplasm of eukaryotic cells, during the gene transcription of SRY (Sry circRNA is another highly expressed circRNA found in mouse testis) [4-6]. These findings did not generate much attention at first until the emergence of the next-generation sequencing technology in 2012. Salzman *et al.* discovered that non-linear circular transcripts formed by exon rearrangements existed in normal human primary blood cells, leukemia cells, and HeLa cell lines [7]. Since the first report on circRNA, a large number of circRNA molecules have been identified. Although the presence of circRNAs in human cells has been discovered for more than two decades, it is only recently that their significance in human cells has been appreciated [7-9].

Studies have shown that the expression and role of circRNA are related to the occurrence and development of various human diseases (cardiovascular diseases, diabetes mellitus, cancer), and also related to plant diseases, such as the potato bacterial soft rot. CircRNAs are also related to biological tissue development and cell aging. Expression of specific circRNAs in different biological samples makes them ideal biomarkers for disease diagnosis and to assess tissue development. The elucidation of their mechanisms of action driving diseases could also designate circRNAs as therapeutic targets [10].

2. BIOGENESIS, CHARACTERISTICS, AND FUNCTIONS OF CircRNA

2.1. Biogenesis

According to gene structure annotation information, there are five main origins of circRNAs. These origins are: 1) exon-only circular RNA: Exon circRNA [7]; 2) intron only source: Circular intronic RNA [11]; 3) back-splicing of upstream exons and intron retention: Exon-intron circular RNA [12]; 4) circular RNA from Fusion Gene: f-circRNA [13, 14]; and 5) read-through circular RNA formed by polymerase II: rt-circRNA [15]. Although the origins of circRNAs are diverse, circRNAs are mainly produced by head-to-tail back splicing. The formation of circular RNA is regulated by cis-acting elements and trans-acting elements, through mechanisms that are not completely understood. There are currently three known mechanisms driving the formation of circular RNA Fig. (1).

In flanking intron reverse splicing pair drive Fig. (1A), the intron sequences on both sides of the exon are complementary paired so that the 5' splice donor site of the mRNA precursor directly joins with the 3' splice acceptor site to form a circular RNA. Intron pair drive is also called the di-

*Address correspondence to this author at the College of Computer and Information Engineering, Jiangxi Agricultural University, Nanchang, China; E-mails: aliang1229@126.com, 83577939@qq.com

rect back-splicing model [16]. Studies have found that introns of less than 100 nt with reverse complementary repeat sequences can promote exon circularization. Reverse complementary matches (RCMs), intronic complementary sequences (ICSs), and Alu elements are complementary pairing sequences that are abundant on the flanking introns of circularizable exons. They effectively promote the pairing of flanking introns and circularization [17, 18].

In RNA binding protein drive Fig. (1B), RNA-binding proteins promote the formation of tissue-specific circular RNA. RNA binding proteins participate in the formation of circular RNA by binding specific motifs in flanking intron sequences [19-21]. For example, in *Drosophila melanogaster* and human, the second exon of the MBL/MBNL1 gene can form circular RNA that depends on a specific binding motif on flanking introns in this gene [19]. Heterogeneous nuclear rib nucleoprotein L (HNRNPL) promotes the

formation of circular RNA by binding these flanking sequences [22].

In lasso driven cyclization Fig. (1C), when the mRNA precursor undergoes GU/AG splicing, exon skipping produces an intron-exon-containing lasso intermediate, which then undergoes back-splicing to form a circular RNA. Current studies in *Arabidopsis thaliana*, *Solanum lycopersicum*, and *Zea mays* have shown that circular RNAs driven by lasso are widespread in plants [16, 23, 24].

2.2. Characteristics

CircRNAs have unique characteristics that distinguish them from linear non-coding RNAs. They are widely expressed in human cells, and their expression sometimes exceeds 10 times the level of their linear isomers [7, 8]. They are more stable in the human body than linear RNAs beca-

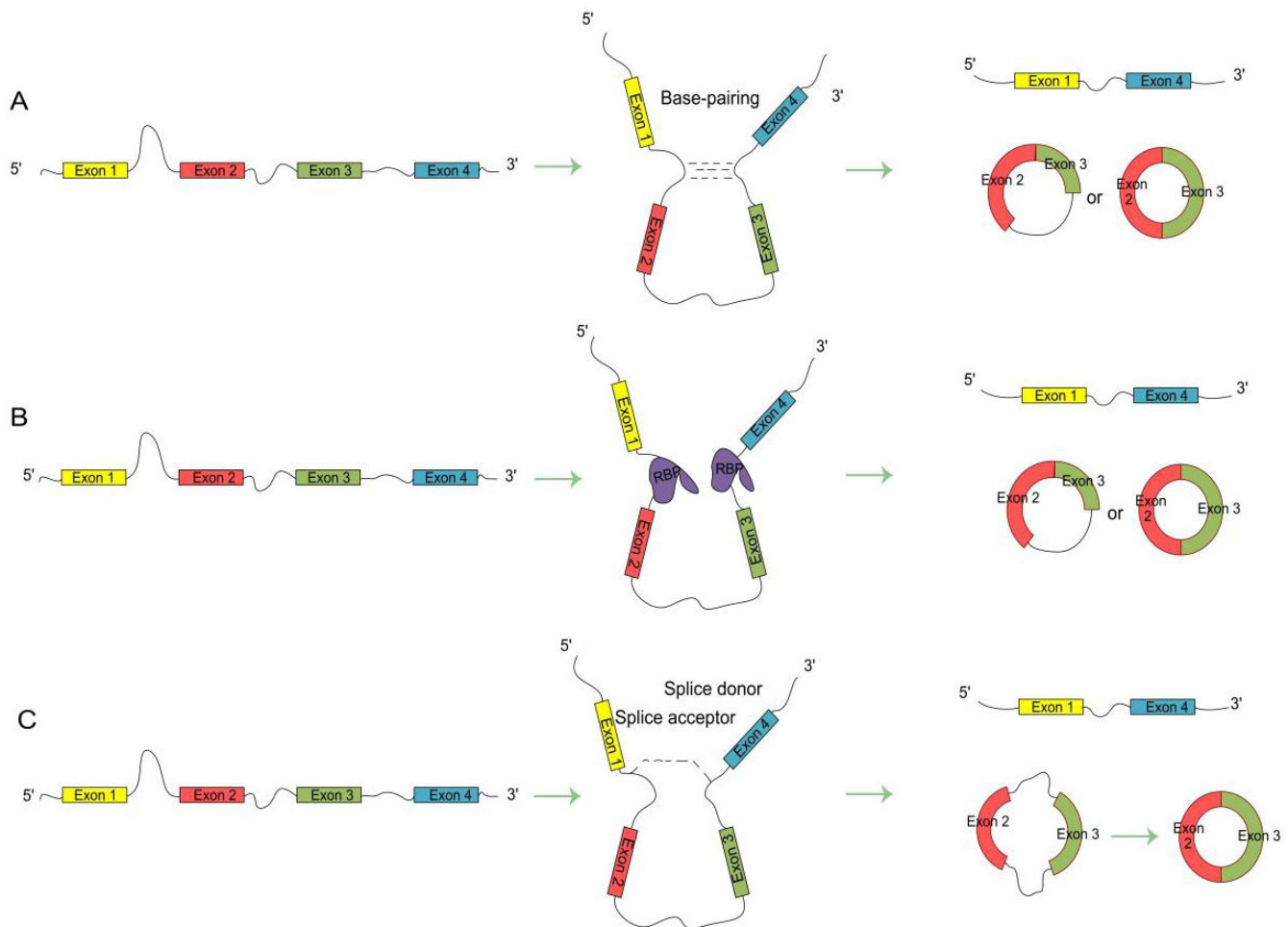


Fig. (1). RNA circularization mechanisms. (A) Introns complement. Complementary pairing elements such as RCMs/ICSs/AIU promote flanking intron pairing and cyclization; (B) RNA binding protein drives cyclization. RNA binding proteins promote cyclization by binding to flanking introns of circularized exons; (C) Lasso driven cyclization. The exon skipping event that occurs during the splicing of mRNA precursor results in the formation of intermediate containing intron and exon, thus forming circular RNA. (A higher resolution / colour version of this figure is available in the electronic copy of the article).

use they are not easily degraded by RNase R [9]. CircRNAs have a tiny response element that can interact with miRNAs to regulate the expression of target genes. Most circRNAs are formed by exons, while a few are from introns only. Most of them are non-coding, highly conserved, and located in the cytoplasm; only a few are located in the nucleus [2]. Most circRNAs can play a regulatory role in transcription and post transcription, only a few of them play a role in transcription [11].

