Am J Transl Res 2022;14(11):8049-8063
www.ajtr.org /ISSN:1943-8141/AJTR0145542

Original Article

Prediction and screening of circRNA in triple-negative breast cancer

Xuanhe Li, Song Zhang, Siyu Sun, Xicheng Yue, Liyu Qian, Jie Tang, Fangqian Jiang, Jianfei Lu, Yifan Cao, Shengwen Meng, Tingjing Yao

Oncology Surgery, The First Affiliated Hospital of Bengbu Medical College, 287 Changhuai Road, Bengbu 233000, Anhui, China

Received July 25, 2022; Accepted October 28, 2022; Epub November 15, 2022; Published November 30, 2022

Abstract: Objective: The purpose of this investigation was to study the expression profile and potential function of circular RNA (circRNA) and long noncoding RNA (lncRNA) in triple-negative breast cancer (TNBC). Methods: RNA sequencing technology was used to detect differentially expressed circRNAs and lncRNAs between TNBC tissues and the adjacent tissue. The potential functions of these different RNAs were analyzed by GO and KEGG enrichment analysis by bioinformatics tools. We also selected and analyzed these key circRNAs and lncRNAs to verify their important functions in TNBC. Results: A total of 139 differentially expressed circRNAs and 1001 lncRNAs were obtained. The co-expression analysis showed that the hub lncRNAs (OIP5-AS1, DRAIC) were associated with several tumors and mainly enriched in tumor metastasis. We also screened 5 circRNA-hosting genes (NTRK2, FNTA, BAPGEF2, MGST2, ADH1B) that were associated with the brain-derived neurotrophic factor (BDNF) receptor signaling pathway and cerebral cortex development, as well as AMPK and TGF-β signaling pathway. Conclusion: We identified a large number of differentially expressed circRNAs and lncRNAs, which provide useful insight in understanding TNBC carcinogenesis.

Keywords: TNBC, circRNA, lncRNA, enrichment analysis

Introduction

Breast cancer (BC) is one of the most common gynecological tumors among women throughout the world [1]. BC can be divided into three broad subtypes according the expression status of estrogen receptor, progesterone-receptor and human epidermal growth factor receptor 2 [2]. Triple-negative breast cancer (TNBC), with negative expression of all three receptors of these receptors, is more aggressive than other BC types [3]. Patients with TNBC have a higher risk of recurrence, earlier and multiple organ metastasis and worse prognosis [4].

Recently, the application of high throughput sequencing (NGS) technology has facilitated the discovery and investigation of genes in patients with cancer [5]. Reports have found several gene alternations and abnormal pathways that are associated with the occurrence of TNBC [6, 7]. In addition, the non-coding RNAs and mRNAs play important roles in regulation of cancer initiation and progression [8]. lncRNAs are long-stranded RNAs without coding function, with a transcription lengths varying from 200 bp to 1000 kb [9]. There are two main regulation modes of lncRNAs; one is cis regulation, where lncRNA regulates the expression of its neighboring genes [10]. The cis-regulated target genes are mainly predicted based on the positional relationship, which are defined as abnormally expressed lncRNAs and abnormally expressed mRNAs upstream and downstream of the chromosome by 100 kbp [11]. Cis-acting elements are sequences that affect gene expression in the flanking sequences of genes, including promoters, enhancers, regulatory sequences, and induction elements. They participate in the regulation of gene expression in the nucleus, and are usually transcribed into non-coding RNA [12]. The second type of regulation is trans regulation, where lncRNAs regulate the expression of genes across chromosomes. Their target genes mainly depend on the amount of free energy between lncRNA and

mRNA sequences. Studies have shown that lncRNAs function in DNA epigenetics and participate in multiple biological regulatory processes [13, 14].

Moreover, circular RNAs (circRNAs) also play important roles in cancer research. As a new category of non-coding RNAs, circRNAs have had been used as diagnostic and prognostic biomarkers for their high stability and tissue/stage specificity [15]. Four types of circRNAs, including all-exon circRNA, EICircRNA with a combination of introns and exons, Lasso-type ciRNA composed of introns and viral RNA circRNA, are produced by circularization of genome [16]. Except for a small part of the ciRNAs or EICircRNA produced by intron circularization located in the nucleus, most of the circRNAs are mainly located in the cytoplasm. Different types of circRNAs have different biological functions, such as ciRNAs or EICircRNAs located in the nucleus which are mainly involved in transcriptional regulation. Since the competing endogenous RNA (ceRNA) hypothesis was proposed in 2011 [17], circRNAs are considered to compete with mRNA for the target binding site of miRNA to regulate multiple malignancies [18]. Previous studies have demonstrated that several circRNAs participate in the carcinogenesis and development of TNBC [19, 20].

Although the important role of non-coding RNAs exists in TNBC, the systematic analysis of their expression profile and function is still rare. The aim of this study was to screen different expression profiles using RNA-sequencing of a lncRNA library and establish the interrelation between non-coding RNAs and mRNAs. The key molecules were further analyzed for exploration of their function and potential marker effects in TNBC.

Materials and methods

Clinical specimen acquisition

Four pairs of TNBC tissues and corresponding adjacent tissues from patients who underwent surgical treatment in the First Affiliated Hospital of Bengbu Medical College (China) were collected for this study. The patients did not receive any treatment before the surgery, and they all signed the informed consent. The Ethics Committee of Bengbu Medical College approved this study.

Preparation of sequencing library and circRNA sequencing

Total RNA was extracted from TNBC and the adjacent tissues, the chain-specific library was constructed by the method of rRNA depletion. After the library passed a quality check, Illumina Novaseq™ 6000 was used for RNA sequencing, and the paired-end read length was 2*150 bp (PE150) [21]. We obtained the clean data after filtering the unqualified sequences, which were used for further statistical analysis.

Gene expression profile research

The latest transcript assembly software StringTie was used to assemble and quantify the read out, and R package edgeR was used for the analysis of different statistics and for illustration. Moreover, the abundance value of gene expression was calculated by FPKM (Fragments Per Kilobase of exon model per Million mapped reads), from which we calculated the expressed abundance for each gene in different samples.

Analysis of differentially expressed genes (DEGs)

The DEGs were obtained by edgeR analysis, and the screening conditions were absolute value of log2 fold change greater than 1 and *p* value less than 0.05. Importantly, the different threshold was defined according to the results of the initial operation.

Cluster analysis of DEGs

The clustering analysis of gene regulation patterns under different experimental conditions was defined as differential gene cluster analysis. Summarizing the same points of sample gene expression profiles for gene cluster analysis can directly display the expression levels of genes in different samples, and then obtain relevant information. Log10 (FPKM+1) was used for exhibition of the gene value.

Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis of DEGs

As an internationally standardized gene function classification system, GO describes the molecular function, cell composition and biological process of the DEGs. In the meantime,

Table 1. Gene expression distribution in triple-negative breast cancer tissues and adjacent tissue samples

Sample	Exp gene	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Sd.	Sum.
CT_1	74855	0.00	0.00	0.31	2.53	1.54	15696.66	90.51	189165.55
CT_2	74855	0.00	0.00	0.39	3.45	1.54	41097.28	186.47	258037.31
CT_3	74855	0.00	0.00	0.19	3.85	1.22	41724.94	214.09	286441.20
CT_4	74855	0.00	0.00	0.23	4.02	1.33	61065.60	266.09	301013.33
NT_1	74855	0.00	0.00	0.26	2.18	1.42	13033.19	65.12	163111.13
NT_2	74855	0.00	0.00	0.34	2.53	1.55	23007.14	100.65	189606.81
NT_3	74855	0.00	0.00	0.21	2.60	1.15	23555.05	112.93	194271.78
NT_4	74855	0.00	0.00	0.31	2.89	1.52	23053.36	122.89	216599.42

