Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Construction of a ceRNA network combining transcription factors in eutopic endometrial tissue of tubal factor infertility and endometriosis-related infertility

Junzui Li (linchuangkeyan@yeah.net)
Xiamen University https://orcid.org/0000-0002-6640-8215
Lulu Ren
Xiamen University Medical College
Cui Yang
Xiamen University Medical College
Rongfeng Wu
Xiamen University Medical College
Zhixiong Huang
Xiamen University School of Life Sciences
Qionghua Chen
Xiamen University Medical College

Research article

Keywords: TFI, endometriosis-related infertility, DEGs, ceRNA network, TFs

Posted Date: December 19th, 2019

DOI: https://doi.org/10.21203/rs.2.19312/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Abstract

Purpose Although tubal factor infertility (TFI) and endometriosis-related infertility all can result in female infertility, the pathogenesis of TFI and endometriosis-related infertility were different. The pathophysiologic mechanisms of TFI and endometriosis-related infertility have not been investigated thoroughly. Thus, the aim of the study is to identify the potential crucial genes, pathways, transcription factors (TFs) and long non-coding RNAs (lncRNAs) associated with TFI and endometriosis-related infertility, and further analyze the molecular mechanism implicated in genes.

Methods 3 patients with TFI and 3 patients with endometriosis-related infertility were recruited, and microarray hybridization of the eutopic endometrial tissue during the window of implantation (WOI) was performed to examine the expression of mRNAs and lncRNAs. First, differentially expressed genes (DEGs) and differentially expressed lncRNAs (DELs) were screened out based on $P < 0.05$ and fold change (FC) ≧ 2. Second, gene ontology, pathway and TFs enrichment analyses and PPI network construction of DEGs were performed. Then, important modules and Hub genes of PPI network were further analysed and a ceRNA network combining TFs based on DEGs and DELs was constructed. Finally, we selected the lapping genes of the important modules, Hub genes and the ceRNA network as potential key genes associated with TFI and endometriosis-related infertility and constructed a new ceRNA network on the base of potential key genes.

Results 508 DEGs and 576 DELs were screened out. The gene ontology and pathway of DEGs were mainly enriched in transmembrane transporter activity and the immune system (eg, passive transmembrane transporter activity, Intestinal immune network for IgA production and so on). In addition, a ceRNA network (based on potential key genes) combining TFs including 2 mRNAs (PLAU and LDLR), 5 miRNAs (hsa-miR-301b-3p, hsa-miR-27a-3p, hsa-miR-20b-5p, hsa-miR-193a-3p and hsa-miR-17-5p), 8 lncRNAs (eg. LMO7-AS1, ITFG3 and MGST1, etc.) and 10 TFs (eg. SRF, FOX, which target to mRNA and eg. POU3F2, HNF1A, which target to miRNA) was successfully built.

Conclusions In conclusion, pathophysiologic mechanisms of TFI and endometriosis related infertility may be related to the transmembrane transporter activity and the immune system. These potential key RNAs and TFs might have clinical utility for the diagnosis and prognosis prediction in TFI and endometriosis-related infertility. The results of the current study might lay the foundation for future basic and clinical research.

Introduction

Tubal factor infertility (TFI) and endometriosis-related infertility are common gynecological reproductive system diseases in the world, both of which seriously affect women's quality of life.[1,2] Although the two diseases have different pathogenesis, they can result in female infertility by affecting the uterine endometrium.[3] The fallopian tube is one of the important components of the female reproductive system. It has the important role of transporting sperm, ingesting eggs and transporting fertilized eggs to the

uterine cavity.[4] TFI refers to women who are infertile caused by tubal adhesions, blockages, insufficient ovulation of the ovaries and other reasons. Recent studies have shown that patients with TFI are significantly thinner in endometrial thickness, which may be one of the pathological mechanisms leading to infertility.[5,6] Endometriosis is a common gynecological disease in women and it refers to the presence of the endometrium outside the uterus. There are various mechanisms for the pathogenesis of this disease, and the endometrial implantation theory is generally accepted. Recent studies have shown that pathological changes in the eutopic endometrium of endometriosis may be associated with cell proliferation, immunity, and implantation.[7]

The endometrium plays an important role in the implantation of blastocysts. Under normal circumstances, the endometrium is a tissue that undergoes periodic changes in hormone regulation[8], and it can only accept embryo implantation during the window of implantation (WOI). During the WOI, the activated blastocyst interacts with the endometrium in the receiving state to promote blastocyst implantation into the uterus. This process begins on days 4-5 after fertilization and ends approximately on days 9-10, which is equivalent to the 19th to 24th day of the menstrual cycle. During this period, a series of morphological and biochemical changes in the endometrium are inseparable from changes in maternal secretion and changes in the expression of numerous molecules in endometrial cells.[9]

Recent studies have shown that with the rapid development of gene chips, high-throughput microarray have been applied to various aspects of disease pathogenesis, such as transcription, translation, and epigenetic regulation.[10-12] However, in the study of TFI and endometriosis-related infertility, the microarray technology of the two has not yet been popularized. In addition, the eutopic endometrium is also an important tissue in genetic analysis, which plays an important role in TFI and endometriosis-related infertility. However, the pathophysiologic mechanisms of TFI and endometriosis-related infertility have not been investigated thoroughly.

Here we selected patients who had been diagnosed with TFI and endometriosis-related infertility respectively in the Department of Obstetrics and Gynecology, the First Affiliated Hospital of Xiamen University (n=6, 3 patients per disease). In the endometrial window phase, the eutopic endometrial tissue was selected for microarray hybridization, and the DEGs and DELs of the two diseases were screened out, and the GO, KEGG pathway and TFs of DEGs were further analyzed by comprehensive bioinformatics methods. In addition, a ceRNA network was further constructed to identify the potential key genes. Excitingly, it is anticipated that the novel DEGs and pathways between FTI and endometriosis-related infertility identified in this study may help to seek potential molecular mechanisms. The workflow of this study was shown in Fig. 2A.

Materials And Methods

Ethics Statement

All procedures were performed according to standard protocols or the manufacturers' instructions and the study was approved by the ethics committees of the First Affiliated Hospital of Xiamen University.

Sample Collection

A total of 6 subjects (3 TFI patients and 3 endometriosis-related infertility patients) were recruited to participate, and all of the patients were of reproductive age with normal menstrual cycles. Herein, 3 patients with endometriosis were all patients whose endometrium occur in the ovary (group EMS). Meanwhile, 3 patients with TFI refer to women who have been diagnosed with tubal obstruction by fallopian tube angiography before (group TFI). In addition, all patients in our study gave informed consent. The follicular were monitored by ultrasonography from the ninth day of the menstrual cycle. On the 6th-7th day after ovulation (the LH test showed double line and the ovulation sign was detected by B ultrasonic examination), endometrium was obtained after these checks performed. The tissue was divided into two parts: the first part was fixed by 10% formaldehyde and was sent for pathological examination and endometrial lesions were excluded. The second part was stored in liquid nitrogen, and then the total RNA was extracted to explore lncRNA and mRNA expression profiles by human lncRNA Array v4.0.

RNA Extractionand Quality Control

According to the manufacturer's instructions, total RNA extracted from tissue was performed using Trizol reagent (Invitrogen). RNA quantity and quality were assessed by spectrophotometric measurement (Nanodrop ND-1000). Denaturing agarose gel electrophoresis was used to evaluate the RNA integrity. RNA was stored at -80°C for further experiments.

Microarray Analysis and Data Processing

Human LncRNA Microarray V4.0 (lncRNAs (40173) + coding genes (20730), 8 × 60 K Array Star, Rockville, MD, USA) designed for the global profiling of human lncRNAs and protein-coding transcripts was used in our study. The microarray can detect about 40,173 lncRNAs and 20,730 coding transcripts, which are derived from authoritative databases (RNA-seq, GENECODE, RefSeq, UCSC, UCR, ect) and other landmark publications. Sample labeling and array hybridization were performed in terms of the Agilent One-Color Microarray-Based Gene Expression Analysis protocol (Agilent Technology) with minor modification. In brief, mRNA samples from endometrium were purified to remove rRNA and were amplified and transcribed into fluorescent cRNA along the entire length of the transcripts without 3' bias utilizing a random priming method (Arraystar Flash RNA Labeling Kit, Arraystar). RNeasy Mini Kit (Qiagen) was employed to purify the labeled cRNA. After hybridization and washing, the arrays were scanned with Agilent DNA Microarray Scanner (part number G2505C). Raw data were extracted using Agilent Feature Extraction software (version 11.0.1.1). Furthermore, quantile normalization and subsequent data processing were done through GeneSpring GX v12.1 software package (Agilent Technologies). Log2 transformation was performed on the mRNA and lncRNA expression data. The average RNA expression value was used when duplicated data were found. Microarray hybridization and

expression data collection were performed by Kangchen Bio-tech, Shanghai, China. The microarray analysis was conducted in School of Medicine, Xiamen University, Xiamen, Fujian, China. Finally, we obtained lncRNA and mRNA expression profiles.

Identification of DEGs

Herein, R software and *limma* package (http://www.bioconductor.org/packages/release/bioc/html/limma.html) in Bioconductor were used to identify DEGs by comparing mRNA expression differences in eutopic endometrium tissues from TFI and endometriosis-related infertility respectively.[13] Normalized data of mRNAs in two groups of samples (TI and EMS) was shown in Fig. 1C. DEGs were identified by fold-change screening at a threshold of 2.0-fold or greater, and a *p*-value < 0.05.

