



# Differential Expression and Bioinformatics Analysis of circRNA in Non-small Cell Lung Cancer

Qiuwen Sun<sup>1†</sup>, Xia Li<sup>2†</sup>, Muchen Xu<sup>2</sup>, Li Zhang<sup>3</sup>, Haiwei Zuo<sup>4</sup>, Yong Xin<sup>2,5</sup>, Longzhen Zhang<sup>2,5</sup> and Ping Gong<sup>1,2\*</sup>

<sup>1</sup> School of Medical Imaging, Xuzhou Medical University, Xuzhou, China, <sup>2</sup> Department of Radiation Oncology, Affiliated Hospital of Xuzhou Medical University, Xuzhou, China, <sup>3</sup> School of Information and Control Engineering, University of Mining and Technology, Xuzhou, China, <sup>4</sup> School of Medical Information and Engineering, Xuzhou Medical University, Xuzhou, China, <sup>5</sup> Cancer Institute of Xuzhou Medical University, Xuzhou, China

## OPEN ACCESS

### Edited by:

William C. Cho,  
QEHI, Hong Kong

### Reviewed by:

Ji Heon Noh,  
Chungnam National University,  
South Korea  
Tuba Denkçeken,  
Sanko University, Turkey

### \*Correspondence:

Ping Gong  
gongping@xzhmu.edu.cn

†These authors have contributed  
equally to this work

### Specialty section:

This article was submitted to  
RNA,  
a section of the journal  
Frontiers in Genetics

Received: 24 July 2020

Accepted: 27 October 2020

Published: 24 November 2020

### Citation:

Sun Q, Li X, Xu M, Zhang L, Zuo H,  
Xin Y, Zhang L and Gong P (2020)  
Differential Expression and  
Bioinformatics Analysis of circRNA in  
Non-small Cell Lung Cancer.  
Front. Genet. 11:586814.  
doi: 10.3389/fgene.2020.586814

Circular RNA (CircRNA) plays an important role in tumorigenesis and progression of non-small cell lung cancer (NSCLC), but the pathogenesis of NSCLC caused by circRNA has not been fully elucidated. This study aimed to investigate differentially expressed circRNAs and identify the underlying pathogenesis hub genes of NSCLC by comprehensive bioinformatics analysis. Data of gene expression microarrays (GSE101586, GSE101684, and GSE112214) were downloaded from Gene Expression Omnibus (GEO) database. Differentially expressed circRNAs (DECs) were obtained by the “limma” package of R programs and the overlapping operation was implemented of DECs. CircBase database and Cancer-Specific CircRNA database (CSCD) were used to find miRNAs binding to DECs. Target genes of the found miRNAs were identified utilizing Perl programs based on miRDB, miRTarBase, and TargetScan databases. Functional and enrichment analyses of selected target genes were performing using the “cluster profiler” package. Protein-protein interaction (PPI) network was constructed by the Search Tool for the STRING database and module analysis of selected hub genes was performed by Cytoscape 3.7.1. Survival analysis of hub genes were performed by Gene Expression Profiling Interactive Analysis (GEPIA). Respectively, 1 DEC, 249 DECs, and 101 DECs were identified in GSE101586, GSE101684, and GSE112214. A total of eight overlapped circRNAs, 43 miRNAs and 427 target genes were identified. Gene Ontology (GO) enrichment analysis showed these target genes were enriched in biological processes of regulation of histone methylation, Ras protein signal transduction and covalent chromatin modification etc. Pathway enrichment analysis showed these target genes are mainly involved in AMPK signaling pathway, signaling pathways regulating pluripotency of stem cells and insulin signaling pathway etc. A PPI network was constructed based on 427 target genes of the 43 miRNAs. Ten hub genes were found, of which the expression of MYLIP, GAN, and CDC27 were significantly related to NSCLC patient prognosis. Our study provide a deeper understanding the circRNAs-miRNAs-target genes by bioinformatics analysis, which may provide novel insights for unraveling pathogenesis of NSCLC. MYLIP, GAN, and CDC27 genes might serve as novel biomarker for precise treatment and prognosis of NSCLC in the future.

**Keywords:** circRNA, expression, pathway, bioinformatics, non-small cell lung cancer

## 1. INTRODUCTION

Lung cancer has become one of the most serious malignant tumors in the world. The incidence rate of lung cancer ranks first among men and women in second place (Hu et al., 2018). There are two types of lung cancer, non-small cell lung cancer (NSCLC), and small cell carcinoma (SCLC). NSCLC accounts for approximately 85% and about 75% of NSCLC patients were in the advanced stage when they were discovered. Despite new developments for NSCLC in diagnosis and treatment, the overall survival rate remains poor and patients with advanced or metastatic have a worse prognosis (Cheung and Juan, 2017). Chemotherapy and biological targeted therapy are the best ways for the advanced stage patients of NSCLC. Thus, the identification of effective biomarkers or therapeutic targets of NSCLC is of great significance in reducing mortality and improving clinical prognosis (Hong et al., 2015).

Circular RNAs (circRNAs) are a special type of non-coding RNA molecule, which has brought great interests to researchers. Recently, circRNA has become a hotspot in RNA and transcriptome research. Compared with the traditional linear RNA, circRNA molecule has a closed circular structure and is not affected by RNA exonucleases (Chen et al., 2019). The expression of circRNAs is more stable and less prone to degradation. Studies have shown that circRNA is rich in microRNA (miRNA) binding sites and can act as miRNA sponges in cells, which can abolish the inhibition of miRNAs on their target genes and increase the expression levels of target genes (Hansen et al., 2013; Zhang et al., 2018). Hansen et al. previously identified a highly expressed circRNA in human and mouse brain. In their study, CIRS-7 acts as a miR-7 sponge and term this circRNA as circular transcript ciRS-7 (circRNA sponge for miR-7). CiRS-7 contains more than 70 selectively conserved miRNA target sites and it can strongly suppress the activity of miR-7, resulting in increased expression levels of miR-7 targets (Hansen et al., 2013). Emerging evidence suggest that circRNAs are closely related to human diseases, especially cancers, and can serve as better biomarkers because of their abundance and stability. For example, Yang et al. (2020). found that circPTK2 exert critical roles in colorectal cancer (CRC) growth and metastasis, it may serve as a potential therapeutic target for CRC metastasis. The same study also revealed that circPTK2 regulates OGD-activated microglia-induced neuronal apoptosis via miR-29b-SOCS-1-JAK2/STAT3-IL-1 $\beta$  signaling (Wang H. et al., 2019). In lung cancer, a group of circRNAs have also been found to be significantly dysregulated and several NSCLC-related circRNAs are identified. For example, Cheng et al. revealed that circTP63 was upregulated in lung squamous cell carcinoma (LUSC) tissues and its upregulation was correlated with larger tumor size and higher TNM stage in LUSC patients. *In-vitro* and *in-vivo*, elevated circTP63 promotes cell proliferation. Mechanistically, circTP63 shares miRNA response elements with FOXM1 (Cheng et al., 2019). Similarly, circPTK2 was significantly downregulated in NSCLC, circPTK2 over expression augmented TIF1 $\gamma$  expression, inhibited TGF- $\beta$ -induced EMT and NSCLC cell invasion, whereas circPTK2 knockdown had the opposite effects (Wang et al., 2018). With the development of bioinformatics technology,

there are more and more bioinformatics researches based on circRNA. However, the novel circRNAs and the function they played were still needed to explore.

In the present study, we analyzed the latest expression profile of circRNAs in lung cancer by using complex and comprehensive bioinformatics methods (Qu et al., 2018). According to the biological function of circRNA, the miRNA and target genes that bind to circRNA were predicted. Enrichment analysis of target genes provided novel insights in the treatment and prognosis of NSCLC. Besides, we performed the survival analysis to identify whether hub genes could be used as prognosis biomarkers for NSCLC. We collected the expression microarrays of circRNAs in NSCLC from Gene Expression Omnibus (GEO) datasets. Differentially expressed circRNAs (DECs) were identified by “limma” package in R software. Subsequently, the miRNAs binding to the significantly downregulated expressed circRNA were obtained by CSCD database. After predicting the miRNA-target genes, we evaluated the main functional pathways by Gene Ontology (GO) enrichment analyses and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses. Based on the differentially expressed circRNAs-miRNAs-target genes identified, the protein-protein interaction (PPI) network was constructed by the Search Tool for the Retrieval of Interacting Genes (STRING) database, then the visualization was performed by Cytoscape 3.7.1. Finally, Gene Expression Profiling Interactive Analysis (GEPIA) was used to analyse the survival of hub genes. These results provide a new direction for the mechanism and clinical treatment of NSCLC (Zhang et al., 2017). A flowchart of this study is shown in **Figure 1**.

