



OPEN The shared mechanism and potential diagnostic markers for premature ovarian failure and dry eye disease

Xi Long^{1,2}, Zixuan Wu¹, Pengfei Jiang³, Kang Tan¹, Pei Liu¹ & Qinghua Peng^{1,2}✉

Premature ovarian failure (POF), which is often comorbid with dry eye disease (DED) is a key issue affecting female health. Here, we explored the mechanism underlying comorbid POF and DED to further elucidate disease mechanisms and improve treatment. Datasets related to POF (GSE39501) and DED (GSE44101) were identified from the Gene Expression Omnibus (GEO) database and subjected to weighted gene coexpression network (WGCNA) and differentially expressed genes (DEGs) analyses, respectively, with the intersection used to obtain 158 genes comorbid in POF and DED. Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO) analyses of comorbid genes revealed that identified genes were primarily related to DNA replication and Cell cycle, respectively. Protein–Protein interaction (PPI) network analysis of comorbid genes obtained the 15 hub genes: CDC20, BIRC5, PLK1, TOP2A, MCM5, MCM6, MCM7, MCM2, CENPA, FOXM1, GINS1, TIPIN, MAD2L1, and CDCA3. To validate the analysis results, additional POF- and DED-related datasets (GSE48873 and GSE171043, respectively) were selected. miRNAs-lncRNAs-genes network and machine learning methods were used to further analysis comorbid genes. The DGIdb database identified valdecoxib, amorfrutin A, and kaempferitrin as potential drugs. Herein, the comorbid genes of POF and DED were identified from a bioinformatics perspective, providing a new strategy to explore the comorbidity mechanism, opening up a new direction for the diagnosis and treatment of comorbid POF and DED.

Keywords Premature ovarian failure, Dry eye disease, Shared mechanism, Potential diagnostic markers, Bioinformatics

Abbreviations

POF	Premature ovarian failure
POI	Premature ovarian insufficiency
DED	Dry eye disease
DES	Dry eye syndrome
OSD	Ocular surface disease
DTS	Dysfunctional tear syndrome
GEO	Gene Expression Omnibus
WGCNA	Weighted gene co-expression network analysis
DEGs	Differentially expressed genes
KEGG	Encyclopedia of Genes and Genomes
GO	Gene ontology
BP	Biological process
CC	Cellular component
MF	Molecular function
PPI	Protein–Protein interaction
MCODE	Molecular complex detection
CDC	Cell division cycle

¹Hunan University of Chinese Medicine, Changsha, China. ²The First Affiliated Hospital of Hunan University of Chinese Medicine, Changsha, China. ³Quzhou Hospital of Zhejiang Medical and Health Group, Quzhou, China. ✉email: pengqinghua@hnuucm.edu.cn

MCM	Minichromosome maintenance
MCC	Maximal clique centrality
MNC	Maximal neighborhood component
LASSO	Least absolute shrinkage and selection operator
SVM-RFE	Support vector machine recursive feature elimination
CV	Cross-validation
ROC	Receiver operating characteristic
AUC	An area under the curve

Premature ovarian failure (POF), formerly known as premature ovarian insufficiency (POI) or premature menopause, is characterized by the cessation of ovarian function before 40 years of age¹. This condition is marked by hypoestrogenism and diminished follicular reserves, leading to menstrual irregularities, infertility, and diminished health-related quality of life². POF is a diverse disease frequently overlooked during diagnosis, with a multifaceted etiology that includes genetic, iatrogenic, infectious, and autoimmune origins. Clinically, POF symptoms predominantly stem from estrogen deficiency and present as amenorrhea, oligomenorrhea, vasomotor symptoms (such as hot flashes and night sweats), sleep disturbances, vulvovaginal atrophy, urinary frequency changes, recurrent infections, and mood disturbances, including irritability and emotional instability³. Women with POF have increased risks of cardiovascular disease, dementia, cognitive decline, Parkinsonism, dry eye syndrome, and osteoporosis^{4,5}. The psychological impact of POF is profound, necessitating counseling to address infertility⁶, self-esteem issues, and increased anxiety and depression.

Dry eye disease (DED), also known as keratoconjunctivitis sicca (KCS), dry eye syndrome (DES), ocular surface disease (OSD), or dysfunctional tear syndrome (DTS), is a common ophthalmic disorder. This syndrome, particularly common in individuals over the age of 48 years, arises from an imbalance in the precorneal tear film, affecting roughly one in seven individuals in this age group. DED originates from dysfunction within the nasolacrimal unit (encompassing the nasolacrimal glands, corneal surface, and eyelids), which leads to compromised or inadequate tear film production⁷. A physiologically intact tear film is crucial for clear vision because it collaborates with the cornea to focus light on the retina. Additionally, it plays a vital role in lubricating the eye, clearing ocular debris, and sustaining nutrition and oxygenation of ocular structures⁸. Patients with DED often experience symptoms such as ocular burning, blurred vision, or pain, which substantially diminish their quality of life. Activities that require visual focus, such as reading or computer work, can be challenging⁹. Despite the availability of treatments to alleviate the symptoms of DED, they often fail to address the underlying causes. A diverse array of factors, including autoimmunity, hormonal imbalances, and adverse environmental conditions, have been implicated in the development of DED¹⁰. Notably, dry eye symptoms may occasionally signal undiagnosed systemic diseases, and timely interventions could avert potentially life-threatening complications.

POF is frequently comorbid with DED, which markedly affects the physical and mental health of the affected women. Numerous studies have established a linkage between POF and DED^{11,12}. The diminished estrogen levels resulting from ovarian hypoplasia adversely affect lacrimal gland secretion, leading to the characteristic symptoms of DED¹³. Additionally, certain inflammatory factors are aberrantly expressed under both conditions, suggesting a shared pathophysiological basis. Investigating these comorbid mechanisms is crucial for understanding the development and progression of both POF and DED^{14,15}. Moreover, such explorations could offer novel perspectives and approaches for treating these conditions. The POF and DED Initiative, which utilizes high-throughput transcriptome sequencing and enrichment with detailed clinical annotations, presents an unparalleled opportunity to examine the transcriptional landscape and molecular pathways of POF and DED. Bioinformatics analyses of these comprehensive datasets have yielded valuable insights into their intricate pathophysiological mechanisms. However, a notable gap remains regarding the use of bioinformatics to explore the roles of differentially expressed genes (DEGs) in POF and DED. In the current study, we aimed to fill this gap by analyzing the POF and DED-related Gene Expression Omnibus (GEO) datasets to enhance our molecular understanding of POF and DED and identify new therapeutic targets, as illustrated in Fig. 1.

Results

GEO information

The DED- and POF-related datasets were screened from GEO to obtain gene chip datasets GSE44101 (microarray platform GPL1261), GSE39501 (microarray platform GPL6887), GSE171043 (microarray platform GPL28441), and GSE48873 (microarray platform GPL13607), which were eligible for the study. These datasets contain both normal and diseased samples. After screening, we selected the DED-related dataset GSE44101 for analysis and GSE171043 for validation, the POF-related dataset GSE39501 for analysis, and GSE48873 for validation.

Weighted gene coexpression network analysis (WGCNA) of DED and POF

Data from the GSE44101 and GSE39501 datasets were analyzed after preprocessing to obtain sample clusters. The β values of 20 and 9 were selected as the soft threshold to construct a coexpression network respectively (Fig. 2A, B), which fulfilled the following conditions: (i) the scale-free fitting index was > 0.8 , (ii) the average connectivity of the network was good, and (iii) the network conformed to the non-scale topological network distribution. Module identification was performed using a dynamic tree-chopping algorithm and by merging similar modules. The gene sets within the modules were highly related. The final GSE44101 dataset obtained 33 modules, while the GSE39501 dataset obtained 50 modules, respectively (Fig. 2C, D). Upon correlating the modules of the GSE44101 and GSE39501 datasets with the clinical characteristics (treat and control) of the samples and setting the MM threshold to 0.8, GS threshold to 0.1, and weight threshold to 0.1, we found that the GSE44101 dataset darkolivegreen4 module was highly correlated with the clinical characteristics ($r = 0.99$, $P = 0.00$) of the tissues.

