**REPUBLIC OF TURKEY**
**YILDIZ TECHNICAL UNIVERSITY**
**GRADUATE SCHOOL OF SCIENCE AND ENGINEERING**

# SURVEY OF EXPRESSION PROFILES OF CIRCULAR RNAS AND THEIR PARENT GENES IN CONTEXT OF TISSUE SPECIFICITY

**Elif İrem KELEŞ**

MASTER OF SCIENCE THESIS
Department of Bioengineering
Bioengineering Program

Supervisor
Assoc.Prof.Dr. Alper YILMAZ

July, 2022

**REPUBLIC OF TURKEY**
**YILDIZ TECHNICAL UNIVERSITY**
**GRADUATE SCHOOL OF SCIENCE AND ENGINEERING**

# SURVEY OF EXPRESSION PROFILES OF CIRCULAR RNAS AND THEIR PARENT GENES IN CONTEXT OF TISSUE SPECIFICITY

A thesis submitted by Elif İrem KELEŞ in partial fulfillment of the requirements for the degree of **MASTER OF SCIENCE** is approved by the committee on  in Department of Bioengineering, Bioengineering Program.

Assoc.Prof.Dr. Alper YILMAZ
Yildiz Technical University
Supervisor

**Approved By the Examining Committee**

Assoc.Prof.Dr. Alper YILMAZ, Supervisor
Yildiz Technical University _____

Prof.Dr. Musa TÜRKER, Member
Yildiz Technical University _____

Prof.Dr. Kazım Yalçın ARĞA, Member
Marmara University _____

I hereby declare that I have obtained the required legal permissions during data collection and exploitation procedures, that I have made the in-text citations and cited the references properly, that I haven't falsified and/or fabricated research data and results of the study and that I have abided by the principles of the scientific research and ethics during my Thesis Study under the title of Survey of Expression Profiles of Circular RNAs and their Parent Genes in Context of Tissue Specificity supervised by my supervisor, Assoc.Prof.Dr. Alper YILMAZ. In the case of a discovery of false statement, I am to acknowledge any legal consequence.

Elif İrem KELEŞ

Signature

*Dedicated to my lovely family*
*and my fiancée*

# ACKNOWLEDGEMENTS

First of all, I would like to thank my dear advisor, Assoc. Dr. Alper Yılmaz, who has always provided support and assistance during this difficult and long process.He is always with us in all our works with his patience and sacrifice and supports us.

I would also like to thank my dear family who have always stood by me with their prayers and support throughout this process.

I would like to thank my dear friend Büşra Lüleci, who with worthy opinion whenever I got stuck, was the saver of my hard days, throughout my entire undergraduate and graduate life.

Finally, I would like to thank my precious fiancée, Mert TARCAN, who stood by me in my most difficult times, sacrificed and provided all kinds of support to complete my thesis.

Elif İrem KELEŞ

# TABLE OF CONTENTS

**LIST OF ABBREVIATIONS** vii

**LIST OF FIGURES** viii

**LIST OF TABLES** x

**ABSTRACT** xi

**ÖZET** xiii

**1 INTRODUCTION** 1

1.1 Literature Review . . . . . 1

1.2 Objective of the Thesis . . . . . 2

1.3 Hypothesis . . . . . 2

**2 GENERAL INFORMATION** 4

2.1 circRNA . . . . . 4

2.2 Types of Circular RNAs . . . . . 5

2.3 Localization of circRNAs . . . . . 7

2.4 Biological Functions of circRNAs . . . . . 8

2.4.1 miRNA sponge . . . . . 9

2.4.2 Regulation of Transcription and Post-transcription . . . . . 10

2.4.3 Interaction with RNA binding proteins . . . . . 11

2.4.4 Translation of circRNAs . . . . . 12

2.5 Detection and sequencing of circRNAs . . . . . 13

2.5.1 Identifying Backsplicing . . . . . 13

2.5.2 Genomic Methods . . . . . 14

2.5.3 Computional Approaches for circRNA analysis . . . . . 15

2.6 Disease and circRNA Relation . . . . . 20

2.7 Biogenesis of Circular RNAs . . . . . 27

2.7.1 Trans-acting proteins and RNAs . . . . . 32

2.7.2 RNA binding proteins . . . . . 34

2.7.3 Cis-Regulatory Elements . . . . . 38

**3 METHOD** **42**
3.1 Datasets . . . . . 42
3.2 Database Information . . . . . 42
3.3 Analysis of Available Data . . . . . 42
3.4 Filtration of Tissue-specific circRNAs . . . . . 43
3.5 Calculation of Z-score of Datasets . . . . . 43
3.6 Categorization of Tissue-specific CircRNAs . . . . . 43

**4 RESULTS AND DISCUSSION** **46**

**REFERENCES** **61**

**PUBLICATIONS FROM THE THESIS** **70**

# LIST OF ABBREVIATIONS

| | |
|---|---|
| ADAR | Adenosine Deaminase Acting On RNA |
| BSJ | Back-splice junction |
| cDNA | complementary DNA |
| circRNA | circular ribonucleic acid |
| ciRNA | intronic circular ribonucleic acid |
| CRC | Colorectal cancer |
| DLPFC | Dorsolateral prefrontal cortex |
| EIciRNA | Exon-intron circular ribonucleic acid |
| ecircRNA | Exonic circular ribonucleic acid |
| ecLIP | Enhanced crosslinking and immunoprecipitation |
| Fmn | The mouse formin gene |
| IRES | internal ribosome entry site |
| MBL | Mannose-binding Lectin protein |
| miRNA | micro RNA |
| PFC | preforental cortex |
| PCR | Polymerase Chain Reaction |
| pre-mRNA | precursor messenger RNA |
| RACE | Rapid Amplification of cDNA Ends |
| RBNS | RNA Bind-n-Seq |
| RBP | RNA-binding protein |
| RNase | Ribonuclease |
| tricRNA | TRNA intronic circular RNA |

# LIST OF FIGURES

**Figure 2.1** Linear splicing and circularization . . . . . 5
**Figure 2.2** Types of circRNAs likely to be produced by backsplicing . . . . . 6
**Figure 2.3** Various non-coding RNAs and their functions . . . . . 8
**Figure 2.4** Schematic represantation of circular RNA functions . . . . . 9
**Figure 2.5** CDR1as shows sponge effect by binding to miR-7 . . . . . 10
**Figure 2.6** Differences between roles of circRNAs and cognate linear RNAs in cell proliferation . . . . . 11
**Figure 2.7** CircRNAs are capable of encoding proteins . . . . . 12
**Figure 2.8** Several mechanisms for backsplicing . . . . . 13
**Figure 2.9** Detection of circRNAs with northern blot and RNAse protection . . 14
**Figure 2.10** Computional approaches for circRNAs . . . . . 16
**Figure 2.11** Possible false positive examples in circRNA detections . . . . . 17
**Figure 2.12** The interface of circAtlas 2.0 database . . . . . 19
**Figure 2.13** The mechanisms of action of circRNAs in certain diseases . . . . . 21
**Figure 2.14** Some of identified circRNAs in some body systems . . . . . 21
**Figure 2.15** Biomarker and therapeutic agent potential of circRNAs in the treatment and diagnosis of schizophrenia . . . . . 22
**Figure 2.16** miRNA and protein sponge functions of circRNAs in healthy and tumor tissues . . . . . 24
**Figure 2.17** Schematic representation of circMTO1 function as a miRNA sponge 25
**Figure 2.18** Targets of circRNAs in cardiovascular disease . . . . . 26
**Figure 2.19** Canonical splicing, backsplicing and trans-splicing . . . . . 27
**Figure 2.20** Biogenesis and types of circRNAs . . . . . 29
**Figure 2.21** Competitive splicing resulting in linear and circular RNA . . . . . 30
**Figure 2.22** Models for the production of circular RNA . . . . . 30
**Figure 2.23** circRNA biogenesis mechanism . . . . . 31
**Figure 2.24** The summary of the effects of trans-acting factors in biogenesis of circRNA . . . . . 33
**Figure 2.25** RBPs could bind to enhance or reduce backsplicing . . . . . 35
**Figure 2.26** Reverse complementary sequences, ADAR and QKI effect on circRNA biogenesis . . . . . 35

**Figure 2.27** QKI could increase circRNA number from epithelial to mesenchymal 36
**Figure 2.28** RBPs interacting with pre-mRNA can modulate the backsplicing of circRNAs . . . . . 38
**Figure 2.29** Location of intronic complementary sequences can promote or not circularization . . . . . 39
**Figure 4.1** hsa-TPM2 circRNA expression profile . . . . . 47
**Figure 4.2** hsa-TTN expression profile . . . . . 48
**Figure 4.3** hsa-TNNC1 expression profile . . . . . 49
**Figure 4.4** hsa-TPM1 expression profile . . . . . 50
**Figure 4.5** hsa-TNNI2 expression profile . . . . . 51
**Figure 4.6** hsa-CTRC expression profile . . . . . 52
**Figure 4.7** Biological features for each RBP of interest . . . . . 59

# LIST OF TABLES

**Table 2.1** circRNA detection tool lists with three different approaches . . . . . 18
**Table 2.2** Different functions of selected circRNAs in various cancers . . . . . 23
**Table 3.1** unmatched circRNAs and their parent genes . . . . . 44
**Table 4.1** Expression and gene percentage values of circRNAs that are not present in the same tissue with their parent gene . . . . . 53
**Table 4.2** Only 10 samples from the top 5 ranked RBPs . . . . . 55
**Table 4.3** RBPs binding to target genes as a result of RNA bind-N-seq (RBNS) 56
**Table 4.4** RBPS binding to target genes as a result of eCLIP experiment . . . . 58

# ABSTRACT

## Survey of Expression Profiles of Circular RNAs and their Parent Genes in Context of Tissue Specificity

Elif İrem KELEŞ

Department of Bioengineering
Master of Science Thesis

Supervisor: Assoc.Prof.Dr. Alper YILMAZ

Non-coding RNAs are group of RNA that have various roles at transcriptional and post-transcriptional levels and they comprise majority of human human. CircRNAs are non-coding RNA groups which gained attention recently. circular RNAs are single stranded, circular and covalently closed RNA molecules. Because of its circular structure, CircRNAs cannot code proteins. Upon their initial discovery circRNAs were considered as splicing error in viroids. Ever increasing researches showed that CircRNAs interact with other non-coding RNAs (e.g. miRNA) and proteins (e.g. RBPs), and they are associated with epigenetic changes. Although biological functions of circRNAs are brought to light by ongoing research, uncertainties about their biogenesis still exist. CircRNAs are produced from pre mRNAs by mechanism called back splicing, which is considered a distinct splicing mechanism compared to alternative splicing. In order to have circular form, two edges of circRNAs (5' donor and 3' acceptor site) need to be connected. Exons and flanking introns forming the circRNA may affect both structure and type of circRNA. Because formation of circRNA depends on many factors, biogenesis of circRNAs is yet to be fully understood. Since circRNAs are constructed from pre-mRNAs, it is at least expected that a correlation exists between parent gene expression and circRNA level. In this study, gene expression and circRNA expression data for various tissues were retrieved from circAtlas 2.0 database. After filtering genes and circRNAs that are expressed in specific tissues a striking result was observed. For certain circRNAs, parent gene of the circRNA have been expressed in different tissues. In our study we focused on Uterus tissue to investigate the phenomenon. Samples which contain circRNA expression but parent gene expressions have been detected and

RBPs which have potential of binding to region of parent gene transcript that constitute the circular RNA were identified by integrating results of e-CLIP and RNA bind N seq data acquired from ENCODE project. Our results shed light on biogenesis of circRNAs. RBPs have potential to interact with certain parent gene transcripts and direct all of them to back splice into circRNAs. Thus, the potential of circRNA in diagnosis and treatment of diseases in uterus and role of parent gene in circRNA biogenesis will have been much better understood.

**Keywords:** circRNA, expression profile, tissue-specific

# ÖZET

## Dokuya Özgü Halkasal RNA'lar ve Kaynak Genlerinin İfade Profillerinin Araştırılması

Elif İrem KELEŞ

Biyomühendislik Anabilim Dalı
Yüksek Lisans Tezi

Danışman: Doç.Dr. Alper YILMAZ

Kodlamayan RNA'lar özellikle insan genomunun büyük bir parçasını oluşturan transkripsiyonal ve post transkripsiyonal seviyede çeşitli rollere sahip RNA grubudur. CircRNA'lar ise yeni keşfedilmiş, tek dizili ve kovalent olarak kapatılmış halkasal yapıdaki kodlamayan RNA grubudur. Dairesel olması sebebiyle protein kodlayamamaktadır. Viroidlerdeki ilk keşfinde splicing hatası olduğu düşünülürken, daha sonraki araştırmalarla diğer kodlamayan RNA'larla (miRNA gibi) ve proteinlerle (RBP gibi) de etkileşimde olduğu ve böylece de epigenetik değişikliklere yol açtığı görülmüştür. Bu işlevsellikleri sebebiyle circRNA'lar dikkatleri üzerine çekmiş ve birçok çalışmanın kaynağı olmuştur. Biyolojik fonksiyonları belirlenmiş olsa da, biyogeneziyle alakalı belirsizlikler varlığını sürdürmektedir. circRNA'lar pre mRNA'lardan; alternatif splicing den farklı olarak backsplicing mekanizması ile üretilmektedir. Dairesel bir forma sahip olması için iki ucun (5' donor and 3' acceptor site) birleştirilmesi gerekmektedir. Bu birleşme çeşitli termodinamiksel faktörlerden etkilenmektedir. CircRNA yı oluşturan ekzonlar ve flanking intronlar hem dairesel yapı oluşumunu hem de oluşacak circRNA nın türünü etkileyebilmektedir. circRNA oluşumu birçok faktöre bağlı olduğundan biyogenezle alakalı kesin hükümler verilememektedir. Ancak circRNA'lar pre mRNA'lardan oluştuğu için gen ekspresyonu ile arasında korelasyon olması beklenmektedir. Bu çalışmada çeşitli dokuları kapsayan ve bir milyonun üzerinde circRNA ve gen ekspresyonlarını içeren circAtlas 2.0 veritabanı kullanılmıştır. Yalnızca bir dokuda ifade edilen genler seçilmiştir. Bunun sonucunda circRNA'nın ifade edildiği dokuda parent gen ekspresyonunun bulunmadığı genin başka bir dokuda ifade edildiği görülmüştür. Elde edilen verilerden örnek

bir grup olarak bu durumu en fazla içeren rahim dokusu seçilmiştir. Bu dokuda circRNA ekspresyonu olduğu halde parent gen ekspresyonu olmayanlar belirlenmiş ve bu gen bölgesine bağlanabilecek potansiyele sahip RBPler ENCODE'daki eCLIP ve RNA bind-N-seq deneylerinden elde edilerek kullanılmıştır. Bu çalışma sonucunda elde edilecek bilgilerle circRNA'ların rahimdeki ekspresyonları, fonksiyonları ve circRNA biyogeneziyle alakalı yeni bakış açıları geliştirilmesi amaçlanmıştır. Böylece circRNA'lar rahimle ilgili hastalıkların teşhis ve tedavisi için potansiyel olma durumunda olup, circRNA biyogenezinde parent genin rolü de daha iyi anlaşılmış olacaktır.

**Anahtar Kelimeler:** circRNA, ekspresyon profili, dokuya özgü

# 1
# INTRODUCTION

## 1.1 Literature Review

Between 70% and 90% of mammalian genome transcribes to RNA. However, only 1% of this amount is converted to protein and remaining part becomes expressive in cell as non-coding RNA in order to take part in stages of transcriptional and post transcriptional regulation, epigenetic regulation and cell cycle [1].

circRNAs are one of the members of long non coding RNAs and one of the most popular RNAs. Because of unusual circular forms of them, firstly, they have been subjected to definitions such as splicing error or transcriptional noise. However, it was discovered that unlike linear RNA, circRNAs are generated by a different mechanism of splicing, it has been also proven that circRNAs are also a part of cellular processes.

As studies on circRNA increased, it has been proven that there is a special splicing mechanism, called backsplicing, and as a result of this mechanism, different kinds of circRNA are generated and these RNAs have many different biological functions. circRNAs were shown to interact with DNA, other non-coding RNAs and various proteins in cell. Consequently, circRNAs also affect the processes in which these structures (DNA, protein, non-coding RNA) take part [2]. Despite all these crucial functions, uncertainties regarding circRNAs still remain.

Circular form of circRNAs is generated by backsplicing on the ends of 5' and 3' of the pre-mRNA. So far, two basic methods have been identified: lariat driving and direct backsplicing. During this process, different factors that affect the formation of circRNA such as cis acting and trans acting have been determined. RNA binding proteins (RBPs) can stimulate the formation of circular structure by binding to flanking introns of exons that form circular RNAs, or they can inhibit the formation of circular structure by disrupting the binding of necessary factors to pre-mRNA. As a result of these processes, different circRNAs can be obtained from the same pre-mRNA. Most circRNAs only consist of exons but some circRNAs might also have introns. By

joining of the ends, a loop structure is formed, thus, the degradation of the ends by exonucleases becomes unlikely, rendering circRNAs as stable molecules. Since circRNAs are generated from mRNAs, it can be predicted that circRNAs show similar expression patterns with their parent genes.

Widespread expression is a type of expression pattern. On the other hand, cell, tissue or disease specific expression is more noteworthy expression pattern. Also, a significant number of genes expressed in tissues can produce circRNAs and thus, expression profiles of circRNAs are expected to be correlated with expression profiles of their parent genes.

Use of circRNA as biomarker is encouraged because of their tissue, cell and disease specific expressions and also their stable structure, their abundance in blood and saliva [3, 4]. As a result of previous research, circRNAs that are associated with common diseases such as various types of cancer, Alzheimer's, diabetes were determined and difference in their expression related to disease have been determined [5–7].

Therefore, elucidating the uncertainties in circRNA biogenesis, the tissue specificity, analyzing the profiles of expression related to parent gene will bring a new perspective to circRNAs and let new therapeutic approaches to diseases be developed.

## 1.2 Objective of the Thesis

Recently discovered circRNAs are the latest members of the family of non-coding RNA and therefore, there are no certain and generalizable knowledge about their biogenesis. Their existence in fluid of body such as blood, saliva, urine makes it easy to use as biomarker. In addition, their existence in different tissues with different expression patterns shows their tissue specificity. In this study, tissue-specific profiles of circRNAs will be studied in relation to expression profile of parent gene in order to elucidate biogenesis of circRNAs.

## 1.3 Hypothesis

The studies focused on non-coding RNAs are increasing because of importance and potential functions of non-coding RNAs in cellular events.

Discovery of circRNAS have revealed questions about biogenesis. Linear RNAs are constructed from pre-mRNAs by alternative splicing. Forming of circRNAs by a mechanism other that that of linear RNAs, makes circRNAs specific. There are a lot of factors that affect biogenesis, therefore the mechanism of biogenesis cannot be

generalized. However, in depth analysis of tissue specific expression of parent gene and its associated circRNA have potential to shed light on biogenesis of circRNAs.

# 2
# GENERAL INFORMATION

## 2.1 circRNA

Circular RNAs are covalent and closed loop structure RNAs, unlike known linear RNAs [8]. The first discovered RNA in circular structure is a plant viroid formed by host cellular enzymes, contrary to the circular RNA formation mechanism. This type of viroids has been discovered incidentally in the 1990s. Then, having an important area in eukaryotes has been detected with high-throughput RNA sequencing technology and bioinformatic analyses [9]. When eukaryotic circular RNAs were discovered first, they were thought to be a byproduct of abnormal splicing without any function [8, 10, 11]. However, their high expression levels and having tissue and cell specificity suggest that they may have important biological roles [12]. Moreover, the complex biogenesis of circRNAs and their relationship with many elements indicate that circRNA splicing is not accidentally but a feature of gene expression [13].

The resource of most human and mouse circRNAs is protein-coding genes [14]. circRNA forming genes have been observed that having significant differences from other genes. These genes contain longer exons than non producing circRNA genes. Single exon circRNA exons ve flanking introns are unusually longer than multi-exon circRNAs [15].

circRNAs in various organisms have been revealed to be produced by back-splicing mechanisms and can be arranged from all regions of the genome, obtaining mostly from exons but, less commonly, from intragenic, antisense, intronic, intergenic regions [16]. Circular RNAs are identified according to the originated gene regions and they are accordingly located in the nucleus or cytoplasm in the cell.

As a member of the noncoding RNA family, circRNAs are emerging factor that are present in different species from Archaea to mammals [17].They are composed of the back-splicing mechanism that a forming of phosphodiester bonds of 5' upstream and 3' downstream sites and resulting in the formation of a circular loop structure. Thus,

that back-splicing mechanism is an out of the ordinary type of alternative splicing [18]. Non-coding RNA types formed by that splice mechanism have no free ends, unlike linear RNAs (Figure 2.1). However, since the 5' and 3' splice regions have been used for linear RNA formation, it has been observed a decrease in the proportion of circular RNA formation [18].

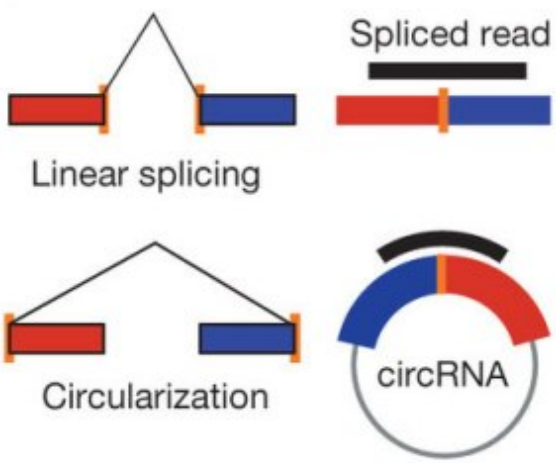


**Figure 2.1** Linear splicing and circularization

Lacking accessible ends, circRNAs are resistant to exonuclease RNase R activity, and so they are are more stable than linear isoforms. Due to its circular form and lack of free ends, it could not be detected with the classical RNA dataset, and so analyzed with RNase R treatment in recently analysis [18]. In addition, they can be used as a biomarker due to this feature and they can play different roles in the psychological and pathological processes of various diseases [19, 20].

circRNAs are found in body fluids such as blood and saliva. Also, they are tissue-specific and differentially expressed in various tissues.They are highly expressed in the nervous system and due to the slow division of neurons, some circular RNAs are discovered more abundant here[18].

In addition to all these features, circRNAs are tissue-, species- and developmental stage-specific, indicating that they have crucial functions in numerous physiological and pathophysiological processes [21–23].

## 2.2 Types of Circular RNAs

Recent studies have revealed that circRNAs are a type of long non-coding RNA ranging from 100 nt to 4 kb in length [13, 24, 25]. circRNAs can be expressed

from all protein-coding genes but the amount of circular RNA produced from some genes is more than the linear RNA, suggesting that the function of these genes is to produce circular RNA, not protein-coding mRNA [26]. Also, protein-coding genes are evolutionarily conserved and production of circular RNA from a certain exon is species-specific [26]. They can be generated from exons, introns, or both regions of genes and classified accordingly. Circular RNAs are divided into 4 subgroups; ecircRNAs, ciRNAs, EIciRNAs, tricRNAs (Figure 2.2). It is known that most circRNAs are derived from exons although the types of circRNAs differentiate according to tissues and species [27].

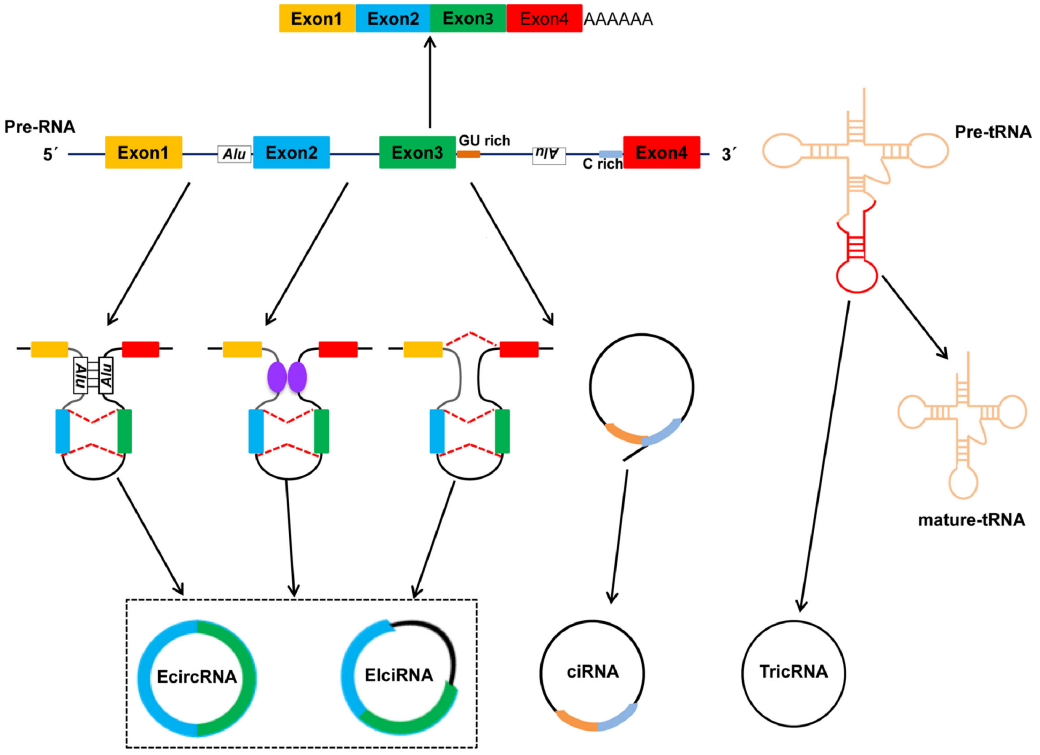


**Figure 2.2** Types of circRNAs likely to be produced by backsplicing. CircRNAs are usually classified into four types: EcircRNA, EIciRNA, ciRNA tricRNA [28]

The vast majority of circRNAs formed through different mechanisms are exonic circRNA (ecircRNAs). They can be formed from single or several exons [29]. While the half-life of mRNAs is approximately 10 hours, the half-life of exonic circRNAs is more than 48 hours, and therefore they have a very stable structure [14, 30]. Also,they are resistant to debranching enzymes as they do not contain 2'-5' carbon linkages [19].

Although exonic circRNA is normally localized in the cytoplasm, exporting of these from the nucleus can be during mitosis [31–33]. They are involved in post-transcriptional gene regulation while they are in the cytoplasm.

ciRNAs and EIciRNAs are located in the nucleus to regulate their parent gene transcription [24]. As an example of the function of EIcircRNA; circEIF3J and circPAIP2 play a role in regulating the expression of host genes by binding to RNA Pol II [27].

While circular intronic RNAs (ciRNAs) originate from only introns, exonic-intronic circRNAs (EIciRNAs) are composed of both exons and introns [29]. ciRNAs are generated during the canonical splicing and derived from the failing of intronic lariat debranching. According to the bioinformatic analysis and experimental results, ciRNA production is dependent on the RNA motif containing the 7-nt GU rich region near the 5' ss and the 11 nt C rich region at the branch point site [24, 34]. Since ciRNA is involved in gene expression, its absence reduces parent gene expression [35]. For instance, one of the ciRNAs, ci-ankrd52, accumulates at the transcription sites and associates with Pol II, and acts as a positive regulator of Pol II transcription. Except of this function, some ciRNAs interact with RBPs and act as molecular sponges in the nucleus and it can also contains miRNA target sites. In addition, ciRNA expression is positively correlated with their parent mRNA expression and also has tissue and cell specificity [34].