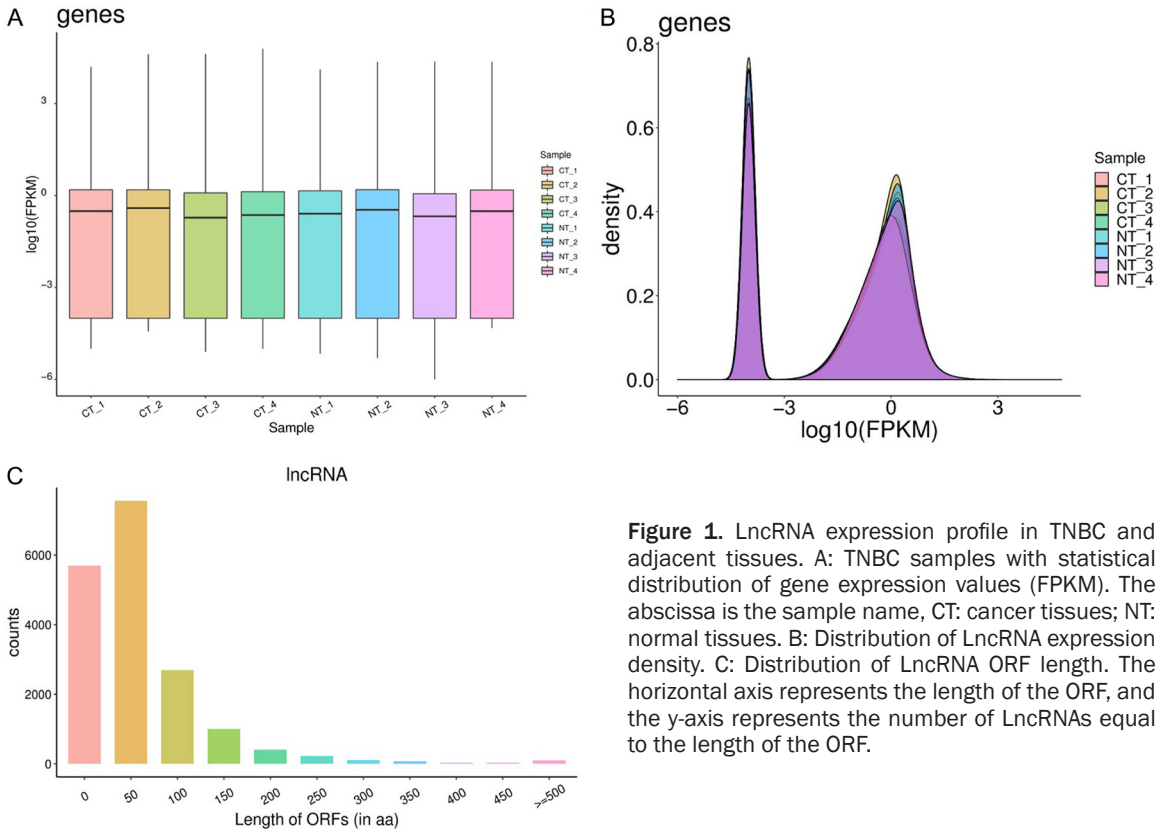

Figure 1. LncRNA expression profile in TNBC and adjacent tissues. A: TNBC samples with statistical distribution of gene expression values (FPKM). The abscissa is the sample name, CT: cancer tissues; NT: normal tissues. B: Distribution of LncRNA expression density. C: Distribution of LncRNA ORF length. The horizontal axis represents the length of the ORF, and the y-axis represents the number of LncRNAs equal to the length of the ORF.

we further investigated the biological function through the KEGG database, which could enrich the DEGs into multiple signal pathways. R ggplot2 was used to analyze the enrichment results as a scatter diagram.

Co-expression analysis of differentially expressed lncRNAs and genes

In order to associate functional genes to each DEGs, co-expression of the top 100 differentially expressed lncRNAs and mRNAs were performed with the criterion of Pearson correlation coefficient larger than 0.99 by WGCNA database, followed by construction of integrated networks using Cytoscape 3.2.0 (http://cytoscape.org/). The hub nodes were identified based on the degree and betweenness centrality. Finally, the hub lncRNAs and mRNAs in this network were functionally analyzed by the DAVID database.

Results

The overall gene expression level of TNBC

A total of 74,855 genes were identified in TNBC and adjacent tissues, and the gene expression

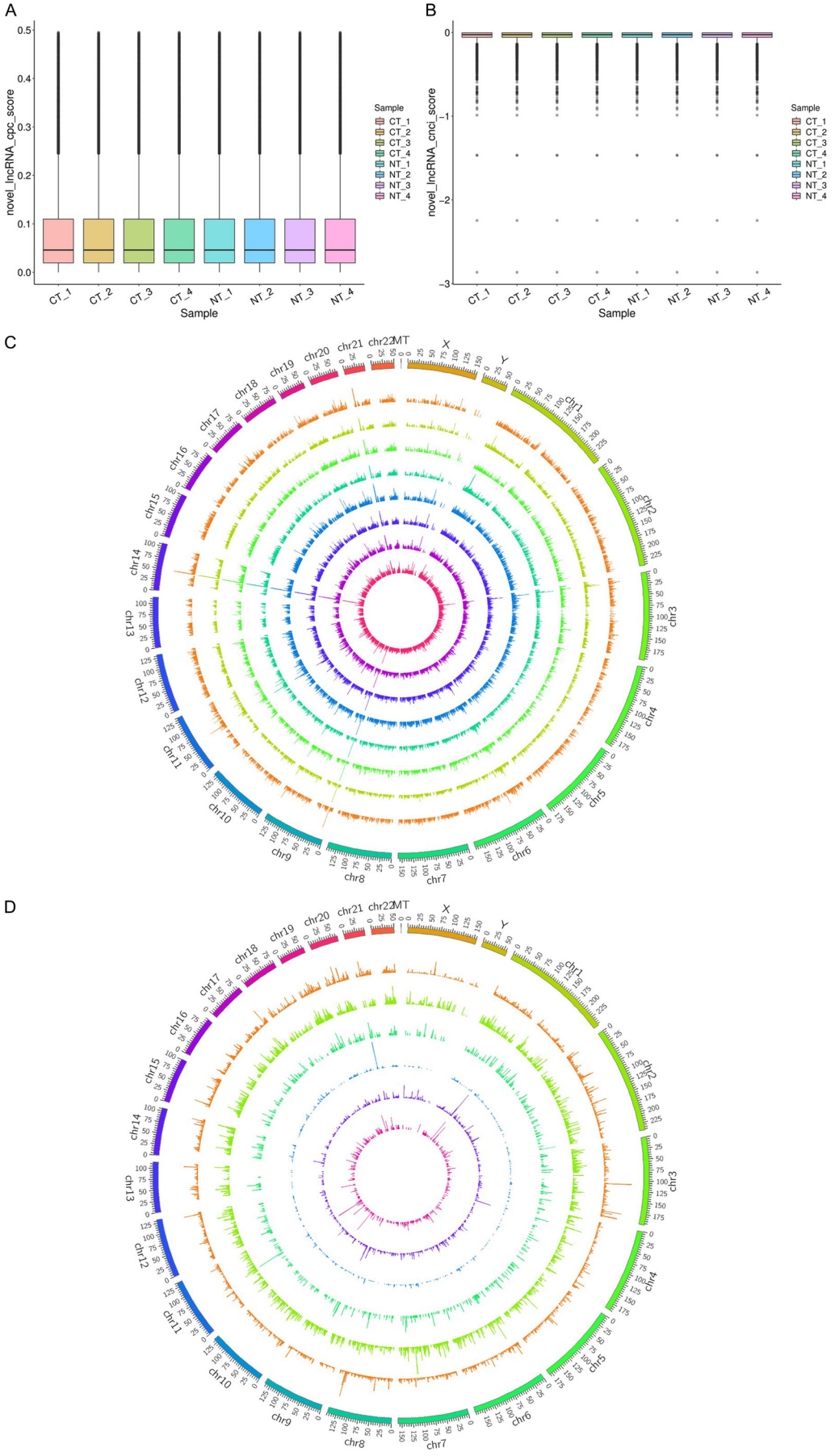
A
novel_lncRNA_cpc_score
Sample
CT_1
CT_2
CT_3
CT_4
NT_1
NT_2
NT_3
NT_4
B
novel_lncRNA_cncl_score
C
D

Figure 2. Analysis of the overall expressed level of lncRNA in TNBC. A: Score of lncRNA CPC in each sample, CPC: Coding Potential Calculator. B: Score of lncRNA CNCI in each sample, CNCI: Coding-Non-Coding Index. C: Visualization of lncRNA genome mapping in different CT samples. D: Visualization of lncRNA genome mapping in different NT samples. Every 25 MB of each chromosome was taken as a basic unit, the level of lncRNAs in each segment of genome were counted and mapped in different samples, and the mapping number of lncRNAs was counted in different parts of the lncRNA genome. The detected lncRNAs were represented by the larger inner green ring, and the smaller inner loops indicated differentially expressed lncRNAs folding changes >2 and *P* value <0.05. The up and down regulation of lncRNAs have been marked with red and blue bars.