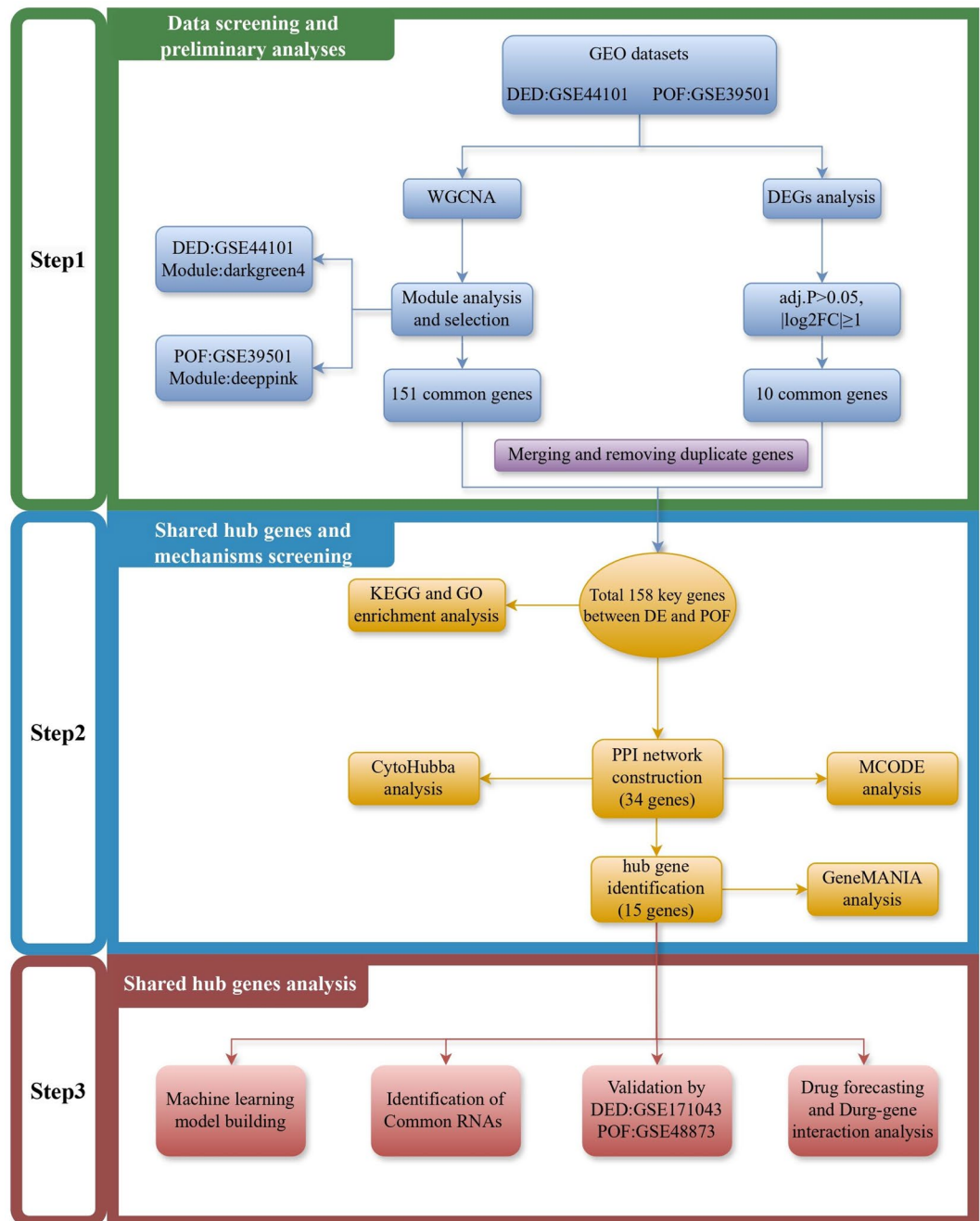


Figure 1. Schematic diagram of the general flow of this research work. In order to gain a more penetrating understanding of the mechanisms of comorbidity between POF and DED, we began our research journey using data from the GEO database. Specifically, this study was divided into three main parts. Firstly, in the first part we identified the GSE44101 and GSE39501 datasets as the primary cohorts through a careful matching strategy, and identified the GSE171043 and GSE48873 datasets for validation. We performed WGCNA analysis on GSE44101 and GSE39501 to identify the key modules and module genes, followed by differential expression analysis, respectively. The 158 common genes of POF and DED were obtained by merging and removing duplicate genes. Subsequently, in the second part, in order to reveal the functional basis of these common key genes, we performed exhaustive GO and KEGG analyses. Meanwhile, in order to further screen the hub genes, we executed a set of step-by-step screening of 15 hub genes after covering PPI, MCODE, Cytoscape, and GeneMANIA analyses. In the third part, the hub genes were further analysed and validated through seamless integration with 4 different methods and 2 datasets. This comprehensive approach not only enriches our understanding of the mechanisms of POF and DED comorbidity, but also lays a solid foundation for identifying prospective targeted therapies and interventions for the treatment of POF and DED comorbidity.

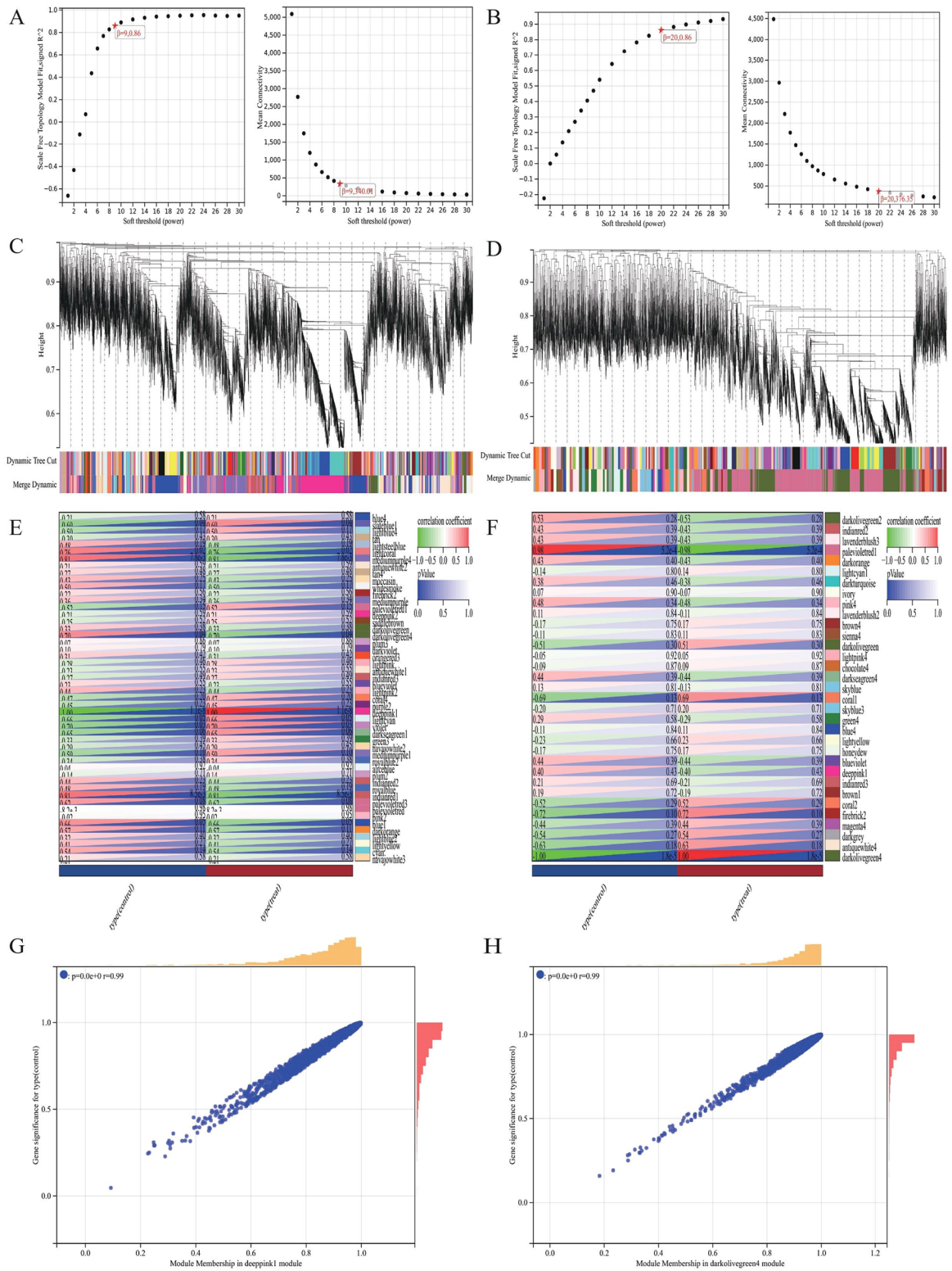


Figure 2. Weighted Gene Co-expression Network Analysis. **(A)** Selection of soft thresholds (β) for WGCNA analysis of POF. **(B)** Selection of soft thresholds (β) for WGCNA analysis of DED. **(C)** This diagram shows the degree of correlation within the module of the POF dataset. **(D)** This diagram shows the degree of correlation within the module of the DED dataset. **(E)** Relationship between modules and clinical features in POF dataset. **(F)** Relationship between modules and clinical features in DED dataset. Rows represent different modules and columns represent clinical features. Each cell contains the correlation between modules and clinical features and the corresponding p -value. **(G)** Module membership in deeppink1 module in POF dataset. **(H)** Module membership in darkolivegreen4 module in DED dataset.