Both intronic and exonic circRNAs impact in regulation of gene expression [34]. Finally, fourth type circRNAs is TricRNA. The intron-containing pre-tRNA can be split into two parts by the tRNA splicing endonuclease complex in the BHB motif, and then the intron ends form mature tRNA and circular intronic tRNA(TricRNA) [1, 29].

## 2.3 Localization of circRNAs

Linear mRNAs are mostly located in the cytoplasm, but circRNAs can be located both in the nucleus and cytoplasm. Although circRNAs often contain more than one exon, EIciRNAs in the nucleus have been found to contain intron. circRNAs in the nucleus can organise transcription and splicing of their parent genes. A ciRNA, ci-ankrd52, exists in the nucleus and together with Pol II positively regulates the translation of the encoding gene.

ElciEIF3J, which is an EIciRNA, is also found in the nucleus and takes part in the transcription of the parental gene by interacting with the U1 snRNP [24].

While circular RNAs (EIciRNA and ciRNA) containing introns are found in the nucleus, most circRNAs are exported from the nucleus to the cytoplasm by different RNA helicases. ATP - dependent RNA helicases as DDX39A can transport relatively smaller circular RNAs, while some of the spliceosome RNA helicases like as DDX39B transport

larger circular RNAs from the nucleus to the cytoplasm.

## 2.4 Biological Functions of circRNAs

CircRNAs have cellular functions such as miRNA sponge, regulation of translation, and protein translation [19]. CircRNAs can also interact with different proteins except for these functions (Figure 2.3).

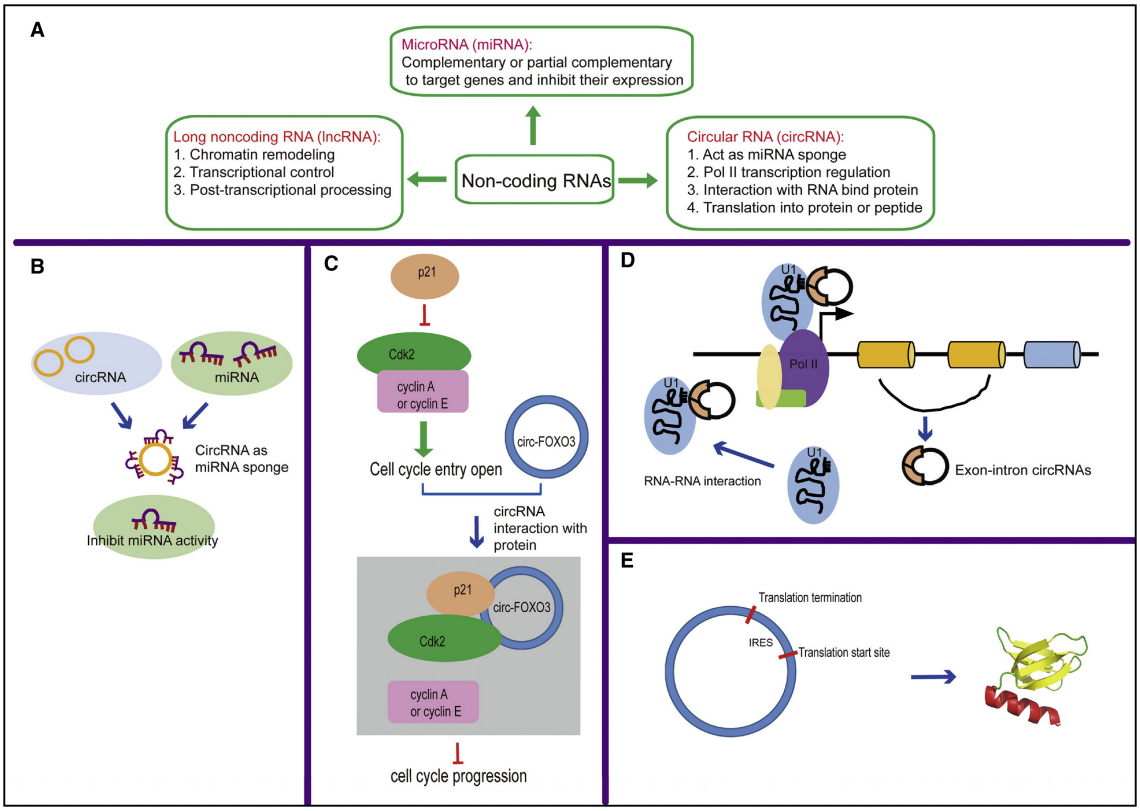


**Figure 2.3** (A) Various non-coding RNAs and their functions (B) miRNA sponge functions of circRNAs (C) repression of cell-cycle progression by interaction with circRNA and RBP (D) circRNAs can bind to RNA Pol II to regulate host gene transcription (E) CircRNAs having IREs could be translated [27]

These circRNA transcripts are regulatory molecules that interact with all RNA types coding or non-coding, DNA, and proteins in the cell [26]. They are one of the significant factors in the regulation of the expression of protein-coding genes or non-coding RNAs [36]. It has been observed that circRNAs are required in the regulation and control during the transcription or translation and post-translational stages [26]. Circular RNAs with different functions can be originated from the same gene (Figure 2.4) [26].

Hence, the formation of a circular structure has a mechanism by which the functions of eukaryotic genes are regulated. It is also involved in regulation of parental gene

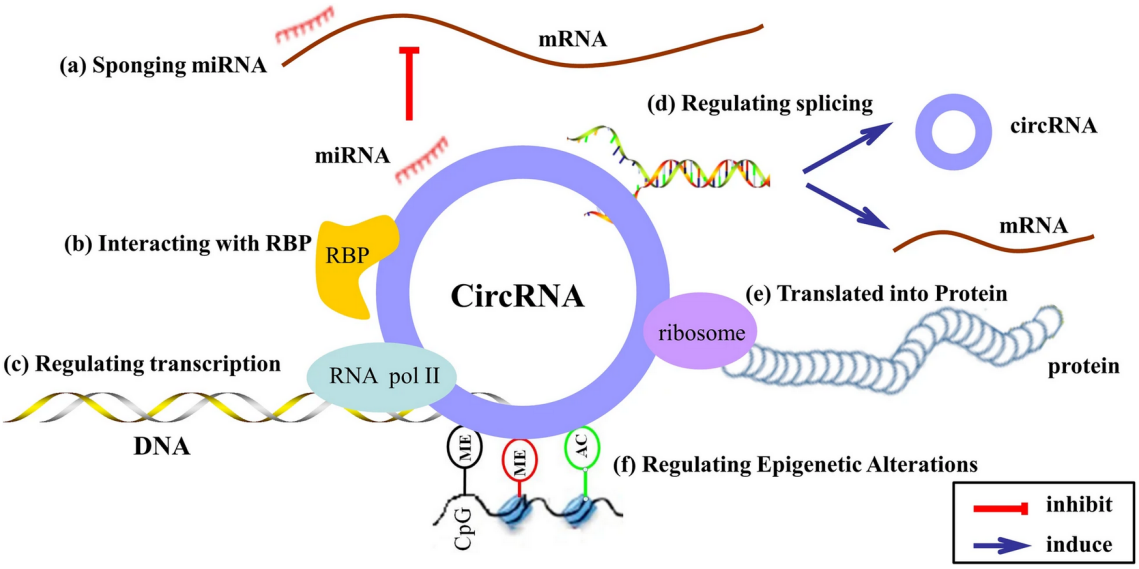


**Figure 2.4** Schematic represantation of circular RNA functions [37]

expression, encoding protein, RNA-binding protein interaction and as miRNA sponge and as competing endogenous RNA [27].

### 2.4.1 miRNA sponge

As a result of many studies, circRNAs have been revealed that regulate mRNA expression by showing the miRNA sponge effect. Also, circular RNAs in the cytoplasm are ceRNAs (competing endogenous RNAs) defined as miRNA sponges by binding miRNAs and preventing them from binding to target mRNAs. circRNAs can affect many different biological functions with the miRNA sponge effect in the cell [2]. The same circRNA can bind to multiple miRNAs [38]. However, each circRNA has properties specific to the miRNA to which it will bind. They must have high stability to be unaffected by circRNA-miRNA binding. Although the stability of circRNAs is high, they have been observed to degrade under some normal or stress conditions in recent studies. On the other hand, miRNA containing target regions for circRNA can cause degradation [2].

Mouse circSry is associated with mir-138, which is effective in testis development and contains 16 target sites.

circZNF1 functions as the miR-23b-3p sponge during human epidermal stem cell differentiation. A number of circular RNAs such as circMAT2B97 and circASAP1 can cause hepatocellular carcinoma to progress by affecting pathways associated with these circular RNAs such as circMAT2B–miR-338-3p–PKM2 or circASAP1–miR-326–miR-532-5p–MAPK1–CSF1. Additionally, CDR1as is one of the most conspicuous examples of circRNAs abundantly expressed in mammalian brains

and contains more than 70 binding sites for miR-7. With the reduced levels of CDR1as expression, a decrease observes in miR-7 and mRNA levels related to it, demonstrating that CDR1as acts as a ceRNA sponge (Figure 2.5) [18].

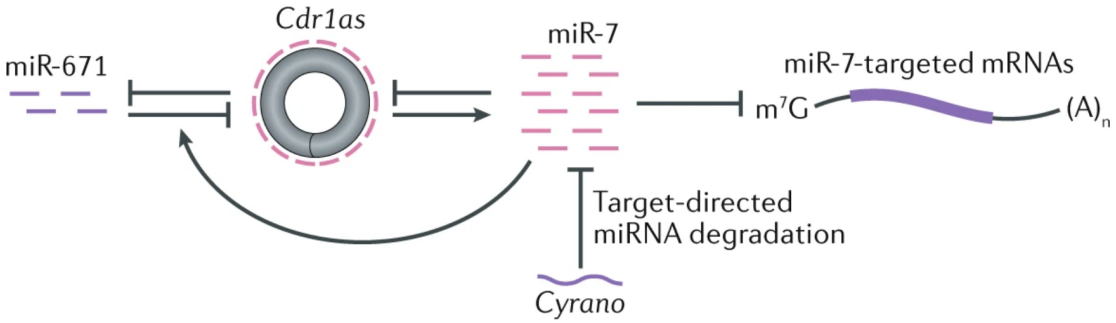


**Figure 2.5** CDR1as shows sponge effect by binding to miR-7 [18]

Circular RNAs have a tissue-specific and parallel expression profile with their parent genes. The results about mouse sry circularization [32] suggests that some circular RNAs that have quite low expression levels were products of splicing errors and therefore do not have biological functions [39].

circRNAs also play a role in gene expression by interacting with RNA binding proteins. circRNA can interact with different proteins. Exonic-intronic circRNAs affect the transcription of parent genes by interacting with Pol II and U1 small nuclear ribonucleoproteins in the promoters of their parent genes [18]. circRNAs act independently of their cognate linear RNAs in the regulation of cell proliferation and so, while deletion of circular RNAs does not change cell proliferation, deletion of linear RNAs can alter cell proliferation (Figure 2.6).

Due to the lack of free 5' end, circRNAs can not be translated into proteins, except in exceptional conditions [40].

### 2.4.2 Regulation of Transcription and Post-transcription

Transcription can be regulated by different mechanisms of circularization [19]. For example, an exonic circRNA containing splice acceptor upstream is generated from the Fmn (the mouse Formin gene) transcript. As a result of this splicing, a non-coding linear RNA transcript can be left and the expression level of the protein is reduced. This function of exonic circRNA has been accepted as "mRNA trap" [12, 19].

EIciRNAs enhance the expression levels of the genes through transcription control, which requires interaction with parent genes in cis form. ci-ankrd52 as exonic intronic circular RNA acts as a positive regulator of the expression of the parent gene ANKRD52

by interacting with Pol II during transcription. This can be only one example of the many transcriptional control functions of circular RNA.

The transcription mechanism interacts with EIciRNAs such as circEIF3J and circPAIP2, stimulating the transcription of its parent genes. The complex consisting of U1snRNA and these EIciRNAs, interacts with U1A, U1C proteins, U2 snRNA, and RNA polymerase II and plays a cis regulator role in the parent gene of these circular RNAs. In addition, circular RNAs interacting with regulatory proteins can affect the translation of targeted mRNAs. As stated, CircPABPN1 downregulating the translation of the parent gene acts as a positive regulator of PABPN1 by inhibiting the huR protein. Backsplicing mechanism supports post-transcriptional regulation as well as transcriptional mechanism. Through the MBL protein binding sites containing in the circMbl introns, this gene can creates negative feedback on its own expression [2].

### 2.4.3 Interaction with RNA binding proteins

Exonic circRNAs can sequester RBPs like linear non-coding RNA transcripts and they can also bind to multiple RBPs as scaffolds through the increased stability of circRNAs.

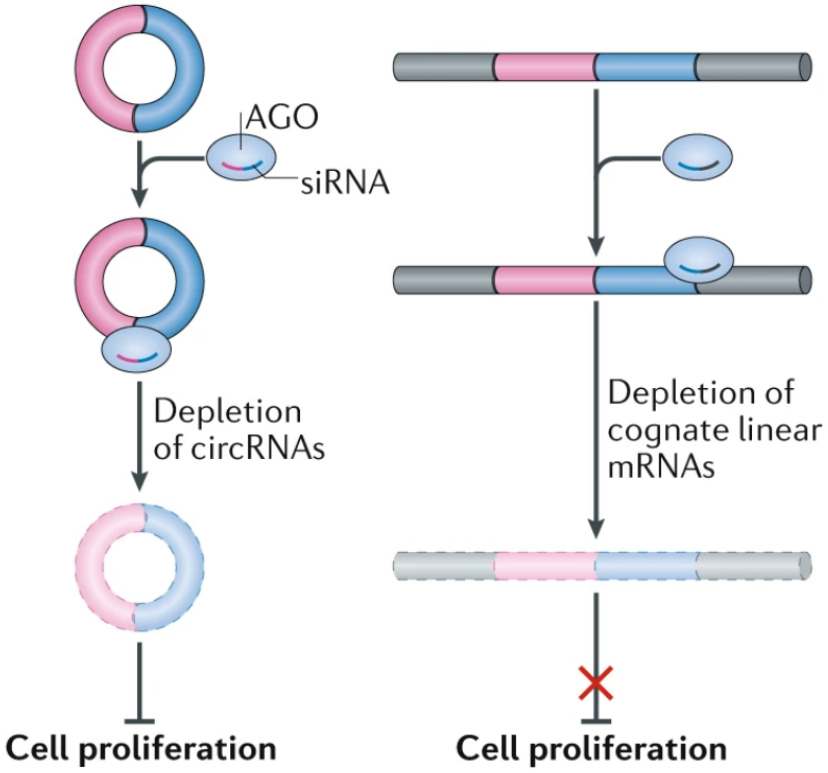


**Figure 2.6** Differences between roles of circRNAs and cognate linear RNAs in cell proliferation. Depletion of circRNAs can not affect cell proliferation but depletion of cognate linear RNAs affect this process [18]

They can bind simultaneously to RBPs and DNA or RNA regions complementing circRNA sequence [19, 33, 41].

### 2.4.4 Translation of circRNAs

The vast majority of circular RNAs are transcribed from the same gene as linear mRNAs. Circular RNAs also have the capability to produce proteins (Figure 2.7) and therefore can synthesize proteins having different functions than their canonical counterparts.These non-coding proteins have different regulatory roles [42].

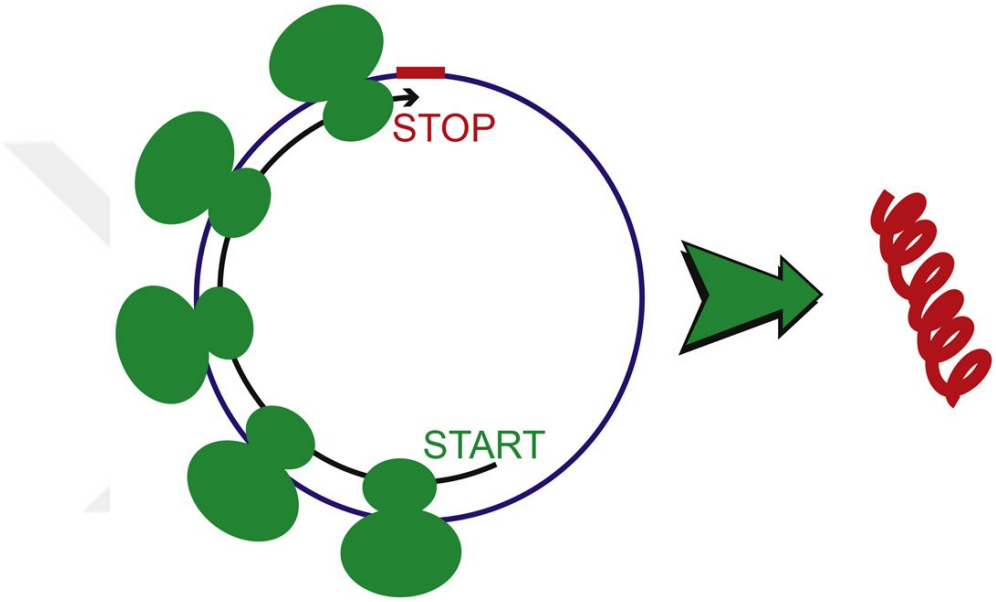


**Figure 2.7** circRNAs are capable of encoding proteins [43]

circRNAs have not start and stop codons and so they are lack of protein coding capacity. Thereby, they can be translated to protein thanks to the IRES region added by various engineering approaches [19, 44].

In circRNAs, a translation could also occur independently of IRES but with the presence of sequences that allow translation to be initiated [43].

In recent studies, circ-ZNF609 has been found that contains the open reading frame necessary for protein synthesis providing an example of the coding capacity of circRNAs. CircRNAs related to translation in drosophila have been defined similarly [43, 45]. To date, any circRNAs produced in human fibroblasts and containing a large amount of ATGs could not determined naturally translated into proteins [46]. However, circRNAs can be made to be capable of protein synthesis with various engineering approaches [46].

## 2.5 Detection and sequencing of circRNAs

Circular RNAs obtained from the cytoplasm of several eukaryotic cells were first observed with electron microscopy, but lariat which are byproducts of splicing and circRNAs could not be easily distinguished with this method [47, 48].

Circular RNAs can not be easily separated based on their size and electrophoretic mobility like other RNAs. Due to the absence of the 3' and 5' ends, amplification and fragmentation techniques that require polyadenylated free ends such as rapid amplification of cDNA ends (RACE) or RNA-Seq cannot be used.Therefore, different approaches have been developed for circRNA detection.

### 2.5.1 Identifying Backsplicing

Backsplicing is the most distinctive hallmark of exonic circRNA formation, but backsplice sequences can also occur as a result of different mechanisms (Figure 2.8) such as reverse transcriptase template switching; tandem duplication and RNA trans-splicing [19].

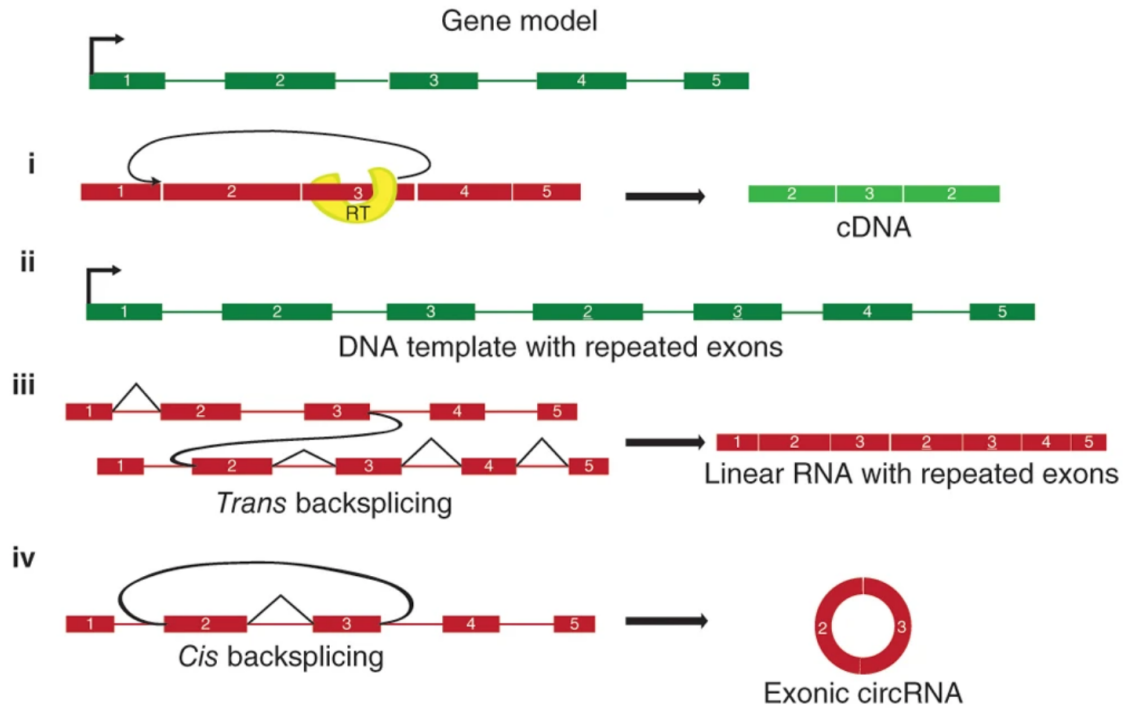


**Figure 2.8** Several mechanisms for backsplicing.i) Reverse transcriptase. (ii) Tandem duplications in the DNA template. (iii) Trans-backsplicing. (iv) Exonic circRNAs can be generated by cis-backsplicing from exons in the same RNA [19]

All types of circular transcripts identified to date have been determined to have a lower expression level than linear forms and require nested PCR experiments for detection

[42]. Detection of circRNA in the RNA pool can be analyzed with a northern blot or RNAse protection (Figure 2.9) [19].

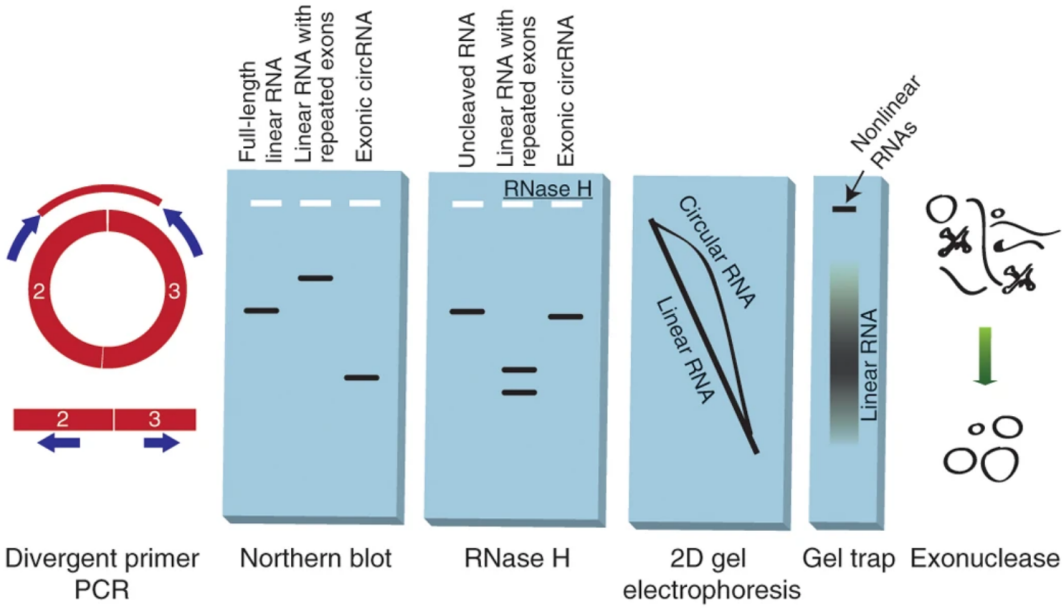


**Figure 2.9** Detection of circRNAs with northern blot and RNAse protection [19]

The migration of exonic circRNAs in the gel crosslinking is more slowly compared to exonic linear RNAs of the same length. However, exonic circRNAs contain fewer nucleotide sequences than other backsplice sequences from the same gene and therefore move faster through the gel. Therefore, the product is linearized by RNase H degradation or hydrolysis to obtain more conclusive information. Evidence of circularity is strengthened by 2D electrophoresis or gel electrophoresis [31, 49, 50]. Different enzymatic methods can also be used in which the linear structure is degraded but the structure of the circRNA is preserved; such as RNAse R exonuclease, and terminator exonuclease treatment. Higher efficiency can be obtained is thought when the above methods are applied in combination. Accurate detection is important because lariate RNA is biochemically different from exonic circRNA. Lariate RNAs contain 2'-5' carbonic linkages contrarily to the 3'-5' linkages of exonic circRNAs [19].

### 2.5.2 Genomic Methods

Two methods can be used as a genomic approach, that one is candidate junction lists generated from transcript models and the other identifying junctions that match the genomic sequence [42, 46]. CircleSeq has been developed as a method that uses both RNase R, which is one of the biochemical approaches, and sequencing. Firstly, RNase R digestion is utilized to detect RNase R resistant species. rRNA depletion can be applied in mammals. In this approach, the readings containing backsplices

are defined by algorithm primarily and the reads of circular types are enriched in the RNase R applied sample. Exonic circRNA detection with this method is greater confidence than other methods [19].

More than 183,000 human, 96,000 macaque, and 82,000 mouse circRNAs have been identified in recent computational analyzes with the detection of back splice connections from RNA-Seq data [18].

RNA-Seq is thought as an imperative tool for differential splicing events of mRNA and differential gene expression analyses at the transcriptome level [1, 51]. It is also an important tool for finding expression, sequence, and structural alterations in transcriptomes [42].

The expression patterns of circular RNAs in mammals varies according to tissues and cell types. High expression level of circRNAs is observed in brain and human platelets despite generally having low expression levels.In addition, circular RNA enrichment is observed during various cell differentiation. Most circRNAs are thought to express less than about 3% of the linear transcript from the same gene [52, 53]. Remarkably, the expression of circular RNA is higher levels than their cognate linear mRNAs, indicating expressions of circRNAs and their linear isoforms are independent [18].

Next-generation sequencing is one of the most common methods used in circRNA sequencing [1]. The low-cost and high capacity of this method makes significant contributions to the discovery of circRNAs and consequently determining functional sequence variants [54].

Nanopore technology is one of the most efficient methods used for detailed analysis of circRNAs. Moreover, this method can also detect low abundance of circRNAs. In this method, full-length complementary (cDNA) circRNA sequences are created using circular reverse transcription to amplify circRNAs. Nanopore technology can be used for sequencing, and appropriate algorithms for expression detection. It has also been shown that the this method is more efficient than the Illumina RNA-Seq method The improvement of circRNA detection in body fluids will make important contributions to research and offer new perspectives on therapeutic approaches [1, 55].