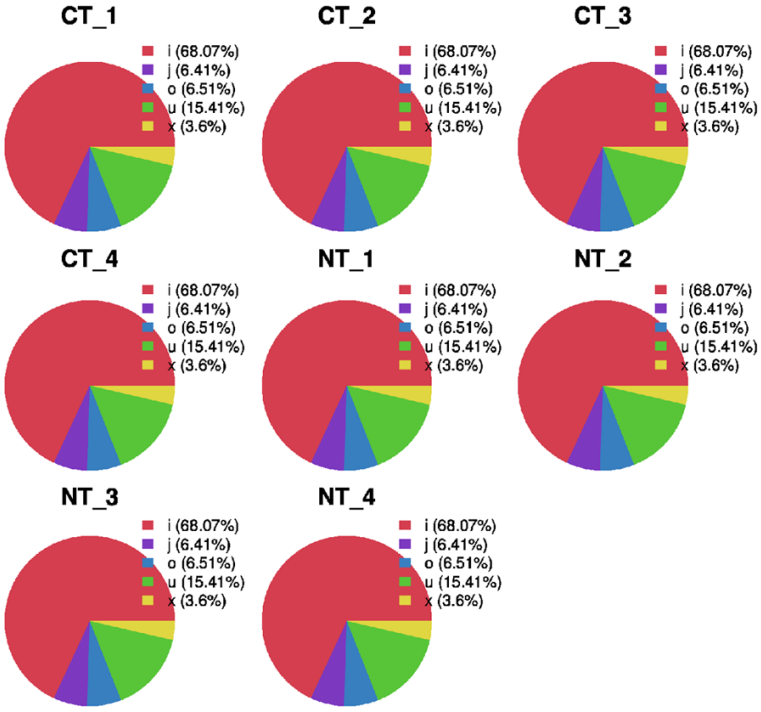

Figure 3. Pie chart illustration the ratios of lncRNA class code. j: Potentially novel isoform (fragment): at least one splice junction is shared with a reference transcript; i: A transfrag falling entirely within a reference intron; o: Generic exonic overlap with a reference transcript; u: Unknown, intergenic transcript; x: Exonic overlap with reference on the opposite strand.

of each sample was shown as the FPKM value (**Table 1**; **Figure 1A**). We further compared the trends of lncRNA expression in each sample, which is shown on the density distribution graph (**Figure 1B**). Lastly, the statistical histogram was used to show the length distribution of the open reading frame (ORF) region of lncRNAs (**Figure 1C**), suggesting that short fragments (<300) had the large number of counts.

Transcripts assembling and lncRNA screening

We used the String software to assemble the Reads. After removing the known mRNA and transcripts smaller than 200 bp, lncRNAs were predicted in the residual transcripts by CPC (Coding Potential Calculator) and CNCI (Coding-Non-Coding Index) software. In the residual transcripts, novel mRNAs that had the encoding functions were also filtered and we eventually harvested the lncRNAs sequences. Statistical calculations of the predicted scores of lncRNAs in samples were performed by CPC and CNCI, as shown in **Figure 2A** and **2B**. Furthermore, ciros (ww.circos.ca) was used to map the lncRNAs, showing the distribution of lncRNA candidates in the chromosome (**Figure 2C**, **2D**). Lastly, the ratio of lncRNA's class code in each sample was showed as the pie chart (**Figure 3**).

Differential expression of lncRNAs in TNBC

A total of 1001 differentially expressed lncRNAs were obtained, which were constituted of 555 up-regulated and 446 down-regulated lncRNAs (**Figure 4A**, **4B**), with the condition of LOG|FC| >1, *P* value <0.05. In the meantime, according to the degree of similarity in the gene expression profiles, clustering analysis of the top 100 differentially expressed lncRNAs were shown using a heat map. The value of Log10 (FPKM+1) was used for intuitively reflecting the clustering pattern (**Figure 4C**).

GO and KEGG analysis of DEGs

Scatter plot was used to show the results of the GO enrichment based on the calculation of q value (**Figure 5A**). The results showed that the differentially expressed lncRNA-target genes mainly enriched in N-acyltransferase and CCR chemokine receptor binding. KEGG analysis showed that these lncRNA-target genes were

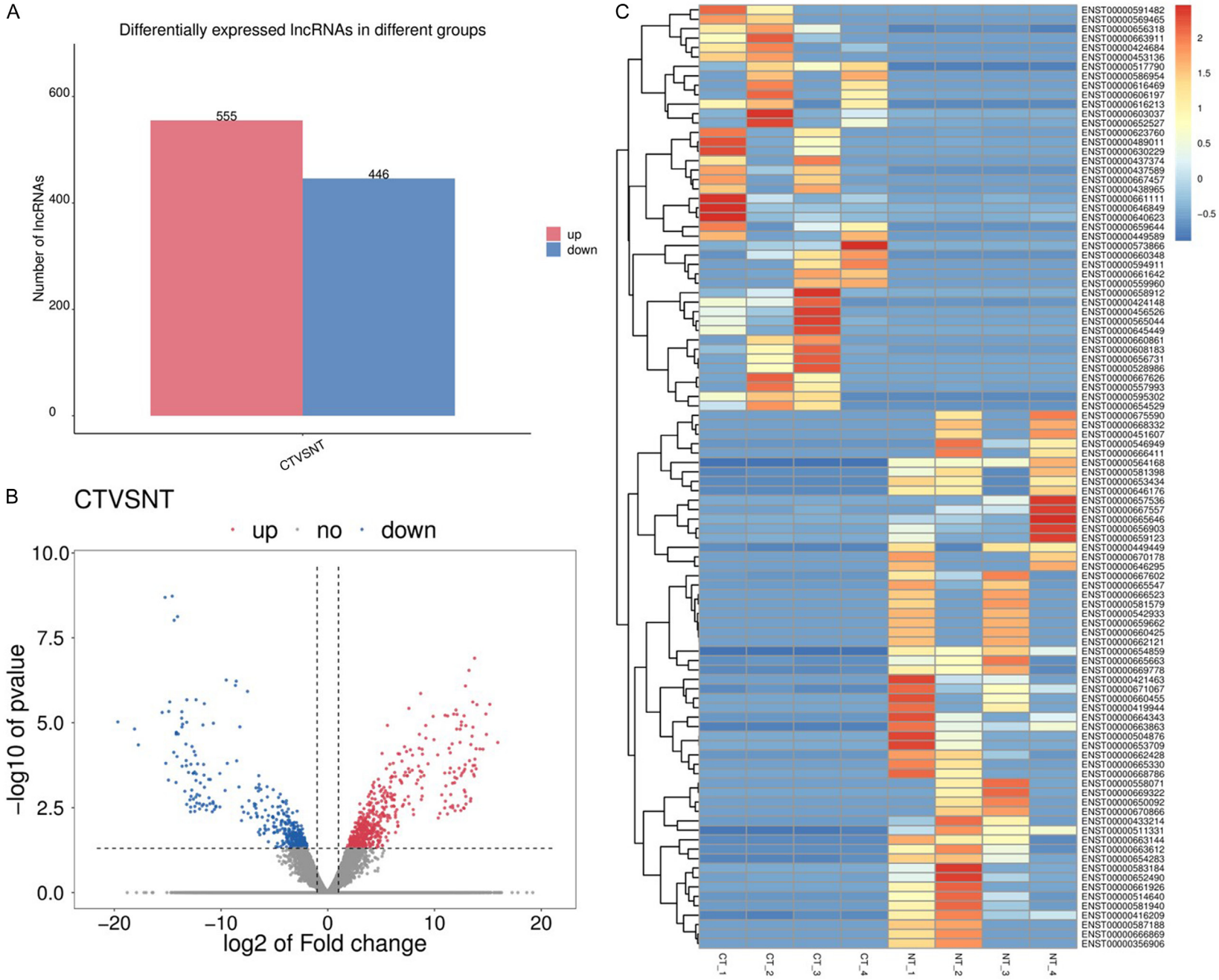
A
Differentially expressed lncRNAs in different groups
Number of lncRNAs
600
400
200
0
555
446
CTVSNT
up
down
B
CTVSNT
up
no
down
10.0
7.5
5.0
2.5
0.0
−log10 of pvalue
−20
−10
0
10
20
log2 of Fold change
C
CT_1
CT_2
CT_3
CT_4
NT_1
NT_2
NT_3
NT_4
2
1.5
1
0.5
0
−0.5

Figure 4. The differentially expressed lncRNAs in TNBC. A: Histogram showing the number of up and down expressed lncRNAs in TNBC. B: Volcano plot showing the up and down expressed lncRNAs in TNBC. The x-axis represents the normalized difference (TNBC group/normal group). The y-axis represents the normalized *P* value. C: Heat map showing the clustering analysis of the differentially expressed lncRNAs in TNBC. Different colors represent different gene expression levels, and the color range from blue to white to red represents the expressed level (log10 (FPKM+1)) from low to high.

mainly enriched in complement and coagulation cascade pathways and caffeine metabolism (**Figure 5B**).