2.3. Functions

CircRNAs have multiple binding sites for miRNA, which results in their sponge-like function for miRNA [25]. In cancers, accumulating evidences show that circRNAs, such as circRNA-Ctfr, for example, act as sponges for miRNA and participate in transcriptional regulation of target genes [26]. Researchers have found such miRNA sponge function in Rice, Wheat, Tomato, Tangerine, and Arabidopsis. Some circRNAs also have one or more binding sites for RNA binding proteins [27].

CircRNAs can regulate the alternative splicing of host genes. In *Arabidopsis thaliana*, circRNA can strongly bind to the DNA locus of the host gene to form an R-loop structure of RNA: DNA heterozygote. The R-loop structure can inhibit the transcription of this region. A cross exon alternative splicing event occurs, which promotes the production of an alternative splicing transcript variant (SEP3.3), affecting the flowering phenotype [28]. In addition, the data from bamboo showed that the frequency of alternative splicing events in genes producing circular RNA was significantly higher than that of randomly selected genes, indicating that circular RNA could regulate the selective splicing process of linear genes [29].

CircRNAs can regulate the expression of host genes. Early studies have shown that EIciRNA can interact with U1 snRNA to form the complex EIciRNA-U1 snRNP. The complex interacts with the PolII transcription complex to promote the expression of host genes [12]. In plants' research, Lu *et al.* also found that circRNAs and their linear isomers can inhibit the post-transcriptional expression of host genes by constructing a genetic transformation system of rice over-expression [30].

CircRNAs have a potential translational function. There are a large number of consistent m6A motifs on circRNAs in human cells. With the participation of various proteins, one m6A site can activate the translational function of circRNA [31]. In *Drosophila melanogaster*, a particular class of circRNA was found to use the promoter of host genes to combine with ribosomes and promote translation [32]. Recently, it has been found that a high concentration of RNA cyclized *in vitro* can induce the expression of proteins in mice, which indicates that circRNA can be used as an effective protein expression tool [33].

CircRNAs can induce the formation of pseudogenes. Pseudogenes are derived from linear mRNAs by reverse transcription and integrated into the genome. Linear mRNAs

can form pseudogenes that have the same sequences as the exons. Stable circRNAs can also be reverse transcribed to produce pseudogenes that have reversed sequences then integrated into the genome [34].

3. DETECTION OF CircRNAs

3.1. Methods Based on RNA Sequencing Data

Since the first report of 80 circRNAs found by Salzman using RNA-Seq sequencing data in 2012, the next-generation sequencing technology has been continuously improved, and multiple ring RNA tools have been generated [7]. RNA next-generation sequencing technology not only generates high throughput and accurate information, but can also identify new circRNAs. It is a cost-effective method and its rapid development in bioinformatics, increased the number of circRNAs that have been found in mammals, humans, and plants [35, 36].

CircRNA is different from linear RNA, which has a closed-loop structure that is more stable than linear RNA, because in contrast to linear RNA, circRNA is not easily degraded by RNase R. Current studies have found that circRNA molecules are rich in miRNA binding sites, which can quench miRNAs and decrease their inhibitory effects on target genes. Detection and sequencing of circRNAs in organisms can be improved by enrichment of low abundance circRNA. This can be achieved by removing rRNAs and by digesting linear RNAs. After construction of the sequence database of circRNAs, analyses of distribution, annotation, expression, conservation, and interaction with miRNA can be carried out to further explore the functions of these circRNAs.

At present, there are relatively few studies and comprehensive annotations on circRNAs. Because the traditional detection technologies could only detect the known circRNAs, information on large number of specifically expressed circRNAs are missing. Next-generation sequencing data is used to identify circRNA, using software that can not only identify known circRNAs but can also predict new circRNAs Table 1. The emergence of RNA next-generation sequencing data makes it possible to sequence hundreds of millions of short fragments, of which single base pair analysis provides a method for accurate and effective identification of circRNAs.

Various software packages can be used to detect circRNAs from RNA-Seq data. KNIFE, NCL Scan and PTES Finder rely on providing gene annotation information to construct circRNA sequences for circRNA detection [39, 44, 45]. Find-circ and UROBORUS can be used in combination to map unmapped reads to the genome, extract the first and last 20 bp anchors from unmapped reads, and then obtain reverse splicing events from the location information of these anchors [9, 41]]. NCL Scan and PTES Finder use the estimated segmented anchor location information to create the putative circRNA sequence [44, 45]. CircRNA-finder, CIRC explorer, DCC, Map splice, and Segemehl detect and analyze alternative splicing events by splicing alignment algo-

Table 1. Detection methods of circRNAs based on sequencing data.

Method	Approach	Dataset or Database	Language	URL
Find-circ [9]	Find-circ takes two 20 bp of the reads that are not mapped and map them to the genome again. Next, the GU/AG cleavage site is found by short sequence alignment to infer the potential circular RNA sequence.	human hg19, mouse mm9, C. elegans ce6, UCSC genome browser	python	https://www.nature.com/articles/nature11928/
CIRI [37]	The fasta sequence of the input genome is compared with the same file generated by the sequencing data to detect junction reads: paired chiasitic clipping signals at the junction points of the covered circular RNA. Compare junction reads. The conservative splicing sites of PEM and GT-AG are filtered, and the dynamic programming algorithm performs the final circular RNA prediction.	ENCODE RNA-seq data	perl	https://doi.org/10.1186/s13059-014-0571-3
DCC [38]	DCC uses the output from the STAR read mapper to systematically detect back-splice junctions in next-generation sequencing data. DCC applies a series of filters and integrates data across replicate sets to arrive at a precise list of circRNA candidates.	rRNA-depleted total RNA-seq data	python	https://doi.org/10.1093/bioinformatics/btv656
KNIFE [39]	Detect and quantify circular and linear RNA splicing events at the annotated and unannotated exon boundaries, including intergenic regions of the genome, thereby improving the sensitivity and specificity of circular RNA detection.	ENCODE poly(A) + and poly(A) - RNA-Seq data	Python/perl	https://link.springer.com/article/10.1186/s13059-015-0690-5
CIRC explorer [40]	Identify the linker reads from the reverse splicing exons, realign the connected reads with the existing gene annotations, determine the precise locations of the splicing sites of the downstream donor and upstream acceptor, and use a custom algorithm to adjust the mapping errors based on RefSeq exon annotations	RNase R- treated RNA-seq from H9 human embryonic stem cells (hESCs)	python	https://doi.org/10.1016/j.cell.2014.09.001
UROBORUS [41]	The artificial paired-end seed (20 bp) is first extracted from two ends of reads in an unmapped.sam file, and then aligned to the reference genome. The UROBORUS pipeline designed algorithm to deal with BMJ and UMJ reads, and detect more circRNA supported reads.	RNA-seq data of glioma samples in Hg19	perl	https://doi.org/10.1093/nar/gkw075
MAP splice [42]	In the 'tag alignment' phase, candidate alignments of the mRNA tags to the reference genome G are determined. In the 'splice inference phase,' splice junctions that appear in the alignments of one or more tags are analyzed to determine a splice significance score based on the quality and diversity of alignments that include the splice.	a synthetic noise-free RNA-seq data	python	https://doi.org/10.1093/nar/gkq622
Segemehl [43]	Segemehl is a single-end RNA-seq data segmentation read mapping algorithm, which combines seed mapping based on error ESA and fast bit vector comparison. It can accommodate multiple splits in one read and does not make a priori assumptions about the transcript structure. It is implemented in the Segemehl mapping tool, which can easily identify conventional splice junctions, collinear and non-collinear fusion transcripts, and trans-spliced RNA	RefSeq database Drosophila RNA-seq dataset human melanoma transcriptome dataset	python	https://link.springer.com/article/10.1186/gb-2014-15-2-r34