### 2.5.3 Computional Approaches for circRNA analysis

Our knowledge of circRNA expression profile has increased with the application of bioinformatic analyzes and microarrays [56]. circRNA prediction is made after appropriate computational approaches, then miRNA and RBP binding sites can be detected (Figure 2.10)

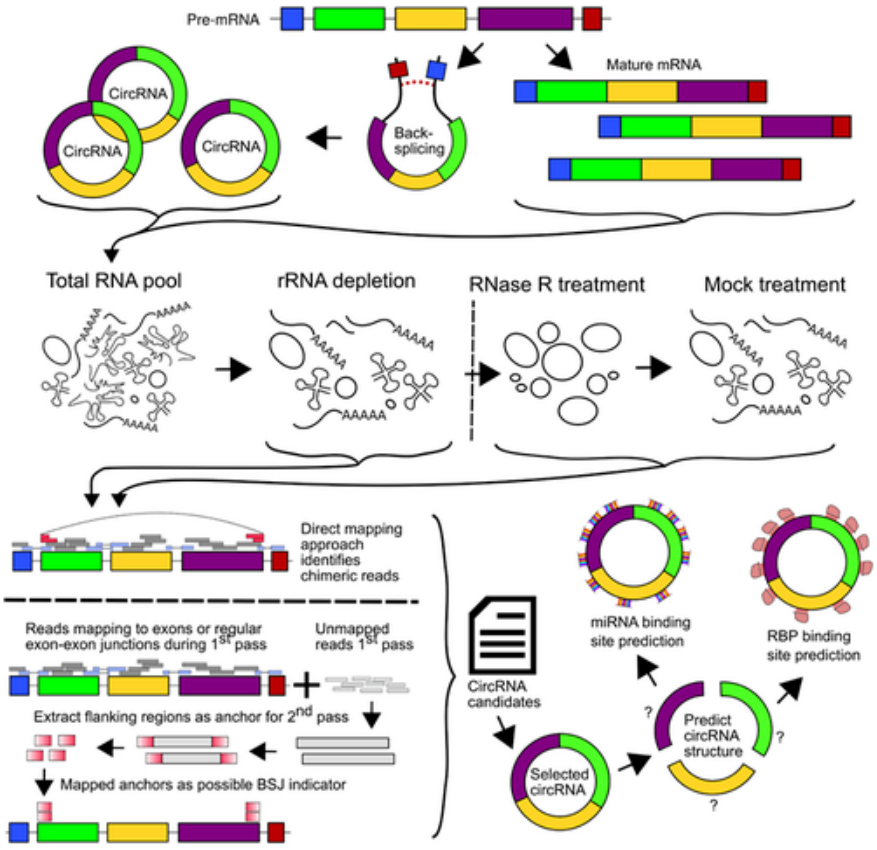


**Figure 2.10** Computional approaches for circRNAs [57]

Although high throughput RNA-Seq is generally used to detect circRNA, microarray-based approaches can also be used. Circleseq is the most recent of these approaches. It can also detect the properties of exons that can form circRNA. Circleseq has found that circular structures can have around 5 exons longer than the average exon length [13].

Sequencing libraries are created from rRNA and treated with RNAse R. In RNAse R-treated sequencing libraries consuming linear RNAs facilitates circRNA detection [57–59]. In contrast to conventional RNA-Seq reads, circRNA BSJs (back splice junction) are targeted specifically in this process. Sequencing libraries selected with polyadenylation cannot be used but a new approach that is a combination of RNAse R, polyadenylation, and poly-A depletion has recently been proposed [57]. Sequencing adapter residues are removed from the acquired RNA-Seq data to improve efficiency and performance of read [60, 61]. But despite this, multiple false-positive results (Figure 2.11) can occur as a result of sequencing reads.

There are various detection tools based on different algorithmic strategies. Table 2.1 describes briefly the tools such as circFinder, find-circ, CIRI, updated version of CIRI,

CIRI2, and CircExplorer.

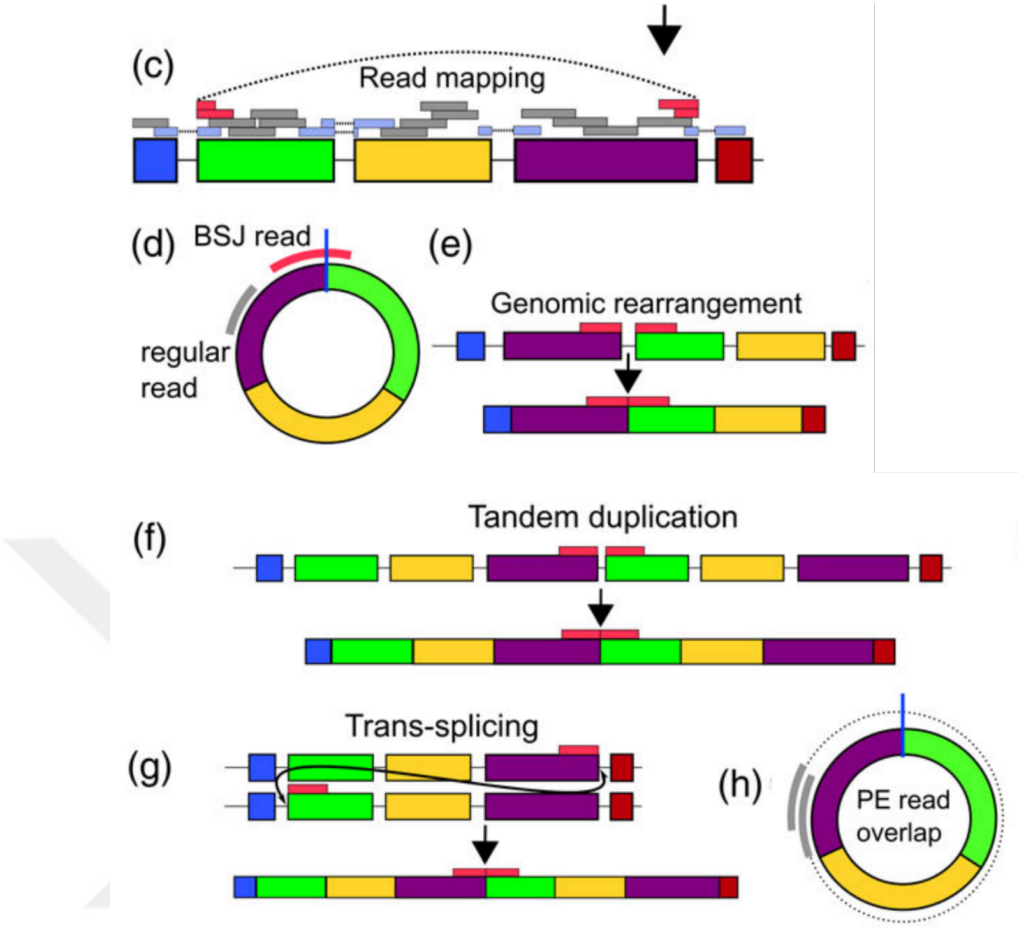


**Figure 2.11** Possible false positive examples in circRNA detections. (d) read mapping: BSJ can be occured from red ends. (e) Scrambled exons can mimic the real BSJ reading. (f) tandem duplication can cause false positives. (g) trans splicing events can also cause false positives to mimic BSJ. (h) Overlapping paired-end may form a circular structure not involving BSJ [57]

**Table 2.1** circRNA detection tool lists with three different approaches [60]

| Tool | Reference | Language | Requirements |
|---|---|---|---|
| Indirect and Multi-Stage Approaches | | | |
| MapSplice | Wang et al. (2010) | Phyton | Unmapped sequencing reads |
| find_circ | Memczak et al.(2013) | Phyton | Expects not mapping reads |
| CIRCfinder | Zhang et al.(2013) | Phyton | Limited to intronic circRNAs |
| CircExplorer | Zhang et al.(2014) | Phyton | Unmapped sequencing reads |
| Approaches Directly Employing Chimeric Reads | | | |
| circRNA finder | | Perl | Unmapped sequencing reads |
| CIRI | Gao et al. (2015) | Perl | Reads pre-mapped with BWA |
| DCC | Cheng, Metge, and Dieterich(2016) | Phyton | Unmapped sequencing reads |
| CircExplorer | Zhang at all (2016) | Phyton | Unmapped sequencing reads |
| Tools Using Statistical Approaches | | | |
| KNIFE | Szabo et al. (2015) | Phyton,Perl,R | Unmapped sequencing reads |

circRNA-finder is one of the early circRNA detection approaches. Some mismatches can occur when sequencing reads are aligned to the reference genome through the all-read length. Intron regions can be left as gaps especially in eukaryotic organisms, as introns are removed during the formation of mature mRNA from pre-mRNA. Various software is required to align such readings correctly. Since both linear and circular RNAs could have the same sequence, the most important step in the computational approach is the correct detection of this structure.

circRNA databases originated with obtained data from different studies. Circbase [62], circRNADb [63], tissue specific circRNA database TSCD [64], circbank[65], circFunbase [66], circAtlas 2.0 [67] are the most used databases. These web-based databases can provide information about miRNA-circRNA, circRNA-disease and RBP-circRNA relationships and more. circNet is database containing only human sourced data. Computational analyses based only on sequencing are limited analyses that are not very reliable for long circRNAs.

Another database circAtlas 2.0 is one of the most comprehensive and updated databases developed so far. Over one million circRNAs were identified across 6 species (human, Macaca, mouse, rat, pig, chicken) from 1070 RNA-Seq samples in 19 different tissues 2.12. That database provides expression values of circRNAs across all tissues. Also provides RBP, miRNA, and parent gene information [67].

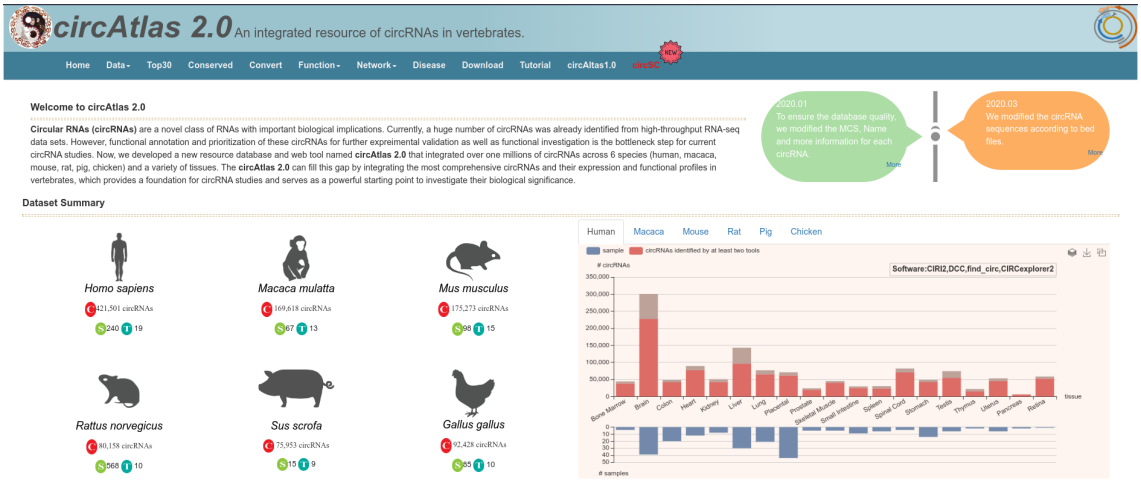


**Figure 2.12** The interface of circAtlas 2.0 database [67]

circbase is a database and website including datas from Homosapiens, Mus musculus, Caenorhabditis elegans, and Latimeria samples [62].

TSCD (Tissue-Specific CircRNA Database) is a database containing 302,853 human and mouse tissue specific circRNAs and associated with them miRNAs and RBPs [64].

circbank is one of the circRNA databases that also includes downloadable features that differ from other databases such as the coding potential and mutations [65].

Computational approaches are more appropriate for circRNAs less than 500 bp and greater than 150 bp in length as they may not include the entire circRNA sequence. When exceeding the 500 bp length limit, algorithms can technically fail. Although algorithmic predictions may only be sufficient for the first step, and then laboratory studies will be required for validation [57]. Although computional analyzes save time and space, circRNA validation with wet lab analysis will be more dependable. The qPCR method can be used to detect the expression variation of selected circRNA.For this, a primer covering the BSJ region of the circRNA should be used. Because BSJ is specific to circRNA, only the primers amplified circular RNA during PCR. Primer design and selection is also an important step for pcr and circtool software includes one of the packages used for this. The primer design becomes easier with the use of this package [57, 68].

As circRNA studies increase, computational approaches increase at the same rate.Despite some deficiencies, the computational approach is the most preferred approach among bioinformatics methods.

## 2.6 Disease and circRNA Relation

circRNAs have been discovered as a therapeutic potential, and a notable biomarker in major human diseases [13, 55, 69]. circRNAs play a role in various diseases by interacting with proteins, RBPs, miRNAs, and indirectly with mRNAs. (Figure 2.13).

Being easily obtained from various body fluids makes them possible to use for the detection and progression evaluation of diseases [70].

Circular RNAs are involved in many different cellular functions in various body systems such as neurologic, cell proliferation and transformation, and immunity in multiple body systems. (Figure 2.14). Accordingly, abnormal regulation of circular RNAs causes various diseases such as atherosclerosis, neurological disorders, and cancer [18, 69, 71].

The development, pathogenesis, and source of many diseases have been elucidated with new generation high throughput DNA and RNA sequencing technologies.

circRNAs are a potential risk factor for certain diseases such as schizophrenia. Schizophrenia is a psychological disorder caused by both environmental and molecular factors that affect the molecular pathways in cells related to brain development [36].

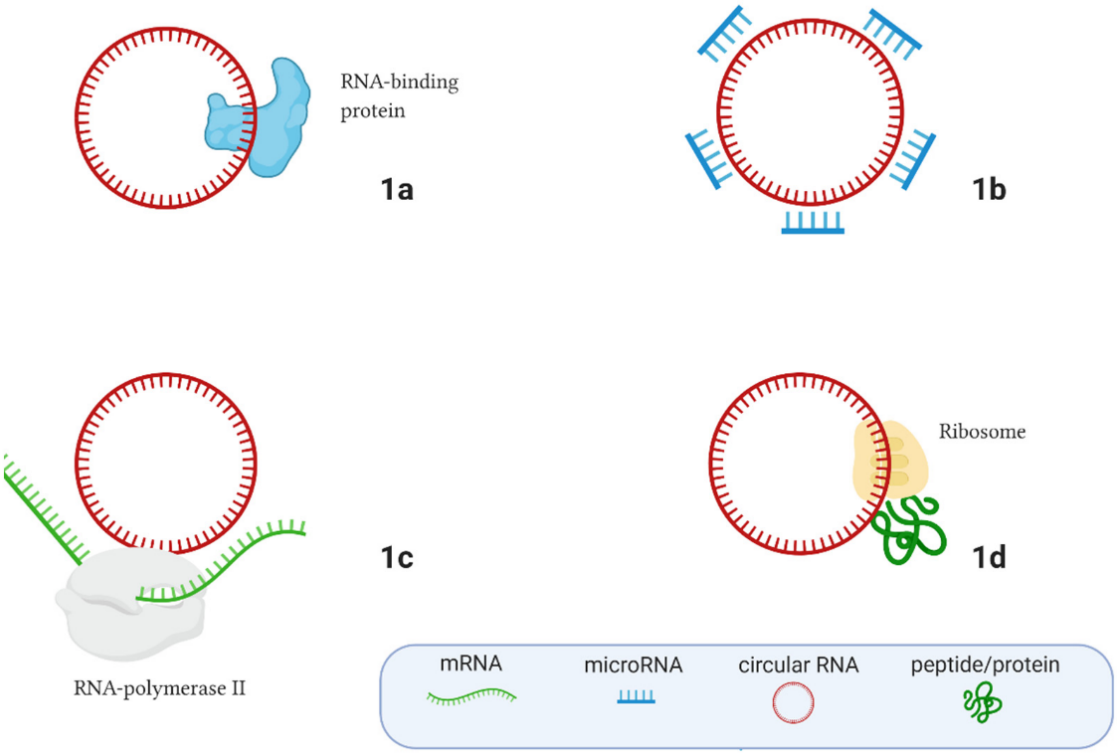


**Figure 2.13** The mechanisms of action of circRNAs in certain diseases [36]

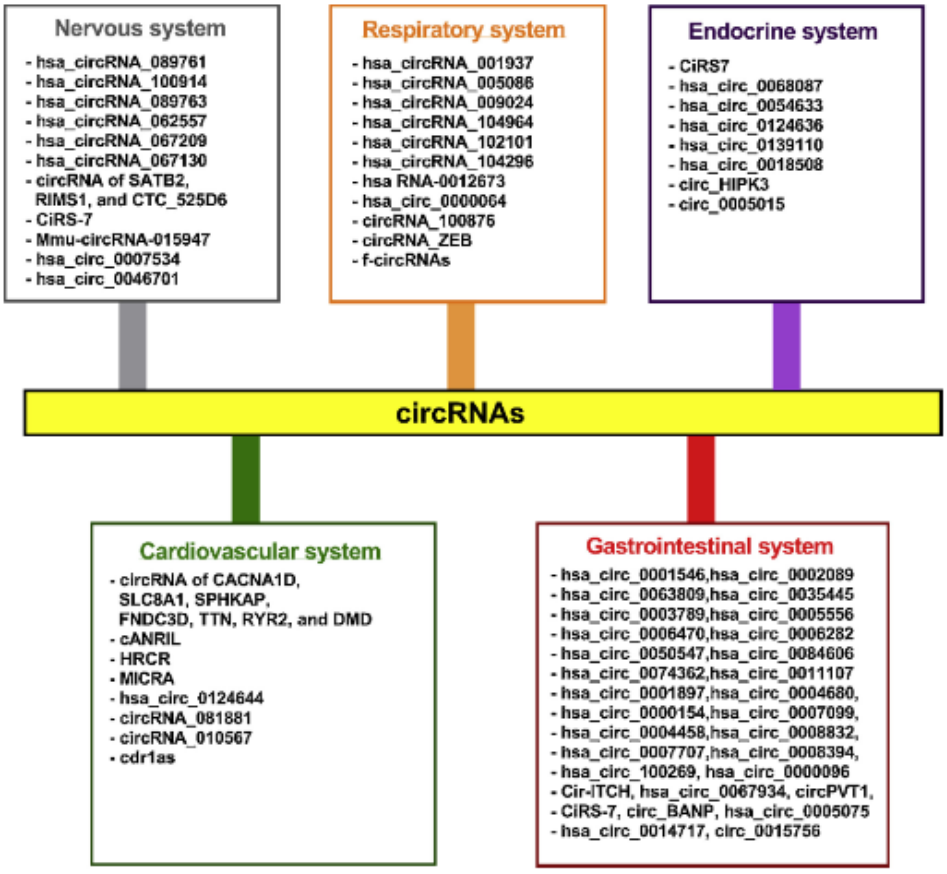


**Figure 2.14** Some of identified circRNAs in some body systems [3]

They can be detected and silenced with short interfering RNAs.It has been shown that circHomer1a which is the host gene of the Homer protein, is downregulated in the schizophrenic postmortem prefrontal cortex (PFC) of the human brain and plays an important role in the cognition process. While RBPs are thought to be a post-transcriptional molecule found in the nucleus and cytoplasm, modern biology has shown that it is effective in the regulation of gene expression in multiple diseases such as cancer, psychiatric and neurodegenerative diseases. circRNAs are a part of molecular processes in diseases interacting with RBPs.Various studies have shown that gene expression changes in different cellular processes interaction RBP or miRNA with circRNA can cause schizophrenia. The circRNAs which have been shown previously to play a role in schizophrenia and include target sequences for miR-34a-5p and miR-449a molecules, have been highly expressed. Thus, it has been accepted that circRNAs could be used as biomarkers for providing information about disease mechanisms (Figure 2.15).

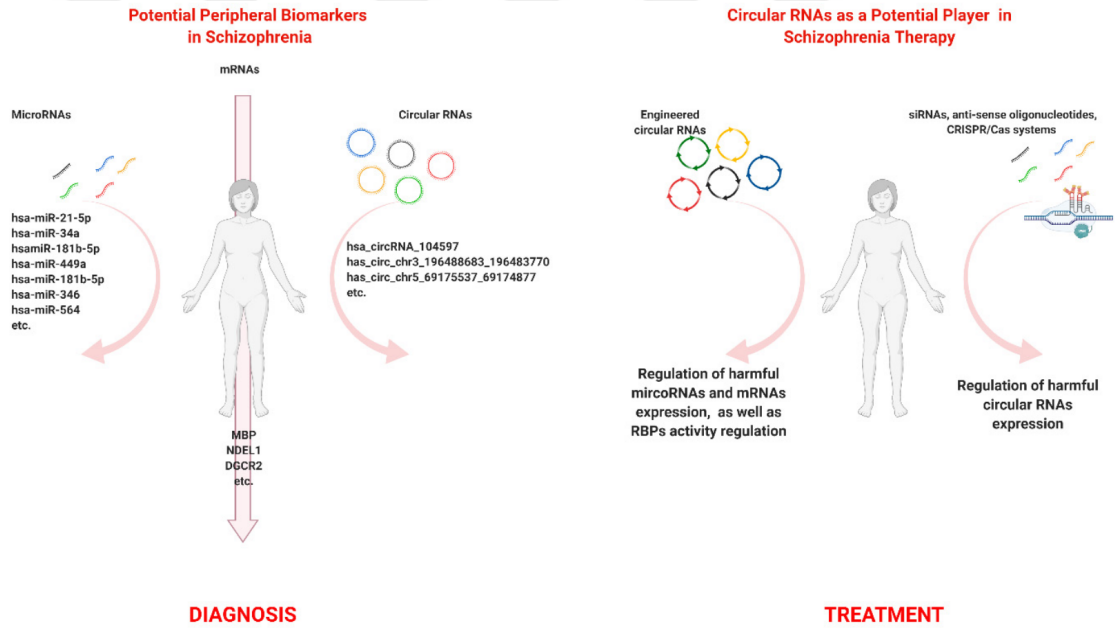


**Figure 2.15** Biomarker and therapeutic agent potential of circRNAs in the treatment and diagnosis of schizophrenia [36]

In addition,circRNAs have extracted from blood plasma exosomes of healthy and schizophrenic patients and they have identified by sequencing. In the analyzes performed on healthy and schizophrenic DLPFC, it was observed that 390 of 574 circRNAs were downregulated and interacted with miRNAs. miRNAs are non-coding elements that are active in the regulation of gene expression in different cell processes. The circRNAs which have been shown previously playing a role in schizophrenia and including target sequences for miR-34a-5p and miR-449a molecules, have highly expressed. Thus, it was predicted that circRNAs could be used as biomarkers by providing information about disease mechanisms.

**Table 2.2** Different functions of selected circRNAs in various cancers [21]

| Cancer | circRNA | Dysregulation | Type of tissue/cell lines | RBPs | Participation |
|---|---|---|---|---|---|
| breast cancer | circFOXO3 | down | BC tissues and cells | MDM2,p53 | apoptosis |
| breast cancer | circMTO1 | down | monastrol-resistant cell lines | TRAF4 | monastrol resistance |
| liver cancer | circZKSCAN1 | down | HCC tissues | FMRP | tumor quiescence |
| Gastric cancer | circAGO2 | up | GC tissues and cells | HuR | growth, invasion, metastasis |
| Glioblastoma | CDR1as | down | GBM tissues | MDM2,p53 | colony formation, proliferation, cell cycle progression, apoptosis |

According to the checks over after treatments, it was observed that hsa_circRNA_104597 which is the host gene PLEKHA2 has upregulated in the peripheral blood cells.Dysregulation of hsa_circRNA_104597 has been associated with schizophrenia and it has potential biomarker for this disease [72].

The circ_0005276 circRNA and FUS protein complex is a significant example that affects prostate cancer development by regulating the transcription of XIAP (X-linked apoptosis protein). In addition to circ_0005276-FUS complex, the circ_foxo3-p21-CDK2 complex have a part in inhibiting cancer progression in mouse. In one of the studies on prostate cancer, the level of circRNA expression was associated with disease progression. At another example of the role of circRNAs in prostate cancer, the level of circRNA expression was associated with disease progression.

It was also observed that approximately 11% of the abundant circRNAs are required for cell proliferation and circCSNK1G3 acts independently of their linear counterparts during that disease progression [73].

In recent research, a variety of circRNAs playing o role in cancer has been identified. Upregulated or downregulated circRNAs can cause cancer by affecting various cellular processes (Table 2.2).

It has been demonstrated that circRNAs have crucial functions as biomarker and therapeutic agents in carcinogenesis, progression and chemotherapy resistance [10, 21, 36]. hsa-circ-0009910 and circ-ZNF609 are circRNA types having a critical role in cancer.

The level of hsa-circ-0002577 observed in endometrium cancer patients was 2.4 higher than in healthy ones.The WDR26 gene is associated with this circRNA and its upregulation stimulates the activation of the PI3K/AKT pathway, causing metastasis. Accordingly, overexpression of hsa-circ-0002577 has been correlated with tumor [10]. The miRNA sponge function of circRNA is also effective in cancer development (Figure 2.16) [21].

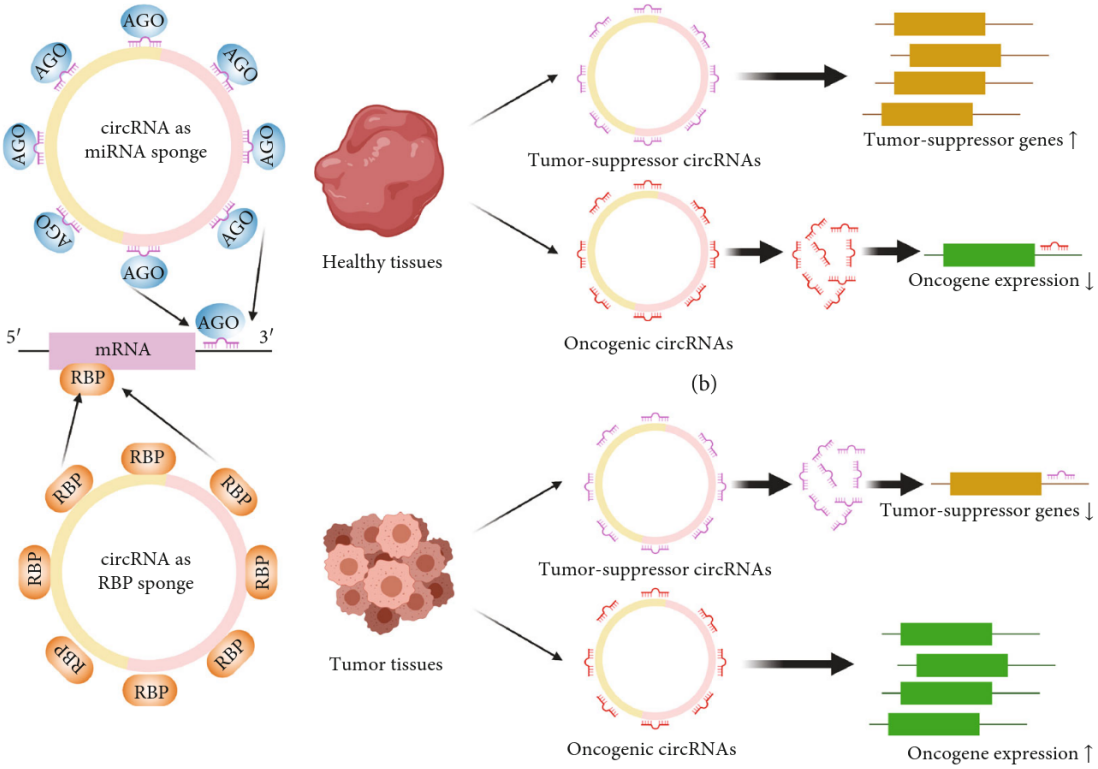


**Figure 2.16** miRNA and protein sponge functions of circRNAs in healthy and tumor tissues. (a) circRNAs affect gene regulation by binding with RBPs and miRNAs (b)Tumor suppressor circRNAs bind to tumor suppressor miRNAs and tumor suppressor gene expression increases in healthy tissue. Oncogene circRNAs in healthy tissue reduce oncogene expression by attaching oncogene miRNAs. (c)In tumor tissues, unlike in healthy tissues, tumor supressor circRNAs reduce the expression of the tumor suppressor gene with the miRNAs to which they bind. Oncogenic circRNAs increase oncogene gene expression by oncogene miRNA binding [13]

ciRS-7 is one of the novel oncogenes and promoting agent in multiple cancer types such as gastric cancer, colorectal cancer, breast cancer, lung cancer.It contains multiple miRNA binding sites and acts as an competiting endogenous RNA (ceRNA) or miRNA sponge in the regulation of transcriptional activity [74]. As an example of this, FOXO3-derived circRNAs are miRNA sponges containing 8 different binding sites to inhibit the growth of breast cancer. Another example of the tumor suppressor function of circRNAs is circMTO1 which binds miRNA-9, which upregulates the expression of p21 in HCC [75]. Except for this suppressive function, it can either promote or inhibit tumor development by binding to different miRNAs (Figure 2.17) [75].