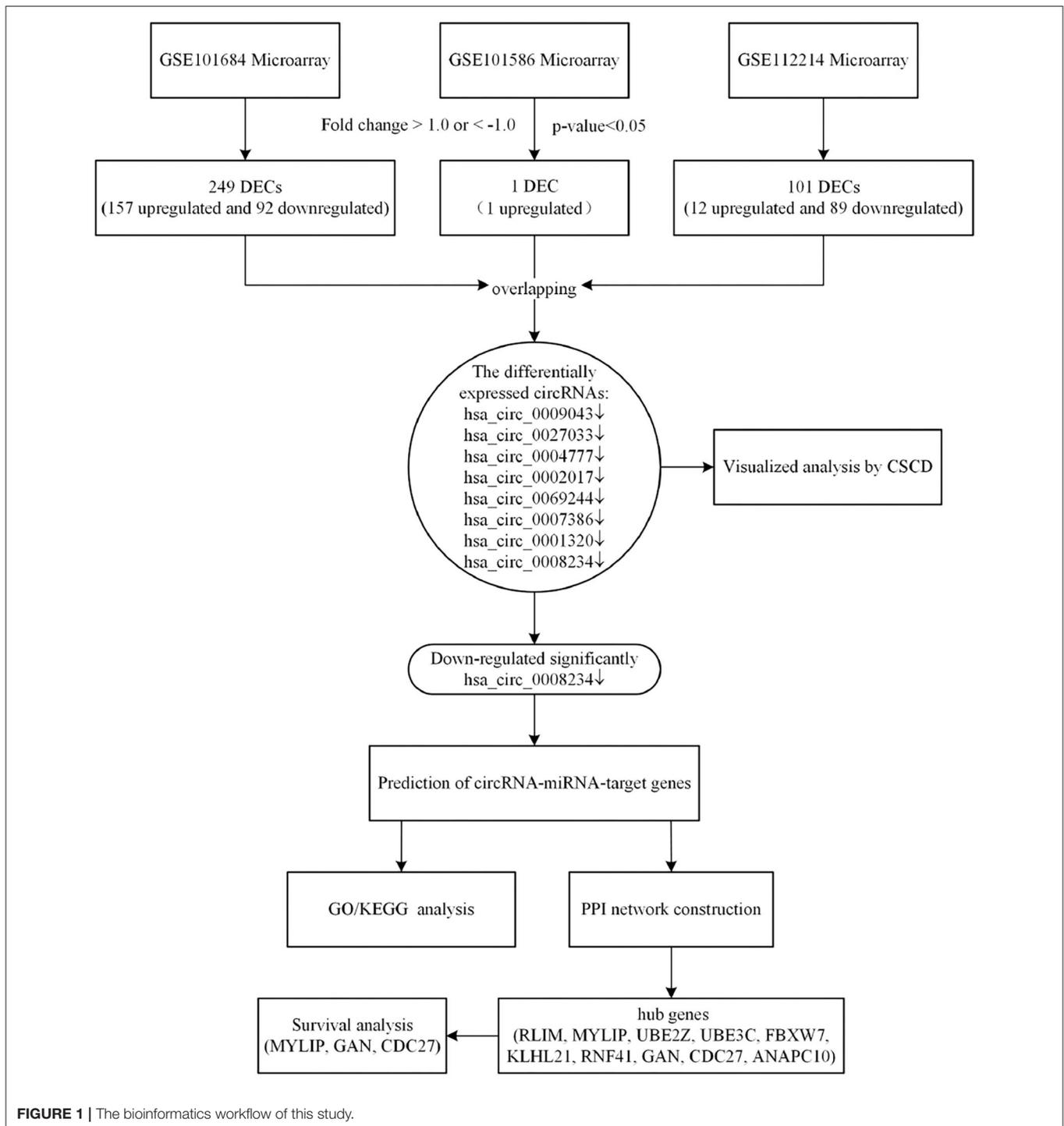
## 2. MATERIALS AND METHODS

### 2.1. Data Acquisition

The Gene Expression Omnibus database (GEO, <https://www.ncbi.nlm.nih.gov/geo/>) plays an important role in many fields, including comparative genomic analysis, proteomics, non-coding RNA, single nucleotide polymorphism genome and gene methylation status analysis (Liu et al., 2017; Sayers et al., 2018). A search for the microarray that met the requirements from the GEO dataset was performed by using “lung cancer” OR “lung tumor” AND “circRNA.” The inclusion and exclusion criteria were shown as follows: (1) microarray expression profiling of circRNAs, (2) including human NSCLC lung samples and matched adjacent samples, (3) raw matrix data and platform annotation data. After excluding the datasets with the latest time, 3 datasets [GSE101586 (GPL19978 platform), GSE101684 (GPL21825 platform), GSE112214 (GPL19978 platform)] were downloaded.

### 2.2. Data Processing

The raw microarray data were read by the R software and the expression values of the circRNAs were subjected to a base 2 logarithmic conversion. The same circRNAs in the expression matrix were combined by averaging the expression values. The circRNAs probe information were annotated by the platform file GPL21825 and GPL19978 (platform file of the microarrays). All related annotation such as probe name, platform information



and sample information etc. were removed and the related gene expression level such as gene name, sample name were kept of the obtained original microarray data.

### 2.3. Identification of Differentially Expressed circRNAs

The ID of the gene matrix data were converted to gene names by Perl script. In order to minimize the heterogeneity

among different datasets, normalization and log<sub>2</sub> conversion were applied for the raw microarray data by using the “limma” package in the R software in our analyses (Liu et al., 2017). The same circRNAs in the expression matrix were combined by averaging the expression values. Subsequently, the fold-change and Student’s t testing were used for screening for DEGs in the three sets, respectively. The criterion for DEGs was  $|\log_2\text{FoldChange}| > 1$  and the adjusted  $P - \text{value} <$

0.05. Furthermore, volcano plots were performed based on the circRNAs information (Tian et al., 2020).

## 2.4. Prediction of circRNA-miRNA-Target Gene Interactions

The circBase database and Cancer-Specific CircRNA database (CSCD, <https://gb.whu.edu.cn/CSCD/>) were used to predict miRNA binding sites (MREs) (Rong et al., 2017), CircBase provided the chromosomes location of the circRNAs, the chromosomal location and length of the RNA required for the study. Interactions between miRNA and target genes were predicted based on the miRDB, miRTarBase, and TargetScan databases. Only target genes recognized by all three databases were considered as candidate target genes (Xiong et al., 2018).

## 2.5. GO and KEGG Pathway Enrichment Analyses

In order to understand the potential functions of circRNAs, GO, and KEGG pathway enrichment analysis of target genes were performed by using the “cluster profiler” package in the R software. GO analysis cover three areas, cellular components (CC), molecular functions (MF), and biological processes (BP), each category explains the biological function of genes at different levels (Zhang et al., 2019). The KEGG pathway enrichment analysis is used for the degree of enrichment of differential genes in pathway terms.

## 2.6. Construction PPI Network and Module Analysis

Based on the target genes identified, the PPI network was constructed by using the STRING database and the visualization was performed by Cytoscape 3.7.1 (Doncheva et al., 2019). The confidence score was set as  $score > 0.4$ . The Molecular Complex Detection (MCODE) and Cytohubba in Cytoscape were used to screen modules of hub genes from the PPI network with  $MCODEscore > 2$  and  $Cytohubbadegrees > 10$ .