(Table 1) contd....

Method	Approach	Dataset or Database	Language	URL
NCL scan [44]	Map RNA sequence reads with reference genomes and known transcripts, eliminate collinearity matching reads, and connect the two ends of each unmapped read to generate a continuous sequence. Use BLAT to align each linked sequence with the reference genome. According to the corresponding BLAT comparison results and GENCODE comments, use the assumed NCL connection points to make a "hypothetical NCL reference"	hg19 /GRCh37	python	https://doi.org/10.1093/nar/gkv1013
PTES Finder [45]	PTES Finder identifies putative PTES structures by mapping RNAseq reads to sequence models generated using existing transcript annotation. It then applies a series of mapping and alignment filters to systematically remove known classes of false positives.	sample SRR4497 5A in human fibroblast total RNA	perl	http://link.springer.com/article/10.1186/s12859-016-0949-1
CircRNA-finder [46]	STAR compares the reference genome and runs circRNA-finder to get the circular RNA file whose splicing site meets the GT-AG splicing signal	Drosophila total RNA-sequencing data	perl	https://doi.org/10.1016/j.celrep.2014.10.062
CircRNAFisher [47]	CircRNAFisher is a systematic calculation pipeline suitable for whole-genome circRNA identification and annotation from scratch. CircRNAFisher combines BSJ search with a series of statistical filters to detect candidate circRNAs. It can also combine BSJ overlapping reading fragments with inconsistent BSJ Read the fragment to estimate the P value of the identified circRNA	A549 and MCF7 in RNA-Seq data from the ENCODE project; human hg19 and UCSC database	perl	https://www.nature.com/articles/s41401-018-0063-1
PcircRNA-finder [48]	PcircRNA-finder collects all backsplice sites by chiasmic clipping mapping of PE reads based on available main fusion detection methods.	rRNA-/RNAase R RNA-Seq data	python	https://doi.org/10.1093/bioinformatics/btw496

rithm [38, 40, 42, 43, 46]. CircRNA-finder needs to pair the end sequencing data and relies on the RNA-Seq splicing alignment software STAR [49]. PcircRNA-Finder is the software specially designed for the identification of plant circRNA, which accurately distinguishes and predicts the downstream donors and upstream receptors of potential circRNA in plants with high sensitivity [50].

Although several software and algorithms are developed for circRNA detection using RNA next-generation sequencing data, there is no uniform measurement standard and there is still a large space for improvement of this software.

3.2. Machine Learning Methods Based on Sequence Feature

The existing circRNA recognition tools that rely on RNA next-generation sequencing data usually have shortcomings of lack of accuracy, low repeatability between different methods, and high false positive or false negative rate [51]. In order to solve this problem, a growing number of researchers use machine learning methods to establish models

to distinguish circRNA according to the inherent characteristics of sequences without relying on sequencing data Table 2. In recent years, the combination of sequence features and machine learning has been successfully applied to the prediction of gene regulatory sites and splicing sites [52].

The contextual regression model is trained to predict the formation of circRNA from a random genomic locus on the human genome, with potential biogenesis factors of circRNA as the features of the training data. The features are inputted into a neural network that generates a contextual weight for each feature that represents the importance of the features. Then, the features are weighted by the corresponding weights to make an easier separation of samples by summing the weighted features to get the prediction [53]. Liu *et al.* uncovers a potential new link between circRNA biogenesis and flanking CpG islands, which suggests a potential correlation between DNA methylation and circRNA biogenesis. The HELM method is to learn a sparse encoder in an unsupervised way and transform the original input into a higher-level representation. The random perturbation of the output matrix is the input for supervised feature classification.

Table 2. Sequence feature approaches for circRNA classification

Method	Approach	Dataset or Database	Language	URL
Context regression Model [53]	The features are inputted into a neural network that generates a contextual weight for each feature representing its importance. Features are then organized by corresponding weights to make an easier separation of samples. In classification or regression tasks, the weighted features are then summed to yield the prediction.	CircNet	python	https://doi.org/10.1093/bioinformatics/btz705
Helm [54]	The goal of HELM is to learn the sparse encoder in an unsupervised way and transform the original input into a higher-level representation. The random disturbance of the output matrix is the input of the supervised feature classification. There is only one hidden layer for supervised feature classification. The goal is to minimize the norm of training error and output weight.	Circbase and GENCODE	python	https://doi.org/10.1007/s00438-017-1372-7
Circdeep [55]	Constructing RCM descriptors that provide the possibility of circularization of flanking sequences and query sequences, integrating ACNN and BLSTM to process different input data types. Integrate shared representations among different modes by fusion of RCM descriptors, ACNN-BLSTM sequence descriptors, and conservation descriptors into high-level abstract descriptors. Add feature fusion learning to heterogeneous descriptors, train a single DNN, and finally predict circRNA.	GENCODE and human circRNAs from the database circRNADB	python	https://doi.org/10.1093/bioinformatics/btz537
PredcircRNA [56]	PredcircRNA focuses on distinguishing circRNAs from other lncRNAs through multiple kernel learning. First, by distinguishing features from different sources extracted from the transcript, and second, by a calculation method based on a multiple kernel learning framework, used to fuse these heterogeneous features. Finally, perform 5-fold cross-validation.	human circRNA data from circbase database and GENCODE	python	https://doi.org/10.1039/c5mb00214a
WebCircRNA[57]	WebCircRNA is a user-friendly web server that predicts if coding and noncoding RNAs have circRNA isoforms and whether circRNAs are expressed in stem cells. The predictions are made by random forest models using sequence-derived features as input. The output scores are converted to fractiles, which are used to assess the circRNA and stem cell potential.	CircBase and GENCODE v19	python	https://doi.org/10.3390/genes9110536
CircLGB[58]	CircLGB, a machine learning-based framework to discriminate circRNA from other lncRNAs. CircLGB integrates commonly used sequence-derived features and three new features containing adenosine to inosine (A-to-I) deamination, A-to-I density and the internal ribosome entry site. circLGB categorizes circRNAs by utilizing a LightGBM classifier with feature selection.	CircBase, LNCipedia, CircRNadb and GENCODE	python	https://doi.org/10.3389/fgene.2020.00655
CirRNAPL[59]	First, the method extracts the features of nucleic acid composition and structure of the circRNA sequence. Four features of the sequence data were extracted, including Ribonucleic acid composition, Autocorrelation, Pseudo-ribonucleic acid composition, and Predicted structure composition. Then, the extreme learning machine based on particle swarm optimization is used as the classification algorithm. The classifier CirRNAPL is constructed by a tenfold cross-validation method to identify the RNA sequence to be labeled.	CircBase and GENCODE v19	Java	https://doi.org/10.1016/j.csbj.2020.03.028
DeepciRGO[60]	DeepciRGO predicts the function of circRNA by integrating multiple biological information related to circRNA. Firstly, constructing a global heterogeneous network according to circRNA co-expressions, circRNA-protein associations, and protein-protein interactions. Then, the latent topological features of the global network are extracted through HIN2Vec and are further fed into a deep neural network classifier. Finally, circRNAs are annotated with GO terms through the trained classifier.	StarBase v2.0 and CSCD	python	https://doi.org/10.1186/s12859-020-03748-3