Contrary to the previously mentioned circRNAs with tumor suppressor function, some circRNAs facilitate tumorigenesis. For example, circRNA named hsa-circ-000984 sponges miR-106b in CRC (colorectal cancer) cells and promotes the cancer metastasis [77].

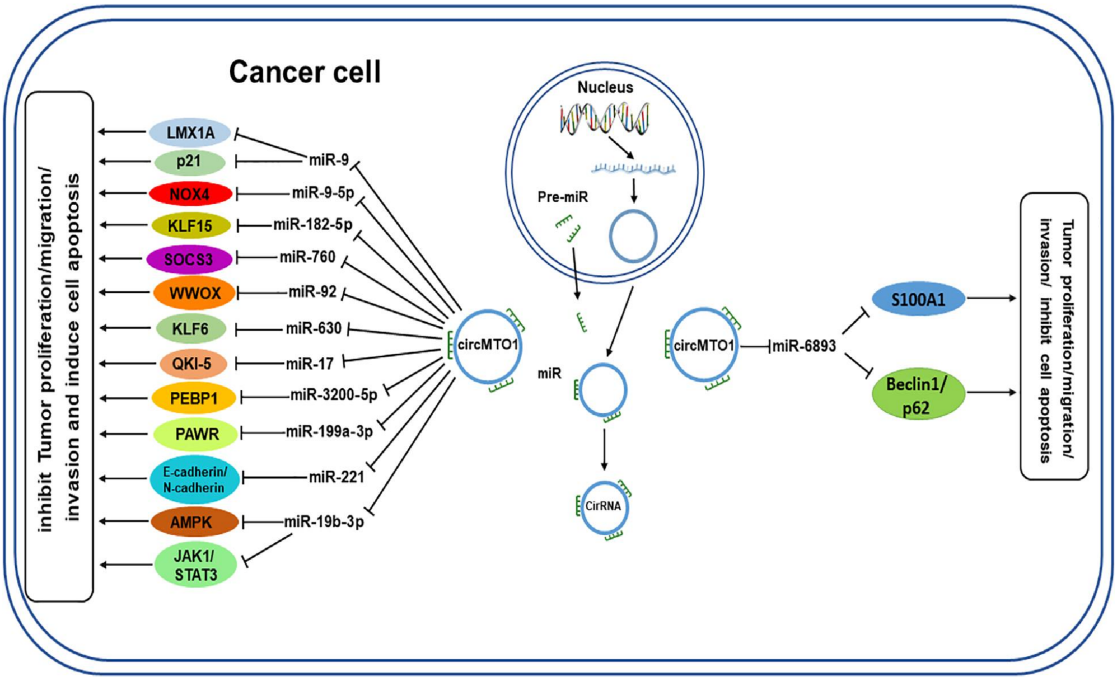


**Figure 2.17** Schematic representation of circMTO1 function as a miRNA sponge. CircMTO1 can act as a miRNA sponge and can promote or inhibit tumor progression [76]

According to the results of the studies, circRNA-RBP interaction has emerged in various stages of cancer such as cell death, metastasis, invasion, and survival [21].

Although the aberrant expression of RBPs is a candidate for biomarker, the detection of circRNAs from body fluids seems more favorable for clinical application. However, more data on circRNAs are still needed [21].

circRNAs are produced abundantly in neuronal tissues, especially in the brain [8]. Recent studies have shown that circRNAs such as Cdr1as and circ-000950 are involved in the pathogenesis of neurodegenerative diseases via sponging mir-103 [78]. Overexpression of circ-0000950 inhibits the expression of anti-inflammatory miRNAs by binding and thus the target genes of these miRNAs have been expressed in a way that enhances neurotoxicity and inflammation [78].

CircRNAs also have pathological and psychological roles in the cardiovascular system (Figure 2.18). Due to the stability of circRNAs in circulation, the observed changes can be considered as biomarkers for CVDs. Another function of circRNAs in the cardiovascular system may be to protect against heart failure by acting as a sponge for miR-223 [27].

Apart from circRNAs produced naturally in the cell, artificial circular RNAs can also be engineered and play a role in various diseases processes.

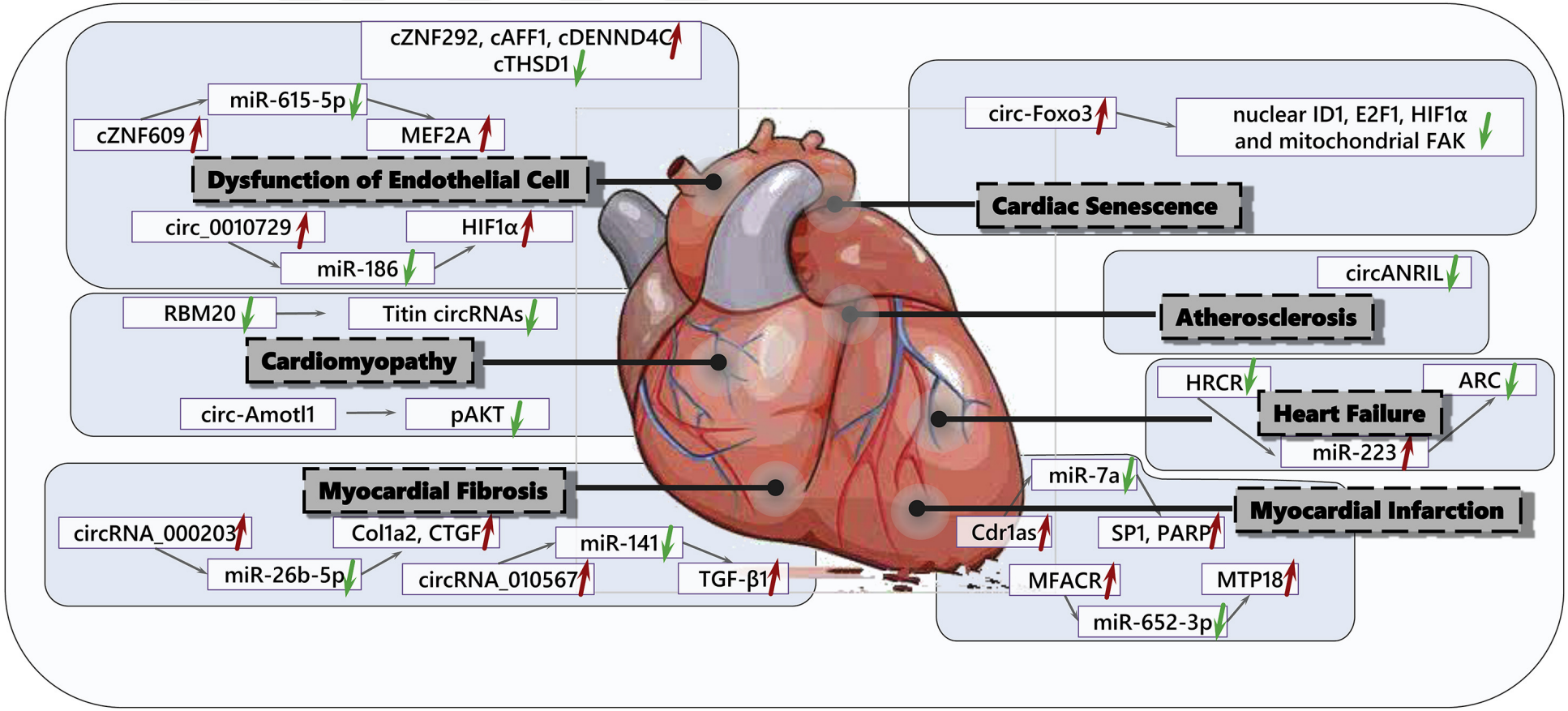


**Figure 2.18** Targets of circRNAs in cardiovascular disease [27]

## 2.7 Biogenesis of Circular RNAs

DNA comprises codes that are required for proteins to perform biological functions in living organisms. These codes are copied into RNA molecules and coded protein regions and noncoding regions are found in the first occurring RNA molecules. Introns, non-coding regions, are removed from this structure known as pre-mRNA and exons are sequentially added to each other. This process is called splicing. There are two different types of splicing: canonical splicing, in which linear RNA is produced, and backsplicing, in which circular RNA is produced. If the pre mRNA splice sites are combined linearly, mature mRNA is composed, but if the upstream region (5') of one exon and the downstream region (3') of the other exon are combined by a phosphodiester bond, circular RNA is formed and this mechanism is called as "back-splicing" [26, 79]. Recent studies have revealed that circRNA biogenesis is different from linear RNA biogenesis (Figure 2.19).

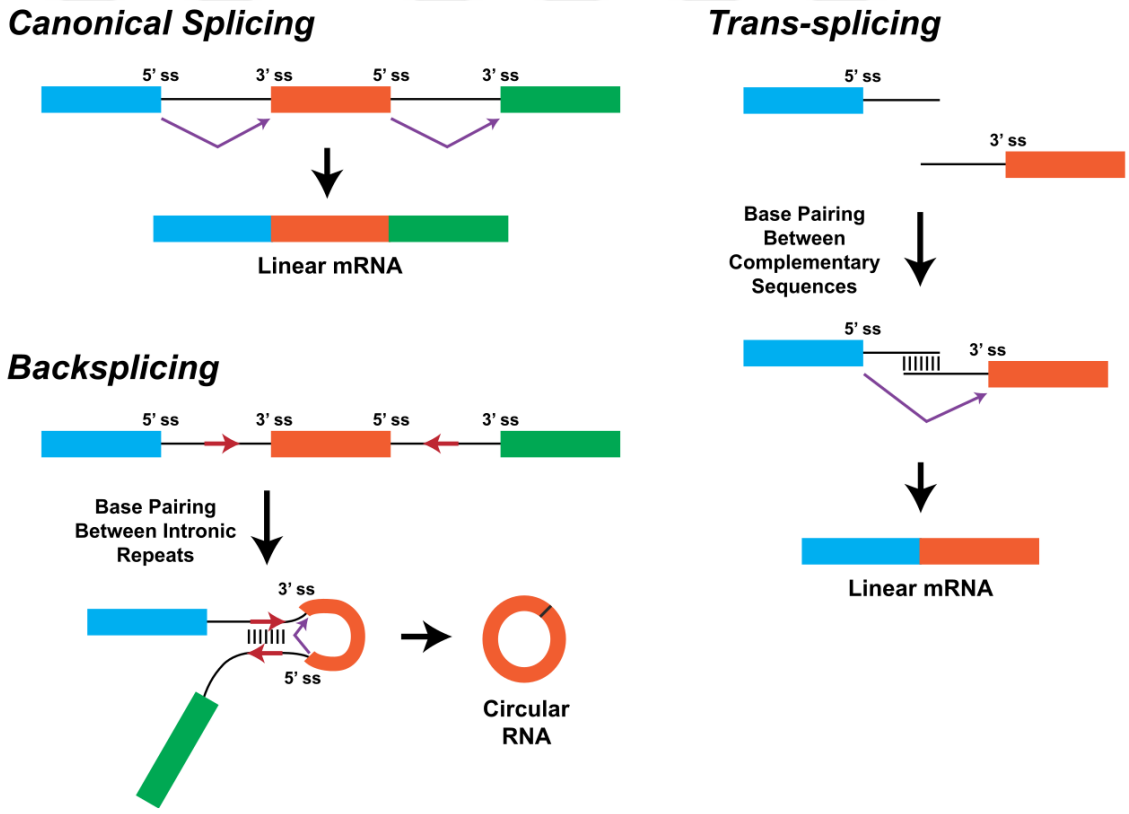


**Figure 2.19** Canonical splicing, backsplicing and trans-splicing[26]

The biogenesis mechanisms of various types of circRNAs are also diversely constituted. Exonic circRNAs are generated by through lariat-driven and back splicing. When the donor and acceptor sites get closer to occuring a lariat containing the exons, the introns are removed, and exons are bound via 5'-3' phosphodiester bond (Figure 2.20A). Another way to generate ecircRNA is back splicing by bridging introns and exons through RBP interactions (Figure 2.20B). ALU sequences pairing is facilitate

connecting of upstream splicing receptor and downstream splicing donor. The introns can be removed during circularization for formation of ecircRNA or EIciRNA (Figure 2.20C). ciRNA biosynthesis is originated a 7-nt GU-rich element and an 11-nt C-rich element in the introns (Figure 2.20D). Whereas ecircRNA are essentially located in the cytoplasm, EIciRNAs and ciRNAs playing a vital role in transcriptional regulation are found in the nucleus. Generation of tricRNA. pre-tRNA are divided into two parts by the tRNA splicing enzymes. 3'–5' phosphodiester bond generate tricRNAs, and the other part generates tRNAs (Figure 2.20E). tRNA intronic circular (tric)RNAs were discovered as non-coding circular RNAs in bacteria and eukaryotic cells. The importance of various trans and cis acting elements for tricRNA biogenesis has been identifed [80]. Except mentioned types of circular RNAs, pre-rRNAs in bacteria are formed in a circular process similar with tricRNA formation.

For instance, contrary to mature linear RNA production, capping is not required, but Pol II transcription is needed for back splicing. Also, different circRNAs from the same sequence can be produced, as in alternative splicing[18].

There are some factors that affect back-splicing as well as factors affecting splicing too. These factors are RBPs, kinetic conditions, exon length, and special motifs. Splice regions are involved in special motifs for circular RNA formation, but no special motifs are required except for these regions. However, single exons need a certain length in the formation of circular RNA[18]. Although back splicing efficiency is lower than canonical splicing, that can be increased under favorable kinetic conditions[18].

Pre-mRNAs can be spliced through two different splicing mechanisms and therefore the same transcript can compete for both canonical and back splicing (Figure 2.21) [57, 82]. Therefore these mechanisms are related to each other for circular RNA formation.

Backsplicing generated from intronic or exonic sequences can be accomplished by a "lariat-driven" or "direct back-splicing mechanism" (Figure 2.22) [2]. Circular RNA biogenesis is divided according to which step comes first. Although both splicing mechanisms can occur, direct backsplicing is thought to occur more frequently [19].

Lariat is a loop structure converting an intron with standard splicing. When skipping happens inter exons during alternative splicing, lariat structure occurs inter skipped exons [79]. All skipped exons can not generate circular RNA, but a decrease in mature linear mRNA formation could observe if backsplicing of exons increase. Although there is a correlation between exon skipping and circRNA formation, the efficiency of this mechanism can also be affected by various regulators. Therefore, there is no exact mechanism for circular RNA biogenesis [18].

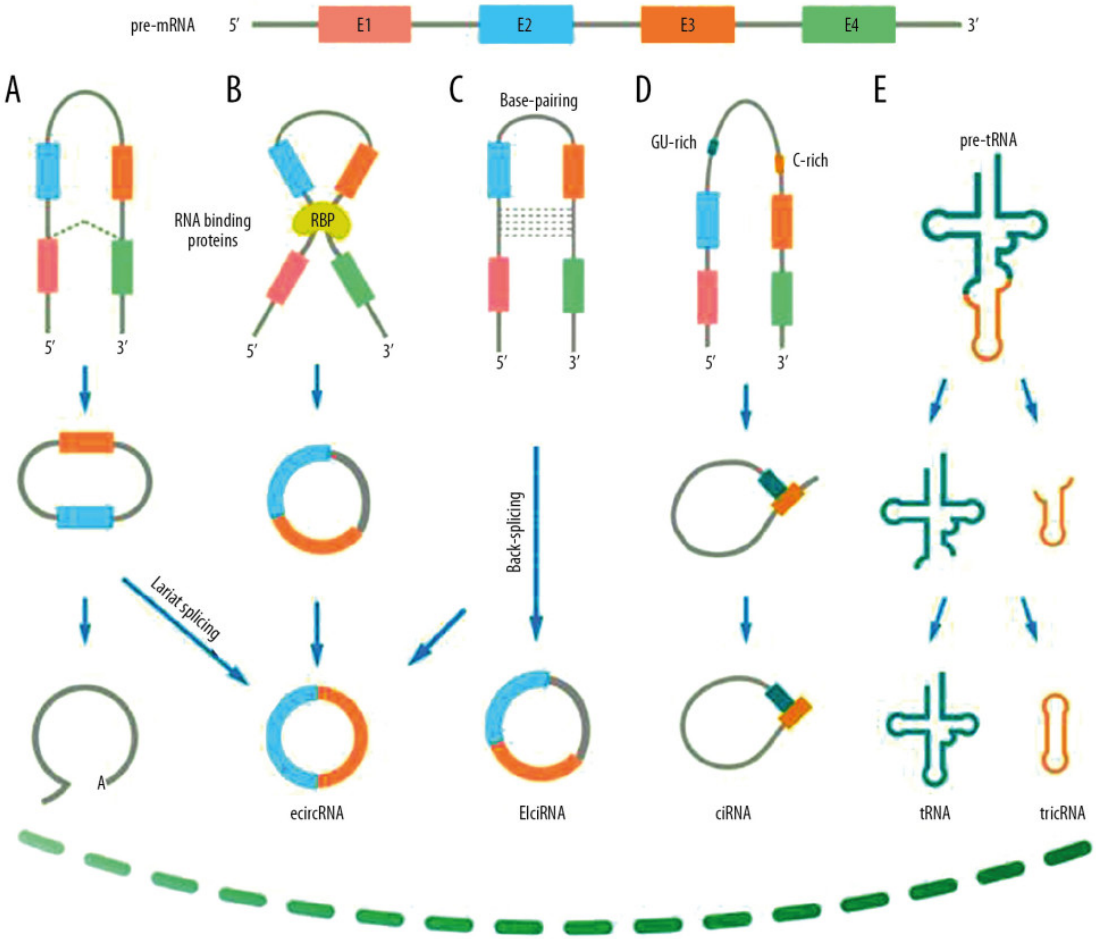


**Figure 2.20** Biogenesis and types of circRNAs
(A)Lariat driven circularization. The lariat consists of exons and introns due to the covalent bond between hydroxyl and phosphate in the 3' upstream and 5' downstream regions of the pre mRNA. EcircRNA is formed through interaction of hydroxyl and phosphate groups inter exons and introns.(B) RNA-binding protein (RBP) dependent circularization.RBPs can promote ecircRNA formation by interacting with flanking introns.(C) Base-pairing-driven circularization. Base pairing of complementary or non-complementary repetitive sequences in flanking introns enables the formation of ecirCRNA or EIciRNA. (D) Biosynthesis of ciRNA. Formation of ciRNAs essentially depends on a 7-nt GU-rich element and an 11-nt C-rich element.(E)Generation of tricRNA. pre-tRNA are divided into two parts by the tRNA splicing enzymes. tricRNAs are originated 3'- 5' phosphodiester bond, and tRNAs are generated from the remaining part [81]

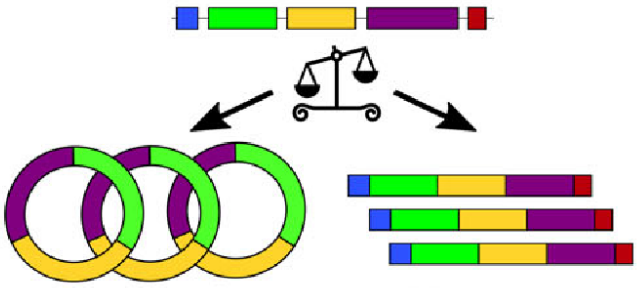

**Figure 2.21** Competitive splicing resulting in linear and circular RNA [57]

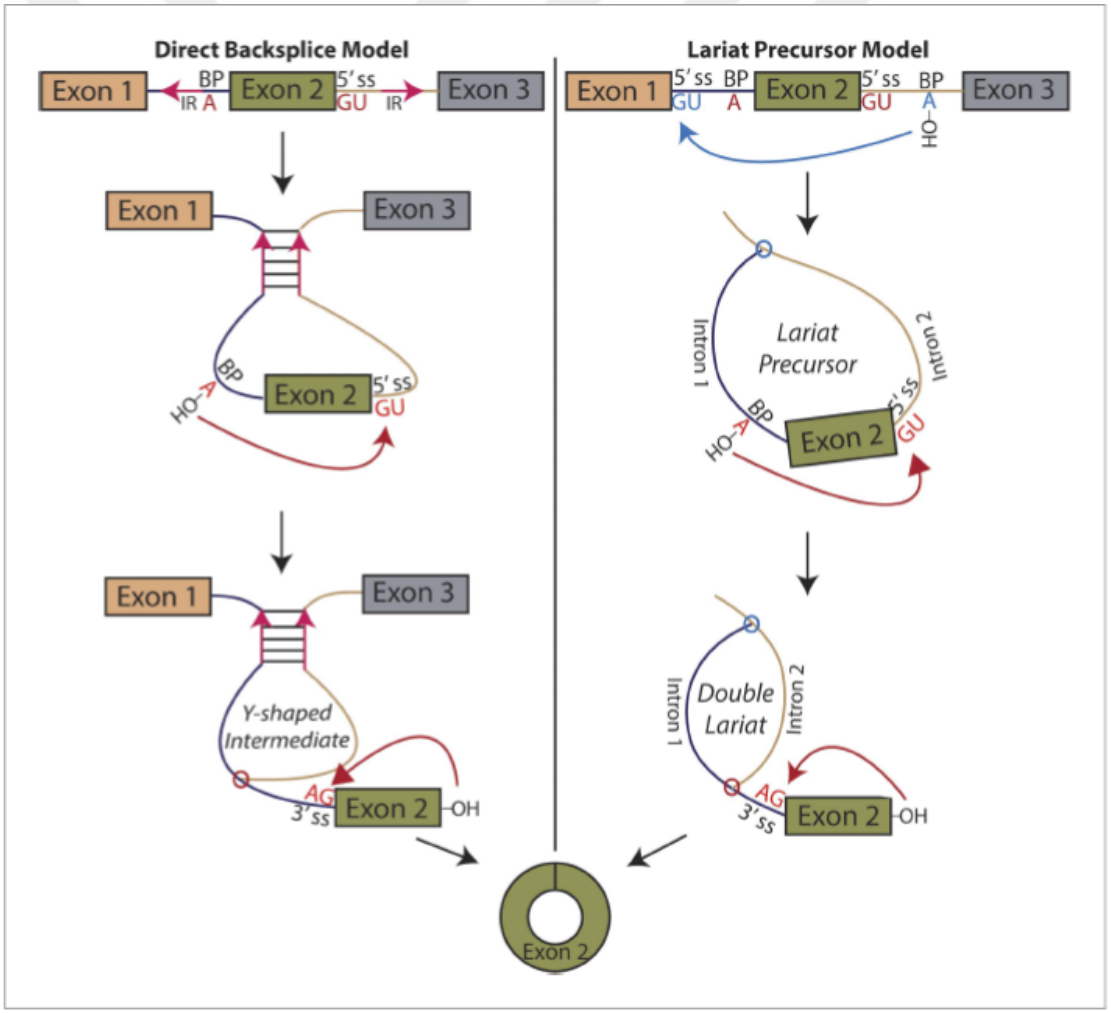


**Figure 2.22** The biogenesis of ecircRNAs is enabled through either complementary base pairing (left) or the formation of a lariat precursor (right). [79]

Initially, if canonical splicing occurs, the pre-mRNA produces a linear RNA inter non-consecutive exons and a long lariat consisting of exons and introns, that is afterwards back-spliced to produce a circRNA (Figure 2.23A). This process is called as "lariat-driven circularization" or "exon skipping" model. Conversely, if the back splicing occurs first, circular RNA formation will occur with an intermediate that needs further processing containing introns and exons to facilitate linear RNA formation from the pre-mRNA (Figure 2.23B). This process is known as "direct back-splicing".