## 2.7. Survival Analysis of Hub Genes

GEPIA (<http://gepia.cancer-pku.cn/>) is an online tool for analyzing gene expression in cancer and healthy samples, which we used to analyze the following survival times: 50, 100, 150, 200, and 250 months (Tang et al., 2017). We performed the survival analysis to identify whether hub genes could be used as prognosis biomarkers for NSCLC.

# 3. RESULTS

## 3.1. CircRNA Profile Overview

Data from each microarray of NSCLC and adjacent normal mucosa tissues were separately analyzed by R program to screen differentially expressed circRNAs. Overall, there were 249 DECs in GSE101684, of which expression levels of 157 circRNAs were increased ( $|\log_2\text{FoldChange}| > 1, p < 0.05$ ) and 92 circRNAs were decreased ( $|\log_2\text{FoldChange}| > 1, p < 0.05$ ). There was only 1 DEC in GSE101586 and it was up-regulated.

Since GSE101586 dataset has 1 DEC, we only provide volcano maps of GSE101684 and GSE112214 (Figure 2). The volcano

map shows the number and distribution of the probes in the same plane, the horizontal axis represents the normalized difference (tumor group/normal group), the red dot indicates the up-regulated circRNAs and the green dot indicates the down-regulated circRNAs. DECs were classified and circRNA expression level were evaluated of different samples by heatmap.

In order to make the study more rigorous, we screened DECs of GSE101586, GSE101684, and GSE112214 microarrays, since there is no intersection between GSE101586 dataset and the other two datasets, we took the intersection of GSE101684 and GSE112214 datasets to obtain a total of eight DECs (Figure 3). The eight overlapped circRNAs were all down-regulated and hsa\_circ\_0008234 is the most significantly down-regulated.

## 3.2. Interaction of circRNA-miRNA-Target Genes

Bioinformatics analyses show that in mammalian cells, some circRNAs play an important role in miRNA sponges. In this study, the CircBase database gives the location of the gene and chromosome where the circRNA is located, which samples can find the circRNA and fasta base sequences. The basic characteristics of the eight circRNAs are listed in Table 1. The structural patterns of the eight circRNAs are shown in Figure 4. The fluorescent green part represents the protein that circRNA may encode, namely Open Reading Frame (ORF). The blue part represents the approximate location of circRNA binding to the protein and the position of the red small triangle represents the binding position of the circRNA to the miRNA.

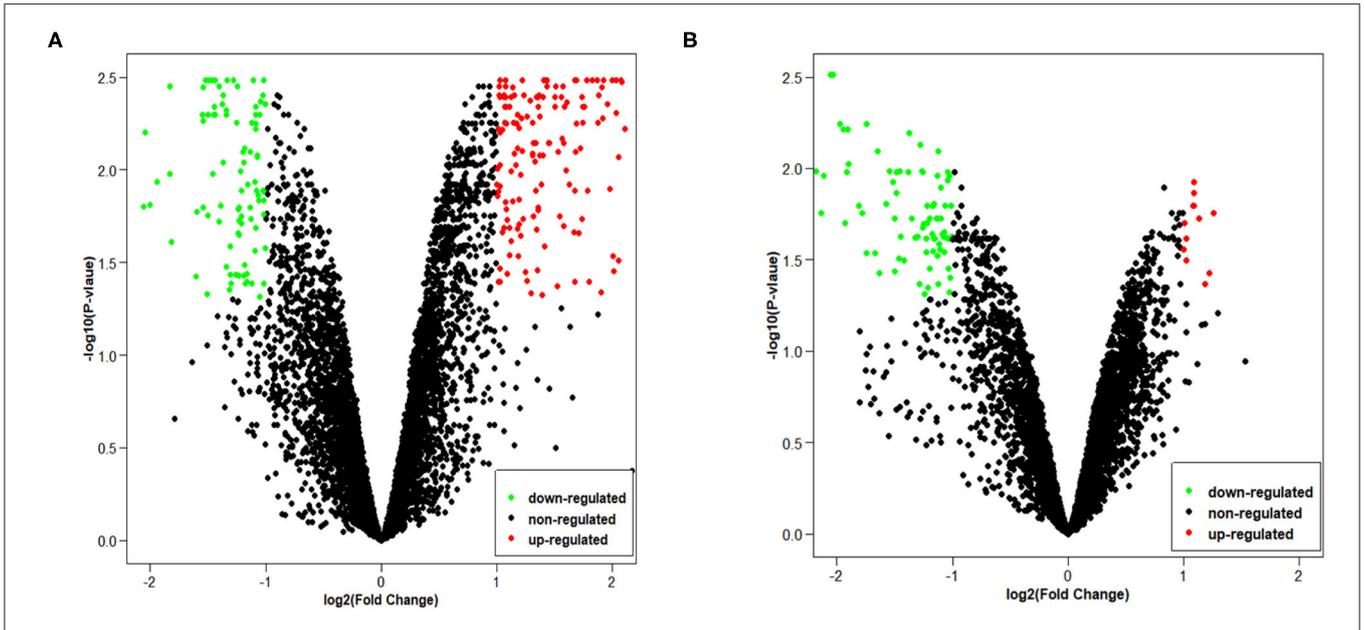
In this study, we chose hsa\_circ\_0008234, which was down-regulated significantly among eight overlapping DECs, for subsequent target gene prediction and bioinformatics analysis. A total of 43 miRNAs binding to hsa\_circ\_0008234 were identified.

Based on the miRDB, miRTarBase, and TargetScan databases, the miRNA-target genes contained in the three databases were selected by the miRNA name. This prepares for the next bioinformatics analysis. Totally, 427 target genes were binded to 43 miRNAs, respectively. Table 2 gives the target genes of the first 10 miRNAs in hsa\_circ\_0008234.

