Online Databases and Non-coding RNAs in Cardiovascular Diseases

3

Rui Chen, Chao Shi, Jianhua Yao, and Wei Chen

Abstract

Cardiovascular disease is characterized by its highest morbidity and mortality. One of the main pathological basis of this disease is the dysregulation of gene expression. Non-coding RNA (ncRNA) is a kind of functional RNA, which is transcript from DNA but not translated into proteins. More and more studies have established the important roles of ncRNAs, including transcription, RNA maturation, translation, protein degradation, and their involvement in the pathogenesis of diseases such as cancer and cardiovascular diseases. This chapter will focus on the biological functions of ncRNAs and their advances in cardiovascular disease. With the development of sequencing and computer technology, more and more databases can be easily obtained on the internet. In another part of this chapter, we will summarize some commonly used non-coding RNA databases, which can be easily and quickly used for relevant research.

R. Chen
Shanghai Applied Radiation Institute, School of Environmental and Chemical Engineering, Shanghai University, Shanghai, China

C. Shi
Cardiac Regeneration and Ageing Lab, Institute of Cardiovascular Sciences, School of Life Science, Shanghai University, Shanghai, China

J. Yao (✉)
Department of Cardiology, Shanghai Tenth People's Hospital, Tongji University School of Medicine, Shanghai, China
e-mail: yaojianhua@tongji.edu.cn

W. Chen (✉)
Emergency Department, Shanghai Tongji Hospital, Tongji University School of Medicine, Shanghai, China

1 Introduction

Cardiovascular disease was one of the diseases with the highest morbidity and mortality in the world. However, the basic research on cardiovascular disease was still limited [1, 2]. In-depth basic research could provide a guiding direction for new methods in treating cardiovascular diseases [3]. Therefore, exploring the molecular and cellular mechanisms in the pathogenesis of cardiovascular disease could help us understand the disease process, discover new biomarkers, and find new therapeutic targets [4, 5]. Over the past decade, the development of high-throughput sequencing technologies given us an opportunity to expand our understanding of human transcriptomes [6, 7]. Previous studies have suggested that proteins are the main regulators of cardiovascular

disease [8, 9]. However, with the understanding of the complexity human transcriptomes, it was revealed that most human genomes can transcribe RNA, of which more than 98% were untranslated proteins [10]. The number of non-coding RNAs was far exceeds the amount of mRNA encoding proteins [11, 12].

Although the expression level of ncRNAs were generally low and did not encode proteins, they could show strong tissue specificity and express abnormalities in various human diseases and biological processes such as organ development, internal and external environmental stimulation and disease occurrence [13, 14, 15]. So far, more and more studies have shown the important regulatory functions of ncRNAs in participate in biosynthetic steps such as transcription [16], RNA maturation [17], translation [18], protein degradation [19], and found ncRNAs play an important role in pathological processes such as tumors and cardiovascular diseases [20]. Based on their size, these non-coding RNAs were classified into microRNAs (miRNAs, <200 nucleotides), and long non-coding RNAs (lncRNAs, >200 nucleotides), in which lncRNAs can also exhibit a circular shape called circular RNAs (circRNAs). In this section, we will illuminate the important role of non-coding RNAs in the development of cardiovascular diseases through their special biological function. The main classification of non-coding RNA and its formation in the cell is shown in Fig. 3.1.

2   miRNAs in Cardiovascular Diseases

A miRNA was an endogenous non-coding RNA, typically 18–25 nucleotides, which regulated the expression of a target gene by binding to the 3′UTR of the mRNA and inhibiting its translation. The precursor of the miRNA first existed in the nucleus. Then after the transcription and subsequently matured by several enzymatic reactions, mRNAs transferred to the cytoplasm to exert their biological functions. miRNAs exerted their regulatory functions by recruiting specific silencing proteins to form an RNA-induced silencing complex (RISC). It was speculated that about 60% of mRNA will be target by miRNA, while a miRNA may target more than 100 mRNAs in human. Specific miRNAs have significantly different expression levels in cardiac tissue and cardiovascular, and played as regulators of cardiovascular function, including cardiovascular cell differentiation, growth, proliferation and apoptosis, angiogenesis and cell contractility regulation. Meanwhile, the pathological changes in the cardiovascular system always accompanied with some specific miRNAs changes. These miRNAs changes have been confirmed to be associated with cardiovascular diseases such as arrhythmia, cardiac hypertrophy, fibrosis, myocardial infarction and heart failure. Recent studies indicated that miRNAs with specific changes in the progression of heart valve disease play key role in the processes of disease progression, such as fibrosis, calcification, matrix degradation remodeling, and inflammation. In addition, some miRNAs could also regulate extracellular body through exosomes and participate in circulation in the exocrine [21–23]. Further studies revealed that these miRNAs can be used as biomarkers for cardiovascular diseases for clinical diagnosis and personalized medicine.

3   lncRNAs in Cardiovascular Diseases

lncRNAs were the heterogeneous RNA transcripts, which contain more than200 nucleotides in length. lncRNAs could be classified into sense, antisense, intron, genomic and divergent lncRNAs based on their relative genomic location. lncRNAs were involved in many biological processes, such as chromatin structural changes, transcription, post-transcriptional processing, intracellular trafficking, and regulation of enzyme activity.