Enrichment Analyses of DEGs

To understand the biological roles of these DEGs associated with TFI and endometriosis-related infertility, pathway analysis (based on Kyoto Encyclopedia of Genes and Genomes, http://www.genome.jp/kegg/) and Gene Ontology (GO) (http://www.geneontology.org) analysis were assessed by the clusterProfiler package in R software to explore the significant pathways related to DEGs14. adjust. P < 0.05 was used as thresholds to define markedly enriched GO terms/pathways. GO is a series of semantics used to describe the characteristics of genes. These semantics are mainly divided into three kinds: Biological Process (BP) , an orderly biological process with many steps, such as cell growth, differentiation and maintenance, apoptosis and signal transduction, Molecular Function (MF), which mostly refers to the functions of individual gene products, such as binding activity or catalytic activity, and Cellular Component (CC), which is used to describe the location of gene in cells, such as cytoplasm, nucleus, endoplasmic reticulum or mitochondria. KEGG is a database for systematic analysis of metabolic pathways of gene products in cells, and it is the most commonly used metabolic pathway analysis. ClusterProfiler[14] package (http://www.bioconductor.org/packages/release/bioc/html/clusterProfiler.html) in Bioconductor implements methods to analyze and visualize functional profiles (GO and KEGG) of gene and gene clusters.

TFs-DEGs Network Construction

Genes are often regulated by TFs at the transcription level. Therefore, DAVID (Database for Annotation, Visualization and Integrated Discovery) 6.8 (https://david.ncifcrf.gov/) was used for TFs enrichment analyses to investigate DEGs at the transcriptional level. Herein, the *p*-value (Benjamini) ≥ 0.05 was taken as the exclusion standard and the TFs-mRNAs network was constructed by Cytoscape[15] software. DAVID is a gene functional classification tool that integrates a set of functional annotation tools for investigators to analyze biological functions behind massive genes. Currently, the DAVID database is mainly used for the analysis of functions, pathways and TFs enrichment of genes.

Analysis of the Important Modules and Hub Genes

To analysis the important modules and Hub genes associated with endometriosis-related infertility, we constructed a PPI network of DEGs through the STRING 11.0 (https://string-db.org/) online database with the basic settings (meaning of network edges: evidence, minimum required interaction score: medium confidence) and visualized the results by Cytoscape software. Then, the important modules in the PPI network was analyzed by Molecular Complex Detection (MCODE) plug-in with the max depth, haircut on, node score cut-off, degree cut-off and k-core set as 100, 0.2, 0.2, 10 and 2, respectively. Gene ontology and pathway enrichment analysis of genes in each module (module 1, module 2 and module 3) were performed by DAVID database and the cut-off criterion of *p*-value (Benjamini) was set as < 0.05.

Moreover, in this study, we screened out the top 50 Hub genes in the PPI network by cytoHubba[16] plug-in and predicted their TFs by iRegulon[17] plug-in. The top 5 TFs in the NES scores were displayed. STRING, a database that predicts interactions between proteins or genes, includes both direct physical interactions and indirect functional correlations between proteins. MCODE plug-in can discover closely related regions in PPI network, which may represent molecular complexes. The MCC algorithm of CytoHubba plug-in can help us discover key genes and subnetwork in PPI network. IRegulon, a plug-in that predict TFs of gene sets, predicts TFs by calculating motif enrichment analysis and uses multiple position weight matrices (PWM) to score each motif.

Identification of DELs

The normalized expression of lncRNAs in two groups of samples (TI and EMS) was shown in Fig. 1D. Moreover, normalized data of lncRNAs was used for differential analysis by *limma* package in R software and the DELs were identified by fold-change screening at a threshold of 2.0-fold or greater, and a *p*-value < 0.05.

CeRNA Network Construction

It is well known that the DEGs is regulated by miRNAs and lncRNAs at transcriptional level. The combination of miRNAs and mRNAs can promote the degradation of mRNAs. At the same time, lncRNAs can competitively combine with mRNAs to further inhibit the degradation of mRNAs by miRNAs. The interaction network of lncRNA, miRNAs and mRNAs are called competing endogenous RNA (ceRNA) network. Herein, in order to better understand the physiological changes of eutopic endometrium between patients with TFI and endometriosis-related infertility, we used R software and Perl software to construct the ceRNA network (DEGs-miRNAs-DELs) through the miRDB database (miRNA-mRNA, http://mirdb.org/), miRTarBase database (miRNA-mRNA, http://mirtarbase.mbc.nctu.edu.tw/index.html), TargetScan database (miRNA-mRNA, http://targetscan.org/vert_72/) and miRcode database (lncRNA-miRNA, http://www.mircode.org/index.php) database.

Herein, first, DELs-miRNAs interaction pairs were integrated using the miRcode database. Second, target genes of miRNA signatures were obtained using three databases: miRDB, miTarBase, and TargetScan. Genes present in all three databases were regarded as target genes of miRNAs. Comparing predicted target genes of miRNAs with the DEGs, only the remaining overlapping genes and their interaction pairs

were used for further analysis. Then, the ceRNA networks (DEGs-miRNAs-DELs) were constructed by Cytoscape software according to DELs-miRNAs pairs and miRNAs-mRNAs (overlapping genes) pairs.

TFs Prediction of MRNA in CeRNA Network

MRNA and miRNA can be regulated by TFs at the transcription level. Herein, in order to predict the upstream regulatory mechanism of mRNAs related to endometriosis-related infertility, we predicted the TFs of mRNAs in ceRNA network by iRegulon plug-in in Cytoscape software and we only showed the top 10 TFs in the NES scores in this study.

Screening of Potential Key Genes

In order to screen out the potential key genes related to endometriosis-related infertility. In this study, we selected the lapping genes of important modules, Hub genes and ceRNA network as the potential key genes. At the same time, we also predicted the TFs of potential key genes and miRNAs (target to potential key genes) by iRegulon, respectively. The top 5 TFs in the NES scores were displayed.

TFBs Predition of Potential Key Genes

In this study, we predicted transcription factor binding sites (TFBs) for these potential key genes that bind to TFs by JASPAR 2020 online software (http://jaspar.genereg.net/). JASPAR[18] is a collection of transcription factor DNA-binding preferences, modeled as matrices. These can be converted into Position Weight Matrices (PWMs or PSSMs), used for scanning genomic sequences. JASPAR is an open-access database of curated, non-redundant transcription factor (TF) binding profiles stored as position frequency matrices (PFMs) and TF flexible models (TFFMs) for TFs across multiple species in six taxonomic groups. The higher the score, the greater the likelihood that the TF will bind to the DNA binding site (motif).

Results

MRNA and LncRNA Expression Profiles

Through comparing the quality of expression values of microarray before and after normalization, we found that the medians of expression value of 6 samples were in a straight line (Fig. 1A and Fig. 1B). Microarray analysis was used to assess the expression levels of mRNAs in eutopic endometrial tissue of TFI and endometriosis-related infertility. We identified 17965 mRNAs: 404 were up-regulated, 549 were down-regulated (FC $\geq$ 1.0, *p*-Value < 0.05), while 17012 were not differential expressed. The variation of mRNA expression between the TFI and endometriosis-related infertility is shown with a scatter plot (Fig. 1C).

Using microarray analysis, we identified 18912 lncRNAs in eutopic endometrial tissue of TFI and endometriosis-related infertility: 498 were up-regulated, 705 were down-regulated ($\geq$ 1.0 fold, p < 0.05),

while 17709 were not differential expressed. The variation of lncRNA expression between the TFI and endometriosis-related infertility is shown with a scatter plot (Fig. 1D).

Identification of DEGs

In this study, 508 DEGs (FC ≧ 2.0, *p*-Value < 0.05) between TFI and endometriosis-related infertility were screened out, including 203 up-regulated mRNAs and 305 down-regulated mRNAs. The volcano plots display the expression profile for all detected mRNAs and the top 10 genes with *p*-value from small to large are shown in Fig. 2B. In addition, the top 50 most significantly differentially expressed mRNAs are shown in Fig. 2C, and the clustering analysis revealed the relationships among mRNA expression patterns in different samples.

Enrichment Analyses of DEGs

GO functional enrichment analysis showed that DEGs were mainly involved in GO terms about transmembrane transporter activity (eg, passive transmembrane transporter activity, organic anion transmembrane transporter activity, G protein-coupled receptor binding, cation transmembrane transporter activity and so on; Fig. 3A). The gene transcription profile was significantly different between TFI and endometriosis-related infertility. KEGG pathway enrichment analysis were mainly enriched in immune system (eg, Cell adhesion molecules (CAMs), Intestinal immune network for IgA production, Viral protein interaction with cytokine and cytokine receptor and so on; Fig. 3B).

TFs-DEGs Network Construction

In this study, 52 TFs were enriched, including RSRFC4, LHX3, SRY, S8, and so on (Table 1). So, the visualization of TFs-mRNAs network was shown in Fig. 3C.

Analysis of the Important Modules and Hub Genes

A total of 495 nodes and 1040 edges were mapped in the PPI network of the identified DEGs and these DEGs demonstrated significant interactions of DEGs each other. In addition, 3 modules were selected by MCODE plug-in (Fig. 4A). As presented in Fig. 4B, GO functional enrichment analysis and KEGG pathway enrichment analysis of genes in each module showed that these genes were mainly enriched in immunity (eg, chemokine activity, cell chemotaxis, chemotaxis, inflammatory response as well as immune response and so on). The top 50 Hub genes screened out by cytoHubba were selected and sequentially ordered as table 2 (ranked by MCC). The TFs in the top 5 of NES were RAD21, NFIC, BCL6, RXRA and IKZF1, respectively; Fig. 4C).