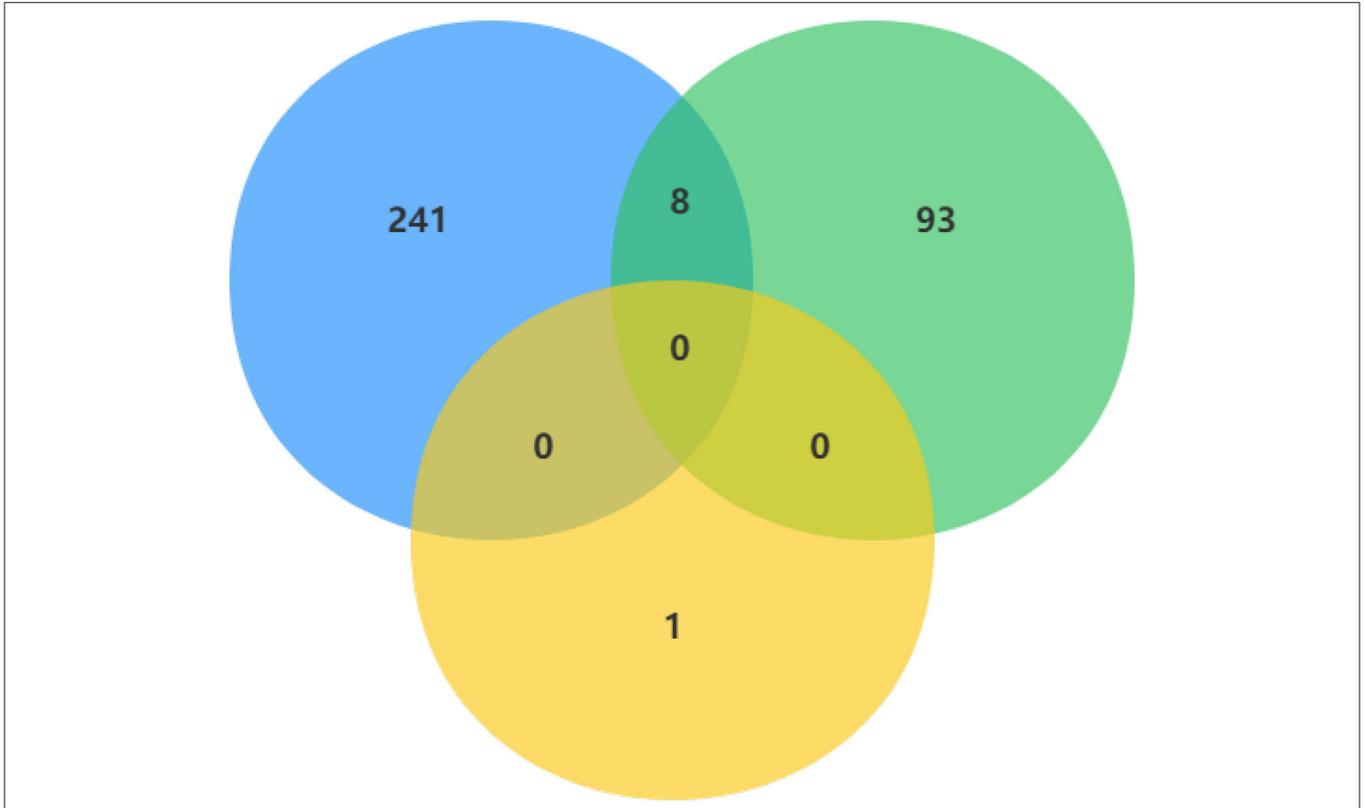
## 3.3. GO and KEGG Pathway Analyses

GO analysis was performed on hsa\_circ\_0008234 (Figure 5), which was down-regulated significantly in overlapping circRNAs. The results showed that the target genes were mainly involved in the biological processes such as “covalent chromatin modification,” “histone modification,” and “peptidyl-lysine modification” etc. As for molecular function, these genes showed enrichment in “chromatin,” “adherens junction,” and “cell-cell junction” etc. Besides, cell component indicated enrichment predominantly at “proximal promoter sequence-specific DNA binding,” “DNA-binding transcription activator activity, RNA polymerase II-specific,” and “RNA polymerase II proximal promoter sequence-specific DNA binding” etc.

KEGG pathway analysis of hsa\_circ\_0008234 showed that six significant enrichment pathways include “Insulin signaling pathway,” “Signaling pathways regulating pluripotency of stem cells,” “AMPK signaling pathway,” “Neurotrophic signaling



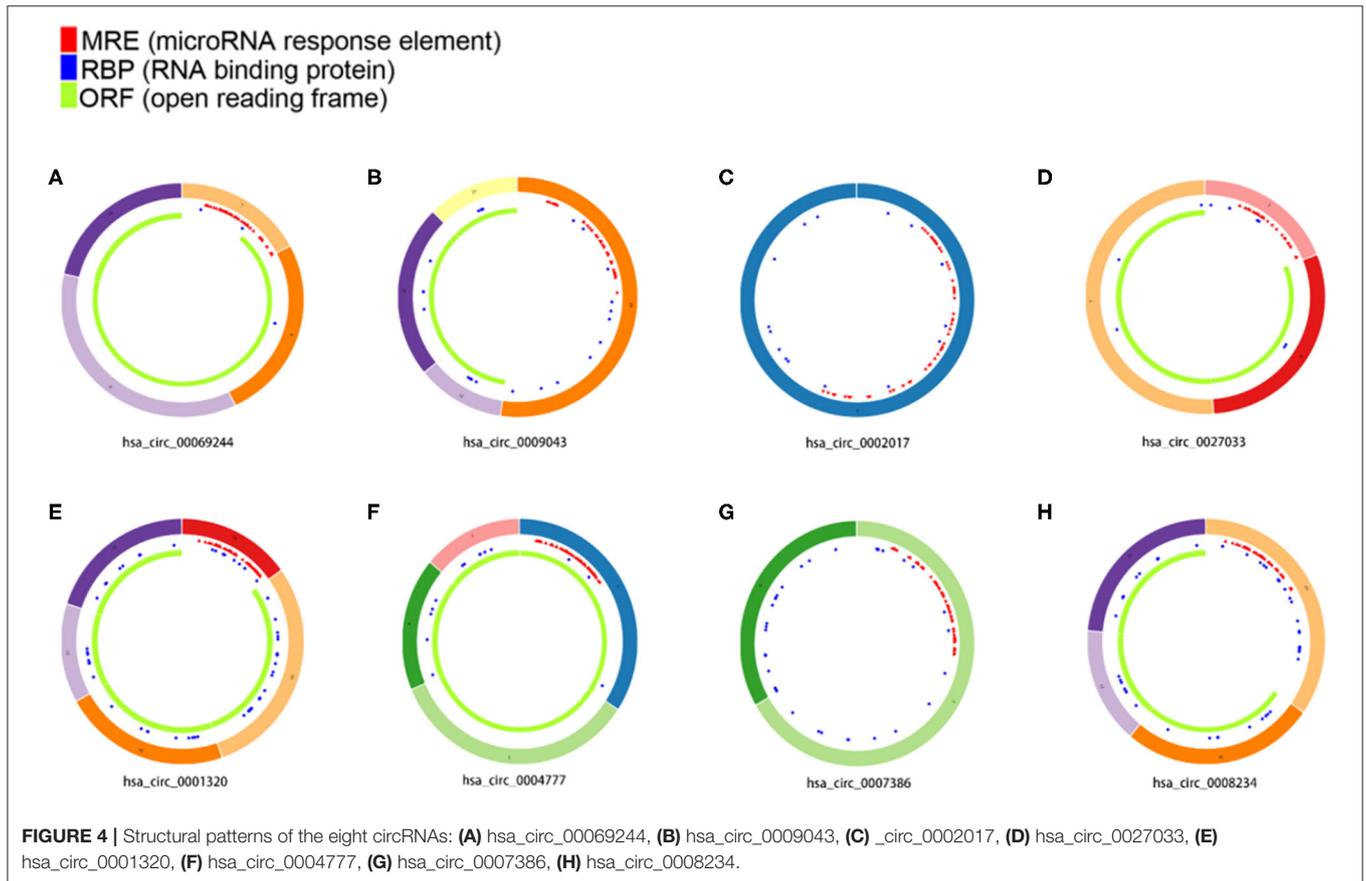
**FIGURE 2 |** The volcano plots of two chips: **(A)** GSE101684, **(B)** GSE112214.



**FIGURE 3 |** DEGs were obtained by overlapping GSE101586, GSE101684, and GSE112214 microarrays.

**TABLE 1** | Basic characteristics of the eight differentially expressed circRNAs.

CircRNA ID	Position	Genomic length	Strand	Gene symbol	Fold change
hsa_circ_0009043	Chr2:72945231-72960247	15,016	–	EXOC6B	–1.01962
hsa_circ_0027033	Chr12:56962758-56965639	2,881	+	RBMS2	–1.02682
hsa_circ_0004777	Chr1:204082042-204083733	1,691	+	SOX13	–1.08336
hsa_circ_0002017	Chr2:36623756-36623930	174	+	CRIM1	–1.09055
hsa_circ_0069244	Chr4:16587544-16760883	173,339	–	LDB2	–1.13355
hsa_circ_0007386	Chr2:36668400-36669878	1,478	+	CRIM1	–1.19527
hsa_circ_0001320	Chr3:71064699-71102924	38,225	–	FOXP1	–1.28156
hsa_circ_0008234	Chr3:71090478-71102924	12,446	–	FOXP1	–1.33598



pathway,” “Thyroid hormone signaling pathway,” and “Prostate cancer” (Figure 6).

### 3.4. PPI Network Construction, Module Analysis of Selected Hub Genes

PPI networks were constructed on the basis of STRING database. For 427 differentially expressed target genes, PPI network was mapped in Figure 7A. In total, 370 nodes and 804 edges were displayed in the PPI network. With the  $k - core = 2$  as the screening condition, the top 10 hub genes were selected according to the MCC algorithm in Cytohubba. They were RLIM, MYLIP, UBE2Z, UBE3C, FBXW7, KLHL21, RNF41, GAN, CDC27, and ANAPC10 (Figure 7B).

