Construction of Polycystic Ovarian Syndrome Related lncRNA-mRNA Network Based on ceRNA to Identify Functional lncRNAs in Polycystic Ovarian Syndrome

Yue Ma
Zhengzhou University https://orcid.org/0000-0003-3960-7574
Jun Zhai (bestzhai2005@163.com)
Zhengzhou University First Affiliated Hospital

Research

Keywords: Polycystic ovary syndrome, lncRNA, ceRNA, Biomarkers

Posted Date: August 17th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-58728/v1

Abstract

Background

Polycystic ovary syndrome (PCOS) is a prevalent endocrine and metabolic disorder in women of childbearing age. Recent studies have shown that long non-coding RNA (lncRNA) played a vital role in the development of the PCOS. Competitive endogenous RNA (ceRNA), a novel interacting mechanism, in which lncRNA could interact with micro-RNAs (miRNA) and indirectly interact with mRNAs through competing interactions. However, the mechanism of ceRNA regulated by lncRNA in the PCOS was unclear.

Results

We constructed the global background network based on the assumed lncRNA-miRNA and miRNA-mRNA pairs, which were obtained from lncRNASNP, miRTarBase and StarBase database. Then we calculated differentially expressed genes of PCOS using the data of GSE95728. PCOS related lncRNA-mRNA network (PCLMN) was constructed by hypergeometric test, including 41 mRNA nodes, 41 lncRNA nodes and 203 edges. Topological analyses was performed to determine the crucial lncRNAs with the highest centroid. We further identified the subcellular localization, performed functional module analyses and identified putative transfer factors of the key lncRNAs. Functional enrichment analyses were performed by GO classification and KEGG pathway analyses. Finally, 3 key lncRNAs(LINC00667, H19, AC073172.1) and their ceRNA sub-networks, which were involved in NF-kB signaling pathway, inflammatory, apoptotic and immune-related processes, had been found as the potential PCOS related disease genes.

Conclusions

Based on the result above, we speculate that LINC00667, H19, AC073172.1 and their ceRNA sub-networks played an crucial role in PCOS. All these results can help us discover the molecular mechanism and offer new predictive biomarkers for PCOS.

Introductionce

Polycystic ovary syndrome (PCOS) is one of the most prevalent endocrine and metabolic disorders, affecting up to one in five women of childbearing age[1]. Its cardinal features are hyperandrogenism , menstrual irregularity and polycystic ovary morphology[2]. PCOS is associated with insulin resistance, metabolic syndrome, increased risk of endometrial cancer, ovulatory dysfunction, infertility, pregnancy complications, type 2 diabetes and cardiovascular disease[3-7]. Although the exact pathogenesis of PCOS is complicated and remains unclear, most researchers hold that genetic factors play a key role in PCOS occurrence and development.

Long non-coding RNAs (lncRNAs) are defined as a class of non-coding transcripts with the length of more than 200 nucleotides[8]. Although lncRNAs lack the capacity of code for proteins, an increasing body of evidence suggested that lncRNAs play key roles in many biological processes, including genetic imprinting, X-chromosome in activation, transcriptional and post-transcriptional regulation, recruitment of epigenetic modifiers, control of mRNA decay, organelle biogenesis, and subcellular trafficking[9, 10]. Accumulating evidence suggests that lncRNAs are dysregulated and closely associated with numerous human diseases including PCOS and PCOSrelated diseases.

For instance, results from microarray data analysis showed that some LncRNAs were abnormally expressed in the cumulus cells of PCOS patients, which may lead to the occurrence of PCOS and the development of oocytes[10]. It was reported that the LncRNA HCG26 participated the regulation of granulosa cells proliferation and steroidogenesis[11]. And the LncRNA CTBP1-AS could enhance androgen receptor (AR) level in vivo by facilitating AR transcription, leading to hyperandrogenemia in patients[12]. However, there were only a few lncRNAs that had been verified related to the PCOS, so the high through put experiments data and bioinformatics methods were preformed to predict the potential PCOS relateddisease lncRNAs.

Competitive endogenous RNA (ceRNA) was a novel regulatory mechanism that had been put forward lately[13], it hypothesizes that lncRNAs can regulate other RNA transcripts by using miRNA response elements. ceRNA has been confirmed working in the development of different diseases, and ceRNA networks have been built in lung cancer[14], cardiac hypertrophy[15], implantation failure[16] and so on.However, the ceRNA mechanism related to PCOS remains unclear. Considering the expected role of LncRNAs in the regulation of PCOS and the function of the ceRNA mechanism, we attempt to construct the PCOS related lncRNA-mRNA network (PCLMN) to identify the key lncRNAs that may serve as biomarkers of PCOS.

Materials And Methods

We constructed the PCLMN and then investigated the functional characteristics of PCLMN. The key lncRNAs associated with PCOS were identified, their related modules were analyzed, the functional enrichments were researched.

miRNA-LncRNA and miRNA-mRNA interaction data

The predicted lncRNA-miRNA interact pairs in our study was obtained from the starbase V2.0[17] and lncRNASNP2[18] database. We also applied the lncRNASNP2 tools for predicting the potential miRNAs that these differential expression lncRNAs targeted. And the predicted miRNA-mRNA interact pairs were obtained from miRTarBase[19] and starbase V2.0 database.Next, the global triple network was constructed as the background network to identify gene interactions.

Gene expression profile

Gene Expression Omnibus (GEO) database is a public data source that stores generous gene expression profiles. In this study, The gene expression data of GSE95728 was downloaded from GEO database, which was based on the platform of GPL16956(Agilent-045997 Arraystar human lncRNA microarray V3). It compared the LncRNAs and mRNAs expression profiles in granulosa cells from seven PCOS patients and seven controls (women with normal ovarian reserve).

Identify the differential expression gene and Probe re-annotation

We imported the data into the R-studio and normalized it with RMA algorithm[20]. Bioconductor limma package was applied to identify the differential expression genes(DEGs) between control and PCOS groups, we considered that |log2(fold change)| > 2 and adjusted P-value <0.05 was statistically significant.

The probe annotation sequences supported by the Agilent technologieswere aligned to the human long non-coding transcript sequences and human protein-coding transcript sequences from the GENCODE databaseby running SeqMap program[21]. The alignment results were filtered as follows:

1. Retained the probes matched to one transcript, deleted the probes simultaneously matched to the protein-coding transcripts and long non-coding transcripts. Finally, we obtained two sets of probes-transcripts pairs.
2. For each probes–transcript pair, removed the probes matched to more than one transcripts.
3. Each transcript should be matched to three probes at any rate.

Construction of PCLMN

To construct PCLMN, we mapped all these differentially expressed lncRNAs and mRNAs into the global triple network, then lncRNA-miRNA-mRNA interactions were extracted by hypergeometric test with P < 0.01. The p-value was measured as:

$$P = 1 - \sum_{i=0}^{r-1} \frac{\binom{t}{i}\binom{m-t}{n-i}}{\binom{m}{n}}$$

where m meant the whole number of miRNAs in miRTarBase and starbase V2.0 database.t meant the number of miRNAs interacting with an mRNA, n meant the number of miRNAs interacting with a lncRNA, and r meant the number of miRNAs shared between the mRNA and lncRNA.

Hierarchical clustering

We performed unsupervised hierarchical clustering by Multiple Experiment Viewer (MeV V4.9) software. The data were normalized and processed using hierarchical clustering with Pearson correlation as distance metric selection and average linkage method. The genes with comparable expression profiles were grouped together.