Screening of DELs

576 DELs between TFI and endometriosis-related infertility were screened out, including 225 up-regulated lncRNAs and 351 down-regulated lncRNAs. Among these DELs, the top 10 DELs with *p*-value from small to large were shown in Fig. 2D. In addition, the top 50 most significantly differentially expressed lncRNAs

are shown in Fig. 2E, and the clustering analysis revealed the relationships among mRNA expression patterns in different samples.

Construction of the CeRNA Network commbining TFs and Screen of Potential Key Genes

A ceRNA network (21 mRNAs, 13 miRNAs and 16 lncRNAs) was constructed and is presented in Fig. 5A. At the same time, the top 10 TFs of mRNAs in ceRNA network were visualized in the ceRNA network combining TFs by Cytoscape software according to NES ranking. After counting the interaction numbers of each RNA, we found the lapping genes (potential key genes) of important modules, Hub genes and ceRNA network include 2 mRNAs (PLAU and LDLR). At the same time, the top 10 TFs of potential key genes and miRNAs were visualized according to NES ranking respectively. Herein, a new ceRNA network combining TFs based on potential key genes, including 2 mRNAs (PLAU and LDLR), 10 TFs (eg. SRF, FOS, which target to mRNA And HNF1A, POU3F2, which target to miRNA), 5 miRNAs (hsa-miR-301b-3p, hsa-miR-27a-3p, hsa-miR-20b-5p, hsa-miR-193a-3p and hsa-miR-17-5p) and 8 lncRNAs (eg. LMO7-AS1, ITFG3 and MGST1, etc.) (Fig. 5B and Fig. 5C) was shown in Fig. 5D, and the Transcription Factor Binding site (TFBs) of potential key genes was shown in table 3. Herein, we only show the binding sites with the highest score for each TF.

Discussion

TFI and endometriosis-related infertility are well known to all. However, the underlying molecular mechanisms between TFI and endometriosis-related infertility remain well unknown. In this study, we made use of eutopic endometrial tissue from TFI and endometriosis-related infertility patients with active disease to screen out the DEGs and DELs of the two diseases, and subsequently analyzed for their biological relevance by using GO, KEGG pathway. In addition, a ceRNA network was further constructed to identify the potential key genes. As a result, we found the DEGs between TFI and endometriosis-related infertility to identify the factors for the pathogenetic or therapeutic relevance in endometriosis-related infertility, including “PLAU” “LDLR”. At the same time, the TFs of these 2 potential key genes were “IKZF2”, “TFAP4”, “HSF1”, “FOS”, “SRF”.

Studies are discussed that the high expression of “LDLR” gene leads to the occurrence of hyperlipidemia,[19,20] which further interferes with endothelium-related vasomotor contraction, leading to the disorder of vascular hemodynamic regulation.[21] By stimulating the release of platelet-related p-selectin, it induces leukocyte adhesion, and increases blood viscosity and blood flow resistance.[22] Secondly, the activation of neutrophils can extend pseudopodia and adhesion with other cells (such as white blood cells, endothelial cells, platelets, etc.), blocking the microvessels.[23,24] A large number of studies have shown that inflammation response plays an important role in the pathogenesis of endometriosis-related infertility.[25-28] For example, endometriosis leads to pelvic inflammatory response and produces a variety of cytokines, which can gobble up sperm or fertilized eggs, inhibit oviduct motility, and thus interfere with fertility. Another key gene “PLAU” gene located on 10q22.2 encodes urokinase type plasminogen activator (uPA), transforming plasminogen into plasminogen.[29] And previous studies have

shown that uPA protein is mainly located in the cytoplasm, and both ectopic and eutopic endometrium express uPA protein,[30] which is positively expressed in vascular endothelial cells.[31] It is involved in the occurrence and development of endometriosis.

Making mention of the TFs of these 2 potential key genes, IKZF2 is one of the huge TFs that regulate the development of the lymphatic system as well as plays an important role in the development of T, B lymphocytes and NK cells.[32] FOS can effectively reduce the amount of serum cholesterol, triglycerides, free fatty acids.[33]

TFAP4 is an important downstream factor of c-MYC and widely regulates cell proliferation and differentiation, cell cycle, apoptosis, immune response and other processes. There have been growing evidences that elevated TFAP4 expression significantly correlates with tumor progression and poor prognosis in a number of malignancies, including colorectal cancer,[34,35] gastric cancer[36] and non-small cell lung cancer.[37] As for the mechanism of TFAP4 in different diseases, studies showed that TFAP4 overexpression promoted Wnt/β-catenin pathway activation,[38] PI3K/AKT signaling pathway activation,[39] p53 pathway activation[40] and WNT signaling activation[41] etc. However, whether TFAP4 plays a role in endometriosis-related infertility is challenging.

HSF1 is the master transcriptional regulator of the heat shock response in eukaryotes,[42] which is a signaling pathway in cells that regulates the expression of molecular chaperones in the presence of thermal and other environmental stresses.[43] Research clearly shows that Hsp90 is a major repressor of HSF1 gene. HSF1 has been shown to contribute to cancer progression,[44] organismal lifespan[45] and protect against obesity.[46] By contrast, a lack of HSF1 activity has been suggested to contribute to neurodegenerative diseases.[47] Meanwhile, some studies suggested that heat shock protein may be attached importance to the pathogenesis of EMS

SRF, a member of the MADS transcription factor family, regulates the transcription of a series of genes by binding to serum response element (SRE). It mainly shows rapid gene response to ischemia, hypoxia and other stimuli, and affects cell development, differentiation, proliferation and apoptosis, including mesoderm formation,[48] cardiac development,[49] angiogenesis, oligodendrocyte differentiation,[50] neuronal migration,[51] and circadian regulation.[52] Mitogen-activated protein kinase (MAPK) and RhoA signaling pathway are two important signaling pathways that activate intracellular SRF expression and can promote the transcription of growth-related genes and cell cycle genes. In addition, SRF also plays an important role in the biodynamics of the actin skeleton and participates in the regulation of cell migration ability.

Those crucial genes and TFs could reflect the key biomarkers with mechanistic relevance for endometriosis-related infertility pathogenesis and progression, some of them well established and already exploited for therapeutic purposes. Herein, those DEGs and DELs could contribute to promote the diagnostic and therapeutic interest in endometriosis-related infertility on relevant and innovative

molecules or pathways, which could indicate a new direction of the acquaintance of endometriosis-related infertility. Based the DAVID database, in the GO and KEGG pathway enrichment analyses of DEGs the domain terms including transmembrane transporter activity and immune system, indicating that endometriosis-related infertility is characterized by transmembrane transporter and immune processes in the process of disease. Besides, the TFs of miR in the ceRNA were mainly enriched in “MKX”, “EN1”, “HNF1A”, “POU3F2” and “HDAC2”.

MKX is a member of the three-amino-acid loop (TALE) superclass of atypical homeobox genes belonging to the Iroquois family.[53] MKX is expressed in the embryonic progenitor cell populations of the cartilage, skeletal muscle, tendons and bones, in addition to the tips of the ureteric buds in the metanephric kidneys and the testis cords of the male gonad testis in the mouse embryo.[54-57] Recently, research revealed that MKX was a central transcription factor that regulated AF development, maintenance and regeneration. Meanwhile, MKX is required for the maturation of collagen fibrils and plays a crucial role in tendon maturation by regulating the expression of type I collagen.[58] Additionally, activation of Wnt/β-catenin signaling reduces gene expressions of *Mkx.*[59]

EN1, a neural-specific transcription factor, plays a crucial role in the development of many tissues and organs.[60] It can promote cell survival and cell resistance to apoptotic stress, thereby promoting dopaminergic neuronal-cell longevity throughout adulthood.[61] EN1 is highly expressed in many tumors, such as breast tumors, salivary gland adenoid cystic carcinoma, and adenoid cystic carcinoma.[62]

HNF1A, hepatocyte nuclear factor 1 homeobox A, was subsequently verified to be expressed in several organs including the liver, pancreas, kidney, and intestine.[63] In addition, HNF1A has been reported to inhibit Wnt and NF-κB signalling during hepatocarcinogenesis and hepatocellular carcinoma metastasis by transcriptionally regulating the expression of miR-194.[64] Study also revealed that STAT3-induced upregulation of HNF1A-AS1 promoted OSCC progression by activating NOTCH signaling pathway.[65]POU3F2 belongs to a large family of brain-specific homeobox transcription factors and is expressed in the central nervous system (CNS) during neuronal development and in adult brain. POU3F2 has been shown to regulate the expression levels of critical genes at different stages of neural differentiation, the migration of cortical neuron, the neurogenesis and positioning of cortical neurons, etc.[66] In addition, it is also involved in different types of cancer, such as melanoma, lung cancer, stomach cancer.[67] It is concluded that POU3f2 affects the NF-κB, MAPK/JNK and Myc/Max pathways in cancer pathway reporter array.

HDAC2, a subtype of histone deacetylase (HDAC) that is a group of enzymes that regulate lysine acetylation and thereby protein function, mainly plays a positive role in cardiac hypertrophy. Gradually, HDAC2 was found to be aberrantly expressed in several types of cancer including gastric, colorectal, prostate and Hodgkin’s Lymphoma. Later, study has reported that HDAC2 was recruited to *CDKN1A* promoter by FOXO3a and regulates p21 expression in cerebellar granule neuron.[68] In

hepatocellular carcinoma, it also was shown to be related to P21. Meanwhile, some researchers have found that sumoylation of HDAC2 promotes NF-κB-dependent gene expression.[69]

In conclusion, according to integrated bioinformatical analyses of multiple cohort's profile datasets in the eutopic endometrial tissue between TFI and endometriosis-related infertility patients, we identified the 508 DEGs. Other studies could also use some methods to explore the releation of other genes and endometriosis-related infertility, and futher gain new insights into important biological processes. But our study erplored the potential key gene associated with endometriosis-related infertility and the relevant and innovative molecules or pathways, which could indicate a new direction of the acquaintance of endometriosis-related infertility. As for the relationships between these TFs and endometriosis-related infertility, little is known and more studies are needed. In a word, these candidate genes and pathways might provide novel insight into its occurrence of endometriosis-related infertility, and be therapeutic targets and biomarkers for endometriosis-related infertility treatment.