lncRNA could also regulate other endogenous ncRNAs, particularly miRNAs, through competitive binding. Compared to miRNAs, lncRNAs were less conserved, suggesting that these RNA
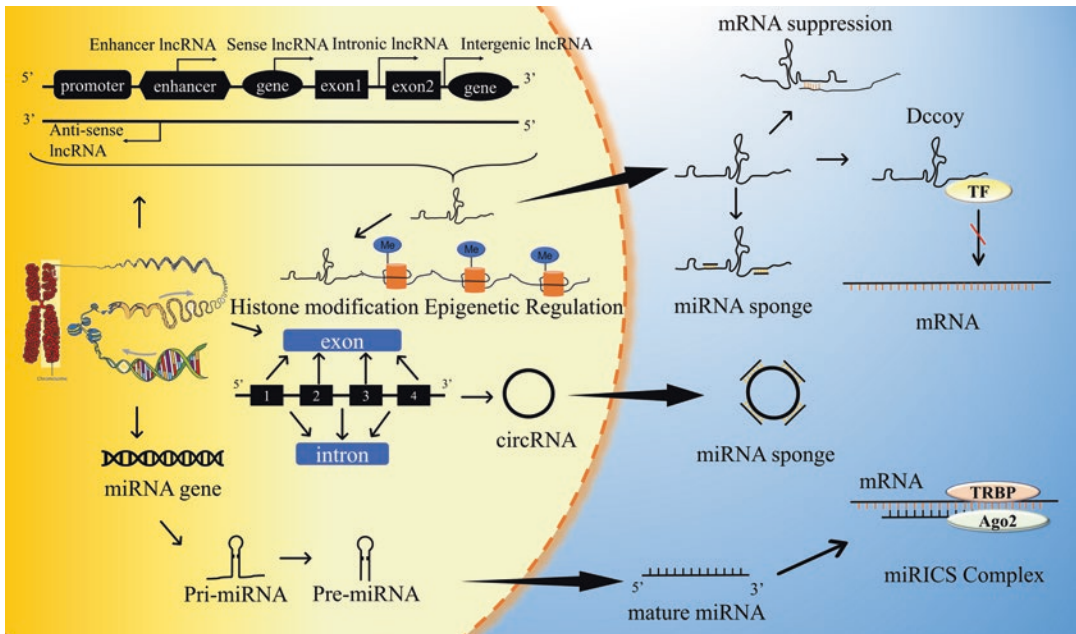
Fig. 3.1 The main classification of non-coding RNA and its formation in the cell

molecules have a species-specific effect. Although the dysregulation of lncRNAs was associated with various human diseases, their mechanism of action still remained unclear. To date, it has been reported that lncRNAs dysregulation was detected in many cardiovascular diseases such as myocardial infarction, myocardial fibrosis, cardiac hypertrophy and heart failure.

4 circRNAs in Cardiovascular Diseases

Circular RNAs (circRNAs) transcripts were first identified in the early 1990s, but knowledge of these species has remained limited, due to their difficult study through traditional methods of RNA analysis. circRNAs were a peculiar group of RNAs, consisting of at least a few hundred nucleotides and relatively stable in their circular state. circRNAs were involved in a wide range of biological processes, that expression disorder of circRNAs might lead to abnormal cellular functions and disease. However, the regulation of circRNAs in cardiovascular diseases remains largely

unexplored. As the develop of RNA-Seq studies, large number of different circRNAs were detected to be expressed in cardiac tissues from human and rodents.

As the roles of non-coding RNAs in cardiovascular diseases were gradually unearthed, a large amount of data was reported every year, including non-coding RNA sites, regulatory mechanisms, etc. Document summarization and data induction became a daunting task when summarizing the rules and guiding the direction of the next step. With the development of electronic computer hardware and software such as artificial intelligence and cloud computing, many non-coding RNAs databases have been established. These databases summarized the existing research data of miRNAs and lncRNAs with the combination of bioinformatics analysis. These gave us a way to generalize the functions reported by miRNAs and lncRNAs and predict other possible sites of their effects. The rational use of these databases could significantly improve the efficiency of our researchers. Below we will introduce some of the more mature online databases.

5   miRNA Database

5.1   deepBase v1.0

deepBase v1.0 [24] can be used to annotation and discovery miRNAs, lncRNAs and circRNAs sequencing, view the expression of various non-coding RNAs, and downloadable data. Searching for miRNA expressions in web pages can be retrieved by gene names and symbols. This function was very useful tool for us to select new genes. Although there were many choices of lncRNA that were specifically expressed in a certain tissue, we could analyze the conservation of a non-coding RNA in the evolution option through this database.

5.2   miRBase

miRBase [25] was an online miRNA database, which was developed by researchers at the University of Chester. The database contained more than 200 species and close to 40,000 miRNAs. It was the most comprehensive miRNA database in which we can do species specific particular miRNA search and browses, such as the numbering of the precursor miRNA, the relative expression, the starting position of the miRNA precursor on the chromosome, and gives confidence to judge. miRbase also assigned names to newly discovered miRNAs, which provided a convenient and fast communication platform for researchers.

5.3   microRNA.org

microRNA.org [26] was a comprehensive database of miRNA target prediction and expression. The target predictions provided by the website were implemented by the MiRanda algorithm, which combined current research reports on miRNAs and provided data from a large number of mammalian tissue synthesis sequencing projects. Through the latest algorithms, users could search for genes that may be regulated by a particular miRNA, or search for a variety of miRNAs that regulate the certain gene, as well as the expression profiles of various miRNAs. The database included human, mouse, rat and other species, by combining with 250 miRNA libraries, provided miRNA function prediction functions during the study of miRNA in cardiovascular-related diseases.

5.4   miRNAMap

miRNAMap [27] validated miRNA target genes by collecting experimentally validated miRNAs in human, mouse, rat and other mammalian genomes. Three algorithms include miRanda, RNAhybrid and TargetScan were used to validate miRNA targets in the 3′-UTR of the gene as well as known miRNA targets. By using filtration of multiple algorithms to speculate the target site of the miRNA in order to reduce the probability of predicting false positives in the miRNA target site prediction.

5.5   Cupid

Cupid [28] was an online database that validated the high-throughput validation analysis in breast cancer cell lines by simultaneously predicting miRNA-target interactions and their mediated endogenous RNA interactions. The database publisher verified the accuracy of the interaction of 500 miRNAs with the target to make sure the accuracy of this database.

5.6   TargetScan

TargetScan [29] was a website for predicting miRNA target genes, which contained human, orangutan, macaque, mouse, rat and other species. The website could display the prediction after inputting the gene name or the ENST label. This website also predicted target genes' possible role through miRNAs.