Only one hidden layer can be used for supervised feature classification. The goal is to minimize the training error and output weight norm [54]. Chaabane *et al.* propose a new method circDeep, to fuses an RCM descriptor, ACNN-BL-

STM sequence descriptor and a conservation descriptor into high-level abstraction descriptors, where the shared representations across different modalities are integrated. The next steps are to add feature fusion learning to heterogeneous de-

scriptors, train a single DNN, and finally predict circRNA [55]. PredcircRNA focuses on distinguishing circRNA from other lncRNAs through multiple kernel learning, the first is to extract features from different sources of the transcript, and then the calculation method based on the multiple kernel learning fuses these heterogeneous features together [56]. For WebCircRNA, Pan *et al.* extracts different features such as sequence composition of *e.g.* RNA secondary structure and conservation to train random forest model, and five-fold cross-validation to assess the performance [57]. CircLGB is a machine-learning framework that includes a Classifier named LightGBM, which has 5 major procedures: (1) Collecting human circRNAs and lncRNAs transcripts, which are combined to construct the circLncRNA dataset; (2) Extracting four groups of sequence-derived features; (3) Sorting the extracted features according to the importance score, and using sequential forward search (SFS) to determine the optimal feature subset; (4) Training the LightGBM classifier for circRNA identification with the optimal feature subset; and (5) Calculating the performance metrics for model evaluation [58]. Niu *et al.* develop a new classifier, CirRNAPL, which extracts the features of nucleic acid composition and structure of the circRNA sequence to optimize the extreme learning machine (ELM), based on the particle swarm optimization algorithm; ELM algorithm randomly assigns input weights and hidden layer thresholds and directly calculates output layer weights by least squares, PSO is used to optimize the input weight and hidden layer deviation of ELM, which can improve the generalizability of the methods [59]. DeepciRGO is constructed using the dependencies between GO classes as background information to predict circRNA functions by integrating multiple interactions and associations; the first step is to extract the topological information of each node in the global network as its feature; and then build a neural network for each GO, consider the functional dependencies between the classes in GO [60].

Due to the relatively low expression of CircRNAs, the specific strategies of CircRNAs detection tools also differ greatly in the results. In addition, the level of circular RNA in all cell lines is very different, which indicates that the expression level of circRNA in different cell lines has different regulatory modes. However, the machine learning methods based on sequence feature only focus on modeling the expression status of circular RNA, and do not consider the expression level of circular RNA. Therefore, the machine learning methods still have more room for development in the study of circular RNA. For example, chen *et al.* model circRNA expression patterns using RNA-seq data as well as integrated sequence and epigenetic features to demonstrate the potential involvement of H3K79me2 in circRNA expression [61]. Integrating sequence features with RNA sequencing data may achieve a more comprehensive understanding of circular RNA in subsequent research.

4. INFLUENCE OF circRNAs ON DISEASES

The human genome can be widely transcribed to generate a large number of non-coding RNAs, such as miRNA, lncRNA, piRNA, circRNA, and others. Many of these are

closely related to the occurrence and development of diseases [62]. Recent studies have shown that circRNAs play an important role in diseases, such as cardiovascular diseases, diabetes mellitus, cancer, and plant-related diseases such as potato bacterial soft rot [27].

4.1. Association Between CircRNAs and Diseases

Based on studies of the structure and function of circRNAs, researchers found that they play a very important role in atherosclerosis, nervous system disorders, diabetes mellitus, and cancer [2]. The antisense noncoding RNA in the ring INK4 locus is the antisense transcript of cyclin-dependent kinase 4 inhibitor protein and its gene variable reading frame fINK4A/ARF1 [63]. SNP can affect cANIL production, thus affecting cANR IL splicing. In turn, cANR IL can affect the inhibition of the INK4a/ARF gene that is mediated by the multi comb family of proteins, therefore promoting the risk of atherosclerosis [64]. Expression of the CircRNA CDR1as was first identified in brain tissue 20 years ago, and the CDR1 gene was once considered as one of the two important genes involved in autoimmune nervous system disorder [65]. CDR1as ability to recruit and interact with miR-7 is the most important basis for CDR1as involvement in the pathogenesis of diseases. It can indirectly regulate the expression of miR-7 targets, thus affecting the occurrence and development of diseases. In an Alzheimer's disease study, it was found that there was maladjustment of the miRNA ring RNA system in the hippocampal CA1 region. The loss of CiRS-7 decreased the activity of "miRNA sponge" and increased the miR-7 level around brain cells, which is one of the mechanisms associated with AD pathogenesis [66]. MiR-7 can directly regulate the α -synuclein expression and play a role in the pathogenesis of Parkinson's disease [67]. In pancreatic β cells, miR-7 inhibitors can induce the subsequent stimulation effect of the mammalian rapamycin target protein signaling pathway in cell proliferation, indicating that miR-7 may be a therapeutic target for diabetes mellitus [68]. Other studies have shown that miR-7 acts as a tumor suppressor, and that its dysfunction and low expression participate in the occurrence and development of different malignant tumors, including breast, liver, and cervical cancers [69]. In addition, it can inhibit the growth of the A549 human non-small cell lung cancer cells by regulating the expression of the apoptotic-related gene β -lymphoma 2 [70]. To reveal the relationship between circRNA and aging, we studied the expression of the ring RNA in different tissues and organs of mice at different growth stages. In the mammary gland, only circUSP3 expression negatively correlated with aging, while in the intestine, the three types of growth related circRNAs, positively correlated. This suggests that circRNAs may be involved in different senescence pathways and play different biological roles [71].

CircRNAs affect plant growth. In kiwifruit, it was found that circRNAs have a specific response to pathogen invasion [72]. In potatoes, differentially expressed circRNAs responded to bacterial infections producing Botox [73]. Several specific circRNAs were found in tomato leaves infected with yellow leaf curl virus disease [74]. There was also a differen-

tial expression of circRNAs in tomato fruits during their development, which indicated that circRNAs could be involved in fruit ripening and coloration [75].

At present, we still do not have a complete understanding of circRNA's functions in health and diseases. Their biological functions and clinical values need to be further explored, and it is undeniable that the research on circRNAs has a bright future for the prevention and treatment of diseases.