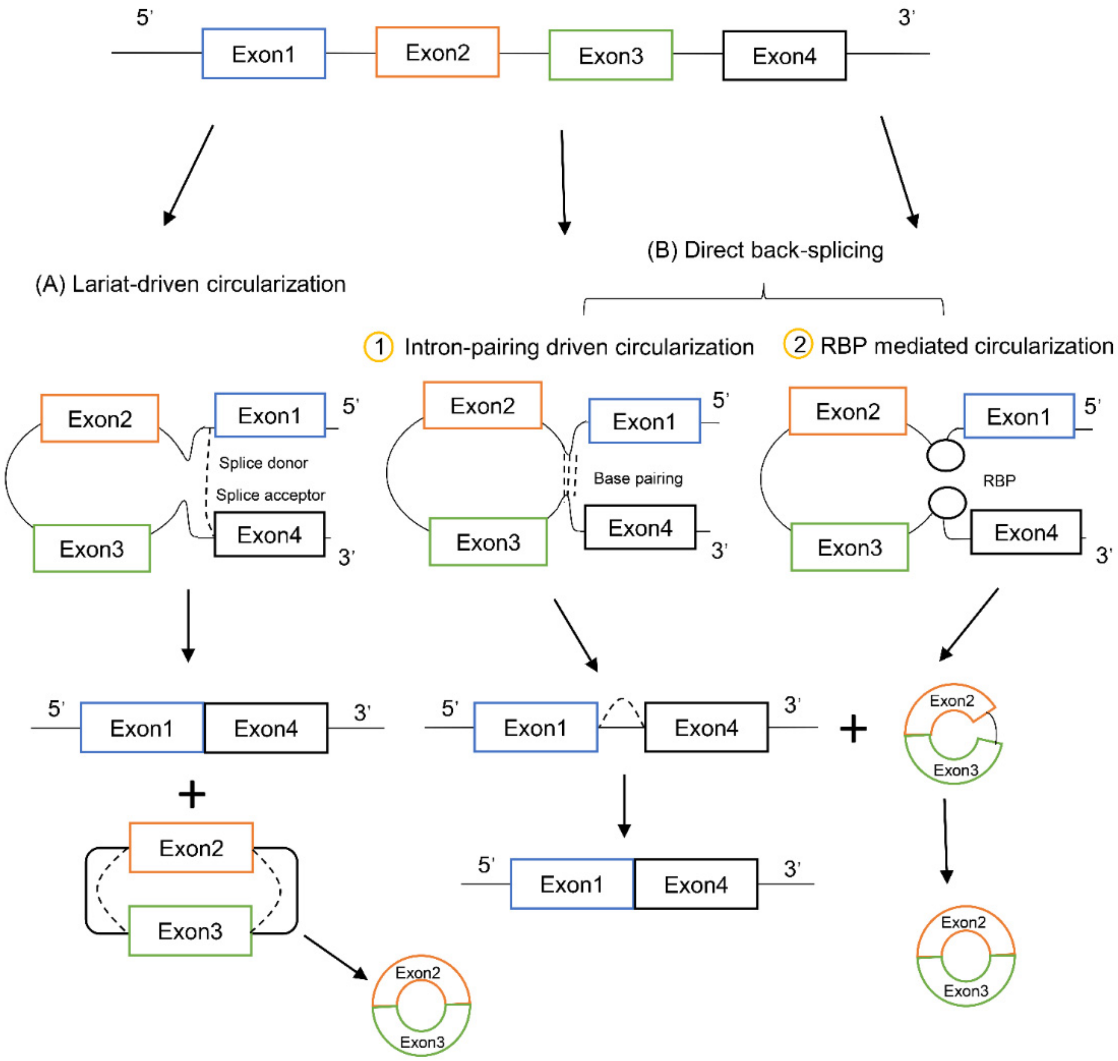


**Figure 2.23** circRNA biogenesis mechanism
(A) Firstly, canonical splicing occurs, a linear RNA forms comprised of exon 1 and exon 4 via exon skipping and a lariat structure containing exon 2 and exon 3 forms and then circRNA produces by back splicing. (B) Firstly, back-splicing occurs, circRNA including introns and exons forms via back-splicing and then is generated linear RNA. There are two types of that mechanism, one is (1)"intron-pairing driven circularization", circularization is promoted by the complementary sequences of the flanking introns. (2) "RBP mediated circularization", circularization is promoted by interaction between RBP and intron binding sites [82]

According to the occurrence of the circularization mechanism, the "direct

back-splicing" process takes place in two different ways, "intron-pairing driven circularization" and "RBP mediated circularization". Intronic sequences are highly effective in circRNA biogenesis in humans and flies [24]. "Intron-pairing driven circularization" induces circularization due to convergence caused by the complementation between introns flanking the exons.

circRNA biogenesis can also be regulated under certain conditions by binding of RBPs to intronic regions."RBP mediated circularization" mechanism promotes the assembly of these exons in a circular form through interaction between RBPs attached to flanking intronic sequences.

To summarize, effective factors in formation of circular RNA can be classified as two main categories trans acting and cis-regulatory elements and a combination of these factors [24].

Circularization is also affected by the properties of exons and flanking introns. Exon length in circular RNAs is more than in other expressed exonic RNAs and flanking introns containing inverted repeats are also longer than average. The circRNAs defined to date always contain GT-AG pair in canonical splice sites [19]. The flanking introns in the 500 bp region of the backsplice sites of the circular RNA-forming genes are supported with repeat elements. Unlike the 500 bp length, some circRNAs contain almost 200 nt and shorter flanking introns [19, 46].

### 2.7.1 Trans-acting proteins and RNAs

Various trans-acting proteins are required as splicing regulators and are responsible for the formation and expression of circular RNA at the post-transcriptional level. Trans-acting factors are divided into four groups according to their different mechanisms in the formation of circRNA: RNA binding proteins, spliceosome factors, cleavage factors, and RNA helicases (Figure 2.24).

The other factors except for cis/trans-acting elements that affect the efficiency of circularization are involved. The circularization efficiency is thermodynamically related to the strengthening of the hairpin between the repeats [26].

#### 2.7.1.1 Spliceosome factors

The spliceosome is a complex composed of five different proteins that are responsible for removing the intervening regions from the precursor mRNA [83]. Spliceosomes in eukaryotes are structures that consist of different small proteins U1, U2, U4, U5, and U6 and can contain binding sites in both exons and flanking introns of pre-mRNA for

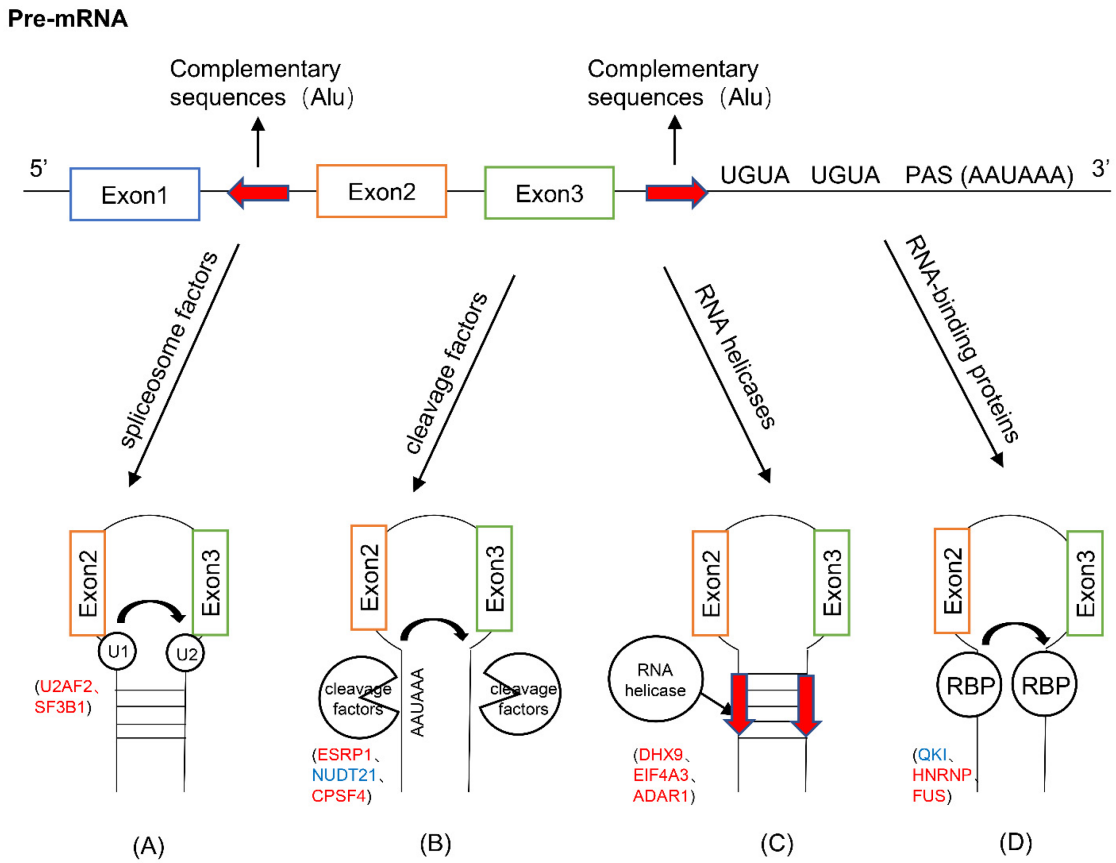


**Figure 2.24** The summary of the effects of trans-acting factors in biogenesis of circRNA

(A) Spliceosome factors can bind to relevant pre-mRNA sites. (B) Cleavage factors can mediate the attaching of cis-acting elements to pre-mRNA. (C) RNA helicases can assemble or unwind flanking complementary Alu sequences. (D) RNA binding proteins bind to flanking introns in pre-mRNA and induce circularization, by bringing together ends of exons. In addition, trans-acting factors in red demonstrate increased expression in cancer, and the blue ones show decreasing expression [82]

that spliceosomal complex. That structure can be effective in circular RNA biogenesis by interacting with binding sites in pre-mRNA. It has been observed that expression changes in spliceosome core components are efficient in circRNA formation. It has been observed that expression changes in spliceosome core components are efficient in circRNA formation. When a U2 snRNP associated with the 5' splice site is consumed, the production of pre-mRNA is converted to the formation of circRNA, and the amount of linear RNA decreases. Accordingly, the production of cirRNA through the spliceosomal complex can affect developmental period of serious diseases as cancer [82].

#### 2.7.1.2 Cleavage Factors

According to the researches, some cleavage factors can promote the formation of circular RNA by cutting the flanking intronic sequences of the exons forming circRNA. These circular RNAs are associated with cellular events such as tumor proliferation and migration in various cancers or carcinomas [82].

#### 2.7.1.3 RNA helicases

Biogenesis of circRNA is affected by RNA helicases activity which is a family of proteins that are responsible for splicing and translation by unwinding the paired RNA sequence thus production of circRNAs is prevented.

### 2.7.2 RNA binding proteins

There is a class of proteins that regulate gene expression by binding to specific regions in the genome. RBPs have similar functions in the genome to these proteins. RNA-binding protein (RBP) is one of the critical proteins that plays an effective role in gene expression by regulating splicing mechanism, gene expression, mRNA stability, translation, and localization [84]. RBPs are involved in post-transcriptional processes such as transcription, stabilization, and translation because of containing RNA binding sites. RBPs are also involved in the control of circRNA biogenesis by binding to cis elements in adjacent introns [18]. Proteins binding intronic repetitive sequences could promote or obstruct backsplicing so as to forming circle structure (Figure 2.25) [18, 26]. Besides RBPs, the properties of sequences also affect circularization. Reverse complementary sequences such as Alu repeats induce circularization.Also, circRNA-producing exons have been found generally contain more RBP binding sites than non-circular exons in host genes and the same expression genes [85]. circRNA localization and function can be regulated by RBPs,too [86]. Additionally, one RBP can contribute to more than one process. Some RBPs can be expressed in all tissues according to their function, while others can be expressed in only one tissue, they are tissue specific [87, 88]. The roles of RBPs change according to their location in the cell [89, 90]. Abnormal changes in RBP expression are associated with diseases as can cause expression changes in target transcripts [91–93]. It is supported with various studies hundreds of RBPs are expressed abnormally under disease conditions.

RBPs involved in circRNA biogenesis have been compiled from various studies. RBPs such as QKI, HNRNPL, FUS, and MBL/MBNL1 can be found in the RBP-related processes of circRNAs [1]. Some RBPs promote circRNA formation, while others can inhibit circRNA formation. ADAR, QKI and MBL that are responsible RBPs in

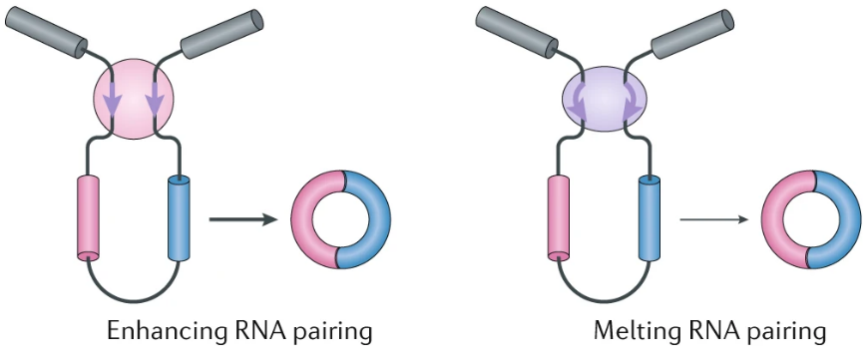


**Figure 2.25** RBPs could bind to enhance or reduce backsplicing[18]

alternative splicing are related with specific circRNA biogenesis (Figure 2.26.

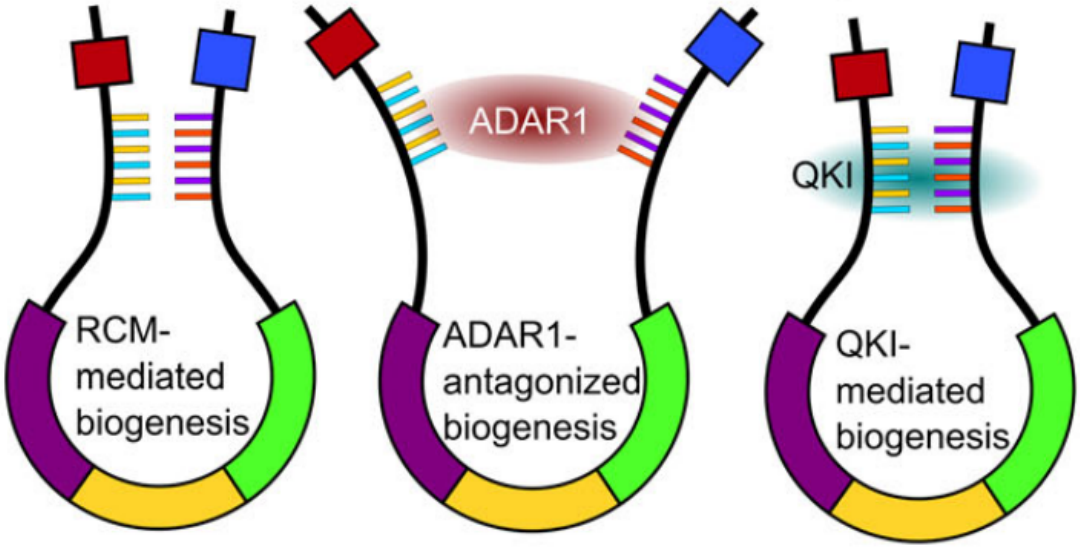


**Figure 2.26** Reverse complementary sequences, ADAR and QKI effect on circRNA biogenesis [57]

While ADAR has a reducing effect on circRNA biogenesis with its weakening effect on base pairing, QKI protein has an promoting effect on that process [57]. ADAR enzyme can convert adenosines which are in the double-stranded regions to inosine and so inhibit circRNA formationc[26].

MBL splicing factor which is one of the essential players in back-splicing regulation of pre-mRNAs through the development stage was found firstly as trans acting factor. The effect of MBL protein on the formation of circular RNA was experimentally analysed on the formation of circular RNA derived from the second exon of the Mbl gene.It was observed that the presence of MBL binding motifs in the flanking introns and the

binding of MBL proteins to these regions are related to formation of circMBL. Morever, downregulation of MBL proteins has decreased the circMbl levels.

MBNL-family proteins are comprised of MBNL1, MBNL2 and MBNL3 is effective in the skeletal, nervous system, and cardiac in human.

QKI is the most typical example of RNA-binding protein involved in circRNA biogenesis seen in in vivo and in vitro models. It has been observed that it promotes circularization of circRNAs in ataxia, schizophrenia and glioblastoma diseases from epithelial to mesenchymal by binding to regions flanking the circRNA-forming exons (Figure 2.27) [94].

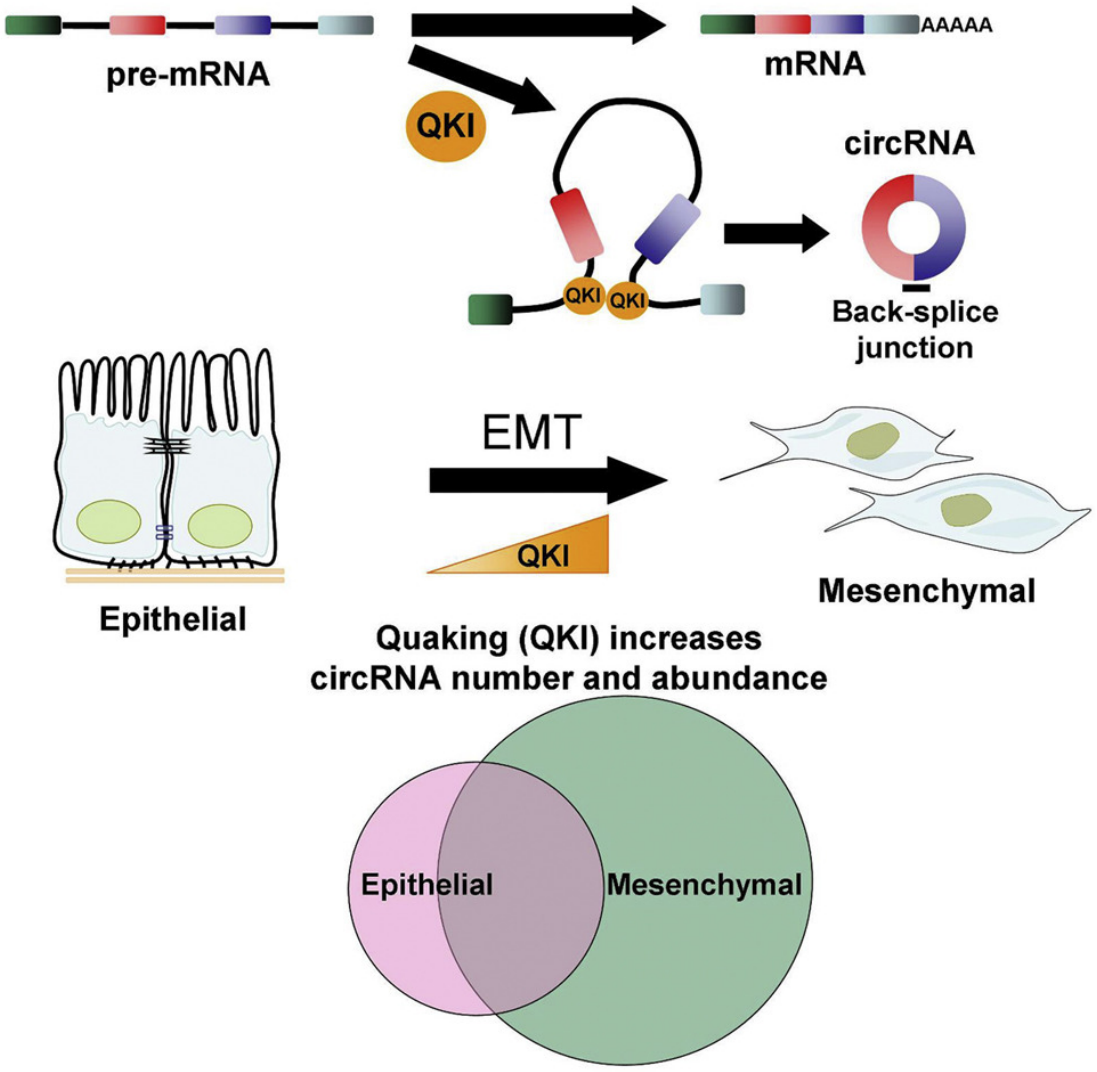


**Figure 2.27** QKI could increase circRNA number from epithelial to mesenchymal.[94]

Therefore, these proteins defined as a pre-mRNA splicing factor [21]. Also, it has been suggested in various studies that QKI protein has an oncogene by promoting

proliferation in colon cancer and a tumor suppressor function by inhibiting progression in prostate cancer [75].

EMT plays important roles in various developmental stages such as cancer progression and tissue fibrosis, embryo development involves differentiation in cell morphology and in gene expression patterns [95].

HnRNPs and SR Proteins: The formation of 490 nucleotide long circRNA that occurs from the Laccase 2 gene, depends on the interaction between SR and hnRNPs proteins having an important role in pre-mRNA splicing with flanking introns. RBPs can suppress or promote the biogenesis of circRNA according to the relation between these proteins and genes.

RBM20: RBM20 is a regulatory protein promoting circRNA formation by binding to the flanking introns of circularizing exons of the Titin gene in cardiac tissue.

DHX9: DHX9 protein facilitates backsplicing and promotes circRNA formation by interacting with Alu repeats.

NOVA2: NOVA1 and NOVA2 of neuronal RBPs are important regulators in alternative splicing, and it has been suggested that loss of function of nova2 leads to a decrease in circRNA level in the embryonic cortex.

FUS: FUS protein emerging splicing factor of motor neurons of mouse embryonic stem cell binds to the flanking introns splice junctions of exons. Brain tissue neuronal cells produce a lot of circRNAs and knockout of FUS affects circRNA biogenesis.

Considering all these examples, different circRNAs can be generated as a result of interaction with RBPs (Figure 2.28).

Despite all referred roles of RBPs, that is unclear how they act in the biogenesis of circRNA [18].

RBPs can also interact with circRNAs except for the formation stage. Interactions of protein and circRNA adjust various cellular physiology such as cell differentiation, cell proliferation, apoptosis, and cancer cell metastasis.Interactions of protein and circRNA adjust various cellular physiology such as cell differentiation, cell proliferation, apoptosis, tumorigenesis, and cancer cell metastasis. The role of between circRNAs and RBPs in tumorigenesis has been determined by updated techniques like RNA pulldown and RNA immunoprecipitation. Clarifying this relationship will provide a new perspective on cancer diagnosis and treatment in the future [21].

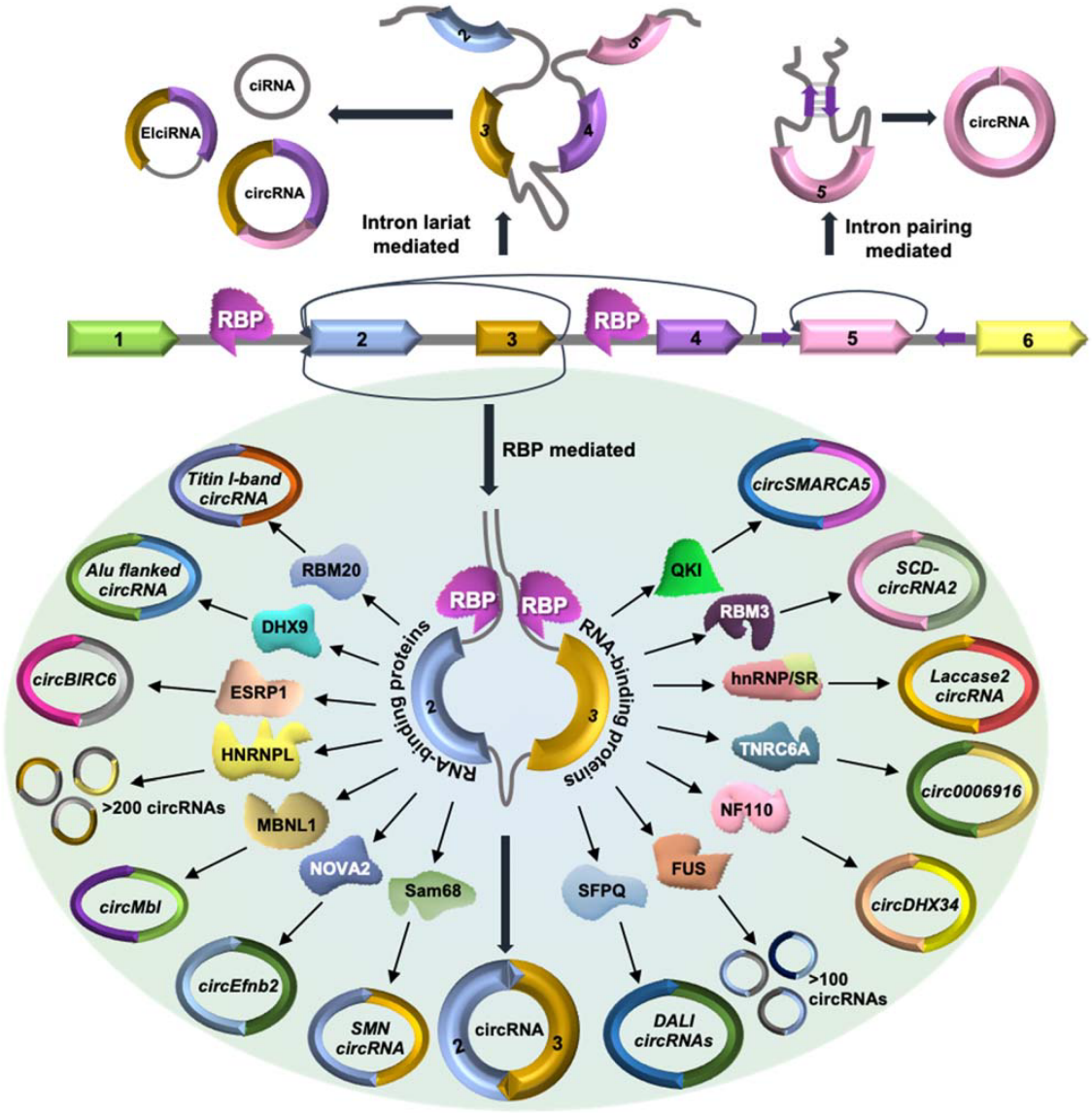


**Figure 2.28** Interaction of pre-mRNAs with RBPs induces the formation of multiple types of circRNAs[84]

### 2.7.3 Cis-Regulatory Elements

Reverse complementary sequences as the Alu element and direct binding sites of some splicing factors are effectual in the circularization of RNA. Some protein factors may play a role in circular RNA biogenesis by disrupting cis-acting elements. DHX9, a nuclear RNA helicase, is able to inhibit the function of Alu elements, which are repetitive sequences. DHX9 interacts with an isoform of ADAR, causing a change in the amount of circRNA production.The loss of DHX9 causes a twofold increase in the amount of circRNA.

### 2.7.3.1 Intronic complementary sequences

RNA pairs that are necessary for circRNA formation usually consist of complementary inverted repeat sequences such as Alu elements and sometimes non-repetitive complementary sequences [18, 82]. Internal exons and flanking intronic repeats have proven to be necessary for the biogenesis of circRNAs.

The flanking introns in the back splice sites of circRNA are longer than the other introns, but with some exceptions. Thus, these complementary sequences in flanking introns are suitable to promote circularization [19].

Also, the location of intronic complementary sequences is important for circularization. While normal splicing occurs in complementary sequences within the same intron and occurs linear RNA, intronic complementary sequences in flanking introns promote circularization (Figure 2.29).

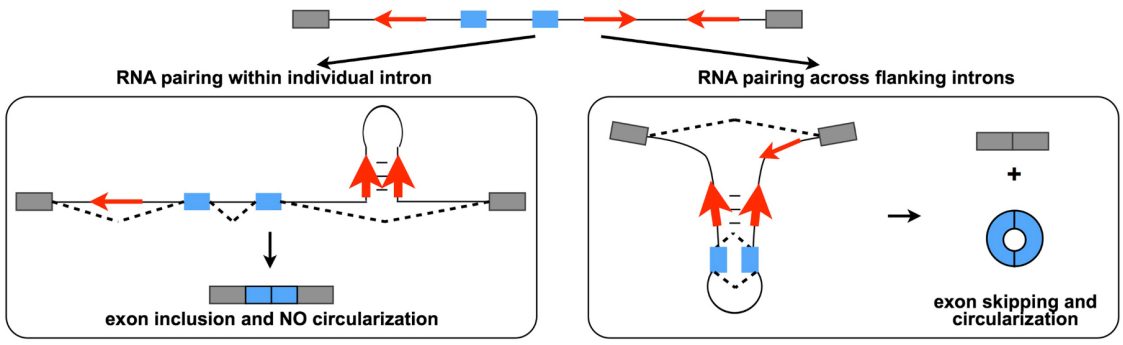


**Figure 2.29** Location of intronic complementary sequences can promote or not circularization. Constitutive splicing occurs in complementary sequences within a single intron (left) Circular RNA is formed as a result of flanking intron RNA pairing (right)[39]

circSry that originated from the mouse Sry gene is circled by intronic complementary sequences(ICS), early studies in mice demonstrated that ICSs boost formation. The ICS required to induce circRNA formation can vary between 20-40 nucleotides [18].