5.7 miRTarBase

miRTarBase [30] was a manually collected, experimentally validated miRNA target gene database, which provided a variety way to search, such as miRNA ID, target genes, KEGG pathways, validated methods, or diseases. The database was incubated with a unique miRNA-target interactions (MTS) number for each miRNA target gene. miRTarBase not only provided highly reliable miRNA target gene information, but also provided information on miRNA and diseases that were reported in the literature.

5.8 Diana-microT-CDS

Diana-microT-CDS [31] was an online database, which provided computer simulations of miRNA-mRNA interactions with a good user interface. This database provided sufficient information on predicting miRNA target genes, such as global scores of interactions and visualized results. microT-CDS was the only miRNA target prediction algorithm available online that was specifically designed to identify miRNA targets in the 3′UTR and CDS regions. The database contained miRNA target information in the mRNA sequences of human, mus musculus, drosophila melanogaster and C. elegans, which contained a wealth of tooltips and convenient menu options that were easy to use. In addition, the new system that combined advanced workflows also supported massive data analysis from Next Generation Sequencing (NGS).

5.9 miRecords

miRecords [32] was an online database that provided animal miRNA-target gene prediction. The database consisted of two parts, one was a proven large, high-quality database, which also includeed artificially screened miRNA targets. The other one was an online target gene prediction algorithm, which the algorithm provided by the database emphasized experimental data support of miRNA targets.

5.10 PicTar

PicTar [33] database used a certain algorithm to identify the target of a microRNA. The searchable website provided detailed predictions of microRNA targets from the following species, including: vertebrate, seven drosophila species, three nematode species, and human non-conserved but co-expressed microRNA targets (e.g.: Express microRNA and mRNA in the same tissue).

5.11 TarBase

TarBase [34] was a database of miRNA target genes prediction website that was supported by experimental data. After 10 years of data compilation, the database provided information on miRNA target genes in various species such as humans and mice. For each target gene data, such as the relevant literature, organization type, and test method were given. In TarBase, each experimental evidence was divided into two categories, low and high. Low represented the traditional experimental method. Compared with the high-throughput sequencing analysis, the reliability of the result was higher. We can filter the miRNA target supported through the low method in order to obtain a high-quality miRNA target gene data set.

5.12 miRWalk

miRWalk [35] was a comprehensive miRNA target gene database containing miRNA target gene information from human, mouse and other species. It was an integrated database that integrated information from miRDB, TargetScan, miRTarBase and other databases. The regulation network between the miRNAs and the target genes visualized the function of graph and gene set enrichment, supported the enrichment analysis and the analysis of reactive pathway, KEGG pathway, and gene ontology. The miRNA databases URLs and related information listed above are shown in Table 3.1.

Table 3.1 microRNA databases

Name	Weblink	Data content
deepBase	http://biocenter.sysu.edu.cn/deepBase/	A platform containing evolution and expression patterns of diverse ncRNAs across 19 species from 5 clades.
miRBase	http://www.mirbase.org/	A searchable database of published miRNA sequences and annotation.
microRNA	http://www.microrna.org/	Predict microRNA targets & target downregulation scores. Experimentally observed expression patterns.
miRNAMap	http://mirnamap.mbc.nctu.edu.tw/	Collect experimental verified microRNAs and their target genes in human, mouse, rat, and other metazoan genomes.
Cupid	http://cupidtool.sourceforge.net/	A method for simultaneous prediction of miRNA-target interactions and their mediated competitive endogenous RNA interactions.
TargetScan	http://www.targetscan.org/vert_72/	Predict biological targets of miRNAs by searching for the presence of conserved 8mer, 7mer, and 6mer sites that match the seed region of each miRNA.
miRTarBase	http://mirtarbase.mbc.nctu.edu.tw/php/index.php	Accumulated more than 360,000 miRNA-target interactions (MTIs).
Diana-microT-CDS	http://diana.imis.athena-innovation.gr/	Specifically trained on a positive and a negative set of miRNA Recognition Elements (MREs) located in both the 3′-UTR and CDS regions.
miRecords	http://c1.accurascience.com/miRecords/	Hosts 2705 records of interactions between 644 miRNAs and 1901 target genes in 9 animal species.
PicTar	https://pictar.mdc-berlin.de/	An algorithm for the identification of microRNA targets.
TarBase	http://carolina.imis.athena-innovation.gr	Containing more than 1,000,000 miRNA-gene interactions.
miRWalk	http://mirwalk.umm.uni-heidelberg.de/	Stores predicted data obtained with a machine learning algorithm including experimentally verified miRNA-target interactions.

6 lncRNA Databases

6.1 LNCipedia

LNCipedia [36–38] was a comprehensive human lncRNA database that integrated lncRNA records from multiple databases and articles and gave them a uniform ID or named new reported lncRNA. The database summarizes lncRNAdb, Broad Institute, Ensembl, Gencode, Refseq, NONCODE, FANTOM and other lncRNA databases.

6.2 lncRNAdb

lncRNAdb [39, 40] was available at the website http://www.lncrnadb.org/. In order to make it easy for the users to compile and update the information of lncRNAs, researchers established this database containing comprehensive lncRNAs and their biological functions. lncRNAdb provided users with plenty information of lncRNAs, including lncRNA sequences, struction, genomic context and so on. Users could search for the published lncRNA names and sequences species, associated protein-coding genes. Also, this database was linked to the UCSC Genome Browser and Noncoding RNA Expression Database, making various sources of lncRNAs. With 4 years development, there existed version 2.0 lncRNAdb. This new version contained 287 eukaryotic lncRNAs, and a more accessible interface for users to search the sequence information and expression data. There were some new features of the version 2.0, including the nucleotide sequence information and an easier way for users to export the information they searched. With deeper study of lncRNAs, there will be more lncRNAs adding to this database and make it better for the future research.

6.3   LncRNAWiki

LncRNAWiki [41] database integrateed the sequences and annotation information of human lncRNAs in GENCODE, NONCODE, and LNCipedia databases. It now has 105,255 non-redundant lncRNAs, of which 103 have full annotation and biological function tests. The database was also an editable public platform that not only allowed users to annotate, update, and organize existing lncRNAs, but also open a permission for all user to share their new detected lncRNAs. This database was an open platform for human lncRNA information.