4.2. Prediction Method of Diseases Associated circRNA

Benefiting from the prediction of association between diseases and circRNAs, many diseases have now been confirmed to be related to circRNAs. Because circRNAs can adsorb miRNAs in cells and act as sponges for miRNAs, it is possible to construct a heterogeneous information network between circRNAs, miRNAs, and diseases to explore their relationships. For heterogeneous network construction, a circRNA co-expression network is first constructed by using the scale-free characteristics of the ENCODE data set and biological network. In addition, miRNA functional similarity network was constructed by using the data from miRNA target gene database. The miRNA functional similarity network was verified, and the disease-related miRNA and human disease network were obtained after cleaning and analysis of the data, using existing databases and the literature. In the second step, using the Meta path-based approach to get the topological features from the circRNA to the disease and using path count and random walk to measure the correlation degree of different nodes, the different network topology characteristics from circRNA to disease were obtained. Finally, using the machine learning methods (Katz method, single class support vector machine method, category weighting strategy, and bagging ensemble learning strategy), the association between circRNA and disease was verified [76].

Wang *et al.* used a deep convolutional neural network method based on multi-source information to predict the relationship between circRNA and disease. In this method, it was first assumed that similar circRNAs are usually associated with diseases with similar phenotypes. Based on this assumption, a digital descriptor using circRNA similarity network and disease semantic similarity network was constructed. Then, using circRNA and disease's biological information, the Gaussian interaction contour kernel similarity network was introduced into the descriptor. The multi-source information such as disease semantic similarity, disease Gaussian interaction contour kernel similarity, and circular Gaussian interaction contour kernel similarity were then fused. Next, the deep learning convolutional neural network (CNN) algorithm was used to automatically and objectively extract the deep features of the circRNA disease descriptor. Finally, the Extreme Learning Machine classifier was used to quickly and accurately predict the potential circRNA disease association [77].

Another way to determine the correlation between circRNA and disease was based on matrix factorization. First, the

potential circRNA disease association was calculated from the circRNA similarity and disease similarity extracted from the disease semantic information. Known associations of circRNA, genetic disease, and diseases associated with circRNA were also extracted. Then, the circRNA disease interaction spectrum was updated by the adjacent interaction spectrum to correct for false-negative association. Finally, the updated circRNA disease interaction spectrum was decomposed by matrix factor to predict the correlation between circRNA and disease [78].

5. DISCUSSION AND PROSPECT

Based on the in-depth study of next-generation sequencing data and the utilization of machine learning methods, our comprehension of circRNAs has gradually deepened. In this review, we briefly describe the biogenesis and functions of circRNAs. In addition, two important strategies for the identification of circRNAs are introduced: sequence feature and RNA next-generation sequencing data. Next-generation sequencing data technology has prompted a variety of identification software, and DNA sequence features have been rapidly applied to the identification of circRNAs, which have become a new hot spot in bioinformatics. With the development of technologies for the identification of circRNAs, the relationship between circRNAs and diseases gradually became apparent. From the discovery of circRNA in plants' viruses in 1976, to human tumors and cell senescence, and to plant soft rot disease and drought response, increasing knowledge about circRNAs as disease factors has been gathered. In addition, algorithms for the identification of disease-related circRNAs will greatly benefit contemporary medicine as correlations between circRNAs and diseases are gradually discovered.

Although significant progress has been made in the study of circRNA, there are still many problems to be solved. There is no unified biogenic mechanism to explain the biogenesis of all the discovered circRNAs [53, 79]. Algorithms used for the prediction of circRNAs generate results with a large deviation, and there is no unified standard to evaluate the accuracy of these results. Therefore, it is still difficult to detect circRNAs in the genome with sufficient sensitivity and specificity [80]. The mechanism of degradation of circular RNA is even less understood than their biogenesis mechanism. How are circRNAs degraded to maintain a balance in organisms [27]? In recent years, many studies have been conducted on human circRNAs, but much less work has been done on the mechanisms of formation, localization and degradation of plant circRNAs. In addition, specific identification tools for plant circRNAs are lacking.

CONCLUSION

In recent years, studies on plant growth and development, abiotic stress, ncRNAs, and regulatory networks have become more specific and documented. Prediction algorithms based on experimental techniques and experimental ideas will be developed and improved. Due to the stability, tissue specificity, and high expression levels of circRNAs, a

variety of known circRNAs have been associated with diseases, making them more likely to become useful biomarkers and therapeutic targets in human diseases. In addition, circRNAs are modulated by environmental stimuli, causing a variety of diseases. Therefore, circRNAs could become important biomarkers to study the relationship between environmental stimuli and induction of human diseases. Today, the relationships between circRNAs and the pathogenesis of diseases are still unclear, but we believe that in the near future, their functions and relationships with diseases will be more thoroughly understood [81].

CONSENT FOR PUBLICATION

Not applicable.

FUNDING

This study is supported by the National Nature Science Foundation of China (Grant Number: 61862032), the Jiangxi Natural Science Foundation (Grant Number: 20202BABL202034), Jiangxi Provincial Department of Education Science and Technology plan Project (Grant Number: GJJ180197).

CONFLICT OF INTEREST

The authors declare no conflict of interest, financial or otherwise.

ACKNOWLEDGEMENTS

The authors thank the reviewers for their useful feedback and constructive comments.