A many different studies have been to prove the necessity of short intronic repeats in circrNA formation. In one of these studies with ZKSCAN1, a vector was created by placing a certain length of exon region between ZKSCAN1 introns and it was concluded that short intronic repeats (30 -40 nt) are required to produce circular RNA as well as exonic elements. According to the results from the ZKSCAN1 and HIPK3 loci, some of the short repeats, but not all, support circular RNA production. ICS length is also one of the factors affecting circular RNA formation. Longer repeats contain more sequences therefore have more base pairing potential for circularization and provide flexibility in the transcript and facilitate base pairing. Nevertheless some

long repeat sequences can strongly inihibit circularization [26]. Disruptions in the order of intronic sequence repeats also significantly affect circular RNA production efficiency [26]. The formation of circRNA occurs by Pol II transcription, therefore Pol II transcription elongation rate (TER) is likely to be one of the effective factors in circRNA formation. High TER can promote circRNA formation by increasing the probability of pairing between the intronic complementary sequence. In addition, high TER could increase the likelihood of backsplicing reducing linear splicing [18].

The circRNAs discovered only after pre-mRNA transcription indicate that a non-negligible proportion of the back splicing could occur post transcriptionally [18]. It is now recognized that some introns are slowly or posttranscriptionally spliced thereby allowing the opportunity for back splicing reactions to occur [96]. Backsplicing is thought to be a slow process, as the pairing of complementary introns takes time. Consequently, linear RNA formation occurs if there is not enough time for base pairing or for spliceosome activity [26]. The importance of repetitive elements was first discovered at the mouse sry locus comprised of a single exon and encircled by very long complementary intronic sequences. Sry circrna generation mechanism is also applicable for the generation of circular RNA from human genes [97]. Although efficiency of circularization may increase based on simple thermodynamics that enhancing the complementary hairpin between the flanking intronic repeats, every pair of inverted repeats can not able to support circularization [26]. Watson- Crick base pairs between intronic coupling are more efficient for circularization. Alu elements constitute more than 10% of the human genome and are repetitive elements facilitating the biogenesis of endogenous circRNAs [24]. Alu elements play a critical role in circularization in some genes such as ZKSCAN1, HIPK3, and EPHB4, while inverted DNA transposon repeats play a role in some other genes. Reverse splicing events originated from junction of downstream splice donor and upstream splice acceptor sites associated with presence of Alu elements in the flanking introns.The inverted ALU repeats increase exonic circRNA formation 5 fold [19]. Even when the flanking introns scrambled, formation of circular structure occurs. Same intronic repeat sequences at various cell are not be able to provide equal efficiency, since producing of circular structure is cell or tissue specific. Even if intron length is not directly related to circular RNA generation, the probability of circular RNA formation may be increased as longer introns may contain more Alu. Additionally, the amount of space and competition between the complementary sequences in the enclosing of introns may result in the formation of different circRNAs. Despite all these factors, not all matched intron sequences support circular structure formation, circRNA biogenesis can be prevented forming the hairpin structure[82]. Also, repetitive sequences and inverted sequences are not the only essential factors in the formation of circular

structures, as repetitive sequences are not present in lower eukaryotic models and circular RNA formation can occur in genes lacking intronic complementary sequences [79]. In addition to all mentioned, the presence of inverted repeats is not sufficient evidence to express that circular RNA will be produced with the same efficiency in all cell types [26]. Another factor in circularization is exon length. Generally, longer exons are circularized more efficiently because of flexibility in the transcript. Although the human genome consists of repeat sequences >45%, circularization is restricted to specific exons because of that mechanism is multifactorial [26].

Especially in mammals, there are abundant non-coding RNAs (lncRNA) longer than 200 base pairs. circRNAs are a group of long non-coding RNAs which are tissue-specific and formed by covalent bonds. Circular RNAs in the genome are grouped categorically and functionally according to their location in the genome and the relationship between protein-coding genes. According to their position in the genome they are collected four groups; intronic circRNA, exonic circRNA, antisense circRNA, intergenic circRNA [98].

Considering the information related to the biogenesis of circRNAs in the literature; even with all the research so far, it is seen that there are still uncertainties. In this study, we aimed to clarify these uncertainties and gain new perspectives on the circRNA biogenesis mechanism.

# 3
# METHOD

## 3.1 Datasets

The information of species, circatlas id, position, strand, circRNA type, host gene, tissue, tissue specificity index, algorithm has been downloaded from the circAtlas [67]database. eCLIP and RNA-bind-N seq experimental data were obtained from ENCODE. The target region information to which RBPs bind have been obtained from eCLIP and RNA-bind-N seq experiments performed on hepG2 and K562 cells. Data containing expression values of 356 different RBPS in 40 different tissues have been used[87]. All these data have been analyzed in R programming using tidyverse packages.

## 3.2 Database Information

circAtlas involves more than one million circRNAs found in 6 different species and their expressions, as well as expressions of their parent genes. IRES sites, miRNA binding sites, and RBP binding sites of circRNAs can be predicted using circAtlas. There are circRNAs detected using maximum 4 different algorithms, CIRI [99, 100], find_circ, CIRCexplorer [101], DCC [102], in this database. And this case has increased the reliability of the information used.

## 3.3 Analysis of Available Data

The obtained data from circAtlas database was processed with the R programming language and aimed to investigate the parent gene expression profiles of tissue-specific circRNAs.These analyzes were performed using tidyverse packages in R programming. In this way, some new approaches have been proposed for uncertainties in circRNA biogenesis.

## 3.4 Filtration of Tissue-specific circRNAs

circRNA species can be exonic, intronic, antisense, intergenic. Among these species, we performed our analysis with exonic, intronic and antisense ones. Among these circRNAs, whose tissue specificity index and the number of tissues were equal to 1 and determined by at least two algorithms and longer than 100 nt [103] were selected. In this way, we made our analysis on 111,704 tissue-specific circRNA transcripts found in only one tissue. In the circAtlas database, circular RNAs have two different expression values, max and mean, according to the tissues in which they are found, and 3 different expression values for gene expression as Q1, Q3, and mean expression. However, these expression values were found separately for the parent gene and circRNA, and these values were combined in a single table over circRNA. Since the mean values of circRNA isoforms are the same in our data, they were not taken into account and were excluded from the analyses.

## 3.5 Calculation of Z-score of Datasets

The z-score was calculated over the mean expression value for gene expression, and those with a z score above 3.5 (which can be considered as the threshold value) were selected.

Thus, the expression values that are higher than the average at a certain rate were selected. z score allows us to compare how far a data is from the mean.

$$z = (x - \mu)/\sigma \tag{3.1}$$

$\mu$ is mean, $\sigma$ is standard deviation and x is sample value

## 3.6 Categorization of Tissue-specific CircRNAs

As a result of these selections, obtained tissue-specific circRNAs can be divided into 3 different categories mainly. Firstly; gene expressions are approximately the same in any two tissues, but their circRNA expressions are different. In the second case, circRNA and the parent gene expression values have a linear correlation. The last group is that circRNA and their parent gene expressions are located in different tissues. In this group, a match was made between the tissues of the circRNAs and the parent genes of them, of which the unmatched ones were selected.(Table 3.1)

**Table 3.1** unmatched circRNAs and their parent genes

| circrna | tissue | host | tissue_circ | match | match2 |
|---|---|---|---|---|---|
| hsa-PDE6A | Retina | ENSG00000132915.10 | Testis | 0 | 0 |
| hsa-PDE6A | Retina | ENSG00000132915.10 | Brain | 0 | 0 |
| hsa-CTRC | Pancreas | ENSG00000162438.11 | Colon | 0 | 0 |
| hsa-CNGB1 | Retina | ENSG00000070729.13 | Placenta | 0 | 0 |
| hsa-CNGB1 | Retina | ENSG00000070729.13 | Brain | 0 | 0 |
| hsa-IMPG1 | Retina | ENSG00000112706.11 | Testis | 0 | 0 |
| hsa-CELA2A | Pancreas | ENSG00000142615.7 | Brain | 0 | 0 |
| hsa-RP1L1 | Retina | ENSG00000183638.5 | Brain | 0 | 0 |
| hsa-IMPG2 | Retina | ENSG00000081148.11 | Stomach | 0 | 0 |
| hsa-IMPG2 | Retina | ENSG00000081148.11 | Brain | 0 | 0 |

Our tissue-specific analysis proceeded from this last group. Among the circRNAs in this group, the highest 50 were selected according to their expression values.

In addition to the average gene expression values given in CircAtlas, the percentage of genes was also calculated. Thus, the gene percentages of the parent genes of the remaining circRNAs in the belonging tisssues have been calculated.

Firstly, the mean value of the gene expressions, and then the percentage of the gene mean values have been calculated. In this way, those whose haveIn addition to the average gene expression values given in CircAtlas, the percentage of gene mean values have also been calculated. In this way, those whose have negligible expression value of parent genes in tissues of circRNAs have been taken. As a result of all these elimination processes, we have 15 circRNAs left. Since the uterus is tissue in which the parent genes of these circRNAs are the most expressed, we preferred to proceed with our thesis over the uterus. The circRNAs that are in different tissues from their expressed parent genes exist colon, liver, placenta, prostate and brain except for the uterus.

circRNA biogenesis is associated with many independent factors such as complementary flanking introns, the effect of RBPs, and thermodynamic conditions. In this study, we tried to demonstrate the relationship between biogenesis and RBPs.

We can call the parent genes of uterus circRNAs as target RNA. RBPs interacting with these target RNAs affect various cellular processes [104]. Therefore, identification of RBP binding sites is critical. Enhanced Crosslinking Immunoprecipitation (eCLIP) and RNA Bind-n-Seq (RBNS) are the most preferred two option to detect the binding site and motifs on the target RNAs of RBPs. The necessary RBP information to identify RBPs that bind to target genes was obtained from these two experimental data. RBPs in eCLIP and RBNS experimental data have intersected with parent gene regions of circRNAs in uterus. Thus, RBPs binding to the chromosomal regions of parent genes expressing circRNAs have been determined. In addition, the z scores of the expression values of these RBPs have been calculated and the first 5th ones in the uterus have been selected among them. Although these RBPs are not specific to the uterus, they are in the top 5 most expressed compared to other tissues.

Thus, as a result of these analyzes, we can say that we have identified RBPs that are thought to play a role in circRNA expression.

# 4
# RESULTS AND DISCUSSION

CircRNAs are a group of non-coding RNAs formed by back-splicing the 5' and 3' ends of the pre mRNA into a loop structure. The circRNA formation mechanism differs according to its genomic origin and exon-intron inclusion.

Exonic circRNAs are formed by direct back splicing and exon skipping and contain single or multiple exons. A bridge is occured with RBP pairing, thereby circular structure is originated in the direct back-splicing stage [44, 105]. Exonic intronic circRNAs (eIcirRNA) are also generated by a similar mechanism as ecircRNA. Unlike exonic circRNAs, eIcircRNAs contain both exons and introns. ciRNA formation requires 7 nt GU rich elements at the 5' splice site and 11 nt C-rich elements at the 3' splice site.

Considering the circRNA formation mechanisms, it has been seen that repetitive flanking introns and non-repetitive but complementary sequences and RBPs are effective in circRNA formation. RBPs such as ADAR, MBL, and QKI are known to play a role in alternative splicing and ecircRNA formation[44].

We aimed to analyse the relationship between circRNA and gene expression, in our research. After analyzing the circRNA-gene expression data from taken circATLAS, we have taken the top 50 expressions of circRNAs according to the gene expression distribution. As a result, we encountered 3 different cases.

First, while gene expression values are close to each other in two different tissues, circRNA expression values are different (Figure 4.1). In some samples, circRNA expression can be seen in only one tissue, even though the main gene expressions are close to each other in different tissues. (Figure 4.2 and Figure 4.3). The fact that similar gene expression values in different tissues have different circRNA expressions can be explained by the tissue specificity of circRNAs (Figures 4.1, 4.2 and 4.3).

In the second case, circRNA and gene expression values may have a linear correlation (Figure 4.4).

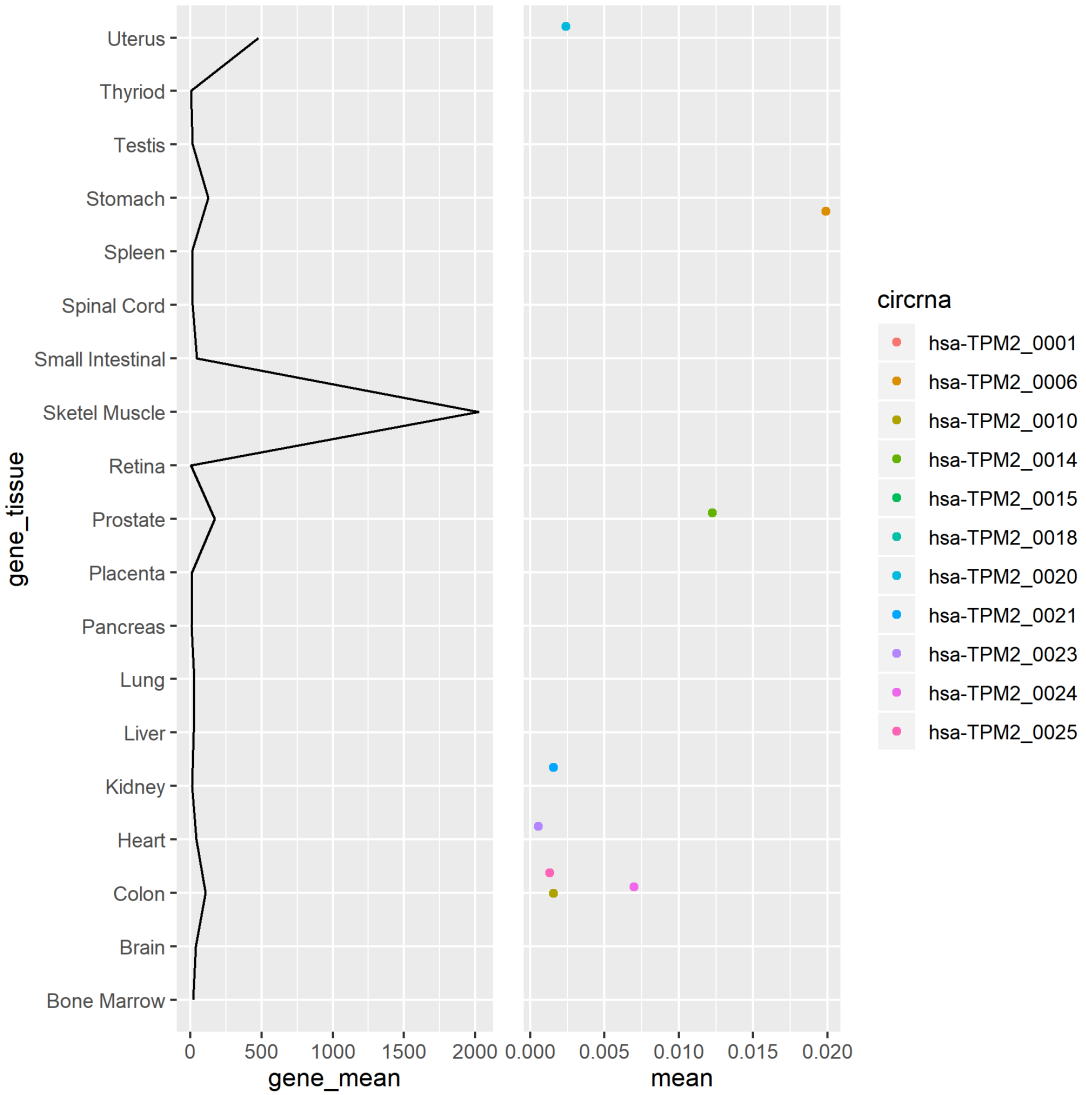


**Figure 4.1** hsa-TPM2 circRNA expression profile
While the parent gene of hsa-TPM circRNA has nearly same expression in both stomach and prostate tissue, its circRNA expression is different. *gene_mean*: mean value of parent gene expression, *mean*: mean expression value of TPM circRNAs

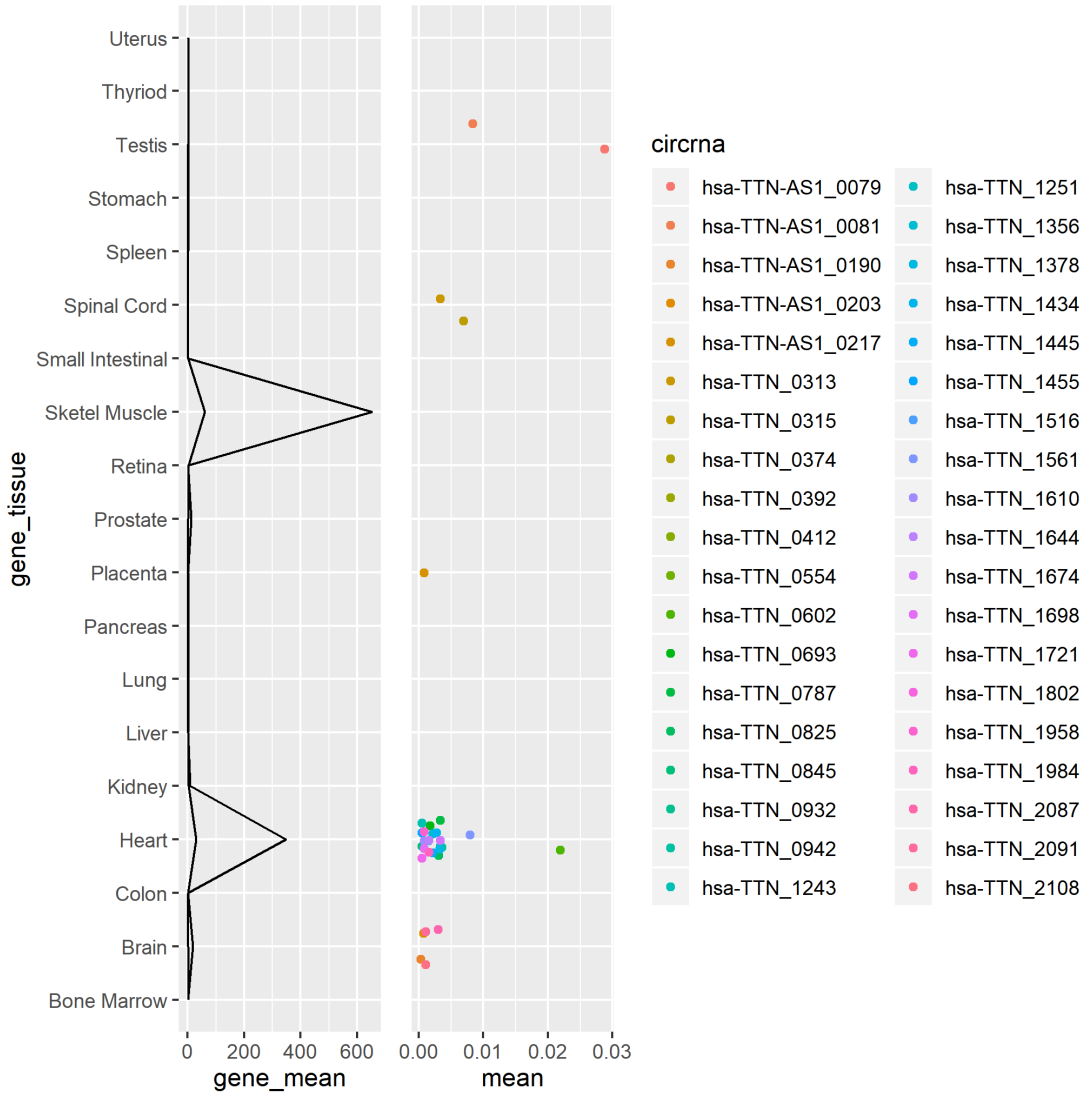


**Figure 4.2** hsa-TTN expression profile
While the parent gene of hsa-TTN circRNA has nearly same expression in both sketel muscle and heart tissue, its circRNA expression has showed accumulation in the heart. *gene_mean*: mean value of parent gene expression, *mean*: mean expression value of TTN circRNAs

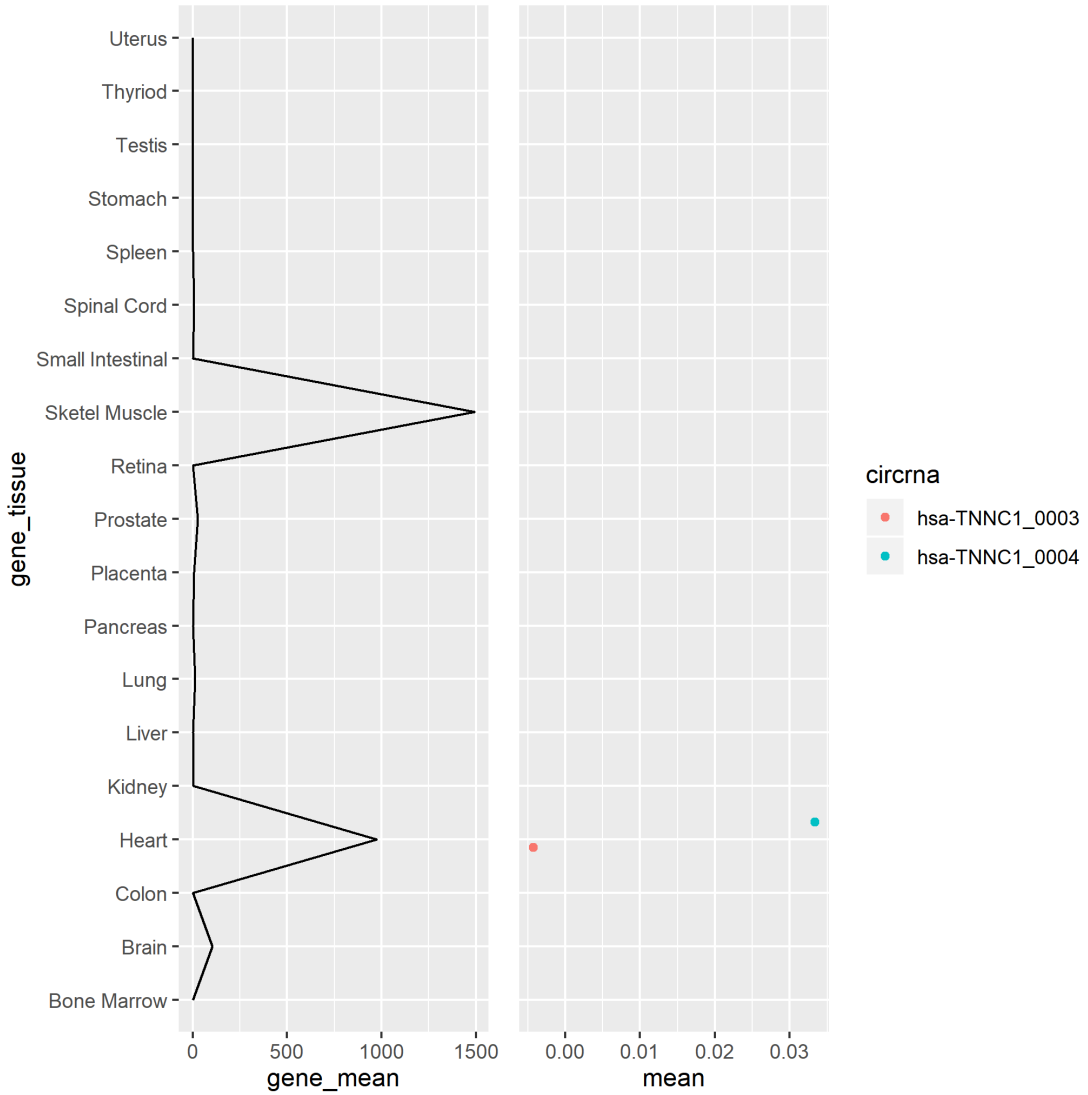


**Figure 4.3** hsa-TNNC1 expression profile
The same way with hsa-TTN, the parent gene of hsa-TNNC1 is expressed nearly the same in the skeletal muscle and heart, the circRNA is expressed in only heart tissue. *gene_mean*: mean value of gene expression, *mean*: mean expression value of TNNC1 circRNAs

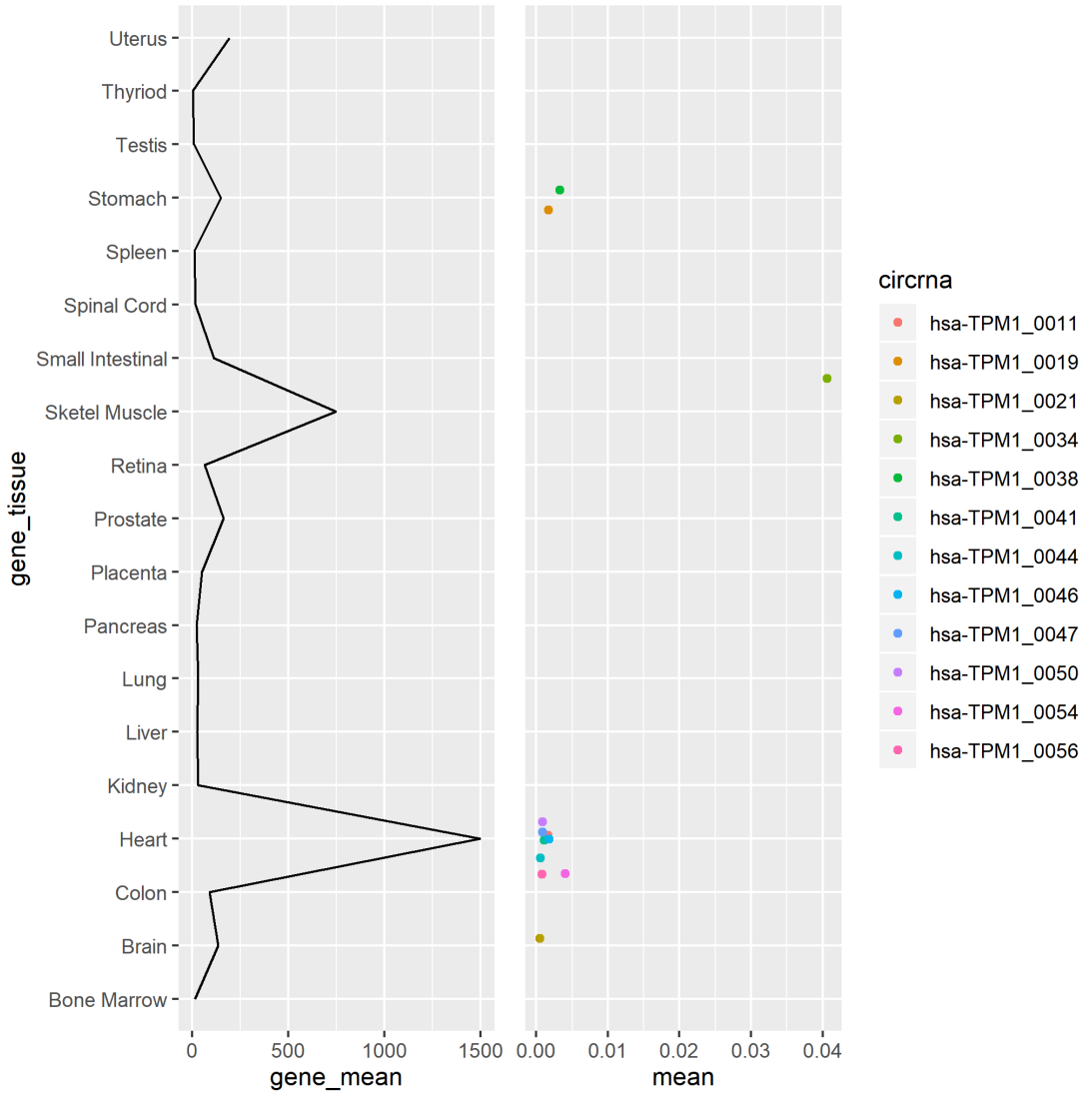


**Figure 4.4** hsa-TPM1 expression profile
CirRNA expression profile is related with its gene expression.The parent gene of hsa-TPM1 has discovered in stomach, skeletal muscle an heart.Its circRNA has exist in linearly with parent gene expression. *gene_mean*: mean value of gene expression, *mean*: mean expression value of TPM circRNAs

In some cases, circRNA can not be expressed, even though the parent gene is highly expressed(Figure 4.5).