The GSE39501 dataset deeppink1 module was highly correlated with clinical characteristics ($r = 0.99$, $P = 0.00$), as shown in Fig. 2E, F, G, H. Therefore, we selected the green GSE44101 dataset darkolivegreen4 module with the GSE39501 dataset deeppink1 module for further screening of key genes analysis. By analyzing WGCNA, 151 coexpressed genes in DED and POF were obtained by considering the intersection of the two (Fig. 4A).

Analysis of DED and POF DEGs

Differential gene expression analysis of GSE44101 and GSE39501 revealed 1089 DEGs in GSE44101, of which 553 genes were upregulated, and 536 genes were downregulated. A total of 166 DEGs were found in GSE39501, of which 131 were upregulated, and 35 were downregulated (Fig. 3). GSE44101 and GSE39501 had 10 common

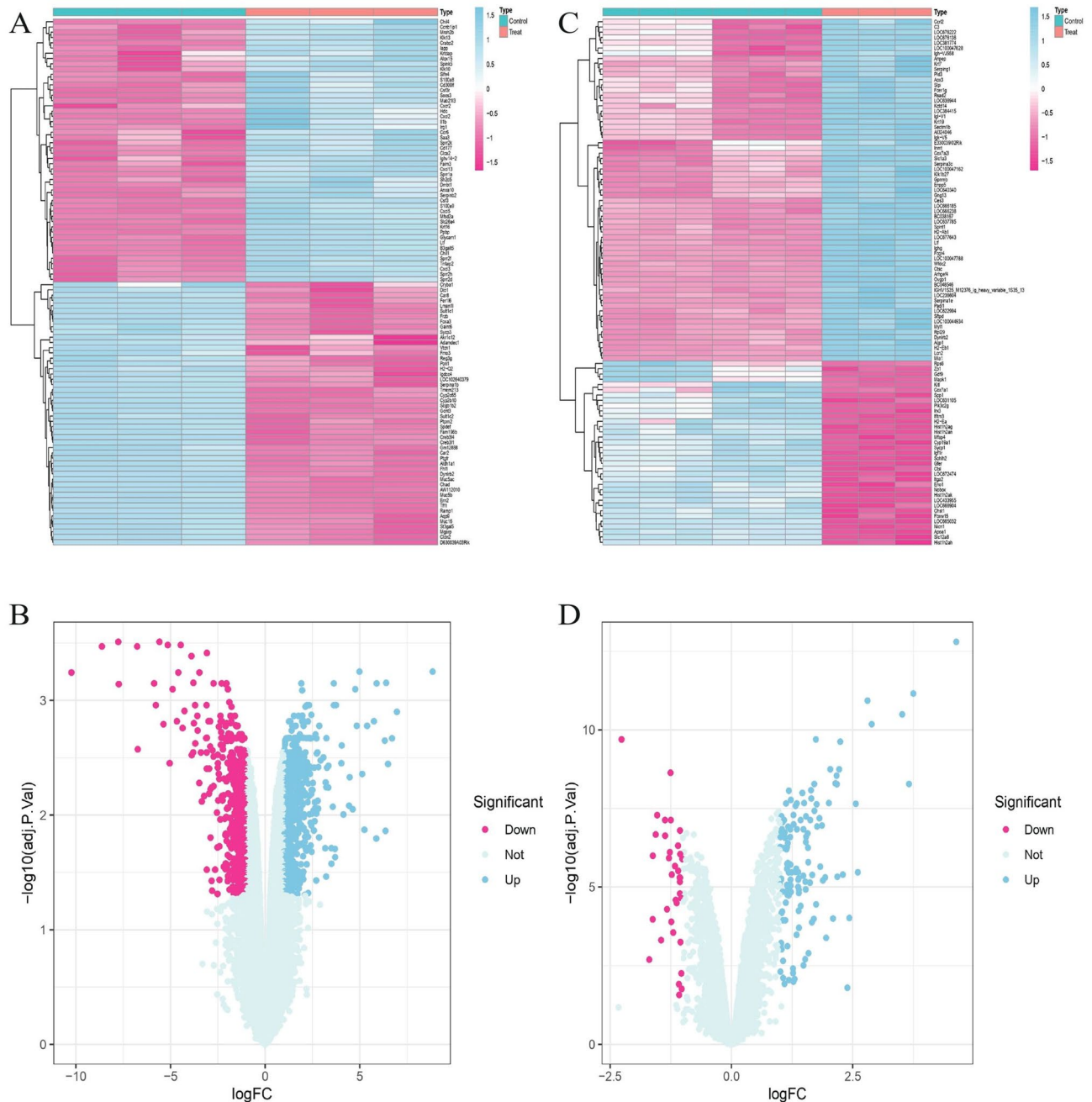


Figure 3. Differentially expressed genes in POF and DED. (A) Volcano plot of differentially expressed genes in DED dataset, with pink representing up-regulated genes and blue representing down-regulated genes. (B) Heatmap of differentially expressed genes in DED dataset. Pink color represents high expression and blue color represents low expression. (C) Volcano plot of differentially expressed genes in POF dataset, with pink representing up-regulated genes and blue representing down-regulated genes. (D) Heatmap of differentially expressed genes in POF dataset. Pink color represents high expression and blue color represents low expression.

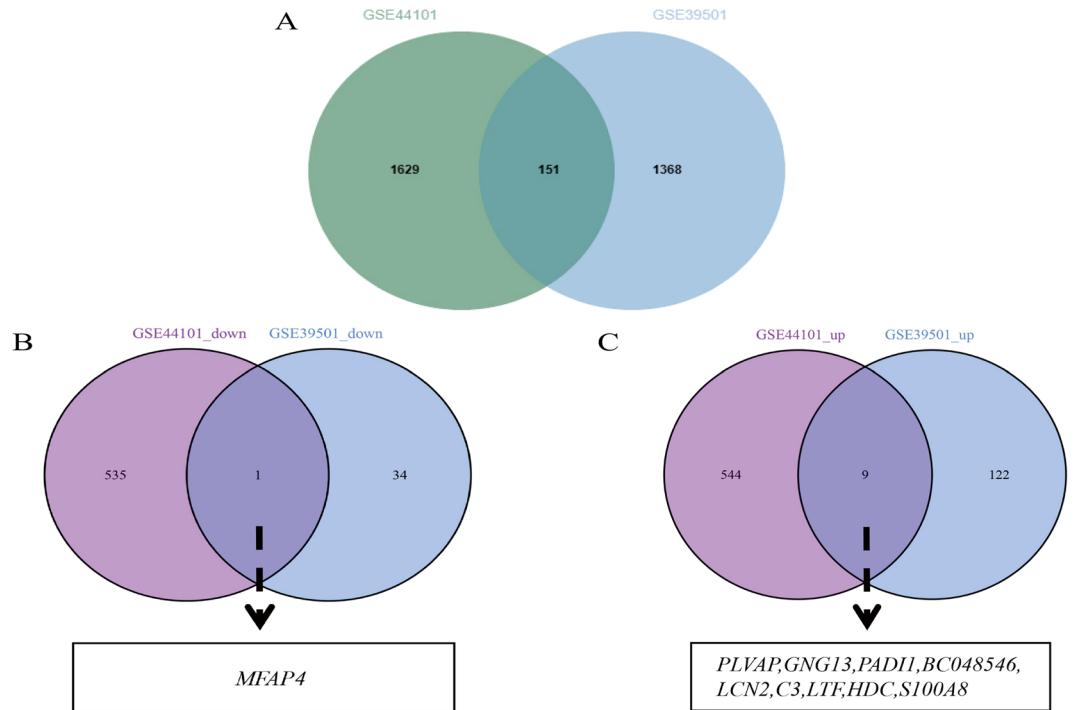


Figure 4. POF and DED co-morbidity-genes screening. (A) Intersecting WGCNA genes of POF and DED. Green color represents DED dataset and blue color represents POF dataset. (B) Intersecting down-regulated DEGs of POF and DED. Purple color represents DED dataset and blue color represents POF dataset. (C) Intersecting up-regulated DEGs of POF and DED. Purple color represents DED dataset and blue color represents POF dataset.

DEGs, of which nine genes were upregulated, namely BC048546, LCN2, LTF, PADI1, HDC, PLVAP, GNG13, S100A8, and C3, and one gene was downregulated, *MFAP4*, as shown in Fig. 4B, C.