Enrichment analysis

We used Gene Ontology (GO) terms of biological process and molecular function to find out the function of target genes from three aspects: biological process, cellular component and molecular function. Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis was applied to explore the significant pathways of DEGs. GO classification and KEGG pathway analyses were performed by DAVID V6.8 tools and Cytoscape V3.8.0 with the ClueGo V2.3.7 plug-in. The p-value was calculated by two-sided hypergeometric test and Benjamini-Hochberg adjustment. We considered the GO terms and KEGG pathways with p-value <0.05 was statistically significant. All the target genes in PCLMN and ceRNA sub-networks were studied by GO and KEGG pathway analyzes, and the results were visualized with Cytoscape.

Topological analysis and selection of key lncRNAs

We performed the topological analysis of differential expression LncRNAs (DELs)and differential expression mRNAs (DEMs) to explore the central nodes of the PCLMN network. Cytoscape with the CentiScaPe V2.2 plug-in was used to analyses the topological parameters including closeness, betweenness and degree. The top 10 genes in degree, betweenness, and closeness were compared and the overlapped genes across all three topological parameters were chosen as the key genes for further ceRNA analysis.

Identification of the cellular localization of key lncRNAs

We predicted the subcellular localization of key lncRNAs via LncLocator[22], which is a public platform based on a stacked ensemble classifier.Using lncRNA sequence information, the platform could provide the possibility of lncRNA distribution in five subcellular localizations, including cytoplasm, nucleus, ribosome, cytosol and exosome. The sequences of the key lncRNAs were downloaded from UCSC genome browser database.

Construction of ceRNA sub-networks

We extracted all the key lncRNAs and its first mRNA neighbors clustered together from the PCLMN. The miRNAs associating the lncRNA and mRNA pairs were extracted from the global triple network and used to construct thelncRNA-miRNA-mRNA triple pairs. Then we constructed the ceRNA networks based on the ceRNA theory and visualized them with Cytoscape software.

Identification of the putative transfer factors

TFs can bind to the DNA-regulatory elements of lncRNAs to activate or inhibit lncRNAs expression. To elaborate the potential linkages between the key lncRNAs, we identify the TFs that might regulate the key lncRNAs. The promoters were defined as the ±2 kb regions of lncRNA transcriptional start site.We used PROMO V3.0.2 with Maximum matrix dissimilarity rate < 3 to scan the predicted TFs [23, 24]. Then we used the Venn diagram to obtain the overlapping TFs targeting all the key lncRNAs.

Identification of functional modules in the PCLMN

It has been proved that LncRNAs participate in biological processes by coming into being as functional modules with other genes. To explore important lncRNA-related modules in network, we conducted Cytoscape with the MCODE plug-in to identify the functional modules (parameters: “Haircut,” “Fluff,” Node Score Cutoff: 0.2).

Results

Identification of differentially expressed lncRNAs and mRNAs

We download the gene expression profile of GSE95728 from GEO database, comparing the LncRNAs and mRNAs expression profiles in granulosa cells from seven PCOS patients and seven controls. 86 DELs and 112 DEMs were identified with FC > 2 and adjusted P < 0.05.

Construction of the global lncRNA-miRNA-mRNA triple network

In total, we obtained 502653 miRNA-mRNA interaction pairs from miRTarBase and starbase database. The human miRNA-lncRNA interactions were downloaded from the starbase database, then the lncRNASNP2 tool was preformed to predict the known miRNAs in the starbase database that could interact with the lncRNAs appeared in previous 86 differential expressed lncRNAs. We downloaded the differential expressed lncRNAs’ sequences from UCSC genome browser. The threshold we set was prediction score > 160 and energy < -20. In total, we obtained 66059 miRNA-lncRNA interaction pairs. All miRNA-mRNA pairs and miRNA-lncRNA pairs were merged to construct global triple network which was used as a background network to construct the PCLMN.

Construction of the PCOS related lncRNA-mRNA network

We mapped all the DELs and DEMs into the global triple network, then hypergeometric test was performed to extract 334 lncRNA-miRNA-mRNA triplets. In conclusion, 203 lncRNA-mRNA pairs including 41 lncRNAs and 41mRNAs were chosen to construct the PCLMN,as shown in Figure 2B.

The functional characteristics and Hierarchical clusteringof PCLMN

We clustered the 82 differentially expressed genes included in PCLMN by unsupervised hierarchical clustering, suggesting a unique gene expression profile between control and PCOS groups, as shown in Figure 2A.

We performed GO classification and KEGG pathway analysis on the DELs and DEMs in PCLMN. The result of GO analysis manifest that 27 GO terms enriched ($P \leq 0.05$ adjusted with Benjamini-Hochberg) listed in Figure 2C and Table 1. And KEGG pathway analysis found 13 pathways ($P \leq 0.05$), as shown in Figure 1D and Table 2.

Many inflammatory reactions were enriched in PCLMN, including Cytokines and Inflammatory Response, chemokine production, neutrophil chemotaxis, leukocyte aggregation, Interleukin-10 signaling pathway, IL-17 signaling pathway. The recent study proved that the presence of chronic low level inflammatory response states is closely related to the occurrence of insulin resistance, hyperandrogenemia, and metabolic syndrome in PCOS. Insulin resistance (IR) played a key role in POCS physiology and pathology, it is reported that inflammatory cytokines might induce insulin resistance by acting directly on insulin-like receptor molecules[25]. A study found that serum CRP and IL level increased with the raise of insulin and insulin resistance index (HOMA-IR), suggesting that inflammatory factors are closely related to IR formation[26]. And inflammatory factor-mediated signaling inhibits the tyrosine kinase activity of insulin receptor (INSR)and interferes with insulin substrate 1(IRS-1) synthesis, blocking INS signaling pathways and leading to IR formation[27]. Meanwhile, regulation of reactive oxygen species biosynthetic process was enriched in PCLMN. The oxidative stress markers in PCOS patients included lipid peroxidation and protein hydroxyl content increased and the antioxidant capacity declined[28]. Reactive oxygen species-induced oxidative stress may contribute to a proinflammatory state that induces insulin resistance and hyperandrogenism[29].

Many immune-related biological processes covering T cell selection, T celldifferentiation in thymus, regulation of cytokine production involved in immuneresponse, regulation of lymphocyte apoptotic process, IL1 and megakaryocytes in obesity, were enriched in PCLMN. Ovulation disorder is one of the common clinical manifestations of PCOS, also the main cause of infertility. The immune system can regulate follicle growth by releasing lymphocytes and macrophages. Lymphocytes can promote follicle growth, ovulation, and luteinization within a certain threshold, but abnormal inflammatory conditions can cause lymphocytes to over secrete inflammatory factors that promote cytotoxicity and induce apoptosis, and then phagocytosis by macrophages, leading to follicular development stagnation or ovulation failure[30].

Table1. Enriched GO terms in PCLMN.