### 3.5. Survival Analysis of Hub Genes

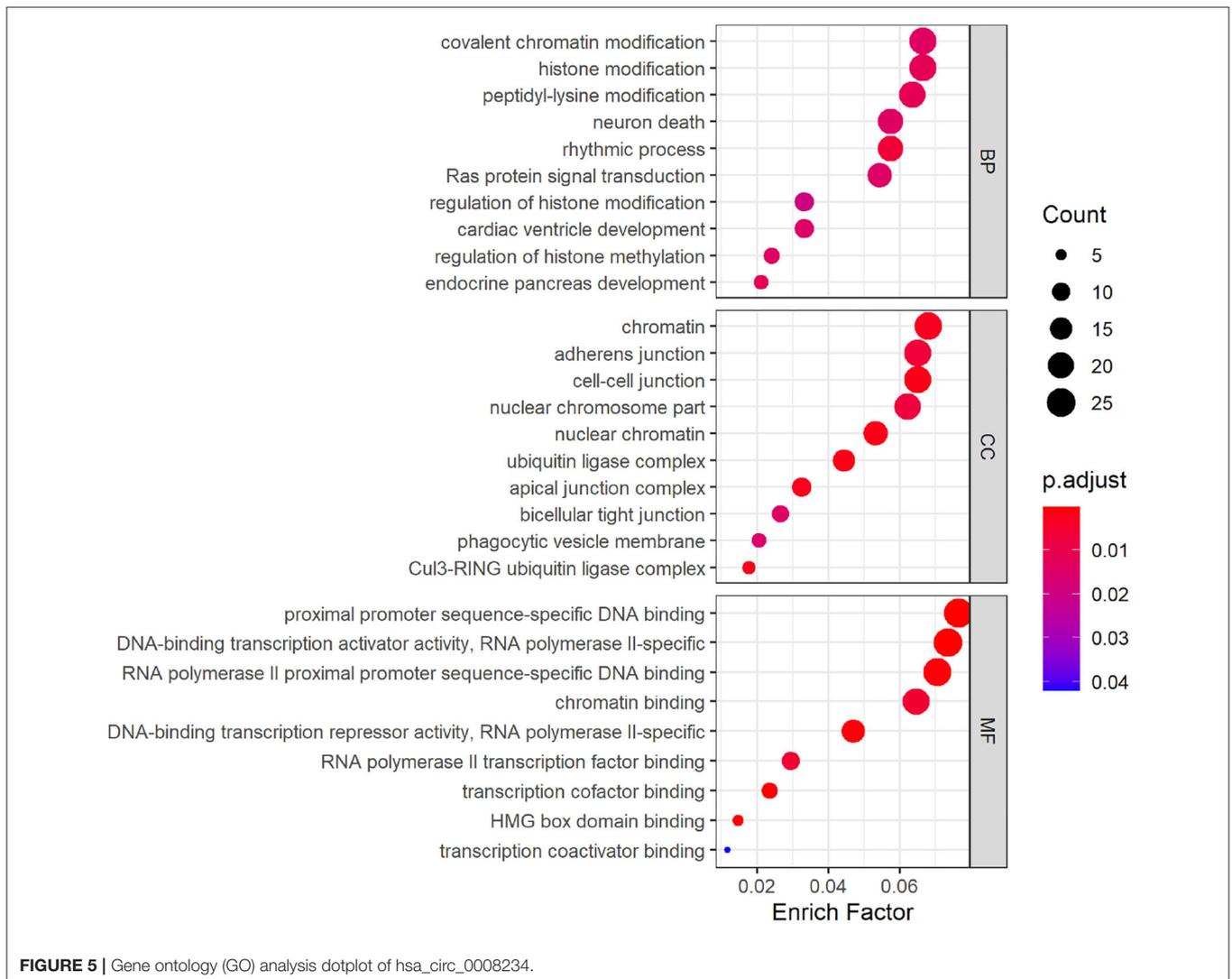
Survival analysis of the hub genes is shown in Figures 8A–J. As shown in the figure below, only MYLIP, CDC27 and GAN among the ten hub genes were found to be significantly associated with patient prognosis ( $P < 0.05$ ) so were chosen as candidate genes.

## 4. DISCUSSION

NSCLC is a type of lung malignant tumor that originates from the bronchial mucosa or glands. NSCLC has a short course and poor prognosis. It is generally found at an advanced stage, which is very difficult to control and treat. CircRNA is a special non-coding RNA with a closed loop structure that abundantly expressed in eukaryotic cells, they are usually expressed in

**TABLE 2** | Target genes of the first 10 miRNAs in hsa\_circ\_0008234.

miRNA	Target-Genes
hsa-miR-1272	RTN4,PHKA1,BCAS2,COIL,G2E3,FAM83C,TMEM167A,SMNDC1,SAR1B
hsa-miR-1322	IRF2BPL,C11orf58,CTCF,UCHL5
hsa-miR-150-5p	PDCD4,MTMR9,PDIA6,HILPDA,GK5,ELK1,XPNP3,AHI1,EREG,ZEB1,AIFM2
hsa-miR-1910-5p	CLIC4,GSK3B,MTHFD2
hsa-miR-2116-3p	MBD6,EREG,ZBTB18,AP4S1,SRRM4,TXNIP,KRBOX4
hsa-miR-296-3p	TCERG1,HILPDA,ZNF138,ZNF117,GRINA,ZNF99
hsa-miR-31-3p	EPB41L4B,SLC30A5,FRS2,PLEKHB2,CRK,CHMP4B,BACH1
hsa-miR-3677-5p	TBC1D12,SLC7A1,NCMAP,HMGXB3
hsa-miR-4264	ESR1,KLHL15,PALM2,TM4SF1,ADIPOR2,SLCO3A1,SENP1,THRB,FOXK2,PODXL
hsa-miR-4520-3p	PRKX,RAB25,H2AFZ,RAB10,MYOCD,ABAT,NUFIP2



tissues or at developmental stages and most sequences are highly conserved (Wolf et al., 2017). Recently, with the development of gene sequencing data and bioinformatics technology, circRNA has been reported in various types of cancer for diagnosis and

treatment. In lung cancer, Cai et al. (2020) analyzed circRNA and gene expression data to detect differentially expressed circRNAs in NSCLC patients, their results explore the pathogenesis of NSCLC to identify novel treatment biomarkers. Liu et al. (2019)