Comparison of lncRNA and mRNA structural characteristics

We know that the structural characteristics (including length distribution, number of exons) and expressed levels of lncRNA and mRNA are quite different. Hence, we mainly focused on the comparative analysis of the structural characteristics and expression levels of them. We found that the number of exons in mRNAs were significantly higher than lncRNAs (**Figure 6A**), and most of lncRNAs (75%) had a number of exons less than 3. The expression levels of lncRNAs and mRNAs in TNBC are showed in **Figure 6B**, **6C**, which suggested that the mRNA levels were slightly higher than lncRNA.

Interaction analysis of lncRNAs and mRNAs

The top 100 differentially expressed lncRNAs and mRNAs were included in WGCNA, and the co-expression networks were established (**Figure 7**). In this network, we found OIP5-AS1 was co-expressed with MATN3, BGN, FN1, P4HA3; BPIFB1, NUN5A, and COL1DA1 was co-expressed with ACD447B4; STMND1 and ABCC11/12 were co-expressed with AC1037-18, AL513327 and ACD93001. We selected hub RNAs because of their higher degree and betweenness centrality, including lncRNAs OIP5-AS1, DRAIC, AC093904, AC044784, LINC02747 and mRNAs UNC5A, BPIFB1, MBOAT2, PITX1, STMND1, KLK4. In addition, functional analysis of hub lncRNAs indicated their roles in cancer-related process, such as cell proliferation and metastasis. The hub mRNAs were mainly enriched in neuron projection development and glycerolipid metabolic pathways.

Screening of key circRNA-hosting genes

We found that several types of circRNA existed in TNBC samples, including all-exon, lasso-type composed of introns and intergenic region types (**Figure 8**). Second, according the screening threshold with LOG|FC| >1, and *P* value <0.05, we identified a total of 139 differently expressed circRNAs, of which 26 were up-regulated and 113 were down-regulated in TNBC (**Figure 9A**, **9B**). The cluster analysis was performed (**Figure 9C**) by heat map according the Log10 (FPKM+1) value. Furthermore, GO and KEGG analysis showed that the differentially expressed circRNA-hosting genes enriched in ubiquitin conjugating enzyme binding and cerebral cortex development through GO analysis, and were mainly enriched in AMPK and TGF-β signaling pathway (**Figure 10**). Eventually, we selected five key hosting genes of circRNAs, including NTRK2, FNTA, BAPGEF2, MGST2, ADH1B, and their target circRNAs, such as circRNA11248, circRNA13256, circRNA10523, were associated with the BDNF receptor signaling and ubiquitin conjugating enzyme binding, as well as the AMPK and TGF-β pathway.

Discussion

The investigation of non-coding RNAs and their functions have become new hotpots in cancer research. As we know, the pathogenesis of TNBC is complex and not fully recognized. It has been demonstrated that various molecules and factors could result in the occurrence of TNBC [22, 23], and ncRNAs play crucial roles in regulating the initiation and progression of cancers [24]. Recently, the differential expression of lncRNAs and circRNAs were comprehensively analyzed based on the RNA-sequencing technology, which has been extensively applied to identify the targets in multiple diseases.

In this study, we discovered the differentially expressed profile of lncRNAs and circRNAs between TNBC and adjacent tissues. For lncRNAs, 1001 differently expressed lncRNAs (555 up-regulated and 446 down-regulated) were involved in TNBC. With the aim to assess the relationship between lncRNAs and mRNAs, we established the co-expression network and further screened the hub RNAs. We showed that the hub lncRNAs OIP5-AS1, DRAIC,

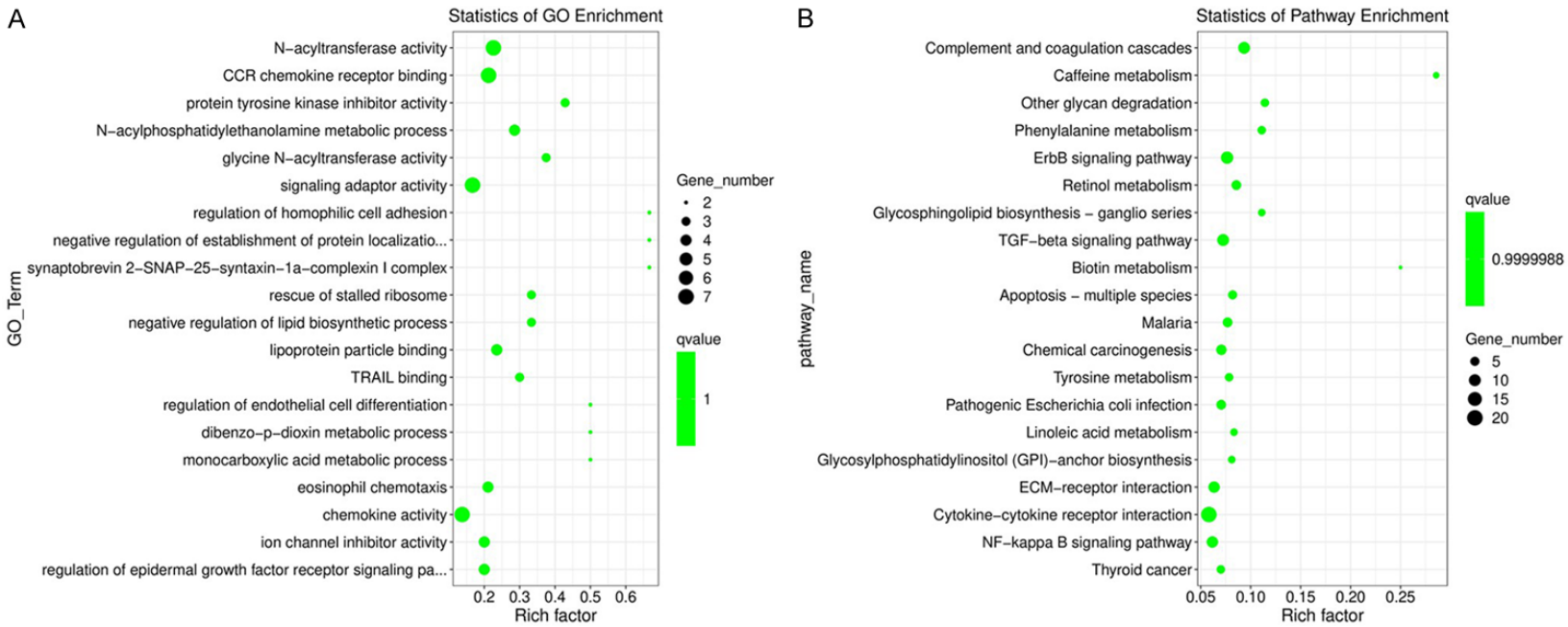

Figure 5. Enrichment analysis of differentially expressed lncRNA-target genes. A: Scatter plot of GO enrichment of differentially expressed genes in TNBC. Horizontal axis that rich factor represents the number of differential genes located in GO/Total number of genes located in GO. The ordinate is the comment of GO function. The size of the dot represents the number of differentially expressed genes, and the color of the dot represents the q value. B: KEGG enrichment analysis of the differentially expressed genes in TNBC. Horizontal axis is the rich factor that represents the number of differential genes located in KEGG/Total number of genes in KEGG. The ordinate is showing the pathways related to genes.