6.4   MONOCLdb

MONOCLdb [42] fully called the Mouse Non-Code Lung interactive database, which was available at https://www.monocldb.org/. The lncRNAs from this database were based on the mice infected by influenza and SARS-CoV viruses.5329 of 20,728 mouse lncRNAs genes showed differently expressed after infected by these two viruses. In this database, developers annotated the difference of the expression by two computational methods, module-based and rank-based, so that users can retrieve the annotations, expression profiles and the functional enrichment of lncRNAs. Also, via MONOCLdb, users could generate the expression heatmaps and the functional enrichment results on-the-fly. This database provided users with association scores between lncRNAs and pathogenicity variables MONOCLdb was an integrative and interactive database containing abundant lncRNAs in order to make convenience for the researchers to learn more about their role during the viruses infection.

6.5   NONCODE

NONCODE [43, 44], located at the web server http://www.bioinfo.org/noncode/, was a database which showed complete collection of the lncRNAs. Based on the third-generation sequenc-

ing method, developers could get plenty correct annotations and establish this database. NONCODE contained 527,336 lncRNAs from 16 different species, including 167,150 human lncRNAs and 130,558 mouse lncRNAs. Compared with other database, NONCODE really had a large number of lncRNAs. Except the large number, NONCODE contained three unique features, including a conservation annotation of the lncRNAs, a relationship annotation between lncRNAs and diseases, and a platform for users to choose the high-quality datasets. This database decided to show the relationship between the lncRNAs and diseases to make a systematic network about the noncoding regions and the diseases. Also, it made more convenient for the disease study users to get their interested information. Although there still existed some uncleared things in the development of this database, the developers will attempt to solve the problem and make NONCODE a better database for the users.

6.6   lncRNome

lncRNome [45] was a comprehensive database that integrated lots of significant annotations for lncRNAs, and now became one of the largest database of lncRNAs in the world. It was available at the website http://genome.igib.res.in/lncRNome/. In lncRNome, each lncRNA has its own page including the information about the annotation sets. Through the links on this page, users could easily get other information of this lncRNA including the sequence, structure, interactions, variations and so on. There existed 937 quadruplex and 40 hairpins motifs, 3716 miRNA binding sites and more than 10,000 binding sites for proteins on the lncRNA in lncRNome. Also, lncRNome provided 345,351 genomic variations associated with lncRNAs and 11,790 epigenetic marks in the promoters of lncRNAs. Furthermore, this database has multiple options for users to search for the lncRNAs, such as thename or known targets. In addition, the representative of the available associated genomic annotations was also one of the features of this database.

6.7   C-It-Loci

C-It-Loci [46] was an database that allowed users to browse the transcripts specific in tissue from human, mouse and zebrafish. It was an available and free web server, which was established in 2015. User can used through the website http://c-it-loci.uni-frankfurt.de/. C-It-Loci made it easier for users to identify ncRNAs and protein-coding genes from various tissues. Not only normal lncRNAs, but also the housekeeping genes and its housekeep lncRNAs were contained in this database. To make convenience for the users with limited knowledge, C-It-Loci allowed users to see the different types of transcripts C-It-Loci had a quick search function on the page, which made it easy for users to compare the difference of two transcripts. C-It-Loci also offered a diagram for the users to analyze the conserved regions and explore the detailed information of the lncRNA they searched. Compared with other databases, C-It-Loci had the latest human genomic assembly, and defined lncRNAs with three organisms. Furthermore, C-It-Loci was the only database that has both protein-coding genes and lncRNAs transcripts. There would be a great prospect to study on lncRNAs through this database.

6.8   MiTranscriptome

MiTranscriptome [47] was a database about genome-wide lncRNA expression, constructed by Matthew K Iyer and his colleagues in 2015. The web server of this database is http://www.mitranscriptome.org/. In order to build this database, they curated 7256 RNA-seq from tumors, normal tissues and cell lines, which was comprising from 25 independent studies. Finally, they got a transcriptome of 91,013 expressed gene, 68% was lncRNAs. Interestingly, 79% of them were unannotated in previous study. Then they used non-parametric differential expression test to filtrate and finally got 7942 lncRNAs that were associated with disease or cancer. These lncRNA could be valuable for the further study of cancer and the development of biomarker. Therefore,

MiTranscriptome provided a foundation for lncRNA genomics, biomarker development and the delineation of cancer.

6.9   slncky Evolution Browser

slncky Evolution Browser [48] was developed in 2016. The website of this database was https://scripts.mit.edu/~jjenny/. slncky was a tool for searching conserved lncRNAs through a sensitive noncoding aligning method. This database could be efficiently used to produce high quality lncRNAs from RNA-sequencing data. It could separateed lncRNAs accurately from coding genes, pseudogenes, assembly artifacts, and also identify novel proteins including small peptides. On the basis of this tool, they develop the database called slncky Evolution Browser. Through this database, they list 233 constrained lncRNAs out of the currently annotated transcripts. With the powerful tool like slncky, this database would contain more valuable lncRNAs which will do a lot to the research of cancer and relative diseases. The lncRNA databases URLs and related information listed above are shown in Table 3.2.