GO Term	P-Value	Associated Genes%
regulation of receptor binding	2.67E-05	10.71
T cell selection	1.64E-04	5.88
positive T cell selection	6.78E-05	7.89
T cell differentiation in thymus	2.20E-05	5.06
granulocyte chemotaxis	4.15E-07	4.32
neutrophil migration	2.80E-07	4.62
neutrophil chemotaxis	1.09E-07	5.41
interleukin-12 production	3.23E-04	4.69
regulation of interleukin-12 production	2.94E-04	4.84
somatic diversification of immune receptors	4.75E-04	4.11
somatic cell DNA recombination	4.75E-04	4.11
chemokine production	4.93E-05	4.12
regulation of cytokine production involved in immune response	5.56E-05	4.00
regulation of chemokine production	4.01E-05	4.35
regulation of leukocyte apoptotic process	4.01E-05	4.35
regulation of lymphocyte apoptotic process	2.67E-04	5.00
heterotypic cell-cell adhesion	3.38E-04	4.62
chemokine production	4.93E-05	4.12
regulation of lipid catabolic process	3.38E-04	4.62
leukocyte aggregation	3.03E-06	21.43
regulation of reactive oxygen species biosynthetic process	5.56E-05	4.00
negative regulation of reactive oxygen species metabolic process	3.53E-04	4.55
regulation of cytokine production involved in immune response	5.56E-05	4.00
regulation of chemokine production	4.01E-05	4.35
regulation of heterotypic cell-cell adhesion	2.39E-05	11.11
negative regulation of reactive oxygen species biosynthetic process	4.42E-05	9.09
positive regulation of cytokine production involved in immune response	3.69E-04	4.48
positive regulation of chemokine production	4.03E-04	4.35
positive regulation of heterotypic cell-cell adhesion	4.65E-06	18.75

unsaturated fatty acid biosynthetic process	3.38E-04	4.62
regulation of lymphocyte apoptotic process	2.67E-04	5.00

Table2. Enriched KEGG Pathways in PCLMN.

GO Term	P-Value	Associated Genes%
Hematopoietic cell lineage	2.09E-04	4.04
IL-17 signaling pathway	8.26E-06	5.32
Intestinal immune network for IgA production	9.35E-03	4.08
Chagas disease (American trypanosomiasis)	5.22E-07	5.88
African trypanosomiasis	5.42E-03	5.41
Malaria	1.42E-05	8.00
Bladder cancer	2.43E-04	7.32
Asthma	3.82E-03	6.45
Allograft rejection	5.71E-03	5.26
Primary immunodeficiency	5.71E-03	5.26
Phosphorylation of CD3 and TCR zeta chains	1.93E-03	9.09
Translocation of ZAP-70 to Immunological synapse	1.44E-03	10.53
Generation of second messenger molecules	4.59E-03	5.88
PD-1 signaling	2.11E-03	8.70
Interleukin-10 signaling	1.11E-05	8.51
FCGR3A-mediated IL10 synthesis	7.26E-03	4.65
Nucleotide-binding Oligomerization Domain (NOD) pathway	6.62E-03	4.88
Allograft Rejection	1.44E-04	4.44
Bladder Cancer	2.43E-04	7.32
IL-3 Signaling Pathway	9.35E-03	4.08
IL1 and megakaryocytes in obesity	2.50E-03	8.00
Photodynamic therapy-induced NF-kB survival signaling	5.56E-08	14.29
Lung fibrosis	8.64E-04	4.76
Hepatitis C and Hepatocellular Carcinoma	9.72E-03	4.00
T-Cell antigen Receptor pathway during Staphylococcus aureus infection	8.25E-04	4.84
Development and heterogeneity of the ILC family	4.07E-03	6.25
Platelet-mediated interactions with vascular and circulating cells	1.15E-03	11.76
LTF danger signal response pathway	2.71E-05	15.00
Cancer immunotherapy by PD-1 blockade	2.30E-03	8.33

Pathogenesis of SARS-CoV-2 Mediated by nsp9/nsp10 Complex	1.93E-03	9.09
COVID-19 AOP	9.07E-08	26.67
Cytokines and Inflammatory Response	2.91E-03	7.41
Type II interferon signaling (IFNG)	5.42E-03	5.41

Topological characteristics of PCLMN and the cellular localization of key lncRNAs

The degree distribution was performed to the PCLMN and all the nodes were found that followed power-law distribution ($R^2 = 0.80$), as shown in Figure 3A, indicating that PCLMN network is a scale-free network. We analyzed the topological characteristics of PCLMN to identify the biological functions of lncRNAs. Parameters of degree, betweenness, and closeness were calculated. Then we selected top ten of each topological parameter and found that three lncRNAs were intersected among the three features, they are LINC00667, AC073172.1, and H19, as shown in Figure 3B. And the subcellular localization of lncRNA played a crucial role in its function and molecular mechanism, ceRNA generally worked in the cytosol and cytoplasm, so it was necessary to confirm the subcellular localization of key lncRNAs.

For LINC00667, we found that it was mainly located in cytosol, as shown in Figure 4A. And there were 13 first mRNA neighbors in PCLMN. GO classification adjusted by Benjamini-Hochberg with $P \leq 0.05$ detected four enriched Biological process, one Molecular function and one Cellular Component, as shown in Figure 4C. KEGG analysis with $P \leq 0.05$ found ten enriched KEGG pathways, as shown in Figure 4D.

NF-kappa B signaling pathway was enriched in the ceRNA sub-network of LINC00667. Nuclear factor kappaB (NF-κB) is a crucial factor regulating inflammation, and following exogenous stimulation-free NF-κB translocates into the nucleus where it induces gene transcription to regulate and promote an inflammatory reaction[31]. It is reported that NF-κB in the PCOS groups weas significantly higher[32], indicating that PCOS maight be associated with chronic inflammation. In addition, GO classification also pointed to inflammatory response, upholding its importance in PCOS.

Biological process also enriched in the negative regulation of apoptotic process. Apoptosis is considered to be the mechanism of follicular atresia and the basis for the cyclical growth and regression of human ovarian follicles[33]. Current studies confirmed that the reduction of some apoptotic markers could be one of the pathogenesis of PCOS[34]. Compared with normally ovulating women, the apoptotic-related moleculer, caspases 3, 8 and 9, were significantly reduced, and the anti-apoptotic regulator, cIAP-2, Hsp27 were over expressed in oocytes of women with PCOS[35, 36].

For H19, we found that it was mainly located in cytosol, as shown in Figure 5A . And 13 mRNAs were picked up from the PCLMN as the first mRNA neighbors. GO classification adjusted by Benjamini-Hochberg with P ≤0.05 detected three enriched Biological process and one Cellular Component, as shown in Figure 5C. KEGG analysis with P ≤ 0.05 found seven enriched KEGG pathways, as shown in Figure 5D.

Biological process enriched in the negative regulation of tumor necrosis factor production. Tumor necrosis factor-α (TNF-α) is a crucial mediator of insulin resistance through its capacity to weaken the tyrosine kinase activity of the insulin receptor[25]. It is reported that TNF-α expression is associated with indexes of insulin resistance, with the increase of insulin sensitivity, the expression of TNF-α decreased with weight loss[37]. And the level of TNF-α in the PCOS group was significantly higher than healthy controls[38].

For AC073172.1, we found that it was mainly located in cytosol, as shown in Figure 6A. And there were 17 mRNAs picked up from the PCLMN as the first mRNA neighbors. GO classification adjusted by Benjamini-Hochberg with P ≤0.05 detected four enriched Biological process and two Molecular function, as shown in Figure 6C. KEGG analysis with P ≤ 0.05 found four enriched KEGG pathways, as shown in Figure 6D.

Biological process enriched in cellular response to lipopolysaccharide, and Toll-like receptor signaling pathway was enriched in the ceRNA sub-network of AC073172.1. Lipopolysaccharide (LPS) is the main pathogenic component in the process of inducing inflammatory response, which is involved in the immune inflammatory response by stimulating the body to release a variety of inflammatory mediators. LPS can bind to Toll-like receptor 4 (TLR4) to activate NF-κB, then promoting transcription of TNF-α ,IL-1β and IL-6[39]. Recent studies found that lipid-induced LPS-mediated inflammation through TLR4 is associated with obesity in PCOS[40]. Saturated fatty acids activate TLR2 and TLR4 signaling, leading to insulin resistance[41], and TLR4 genetic variants have been proved to affect insulin resistance associated with obesity[42]. And this further demonstrated the important role of inflammatory response in PCOS development.