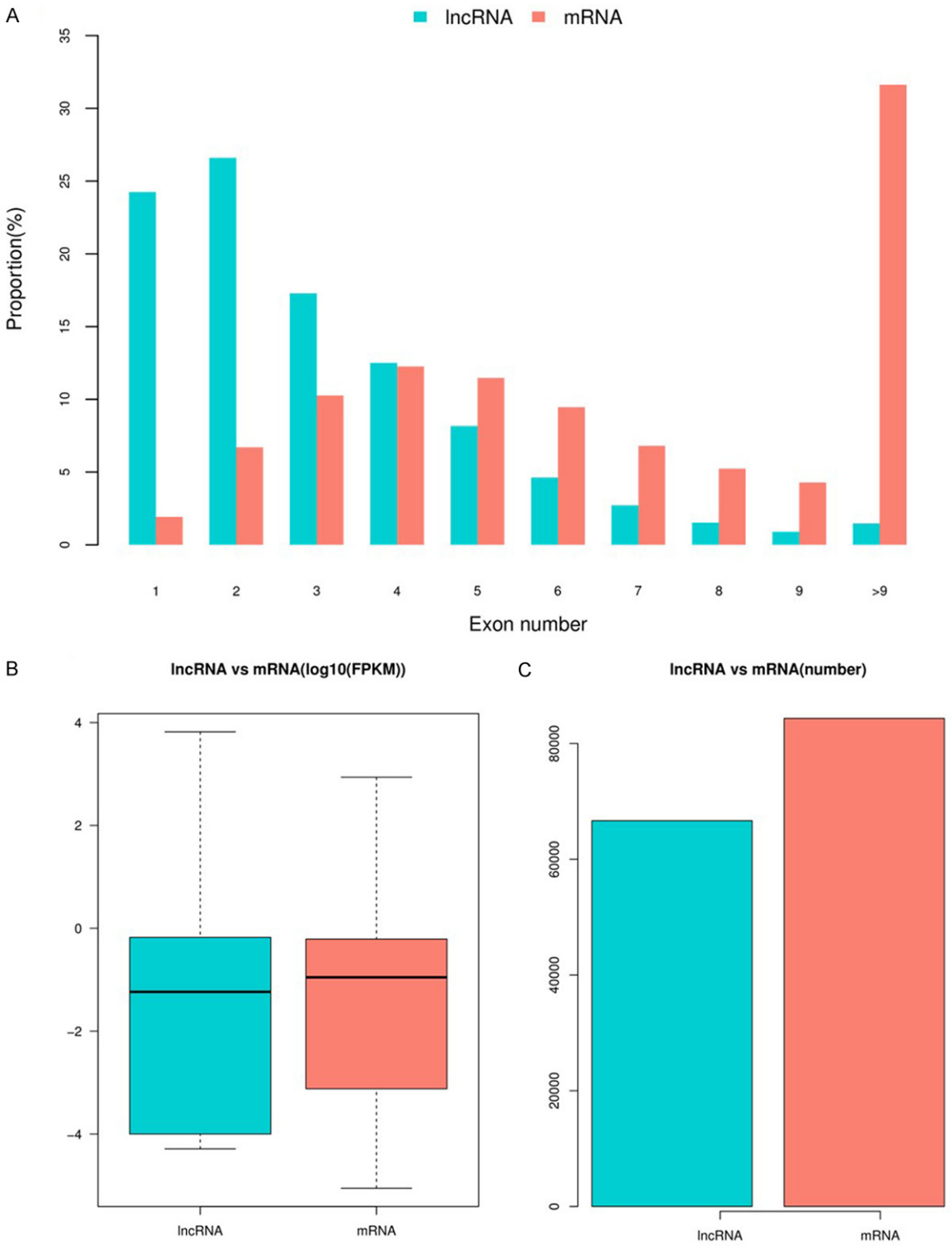

Figure 6. Comparison of lncRNA and mRNA structural characteristics. A: Statistics of the number of lncRNA and mRNA exons in TNBC. B, C: Compassion between the expressed levels of lncRNA and mRNA in TNBC according FPKM value and number.

AC093904, AC044784, LINC02747 were all up-regulated in TNBC, and the hub mRNAs UNC5A, BPIFB1, MBOAT2, PITX1, STMND1, KLK4 were up-regulated while CADM2 was

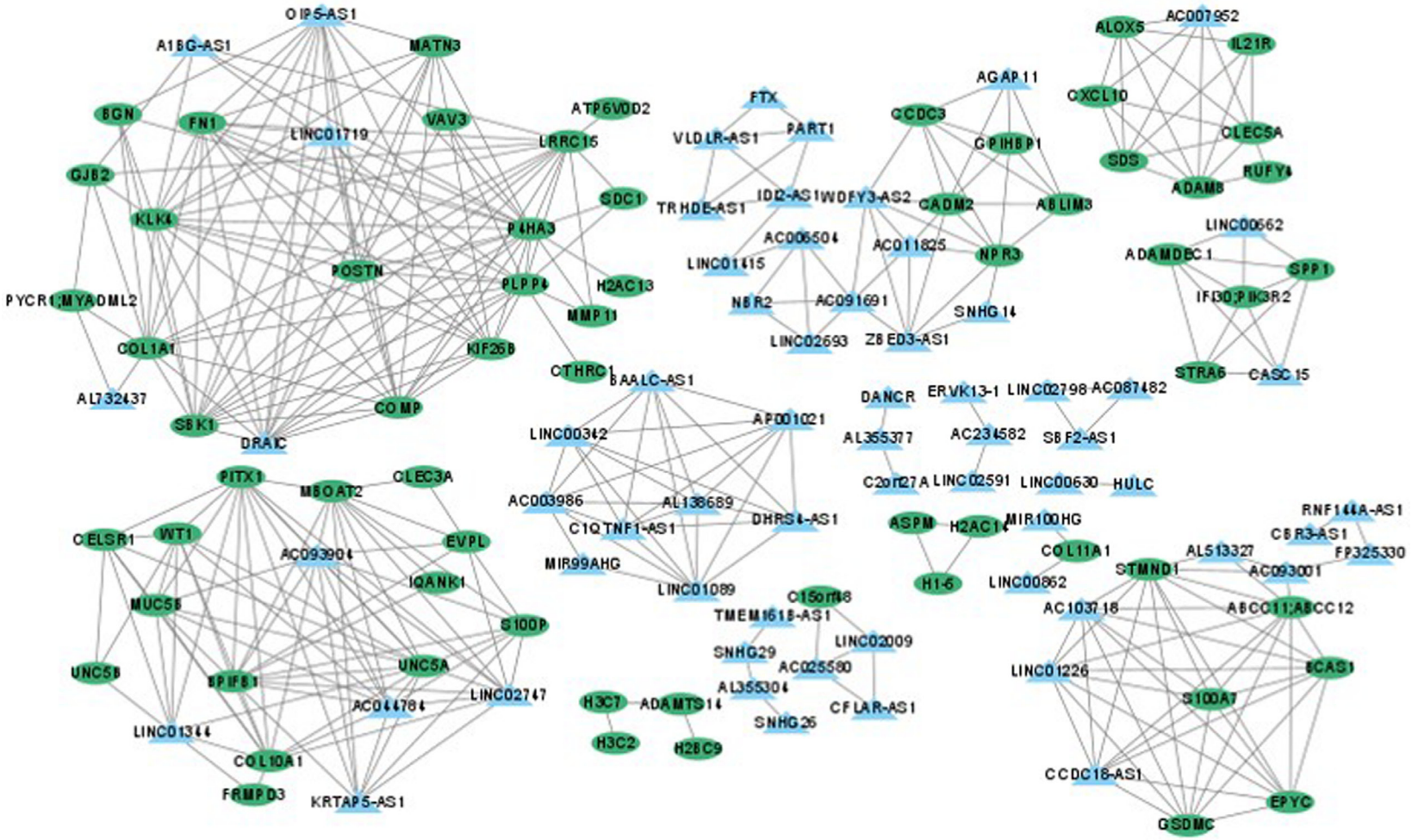

Figure 7. The co-expression network between lncRNAs and mRNAs. The oval represents lncRNAs while the triangle represents the mRNAs.