7   circRNA Databases

7.1   starBase

starBase [49, 50] (http://starbase.sysu.edu.cn/) was established for the researchers to identify the interaction networks of RNAs systematically. The data of the circRNAs in starBase came from 108 CLIP-Seq datasets in 37 studies. Totally 9000 miRNA-circRNAs, 16,000 miRNAs pseudogene and 285,000 protein-RNA relationships wereidentified in this database. With the updated version, starBase V2.0 began to provide users with miRNA-mRNA and miRNA-lncRNA interaction networks. In this new version, developers also identified 10,000 ceRNA pairs and developed miRNA and ceRNA Functions, which made it easier for users to predict the function of the RNAs, which drafted the first interaction maps between miRNAs and circRNAs. Through this

Table 3.2 Long non-coding RNA databases

Name	Weblink	Data content
LNCipedia	https://lncipedia.org/	A public database for long non-coding RNA (lncRNA) sequence and annotation, contains 127,802 transcripts and 56,946 genes.
lncRNAdb	http://lncrnadb.com/	Database that provides comprehensive annotations of eukaryotic long non-coding RNAs.
LncRNAWiki	http://lncrna.big.ac.cn/index.php	Contains 106,063 human long non-coding RNAs.
MONOCLdb	https://www.monocldb.org/	20,728 mouse lncRNA genes.
NONCODE	http://www.noncode.org/	An integrated knowledge database of non-coding RNAs from 17 species.
lncRNome	http://genome.igib.res.in/lncRNome/	Over 17,000 long non-coding RNAs in human.
C-It-Loci	http://c-it-loci.uni-frankfurt.de/	Including the expression profiles of yet-to-be-characterized long non-coding RNAs (lncRNAs).
MiTranscriptome	http://mitranscriptome.org/	Contains over 91,000 genes.
slncky Evolution Browser	https://scripts.mit.edu/~jjenny/	Contains alignments and evolutionary metrics of conserved lncRNAs.

database, users could get more information about the interaction between miRNAs and circRNAs. With the development of technology, the star-Base continued to update their data service for searching. Furthermore, this database will to add RNA data about cancer to improve the understanding the network of circRNAs and miRNAs.

7.2 circBase

circBase [51], a database that you can browse and download the datasets of circRNAs and their expression supporting evidence, also it provided users with the sequencing data of uncovered and novel circRNAs. The available website for this database was http://www.circbase.org/. Before construct this database, developers put forward some expectation on this database, for example, it should provide the genomic context and the available expression of the circRNAs,should summarize the existence and expression for each circRNA. To achieve this goal, they putted together all the datasets from other different laboratories. For the current version of circBase, it contained the simple search interface, various methods of data retrieval and unifying, merging and annotating published datasets. There were three main ways to search for the circRNAs in

circBase, including the simple search, list search and table browser. Through these three methods, users could extend their information of circRNAs from this database. However, circBase doesn't support users to submit the new circRNA data by themselves right now. With the update of circBase, it will be more useful to the researchers in the study of circRNAs.

7.3 CircNet

CircNet [52] was constructed by 464 RNA-seq samples using transcriptome sequencing, which could be available on http://circnet.mbc.nctu.edu.tw/. This database was established by the purpose of extending the catalog of reported circRNAs. CircNet provided lots of information of the circRNAs, including novel circRNAs, integrated miRNA targets, expression profiles of circRNAs etc. When clicked on any of the circRNAs, it will display the complete information of the entire circRNA, including its position on the chromosome, length and sequence information. CireNet was the first database that provided circRNA expression profiles specific in tissue and circRNA-miRNA-gene regulatory networks. CircNet not only contained the latest circRNAs, but also provided a comprehensive analysis

between the reported circRNAs and the novel RNAs. Meanwhile, it made a regulatory network that illustrates the regulation between circRNAs, miRNAs and genes. All in all, this database was a convenient tool for the users to get the information they need easily.

7.4 Circ2Traits

Circ2Traits [53] was a database of circRNAs, which was available on http://gyanxet-beta.com/circdb/. This database was the first comprehensive database containing with the circRNAs of diseased human, which classified the circRNAs by their potential association with diseases. By analyzing the interaction of circRNAs with diseases associated miRNAs they established the association between circRNAs and the diseases. Till now the Circ2Traits contains 1951 human circRNAs with the association of 105 human diseases. What's more, this database contained the complete miRNA-circRNA-mRNA-lncRNA interaction network, in order to help users see the interaction table of each disease clearly. Meanwhile, the ceRNA regulatory network of circRNA was also constructed in this database to further analyze circRNA regulatory pathways. According to the searched circRNA, you could not only get the normal information of that circRNA, but also know its interaction sites SNPs.

7.5 CircBank

CircBank [51] database contained more than 140,000 annotated human circRNAs from many different source, which was a comprehensive database of human circRNA. The available website for this database iwas http://www.circbank.cn/. This database was publicly available with a lot of service, including circRNA modification, circRNA conservation etc. Despite the simple information, there were many new features in this database, including the predicted binding miRNA, circRNA mutation and circRNA methylation. Furthermore, CircBank also put forward

a new nomenclature system based on the host gene name, start position and end position, so that researchers will not be bothered with the ID of circRNA right now. With CircBank, it could be easier for researchers to get enough information of circRNAs they need.

7.6 exoRBase

exoRBase [54], containing circRNA, lncRNA and mRNA, which were derived from RNA-seq data analyses of human blood exosomes. It was established by Prof. Li from Fudan University, who aimed to collect and demonstrate all long RNA species in human blood exosomes. Researchers can visit this database through the website http://www.exoRBase.org/. The first version of this database contained 58,330 circRNAs, 15,501 lncRNAs and 18,333 mRNAs, based on the RNA-seq data from normal samples and disease patients. This database also provided researchers with the annotation, expression level and possible original tissues of the circRNAs. What's more, there were 77 experimental validations from the published articles included in exoRBase. Through the website, users could conveniently browse and download the information of the circRNAs. Different from other database, exoRBase allowed researchers to submit new profiles of the RNAs in human blood exosomes, which will help researchers identify new exosome biomarkers and find their influence on human diseases.

7.7 circRNADb

circRNADb [55] was a diversified-source circRNA database, which built for further study of the circRNAs and its related functions. It contained 32,194 human exonic circRNAs. It was free for the researchers to search the circRNAs they need by the web server at http://reprod.njmu.edu.cn/circrnadb. They could get various kinds of detailed information of the circRNAs, including genomic information, exon splicing, genome

sequence, internal ribosome entry site and open reading frame. The raw circRNAs dataset were collected from related literatures, and only included the circRNA which was supported more than twice. Now there were 16,328 annotated circRNAs with a longer than 100 amino acids open reading frame, and 7170 of them had internal ribosome entry site elements. CircRNADb was designed to be a comprehensive and interactive database, providing advanced search, resource download and many other functions for the users. It could do a lot to the circRNA studies by providing users with the detailed genomic and protein-coding information of each circRNA. Furthermore, developers will update the newly identified circRNAs with their detailed information to the database, in order to build circRNADb a powerful information platform for circRNAs.