Conclusion

In conclusion, physiological of TFI and endometriosis related infertility may be related to the transmembrane transporter activity and the immune system. These potential key RNAs and TFs might have clinical utility for the diagnosis and prognosis prediction in TFI and endometriosis-related infertility. The results of the current study might lay the foundation for future basic and clinical research.

Declarations

Funding

This work was financially supported by National Key R&D Program of China (SQ2017YFSF080005), the National Science Foundation of China (No. 81871145 and No. 81701419). The authors have declared that no competing interests exist.

Availability of Data and Materials

The expression data associated with this study were sequenced by endometrial transcriptome microarray in 6 volunteers.

Authors' contributions

JZL and LLR contributed equally to this work. They mainly achieved all the experimentation. CY, RFW and ZXH participated in data analysis and discussion. QHC organized and supervised the project. All authors read and approved the final manuscript.

Ethics Approval and Consent to Participate

All procedures were performed according to standard protocols or the manufacturers' instructions and the study was approved by the ethics committees of the First Affiliated Hospital of Xiamen University.

Consent for Publication

Not applicable

Competing Interests

The authors declare that they have no competing interests.

References

1. Shen J, Kang Q, Mao Y, et al. (2020) Urinary bisphenol A concentration is correlated with poorer oocyte retrieval and embryo implantation outcomes in patients with tubal factor infertility undergoing in vitro fertilisation. *Ecotoxicol Environ Saf* 187:109816. https://doi.org/10.1016/j.ecoenv.2019.109816
2. Czyzyk A, Podfigurna A, Szeliga A, Meczekalski B (2017) Update on endometriosis pathogenesis. *Minerva Ginecol* 69(5):447-461. https://doi.org/10.23736/S0026-4784.17.04048-5
3. Ding C, Zou Q, Ding J, et al. (2018) Increased N6-methyladenosine causes infertility is associated with FTO expression. *J Cell Physiol* 233(9):7055-7066. https://doi.org/10.1002/jcp.26507
4. Petousis S, Prapas Y, Margioula-Siarkou C, et al. (2018) Unexplained infertility patients present the mostly impaired levels of progesterone receptors: Prospective observational study. *Am J Reprod Immunol* 79(6). https://doi.org/ARTN e12828 10.1111/aji.12828
5. Altmae S, Martinez-Conejero JA, Salumets A, Simon C, Horcajadas JA, Stavreus-Evers A (2010) Endometrial gene expression analysis at the time of embryo implantation in women with unexplained infertility. *Mol Hum Reprod* 16(3):178-187. https://doi.org/10.1093/molehr/gap102
6. Margioula-Siarkou C, Prapas Y, Petousis S, et al. (2017) LIF endometrial expression is impaired in women with unexplained infertility while LIF-R expression in all infertility sub-groups. *Cytokine* 96:166-172. https://doi.org/10.1016/j.cyto.2017.04.009
7. Ren T, Wang S, Sun J, et al. (2017) Endometriosis is the independent prognostic factor for survival in Chinese patients with epithelial ovarian carcinoma. *J Ovarian Res* 10. https://doi.org/ARTN 67 10.1186/s13048-017-0363-y
8. Hong LJ, Liu RZ, Qiao XW, et al. (2019) Differential microRNA Expression in Porcine Endometrium Involved in Remodeling and Angiogenesis That Contributes to Embryonic Implantation. *Front Genet* 10. https://doi.org/ARTN 661 10.3389/fgene.2019.00661
9. Cousins FL, Dorien FO, Gargett CE (2018) Endometrial stem/progenitor cells and their role in the pathogenesis of endometriosis. *Best Pract Res Cl Ob* 50:27-38. https://doi.org/10.1016/j.bpobgyn.2018.01.011
10. Gao S (2018) Data Analysis in Single-Cell Transcriptome Sequencing. *Methods Mol Biol* 1754:311-326. https://doi.org/10.1007/978-1-4939-7717-8_18

11. Li WC, Bai DL, Xu Y, et al. (2019) Identification of differentially expressed genes in synovial tissue of rheumatoid arthritis and osteoarthritis in patients. *J Cell Biochem* 120(3):4533-4544. https://doi.org/10.1002/jcb.27741
12. Chen FL, Huang CL, Wu QM, Jiang LL, Chen ST, Chen LY (2019) Circular RNAs expression profiles in plasma exosomes from early-stage lung adenocarcinoma and the potential biomarkers. *J Cell Biochem*. https://doi.org/10.1002/jcb.29475
13. Ritchie ME, Phipson B, Wu D, et al. (2015) limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 43(7). https://doi.org/ARTN e47 10.1093/nar/gkv007
14. Yu GC, Wang LG, Han YY, He QY (2012) clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. *Omics* 16(5):284-287. https://doi.org/10.1089/omi.2011.0118
15. Shannon P, Markiel A, Ozier O, et al. (2003) Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res* 13(11):2498-2504. https://doi.org/10.1101/gr.1239303
16. Chin CH, Chen SH, Wu HH, Ho CW, Ko MT, Lin CY (2014) cytoHubba: identifying hub objects and sub-networks from complex interactome. *Bmc Syst Biol* 8. https://doi.org/Artn S11 10.1186/1752-0509-8-S4-S11
17. Janky R, Verfaillie A, Imrichova H, et al. (2014) iRegulon: From a Gene List to a Gene Regulatory Network Using Large Motif and Track Collections. *Plos Comput Biol* 10(7). https://doi.org/ARTN e1003731 10.1371/journal.pcbi.1003731
18. Bryne JC, Valen E, Tang MHE, et al. (2008) JASPAR, the open access database of transcription factor-binding profiles: new content and tools in the 2008 update. *Nucleic Acids Res* 36:D102-D106. https://doi.org/10.1093/nar/gkm955
19. Ishibashi S, Brown MS, Goldstein JL, Gerard RD, Hammer RE, Herz J (1993) Hypercholesterolemia in Low-Density-Lipoprotein Receptor Knockout Mice and Its Reversal by Adenovirus-Mediated Gene Delivery. *J Clin Invest* 92(2):883-893. https://doi.org/Doi 10.1172/Jci116663
20. Franciosi S, Sosa MAG, English DF, et al. (2009) Novel cerebrovascular pathology in mice fed a high cholesterol diet. *Mol Neurodegener* 4. https://doi.org/Artn 42 10.1186/1750-1326-4-42
21. Stokes KY, Cooper D, Tailor A, Granger DN (2002) Hypercholesterolemia promotes inflammation and microvascular dysfunction: Role of nitric oxide and superoxide. *Free Radical Bio Med* 33(8):1026-1036. https://doi.org/Pii S0891-5849(02)01015-8 Doi 10.1016/S0891-5849(02)01015-8
22. Stokes KY, Calahan L, Russell JM, Gurwara S, Granger DN (2006) Role of platelets in hypercholesterolemia-induced leukocyte recruitment and arteriolar dysfunction. *Microcirculation* 13(5):377-388. https://doi.org/10.1080/10739680600745877
23. Zhang XY, Cheng R, Rowe D, et al. (2014) Shear-Sensitive Regulation of Neutrophil Flow Behavior and Its Potential Impact on Microvascular Blood Flow Dysregulation in Hypercholesterolemia. *Arterioscl Throm Vas* 34(3):587-593. https://doi.org/10.1161/Atvbaha.113.302868