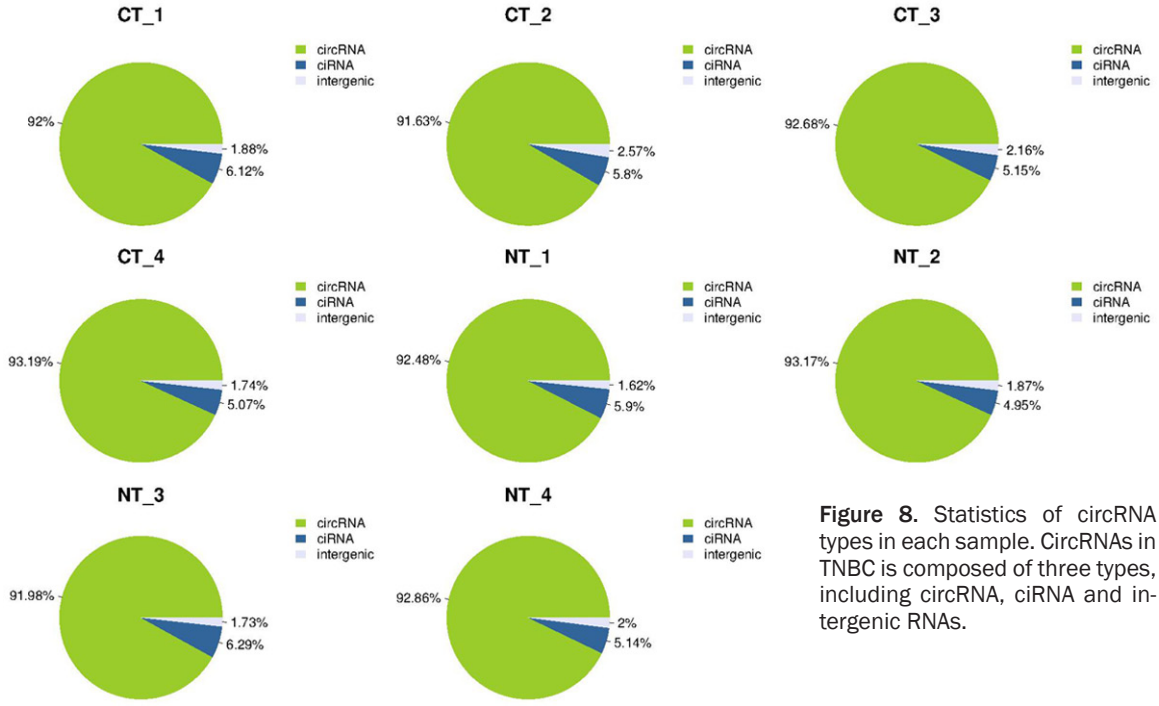

Figure 8. Statistics of circRNA types in each sample. CircRNAs in TNBC is composed of three types, including circRNA, ciRNA and intergenic RNAs.

down-regulated in TNBC. Moreover, enrichment analysis showed that the hub lncRNAs OIP5-AS1 and DRAIC were associated with various cancers and involved in cancer metastasis, while hub mRNAs were mainly enriched in neuron projection development and the glycerolipid metabolic pathway. Several reports have discussed that lncRNA OIP5-AS1 plays multiple roles in different diseases. For example, Xiaowei Niu et al. [25] found that OIP5-AS1 could attenuate myocardial ischaemia/reperfusion injury through the ceRNA mechanism, but it accelerated intervertebral disc degeneration by Zhaoping Che's report [26]. In cancer research, OIP5-AS1 is overexpressed in various tumors [27], including breast cancer. We provided that lncRNA OIP5-AS1 was up-reregulated in TNBC, which is consistent with the consensus. For another, DRAIC showed differentially expressed trends in distinct cancers [28, 29] but was still overexpressed in breast cancer [30]. Our study demonstrated that the expression of lncRNAs seems analogous between breast cancer and TNBC.

We noticed that most of the hub mRNAs were enriched in pathways relative to neuron-related function, which seems amazing. In fact, neurogenesis has been associated with cancer research a long time ago, which aimed to identify functional roles played by the neural system in cancer [31]. Perineural invasion has been observed in multiple cancer types [32], and cancer cells surrounding nerves tend to be more resistant to apoptosis [33]. However, the function of the neural system in the formation of breast cancer, particularly in TNBC, still remains elusive. Renbo Tan's study [34] compared the difference of neural functions between TNBC and non-TNBC, and discovered that neural crest formation is associated with enhancing adaptive immunity in TNBC. We showed that TNBC might be regulated by genes related to neurogenesis, which is a new research orientation for the therapy of TNBC.

Recently, circRNAs seem more prominent due to their specific characterizations, such as high stability, time- and tissue-specificity [15]. Such noncoding RNAs are produced by a non-canonical splicing event known as backsplicing [35]. Hence, another aim of our study was to screen the key circRNAs in TNBC. We observed 139 circRNAs that were differentially expressed in TNBC tissues, of which 26 circRNAs showed a significant up-regulation trend, and 113 circRNAs showed a significant down-regulation trend. The GO analysis showed that the circRNA-hosting genes were mainly enriched in brain-derived neurotrophic factor receptor sig-

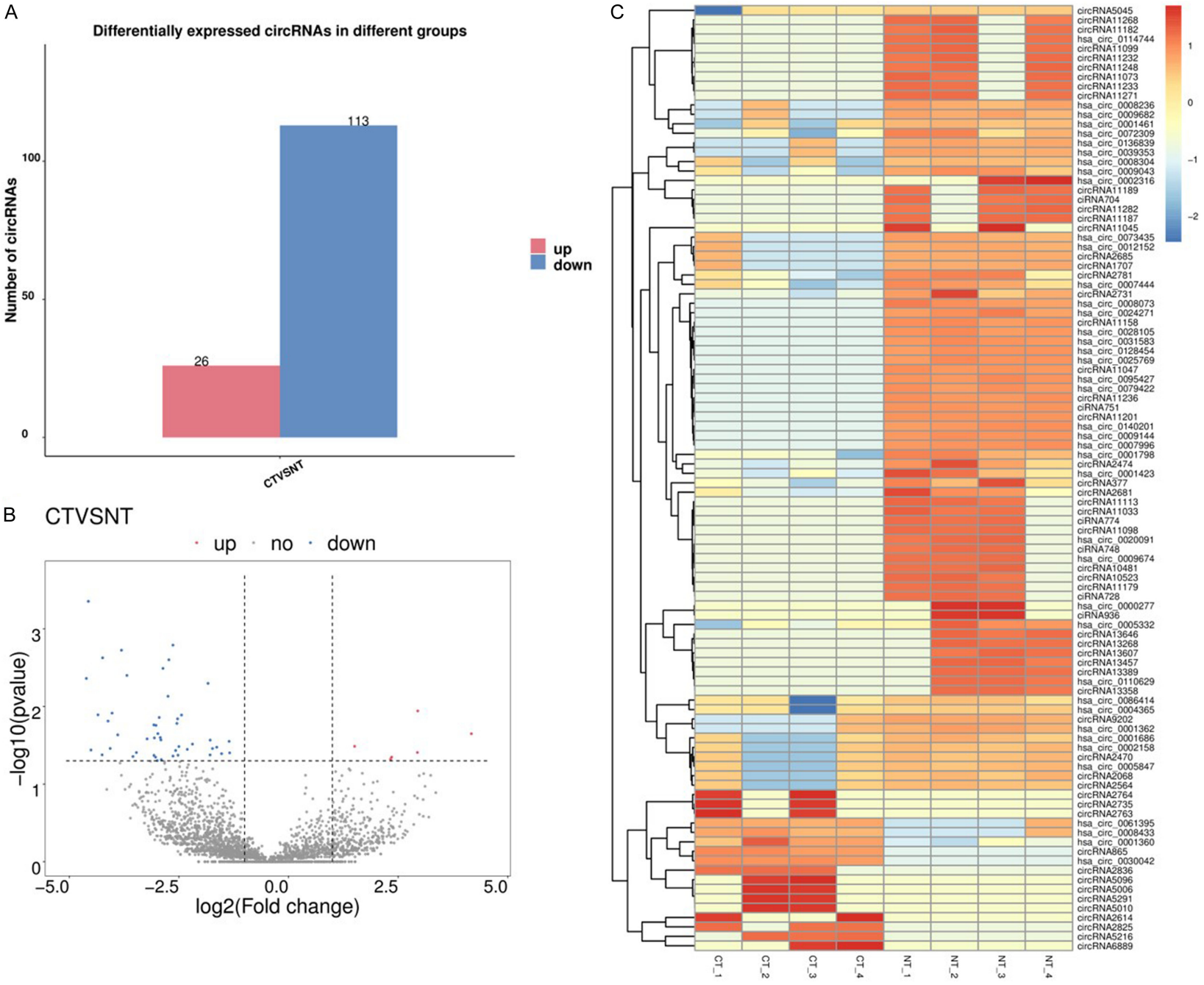
A
Differentially expressed circRNAs in different groups
Number of circRNAs
100
50
0
113
26
CTVSNT
up
down
B
CTVSNT
up
no
down
−log10(pvalue)
3
2
1
0
−5.0
−2.5
0.0
2.5
5.0
log2(Fold change)
C
circRNA5045
circRNA11268
circRNA11182
hsa_circ_0114744
circRNA11099
circRNA11232
circRNA11248
circRNA11073
circRNA11233
circRNA11271
hsa_circ_0008236
hsa_circ_0009682
hsa_circ_0001461
hsa_circ_0072309
hsa_circ_0136839
hsa_circ_0039353
hsa_circ_0008304
hsa_circ_0009043
hsa_circ_0002316
circRNA11189
ciRNA704
circRNA11282
circRNA11187
circRNA11045
hsa_circ_0073435
hsa_circ_0012152
circRNA2685
circRNA1707
circRNA2781
hsa_circ_0007444
circRNA2731
hsa_circ_0008073
hsa_circ_0024271
circRNA11158
hsa_circ_0028105
hsa_circ_0031583
hsa_circ_0128454
hsa_circ_0025769
circRNA11047
hsa_circ_0095427
hsa_circ_0079422
circRNA11236
ciRNA751
circRNA11201
hsa_circ_0140201
hsa_circ_0009144
hsa_circ_0007996
hsa_circ_0001798
circRNA2474
hsa_circ_0001423
circRNA377
circRNA2681
circRNA11113
circRNA11033
ciRNA774
circRNA11098
hsa_circ_0020091
ciRNA748
hsa_circ_0009674
circRNA10481
circRNA10523
circRNA11179
ciRNA728
hsa_circ_0000277
ciRNA936
hsa_circ_0005332
circRNA13646
circRNA13268
circRNA13607
circRNA13457
circRNA13389
hsa_circ_0110629
circRNA13358
hsa_circ_0086414
hsa_circ_0004365
circRNA9202
hsa_circ_0001362
hsa_circ_0001686
hsa_circ_0002158
circRNA2470
hsa_circ_0005847
circRNA2068
circRNA2564
circRNA2764
circRNA2735
circRNA2763
hsa_circ_0061395
hsa_circ_0008433
hsa_circ_0001360
circRNA865
hsa_circ_0030042
circRNA2836
circRNA5096
circRNA5006
circRNA5291
circRNA5010
circRNA2614
circRNA2825
circRNA5216
circRNA6889
CT_1
CT_2
CT_3
CT_4
NT_1
NT_2
NT_3
NT_4
1
0
−1
−2