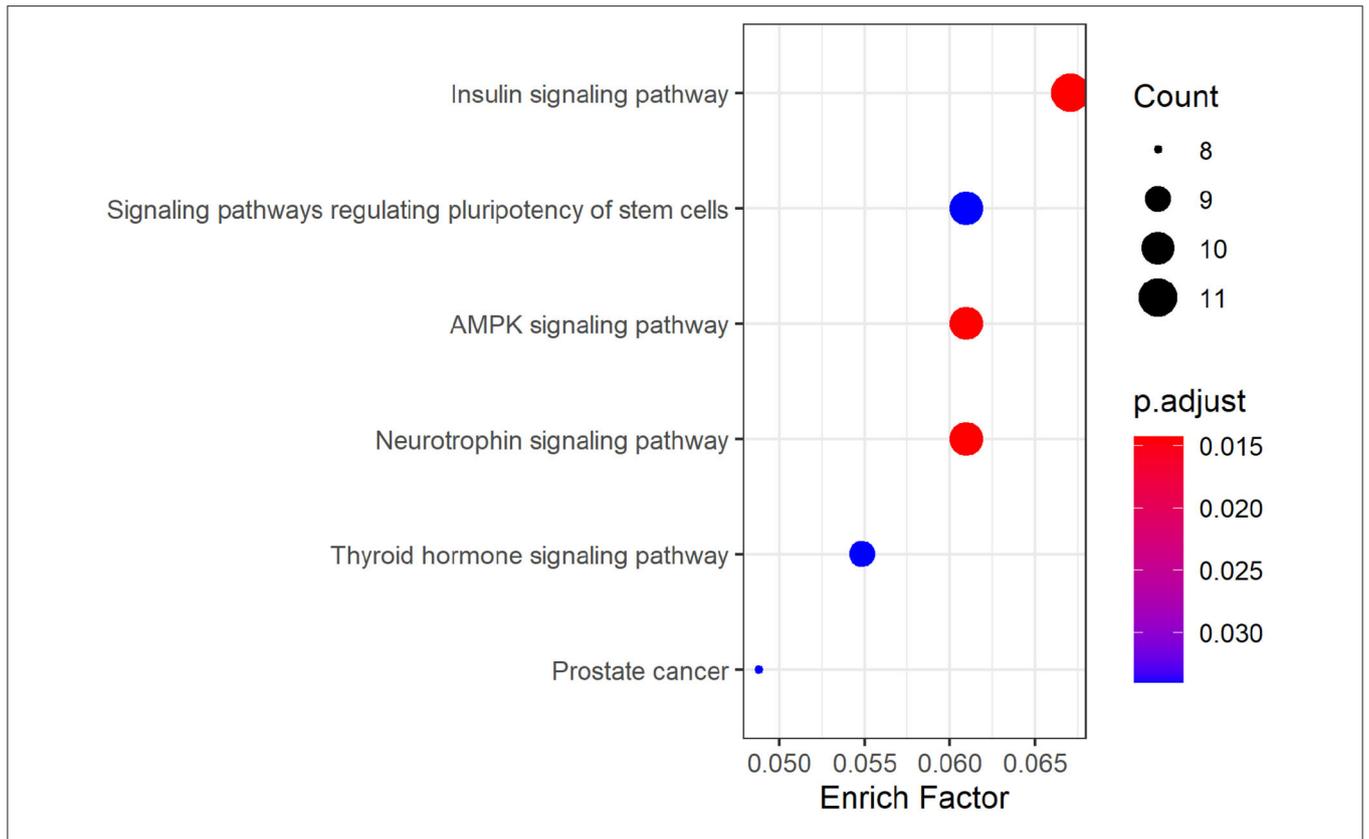


FIGURE 6 | KEGG pathway analysis dotplot of hsa\_circ\_0008234.

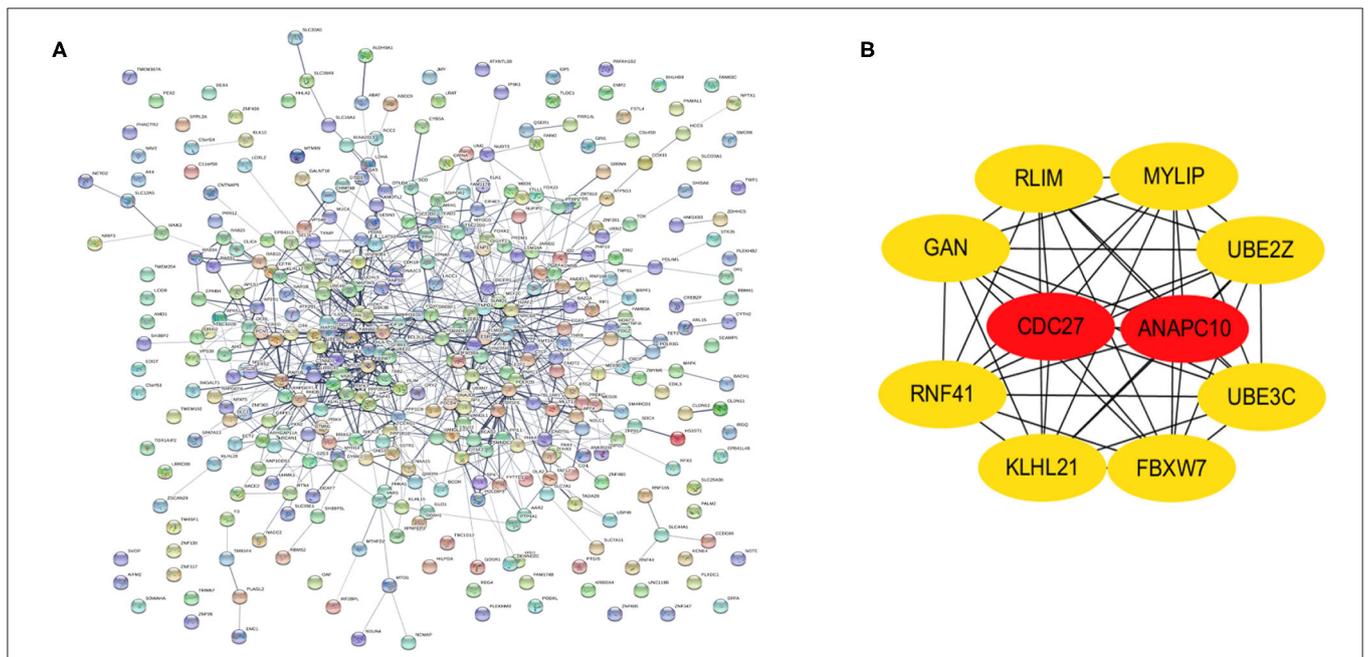
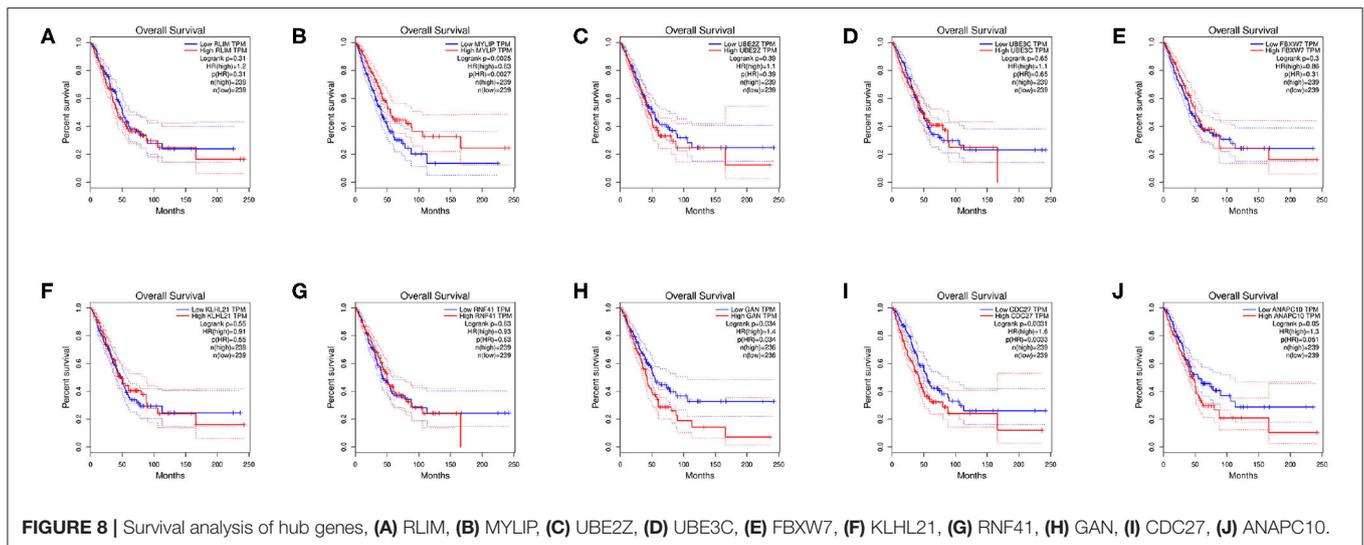


FIGURE 7 | Identification of hub genes from the PPI network with the MCODE and Cytohubba algorithm. (A) PPI network of 370 genes. (B) PPI network of ten hub genes that extracted from the PPI network.



identified potential plasma circRNA biomarkers for the diagnosis of lung adenocarcinoma (LUAD), and constructed a ceRNA network to predict the possible mechanisms and function of circRNAs in LUAD. However, the above two studies only used single chip and the samples were less, these results may not be ideally validated in large sample data. Due to the rapid development of gene sequencing technology, a large amount of biological data has been accumulated and some studies have used multiple microarrays to analyse. For example, Li et al. (2020) used the same public datasets as this study, they took the union of GSE101586, GSE101684, and GSE11214 to identify differentially expressed circRNAs among NSCLC and healthy individuals for early diagnosis of NSCLC. Liang et al. (2020) used the intersection of GSE101586, GSE101684 microarrays and GSE104854 high-throughput sequencing data to exploring the aberrant circRNAs and their potential molecular function in LUAD. Although the two GEO microarrays were same as ours, since high-throughput sequencing data and microarray data are different types, the results obtained by overlapping were different. Meanwhile, previous researches mainly focused on the role of circRNA in the early diagnosis of NSCLC, little is known about the expression of miRNA-target genes that combined with circRNAs in the treatment and prognosis of advanced stage NSCLC patients.