Enrichment analysis and functional annotation

After merging and removing duplicate genes, 158 DED and POF comorbid genes were obtained by combining all genes obtained from WGCNA and differential expression gene analysis. Genes were enriched using R to further clarify the functions of the screened genes and their roles in signaling pathways. The results showed that the identified genes were mainly enriched in DNA replication, cell cycle, glycosylphosphatidylinositol (GPI)-anchor biosynthesis, Chagas disease (American trypanosomiasis), TNF signaling pathway, leishmaniasis, Kaposi sarcoma-associated herpesvirus infection, IL-17 signaling pathway, progesterone-mediated oocyte maturation, and legionellosis, while its major genes functions were mainly in chromosome, chromosomal part, cell cycle, chromosomal region, chromosome, centromeric region, mitotic cell cycle, condensed chromosome, kinetochore, condensed chromosome kinetochore, condensed chromosome, centromeric region, as shown in Fig. 5.

PPI network construction and modular analysis

The genes obtained from WGCNA and differential expression genes analysis were imported into the STRING10.0 database, and the species was set as "homo sapiens" to obtain the genes comorbid in DED and POF. These genes were analyzed using PPI, and the key comorbid genes were screened, as shown in Fig. 6A. Two individual clustering modules of closely linked genes were extracted using the Molecular Complex Detection (MCODE) plugin (Fig. 6B, C). Cluster 1 contained 7 nodes and 19 edges. Cluster 2 contains 10 nodes and 32 edges.

Selection and analysis of hub genes

Furthermore, network analysis using the CytoHubba plugin with MCC, MNC, and Degree value as reference standards identified 15 key hub genes, including CDC20, BIRC5, PLK1, CDCA8, TOP2A, MCM5, MCM6, MCM7, MCM2, CENPA, FOXM1, GINS1, TIPIN, MAD2L1, and CDCA3 (Fig. 6D, E, F). Subsequently, we evaluated the coexpression network and the associated roles of the hub genes using the GeneMANIA database. The network results identified 46.7% coexpression, 22.85% physical interactions, 18.15% predictions, 5.39% pathways, 2.51% shared protein domains, 2.27% genetic interactions, and 2.14% co-localization (Fig. 6G).

Validation of hub gene expression

The expression levels of 15 key genes were validated using the POF-related dataset GSE48873 and the DED-related dataset GSE171043. The t-test was performed on each subset of the dataset. The significance level was set at $P < 0.05$. Herein, we found that PLK1, CDCA8, TOP2A, MCM5, MCM6, MCM2, CENPA, MCM7, TIPIN,

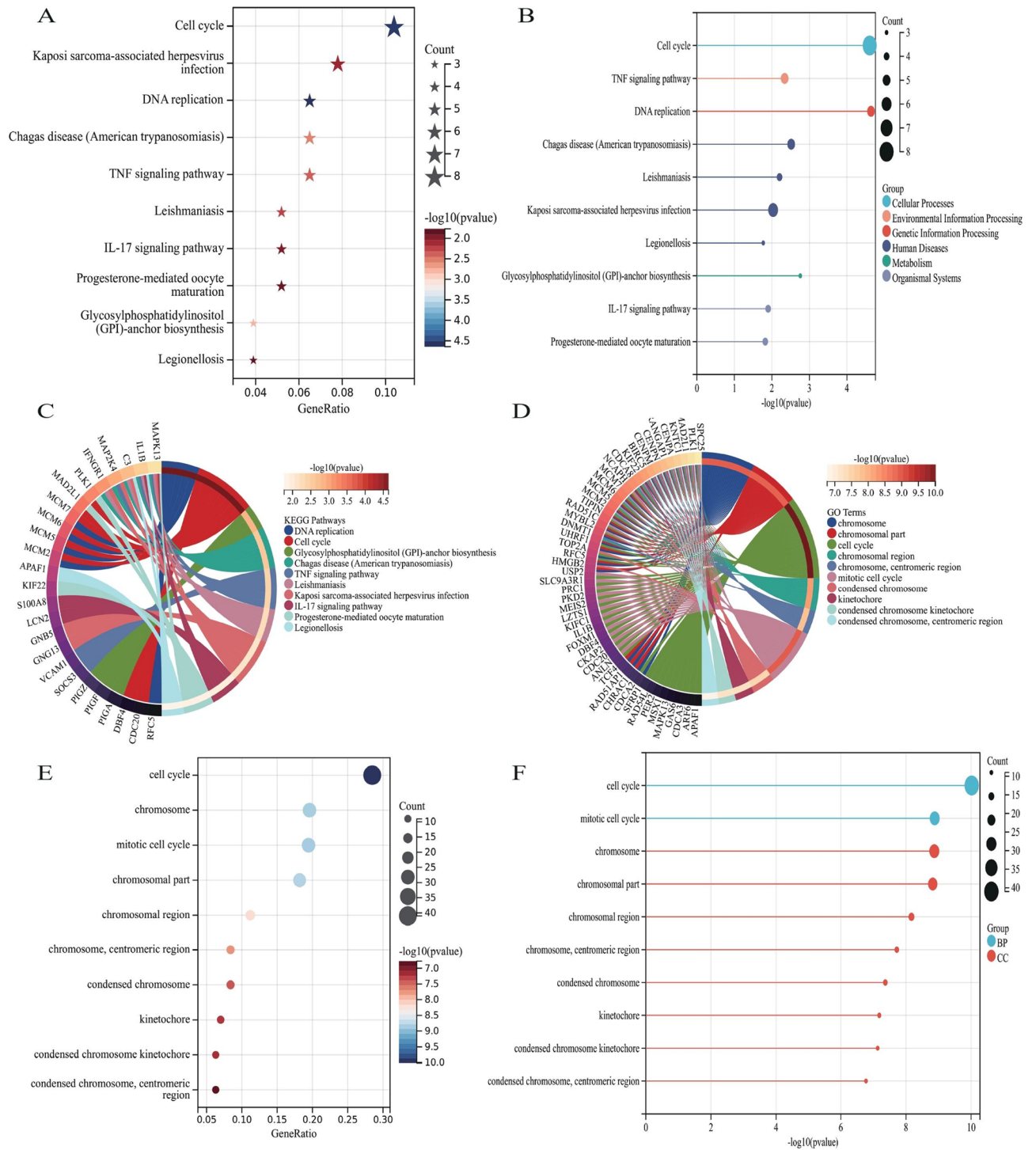


Figure 5. Enrichment analysis and Functional Annotation of comorbid genes between POF and DED. (A) Bubble diagram of KEGG enrichment analysis. (C) Histogram of KEGG enrichment analysis. (D) Circle diagram of KEGG enrichment analysis. (E) Bubble diagram of GO functional analysis. (F) Histogram of GO functional analysis. (G) Circle diagram of GO functional analysis.

GINS1, MAD2L1, CDCA3, BIRC5, FOXM1, and CDC20 were expressed in another POF-related dataset, in which MCM7, TIPIN, GINS1, and MAD2L1 expression levels were significantly higher ($P < 0.05$), and CDC20 expression levels were significantly lower ($P < 0.05$), as shown in Fig. 7. Additionally, CDC20, PLK1, TOP2A, MCM5, MCM6, MCM7, MCM2, CENPA, FOXM1, GINS1, TIPIN, MAD2L1, CDCA3, and CDCA8 were expressed in another DED-related dataset (Fig. 8). Therefore, we hypothesized that CDC20, CDCA8, PLK1, TOP2A, MCM5, MCM6, MCM7, MCM2, CENPA, FOXM1, GINS1, TIPIN, MAD2L1, and CDCA3 were comorbid in POF and DED. TIPIN, GINS1, MAD2L1, and CDC20 were also differentially expressed.