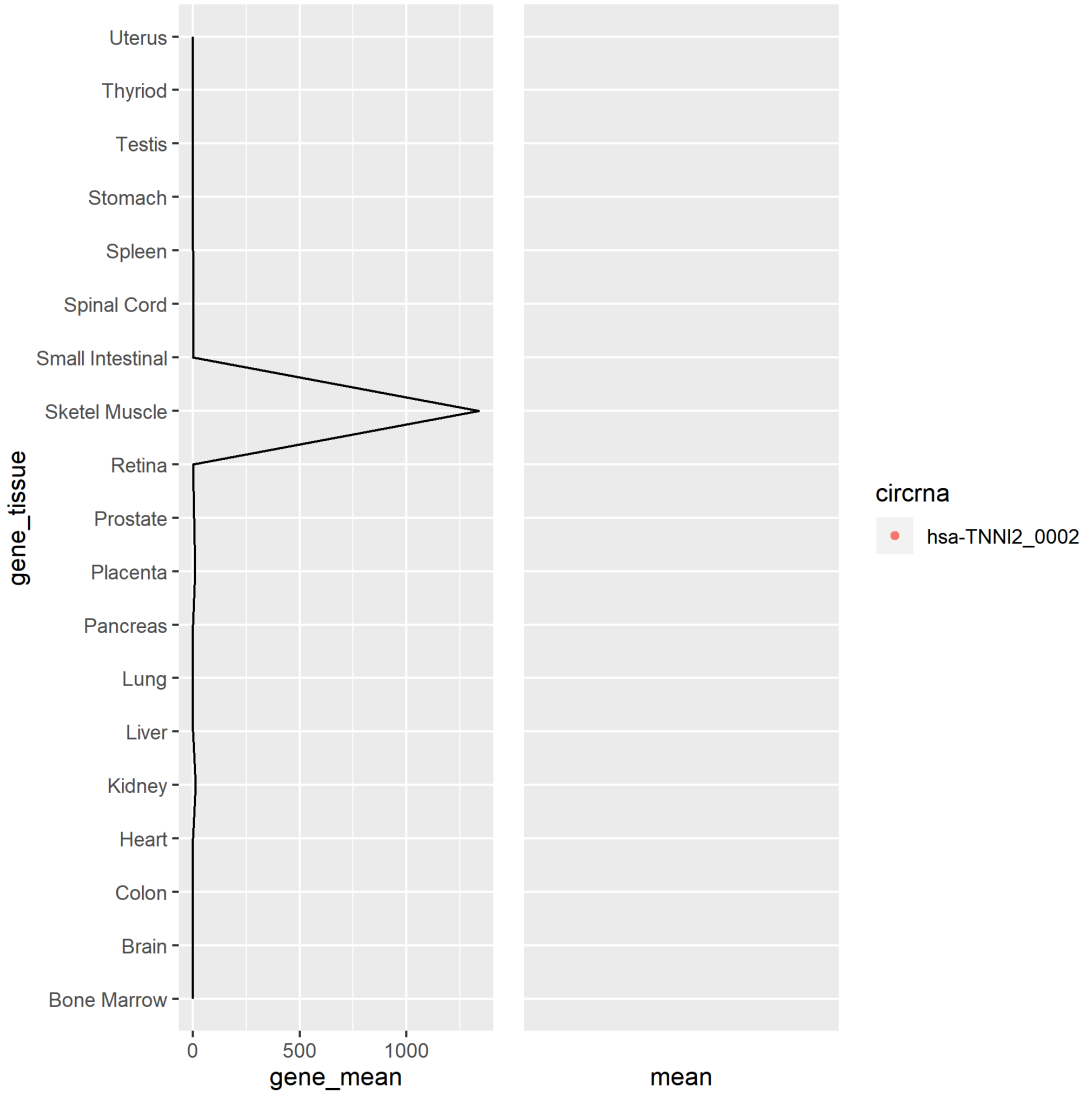


**Figure 4.5** hsa-TNNI2 expression profile
Although parent gene of hsa-TNNI2 circRNA expressed highly, circRNA expression may be absent. *gene_mean*: mean value of gene expression, *mean*: mean expression value of TNNI circRNAs

The last group is that circRNA and their parent gene expressions are located in different tissues(Figure 4.6).

In this study, we tried to shed light on the uncertainties in the biogenesis mechanism by using circRNAs expressed independently of the parent gene. Since circRNA expression is known to be originated from pre-mRNA, it was thought that should be located with the parent gene in the same tissue under normal conditions. While there is the parent gene expression in any tissue, circRNA expression may not be present due to cellular events, that is, the presence of circRNA may not be needed at that time. In another

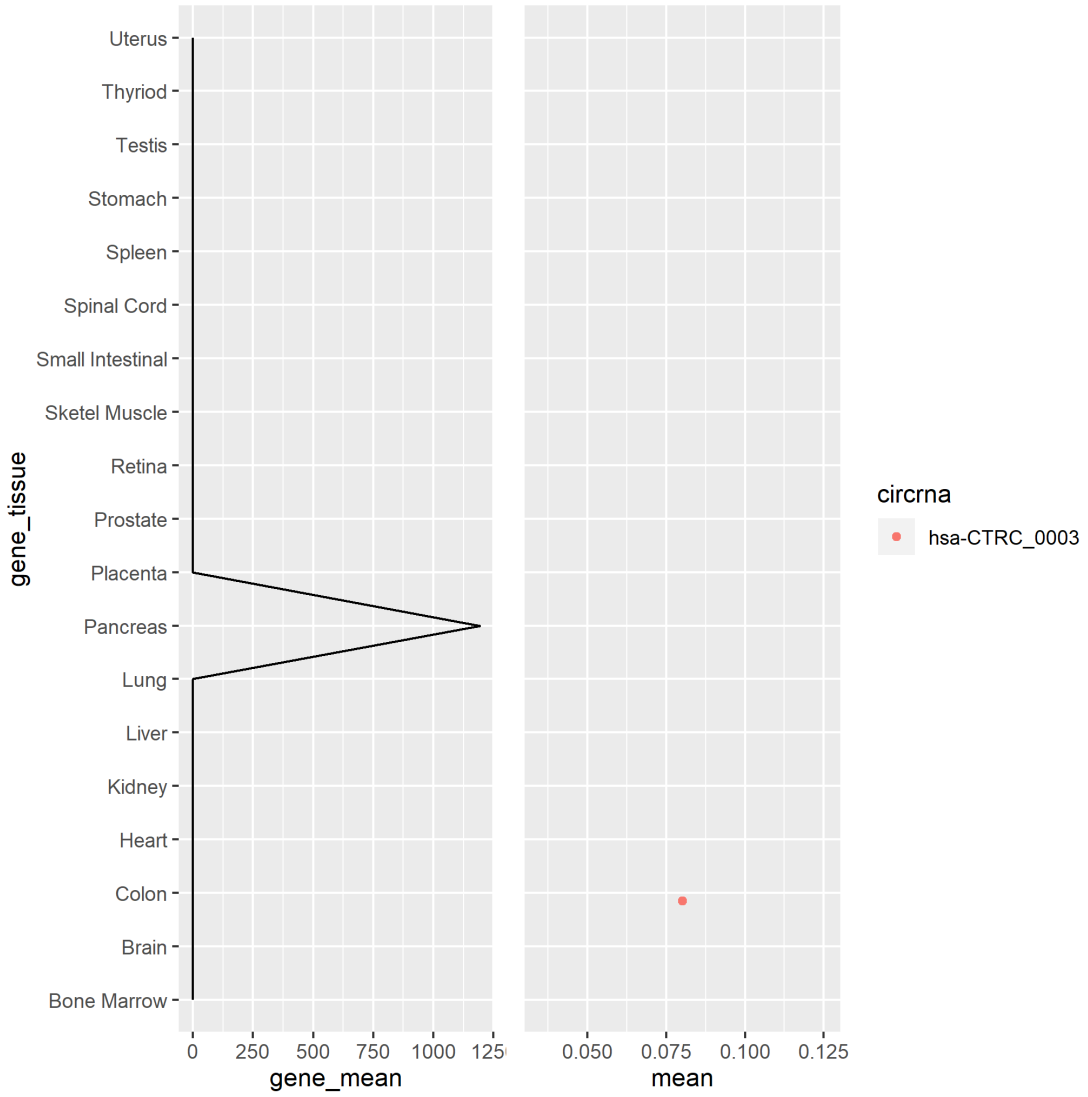


**Figure 4.6** hsa-CTRC expression profile
The parent gene of hsa-CTRC cirCRNA is highly expressed in pancreas, but the circRNA is expressed in colon tissue. *gene_mean*: mean value of gene expression, *mean*: mean expression value of CTRC circRNAs

case, circRNA may not be expressed at the same rate with gene expression. Canonical splicing can also take place from the same pre-mRNA so linear RNA can be occur. However, when the circRNA expression in any tissue, the absence of their parent gene in there has not been associated with the widely known mechanism of biogenesis. Therefore, we have selected samples whose circRNAs and their parent genes are not located in the same tissue from the circAtlas data. (Table 4.1)

**Table 4.1** Expression and gene percentage values of circRNAs that are not present in the same tissue with their parent gene

| circrna | circ_tissue | gene_mean | max | mean | parent_gene | gene_perc[a] |
|---|---|---|---|---|---|---|
| hsa-SLC22A10_0014 | Uterus | 0.00050 | 0.04880 | 0.01220 | ENSG00000184999.11 | 0.0037280 |
| hsa-FABP6_0002 | Uterus | 0.00300 | 0.04800 | 0.01200 | ENSG00000170231.15 | 0.0090736 |
| hsa-CNTNAP2_0054 | Uterus | 0.00200 | 0.03920 | 0.00980 | ENSG00000174469.18 | 0.0031639 |
| hsa-KIF18A_0037 | Uterus | 0.00150 | 0.03920 | 0.00980 | ENSG00000121621.6 | 0.0054618 |
| hsa-CYP2B6_0004 | Prostate | 0.00333 | 0.02175 | 0.00725 | ENSG00000197408.8 | 0.0090101 |
| hsa-FSTL5_0018 | Uterus | 0.00250 | 0.02400 | 0.00600 | ENSG00000168843.13 | 0.0037218 |
| hsa-CTRC_0003 | Colon | 0.00227 | 0.06894 | 0.00406 | ENSG00000162438.11 | 0.0001902 |
| hsa-CDH7_0020 | Uterus | 0.00050 | 0.01200 | 0.00300 | ENSG00000081138.13 | 0.0044218 |
| hsa-KIF18A_0035 | Uterus | 0.00150 | 0.01200 | 0.00300 | ENSG00000121621.6 | 0.0054618 |
| hsa-RP11-110L15_0003 | Liver | 0.00007 | 0.01663 | 0.00237 | ENSG00000257683.1 | 0.0061586 |
| hsa-GRK1_0001 | Colon | 0.00520 | 0.02550 | 0.00150 | ENSG00000185974.6 | 0.0075952 |
| hsa-CPB1_0008 | Placenta | 0.01863 | 0.03712 | 0.00088 | ENSG00000153002.11 | 0.0020336 |
| hsa-PLCL1_0077 | Placenta | 0.00014 | 0.02930 | 0.00070 | ENSG00000115896.15 | 0.0034269 |
| hsa-KCNQ2_0008 | Placenta | 0.00056 | 0.01856 | 0.00044 | ENSG00000075043.18 | 0.0018589 |
| hsa-CYP19A1_0019 | Brain | 0.03659 | 0.01360 | 0.00040 | ENSG00000137869.14 | 0.0086628 |
| hsa-SAG_0001 | Brain | 0.01229 | 0.00389 | 0.00011 | ENSG00000130561.16 | 0.0009209 |

[a] percentage of parent gene expression in circ_tissue out of total 19 tissues

15 circRNAs have been identified, 2 of which were in different isoforms. 6 of them are in the uterus, so the uterus has been chosen as the sample tissue.

As a result of our analysis, it has been observed that the parent genes of uterus-specific circRNAs have been expressed in other than the uterus tissues. In other words, parent gene expression of uterus-specific circRNAs has not been found in uterus tissue. We aimed to relate the reasons why circRNA is expressed in uterus tissue even though not their parent gene there, with the mechanism of circRNA biogenesis.

We identified tissue-specific RBPs that we think may have a function in the biogenesis of circRNAs. More than 600 RBPs have been identified in mammals with studies so far, but most of them have not identified their functions and binding specificities [106].

According to the obtained information from the article [87], the RBPs expressed in a total of 40 different tissues, having higher expression considering z- score in the uterus than in other tissues have been selected. The expression values of these RBPs have been also ranked way including the top 5 ranks of the uterus among other tissues.(Table 4.2)

We have determined the RBPs that bind to our target genes according to the binding sites of the RBPs in the genome that intersect with these top 5 ones and the RNA bind-n-seq experimental data.(Table 4.3)

**Table 4.2** Only 10 samples from the top 5 ranked RBPs

| RBP_name | RBP_Gene | RBP_tissue[a] | RBP_exp | z-score[b] | rank[c] |
|---|---|---|---|---|---|
| CSTF1 | ENSG00000101138 | Uterus | 4.28 | 1.4872558 | 1 |
| SRSF1 | ENSG00000136450 | Uterus | 7.01 | 0.8185294 | 5 |
| ZC3H11A | ENSG00000058673 | Uterus | 6.33 | 1.1306158 | 4 |
| KHSRP | ENSG00000088247 | Uterus | 7.53 | 1.1725409 | 5 |
| PUF60 | ENSG00000179950 | Uterus | 6.84 | 1.1202220 | 3 |
| TIA1 | ENSG00000116001 | Uterus | 7.45 | 2.0632954 | 1 |
| RPS24 | ENSG00000138326 | Uterus | 11.84 | 1.4884188 | 3 |
| FXR1 | ENSG00000114416 | Uterus | 6.38 | 0.9458136 | 4 |
| SERBP1 | ENSG00000142864 | Uterus | 7.02 | 1.2077546 | 3 |
| HNRNPA0 | ENSG00000177733 | Uterus | 7.00 | 1.7997406 | 1 |

[a] Tissue with the highest expression of that RBP gene
[b] z-score value for RBP gene expression among 40 tissues
[c] Rank of tissue according to z-score of RBP gene among 40 tissues

**Table 4.3** RBPs binding to target genes as a result of RNA bind-N-seq (RBNS)

| chr | target_gene_start | target_gene_end | target_GeneID_name | start_RBP | end_RBP | RBP_name | RBP_Gene | RBP_tissue | RBP_exp | z-score | rank |
|---|---|---|---|---|---|---|---|---|---|---|---|
| chr11 | 28042166 | 28129703 | ENSG00000121621 | 28042221 | 28042330 | PUM2 | ENSG00000055917 | Uterus | 5.53 | 1.417317 | 3 |
| chr11 | 28042166 | 28129703 | ENSG00000121621 | 28058011 | 28058124 | PUM2 | ENSG00000055917 | Uterus | 5.53 | 1.417317 | 3 |
| chr11 | 28042166 | 28129703 | ENSG00000121621 | 28117674 | 28117775 | HNRNPC | ENSG00000092199 | Uterus | 8.42 | 1.011953 | 5 |
| chr11 | 28042166 | 28129703 | ENSG00000121621 | 28095862 | 28096008 | KHSRP | ENSG00000088247 | Uterus | 7.53 | 1.172541 | 5 |
| chr11 | 28042166 | 28129703 | ENSG00000121621 | 28087104 | 28087283 | KHSRP | ENSG00000088247 | Uterus | 7.53 | 1.172541 | 5 |
| chr11 | 28042166 | 28129703 | ENSG00000121621 | 28087311 | 28087464 | KHSRP | ENSG00000088247 | Uterus | 7.53 | 1.172541 | 5 |
| chr11 | 28042166 | 28129703 | ENSG00000121621 | 28104840 | 28104953 | KHSRP | ENSG00000088247 | Uterus | 7.53 | 1.172541 | 5 |
| chr11 | 28042166 | 28129703 | ENSG00000121621 | 28104903 | 28105008 | KHSRP | ENSG00000088247 | Uterus | 7.53 | 1.172541 | 5 |
| chr11 | 28042166 | 28129703 | ENSG00000121621 | 28113129 | 28113244 | KHSRP | ENSG00000088247 | Uterus | 7.53 | 1.172541 | 5 |
| chr11 | 28042166 | 28129703 | ENSG00000121621 | 28126538 | 28126678 | KHSRP | ENSG00000088247 | Uterus | 7.53 | 1.172541 | 5 |
| chr11 | 28042166 | 28129703 | ENSG00000121621 | 28129431 | 28129566 | KHSRP | ENSG00000088247 | Uterus | 7.53 | 1.172541 | 5 |
| chr7 | 145813893 | 148118090 | ENSG00000174469 | 145838273 | 145838423 | KHSRP | ENSG00000088247 | Uterus | 7.53 | 1.172541 | 5 |
| chr7 | 145813893 | 148118090 | ENSG00000174469 | 145872847 | 145873002 | PTBP1 | ENSG00000011304 | Uterus | 7.49 | 1.057281 | 4 |
| chr7 | 145813893 | 148118090 | ENSG00000174469 | 145814737 | 145814875 | TARDBP | ENSG00000120948 | Uterus | 7.35 | 1.239750 | 2 |

Considering the intersection of RNA bind-N-seq data and target genes, expressed RBPs in the uterus are PUM2, HNRNPC, KHSRP, PTBP1, and TARDBP. Of these, KHSRP and PUM2 RBPs can bind to more than one region of the gene.These RBPs can have various functions and can be present in the nucleus or cytoplasm (Figure 4.7).

As a result of the intersection of eCLIP experimental data and target genes revealed that the only bound RBP is HNRNPM.(Table 4.4)

K562, is a cell culture gotten from chronic myelogenous leukemia patients. RNA bind-n-seq data includes experimental results using this cell line. In the eCLIP experiment, K562 and the liver cancer cell line HepG2 have been used. If instead there were data from normal human tissues such as the uterus, placenta, etc, more RBPs that bind to target genes could be acquired.

HNRNP is an important member of the RBP family, which plays a significant role in transcriptional and translational regulation, alternative splicing and mRNA stabilization. There are different members of HNRNP family according to the sub groups and the differences in the binding sites. HNRNP M is one of the splicing factors usually found in the nucleus. In addition, HNRNP protein has been shown to increase exon skipping. The role of HNRNPs in regulating gene expression and changes in expression levels have been associated with different diseases, especially cancer [107, 108]. The hnRNP C is one of the most crucial RBP in the HNRNP family and is involved in splicing[108]. KHSRP, (KH-type splicing regulatory protein), is involved in splicing and modulates cellular functions and gene expression. It also promotes maturation of miRNA precursors and affects translation[109].

PUM protein reduce protein expression by mRNA degradation and thus act as a repressor. Since the targets of the PUM protein have important roles in developmental processes, their absence or variation in their expression can cause various diseases[110]. PTBP1 regulates cellular development by taking part in alternative splicing and post-transcriptional gene expressions[111]. As a result, when all these RBPs were evaluated, it was seen that almost all of them have been involved in alternative splicing.

**Table 4.4** RBPS binding to target genes as a result of eCLIP experiment

| chr | target gene start | target gene end | GeneID | name | start_RBP | end_RBP | RBP_name | RBP_Gene | RBP_tissue | RBP_exp | z-score | rank |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| chr11 | 28020619 | 28108156 | ENSG00000121621 | KIF18A | 28027884 | 28027947 | HNRNPM | ENSG00000099783 | Uterus | 7.63 | 1.290003 | 4 |
| chr11 | 28020619 | 28108156 | ENSG00000121621 | KIF18A | 28028161 | 28028211 | HNRNPM | ENSG00000099783 | Uterus | 7.63 | 1.290003 | 4 |
| chr11 | 28020619 | 28108156 | ENSG00000121621 | KIF18A | 28029034 | 28029081 | HNRNPM | ENSG00000099783 | Uterus | 7.63 | 1.290003 | 4 |
| chr11 | 28020619 | 28108156 | ENSG00000121621 | KIF18A | 28031721 | 28031776 | HNRNPM | ENSG00000099783 | Uterus | 7.63 | 1.290003 | 4 |
| chr11 | 28020619 | 28108156 | ENSG00000121621 | KIF18A | 28032273 | 28032318 | HNRNPM | ENSG00000099783 | Uterus | 7.63 | 1.290003 | 4 |
| chr11 | 28020619 | 28108156 | ENSG00000121621 | KIF18A | 28032479 | 28032524 | HNRNPM | ENSG00000099783 | Uterus | 7.63 | 1.290003 | 4 |
| chr11 | 28020619 | 28108156 | ENSG00000121621 | KIF18A | 28033503 | 28033525 | HNRNPM | ENSG00000099783 | Uterus | 7.63 | 1.290003 | 4 |
| chr7 | 146116801 | 148420998 | ENSG00000174469 | CNTNAP2 | 146136841 | 146136863 | HNRNPM | ENSG00000099783 | Uterus | 7.63 | 1.290003 | 4 |
| chr7 | 146116801 | 148420998 | ENSG00000174469 | CNTNAP2 | 146117146 | 146117257 | HNRNPM | ENSG00000099783 | Uterus | 7.63 | 1.290003 | 4 |
| chr7 | 146116801 | 148420998 | ENSG00000174469 | CNTNAP2 | 146117257 | 146117303 | HNRNPM | ENSG00000099783 | Uterus | 7.63 | 1.290003 | 4 |
| chr7 | 146116801 | 148420998 | ENSG00000174469 | CNTNAP2 | 146136627 | 146136707 | HNRNPM | ENSG00000099783 | Uterus | 7.63 | 1.290003 | 4 |
| chr7 | 146116801 | 148420998 | ENSG00000174469 | CNTNAP2 | 146141409 | 146141444 | HNRNPM | ENSG00000099783 | Uterus | 7.63 | 1.290003 | 4 |


**Figure 4.7** Biological features for each RBP of interest
Red: present, Gray: unknown

When the literature information is evaluated; The presence of circRNA has been often thought to require parent gene expression. In accordance with the attained data from circATLAS, the largest and most up-to-date database for circRNAs, we evaluated this argument and made inferences about the biogenesis mechanism. This mechanism is also known to be tissue specific. As a result of the analysis, uterus has been determined as the tissue in which circRNA was expressed the most. The analysis proceeded with the uterine tissue. Thus, it is demonstrated that the acquired data for this tissue could be interpreted as an example of the biogenesis mechanism.

The parent genes of the circRNAs found in the uterus are located other than the uterus,

and the RBPs that are relatively most expressed in the uterus bind to these target genes,parent of the uterus circRNA, regions.

We hypothesize that; all transcripts in the uterus tissue are directed to a back-splicing that differs from pol II-mediated splicing by the interaction of RBPs and various RNAs or proteins, so that parent gene expression of the circRNAs is absent or lowly expressed in the uterus.

# REFERENCES


[1] H. Shen *et al.*, "Circular RNAs: Characteristics, biogenesis, mechanisms and functions in liver cancer," *Journal of Hematology & Oncology*, vol. 14, no. 1, Aug. 2021. DOI: 10.1186/s13045-021-01145-8.

[2] C. Braicu, A.-A. Zimta, D. Gulei, A. Olariu, I. Berindan-Neagoe, "Comprehensive analysis of circular RNAs in pathological states: Biogenesis, cellular regulation, and therapeutic relevance," *Cellular and Molecular Life Sciences*, vol. 76, no. 8, pp. 1559–1577, Feb. 2019. DOI: 10.1007/s00018-019-03016-5.

[3] E. C. S. Lee *et al.*, "The roles of circular RNAs in human development and diseases," *Biomedicine & Pharmacotherapy*, vol. 111, pp. 198–208, Mar. 2019. DOI: 10.1016/j.biopha.2018.12.052.

[4] L. Statello, C.-J. Guo, L.-L. Chen, M. Huarte, "Gene regulation by long non-coding RNAs and its biological functions," *Nature Reviews Molecular Cell Biology*, vol. 22, no. 2, pp. 96–118, Dec. 2020. DOI: 10.1038/s41580-020-00315-9.

[5] H. Xu, S. Guo, W. Li, P. Yu, "The circular RNA cdr1as, via miR-7 and its targets, regulates insulin transcription and secretion in islet cells," *Scientific Reports*, vol. 5, no. 1, Jul. 2015. DOI: 10.1038/srep12453.

[6] W. J. Lukiw, "Circular RNA (circRNA) in Alzheimer's disease (AD)," *Frontiers in Genetics*, vol. 4, 2013. DOI: 10.3389/fgene.2013.00307.

[7] W. Hu *et al.*, "Emerging landscape of circular RNAs in lung cancer," *Cancer Letters*, vol. 427, pp. 18–27, Jul. 2018. DOI: 10.1016/j.canlet.2018.04.006.

[8] W. Chen, E. Schuman, "Circular RNAs in brain and other tissues: A functional enigma," *Trends in Neurosciences*, vol. 39, no. 9, pp. 597–604, Sep. 2016. DOI: 10.1016/j.tins.2016.06.006.

[9] Z. Li *et al.*, "Exon-intron circular RNAs regulate transcription in the nucleus," *Nature Structural & Molecular Biology*, vol. 22, no. 3, pp. 256–264, Feb. 2015. DOI: 10.1038/nsmb.2959.

[10] Y. Wang, L. Yin, X. Sun, "CircRNA hsa_circ_0002577 accelerates endometrial cancer progression through activating IGF1r/PI3k/akt pathway," *Journal of Experimental & Clinical Cancer Research*, vol. 39, no. 1, Aug. 2020. DOI: 10.1186/s13046-020-01679-8.