Identification of the putative transfer factors

We predicted the TFs of the key lncRNAs on PROMO with Maximum matrix dissimilarity rate < 3 , according the sequences of promoters. Finally, We got 27 common TFs targeting all the key lncRNAs, as shown in Figure 7A. According to recent studies, four TFs were associated with PCOS, they were marked in red in Figure , including ER-alpha, PRA, PRB, AR. Sex steroid hormones, such as estrogen and progesterone, play an important role in fertility and reproductive function in the ovary through specific nuclear receptors. Estrogens act by binding to estrogen receptor-alpha (ERα) and beta (ERβ)[43], and both ERα and ERβ are expressed in the human ovary[44]. A study in mice shown that the knockout of ERα leads to the PCOS phenotype with the presence of polycystic ovaries and increased luteinizing hormone (LH) levels[45]. It is reported that mice missing ERα are found to be insulin resistant with impaired glucose tolerance[46], and the insulin resistance in ERα knockout mice was aggravated by accumulation of the bioactive lipid intermediates, inflammation became more severe with high-fat feeding, suggesting

that ERα is also essential to protect tissue inflammation[47]. Progesterone receptor (PR) and androgen receptor(AR) belong to the nuclear hormone receptor family, which are associated with the regulation of eukaryotic gene expression and influencecellular proliferation and differentiation in target tissues. Previous study on development of PRA and PRB knockouts suggested that PRA is necessary for ovulation[48]. PRA and PRB was found to be expressed in human granulosa cells, and the expression of PRA and PRB in PCOS patients was significantly lower than control group[49]. In vitro and in vivo studies in human found that the number of CAG repeats, which encode for an amino-acid sequence in the receptor's transactivation domain, associates inversely with the AR activity[50]. It is reported that short CAG repeats were more frequent in PCOS possibly raising androgenic effects, while longer CAG repeats were more recurrent in the control group, probably playing a protective effect[51]. In brief, sex steroids, TFs, lncRNAs and down-stream factors establish a network regulating the occurrence and development of PCOS.

Identification of functional modules in the PCLMN

We performed Cytoscape with the MCODE clustering algorithm for PCLMN. One functional module that contained H19, LINC00667, MYH11 and BCL11B was identified,as shown in Figure 7B. Further proved the importance of H19 and LINC00667 in the PCLMN. Additionally, BCL11B played an important role in NF-kB pathway, it could enhance TCR/CD28-triggered NF-kB activation[52]. It also proved the importance of inflammatory response in PCOS.

Discussion

In our study, we performed a network-based computational analysis to explore the effective lncRNAs in PCOS. The interactions data from starbase, lncRNASNP2 and miRTarBase database were used to construct a global triple network based on the theory of ceRNA. Then the PCOS related lncRNA-mRNA network (PCLMN) was extracted by mapping the differential expressed genes into the global triple network. This network brought us a new global view of PCOS to identify crucial functional biomarkers. The PCLMN contained 41 mRNA nodes, 41 lncRNA nodes and 203 edges. After that, topology analysis and the MCODE clustering algorithm were performed for PCLMN to identify crucial lncRNAs with central topology structures. Finally, we identify the TFs that might regulate the crucial lncRNAs. According to these analyses, we found that 3 lncRNAs might participate in regulating key pathways in PCOS, such as NF-kB signaling pathway, inflammatory reaction, apoptotic process and immune-related processes. Recent studies have confirmed that the levels of proinflammatory cytokines such as C-reactive protein and IL6, IL8, IL10, IL18 in peripheral blood of PCOS patients are increased,suggesting that PCOS may be a low-grade chronic inflammatory disease[53, 54]. NF-kB is an important transcription factor to initiated and regulated the transcription and expression of various inflammatory mediators, promoted the occurrence and development of inflammatory response, and was also closely related to insulin resistance[55]. Studies shown that PCOS patients had a typical disorder of follicular development, characterized by excessive growth of early growth follicles and dominant follicle selection disorders, which might be associated with abnormal apoptosis regulation[56]. Moreover, studies found that

lncRNAs can synergize with TFs to regulate the biological processes in sex steroid functions. For example, Androgen receptors and PlncRNA-1 could form regulatory circuits that promote prostate cancer[57]. HOTAIR is a direct target of ER-mediated transcriptional repression, and its upregulation promotes ligand-independent ER activities[58]. Thus, we performed motif scanning in transcriptional start site of crucial lncRNAs, the results shown that some key TFs had high binding affinity in transcriptional start site of crucial lncRNAs. We identified a subset of kernel lncRNA-TF pairs that might constitute feedback loops to affect the functional process of sex steroids.

Conclusion

In short, we effectually identified 3 lncRNAs correlated to PCOS according to the ceRNA theory. The crucial lncRNA played an important role in PCOS and could provide new diagnostic markers and therapeutic targets for PCOS.

Declarations

Acknowledgements

Not applicable.

Authors' contributions

Y.M. conducted all analysis, and interpretation of data, and written the manuscript. J.Z. provided critical revision of final manuscript. All authors read and approved the final manuscript.

Funding

Not applicable.

Availability of data and materials

The gene expression datasets used in this study (GSE95728) was acquired from the National Center for Biotechnology Information (NCBI) Gene expression (http://www.ncbi.nlm.nih.gov/geo/).

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors have declared that no competing interest exists.

Reference

1. **Genome-wide association study identifies eight new risk loci for polycystic ovary syndrome**. *Nature Genetics* 2012, **44**(9):1020.

2. None: **Revised 2003 consensus on diagnostic criteria and long-term health risks related to polycystic ovary syndrome**. *Human Reproduction* 2004, **81**(1):19-25.

3. Christakou C, Diamanti-Kandarakis E: **Polycystic ovary syndrome – Phenotypes and diagnosis**. *Scandinavian Journal of Clinical & Laboratory Investigation* 2014, **74**(S244):18-22.

4. S., Ranasinha, A.E., Joham, R.J., Norman, J.E., Shaw, S., Zoungas: **The association between Polycystic Ovary Syndrome (PCOS) and metabolic syndrome: a statistical modelling approach**. *Clinical Endocrinology* 2015.

5. Kakoly NS, Khomami MB, Joham AE, Cooray SD, Misso ML, Norman RJ, Harrison CL, Ranasinha S, Teede HJ, Moran LJ: **Ethnicity, obesity and the prevalence of impaired glucose tolerance and type 2 diabetes in PCOS: a systematic review and meta-regression**. *Human Reproduction Update* 2018(4):4.

6. Bilal M, Haseeb A, Rehman A: **Relationship of Polycystic Ovarian Syndrome with Cardiovascular Risk Factors**. *Diabetes Metab Syndr* 2018:S1871402117304885.

7. Shang K, Jia X, Qiao J, Kang J, Guan Y: **Endometrial Abnormality in Women With Polycystic Ovary Syndrome**. *Reproductive Sciences* 2012, **19**(7):674-683.

8. Mattick, John S: **Non-coding RNAs: the architects of eukaryotic complexity**. *Embo Journal* 2001, **2**(11):986-991.

9. An S, Song JJ: **The coded functions of noncoding RNAs for gene regulation**. *Molecules & Cells* 2011, **31**(6):491-496.