24. Lipowsky HH (2005) Microvascular rheology and hemodynamics. *Microcirculation* 12(1):5-15. https://doi.org/10.1080/10739680590894966
25. Wang XM, Ma ZY, Song N (2018) Inflammatory cytokines IL-6, IL-10, IL-13, TNF-alpha and peritoneal fluid flora were associated with infertility in patients with endometriosis. *Eur Rev Med Pharmaco* 22(9):2513-2518.
26. Jorgensen H, Hill AS, Beste MT, et al. (2017) Peritoneal fluid cytokines related to endometriosis in patients evaluated for infertility. *Fertil Steril* 107(5):1191-+. https://doi.org/10.1016/j.fertnstert.2017.03.013
27. Ahn SH, Edwards AK, Singh SS, Young SL, Lessey BA, Tayade C (2015) IL-17A Contributes to the Pathogenesis of Endometriosis by Triggering Proinflammatory Cytokines and Angiogenic Growth Factors. *J Immunol* 195(6):2591-2600. https://doi.org/10.4049/jimmunol.1501138
28. Margari KM, Zafiropoulos A, Hatzidaki E, Giannakopoulou C, Arici A, Matalliotakis I (2013) Peritoneal fluid concentrations of beta-chemokines in endometriosis. *Eur J Obstet Gyn R B* 169(1):103-107. https://doi.org/10.1016/j.ejogrb.2013.02.010
29. Brungs D, Chen JL, Aghmesheh M, et al. (2017) The urokinase plasminogen activation system in gastroesophageal cancer: A systematic review and meta-analysis. *Oncotarget* 8(14):23099-23109. https://doi.org/10.18632/oncotarget.15485
30. Mahmood N, Mihalcioiu C, Rabbani SA (2018) Multifaceted Role of the Urokinase-Type Plasminogen Activator (uPA) and Its Receptor (uPAR): Diagnostic, Prognostic, and Therapeutic Applications. *Front Oncol* 8:24. https://doi.org/10.3389/fonc.2018.00024
31. Bruse C, Guan YM, Carlberg M, Carlstrom K, Bergqvist A (2005) Basal release of urokinase plasminogen activator, plasminogen activator inhibitor-1, and soluble plasminogen activator receptor from separated and cultured endometriotic and endometrial stromal and epithelial cells. *Fertil Steril* 83:1155-1160. https://doi.org/10.1016/j.fertnstert.2004.09.033
32. Chan SM (2019) The Making of a Leukemic Stem Cell: A Novel Role for IKZF2 in AML Stemness and Differentiation. *Cell Stem Cell* 24(1):5-6. https://doi.org/10.1016/j.stem.2018.12.007
33. Costa GT, de Abreu GC, Guimaraes ABB, de Vasconcelos PRL, Guimaraes SB (2015) Fructo-oligosaccharide effects on serum cholesterol levels. An overview. *Acta Cir Bras* 30(5):366-370. https://doi.org/10.1590/S0102-865020150050000009
34. Jackstadt R, Roh S, Neumann J, et al. (2013) AP4 is a mediator of epithelial-mesenchymal transition and metastasis in colorectal cancer. *J Exp Med* 210(7):1331-1350. https://doi.org/10.1084/jem.20120812
35. Cao J, Tang MH, Li WLL, et al. (2009) Upregulation of Activator Protein-4 in Human Colorectal Cancer With Metastasis. *Int J Surg Pathol* 17(1):16-21. https://doi.org/10.1177/1066896908315813
36. Liu XH, Zhang B, Guo Y, et al. (2012) Down-Regulation of AP-4 Inhibits Proliferation, Induces Cell Cycle Arrest and Promotes Apoptosis in Human Gastric Cancer Cells. *Plos One* 7(5). https://doi.org/ARTN e37096 10.1371/journal.pone.0037096

37. Gong H, Han SR, Yao H, Zhao H, Wang YM (2014) AP-4 predicts poor prognosis in non-small cell lung cancer. *Mol Med Rep* 10(1):336-340. https://doi.org/10.3892/mmr.2014.2209
38. Song J, Xie C, Jiang L, et al. (2018) Transcription factor AP-4 promotes tumorigenic capability and activates the Wnt/beta-catenin pathway in hepatocellular carcinoma. *Theranostics* 8(13):3571-3583. https://doi.org/10.7150/thno.25194
39. Huang T, Chen QF, Chang BY, et al. (2019) TFAP4 Promotes Hepatocellular Carcinoma Invasion and Metastasis via Activating the PI3K/AKT Signaling Pathway. *Dis Markers*. https://doi.org/Artn 7129214 10.1155/2019/7129214
40. Huang Q, Raya A, DeJesus P, et al. (2004) Identification of p53 regulators by genome-wide functional analysis. *Proc Natl Acad Sci U S A* 101(10):3456-3461. https://doi.org/10.1073/pnas.0308562100
41. Lebensohn AM, Dubey R, Neitzel LR, et al. (2016) Comparative genetic screens in human cells reveal new regulatory mechanisms in WNT signaling. *Elife* 5. https://doi.org/ARTN e21459 10.7554/eLife.21459
42. Anckar J, Sistonen L (2011) Regulation of HSF1 Function in the Heat Stress Response: Implications in Aging and Disease. *Annu Rev Biochem* 80:1089-1115. https://doi.org/10.1146/annurev-biochem-060809-095203
43. Richter K, Haslbeck M, Buchner J (2010) The Heat Shock Response: Life on the Verge of Death. *Mol Cell* 40(2):253-266. https://doi.org/10.1016/j.molcel.2010.10.006
44. Dai C, Whitesell L, Rogers AB, Lindquist S (2007) Heat shock factor 1 is a powerful multifaceted modifier of carcinogenesis. *Cell* 130(6):1005-1018. https://doi.org/10.1016/j.cell.2007.07.020
45. Hsu AL (2003) Regulation of aging and age-related disease by DAF-16 and heat-shock factor (vol 300, pg 1142, 2003). *Science* 300(5628):2033-2033.
46. Ma X, Xu L, Alberobello AT, et al. (2015) Celastrol Protects against Obesity and Metabolic Dysfunction through Activation of a HSF1-PGC1alpha Transcriptional Axis. *Cell Metab* 22(4):695-708. https://doi.org/10.1016/j.cmet.2015.08.005
47. Neef DW, Jaeger AM, Thiele DJ (2011) Heat shock transcription factor 1 as a therapeutic target in neurodegenerative diseases. *Nat Rev Drug Discov* 10(12):930-944. https://doi.org/10.1038/nrd3453
48. Arsenian S, Weinhold B, Oelgeschlager M, Ruther U, Nordheim A (1998) Serum response factor is essential for mesoderm formation during mouse embryogenesis. *Embo J* 17(21):6289-6299. https://doi.org/DOI 10.1093/emboj/17.21.6289
49. Parlakian A, Tuil D, Hamard G, et al. (2004) Targeted inactivation of serum response factor in the developing heart results in myocardial defects and embryonic lethality. *Mol Cell Biol* 24(12):5281-5289. https://doi.org/10.1128/Mcb.24.12.5281-5289.2004
50. Stritt C, Stern S, Harting K, et al. (2009) Paracrine control of oligodendrocyte differentiation by SRF-directed neuronal gene expression. *Nat Neurosci* 12(4):418-427. https://doi.org/10.1038/nn.2280
51. Alberti S, Krause SM, Kretz O, et al. (2005) Neuronal migration in the murine rostral migratory stream requires serum response factor. *P Natl Acad Sci USA* 102(17):6148-6153. https://doi.org/10.1073/pnas.0501191102

52. Gerber A, Esnault C, Aubert G, Treisman R, Pralong F, Schibler U (2013) Blood-Borne Circadian Signal Stimulates Daily Oscillations in Actin Dynamics and SRF Activity. *Cell* 152(3):492-503. https://doi.org/10.1016/j.cell.2012.12.027
53. Anderson DM, Arredondo J, Hahn K, et al. (2006) Mohawk is a novel homeobox gene expressed in the developing mouse embryo. *Dev Dyn* 235(3):792-801. https://doi.org/10.1002/dvdy.20671
54. Otabe K, Nakahara H, Hasegawa A, et al. (2015) Transcription Factor Mohawk Controls Tenogenic Differentiation of Bone Marrow Mesenchymal Stem Cells In Vitro and In Vivo. *J Orthop Res* 33(1):1-8. https://doi.org/10.1002/jor.22750
55. Nakahara H, Hasegawa A, Otabe K, et al. (2013) Transcription Factor Mohawk and the Pathogenesis of Human Anterior Cruciate Ligament Degradation. *Arthritis Rheum-Us* 65(8):2081-2089. https://doi.org/10.1002/art.38020
56. Liu WJ, Watson SS, Lan Y, et al. (2010) The Atypical Homeodomain Transcription Factor Mohawk Controls Tendon Morphogenesis. *Mol Cell Biol* 30(20):4797-4807. https://doi.org/10.1128/Mcb.00207-10
57. Ito Y, Toriuchi N, Yoshitaka T, et al. (2010) The Mohawk homeobox gene is a critical regulator of tendon differentiation. *P Natl Acad Sci USA* 107(23):10538-10542. https://doi.org/10.1073/pnas.1000525107
58. Nakamichi R, Ito Y, Inui M, et al. (2016) Mohawk promotes the maintenance and regeneration of the outer annulus fibrosus of intervertebral discs. *Nat Commun* 7:12503. https://doi.org/10.1038/ncomms12503
59. Kishimoto Y, Ohkawara B, Sakai T, et al. (2017) Wnt/beta-catenin signaling suppresses expressions of Scx, Mkx, and Tnmd in tendon-derived cells. *Plos One* 12(7):e0182051. https://doi.org/10.1371/journal.pone.0182051
60. Izquierdo C, Alentorn A, Idbaih A, et al. (2018) Long-term impact of temozolomide on 1p/19q-codeleted low-grade glioma growth kinetics. *J Neuro-Oncol* 136(3):533-539. https://doi.org/10.1007/s11060-017-2677-4
61. Alvarez-Fischer D, Fuchs J, Castagner F, et al. (2011) Engrailed protects mouse midbrain dopaminergic neurons against mitochondrial complex I insults. *Nat Neurosci* 14(10):1260-U1182. https://doi.org/10.1038/nn.2916
62. Drier Y, Cotton MJ, Williamson KE, et al. (2016) An oncogenic MYB feedback loop drives alternate cell fates in adenoid cystic carcinoma. *Nat Genet* 48(3):265-272. https://doi.org/10.1038/ng.3502
63. de Vries AGM, Bakker-van Waarde WM, Dassel ACM, et al. (2015) A novel phenotype of a hepatocyte nuclear factor homeobox A (HNF1A) gene mutation, presenting with neonatal cholestasis. *J Hepatol* 63(5):1295-1296. https://doi.org/DOI 10.1016/j.jhep.2015.08.005
64. Bao C, Li Y, Huan L, et al. (2015) NF-kappaB signaling relieves negative regulation by miR-194 in hepatocellular carcinoma by suppressing the transcription factor HNF-1alpha. *Sci Signal* 8(387):ra75. https://doi.org/10.1126/scisignal.aaa8441