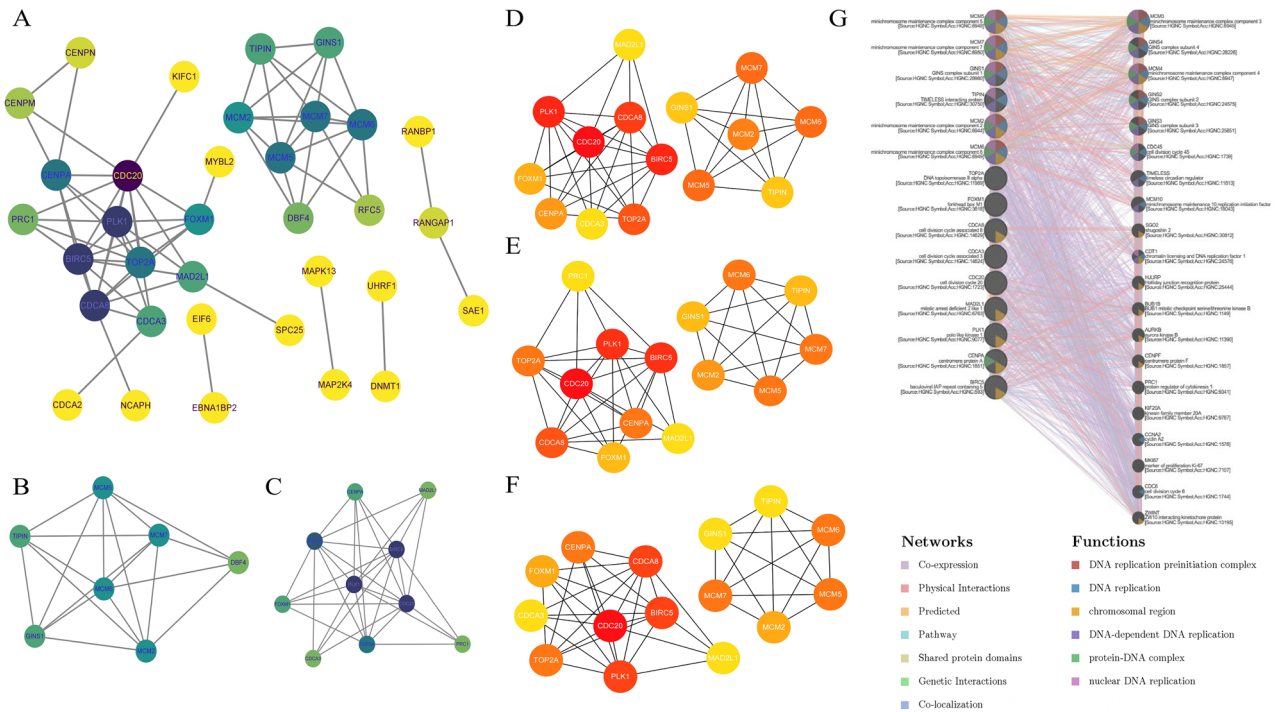


Figure 6. Identification and analysis of hub genes. (A) Protein–protein interaction network from String10.0. The purpler the node represents its higher significance, the yellower it is the opposite. (B) Cluster 1 from MCODE analysis of key genes. (C) Cluster 2 from MCODE analysis of key genes. (D) Identification of hub genes by CytoHubba of MCC ranking method. (E) Identification of hub genes by CytoHubba of MNC ranking method. (F) Identification of hub genes by CytoHubba of Dgree ranking method. (G) Hub genes and their co-expression genes were analyzed by GeneMANIA.

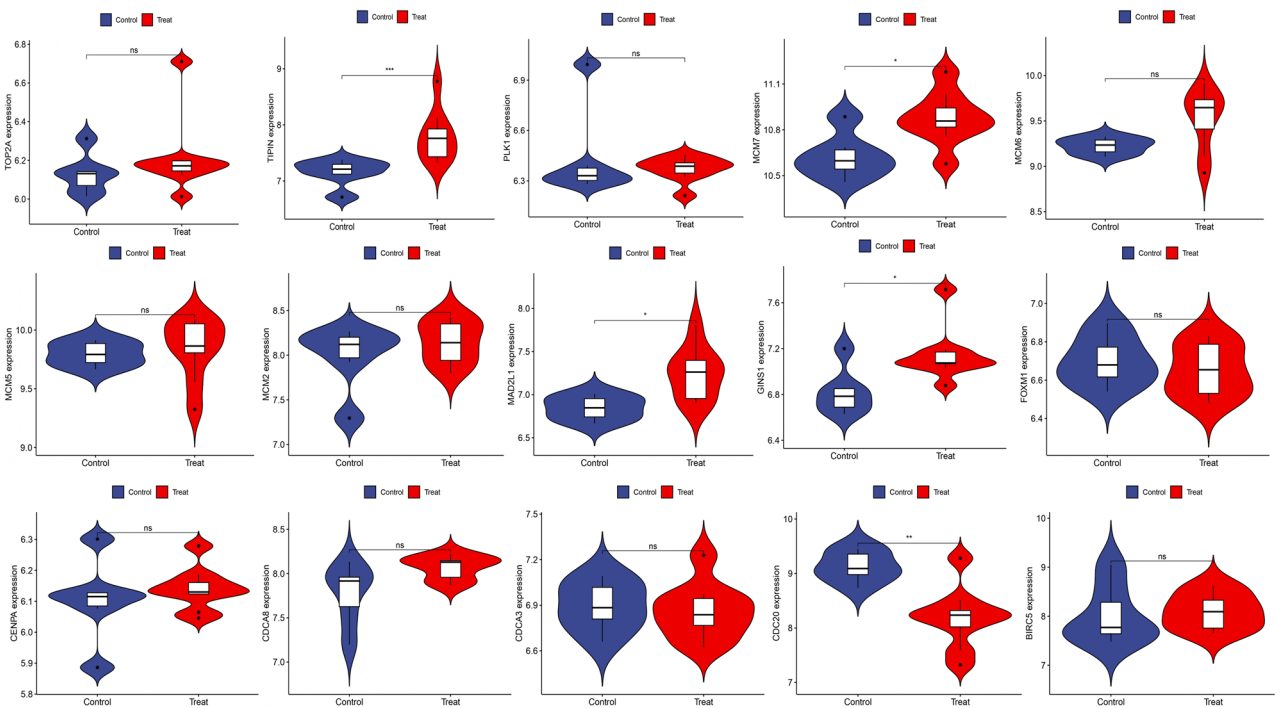


Figure 7. Validation of hub genes expression by POF dataset (GSE48873).

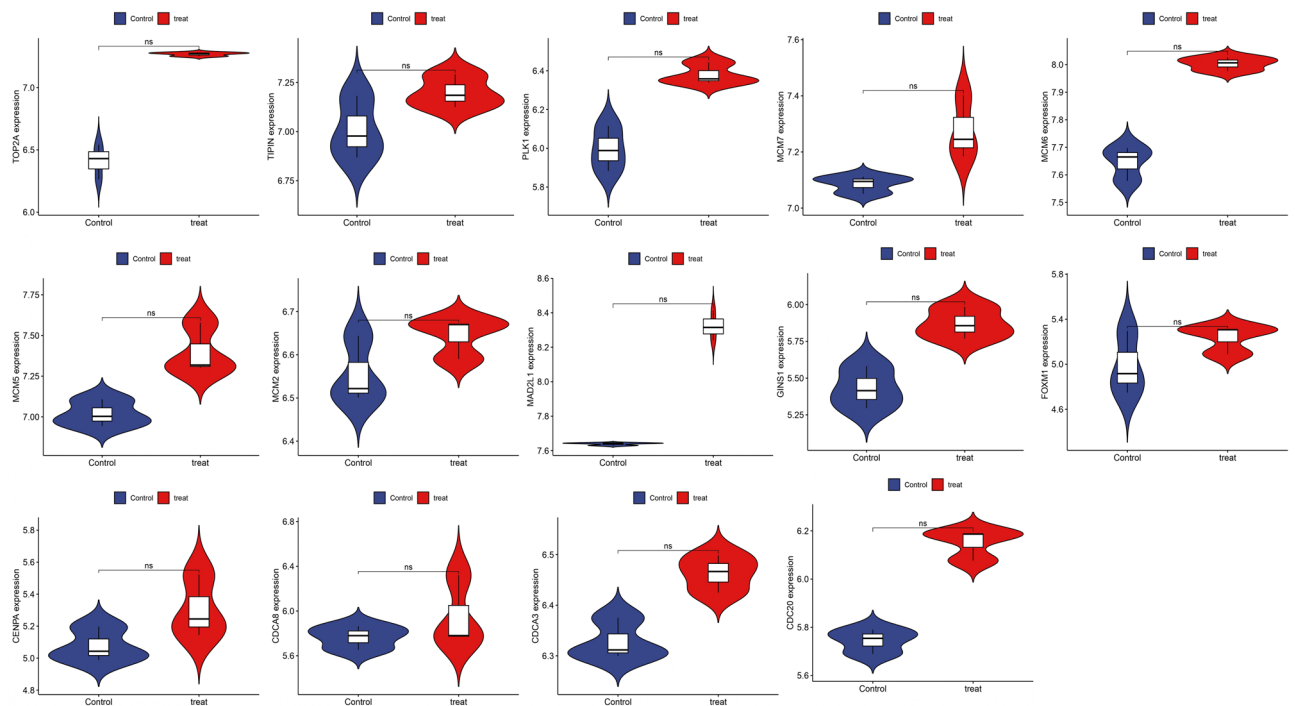


Figure 8. Validation of hub genes expression by DED dataset (GSE171043).