In the current study, we used three microarrays to identify 1 DEC in GSE101586, 249 DECs in GSE101684, and 101 DECs in GSE112214. The 8 DECs were screened out from the three sets of microarrays, including hsa\_circ\_0009043, hsa\_circ\_0027033, hsa\_circ\_0004777, hsa\_circ\_0002017, hsa\_circ\_0069244, hsa\_circ\_0007386, hsa\_circ\_0001320, and hsa\_circ\_0008234. Based on CircBase database and the CSCD database, 43 miRNAs that bind to hsa\_circ\_0008234 (the most significant downregulation among 8 circRNAs) were predicted. Moreover, hsa\_circ\_0008234, was identified previously by Wang et al., who analyzed 40 gallbladder cancer (GBC) patient tissues and normal tissues, suggesting that hsa\_circ\_0008234 has pleiotropic effects,

including promotion of cell proliferation, migration, invasion and inhibition of cell apoptosis in GBC (Wang S. et al., 2019). Additionally, hsa\_circ\_0008234 acted as the sponge of miR-370 to regulate PKLR, resulting in promoting Warburg effect in GBC progression. The prediction of target genes and enrichment analysis were implemented by multiple bioinformatics tools. Enrichment of these genes affected by DECs, which may provide novel insights for unraveling pathogenesis of NSCLC.

As was suggested by GO analysis, target genes of hsa\_circ\_0008234 combined with miRNAs in NSCLC were enriched in biological processes of “covalent chromatin modification,” “histone modification,” “neuron death,” “Ras protein signal transduction,” “cardiac ventricle development,” and “endocrine pancreas development” et al.. Molecular function of GO analysis showed enrichment “inubiquitin ligase complex,” “nuclear chromatin,” “adherens junction,” “bicellular tight junction,” and “phagocytic vesicle membrane” et al.. It is reasonable because frequent cell proliferation and loss of cell adhesion is apparent hallmark of cancers including NSCLC (Hanahan and Weinberg, 2011). In addition, covalent chromatin modification, histone modification and peptidyl lysine modification are the three main epigenetic regulation processes of tumors including lung cancer on the basis of chromatin. These regulatory processes are regulated by a variety of proteins with enzymatic activity and mutations or abnormal expression of these proteins are prone to cause various diseases such as tumors (Liu et al., 2015). KEGG pathway enrichment analysis suggested significant enrichment in pathways including “Insulin signaling pathway,” “AMPK signaling pathway,” and “Neurotrophin signaling pathway” etc.. It was consistent with the fact that AMPK can activate the corresponding pathway or inhibit the corresponding pathway and can significantly antagonize NSCLC (Xu et al., 2018).

PPI network of target genes illustrated the overview of their functional connections, of which top10 hub genes were selected: RLIM, MYLIP, UBE2Z, UBE3C, FBXW7, KLHL21, RNF41,

GAN, CDC27, ANAPC10. Here, we observed that the survival analysis of MYLIP, CDC27, and GAN revealed a correlation between high expression and poor prognosis. MYLIP is a myosin-regulated light chain interacting protein, which belongs to the cytoskeletal protein cluster and participates in the regulation of cell movement and migration (Zhao et al., 2017). The aberrant expression of MYLIP is seen in the occurrence and development of several diseases. Zhao et al. (2017) found that the regulatory mechanisms of miR-19b promoting breast cancer metastasis through directly targeting MYLIP expression and affecting the expression levels of cell adhesion molecules. Here, we observed that survival analysis of MYLIP revealed a correlation between high expression and poor prognosis. Dysregulation of cell cycle progression plays a key role in tumorigenesis. CDC27 is a core subunit that promotes late complexes/loops. Studies have shown that overexpression of CDC27 promotes the proliferation of colorectal cancer and the cell-dependent approach promotes the ability to transfer and form balls (Qiu et al., 2017). Additionally, Xin et al. (2018) demonstrated that the CDC27 expression was obviously increased in gastric cancer tissues, and significantly correlates with EMT-related biomarkers and poor 5-year overall survival. We speculate that CDC27 participates in the development of NSCLC by regulating the cell cycle, apoptosis, and inflammation levels. Gigaxonin (GAN), an E3 ligase adaptor protein. Given the role of GAN in the ubiquitin proteasome complex, it is possible that it could play a role in oncogenesis. Veena et al. found that GAN directly interacts with NF- $\kappa$ B, and this interaction is dependent on the interaction between p16. It is well-known that p16 is an important tumor suppressor gene that renders cancer cells susceptible to cisplatin treatment. They believe that GAN expression could be the biomarker of cancer (Veena et al., 2014). In our study, survival analysis of GAN found that high expression was associated with poor prognosis. The related regulatory mechanism deserves further exploration.

Bioinformatics analysis of this study suggests that differentially expressed circRNAs in NSCLC support the important role of these molecules in the development of NSCLC. Clarifying the underlying mechanisms of the initiation and development of NSCLC would greatly benefit the diagnosis, treatment, and prognosis evaluation. The data in this study provide a new perspective on the mechanisms of lung cancer metastasis and recurrence and highlight potential therapeutic targets that can help improve diagnosis and treatment. Due to the restriction of

data and time, this study did not analyze clinical parameters and prognosis. We need further clinical molecular experiments to better confirm the findings of identified genes and pathways in NSCLC that we studied.

## 5. CONCLUSION

Our study screened differentially expressed circRNAs, miRNAs, and target genes from publicly available microarray data of NSCLC by joint bioinformatics analysis, which can broaden the perspective of the gene study of lung cancer and lay a foundation for further studies on the role of circRNAs in NSCLC. Compared with individual investigation, this study is possible to come up with more reliable and accurate screening results via overlapping relevant data sets. Despite the rigorous bioinformatics analysis of this study, there are still some shortcomings. Further clinical experiments and different gene chip of other platforms are needed to confirm the findings of the identified candidate genes in NSCLC. In conclusion, hub genes including MYLIP, GAN, and CDC27 might serve as biomarkers for precise diagnosis and treatment of NSCLC in the future.

## DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found at: <https://www.ncbi.nlm.nih.gov/gds/?term=GSE101684>; <https://www.ncbi.nlm.nih.gov/gds/?term=GSE112214>; <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE101586>.

## AUTHOR CONTRIBUTIONS

QS and XL carried out data analysis. MX, LZ, HZ, and YX participated in study design and data collection. PG and LZ conceived the study. All authors read and approved the final manuscript. All authors contributed to the article and approved the submitted version.

## FUNDING

This study was supported by the National Natural Science Foundation of China (Nos. 81972845 and 82001987).

## REFERENCES

- Cai, X., Lin, L., Zhang, Q., Wu, W., and Su, A. (2020). Bioinformatics analysis of the circRNA-miRNA-mRNA network for non-small cell lung cancer. *J. Int. Med. Res.* 48, 1–15. doi: 10.1177/0300060520929167
- Chen, X., Mao, R., Su, W., Yang, X., Geng, Q., Guo, C., et al. (2019). Circular RNA circHIPK3 modulates autophagy via MIR124-3p-STAT3-PRKAA/AMPK  $\alpha$  signaling in STK11 mutant lung cancer. *Autophagy* 16, 659–671. doi: 10.1080/15548627.2019.1634945
- Cheng, Z., Yu, C., Cui, S., Wang, H., Jin, H., Wang, C., et al. (2019). circTP63 functions as a ceRNA to promote lung squamous cell carcinoma progression by upregulating foxm1. *Nat. Commun.* 10:3200. doi: 10.1038/s41467-019-11162-4
- Cheung, C. H. Y., and Juan, H.-F. (2017). Quantitative proteomics in lung cancer. *J. Biomed. Sci.* 24:37. doi: 10.1186/s12929-017-0343-y