65. Liu Z, Li H, Fan SM, Lin H, Lian WW (2019) STAT3-induced upregulation of long noncoding RNA HNF1A-AS1 promotes the progression of oral squamous cell carcinoma via activating Notch signaling pathway. *Cancer Biol Ther* 20(4):444-453. https://doi.org/10.1080/15384047.2018.1529119
66. Dominguez MH, Ayoub AE, Rakic P (2013) POU-III Transcription Factors (Brn1, Brn2, and Oct6) Influence Neurogenesis, Molecular Identity, and Migratory Destination of Upper-Layer Cells of the Cerebral Cortex. *Cereb Cortex* 23(11):2632-2643. https://doi.org/10.1093/cercor/bhs252
67. Chen HY, Lee YH, Chen HY, Yeh CA, Chueh PJ, Lin YMJ (2016) Capsaicin Inhibited Aggressive Phenotypes through Downregulation of Tumor-Associated NADH Oxidase (tNOX) by POU Domain Transcription Factor POU3F2. *Molecules* 21(6). https://doi.org/ARTN 733 10.3390/molecules21060733
68. Peng S, Zhao S, Yan F, et al. (2015) HDAC2 selectively regulates FOXO3a-mediated gene transcription during oxidative stress-induced neuronal cell death. *J Neurosci* 35(3):1250-1259. https://doi.org/10.1523/JNEUROSCI.2444-14.2015
69. Wagner T, Kiweler N, Wolff K, et al. (2015) Sumoylation of HDAC2 promotes NF-kappaB-dependent gene expression. *Oncotarget* 6(9):7123-7135. https://doi.org/10.18632/oncotarget.3344

Tables

Table 1 TFs enrichment of DEGs

Term	Count	%	P-value	P-value (Benjamini)
RSRFC4	242	0.348432056	2.61E-08	4.59E-06
LHX3	192	0.276441962	4.68E-07	4.12E-05
SRY	211	0.303798197	1.32E-06	7.74E-05
S8	223	0.32107582	6.42E-06	2.82E-04
HNF3B	203	0.292279782	1.23E-05	4.32E-04
FOXJ2	306	0.440579376	2.07E-05	6.08E-04
NFAT	204	0.293719584	2.19E-05	5.49E-04
GATA6	94	0.135341377	3.21E-05	7.05E-04
HFH1	215	0.309557405	4.98E-05	9.73E-04
TBP	175	0.25196533	7.51E-05	0.001321224
CHX10	200	0.287960377	9.16E-05	0.001464945
POU6F1	203	0.292279782	1.21E-04	0.001775955
STAT	171	0.246206122	1.75E-04	0.002369245
NKX22	210	0.302358395	2.15E-04	0.002702676
GATA	211	0.303798197	2.49E-04	0.002914991
SOX9	207	0.29803899	5.35E-04	0.005865329
SOX5	222	0.319636018	5.79E-04	0.00598294
FOXO1	194	0.279321565	6.30E-04	0.006141281
HFH3	191	0.27500216	7.37E-04	0.006809
CDC5	209	0.300918594	7.84E-04	0.006877164
FREAC2	156	0.224609094	8.53E-04	0.007128416
MEF2	330	0.475134621	0.001034386	0.008245191
MEIS1	213	0.306677801	0.001218986	0.009290158
BACH2	197	0.283640971	0.001882725	0.013724613
HSF2	192	0.276441962	0.001965924	0.013758208
FREAC4	209	0.300918594	0.002021252	0.013602803
FREAC7	221	0.318196216	0.002040433	0.013225948
HNF1	251	0.361390273	0.002129988	0.013313358
FOXD3	160	0.230368301	0.002135593	0.012890891
NKX61	198	0.285080773	0.002236048	0.013046976
IRF7	206	0.296599188	0.002661678	0.015017702
POU3F2	265	0.381547499	0.002848271	0.015565432
TATA	243	0.349871858	0.003916408	0.020711036
HLF	184	0.264923547	0.004254459	0.021828302
BACH1	224	0.322515622	0.004582285	0.022830648
MEIS1AHOXA9	160	0.230368301	0.004985616	0.024138966
AP1	263	0.378667895	0.005554005	0.026144836
PBX1	255	0.36714948	0.005625557	0.025790371
HSF1	140	0.201572264	0.00660878	0.029479958
OCT	206	0.296599188	0.007100544	0.030867407
CDP	264	0.380107697	0.008230457	0.034855009
CART1	211	0.303798197	0.008497903	0.035130508
MSX1	186	0.26780315	0.009824231	0.039604049
FOXO3	127	0.182854839	0.010265814	0.04043525
IRF1	117	0.16845682	0.011240005	0.043246826
TGIF	188	0.270682754	0.011353536	0.042747516
GFI1	196	0.282201169	0.012117478	0.044626822
P53	286	0.411783339	0.012512908	0.045120506
NKX25	277	0.398825122	0.012696098	0.044857056
MEIS1BHOXA9	221	0.318196216	0.013111784	0.045396043
HAND1E47	204	0.293719584	0.013480418	0.045757134
STAT5B	149	0.214530481	0.013547166	0.045115953

Table 2 The first 50 genes provided by MCC method

Rank	Name	Score
1	CXCL12	4.81E+08
2	CXCL2	4.81E+08
3	CCL19	4.81E+08
4	CX3CR1	4.81E+08
5	CCL20	4.81E+08
6	SAA1	4.80E+08
7	PENK	4.79E+08
8	SST	4.79E+08
9	PTGER3	4.79E+08
10	ADRA2A	4.79E+08
10	ADRA2C	4.79E+08
12	NPY1R	4.79E+08
13	TAS2R1	4.79E+08
14	IL6	2235939
15	CCL2	1830768
16	VCAM1	1829852
17	TLR2	1504554
18	TLR1	726654
19	CCL8	725760
20	SELP	368881
21	CCL18	362887
22	SERPINE1	10208
23	HMOX1	7704
24	LCN2	6732
25	LDLR	1607
26	F2	1484
27	REN	1364
28	PLAT	1325
29	PLAU	1212
30	IL33	740
31	SCARB1	727
32	CCL23	720
33	POSTN	554
34	DPP4	495
35	TIMP3	490
36	NT5E	328
37	CYR61	268
38	FGF1	251
39	NRXN1	162
39	FGF7	162
41	PROM1	158
42	IGFBP5	156
43	TGFB2	144
44	HLA-DQA1	138
45	GRIA1	134
45	CYSLTR1	134
47	PIK3R2	133
48	S100A8	129
49	P2RX7	127
50	NMB	126

Table 3 Prediction of TFBs of key crucial genes

mRNA	Transcript	TF	TFBs	Score
PLAU	NM_001145031	IKZF2	\	\
		TFAP4	atcagctgcg	9.35378
		HSF1	attatggaaacttcc	11.6495
		FOS	atgaatcatg	11.3508
		SRF	gcccagagatgc	6.47679
LDLR	NM_000527	HSF1	cctcctgaaatttct	11.141
		FOS	atgagtcccc	8.1038
		SRF	gtccacagaagg	7.04138