7.8   CSCD

CSCD [56] (Cancer-specific circRNA database) was a cancer circRNAs specific database, which was developed by the researchers from Wuhan University. It provided genomic coordinates and gene annotation for each file of cancer-specific circRNA. The available website for this database was http://gb.whu.edu.cn/CSCD/. This database now contained total 272,152 circRNAs, which were identified from both cancer and normal cell lines. There are 950,962 circRNAs recognized from normal samples and 179,909 from both normal and tumor samples. What's more, they identified many circRNAs only in CSCD, which contained many different samples including cancer samples. This database also contained the prediction of the microRNA response element sites and RNA binding protein sites for each circRNA, which could be better used for the researchers to understand the functional effects of circRNAs. The developers also predicted each splicing event in linear transcripts of the circRNAs in order to comprehend the association between the linear splicing and the back-splicing. As the first comprehensive cancer-specific circRNA database, CSCD provided potential open

reading frames in cancer-specific circRNAs, which could significantly contribute to the circRNAs research in cancer.

7.9   circAtlas

circAtlas [57] was a database newly established by Ji et al. in 2019, which was the most abundant and comprehensive circRNAs database from normal samples. They provided many aspects of circRNAs on their circRNA Atlas web server, including their expression patterns, genomic features, conservations and functional annotations. This database was available at http://circatlas. biols.ac.cn/. Researchers could browse, visualize and prioritize the circRNAs they need and get the related information from this database. This database enlarged our knowledge of circRNAs by exploring the landscape of circRNAs in human, macaque and mouse, elucidating their diversities in various tissues. The developers also invented a new method to prioritize disease related circRNAs, which ranked the circRNAs by considering both circAtlas networks and circRNA conservation. As a starting point to investigate the biological importance of circRNAs, cireAtlas will provide a powerful foundation for circRNA studies, and help the circRNA community to annotate and prioritize circRNAs. The circRNA database URLs and related information listed above are shown in Table 3.3.

8   Conclusion

This chapter reviewed the role of ncRNAs in cardiac pathology, which abnormally elevated or decreased expression could lead cardiovascular disease. Although there are a lot of reports on these ncRNAs, but how to quickly find the information you need from tons data was an important issue that needs to be addressed. Therefore, we summarized some online databases that were currently available. These databases have powerful functions, such as predicting ncRNA targets, viewing ncRNA basic information, etc.. In addition, the biological function of the currently

Table 3.3 circRNA databases

Name	Weblink	Data content
starBase	http://starbase.sysu.edu.cn/starbase2/mirCircRNA.php	View the predicted miRNA-circRNA interactions by scanning circRNA sequences overlapping with CLIP-Seq peaks for potential microRNA targets and then output the detailed information.
circBase	http://circrna.org/	Contains thousands of recently showed circular RNAs (circRNAs) that expressed in eukaryotic cells.
CircNet	http://syslab5.nchu.edu.tw/CircNet/	Provides tissue-specific circRNA expression profiles and circRNA-miRNA-gene regulatory networks.
Circ2Traits	http://gyanxet-beta.com/circdb/	A comprehensive knowledge base on potential association of circular RNAs with diseases in human.
CircBank	http://www.circbank.cn/	A comprehensive database of human circRNAs which included more than 140,000 annotated circRNAs from different source.
exoRBase	http://www.exorbase.org/	A repository of circular RNA, long non-coding RNA (lncRNA) and messenger RNA (mRNA) derived from human blood exosomes RNA-seq data analyses.
circRNADb	http://reprod.njmu.edu.cn/circrnadb/circRNADb.php	A comprehensive database of circular RNA molecules in humans, which contains 32,914 annotated exonic circRNAs,
CSCD	http://gb.whu.edu.cn/CSCD/#	A database developed for cancer-specific circRNAs, which contains 272,152 cancer-specific circRNAs.
circAtlas	http://circatlas.biols.ac.cn	circAtlas 2.0 integrates millions of circRNAs across 7 species (human, macaca, mouse, rat, pig, chicken, dog) and a variety of tissues.

known ncRNA was only the tip of the iceberg, and there were still many unknown functions needed to be perfected. Therefore, based on the summarized existing research data, new ncRNA targets can be discovered by design algorithm.

References

1. Sabatine MS, Giugliano RP, Keech AC, Honarpour N, Wiviott SD, Murphy SA, Kuder JF, Wang H, Liu T, Wasserman SM. Evolocumab and clinical outcomes in patients with cardiovascular disease. N Engl J Med. 2017;376(18):1713–22.
2. Fitzgibbons TP, Czech MP. Epicardial and perivascular adipose tissues and their influence on cardiovascular disease: basic mechanisms and clinical associations. J Am Heart Assoc. 2014;3(2):e000582.
3. Liu C-L, Guo J, Zhang X, Sukhova GK, Libby P, Shi G-P. Cysteine protease cathepsins in cardiovascular disease: from basic research to clinical trials. Nat Rev Cardiol. 2018;1
4. Shah MS, Brownlee M. Molecular and cellular mechanisms of cardiovascular disorders in diabetes. Circ Res. 2016;118(11):1808–29.
5. Olson EN. MicroRNAs as therapeutic targets and biomarkers of cardiovascular disease. Sci Transl Med. 2014;6(239):239ps233.
6. Reuter JA, Spacek DV, Snyder MP. High-throughput sequencing technologies. Mol Cell. 2015;58(4):586–97.
7. Kukurba KR, Montgomery SB. RNA sequencing and analysis. Cold Spring Harbor Protoc. 2015;2015(11):pdb.top084970.
8. Sharma S, Garg I, Ashraf MZ. TLR signalling and association of TLR polymorphism with cardiovascular diseases. Vasc Pharmacol. 2016;87:30–7.
9. Castro EA, Peinado AB, Benito PJ, Galindo M, Gonzalez-Gross M, Cupeiro R, Group PS. What is the most effective exercise protocol to improve cardiovascular fitness in overweight and obese subjects? J Sport Health Sci. 2017;6(4):454–61.
10. Mattick JS, Makunin IV. Non-coding RNA. Hum Mol Genet. 2006;15(suppl_1):R17–29.
11. Esteller M. Non-coding RNAs in human disease. Nat Rev Genet. 2011;12(12):861.
12. Palazzo AF, Lee ES. Non-coding RNA: what is functional and what is junk? Front Genet. 2015;6:2.
13. Matsui M, Corey DR. Non-coding RNAs as drug targets. Nat Rev Drug Discov. 2017;16(3):167.
14. Beermann J, Piccoli M-T, Viereck J, Thum T. Non-coding RNAs in development and disease: back-