REFERENCES

- [1] Nigro, J.M.; Cho, K.R.; Fearon, E.R.; Kern, S.E.; Ruppert, J.M.; Oliner, J.D.; Kinzler, K.W.; Vogelstein, B. Scrambled exons. *Cell*, **1991**, *64*(3), 607-613. [http://dx.doi.org/10.1016/0092-8674\(91\)90244-S](http://dx.doi.org/10.1016/0092-8674(91)90244-S) PMID: 1991322
- [2] 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*(11), 675-691. <http://dx.doi.org/10.1038/s41576-019-0158-7> PMID: 31395983
- [3] Diener, T.O. Potato spindle tuber "virus". IV. A replicating, low molecular weight RNA. *Virology*, **1971**, *45*(2), 411-428. [http://dx.doi.org/10.1016/0042-6822\(71\)90342-4](http://dx.doi.org/10.1016/0042-6822(71)90342-4) PMID: 5095900
- [4] Sanger, H. L., et al. Viroids are single-stranded covalently closed circular RNA molecules existing as highly base-paired rod-like structures [J]. *Proceedings of the National Academy of Sciences*, **1976**.
- [5] HSU, COCA-PRADOS. Electron microscopic evidence for the circular form of RNA in the cytoplasm of eukaryotic cells. *Nature*, **1979**.
- [6] Cao, Q.P.; Gaudette, M.F.; Robinson, D.H.; Crain, W.R. Expression of the mouse testis-determining gene Sry in male preimplantation embryos. *Mol. Reprod. Dev.*, **1995**, *40*(2), 196-204. <http://dx.doi.org/10.1002/mrd.1080400208> PMID: 7766412
- [7] 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*(2), e30733. <http://dx.doi.org/10.1371/journal.pone.0030733> PMID: 22319583
- [8] 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*, **2013**, *19*(2), 141-157. <http://dx.doi.org/10.1261/rna.035667.112> PMID: 23249747
- [9] Memczak, S.; Jens, M.; Elefsinioti, A.; Torti, F.; Krueger, J.; Rybak, A.; Maier, L.; Mackowiak, S.D.; Gregersen, L.H.; Munschauer, M.; Loewer, A.; Ziebold, U.; Landthaler, M.; Kocks, C.; le Noble, F.; Rajewsky, N. Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature*, **2013**, *495*(7441), 333-338. <http://dx.doi.org/10.1038/nature11928> PMID: 23446348
- [10] Liu, Y.; Luo, G.; He, D. Clinical importance of S100A9 in osteosarcoma development and as a diagnostic marker and therapeutic target. *Bioengineered*, **2019**, *10*(1), 133-141. <http://dx.doi.org/10.1080/21655979.2019.1607709> PMID: 31055998
- [11] 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. *Mol. Cell*, **2013**, *51*(6), 792-806. <http://dx.doi.org/10.1016/j.molcel.2013.08.017> PMID: 24035497
- [12] Zhaoyong, Li, Chuan, et al. Exon-intron circular RNAs regulate transcription in the nucleus; *Nature Structural & Molecular Biology*, **2015**, *J*
- [13] Guarnerio, J.; Bezzi, M.; Jeong, J.C.; Paffenholz, S.V.; Berry, K.; Naldini, M.M.; Lo-Coco, F.; Tay, Y.; Beck, A.H.; Pandolfi, P.P. Oncogenic Role of Fusion-circRNAs Derived from Cancer-Associated Chromosomal Translocations. *Cell*, **2016**, *165*(2), 289-302. <http://dx.doi.org/10.1016/j.cell.2016.03.020> PMID: 27040497
- [14] Tan, S.; Sun, D.; Pu, W.; Gou, Q.; Guo, C.; Gong, Y.; Li, J.; Wei, Y.Q.; Liu, L.; Zhao, Y.; Peng, Y. Circular RNA F-circEA-2a derived from EML4-ALK fusion gene promotes cell migration and invasion in non-small cell lung cancer. *Mol. Cancer*, **2018**, *17*(1), 138. <http://dx.doi.org/10.1186/s12943-018-0887-9> PMID: 30236141
- [15] Peng, L.; Yuan, X.Q.; Li, G.C. The emerging landscape of circular RNA ciRS-7 in cancer (Review). *Oncol. Rep.*, **2015**, *33*(6), 2669-2674. [Review]. [J]. <http://dx.doi.org/10.3892/or.2015.3904> PMID: 25873049
- [16] Chen, L.L.; Yang, L. Regulation of circRNA biogenesis. *RNA Biol.*, **2015**, *12*(4), 381-388. <http://dx.doi.org/10.1080/15476286.2015.1020271> PMID: 25746834
- [17] Ebbesen, K.K.; Kjems, J.; Hansen, T.B. Circular RNAs: Identification, biogenesis and function. *Biochim. Biophys. Acta*, **2016**, *1859*(1), 163-168. <http://dx.doi.org/10.1016/j.bbagr.2015.07.007> PMID: 26171810
- [18] Liang, D.; Wilusz, J.E. Short intronic repeat sequences facilitate circular RNA production. *Genes Dev.*, **2014**, *28*(20), 2233-2247. <http://dx.doi.org/10.1101/gad.251926.114> PMID: 25281217
- [19] 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*(1), 55-66. <http://dx.doi.org/10.1016/j.molcel.2014.08.019> PMID: 25242144
- [20] Hafner, M.; Landthaler, M.; Burger, L.; Khorshid, M.; Hausser, J.; Berninger, P.; Rothballer, A.; Ascano, M., Jr; Jungkamp, A.C.; Munschauer, M.; Ulrich, A.; Wardle, G.S.; Dewell, S.; Zavolan, M.; Tuschl, T. Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell*, **2010**, *141*(1), 129-141. <http://dx.doi.org/10.1016/j.cell.2010.03.009> PMID: 20371350
- [21] Ji, P.; Wu, W.; Chen, S.; Zheng, Y.; Zhou, L.; Zhang, J.; Cheng, H.; Yan, J.; Zhang, S.; Yang, P.; Zhao, F. Expanded Expression Landscape and Prioritization of Circular RNAs in Mammals. *Cell Rep.*, **2019**, *26*(12), 3444-3460.e5. <http://dx.doi.org/10.1016/j.celrep.2019.02.078> PMID: 30893614
- [22] Fei, T.; Chen, Y.; Xiao, T. Genome-wide CRISPR screen identifies HNRNPL as a prostate cancer dependency regulating RNA splicing [J]. *Proceedings of the National Academy of Sciences of the United States of America*, **2017**. <http://dx.doi.org/10.1073/pnas.1617467114>
- [23] Zhang, X.; Zhang, Y.; Wang, T.; Li, Z.; Cheng, J.; Ge, H.; Tang, Q.; Chen, K.; Liu, L.; Lu, C.; Guo, J.; Zheng, B.; Zheng, Y. A comprehensive map of intron branchpoints and lariat RNAs in plants. *Plant Cell*, **2019**, *31*(5), 956-973.