10. Guttman M, Amit I, Garber M, French C, Lin MF, Feldser D, Huarte M, Zuk O, Carey BW, Cassady JP: **Chromatin Signature Reveals over a Thousand Highly Conserved Large Non-Coding Rnas in Mammals**. *Nature* 2009, **458**(7235):223-227.

11. Liu YD, Li Y, Feng SX, Ye DS, Chen X, Zhou XY, Chen SL: **Long noncoding RNAs: Potential regulators involved in the pathogenesis of polycystic ovary syndrome**. *Endocrinology* 2017.

12. Takayama KI, Horie-Inoue K, Katayama S, Suzuki T, Tsutsumi S, Ikeda K, Urano T, Fujimura T, Takagi K, Takahashi S: **Androgen-responsive long noncoding RNA CTBP1-AS promotes prostate cancer**. *Embo Journal* 2014, **32**(12):1665-1680.

13. Leonardo S: **A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language?** *Cell* 2011, **3**(146).

14. Xiang J, Guan Y, Hui S, Yang L: **Crosstalk in competing endogenous RNA network reveals the complex molecular mechanism underlying lung cancer**. *Oncotarget* 2017, **8**(53):91270-91280.

15. Song C, Zhang J, Liu Y, Pan H, Li CQ: **Construction and analysis of cardiac hypertrophy-associated lncRNA-mRNA network based on competitive endogenous RNA reveal functional lncRNAs in cardiac hypertrophy**. *Oncotarget* 2016, **7**(10):10827-10840.

16. Feng C, Shen JM, Lv PP, Jin M, Wang LQ, Rao JP, Feng L: **Construction of implantation failure related lncRNA-mRNA network and identification of lncRNA biomarkers for predicting endometrial receptivity**. *Int J Biol Sci* 2018, **14**(10):1361-1377.

17. Jian-Hua, Yang, Jun-Hao, Li, Peng, Shao, Hui, Zhou, Yue-Qin, Chen: **starBase: a database for exploring microRNA-mRNA interaction maps from Argonaute CLIP-Seq and Degradome-Seq data**. *Nucleic Acids Research* 2011.

18. Gong J, Liu W, Zhang J, Miao X, Guo AY: **lncRNASNP: a database of SNPs in lncRNAs and their potential functions in human and mouse**. *Nuclc Acids Research* 2015(D1):D181.

19. 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 Research* 2017.

20. Harbron C, Chang KM, South MC: **RefPlus: an R package extending the RMA Algorithm**. *Bioinformatics* 2007, **23**(18):2493.

21. Jiang H, Wong WH: **SeqMap**. *Bioinformatics* 2008, **24**(20):2395-2396.

22. Cao Z, Pan X, Yang Y, Huang Y, Shen H-B: **The lncLocator: a subcellular localization predictor for long non-coding RNAs based on a stacked ensemble classifier**. *Bioinformatics* 2018, **34**(13):2185-2194.

23. Laguna-Bercero MA, Larrea A, Merino RI, Pe?A JI, Orera VM: **Crystallography and thermal stability of textured Co-YSZ cermets from eutectic precursors**. *Journal of the European Ceramic Society* 2008, **28**(12):2325-2329.

24. Domènec F: **Identification of patterns in biological sequences at the ALGGEN server: PROMO and MALGEN**. *Nucleic acids research* 2003, **13**(31).

25. Hotamisligil, Gokhan, S., Peraldi, Pascal: **IRS-1-mediated inhibition of insulin receptor tyrosine kinase activity in TNF-alpha- and obesity**. *Science* 1996.

26. Sarray S, Almawi WY: **Levels of CD40L and other inflammatory biomarkers in obese and non-obese women with polycystic ovary syndrome**. *American Journal of Reproductive Immunology* 2016.

27. Deligeoroglou E, Vrachnis N, Athanasopoulos N, Iliodromiti Z, Sifakis S, Iliodromiti S, Siristatidis C, Creatsas G: **Mediators of chronic inflammation in polycystic ovarian syndrome**. *Gynecological Endocrinology the Official Journal of the International Society of Gynecological Endocrinology* 2012, **28**(12):974-978.

28. Ku??U NK, Var A: **Oxidative stress but not endothelial dysfunction exists in non-obese, young group of patients with polycystic ovary syndrome**. *Acta Obstetricia Et Gynecologica Scandinavica* 2009, **88**(5):612.

29. González F, Rote NS, Minium J, Kirwan JP: **Reactive oxygen species-induced oxidative stress in the development of insulin resistance and hyperandrogenism in polycystic ovary syndrome**. *Journal of Clinical Endocrinology & Metabolism* 2006, **91**(1):336-340.

30. Walusimbi SS, Pate JL: **PHYSIOLOGY AND ENDOCRINOLOGY SYMPOSIUM: Role of immune cells in the corpus luteum**. *Journal of Animal Science* 2013, **91**(4):1650-1659.

31. Yamamoto Y: **IkappaB kinases: key regulators of the NF-kappaB pathway**. *Trends in Biochemical Sciences* 2004, **4**(29).

32. Liu M, Gao J, Zhang Y, Li P, Wang H, Ren X, Li C: **Serum levels of TSP-1, NF-kappaB and TGF-beta1 in polycystic ovarian syndrome (PCOS) patients in northern China suggest PCOS is associated with chronic inflammation**. *Clin Endocrinol (Oxf)* 2015, **83**(6):913-922.

33. Slot, A K: **Luteinizing hormone inhibits Fas-induced apoptosis in ovarian surface epithelial cell lines**. *Journal of Endocrinology* 2006, **188**(2):227.

34. Uyanikoglu H, Sabuncu T, Dursun H, Sezen H, Aksoy N: **Circulating levels of apoptotic markers and oxidative stress parameters in women with polycystic ovary syndrome: a case-controlled descriptive study**. *Biomarkers*, **22**(7):1-18.

35. Cai L, Ma X, Liu S, Liu J, Liu J: **Effects of Upregulation of Hsp27 Expression on Oocyte Development and Maturation Derived from Polycystic Ovary Syndrome**. *Plos One* 2013, **8**(12):e83402.

36. M D: **Granulosa cell survival and proliferation are altered in polycystic ovary syndrome**. *The Journal of clinical endocrinology and metabolism* 2008, **3**(93).

37. Hotamisligil GS, Arner P, Atkinson RL, Spiegelman BM: **Differential Regulation of the p80 Tumor Necrosis Factor Receptor in Human Obesity and Insulin Resistance**. *Diabetes* 1997, **46**(3):451.

38. Nervana S, Maha H, Magda S, Mohamed S: **Clinical Significance of Inflammatory Markers in Polycystic Ovary Syndrome: Their Relationship to Insulin Resistance and Body Mass Index**. *Disease Markers* 2009, **26**(4):163-170.

39. Fitzgeraldbocarsly P: **Natural interferon-alpha producing cells: the plasmacytoid dendritic cells**. *Biotechniques*, **Suppl**(4):16.

40. Frank, González, Robert, Considine, Ola, Abdelhadi, Anthony: **Saturated fat ingestion promotes lipopolysaccharide-mediated inflammation and insulin resistance in PCOS**. *Journal of Clinical Endocrinology & Metabolism* 2018.

41. Yin J, Peng Y, Wu J, Wang Y, Yao L: **Toll-like receptor 2/4 links to free fatty acid-induced inflammation and β-cell dysfunction**. *Journal of Leukocyte Biology* 2014, **95**(1):47–52.