[11] C. Cocquerelle, B. Mascrez, D. Hétuin, B. Bailleul, "Mis-splicing yields circular RNA molecules.," *The FASEB Journal*, vol. 7, no. 1, pp. 155–160, Jan. 1993. DOI: 10.1096/fasebj.7.1.7678559.

[12] C. W. Chao, D. C. Chan, A. Kuo, P. Leder, "The mouse formin (fmn) gene: Abundant circular rna transcripts and gene-targeted deletion analysis," *Molecular medicine*, vol. 4, no. 9, pp. 614–628, 1998.

[13] M. Tang, L. Kui, G. Lu, W. Chen, "Disease-associated circular RNAs: From biology to computational identification," *BioMed Research International*, vol. 2020, pp. 1–21, Aug. 2020. DOI: 10.1155/2020/6798590.

[14] W. R. Jeck *et al.*, "Circular RNAs are abundant, conserved, and associated with ALU repeats," *RNA*, vol. 19, no. 2, pp. 141–157, Dec. 2012. DOI: 10.1261/rna.035667.112.

[15] C. Ragan, G. J. Goodall, N. E. Shirokikh, T. Preiss, "Insights into the biogenesis and potential functions of exonic circular RNA," *Scientific Reports*, vol. 9, no. 1, Feb. 2019. DOI: 10.1038/s41598-018-37037-0.

[16] S. Haque, L. Harries, "Circular RNAs (circRNAs) in health and disease," *Genes*, vol. 8, no. 12, p. 353, Nov. 2017. DOI: 10.3390/genes8120353.

[17] I. Chen, C.-Y. Chen, T.-J. Chuang, "Biogenesis, identification, and function of exonic circular RNAs," *Wiley Interdisciplinary Reviews: RNA*, vol. 6, no. 5, pp. 563–579, Jul. 2015. DOI: 10.1002/wrna.1294.

[18] L.-L. Chen, "The expanding regulatory mechanisms and cellular functions of circular RNAs," *Nature Reviews Molecular Cell Biology*, vol. 21, no. 8, pp. 475–490, May 2020. DOI: 10.1038/s41580-020-0243-y.

[19] W. R. Jeck, N. E. Sharpless, "Detecting and characterizing circular RNAs," *Nature Biotechnology*, vol. 32, no. 5, pp. 453–461, May 2014. DOI: 10.1038/nbt.2890.

[20] X. Geng *et al.*, "Circular RNA: Biogenesis, degradation, functions and potential roles in mediating resistance to anticarcinogens," *Epigenomics*, vol. 12, no. 3, pp. 267–283, Feb. 2020. DOI: 10.2217/epi-2019-0295.

[21] M.-P. Jiang, W.-X. Xu, J.-C. Hou, Q. Xu, D.-D. Wang, J.-H. Tang, "The emerging role of the interactions between circular RNAs and RNA-binding proteins in common human cancers," *Journal of Cancer*, vol. 12, no. 17, pp. 5206–5219, 2021. DOI: 10.7150/jca.58182.

[22] J. Wang, M. Zhu, J. Pan, C. Chen, S. Xia, Y. Song, "Circular RNAs: A rising star in respiratory diseases," *Respiratory Research*, vol. 20, no. 1, Jan. 2019. DOI: 10.1186/s12931-018-0962-1.

[23] D. Han *et al.*, "Circular RNA circMTO1 acts as the sponge of microRNA-9 to suppress hepatocellular carcinoma progression," *Hepatology*, vol. 66, no. 4, pp. 1151–1164, Aug. 2017. DOI: 10.1002/hep.29270.

[24] T. Qin, J. Li, K.-Q. Zhang, "Structure, regulation, and function of linear and circular long non-coding RNAs," *Frontiers in Genetics*, vol. 11, Mar. 2020. DOI: 10.3389/fgene.2020.00150.

[25] E. Lasda, R. Parker, "Circular RNAs: Diversity of form and function," *RNA*, vol. 20, no. 12, pp. 1829–1842, Nov. 2014. DOI: 10.1261/rna.047126.114.

[26] D. Liang, J. E. Wilusz, "Short intronic repeat sequences facilitate circular RNA production," *Genes & Development*, vol. 28, no. 20, pp. 2233–2247, Oct. 2014. DOI: 10.1101/gad.251926.114.

[27] Y. Bei *et al.*, "Circular RNAs as potential theranostics in the cardiovascular system," *Molecular Therapy - Nucleic Acids*, vol. 13, pp. 407–418, Dec. 2018. DOI: 10.1016/j.omtn.2018.09.022.

[28] Y. Zhang, Y. Wang, "Circular RNAs in hepatocellular carcinoma: Emerging functions to clinical significances," *Frontiers in Oncology*, vol. 11, May 2021. DOI: 10.3389/fonc.2021.667428.

[29] X. Zhao, Y. Cai, J. Xu, "Circular RNAs: Biogenesis, mechanism, and function in human cancers," *International Journal of Molecular Sciences*, vol. 20, no. 16, p. 3926, Aug. 2019. DOI: 10.3390/ijms20163926.

[30] B. Schwanhäusser *et al.*, "Global quantification of mammalian gene expression control," *Nature*, vol. 473, no. 7347, pp. 337–342, May 2011. DOI: 10.1038/nature10098.

[31] T. B. Hansen *et al.*, "miRNA-dependent gene silencing involving ago2-mediated cleavage of a circular antisense RNA," *The EMBO Journal*, vol. 30, no. 21, pp. 4414–4422, Sep. 2011. DOI: 10.1038/emboj.2011.359.

[32] R. A. Dubin, M. A. Kazmi, H. Ostrer, "Inverted repeats are necessary for circularization of the mouse testis sry transcript," *Gene*, vol. 167, no. 1-2, pp. 245–248, Dec. 1995. DOI: 10.1016/0378-1119(95)00639-7.

[33] S. Memczak *et al.*, "Circular RNAs are a large class of animal RNAs with regulatory potency," *Nature*, vol. 495, no. 7441, pp. 333–338, Feb. 2013. DOI: 10.1038/nature11928.

[34] Y. Zhang *et al.*, "Circular intronic long noncoding RNAs," *Molecular Cell*, vol. 51, no. 6, pp. 792–806, Sep. 2013. DOI: 10.1016/j.molcel.2013.08.017.

[35] T. B. Hansen *et al.*, "Natural RNA circles function as efficient microRNA sponges," *Nature*, vol. 495, no. 7441, pp. 384–388, Feb. 2013. DOI: 10.1038/nature11993.

[36] A. Nedoluzhko *et al.*, "The biomarker and therapeutic potential of circular rnas in schizophrenia," *Cells*, vol. 9, no. 10, p. 2238, Oct. 2020. DOI: 10.3390/cells9102238.

[37] M. Su *et al.*, "Circular RNAs in cancer: Emerging functions in hallmarks, stemness, resistance and roles as potential biomarkers," *Molecular Cancer*, vol. 18, no. 1, Apr. 2019. DOI: 10.1186/s12943-019-1002-6.

[38] S.-J. Zhang *et al.*, "Identification and characterization of circular RNAs as a new class of putative biomarkers in diabetes retinopathy," *Investigative Opthalmology & Visual Science*, vol. 58, no. 14, p. 6500, Dec. 2017. DOI: 10.1167/iovs.17-22698.

[39] X.-O. Zhang, H.-B. Wang, Y. Zhang, X. Lu, L.-L. Chen, L. Yang, "Complementary sequence-mediated exon circularization," *Cell*, vol. 159, no. 1, pp. 134–147, Sep. 2014. DOI: 10.1016/j.cell.2014.09.001.

[40] S. Werfel, S. Nothjunge, T. Schwarzmayr, T.-M. Strom, T. Meitinger, S. Engelhardt, "Characterization of circular RNAs in human, mouse and rat hearts," *Journal of Molecular and Cellular Cardiology*, vol. 98, pp. 103–107, Sep. 2016. DOI: 10.1016/j.yjmcc.2016.07.007.

[41] T. Romeo, "Global regulation by the small RNA-binding protein CsrA and the non-coding RNA molecule CsrB," *Molecular Microbiology*, vol. 29, no. 6, pp. 1321–1330, Sep. 1998. DOI: 10.1046/j.1365-2958.1998.01021.x.

[42] J. Salzman, C. Gawad, P. L. Wang, N. Lacayo, P. O. Brown, "Circular RNAs are the predominant transcript isoform from hundreds of human genes in diverse cell types," *PLoS ONE*, vol. 7, no. 2, T. Preiss, Ed., e30733, Feb. 2012. DOI: 10.1371/journal.pone.0030733.

[43] N. R. Pamudurti *et al.*, "Translation of CircRNAs," *Molecular Cell*, vol. 66, no. 1, 9–21.e7, Apr. 2017. DOI: 10.1016/j.molcel.2017.02.021.

[44] X. Meng, X. Li, P. Zhang, J. Wang, Y. Zhou, M. Chen, "Circular RNA: An emerging key player in RNA world," *Briefings in Bioinformatics*, bbw045, Jun. 2016. DOI: 10.1093/bib/bbw045.

[45] I. Legnini *et al.*, "Circ-ZNF609 is a circular RNA that can be translated and functions in myogenesis," *Molecular Cell*, vol. 66, no. 1, 22–37.e9, Apr. 2017. DOI: 10.1016/j.molcel.2017.02.017.

[46] J. Salzman, R. E. Chen, M. N. Olsen, P. L. Wang, P. O. Brown, "Cell-type specific features of circular RNA expression," *PLoS Genetics*, vol. 9, no. 9, J. V. Moran, Ed., e1003777, Sep. 2013. DOI: 10.1371/journal.pgen.1003777.

[47] M.-T. HSU, M. COCA-PRADOS, "Electron microscopic evidence for the circular form of RNA in the cytoplasm of eukaryotic cells," *Nature*, vol. 280, no. 5720, pp. 339–340, Jul. 1979. DOI: 10.1038/280339a0.

[48] A. Arnberg, G.-J. V. Ommen, L. Grivell, E. V. Bruggen, P. Borst, "Some yeast mitochondrial RNAs are circular," *Cell*, vol. 19, no. 2, pp. 313–319, Feb. 1980. DOI: 10.1016/0092-8674(80)90505-x.

[49] H. Tabak, G. V. der Horst, J. Smit, A. Winter, Y. Mul, G. M. Koerkamp, "Discrimination between RNA circles, interlocked RNA circles and lariats using two-dimensional polyacrylamide gel electrophoresis," *Nucleic Acids Research*, vol. 16, no. 14, pp. 6597–6605, 1988. DOI: 10.1093/nar/16.14.6597.

[50] C. W. Schindler, J. J. Krolewski, M. G. Rush, "Selective trapping of circular double-stranded DNA molecules in solidifying agarose," *Plasmid*, vol. 7, no. 3, pp. 263–270, May 1982. DOI: 10.1016/0147-619x(82)90007-5.

[51] R. Stark, M. Grzelak, J. Hadfield, "RNA sequencing: The teenage years," *Nature Reviews Genetics*, vol. 20, no. 11, pp. 631–656, Jul. 2019. DOI: 10.1038/s41576-019-0150-2.

[52] R. Ashwal-Fluss *et al.*, "circRNA biogenesis competes with pre-mRNA splicing," *Molecular Cell*, vol. 56, no. 1, pp. 55–66, Oct. 2014. DOI: 10.1016/j.molcel.2014.08.019.

[53] J. E. Wilusz, "Circular RNAs: Unexpected outputs of many protein-coding genes," *RNA Biology*, vol. 14, no. 8, pp. 1007–1017, Sep. 2016. [Online]. Available: https://doi.org/10.1080/15476286.2016.1227905.

[54] E. R. Mardis, "Next-generation DNA sequencing methods," *Annual Review of Genomics and Human Genetics*, vol. 9, no. 1, pp. 387–402, Sep. 2008. DOI: 10.1146/annurev.genom.9.081307.164359.

[55] L. Verduci, S. Strano, Y. Yarden, G. Blandino, "The circRNA–microRNA code: Emerging implications for cancer diagnosis and treatment," *Molecular Oncology*, vol. 13, no. 4, pp. 669–680, Feb. 2019. DOI: 10.1002/1878-0261.12468.

[56] Y. Li *et al.*, "Circular RNA expression profile of alzheimer's disease and its clinical significance as biomarkers for the disease risk and progression," *The International Journal of Biochemistry & Cell Biology*, vol. 123, p. 105 747, Jun. 2020. DOI: 10.1016/j.biocel.2020.105747.

[57] T. Jakobi, C. Dieterich, "Computational approaches for circular RNA analysis," *WIREs RNA*, vol. 10, no. 3, Feb. 2019. DOI: 10.1002/wrna.1528.

[58] M. Sand *et al.*, "Circular RNA expression in basal cell carcinoma," *Epigenomics*, vol. 8, no. 5, pp. 619–632, May 2016. DOI: 10.2217/epi-2015-0019.

[59] Z. Zhong, M. Lv, J. Chen, "Screening differential circular RNA expression profiles reveals the regulatory role of circTCF25-miR-103a-3p/miR-107-CDK6 pathway in bladder carcinoma," *Scientific Reports*, vol. 6, no. 1, Aug. 2016. DOI: 10.1038/srep30919.

[60] T. Jakobi, C. Dieterich, "Deep computational circular RNA analytics from RNA-seq data," in *Methods in Molecular Biology*, Springer New York, 2018, pp. 9–25. DOI: 10.1007/978-1-4939-7562-4_2.

[61] T. B. Hansen, "Improved circRNA identification by combining prediction algorithms," *Frontiers in Cell and Developmental Biology*, vol. 6, Mar. 2018. DOI: 10.3389/fcell.2018.00020.

[62] P. Glažar, P. Papavasileiou, N. Rajewsky, "circBase: A database for circular RNAs," *RNA*, vol. 20, no. 11, pp. 1666–1670, Sep. 2014. DOI: 10.1261/rna.043687.113.

[63] X. Chen, P. Han, T. Zhou, X. Guo, X. Song, Y. Li, "circRNADb: A comprehensive database for human circular RNAs with protein-coding annotations," *Scientific Reports*, vol. 6, no. 1, Oct. 2016. DOI: 10.1038/srep34985.

[64] S. Xia *et al.*, "Comprehensive characterization of tissue-specific circular RNAs in the human and mouse genomes," *Briefings in Bioinformatics*, bbw081, Aug. 2016. DOI: 10.1093/bib/bbw081.

[65] M. Liu, Q. Wang, J. Shen, B. B. Yang, X. Ding, "Circbank: A comprehensive database for circRNA with standard nomenclature," *RNA Biology*, vol. 16, no. 7, pp. 899–905, Apr. 2019. DOI: 10.1080/15476286.2019.1600395.

[66] X. Meng, D. Hu, P. Zhang, Q. Chen, M. Chen, "CircFunBase: A database for functional circular RNAs," *Database*, vol. 2019, Jan. 2019. DOI: 10.1093/database/baz003.

[67] W. Wu, P. Ji, F. Zhao, "CircAtlas: An integrated resource of one million highly accurate circular RNAs from 1070 vertebrate transcriptomes," *Genome Biology*, vol. 21, no. 1, Apr. 2020. DOI: 10.1186/s13059-020-02018-y.

[68] T. Jakobi, A. Uvarovskii, C. Dieterich, "Circtools—a one-stop software solution for circular RNA research," *Bioinformatics*, vol. 35, no. 13, B. Berger, Ed., pp. 2326–2328, Nov. 2018. DOI: 10.1093/bioinformatics/bty948.

[69] R. Li *et al.*, "Potential regulatory role of circular RNA in idiopathic pulmonary fibrosis," *International Journal of Molecular Medicine*, Sep. 2018. DOI: 10.3892/ijmm.2018.3892.

[70] S. H. Aghaee-Bakhtiari, "Online databases and circular RNAs," pp. 35–38, 2018. DOI: 10.1007/978-981-13-1426-1_3.

[71] F. Li *et al.*, "Circular RNA ITCH has inhibitory effect on ESCC by suppressing the wnt/$\beta$-catenin pathway," *Oncotarget*, vol. 6, no. 8, pp. 6001–6013, Feb. 2015. DOI: 10.18632/oncotarget.3469.

[72] G. Yao *et al.*, "*hsa_circRNA_104597*: A novel potential diagnostic and therapeutic biomarker for schizophrenia," *Biomarkers in Medicine*, vol. 13, no. 5, pp. 331–340, Apr. 2019. DOI: 10.2217/bmm-2018-0447.

[73] S. Chen *et al.*, "Widespread and functional RNA circularization in localized prostate cancer," *Cell*, vol. 176, no. 4, 831–843.e22, Feb. 2019. DOI: 10.1016/j.cell.2019.01.025.

[74] J. Chen, J. Yang, X. Fei, X. Wang, K. Wang, "CircRNA ciRS-7: A novel oncogene in multiple cancers," *International Journal of Biological Sciences*, vol. 17, no. 1, pp. 379–389, 2021. DOI: 10.7150/ijbs.54292.

[75] Y.-J. Zhu *et al.*, "Circular RNAs negatively regulate cancer stem cells by physically binding FMRP against CCAR1 complex in hepatocellular carcinoma," *Theranostics*, vol. 9, no. 12, pp. 3526–3540, 2019. DOI: 10.7150/thno.32796.

[76] W. Liu, Y. Xiong, R. Wan, R. Shan, J. Li, W. Wen, "The roles of circMTO1 in cancer," *Frontiers in Cell and Developmental Biology*, vol. 9, Jun. 2021. DOI: 10.3389/fcell.2021.656258.

[77] X.-W. Xu *et al.*, "Circular RNA hsa_circ_000984 promotes colon cancer growth and metastasis by sponging miR-106b," *Oncotarget*, vol. 8, no. 53, pp. 91 674–91 683, Oct. 2017. DOI: 10.18632/oncotarget.21748.

[78] H. Yang *et al.*, "Circular RNA circ_0000950 promotes neuron apoptosis, suppresses neurite outgrowth and elevates inflammatory cytokines levels via directly sponging miR-103 in alzheimer's disease," *Cell Cycle*, vol. 18, no. 18, pp. 2197–2214, Aug. 2019. DOI: 10.1080/15384101.2019.1629773.

[79] S. P. Barrett, P. L. Wang, J. Salzman, "Circular RNA biogenesis can proceed through an exon-containing lariat precursor," *eLife*, vol. 4, Jun. 2015. DOI: 10.7554/elife.07540.

[80] C. A. Schmidt, J. D. Giusto, A. Bao, A. K. Hopper, A. G. Matera, "Molecular determinants of metazoan tricRNA biogenesis," *Nucleic Acids Research*, vol. 47, no. 12, pp. 6452–6465, Apr. 2019. DOI: 10.1093/nar/gkz311.

[81] R.-C. Zou, L.-L. Li, H.-L. Yuan, "Current status of research on the role of circular RNAs in hepatocellular carcinoma and clinical implications," *Medical Science Monitor*, vol. 26, Jun. 2020. DOI: 10.12659/msm.923832. [Online]. Available: https://doi.org/10.12659/msm.923832.

[82] Y. Huang, Q. Zhu, "Mechanisms regulating abnormal circular RNA biogenesis in cancer," *Cancers*, vol. 13, no. 16, p. 4185, Aug. 2021. DOI: 10.3390/cancers13164185.

[83] A. G. Matera, Z. Wang, "A day in the life of the spliceosome," *Nature Reviews Molecular Cell Biology*, vol. 15, no. 2, pp. 108–121, Jan. 2014. DOI: 10.1038/nrm3742.

[84] A. Das, T. Sinha, S. Shyamal, A. C. Panda, "Emerging role of circular RNA–protein interactions," *Non-Coding RNA*, vol. 7, no. 3, p. 48, Aug. 2021. DOI: 10.3390/ncrna7030048.

[85] T. L. H. Okholm *et al.*, "Transcriptome-wide profiles of circular RNA and RNA-binding protein interactions reveal effects on circular RNA biogenesis and cancer pathway expression," *Genome Medicine*, vol. 12, no. 1, Dec. 2020. DOI: 10.1186/s13073-020-00812-8.

[86] W. W. Du, C. Zhang, W. Yang, T. Yong, F. M. Awan, B. B. Yang, "Identifying and characterizing circRNA-protein interaction," *Theranostics*, vol. 7, no. 17, pp. 4183–4191, 2017. DOI: 10.7150/thno.21299.

[87] E. L. V. Nostrand *et al.*, "A large-scale binding and functional map of human RNA-binding proteins," *Nature*, vol. 583, no. 7818, pp. 711–719, Jul. 2020. DOI: 10.1038/s41586-020-2077-3.

[88] P. A. F. Galante *et al.*, "A comprehensive in silico expression analysis of RNA binding proteins in normal and tumor tissue; identification of potential players in tumor formation," *RNA Biology*, vol. 6, no. 4, pp. 426–433, Sep. 2009. DOI: 10.4161/rna.6.4.8841.

[89] Y.-C. T. Yang *et al.*, "CLIPdb: A CLIP-seq database for protein-RNA interactions," *BMC Genomics*, vol. 16, no. 1, Feb. 2015. DOI: 10.1186/s12864-015-1273-2.

[90] S. M. Garcia-Mauriño *et al.*, "RNA binding protein regulation and cross-talk in the control of AU-rich mRNA fate," *Frontiers in Molecular Biosciences*, vol. 4, Oct. 2017. DOI: 10.3389/fmolb.2017.00071.

[91] N. Mittal, N. Roy, M. M. Babu, S. C. Janga, "Dissecting the expression dynamics of RNA-binding proteins in posttranscriptional regulatory networks," *Proceedings of the National Academy of Sciences*, vol. 106, no. 48, pp. 20 300–20 305, Dec. 2009. DOI: 10.1073/pnas.0906940106.

[92] B. Kechavarzi, S. Janga, "Dissecting the expression landscape of RNA-binding proteins in human cancers," *Genome Biology*, vol. 15, no. 1, R14, 2014. DOI: 10.1186/gb-2014-15-1-r14.

[93] A. Ramakrishnan, S. C. Janga, "Human protein-RNA interaction network is highly stable across mammals," *BMC Genomics*, vol. 20, no. S12, Dec. 2019. DOI: 10.1186/s12864-019-6330-9.

[94] S. J. Conn *et al.*, "The RNA binding protein quaking regulates formation of circRNAs," *Cell*, vol. 160, no. 6, pp. 1125–1134, Mar. 2015. DOI: 10.1016/j.cell.2015.02.014.

[95] X. Lai *et al.*, "Epithelial-mesenchymal transition and metabolic switching in cancer: Lessons from somatic cell reprogramming," *Frontiers in Cell and Developmental Biology*, vol. 8, Aug. 2020. DOI: 10.3389/fcell.2020.00760.

[96] P. L. Boutz, A. Bhutkar, P. A. Sharp, "Detained introns are a novel, widespread class of post-transcriptionally spliced introns," *Genes & Development*, vol. 29, no. 1, pp. 63–80, Jan. 2015. DOI: 10.1101/gad.247361.114.

[97] K. Czubak, S. Sedehizadeh, P. Kozlowski, M. Wojciechowska, "An overview of circular RNAs and their implications in myotonic dystrophy," *International Journal of Molecular Sciences*, vol. 20, no. 18, p. 4385, Sep. 2019. DOI: 10.3390/ijms20184385.

[98] X. Li, G. G. Tian, Y. Zhao, J. Wu, "Genome-wide identification and characterization of long noncoding and circular RNAs in germline stem cells," *Scientific Data*, vol. 6, no. 1, Mar. 2019. DOI: 10.1038/s41597-019-0014-9.

[99] Y. Gao, J. Zhang, F. Zhao, "Circular RNA identification based on multiple seed matching," *Briefings in Bioinformatics*, vol. 19, no. 5, pp. 803–810, Feb. 2017. DOI: 10.1093/bib/bbx014.

[100] Y. Gao, J. Wang, F. Zhao, "CIRI: An efficient and unbiased algorithm for de novo circular RNA identification," *Genome Biology*, vol. 16, no. 1, Jan. 2015. DOI: 10.1186/s13059-014-0571-3.

[101] R. Dong, X.-K. Ma, L.-L. Chen, L. Yang, "Genome-wide annotation of circRNAs and their alternative back-splicing/splicing with CIRCexplorer pipeline," in *Epitranscriptomics*, Springer New York, Dec. 2018, pp. 137–149. DOI: 10.1007/978-1-4939-8808-2_10.

[102] J. Cheng, F. Metge, C. Dieterich, "Specific identification and quantification of circular RNAs from sequencing data," *Bioinformatics*, vol. 32, no. 7, pp. 1094–1096, Nov. 2015. DOI: 10.1093/bioinformatics/btv656.

[103] P. Zhang, W. Wu, Q. Chen, M. Chen, "Non-coding RNAs and their integrated networks," *Journal of Integrative Bioinformatics*, vol. 16, no. 3, Jul. 2019. DOI: 10.1515/jib-2019-0027.

[104] E. L. V. Nostrand *et al.*, "Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP)," *Nature Methods*, vol. 13, no. 6, pp. 508–514, Mar. 2016. DOI: 10.1038/nmeth.3810.

[105] S. Misir, N. Wu, B. B. Yang, "Specific expression and functions of circular RNAs," *Cell Death & Differentiation*, vol. 29, no. 3, pp. 481–491, Feb. 2022. DOI: 10.1038/s41418-022-00948-7.

[106] M. Müller-McNicoll, K. M. Neugebauer, "How cells get the message: Dynamic assembly and function of mRNA–protein complexes," *Nature Reviews Genetics*, vol. 14, no. 4, pp. 275–287, Mar. 2013. DOI: 10.1038/nrg3434.

[107] K. O. West, H. M. Scott, S. Torres-Odio, A. P. West, K. L. Patrick, R. O. Watson, "The splicing factor hnRNP m is a critical regulator of innate immune gene expression in macrophages," *Cell Reports*, vol. 29, no. 6, 1594–1609.e5, Nov. 2019. DOI: 10.1016/j.celrep.2019.09.078.

[108] T. Geuens, D. Bouhy, V. Timmerman, "The hnRNP family: Insights into their role in health and disease," *Human Genetics*, vol. 135, no. 8, pp. 851–867, May 2016. DOI: 10.1007/s00439-016-1683-5.

[109] F. Caiazza *et al.*, "KH-type splicing regulatory protein controls colorectal cancer cell growth and modulates the tumor microenvironment," *The American Journal of Pathology*, vol. 189, no. 10, pp. 1916–1932, Oct. 2019. DOI: 10.1016/j.ajpath.2019.07.004.

[110] J. A. Bohn *et al.*, "Identification of diverse target RNAs that are functionally regulated by human pumilio proteins," *Nucleic Acids Research*, vol. 46, no. 1, pp. 362–386, Nov. 2017. DOI: 10.1093/nar/gkx1120.

[111] M. SENOO *et al.*, "RNA-binding protein ptbp1 regulates alternative splicing and transcriptome in spermatogonia and maintains spermatogenesis in concert with nanos3," *Journal of Reproduction and Development*, vol. 66, no. 5, pp. 459–467, 2020. DOI: 10.1262/jrd.2020-060.

# PUBLICATIONS FROM THE THESIS

## Conference Papers

1. E.İ. Keleş, A. Yılmaz, "Expression Profile Survey of Circular RNAs and Their Parent Genes in Context of Tissue Specificity" The 14th International Symposium on Health Informatics and Bioinformatics, 2021, p. 83.