- http://dx.doi.org/10.1105/tpc.18.00711 PMID: 30894459
- [24] Ye, C.Y.; Zhang, X.; Chu, Q.; Liu, C.; Yu, Y.; Jiang, W.; Zhu, Q.H.; Fan, L.; Guo, L. Full-length sequence assembly reveals circular RNAs with diverse non-GT/AG splicing signals in rice. *RNA Biol.*, **2017**, *14*(8), 1055-1063. http://dx.doi.org/10.1080/15476286.2016.1245268 PMID: 27739910
- [25] Mathews, D.H.; Sabina, J.; Zuker, M.; Turner, D.H. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J. Mol. Biol.*, **1999**, *288*(5), 911-940. http://dx.doi.org/10.1006/jmbi.1999.2700 PMID: 10329189
- [26] Su, H.; Tao, T.; Yang, Z.; Kang, X.; Zhang, X.; Kang, D.; Wu, S.; Li, C. Circular RNA cTFRC acts as the sponge of MicroRNA-107 to promote bladder carcinoma progression. *Mol. Cancer*, **2019**, *18*(1), 27. http://dx.doi.org/10.1186/s12943-019-0951-0 PMID: 30782157
- [27] Liu, X.Q.; Gao, Y.B.; Zhao, L.Z.; Cai, Y.C.; Wang, H.Y.; Miao, M.; Gu, L.F.; Zhang, H.X. [Biogenesis, research methods, and functions of circular RNAs]. *Yi Chuan*, **2019**, *41*(6), 469-485. PMID: 31257196
- [28] Conn, V.M.; Hugouvieux, V.; Nayak, A.; Conos, S.A.; Capovilla, G.; Cildir, G.; Jourdain, A.; Tergaonkar, V.; Schmid, M.; Zubieta, C.; Conn, S.J. A circRNA from SEPALLATA3 regulates splicing of its cognate mRNA through R-loop formation. *Nat. Plants*, **2017**, *3*(5), 17053. http://dx.doi.org/10.1038/nplants.2017.53 PMID: 28418376
- [29] Wang, Y.; Gao, Y.; Zhang, H.; Wang, H.; Liu, X.; Xu, X.; Zhang, Z.; Kohonen, M.V.; Hu, K.; Wang, H.; Xi, F.; Zhao, L.; Lin, C.; Gu, L. Genome-wide profiling of circular RNAs in the rapidly growing shoots of moso bamboo (*Phyllostachys edulis*). *Plant Cell Physiol.*, **2019**, *60*(6), 1354-1373. http://dx.doi.org/10.1093/pcp/pcz043 PMID: 30835314
- [30] Lu, T.; Cui, L.; Zhou, Y. *Transcriptome-wide investigation of circular RNAs in rice*; Rna-a Publication of the Rna Society, **2015**, pp. 2076-2087. J
- [31] Yang, Y.; Fan, X.; Mao, M.; Song, X.; Wu, P.; Zhang, Y.; Jin, Y.; Yang, Y.; Chen, L.L.; Wang, Y.; Wong, C.C.; Xiao, X.; Wang, Z. Extensive translation of circular RNAs driven by N⁶-methyladenosine. *Cell Res.*, **2017**, *27*(5), 626-641. http://dx.doi.org/10.1038/cr.2017.31 PMID: 28281539
- [32] Pamudurti, N.R.; Bartok, O.; Jens, M. Translation of CircRNAs [J]. *Mol Cell*, **2017**, *66*(1), 9-21 e7.
- [33] Wesselhoeft, R.A.; Kowalski, P.S.; Parker-Hale, F.C.; Huang, Y.; Bisaria, N.; Anderson, D.G. RNA Circularization Diminishes Immunogenicity and Can Extend Translation Duration *In Vivo*. *Mol. Cell*, **2019**, *74*(3), 508-520.e4. http://dx.doi.org/10.1016/j.molcel.2019.02.015 PMID: 30902547
- [34] Dong, R.; Zhang, X.O.; Zhang, Y.; Ma, X.K.; Chen, L.L.; Yang, L. CircRNA-derived pseudogenes. *Cell Res.*, **2016**, *26*(6), 747-750. http://dx.doi.org/10.1038/cr.2016.42 PMID: 27021280
- [35] Liang, Y.; Qiu, K.; Liao, B.; Zhu, W.; Huang, X.; Li, L.; Chen, X.; Li, K. Seeksv: an accurate tool for somatic structural variation and virus integration detection. *Bioinformatics*, **2017**, *33*(2), 184-191. http://dx.doi.org/10.1093/bioinformatics/btw591 PMID: 27634948
- [36] Li, N.; Yang, J.; Zhu, W.; Liang, Y. MVSC: A multi-variation simulator of cancer genome. *Comb. Chem. High Throughput Screen.*, **2020**, *23*(4), 326-333. http://dx.doi.org/10.2174/1386207323666200317121136 PMID: 32183666
- [37] Gao, Y.; Wang, J.; Zhao, F. CIRI: an efficient and unbiased algorithm for de novo circular RNA identification. *Genome Biol.*, **2015**, *16*, 4. http://dx.doi.org/10.1186/s13059-014-0571-3 PMID: 25583365
- [38] Jun, C.; Franziska, M.; Christoph, D. Specific identification and quantification of circular RNAs from sequencing data [J]. *Bioinformatics*, **2016**, *7*, 1094.
- [39] 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), 126. http://dx.doi.org/10.1186/s13059-015-0690-5 PMID: 26076956
- [40] Zhang, X.O.; Wang, H.B.; Zhang, Y.; Lu, X.; Chen, L.L.; Yang, L. Complementary sequence-mediated exon circularization. *Cell*, **2014**, *159*(1), 134-147. http://dx.doi.org/10.1016/j.cell.2014.09.001 PMID: 25242744
- [41] 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*(9), e87. http://dx.doi.org/10.1093/nar/gkw075 PMID: 26873924
- [42] Kai, W.; Darshan, S.; Zheng, Z. MapSplice: Accurate mapping of RNA-seq reads for splice junction discovery [J]. *Nucleic Acids Research*, **2010**, *18*(18), e178.
- [43] Hoffmann, S.; Otto, C.; Doose, G.; Tanzer, A.; Langenberger, D.; Christ, S.; Kunz, M.; Holdt, L.M.; Teupser, D.; Hackermüller, J.; Stadler, P.F. A multi-split mapping algorithm for circular RNA, splicing, trans-splicing and fusion detection. *Genome Biol.*, **2014**, *15*(2), R34. http://dx.doi.org/10.1186/gb-2014-15-2-r34 PMID: 24512684
- [44] Trees-Juen, C.; Chan-Shuo, W.; Chia-Ying, C. NCLscan: accurate identification of non-co-linear transcripts (fusion, trans-splicing and circular RNA) with a good balance between sensitivity and precision [J]. *Nucleic Acids Research*, **2016**, *3*(3), e29.
- [45] Izuogu, O.G.; Alhasan, A.A.; Alafghani, H.M. PTESFinder: a computational method to identify post-transcriptional exon shuffling (PTES) events. *BMC Bioinformatics*, **2016**, *17*. http://dx.doi.org/10.1186/s12859-016-0881-4
- [46] Westholm, J.O.; Miura, P.; Olson, S.; Shenker, S.; Joseph, B.; Sanfilippo, P.; Celniker, S.E.; Graveley, B.R.; Lai, E.C. Genome-wide analysis of drosophila circular RNAs reveals their structural and sequence properties and age-dependent neural accumulation. *Cell Rep.*, **2014**, *9*(5), 1966-1980. http://dx.doi.org/10.1016/j.celrep.2014.10.062 PMID: 25544350
- [47] Guo-Yi, J.; Duo-Lin, W.; Meng-Zhu, X. CircRNAFisher: a systematic computational approach for de novo circular RNA identification. *Acta Pharmacol. Sin.*, **2018**, *40*. PMID: 30013032
- [48] Chen, L.; Yu, Y.; Zhang, X.; Liu, C.; Ye, C.; Fan, L. PcircRNA_finder: a software for circRNA prediction in plants. *Bioinformatics*, **2016**, *32*(22), 3528-3529. http://dx.doi.org/10.1093/bioinformatics/btw496 PMID: 27493192
- [49] Zeng, X.; Lin, W.; Guo, M.; Zou, Q. A comprehensive overview and evaluation of circular RNA detection tools. *PLOS Comput. Biol.*, **2017**, *13*(6), e1005420. http://dx.doi.org/10.1371/journal.pcbi.1005420 PMID: 28594838
- [50] PlantcircBase. A Database for Plant Circular RNAs [J]. **2017**, *10*, 1126-8.
- [51] Jing, Z.; Xueying, X.; Wanjuan, G. U Identification of circular RNAs using genomic sequence features [J]. *Chinese Journal of Bioinformatics*, **2018**.
- [52] Xiong, H.Y.; Alipanahi, B.; Lee, L.J.; Bretschneider, H.; Merico, D.; Yuen, R.K.; Hua, Y.; Gueroussov, S.; Najafabadi, H.S.; Hughes, T.R.; Morris, Q.; Barash, Y.; Kainer, A.R.; Jovic, N.; Scherer, S.W.; Blencowe, B.J.; Frey, B.J. RNA splicing. The human splicing code reveals new insights into the genetic determinants of disease. *Science*, **2015**, *347*(6218), 1254806. http://dx.doi.org/10.1126/science.1254806 PMID: 25525159
- [53] Liu, C.; Liu, Y.C.; Huang, H.D.; Wang, W. Biogenesis mechanisms of circular RNA can be categorized through feature extraction of a machine learning model. *Bioinformatics*, **2019**, *35*(23), 4867-4870. http://dx.doi.org/10.1093/bioinformatics/btz705 PMID: 31529043
- [54] Chen, L.; Zhang, Y.H.; Huang, G.; Pan, X.; Wang, S.; Huang, T.; Cai, Y.D. Discriminating circRNAs from other lncRNAs using a hierarchical extreme learning machine (H-ELM) algorithm with feature selection. *Mol. Genet. Genomics*, **2018**, *293*(1), 137-149. http://dx.doi.org/10.1007/s00438-017-1372-7 PMID: 28913654
- [55] 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*(1), 73-80. http://dx.doi.org/10.1093/bioinformatics/btz537 PMID: 31268128
- [56] Pan, X.; Xiong, K. PredcircRNA: computational classification of