ground, mechanisms, and therapeutic approaches. Physiol Rev. 2016;96(4):1297–325.

15. Li Y, Liang Y, Zhu Y, Zhang Y, Bei Y. Noncoding RNAs in cardiac hypertrophy. J Cardiovasc Transl Res. 2018;11(6):439–49.

16. Li W, Notani D, Rosenfeld MG. Enhancers as non-coding RNA transcription units: recent insights and future perspectives. Nat Rev Genet. 2016;17(4):207.

17. Holdt LM, Stahringer A, Sass K, Pichler G, Kulak NA, Wilfert W, Kohlmaier A, Herbst A, Northoff BH, Nicolaou A. Circular non-coding RNA ANRIL modulates ribosomal RNA maturation and atherosclerosis in humans. Nat Commun. 2016;7:12429.

18. Gong C, Li Z, Ramanujan K, Clay I, Zhang Y, Lemire-Brachat S, Glass DJ. A long non-coding RNA, LncMyoD, regulates skeletal muscle differentiation by blocking IMP2-mediated mRNA translation. Dev Cell. 2015;34(2):181–91.

19. Sengupta S. Noncoding RNAs in protein clearance pathways: implications in neurodegenerative diseases. J Genet. 2017;96(1):203–10.

20. Anastasiadou E, Jacob LS, Slack FJ. Non-coding RNA networks in cancer. Nat Rev Cancer. 2018;18(1):5.

21. Goldberg L, Tirosh-Wagner T, Vardi A, Abbas H, Pillar N, Shomron N, Nevo-Caspi Y, Paret G. Circulating microRNAs: a potential biomarker for cardiac damage, inflammatory response, and left ventricular function recovery in pediatric viral myocarditis. J Cardiovasc Transl Res. 2018;11(4):319–28.

22. Wang L, Lv Y, Li G, Xiao J. MicroRNAs in heart and circulation during physical exercise. J Sport Health Sci. 2018;7(4):433–41.

23. Batacan RB Jr, Duncan MJ, Dalbo VJ, Buitrago GL, Fenning AS. Effect of different intensities of physical activity on cardiometabolic markers and vascular and cardiac function in adult rats fed with a high-fat high-carbohydrate diet. J Sport Health Sci. 2018;7(1):109–19.

24. Yang J-H, Shao P, Zhou H, Chen Y-Q, Qu L-H. deepBase: a database for deeply annotating and mining deep sequencing data. Nucleic Acids Res. 2009;38(suppl_1):D123–30.

25. Kozomara A, Griffiths-Jones S. miRBase: annotating high confidence microRNAs using deep sequencing data. Nucleic Acids Res. 2013;42(D1):D68–73.

26. Betel D, Wilson M, Gabow A, Marks DS, Sander C. The microRNA. org resource: targets and expression. Nucleic Acids Res. 2008;36(suppl_1):D149–53.

27. Hsu S-D, Chu C-H, Tsou A-P, Chen S-J, Chen H-C, Hsu PW-C, Wong Y-H, Chen Y-H, Chen G-H, Huang H-D. miRNAMap 2.0: genomic maps of microRNAs in metazoan genomes. Nucl Acids Res. 2007;36(suppl_1):D165–9.

28. Chiu H-S, Llobet-Navas D, Yang X, Chung W-J, Ambesi-Impiombato A, Iyer A, Kim HR, Seviour EG, Luo Z, Sehgal V. Cupid: simultaneous reconstruction of microRNA-target and ceRNA networks. Genome Res. 2015;25(2):257–67.

29. Lewis BP, Burge CB, Bartel DP. Conserved seed pairing, often flanked by adenosines, indicates that thou-

sands of human genes are microRNA targets. Cell. 2005;120(1):15–20.

30. Chou C-H, Shrestha S, Yang C-D, Chang N-W, Lin Y-L, Liao K-W, Huang W-C, Sun T-H, Tu S-J, Lee W-H. miRTarBase update 2018: a resource for experimentally validated microRNA-target interactions. Nucleic Acids Res. 2017;46(D1):D296–302.

31. Paraskevopoulou MD, Georgakilas G, Kostoulas N, Vlachos IS, Vergoulis T, Reczko M, Filippidis C, Dalamagas T, Hatzigeorgiou AG. DIANA-microT web server v5.0: service integration into miRNA functional analysis workflows. Nucleic Acids Res. 2013;41(Web Server issue):W169–73.

32. Xiao F, Zuo Z, Cai G, Kang S, Gao X, Li T. miRecords: an integrated resource for microRNA–target interactions. Nucl Acids Res. 2008;37(suppl_1):D105–10.

33. Krek A, Grün D, Poy MN, Wolf R, Rosenberg L, Epstein EJ, MacMenamin P, Da Piedade I, Gunsalus KC, Stoffel M. Combinatorial microRNA target predictions. Nat Genet. 2005;37(5):495.

34. Karagkouni D, Paraskevopoulou MD, Chatzopoulos S, Vlachos IS, Tastsoglou S, Kanellos I, Papadimitriou D, Kavakiotis I, Maniou S, Skoufos G. DIANA-TarBase v8: a decade-long collection of experimentally supported miRNA–gene interactions. Nucleic Acids Res. 2017;46(D1):D239–45.