Figures

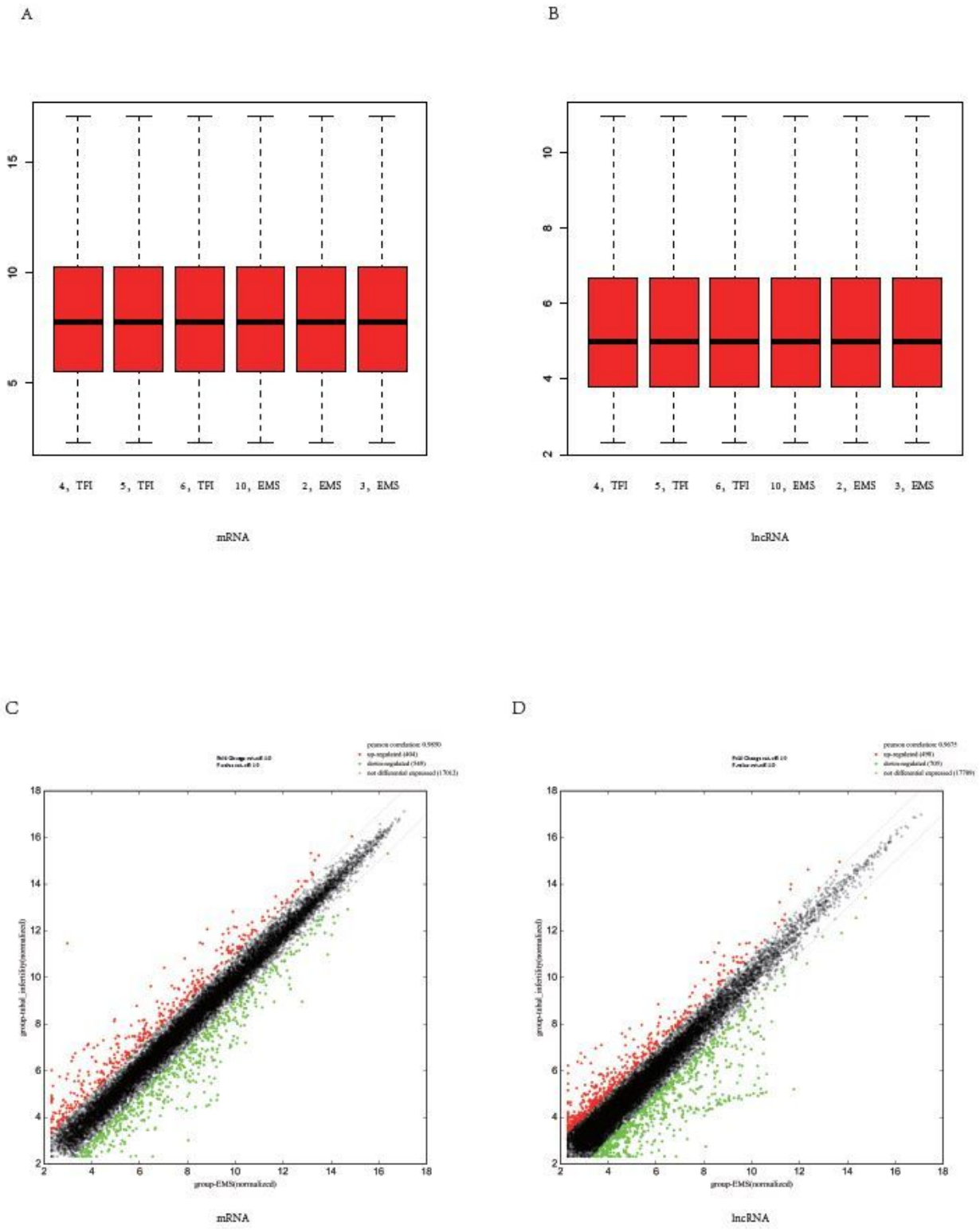

Figure 1

Normalization of data and their expression variation Notes: (A) Normalization of raw data on mRNAs and lncRNAs. (B) The mRNA and lncRNA expression variation between TFI and endometriosis-related infertility. The values shown on the x-axis and y-axis in the scatter plot are the normalized signal values of each sample (log2 scale). The green lines are fold-change lines (the default fold-change value given is

2.0). The lncRNAs (mRNAs) above the top green line and below the bottom green line showed an FC > 2.0 in expression between the 2 compared samples.

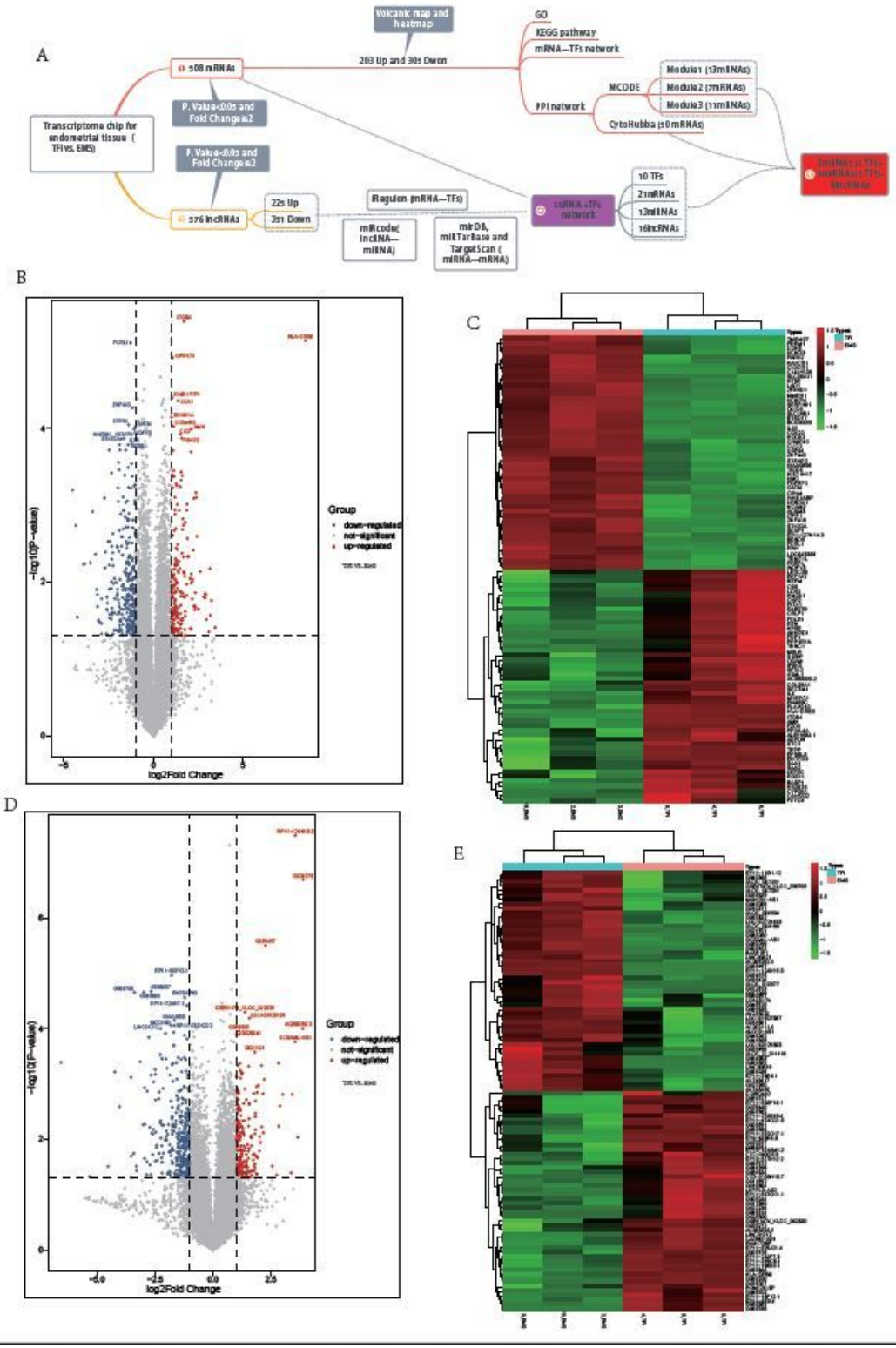

Figure 2

Workflow of this study and expression of mRNAs and lncRNAs Notes: (A) Workflow of this study. (B) Expression of mRNAs between two sets of samples (TFI VS. endometriosis-related infertility). (C) Hierarchical clustering heatmap of DEGs (Top 50 up-regulated and top 50 down-regulated with FC). (D)

Expression of lncRNAs between two sets of samples (TFI VS. endometriosis-related infertility). (E) Hierarchical clustering heatmap of DELs (Top 50 up-regulated and top 50 down-regulated with FC). Firebrick represent the relative expression of mRNAs or lncRNAs is up-regulated and navy represent the relative expression of mRNAs or lncRNAs is down-regulated.

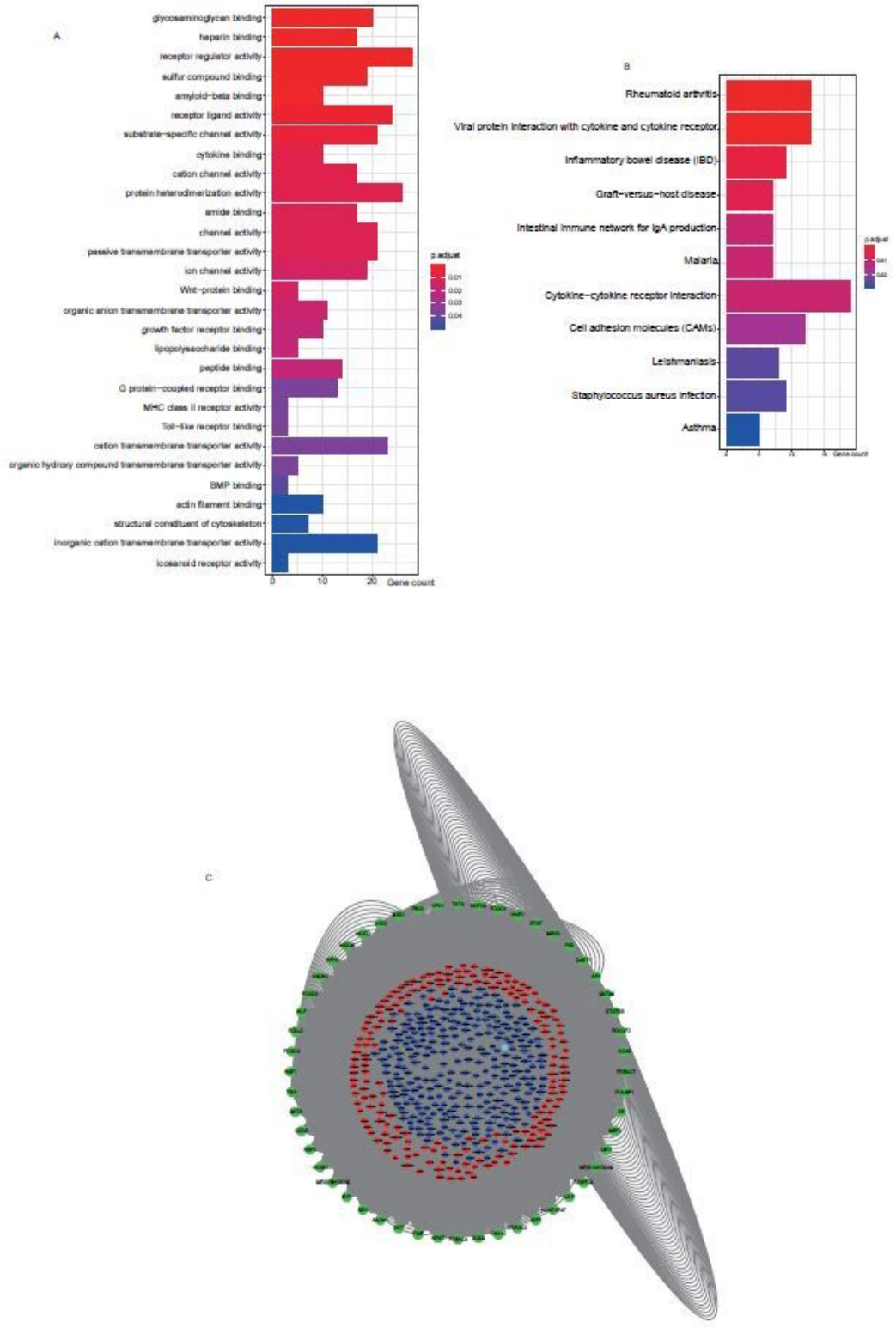

Figure 3

Enrichment analysis of DEGs Notes: (A) GO enrichment analysis of DEGs. (B) KEGG pathway enrichment analysis of DEGs. (C) TFs enrichment analysis of the DEGs. Green represent TFs, blue represent down-regulated DEGs between TFI and endometriosis-related infertility. blue represent up-regulated DEGs between TFI and endometriosis-related infertility.