Figure 9. The differentially expressed circRNAs in TNBC. A: Histogram showing the number of up and down expressed circRNAs in TNBC. B: Volcano plot showing the up and down expressed circRNAs in TNBC. The x-axis represents the normalized difference (TNBC group/normal group). The y-axis represents the normalized *P* value. C: Heat map showing the clustering analysis of the differentially expressed circRNAs in TNBC. Different colors represent different gene expression levels, and the color range from blue to white to red represents the expressed level (log10 (FPKM+1)) from low to high.

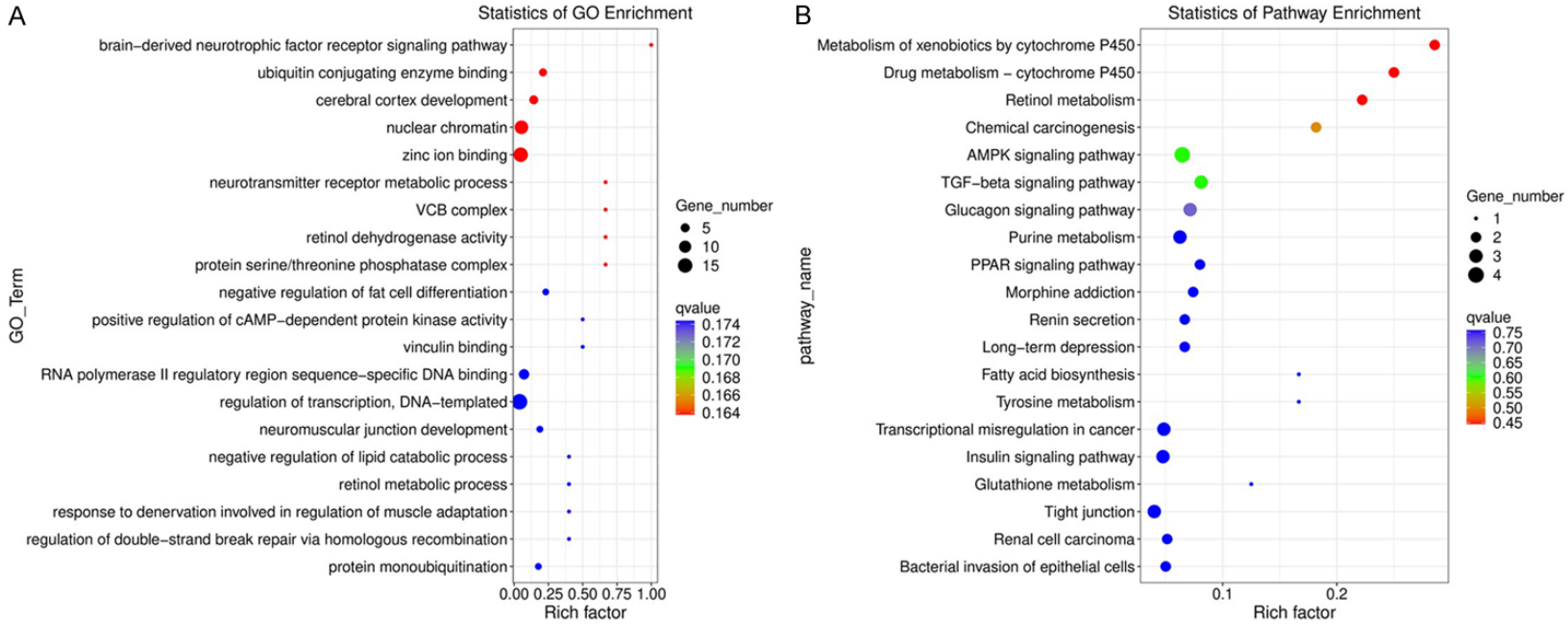

Figure 10. Enrichment analysis of differentially expressed circRNA-hosting genes in TNBC. A: Scatter plot of GO enrichment of differentially expressed genes in TNBC. Horizontal axis that rich factor represents the number of differential genes located in GO/Total number of genes located in GO. The ordinate is the comment of GO function. The size of the dot represents the number of differentially expressed genes, and the color of the dot represents the q value. B: KEGG enrichment analysis of the differentially expressed genes in TNBC. Horizontal axis that rich factor represents the number of differential genes located in KEGG/Total number of genes in KEGG. The ordinate showing the pathways related to genes.

naling pathways and cerebral cortex development, while AMPK and TGF-β signaling pathway was seen in KEGG analysis. Hence, the relationship between neurological function and TNBC are mentioned again from the enrichment of circRNA. TGF-β signaling pathway is implicated in the dysregulation of carcinogenesis, and the elevated TGF-β activity is associated with poor clinical outcome [36]. Importantly, cell autophagy driven by TGF-β further initiates cell proliferation and apoptosis. Moreover, as an energy sensor, AMPK regulates protein and lipid metabolism according to the alterations of energy supply. It has been proven that the AMPK signaling pathway play an important role in regulating the growth and the development of drug resistance in TNBC cells [37]. The new circRNAs and their hosting genes screened in this study might be the potential targets for the drug therapy.

In summary, this study provided an expression profile of circRNA and lncRNA in TNBC using RNA-sequencing, and further analyzed the co-expression interrelation between lncRNAs and mRNAs. The new circRNA-hosting genes could be used as potential markers for the patients with TNBC. GO and KEGG were used to annotate the function of differentially expressed or hub RNAs for the TNBC. However, the detailed function and mechanism of these hub lncRNAs and mRNAs, as well as key circRNAs, need to be further studied with *in vitro* and *in vivo* experiments.

Acknowledgements

This work was supported by funding for the Study on the role and mechanism of circular rna-csmarca5 in triple negative breast cancer, KJ2020A0577.

Disclosure of conflict of interest

None.