42. Steinhardt AP, Aranguren F, Tellechea ML, Rosso LAG, Brites FD, Martínez-Larrad MT, Serrano-Ríos M, Frechtel GD, Taverna MJ: **A functional nonsynonymous toll-like receptor 4 gene polymorphism is associated with metabolic syndrome, surrogates of insulin resistance, and syndromes of lipid accumulation**. *Metabolism-clinical & Experimental* 2010, **59**(5):711-717.

43. Alexander H: **Dissecting physiological roles of estrogen receptor alpha and beta with potent selective ligands from structure-based design**. *Molecular endocrinology (Baltimore, Md)* 2004, **7**(18).

44. Lange U, Neumann E: **Impact of isotype selective estrogen receptor agonists on ovarian function**. *Experimental & Clinical Endocrinology & Diabetes*:-.

45. David, Schomberg, John, Couse, Abir, Mukherjee, Dennis: **Targeted Disruption of the Estrogen Receptor-α Gene in Female Mice: Characterization of Ovarian Responses and Phenotype in the Adult**. *Endocrinology* 1999.

46. P A H: **Increased adipose tissue in male and female estrogen receptor-alpha knockout mice**. *Proceedings of the National Academy of Sciences of the United States of America* 2000, **23**(97).

47. Ribas V, Nguyen MTA, Henstridge DC, Nguyen AK, Beaven SW, Watt MJ, Hevener AL: **Impaired oxidative metabolism and inflammation are associated with insulin resistance in ERalpha-deficient mice**. *American Journal of Physiology Endocrinology & Metabolism* 2010, **298**(2):E304.

48. Shimada: **Expression of two progesterone receptor isoforms in cumulus cells and their roles during meiotic resumption of porcine oocytes**. *Journal of Molecular Endocrinology* 2004, **33**(1):209-225.

49. Artimani, Tayebeh, Shabab, Nooshin, Aflatoonian, Reza, Mehdizadeh, Mehdi, Saidijam, Massoud: **Estrogen and progesterone receptor subtype expression in granulosa cells from women with polycystic ovary syndrome**. *Gynecological endocrinology: the official journal of the International Society of Gynecological Endocrinology* 2015.

50. Zitzmann M, Depenbusch M, Gromoll J, Nieschlag E: **X-chromosome inactivation patterns and androgen receptor functionality influence phenotype and social characteristics as well as pharmacogenetics of testosterone therapy in Klinefelter patients**. *J Clin Endocrinol Metab* 2004, **89**(12):6208-6217.

51. Schuring AN, Welp A, Gromoll J, Zitzmann M, Sonntag B, Nieschlag E, Greb RR, Kiesel L: **Role of the CAG repeat polymorphism of the androgen receptor gene in polycystic ovary syndrome (PCOS)**. *Exp Clin*

Endocrinol Diabetes 2012, **120**(2):73-79.

52. Cismasiu Valeriu B, Duque J, Paskaleva E, Califano D, Ghanta S, Young Howard A, Avram D: **BCL11B enhances TCR/CD28-triggered NF-κB activation through up-regulation of Cot kinase gene expression in T-lymphocytes**. *Biochemical Journal* 2009, **417**(2):457-466.

53. Ilhan T, Arslan BC, Zeynep C, Erdem T, Tayfun S, Can D: **Endothelial dysfunction in young women with polycystic ovary syndrome: relationship with insulin resistance and low-grade chronic inflammation**. *Journal of Clinical Endocrinology & Metabolism* 2004(11):5592.

54. Blum CA, Beat ML, Peter H, Marius K, Christian S, Christian DG, Ulrich K, Puder JJ: **Low-grade inflammation and estimates of insulin resistance during the menstrual cycle in lean and overweight women**. *Journal of Clinical Endocrinology & Metabolism* 2005(6):3230.

55. Lan W, Petznick A, Heryati S, Rifada M, Tong L: **Nuclear Factor-κB: Central Regulator in Ocular Surface Inflammation and Diseases**. *Ocular Surface* 2012, **10**(3):137-148.

56. Maciel GA, Baracat EC, Benda JA, Markham SM, Hensinger K, Chang RJ, Erickson GF: **Stockpiling of transitional and classic primary follicles in ovaries of women with polycystic ovary syndrome**. *J Clin Endocrinol Metab* 2004, **89**(11):5321-5327.

57. Fang Z, Xu C, Li Y, Cai X, Ren S, Liu H, Wang Y, Wang F, Chen R, Qu M: **A feed-forward regulatory loop between androgen receptor and PlncRNA-1 promotes prostate cancer progression**. *Cancer Letters* 2016:62-74.

58. Xue X, Yang YA, Zhang A, Fong KW, Kim J, Song B, Li S, Zhao JC, Yu J: **LncRNA HOTAIR enhances ER signaling and confers tamoxifen resistance in breast cancer**. *Oncogene* 2016.

Figures

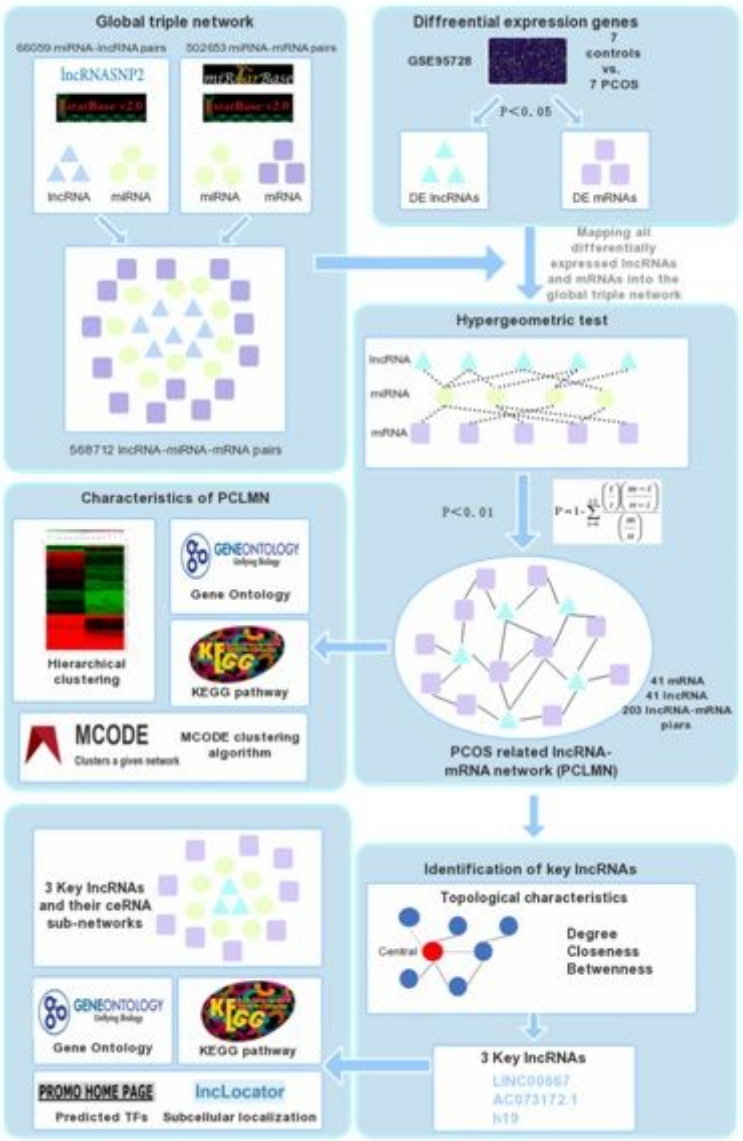

Figure 1

Workflow of this study. First, we constructed the global background network based on the assumed lncRNA-miRNA and miRNA-mRNA pairs. Second, we constructed PCLMN by hypergeometric test, and performed topological analysis to determine the crucial lncRNAs with the highest centroid. Third, We identified the subcellular localization, performed functional module analyses and identified putative transfer factors of the key lncRNAs.