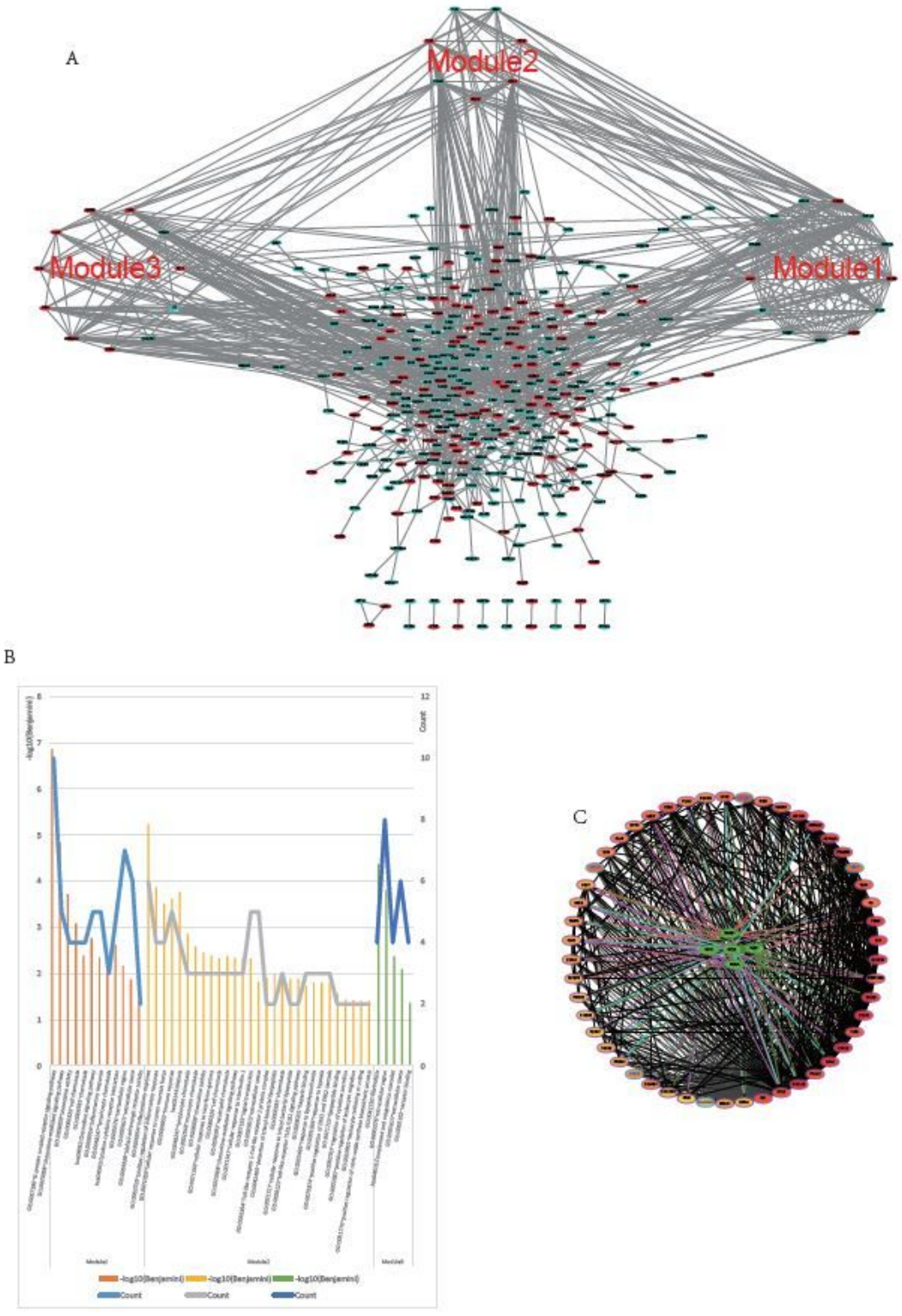

Figure 4

Important modules and Hub genes in PPI network Notes: (A) Construction of PPI network and analysis of important modules in PPI network. (B) GO and KEGG pathway enrichment analysis of important modules. (C) Screening of Hub genes and their TFs. Red represent Hub gens and green represent TFs. The redder the color, the stronger their interaction.

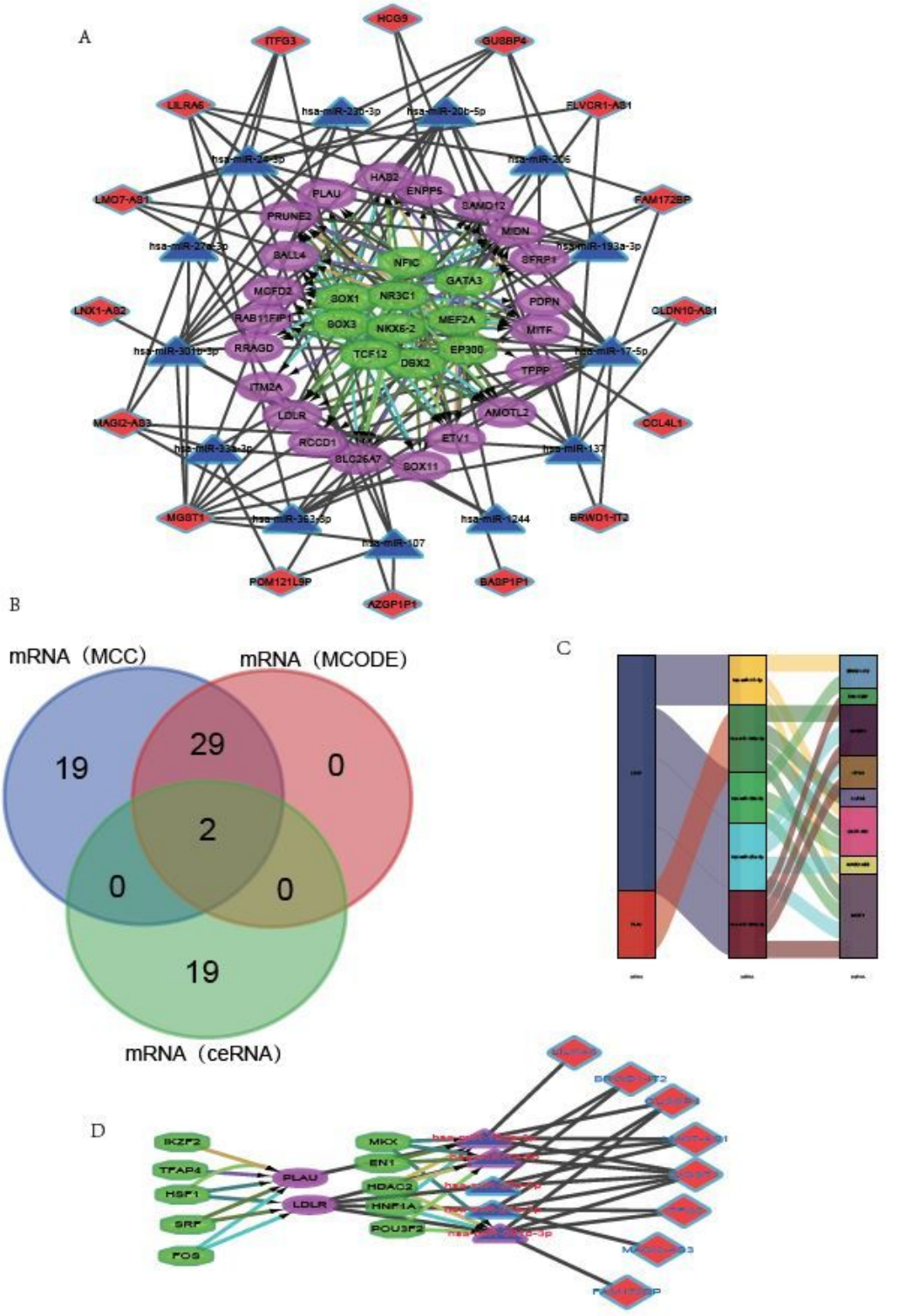

Figure 5

Construction of ceRNA network combining TFs and screen of potential key genes Notes: (A) A ceRNA network combining TFs based on DEGs and DELs. Red represent lncRNAs. Navy blue represent miRNAs. Violet represent mRNAs. Green represent TFs. (B) Venn diagram of MCC (cytoHubba), MCODE and ceRNA network. The lapping genes represent the potential key genes. (C) Sankey diagram for the ceRNA network based on potential key genes. Each rectangle represents a gene, miRNA or lncRNA, and the connection degree of each gene is visualized based on the size of the rectangle. (D) CeRNA network combining TFs based on potential key genes. Red represent lncRNAs, navy blue represent miRNAs, violet represent potential key genes and green represent TFs.