Identification of common RNAs and construction of microRNAs (miRNAs)- long noncoding RNAs (lncRNAs) shared genes network

A total of 210 miRNAs and 296 lncRNAs associated with DED and POF were identified from the three databases, namely, miRanda, miRDB, and TargetScan. When the corresponding database matched the relevant miRNAs, the score was marked as 1. Upon matching all three databases were matched, a score of 3 points was assigned. miRNAs were matched using the SpongeScan database to obtain the corresponding lncRNA data. The miRNA-lncRNA-gene network was constructed by intersecting these non-coding RNAs with the shared genes. The network comprised 57 lncRNAs, 174 miRNAs, and some common genes, including the 14 hub genes (BIRC5, FOXM1, CDCA8, CDCA3, MCM6, CENPA, TOP2A, MAD2L1, PLK1, MCM5, TIPIN, MCM2, GINS1, and MCM7) (Fig. 9).

Drug forecasting and drug-gene interaction analysis

The prediction was performed using the DGIdb database, and a total of 293 predicted drugs were obtained, of which BIRC5 corresponded to 36 predicted drugs, PLK1 corresponded to 176 predicted drugs, and TOP2A corresponded to 80 predicted drugs. We constructed a drug-gene interaction network (Fig. 10A) and found that idarubicin hydrochloride and myricetin were common candidates for TOP2A and PLK1. Doxorubicin, fluorouracil, paclitaxel, epirubicin, camptothecin, and genistein were common TOP2A and BIRC5 candidates. According to the above studies, we screened these drugs based on their interaction scores after excluding antineoplastic drugs. The drugs with the highest interaction scores for each hub gene are listed in Table 1. The BIRC5-predicted drug was valdecoxib, the PLK1-predicted drug was amorfrutin A, and the TOP2A-predicted drug was kaempferitrin. Molecular docking between the hub genes and their predicted drugs revealed that the binding energies of the hub genes and their counterparts were < 0 , suggesting that both could bind spontaneously. It is generally believed that the lower the energy, the more stable the ligand-receptor binding conformation and the greater the possibility of action. Analysis of the molecular docking results showed that BIRC5 had the lowest binding energy with valdecoxib (-5.51 kJ/mol), indicating that the ligand and receptor have the most stable conformation. The docking results are shown in Fig. 10B, C, D.

Machine learning model building

The 15 identified genes were explored as potential markers using the least absolute shrinkage and selection operator (LASSO) regression algorithm and support vector machine recursive feature elimination (SVM-RFE) analysis (Fig. 11). In the LASSO regression algorithm analysis of hub genes in GSE39501, five featured genes, namely CDC20, CDCA3, MCM5, PLK1, and TOP2A, were considered as characteristic genes that significantly correlated with the samples. Likewise, in GSE44101, two featured genes, namely, MAD2L1 and TIPIN, were considered characteristic genes. Notably, in the SVM-RFE analysis, all hub genes in GSE39501 and GSE44101 were identified as characteristic genes. We combined these two results for receiver operating characteristic (ROC) analysis of hub genes, revealing that CDC20, CDCA3, MCM5, PLK1, TOP2A, MAD2L1, and TIPIN possessed good diagnostic performance as characteristic genes, with an area under the curve (AUC) of 1.

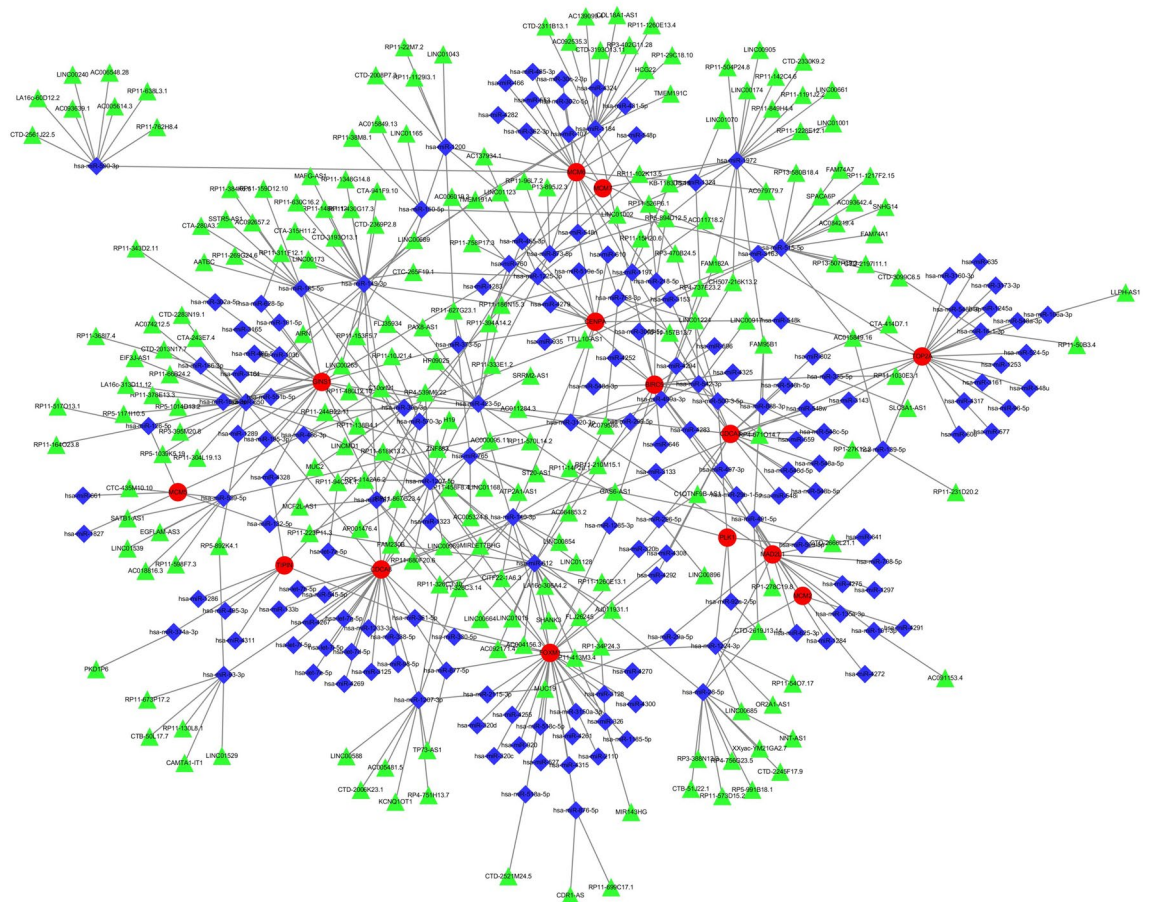


Figure 9. miRNAs-lncRNAs shared genes network. Red circles are mRNAs, blue quadrangles are miRNAs, and green triangles are lncRNAs.

Discussion

POF, also known as pathological ovarian aging, is a reproductive disorder that markedly affects women during their fertile years. The cardinal facets of ovarian aging include a diminishing count and compromised follicle quality, thereby resulting in menstrual cycle irregularities or outright amenorrhea². This progressive decline in fertility ultimately culminates in a complete loss of reproductive capacity, concurrently entwined with the onset of multifaceted systemic ailments culminating in menopause. Women with POF have an increased susceptibility to diverse systemic and ocular disorders¹³. DED, a prevalent ocular affliction with global ramifications, is characterized by inadequate tear production and tears of suboptimal quality, precipitating discomfort, visual disturbances, and consequential harm to the ocular surface¹⁶. Various risk factors, such as age, hormonal fluctuations, and autoimmune conditions, have been identified in DED¹⁰. Some studies have suggested that DED is associated with comorbidities. Recent empirical data indicate an elevated prevalence of DED among individuals with POI when compared with those with normal ovarian function¹². However, the underlying mechanisms and the potential interconnections between these two conditions remain unclear. Against the backdrop of an evolving narrative on tumorigenesis, the scientific focus is progressively veering toward non-oncological realms. The collective interplay between comorbid genes in POF and DED represents an uncharted territory awaiting further exploration. A nuanced examination of the distinctive patterns of comorbidity amid the progression of POF and DED may elicit crucial insights into their pathogenic roles, thereby fostering the development of precisely targeted therapeutic interventions.