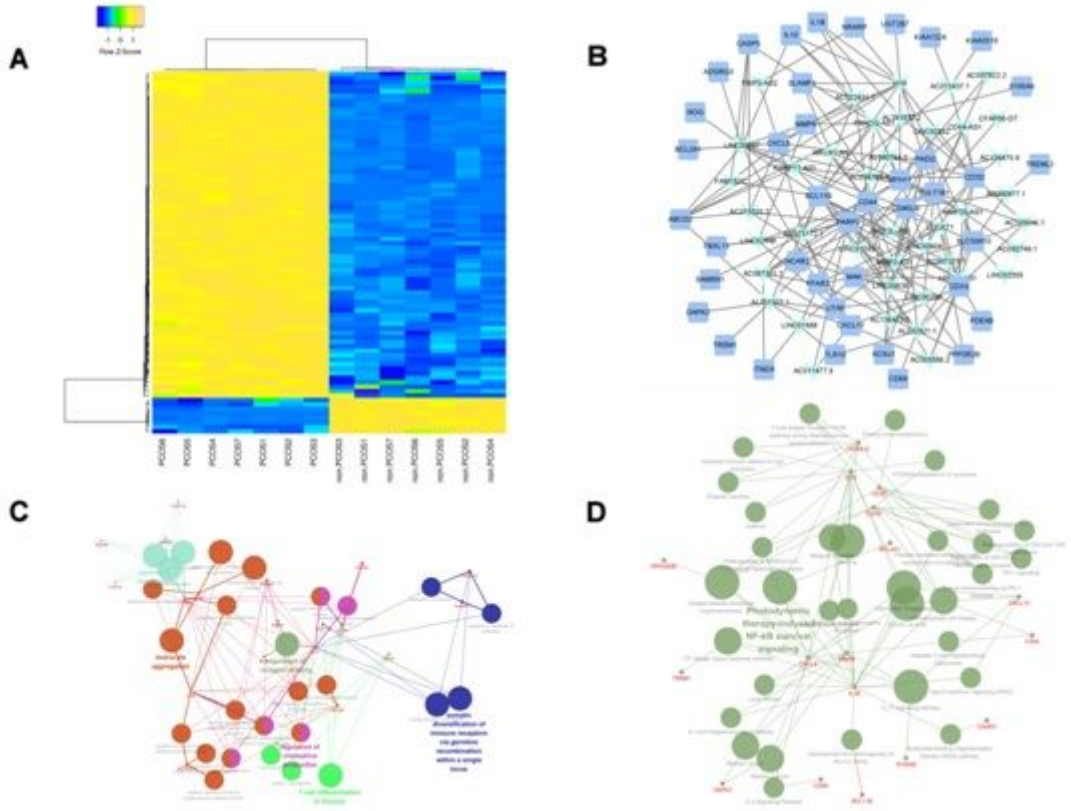

Figure 2

PCOS related lncRNA-mRNA network (PCLMN) and enrichment analysis. A. The clustering of the PCLMN genes. Unsupervised clustering of PCLMN genes could discriminate between control and PCOS groups. B. The blue triangle represents lncRNA and the green rounded rectangle represents mRNA. Grey edges indicate lncRNA-mRNA interactions. C. GO analysis of PCLMN. D. KEGG pathway analysis of PCLMN.

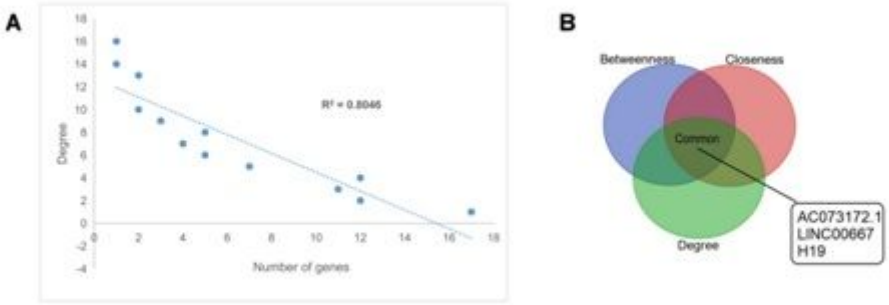

Figure 3

Topology features of ODLMN. A. Degree distributions of the network. All degrees followed a power-law distribution. B. There are four genes existed simultaneously in top 10 lists of degree, betweenness, and closeness.

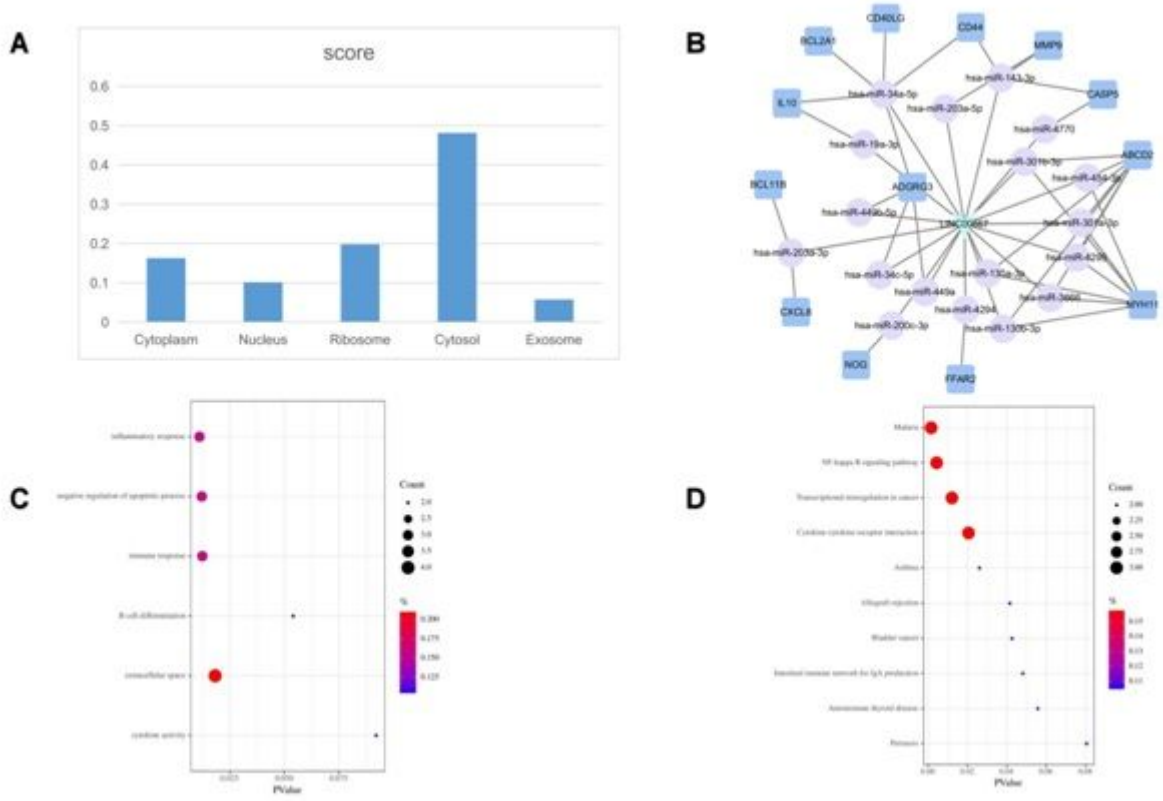

Figure 4

The ceRNA sub-network of LINC00667. A. The cellular location of LINC00667. B. The ceRNA network of LINC00667. C. Biological process enrichment of LINC00667. D. KEGG enrichment of LINC00667.