- Doncheva, N., Morris, J., Gorodkin, J., and Jensen, L. (2019). Cytoscape stringapp: network analysis and visualization of proteomics data. *J. Proteome Res.* 18, 623–632. doi: 10.1021/acs.jproteome.8b00702
- Hanahan, D., and Weinberg, R. A. (2011). Hallmarks of cancer: the next generation. *Cell* 144, 646–674. doi: 10.1016/j.cell.2011.02.013
- Hansen, T. B., Jensen, T. I., Clausen, B. H., Bramsen, J. B., Finsen, B., Damgaard, C. K., et al. (2013). Natural RNA circles function as efficient microRNA sponges. *Nature* 495, 384–388. doi: 10.1038/nature11993
- Hong, Q., Wu, G., Qian, G., Hu, C., Zhou, J., Chen, L., et al. (2015). Prevention and management of lung cancer in china. *Cancer* 121(Suppl. 17), 3080–3088. doi: 10.1002/cncr.29584
- Hu, W., Bi, Z.-Y., Chen, Z., Liu, C., Li, L.-L., Zhang, F., et al. (2018). Emerging landscape of circular RNAs in lung cancer. *Cancer Lett.* 427, 18–27. doi: 10.1016/j.canlet.2018.04.006

- Li, L., Sun, D., Li, X., Yang, B., and Zhang, W. (2020). Identification of key circRNAs in non-small cell lung cancer. *Am. J. Med. Sci.* doi: 10.1016/j.amjms.2020.08.008. [Epub ahead of print].
- Liang, L., Zhang, L., Zhang, J., Bai, S., and Fu, H. (2020). Identification of circRNA-miRNA-mRNA networks for exploring the fundamental mechanism in lung adenocarcinoma. *Oncotargets Ther.* 13, 2945–2955. doi: 10.2147/OTT.S235664
- Liu, J., Li, H., Sun, L., Wang, Z., Xing, C., and Yuan, Y. (2017). Aberrantly methylated-differentially expressed genes and pathways in colorectal cancer. *Cancer Cell Int.* 17:75. doi: 10.1186/s12935-017-0444-4
- Liu, W., Han, F., Jiang, X., Yin, L., Chen, H., Li, Y., et al. (2015). Epigenetic regulation of ANKRD18B in lung cancer. *Mol. Carcinog.* 54, 312–321. doi: 10.1002/mc.22101
- Liu, X., Yang, Y.-E., Zhang, M., Li, R., Yin, Y., and Qu, Y.-Q. (2019). A two-circular RNA signature as a noninvasive diagnostic biomarker for lung adenocarcinoma. *J. Transl. Med.* 17:50. doi: 10.1186/s12967-019-1800-z
- Qiu, L., Tan, X., Lin, J., Liu, R.-Y., Chen, S. X., Geng, R., et al. (2017). CDC27 induces metastasis and invasion in colorectal cancer via the promotion of epithelial-to-mesenchymal transition. *J. Cancer* 8, 2626–2635. doi: 10.7150/jca.19381
- Qu, D., Yan, B., Xin, R., and Ma, T. (2018). A novel circular RNA hsa\_circ\_0020123 exerts oncogenic properties through suppression of miR-144 in non-small cell lung cancer. *Am. J. Cancer Res.* 8, 1387–1402.
- Rong, D., Sun, H., Li, Z., Liu, S., Dong, C., Fu, K., et al. (2017). An emerging function of circRNA-miRNAs-mRNA axis in human diseases. *Oncotarget* 8, 73271–73281. doi: 10.18632/oncotarget.19154
- Sayers, E. W., Barrett, T., Benson, D. A., Bolton, E., Bryant, S. H., Canese, K., et al. (2018). Database resources of the national center for biotechnology information. *Nucl. Acids Res.* 46, D8–D13. doi: 10.1093/nar/gkr1184
- Tang, Z., Li, C., Kang, B., Gao, G., Li, C., and Zhang, Z. (2017). Gepia: a web server for cancer and normal gene expression profiling and interactive analyses. *Nucl. Acids Res.* 45, W98–W102. doi: 10.1093/nar/gkx247
- Tian, J., Fu, Y., Li, Q., Xu, Y., Xi, X., Zheng, Y., et al. (2020). Differential expression and bioinformatics analysis of CircRNA in PDGF-BB-induced vascular smooth muscle cells. *Front. Genet.* 11:530. doi: 10.3389/fgene.2020.00530
- Veena, M., Wilken, R., Zheng, J., Gholkar, A., Venkatesan, N., Vira, D., et al. (2014). p16 protein and gigaxonin are associated with the ubiquitination of nfkb in cisplatin-induced senescence of cancer cells. *J. Biol. Chem.* 289, 34921–34937. doi: 10.1074/jbc.M114.568543
- Wang, H., Li, Z., Gao, J., and Liao, Q. (2019). Circular RNA circPTK2 regulates oxygen-glucose deprivation-activated microglia-induced hippocampal neuronal apoptosis via miR-29b-SOCS-1-JAK2/STAT3-IL-1 $\beta$  signaling. *Int. J. Biol. Macromol.* 129, 488–496. doi: 10.1016/j.ijbiomac.2019.02.041
- Wang, L., Tong, X., Zhou, Z., Wang, S., Lei, Z., Zhang, T., et al. (2018). Circular RNA hsa\_circ\_0008305 (circPTK2) inhibits TGF- $\beta$ -induced epithelial-mesenchymal transition and metastasis by controlling TIF1 $\gamma$  in non-small cell lung cancer. *Mol. Cancer* 17:140. doi: 10.1186/s12943-018-0889-7
- Wang, S., Zhang, Y., Cai, Q., Ma, M., Jin, L. Y., Weng, M., et al. (2019). Circular RNA FOXP1 promotes tumor progression and Warburg effect in gallbladder cancer by regulating PKLR expression. *Mol. Cancer* 18:145. doi: 10.1186/s12943-019-1078-z
- Wolf, F. A., Angerer, P., and Theis, F. J. (2017). SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol.* 19:15. doi: 10.1101/174029
- Xin, Y., Ning, S., Zhang, L., and Cui, M. (2018). CDC27 facilitates gastric cancer cell proliferation, invasion and metastasis via twist-induced epithelial-mesenchymal transition. *Cell. Physiol. Biochem.* 50, 501–511. doi: 10.1159/000494164
- Xiong, D., Dang, Y., Lin, P., Wen, D., He, R.-Q., Luo, D., et al. (2018). A circRNA-miRNA-mRNA network identification for exploring underlying pathogenesis and therapy strategy of hepatocellular carcinoma. *J. Transl. Med.* 16:220. doi: 10.1186/s12967-018-1593-5
- Xu, F., Cui, W., Wei, Y., Cui, J., Qiu, J., Hu, L., et al. (2018). Astragaloside IV inhibits lung cancer progression and metastasis by modulating macrophage polarization through AMPK signaling. *J. Exp. Clin. Cancer Res.* 37:207. doi: 10.1186/s13046-018-0878-0
- Yang, H., Li, X., Meng, Q., Sun, H., Wu, S., Hu, W., et al. (2020). CircPTK2 (hsa\_circ\_0005273) as a novel therapeutic target for metastatic colorectal cancer. *Mol. Cancer* 19:13. doi: 10.1186/s12943-020-1139-3
- Zhang, R., Li, Y., Liu, Q. M., Gao, Y., Du, J., Ma, J., et al. (2019). Differential expression profiles and functional prediction of circular RNAs and long non-coding RNAs in the hippocampus of NRF2-knockout mice. *Front. Mol. Neurosci.* 12:196. doi: 10.3389/fnmol.2019.00196
- Zhang, S., Zeng, X., Ding, T., Guo, L., Li, Y., Ou, S., et al. (2018). Microarray profile of circular RNAs identifies hsa\_circ\_0014130 as a new circular RNA biomarker in non-small cell lung cancer. *Sci. Rep.* 8:2878. doi: 10.1038/s41598-018-21300-5
- Zhang, S., Zhu, D., Li, H., Li, H., Feng, C., and Zhang, W. (2017). Characterization of CircRNA-associated-ceRNA networks in a senescence-accelerated mouse prone 8 brain. *Mol. Ther.* 25, 2053–2061. doi: 10.1016/j.ymthe.2017.06.009
- Zhao, L., Zhao, Y., He, Y., and Mao, Y. (2017). miR-19b promotes breast cancer metastasis through targeting MYLIP and its related cell adhesion molecules. *Oncotarget* 8, 64330–64343. doi: 10.18632/oncotarget.19278

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Sun, Li, Xu, Zhang, Zuo, Xin, Zhang and Gong. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.