In the current study, we conducted data mining and analyzed the comorbidity mechanism of POF and DED based on bioinformatics methods, such as WGCNA and DEGs. We used the CytoHubba plug-in to perform an in-depth analysis of the PPI network and identified 15 key hub genes: CDC20, BIRC5, PLK1, CDCA8, TOP2A, MCM5, MCM6, MCM7, MCM2, CENPA, FOXM1, GINS1, TIPIN, MAD2L1, and CDCA3. Accordingly, hub genes were mainly covered by the cell division cycle (CDC) and minichromosome maintenance (MCM) gene families. Several studies have found that CDC family genes are associated with apoptosis¹⁷, which induces cellular senescence owing to its apoptotic mechanism, with the induction of cell senescence possibly resulting in numerous ocular surface disorders, including DED¹⁸. Simultaneously, cellular senescence, as one of the key causes of DED, is associated with alterations in the structure and function of the lacrimal gland, such as atrophy and reduced protein secretion¹⁹, with a higher probability of lacrimal gland lesions in older females²⁰. The MCM gene family is widely found in eukaryotes such as humans, plays a role in regulating cell division, and is a key factor in initiating DNA replication²¹. Abnormalities in DNA replication are key to inducing reduced tearing

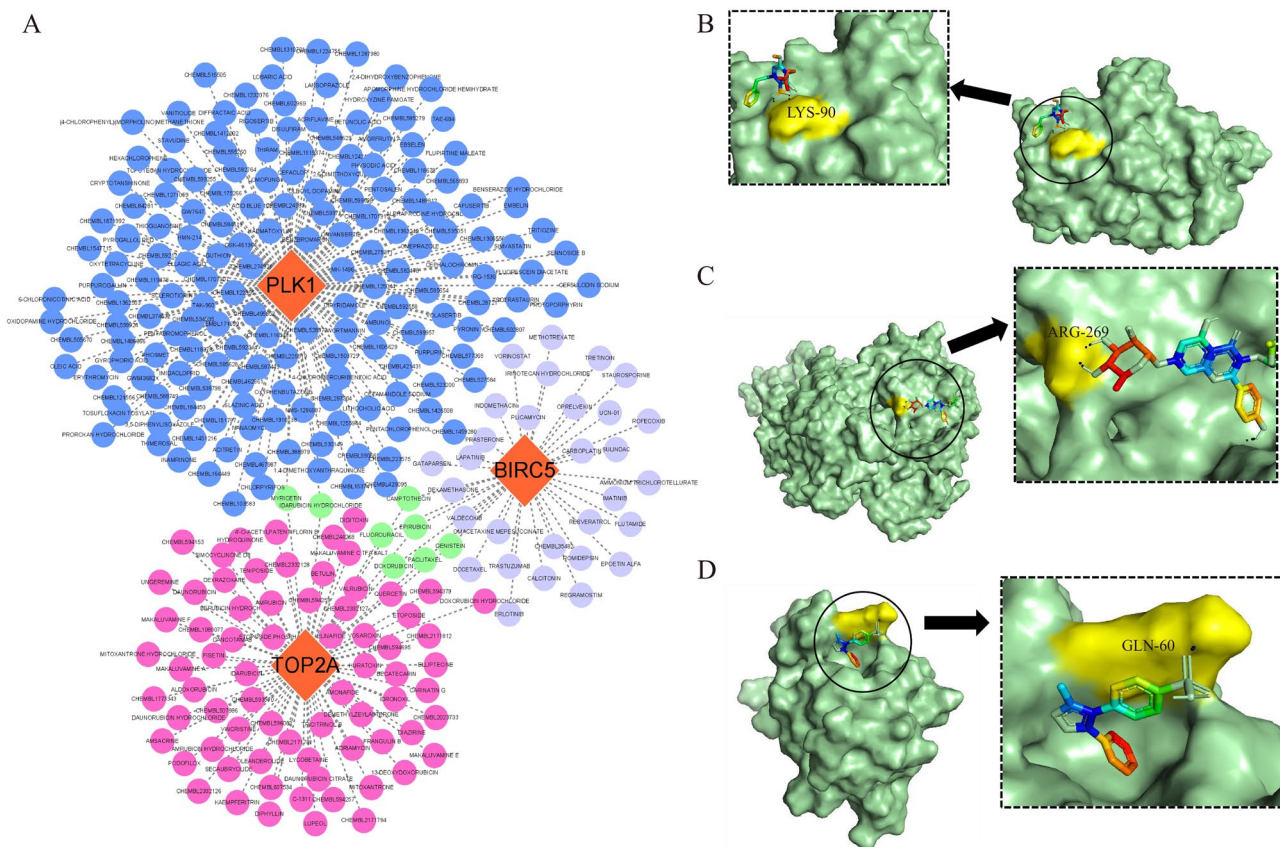


Figure 10. Drug-gene interaction network and docking analysis. (A) Drug-gene interaction network. Orange circles represent genes, blue circles represent predictive drugs for PLK1, purple circles represent predictive drugs for BIRC5, pink circles represent predictive drugs for TOP2A, and green circles represent predictive drugs associated with tow genes. (B) The docking of BIRC5 with VALDECOXIB, the docking energy is -5.51 kJ/mol. (C) The docking of PLK1 with AMORFRUTIN A, the docking energy is -1.19 kJ/mol. (D) The docking of TOP2A with KAEMPFERITRIN, the docking energy is -1.84 kJ/mol.

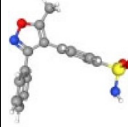
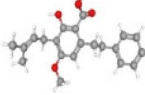
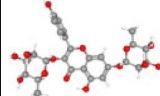
Gene	Drugs	Interaction score	Molecular formula	3D Conformer	Molecular docking score (KJ/mol)
BIRC5	VALDECOXIB	1.72	C16H14N2O3S		-5.51
PLK1	AMORFRUTIN A	0.09	C21H24O4		-1.19
TOP2A	KAEMPFERITRIN	1.55	C27H30O14		-1.84

Table 1. Drug-gene docking analysis.

and episodes of DED, which is consistent with the results of gene enrichment analysis. Primordial follicle depletion or dysfunction is key to the development of POF caused by MCM5, increasing the proliferation of primary granulosa cells, which, in turn, promotes the activation of primordial follicles in neonatal mouse ovaries²². The growth and development of mouse oocytes were found to be associated with three proteins of the MCM family: MCM2, MCM6, and MCM7²³. The expression levels of these pivotal genes were validated using the POF-related dataset GSE48873 and DED-related dataset GSE171043, suggesting that CDC20, CDCA8, PLK1, TOP2A, MCM5,

strategies to mitigate the effects of POF and DED. This integrative approach sheds light on the complex regulatory networks that drive disease comorbidity and paves the way for innovative interventions in precision medicine.

Further analysis revealed that the potential therapeutic agents used were valdecoxib for BIRC5, amorfrutin A for PLK1, and kaempferitrin for TOP2A. Valdecoxib is a non-steroidal anti-inflammatory drug that reportedly inhibits phagocytosis by macrophages⁵¹ and can prevent apoptosis induced by endoplasmic reticulum stress⁵². BIRC5 is a cell cycle-related regulator of apoptosis⁵³. Valdecoxib has high structural stability with BIRC5 docking, as well as a high potential to act via a mechanism possibly related to anti-apoptosis. Amorfrutin A exerts potent anti-inflammatory effects and can be used to treat inflammatory diseases⁵⁴. PLK1 is closely associated with inflammation and plays a role in inflammatory response via ROS-mediated NLRP3 inflammatory vesicles. Kaempferitrin is the main compound in *Chenopodium ambrosioides* extract, a natural herbal ingredient with anti-inflammatory and antioxidant activities⁵⁵. TOP2A, topoisomerase II α , was found to be closely associated with sepsis-induced acute lung injury, which involves a variety of responses, including inflammatory responses and apoptosis⁵⁶. Importantly, the mechanism underlying comorbid POF and DED cannot be distinguished from the occurrence of an inflammatory response, especially when an autoimmune-related inflammatory response is common. Moreover, amorfrutin A, PLK1, kaempferitrin, or TOP2A can be explored to achieve anti-inflammatory responses.

There have been noteworthy efforts to utilize bioinformatics analyses to elucidate the links between metabolic processes and eye and gynecological diseases. Wang et al. highlighted MMP1 and MMP9 as promising predictive indicators and therapeutic targets for uveal melanoma⁵⁷. Hu et al. developed a bioinformatics model for thyroid-eye disease and identified 11 hub genes (ATP6V1A, PTGES3, and PSMD12) with potential implications in these conditions⁵⁸. In a seminal investigation⁵⁹, Qin et al. reported a breakthrough with the identification of a CXCL10-based model exhibiting excellent diagnostic precision for POF. This model, encapsulating CXCL10 alongside Itga2 and Raf1, not only demonstrated robust diagnostic accuracy but also stands poised as a prospective and invaluable diagnostic biomarker. Thus, the collective expression profiles of these genes represent a reservoir of pertinent diagnostic information for POF. Furthermore, comprehensive research endeavors by Liu et al. shed light on a distinct set of six pivotal genes: SRSF1, PDIA5, NEURL1B, UNK, CELF2, and CFL2. Augmenting this discovery, the investigation also identified five hub miRNAs—hsa-miR-27a-3p, hsa-miR-24-3p, hsa-miR-22-3p, hsa-miR-129-5p, and hsa-miR-17-5p. Intriguingly, these findings suggest a regulatory nexus wherein the expression of these key genes appears to be orchestrated by two central hub miRNAs: hsa-miR-27a-3p and hsa-miR-17-5p⁶⁰. These findings not only augment our understanding of the intricate molecular landscape associated with POF but also herald a paradigm shift in diagnostic possibilities. However, the exploration of the mechanisms of co-morbidity between the two diseases, POF and DED, is still in a lacklustre stage.