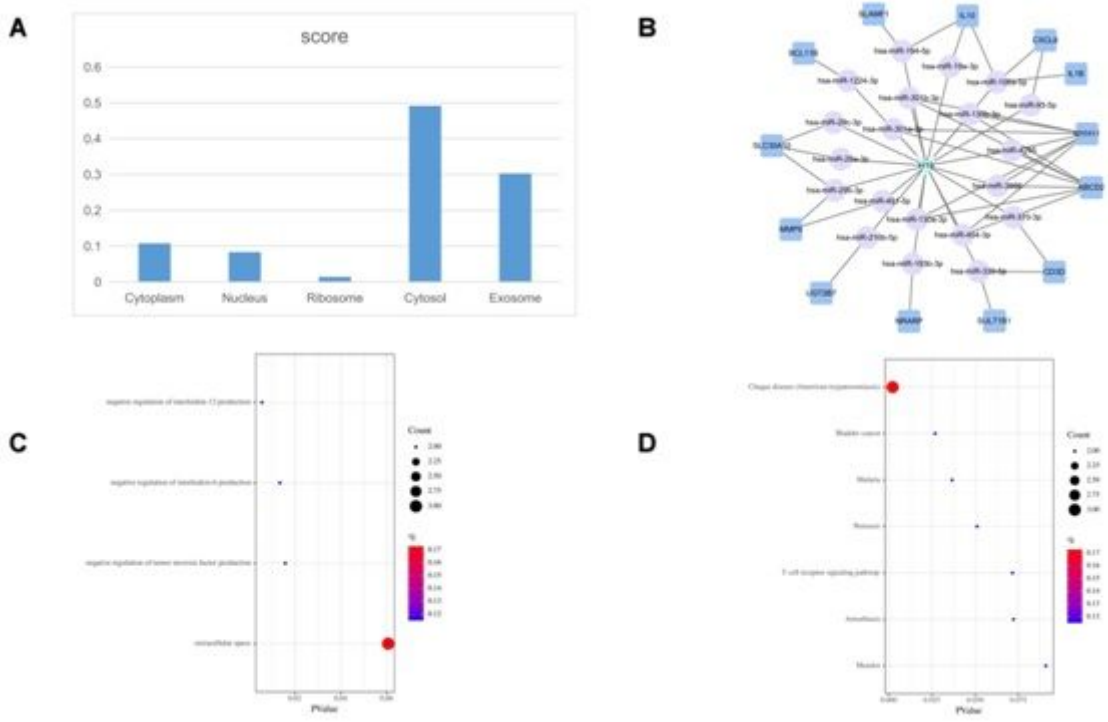

Figure 5

The ceRNA sub-network of H19. A. The cellular location of H19. B. The ceRNA network of H19. C. Biological process enrichment of H19. D. KEGG enrichment of H19.

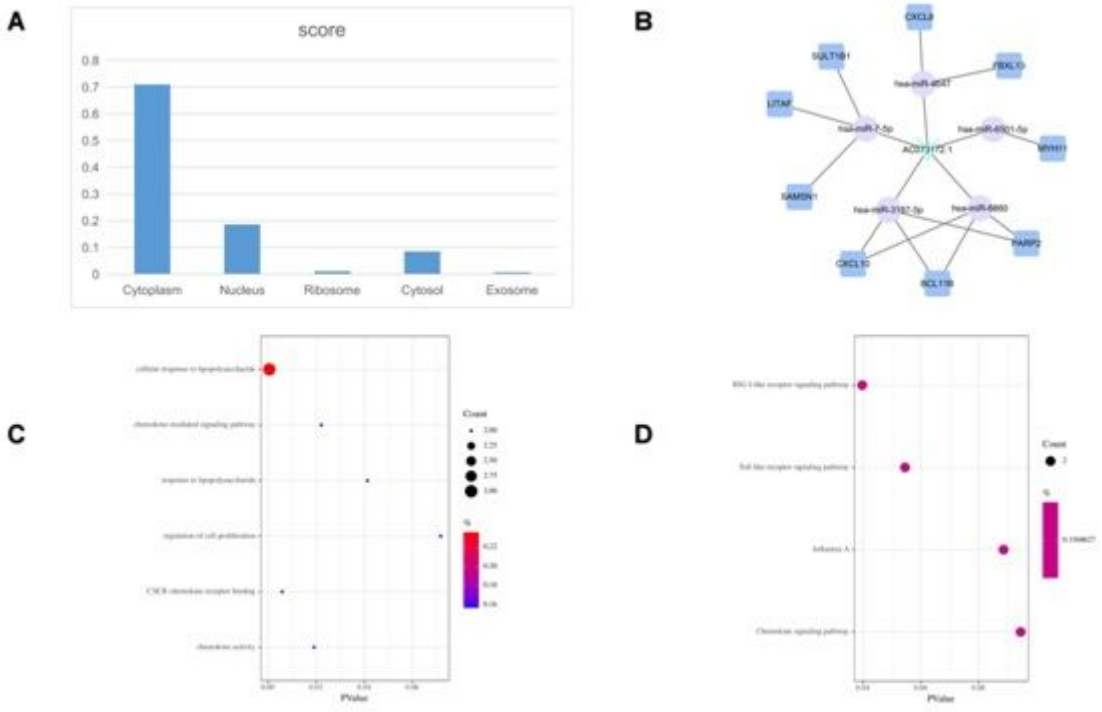

Figure 6

The ceRNA sub-network of AC073172.1. A. The cellular location of AC073172.1. B. The ceRNA network of AC073172.1. C. Biological process enrichment of AC073172.1. D. KEGG enrichment of AC073172.1.

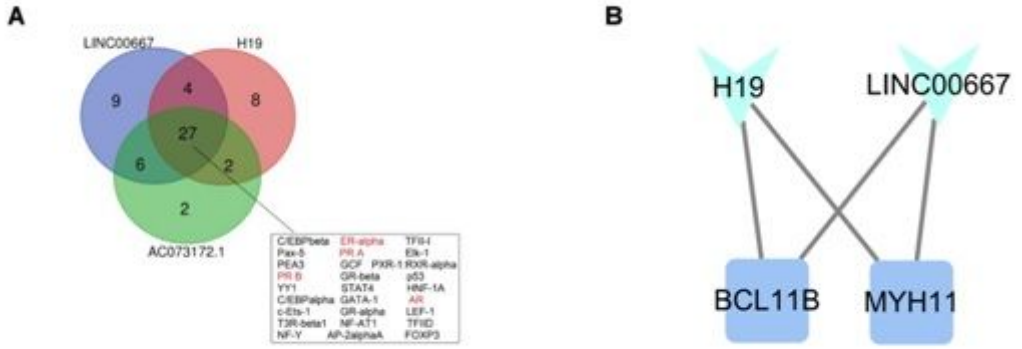

Figure 7

Predicted transcription factors and module analysis of PCLMN. A. The Venn of predicted transcription factors (TFs) of the three key lncRNAs. There are 27 common TFs and TFs involved in the sexual hormone functions were marked in red. B. The view of module.