Address correspondence to: Tingjing Yao, Oncology Surgery, The First Affiliated Hospital of Bengbu Medical College, 287 Changhuai Road, Bengbu 233000, Anhui, China. Tel: +86-13855200468; E-mail: 13855200468@163.com

References

[1] Harbeck N and Gnant M. Breast cancer. Lancet 2017; 389: 1134-1150.

[2] Reis-Filho JS and Lakhani SR. Breast cancer special types: why bother? J Pathol 2008; 216: 394-8.

[3] Bauer KR, Brown M, Cress RD, Parise CA and Caggiano V. Descriptive analysis of estrogen receptor (ER)-negative, progesterone receptor (PR)-negative, and HER2-negative invasive breast cancer, the so-called triple-negative phenotype: a population-based study from the California cancer registry. Cancer 2007; 109: 1721-8.

[4] Fornier M and Fumoleau P. The paradox of triple negative breast cancer: novel approaches to treatment. Breast J 2012; 18: 41-51.

[5] Nakagawa H and Fujita M. Whole genome sequencing analysis for cancer genomics and precision medicine. Cancer Sci 2018; 109: 513-522.

[6] Coussy F, Lavigne M, De Koning L, Botty RE, Nemati F, Naguez A, Bataillon G, Ouine B, Dahmani A, Montaudon E, Painsec P, Chateau-Joubert S, Laetitia F, Larcher T, Vacher S, Chemlali W, Briaux A, Melaabi S, Salomon AV, Guinebretiere JM, Bieche I and Marangoni E. Response to mTOR and PI3K inhibitors in enzalutamide-resistant luminal androgen receptor triple-negative breast cancer patient-derived xenografts. Theranostics 2020; 10: 1531-1543.

[7] Tariq H, Gul A, Zubair M, Jaffer SR, Zafar N and Sadaf G. Targeted next-generation sequencing for reliable detection of genetic status in breast cancer. J Coll Physicians Surg Pak 2021; 30: 837-840.

[8] Liu Y, Liu X, Lin C, Jia X, Zhu H, Song J and Zhang Y. Noncoding RNAs regulate alternative splicing in cancer. J Exp Clin Cancer Res 2021; 40: 11.

[9] Quinn JJ and Chang HY. Unique features of long non-coding RNA biogenesis and function. Nat Rev Genet 2016; 17: 47-62.

[10] Chaudhary LN, Wilkinson KH and Kong A. Triple-negative breast cancer: who should receive neoadjuvant chemotherapy? Surg Oncol Clin N Am 2018; 27: 141-153.

[11] Kader YA, El-Nahas T and Sakr A. Adjuvant chemotherapy for luminal A breast cancer: a prospective study comparing two popular chemotherapy regimens. Onco Targets Ther 2013; 6: 1073-7.

[12] Huang L, Liu Q, Chen S and Shao Z. Cisplatin versus carboplatin in combination with paclitaxel as neoadjuvant regimen for triple negative breast cancer. Onco Targets Ther 2017; 10: 5739-5744.

[13] Dykes IM and Emanueli C. Transcriptional and post-transcriptional gene regulation by long non-coding RNA. Genomics Proteomics Bioinformatics 2017; 15: 177-186.

[14] Ransohoff JD, Wei Y and Khavari PA. The functions and unique features of long intergenic

non-coding RNA. Nat Rev Mol Cell Biol 2018; 19: 143-157.
[15] Vo JN, Cieslik M, Zhang Y, Shukla S, Xiao L, Zhang Y, Wu YM, Dhanasekaran SM, Engelke CG, Cao X, Robinson DR, Nesvizhskii AI and Chinnaiyan AM. The landscape of circular RNA in cancer. Cell 2019; 176: 869-881, e13.
[16] Galupa R and Heard E. X-chromosome inactivation: new insights into cis and trans regulation. Curr Opin Genet Dev 2015; 31: 57-66.
[17] Salmena L, Poliseno L, Tay Y, Kats L and Pandolfi PP. A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language? Cell 2011; 146: 353-8.
[18] Guan YJ, Ma JY and Song W. Identification of circRNA-miRNA-mRNA regulatory network in gastric cancer by analysis of microarray data. Cancer Cell Int 2019; 19: 183.
[19] Zheng X, Huang M, Xing L, Yang R, Wang X, Jiang R, Zhang L and Chen J. The circRNA circSEPT9 mediated by E2F1 and EIF4A3 facilitates the carcinogenesis and development of triple-negative breast cancer. Mol Cancer 2020; 19: 73.
[20] Yang R, Xing L, Zheng X, Sun Y, Wang X and Chen J. The circRNA circAGFG1 acts as a sponge of miR-195-5p to promote triple-negative breast cancer progression through regulating CCNE1 expression. Mol Cancer 2019; 18: 4.
[21] Morabito C, Aiese Cigliano R, Maréchal E, Rébeillé F and Amato A. Illumina and PacBio DNA sequencing data, de novo assembly and annotation of the genome of Aurantiochytrium limacinum strain CCAP_4062/1. Data Brief 2020; 31: 105729.
[22] Derakhshan F and Reis-Filho JS. Pathogenesis of triple-negative breast cancer. Annu Rev Pathol 2022; 17: 181-204.
[23] Da Costa Vieira RA, Biller G, Uemura G, Ruiz CA and Curado MP. Breast cancer screening in developing countries. Clinics (Sao Paulo) 2017; 72: 244-253.
[24] Yan H and Bu P. Non-coding RNA in cancer. Essays Biochem 2021; 65: 625-639.
[25] Niu X, Pu S, Ling C, Xu J, Wang J, Sun S, Yao Y and Zhang Z. lncRNA Oip5-as1 attenuates myocardial ischaemia/reperfusion injury by sponging miR-29a to activate the SIRT1/AMPK/PGC1α pathway. Cell Prolif 2020; 53: e12818.
[26] Che Z, Xueqin J and Zhang Z. LncRNA OIP5-AS1 accelerates intervertebral disc degeneration by targeting miR-25-3p. Bioengineered 2021; 12: 11201-11212.
[27] Zhong J, Chen J, Wang B, Zhou Z, Shen Y, Gong Y, Tan F and Yuan C. OIP5-AS1: a fascinating long noncoding rna in carcinoma. Curr Pharm Des 2021; 27: 4699-4706.
[28] Liu Z, Yang S, Zhou S, Dong S and Du J. Prognostic value of lncRNA DRAIC and miR-3940-3p in lung adenocarcinoma and their effect on lung adenocarcinoma cell progression. Cancer Manag Res 2021; 13: 8367-8376.
[29] Saha S, Kiran M, Kuscu C, Chatrath A, Wotton D, Mayo MW and Dutta A. Long noncoding RNA DRAIC inhibits prostate cancer progression by interacting with IKK to inhibit NF-κB activation. Cancer Res 2020; 80: 950-963.
[30] Li S, Jia H, Zhang Z and Wu D. DRAIC promotes growth of breast cancer by sponging miR-432-5p to upregulate SLBP. Cancer Gene Ther 2022; 29: 951-960.
[31] Entschladen F, Palm D, Lang K, Drell TL 4th and Zaenker KS. Neoneurogenesis: tumors may initiate their own innervation by the release of neurotrophic factors in analogy to lymphangiogenesis and neoangiogenesis. Med Hypotheses 2006; 67: 33-5.
[32] Kuol N, Stojanovska L, Apostolopoulos V and Nurgali K. Role of the nervous system in cancer metastasis. J Exp Clin Cancer Res 2018; 37: 5.
[33] Dai H, Li R, Wheeler T, Ozen M, Ittmann M, Anderson M, Wang Y, Rowley D, Younes M and Ayala GE. Enhanced survival in perineural invasion of pancreatic cancer: an in vitro approach. Hum Pathol 2007; 38: 299-307.
[34] Tan R, Li H, Huang Z, Zhou Y, Tao M, Gao X and Xu Y. Neural functions play different roles in triple negative breast cancer (TNBC) and non-TNBC. Sci Rep 2020; 10: 3065.
[35] Patop IL, Wüst S and Kadener S. Past, present, and future of circRNAs. EMBO J 2019; 38: e100836.
[36] Syed V. TGF-β signaling in cancer. J Cell Biochem 2016; 117: 1279-87.
[37] Cao W, Li J, Hao Q, Vadgama JV and Wu Y. AMP-activated protein kinase: a potential therapeutic target for triple-negative breast cancer. Breast Cancer Res 2019; 21: 29.