35. Dweep H, Sticht C, Pandey P, Gretz N. miRWalk–database: prediction of possible miRNA binding sites by "walking" the genes of three genomes. J Biomed Inform. 2011;44(5):839–47.

36. Volders P-J, Helsens K, Wang X, Menten B, Martens L, Gevaert K, Vandesompele J, Mestdagh P. LNCipedia: a database for annotated human lncRNA transcript sequences and structures. Nucleic Acids Res. 2012;41(D1):D246–51.

37. Volders P-J, Verheggen K, Menschaert G, Vandepoele K, Martens L, Vandesompele J, Mestdagh P. An update on LNCipedia: a database for annotated human lncRNA sequences. Nucleic Acids Res. 2014;43(D1):D174–80.

38. Volders P-J, Anckaert J, Verheggen K, Nuytens J, Martens L, Mestdagh P, Vandesompele J. LNCipedia 5: towards a reference set of human long non-coding RNAs. Nucleic Acids Res. 2018;47(D1):D135–9.

39. Amaral PP, Clark MB, Gascoigne DK, Dinger ME, Mattick JS. lncRNAdb: a reference database for long noncoding RNAs. Nucleic Acids Res. 2010;39(suppl_1):D146–51.

40. Quek XC, Thomson DW, Maag JL, Bartonicek N, Signal B, Clark MB, Gloss BS, Dinger ME. lncRNAdb v2.0: expanding the reference database for functional long noncoding RNAs. Nucleic Acids Res. 2014;43(D1):D168–73.

41. Ma L, Li A, Zou D, Xu X, Xia L, Yu J, Bajic VB, Zhang Z. LncRNAWiki: harnessing community knowledge in collaborative curation of human long non-coding RNAs. Nucleic Acids Res. 2014;43(D1):D187–92.

42. Josset L, Tchitchek N, Gralinski LE, Ferris MT, Eisfeld AJ, Green RR, Thomas MJ, Tisoncik-Go J, Schroth GP, Kawaoka Y. Annotation of long non-

coding RNAs expressed in collaborative cross founder mice in response to respiratory virus infection reveals a new class of interferon-stimulated transcripts. RNA Biol. 2014;11(7):875–90.

43. Liu C, Bai B, Skogerbø G, Cai L, Deng W, Zhang Y, Bu D, Zhao Y, Chen R. NONCODE: an integrated knowledge database of non-coding RNAs. Nucleic Acids Res. 2005;33(suppl_1):D112–5.

44. Zhao Y, Li H, Fang S, Kang Y, Wu W, Hao Y, Li Z, Bu D, Sun N, Zhang MQ. NONCODE 2016: an informative and valuable data source of long non-coding RNAs. Nucleic Acids Res. 2015;44(D1):D203–8.

45. Bhartiya D, Pal K, Ghosh S, Kapoor S, Jalali S, Panwar B, Jain S, Sati S, Sengupta S, Sachidanandan C. lncRNome: a comprehensive knowledgebase of human long noncoding RNAs. Database. 2013;2013

46. Weirick T, John D, Dimmeler S, Uchida S. C-It-Loci: a knowledge database for tissue-enriched loci. Bioinformatics. 2015;31(21):3537–43.

47. Iyer MK, Niknafs YS, Malik R, Singhal U, Sahu A, Hosono Y, Barrette TR, Prensner JR, Evans JR, Zhao S. The landscape of long noncoding RNAs in the human transcriptome. Nat Genet. 2015;47(3):199.

48. Chen J, Shishkin AA, Zhu X, Kadri S, Maza I, Guttman M, Hanna JH, Regev A, Garber M. Evolutionary analysis across mammals reveals distinct classes of long non-coding RNAs. Genome Biol. 2016;17(1):19.

49. Yang J-H, Li J-H, Shao P, Zhou H, Chen Y-Q, Qu L-H. starBase: a database for exploring microRNA–mRNA interaction maps from Argonaute CLIP-Seq and Degradome-Seq data. Nucleic Acids Res. 2010;39(suppl_1):D202–9.

50. Li J-H, Liu S, Zhou H, Qu L-H, Yang J-H. star-Base v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein–RNA interaction networks from large-scale CLIP-Seq data. Nucleic Acids Res. 2013;42(D1):D92–7.

51. Glažar P, Papavasileiou P, Rajewsky N. circBase: a database for circular RNAs. RNA. 2014;20(11):1666–70.

52. Liu Y-C, Li J-R, Sun C-H, Andrews E, Chao R-F, Lin F-M, Weng S-L, Hsu S-D, Huang C-C, Cheng C. CircNet: a database of circular RNAs derived from transcriptome sequencing data. Nucleic Acids Res. 2015;44(D1):D209–15.

53. Ghosal S, Das S, Sen R, Basak P, Chakrabarti J. Circ2Traits: a comprehensive database for circular RNA potentially associated with disease and traits. Front Genet. 2013;4:283.

54. Li S, Li Y, Chen B, Zhao J, Yu S, Tang Y, Zheng Q, Li Y, Wang P, He X. exoRBase: a database of circRNA, lncRNA and mRNA in human blood exosomes. Nucleic Acids Res. 2017;46(D1):D106–12.

55. Chen X, Han P, Zhou T, Guo X, Song X, Li Y. circRNADb: a comprehensive database for human circular RNAs with protein-coding annotations. Sci Rep. 2016;6:34985.

56. Xia S, Feng J, Chen K, Ma Y, Gong J, Cai F, Jin Y, Gao Y, Xia L, Chang H. CSCD: a database for cancer-specific circular RNAs. Nucleic Acids Res. 2017;46(D1):D925–9.

57. Ji P, Wu W, Chen S, Zheng Y, Zhou L, Zhang J, Cheng H, Yan J, Zhang S, Yang P. Expanded expression landscape and prioritization of circular RNAs in mammals. Cell Rep. 2019;26(12):3444–3460.e5.