- circular RNA from other long non-coding RNA using hybrid features. *Mol. Biosyst.*, **2015**, *11*(8), 2219-2226.
<http://dx.doi.org/10.1039/C5MB00214A> PMID: 26028480
- [57] Pan, X.; Xiong, K.; Anthon, C.; Hyttel, P.; Freude, K.K.; Jensen, L.J.; Gorodkin, J. WebCircRNA: Classifying the Circular RNA Potential of Coding and Noncoding RNA. *Genes (Basel)*, **2018**, *9*(11), E536.
<http://dx.doi.org/10.3390/genes9110536> PMID: 30404245
- [58] 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*, 655-31.
<http://dx.doi.org/10.3389/fgene.2020.00655> PMID: 32849764
- [59] 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.
<http://dx.doi.org/10.1016/j.csbj.2020.03.028> PMID: 32308930
- [60] Deng, L.; Lin, W.; Wang, J.; Zhang, J. DeepciRGO: functional prediction of circular RNAs through hierarchical deep neural networks using heterogeneous network features. *BMC Bioinformatics*, **2020**, *21*(1), 519-537.
<http://dx.doi.org/10.1186/s12859-020-03748-3> PMID: 33183227
- [61] Chen, J.B.; Dong, S.S.; Yao, S.; Duan, Y.Y.; Hu, W.X.; Chen, H.; Wang, N.N.; Chen, X.F.; Hao, R.H.; Thynn, H.N.; Guo, M.R.; Zhang, Y.J.; Rong, Y.; Chen, Y.X.; Zhou, F.L.; Guo, Y.; Yang, T.L. Modeling circRNA expression pattern with integrated sequence and epigenetic features demonstrates the potential involvement of H3K79me2 in circRNA expression. *Bioinformatics*, **2020**, *36*(18), 4739-4748.
<http://dx.doi.org/10.1093/bioinformatics/btaa567> PMID: 32539144
- [62] Lukiw, W.J. Circular RNA (circRNA) in Alzheimer's disease (AD). *Front. Genet.*, **2013**, *4*, 307.
<http://dx.doi.org/10.3389/fgene.2013.00307> PMID: 24427167
- [63] Salzman, J.; Chen, R.E.; Olsen, M.N.; Wang, P.L.; Brown, P.O. Cell-type specific features of circular RNA expression. *PLoS Genet.*, **2013**, *9*(9), e1003777.
<http://dx.doi.org/10.1371/journal.pgen.1003777> PMID: 24039610
- [64] Burd, C.E.; Jeck, W.R.; Liu, Y.; Sanoff, H.K.; Wang, Z.; Sharpless, N.E. Expression of linear and novel circular forms of an INK4/ARF-associated non-coding RNA correlates with atherosclerosis risk. *PLoS Genet.*, **2010**, *6*(12), e1001233.
<http://dx.doi.org/10.1371/journal.pgen.1001233> PMID: 21151960
- [65] Dropcho, E.J.; Chen, Y.T.; Posner, J.B.; Old, L.J. Cloning of a brain protein identified by autoantibodies from a patient with paraneoplastic cerebellar degeneration. *Proc. Natl. Acad. Sci. USA*, **1987**, *84*(13), 4552-4556.
<http://dx.doi.org/10.1073/pnas.84.13.4552> PMID: 3474619
- [66] Lonskaya, I.; Shekoyan, A.R.; Hebron, M.L.; Desforges, N.; Al-garzae, N.K.; Moussa, C.E. Diminished parkin solubility and colocalization with intraneuronal amyloid- β are associated with autophagic defects in Alzheimer's disease. *J. Alzheimers Dis.*, **2013**, *33*(1), 231-247.
<http://dx.doi.org/10.3233/JAD-2012-121141> PMID: 22954671
- [67] EUNSUNG. JUNN, KANG-WOO, *et al.* Repression of α -synuclein expression and toxicity by microRNA-7 [J]. *Proceedings of the National Academy of Sciences of the United States of America*, **2009**.
- [68] Y WANG, J. MicroRNA-7 Regulates the mTOR Pathway and Proliferation in Adult Pancreatic -Cells [J]. *Diabetes.*, **2012**.
- [69] Xu, K.; Chen, Z.; Qin, C. miR-7 inhibits colorectal cancer cell proliferation and induces apoptosis by targeting XRCC2 [J]. *Oncotargets & Therapy*, **2014**, *7*(default), 325-32.
- [70] Xiong, S.; Zheng, Y.; Jiang, P.; Liu, R.; Liu, X.; Chu, Y. MicroRNA-7 inhibits the growth of human non-small cell lung cancer A549 cells through targeting BCL-2. *Int. J. Biol. Sci.*, **2011**, *7*(6), 805-814.
<http://dx.doi.org/10.7150/ijbs.7.805> PMID: 21750649
- [71] Jia-Feng, M.; Xiao-Gen, M. A Designation Circular RNA Expression and the Relationship with Aging. *Xiandai Shengwu Yixue Jinzhan*, **2019**.
- [72] Zupeng, W.; Yifei, L.; Dawei, L. Identification of Circular RNAs in Kiwifruit and Their Species-Specific Response to Bacterial Canker Pathogen Invasion [J]. *Frontiers in Plant ence*, **2017**, *8*
- [73] Zhou, R.; Zhu, Y.; Zhao, J.; Fang, Z.; Wang, S.; Yin, J.; Chu, Z.; Ma, D. Transcriptome-Wide Identification and Characterization of Potato Circular RNAs in Response to *Pectobacterium carotovorum* Subspecies *brasiliense* Infection. *Int. J. Mol. Sci.*, **2017**, *19*(1), E71.
<http://dx.doi.org/10.3390/ijms19010071> PMID: 29280973
- [74] Wang, J.; Yang, Y.; Jin, L.; Ling, X.; Liu, T.; Chen, T.; Ji, Y.; Yu, W.; Zhang, B. Re-analysis of long non-coding RNAs and prediction of circRNAs reveal their novel roles in susceptible tomato following TYLCV infection. *BMC Plant Biol.*, **2018**, *18*(1), 104.
<http://dx.doi.org/10.1186/s12870-018-1332-3> PMID: 29866032
- [75] Tan, J.; Zhou, Z.; Niu, Y. Identification and Functional Characterization of Tomato CircRNAs Derived from Genes Involved in Fruit Pigment Accumulation [J]. *entific Reports*, **2017**, *7*(1), 8594.
- [76] Yizhou, Sun; Jia, Wei. Meta-Path-Based Search and Mining in Heterogeneous Information Networks [J]. **2013**, *(4)*, 329-38.
- [77] Wang, L.; You, Z.H.; Huang, Y.A.; Huang, D.S.; Chan, K.C.C. An efficient approach based on multi-sources information to predict circRNA-disease associations using deep convolutional neural network. *Bioinformatics*, **2020**, *36*(13), 4038-4046.
<http://dx.doi.org/10.1093/bioinformatics/btz825> PMID: 31793982
- [78] Hang, W.; Bin, L. iCircDA-MF: identification of circRNA-disease associations based on matrix factorization. *Brief. Bioinform.*, **2019**.
- [79] Chengyu, L.; Yu-Chen, L.; Hsien-Da, H. Biogenesis mechanisms of circular RNA can be categorized through feature extraction of a machine learning model [J]. *Bioinformatics*, **23**(23)
- [80] Szabo, L.; Salzman, J. Detecting circular RNAs: bioinformatic and experimental challenges. *Nat. Rev. Genet.*, **2016**, *17*(11), 679-692.
<http://dx.doi.org/10.1038/nrg.2016.114> PMID: 27739534
- [81] Lai, X.; Bazin, J.; Webb, S.; Crespi, M.; Zubieta, C.; Conn, S.J. CircRNAs in Plants. *Adv. Exp. Med. Biol.*, **2018**, *1087*, 329-343.
http://dx.doi.org/10.1007/978-981-13-1426-1_26 PMID: 30259378