The objective of the current study was to bridge this gap using a novel methodology. We enriched our analysis by incorporating extensive POF and DED datasets from a continuously refined GEO database. Our predictive model for comorbidity provides a theoretical framework and charts a course for forthcoming metabolic research and the therapeutic modulation of metabolic perturbations in POF and DED. The novelty of comorbidity research lies in the fact that it breaks away from the traditional concept of treating different diseases as independent entities in the field of medicine, and takes into account the correlation and interaction between different diseases. By studying the phenomenon of comorbidity, the interactions and influences between diseases can be better understood, providing more comprehensive guidance for diagnosis and treatment in clinical medicine. By taking multiple factors into account, the complexity and diversity of diseases can be better grasped, providing more effective strategies for prevention and treatment. Besides, the novelty of this study is mainly reflected in the use of cutting-edge and rich bioinformatics analyses in order to comprehensively and systematically reveal the association and interactions between POF and DED. Through big data analysis, potential associations hidden in massive data were mined to provide more precise and in-depth analyses for co-morbidity studies. Meanwhile, through the mining of mechanisms combined with bidirectional causality, it provides theoretical basis and practical guidance for in-depth understanding and intervention of POF and DED comorbidity. Nevertheless, our investigation has some limitations. The fundamental mechanisms underlying POF and DED require further empirical validation through comprehensive *in vitro* and *in vivo* studies. In addition, the prognostic implications of comorbid genes remain elusive, presenting a valuable opportunity for future research. However, the intricate relationship between these comorbid genes in POF and DED remains elusive. Unraveling this relationship could provide critical insights into the pathophysiology of these conditions and reveal novel targets for therapeutic intervention. This represents fertile ground for future inquiry, with the potential to substantially advance our understanding of the molecular underpinnings and prognostic markers of comorbid POF and DED. Future investigations should focus on detailed functional analyses of the identified hub genes and their interactions. This includes exploring their roles in key signaling pathways, regulatory networks, and disease processes. Additionally, large-scale clinical studies are necessary to validate the prognostic value of these genes and translate these findings into clinical practice.

Materials and methods

Data source and preprocessing

GEO (<http://www.ncbi.nlm.nih.gov/geo/>) is an extensive and publicly available gene expression database created and maintained by the National Center for Biotechnology Information (NCBI), containing high-throughput gene expression data submitted by research institutions worldwide. We used the inputs “dry eye” and “premature ovarian failure” to retrieve gene expression datasets for DED and POF. Inclusion criteria were as follows: (i) the sample size of each group was ≥ 3 , (ii) the samples in the normal and control groups were included, and (iii) the two datasets used for analyses were of the same species, and the dataset used for validation and analyses, respectively, should preferably be of different species. (iv) The data set must have access to raw and complete

data that can be exported for analysis. After screening the dataset, we preprocessed the matrix files obtained from the platform and matched the probes with their gene symbols according to the corresponding annotation files of the platform. Finally, we obtained gene matrices with row names as gene symbols and column names as sample names for subsequent analyses.

WGCNA

WGCNA considers biological functions as a whole using a gene network construction algorithm based on the expression similarity between genes. To identify DED- and POF-related gene coexpression modules, we used gene expression profiles to eliminate the top 50% genes with the smallest MAD and removed outlier genes and samples using the good-samples-genes method of the R package WGCNA, which was used to construct scale-free coexpression networks using a power function to construct a weighted neighbor-joining matrix. β is a soft-threshold parameter that emphasizes inter-gene strong correlations and penalizes weak correlations. Genes with similar expression profiles were grouped into gene modules, and the minimum size of the gene dendrogram was 30. The sensitivity was set to 3. To further analyze the modules, we computed the similarity of genes characterized by the modules, selected the cut lines of the module dendrograms, and merged the modules with distances < 0.25 . Modules with high correlation coefficients for DED and POF were selected, and genes from these modules were obtained for further analysis. We used an online tool (<https://jvenn.toulouse.sra.fr/app/example.html>) to determine the intersection of obtained genes using a Venn diagram, which can be used to obtain genes coexpressed in DED and POF.

Identification of DEGs

DEGs in DED and POF were identified using the R software limma package. Specifically, datasets were processed, converted into expression matrices, and grouped. The datasets were normalized and analyzed by using the limma software package and set adjusted with $P < 0.05$ and $\log_2|FC$ (fold change) ≥ 1 . The analysis identified DEGs, along with relevant information regarding their upregulated and downregulated expression. Considering the intersection between upregulated and downregulated expressed genes using the Venn diagram, genes common to DED and POF, respectively, were identified.

Enrichment analysis and functional annotation

To further elucidate the functions of the screened genes, we performed KEGG enrichment analysis and GO functional annotation of genes associated with the comorbidity of DED and POF, which were obtained using WGCNA and DEG analyses. For gene set functional enrichment analysis, we employed the KEGG rest API (<https://www.kegg.jp/kegg/rest/keggapi.html>) to obtain the latest gene annotations of the KEGG Pathway, and the R package clusterProfiler (version 3.14.3) to perform enrichment analysis. Enrichment analysis was performed with GO annotations of genes in the R package org.Hs.eg.db (version 3.1.0), including biological processes (BP), cellular components (CC), and molecular functions (MF), to obtain gene set enrichment results. A minimum gene set of 5 and a maximum gene set of 5000 were set, with $P < 0.05$ and $FDR < 0.1$ deemed statistically significant.

PPI network construction and modular analysis

To clarify the interaction relationship between the genes derived from the WGCNA and DEG analyses, STRING (<https://string-db.org>), an online tool for retrieving interacting genes, was used to construct a PPI network. The species was set as "homo sapiens," and the minimum required interaction score was set as high confidence (0.9), with obtained protein data imported into Cytoscape (<http://www.cytoscape.org>; version 3.8.2) for PPI network analysis and visualization. Network analysis was then performed using relevant tools (Analyze Network). The PPI network was analyzed for key functional modules using the MCODE plug-in to obtain closely functioning subnetworks in the network, with the following specific parameters: degree cutoff = 2, max depth = 100, node score cutoff = 0.2, and K-core = 2.

Identification and analysis of hub genes

The PPI genes were analyzed according to the degree-value algorithm using the CytoHubba plugin to understand the importance of individual genes in this network and identify hub genes. Using MCC, MNC, and Degree values as reference standards, we selected the top 15 hub genes and used GeneMANIA (<https://genemania.org/>) to generate the coexpression networks for identified hub genes.

Validation of hub genes

To further strengthen the confidence of the results, we used R software to validate the expression of hub genes using the GSE171043 and GSE48873 datasets. Statistical significance was set at $P < 0.05$.

Identification of common miRNAs and lncRNAs

Noncoding RNA transcripts, such as miRNAs and lncRNAs, play crucial roles in genetic regulation⁶¹. MiRNAs influence gene expression by either enhancing or inhibiting mRNA degradation and translation⁶². Conversely, lncRNAs are noncoding RNA molecules that typically comprise ~200 nucleotides and regulate various physiological and biochemical cellular processes by mediating chromosomal changes, transcriptional activation, and interference⁶³. Recent studies have highlighted the extensive crosstalk between miRNAs and lncRNAs, which involves competition for binding between miRNAs, lncRNAs, and other regulatory targets⁶⁴. Notably, competitive endogenous RNAs have been identified in which lncRNAs function by sequestering miRNAs. Therefore, in this

study, we aimed to investigate whether specific miRNAs and lncRNAs exhibit shared regulatory mechanisms and developmental processes in DED and POF using the Perl software.

Drug forecasting and drug-gene interaction analysis

To further understand the value of hub genes in clinical settings, we utilized the DGIdb database (<https://dgidb.genome.wustl.edu/>) for drug prediction and their interactions with the identified hub genes, thus exploring the treatment of comorbid POF and DED. Docking was performed after converting the compound and target protein formats to *pdbqt format using AutoDock software, setting up the folder and working directory, respectively, placing the prepared files, ensuring that the software was running, and adopting the default parameters of AutoDock Vina for docking. After obtaining the docking results, the docking results were converted from *pdbqt format to *pdb format using Open Babel (<https://openbabel.org/>) and then imported into PyMOL (<https://pymol.org/2/>) for visualization.

Machine learning model building

To further explore the diagnostic value of intersecting genes, we constructed a reliability predictive model using machine learning. Regression analysis was performed using LASSO with the glmnet package in R. Hub gene expression values in the GSE44101 and GSE39501 datasets were used as predictor variables, and their sample groups were used as response variables. Cross-validation (cv) of the model was performed using the cv.glmnet function, and CV plots were plotted to determine the optimal regularization parameters. Based on the coefficients of the LASSO regression, genes with non-zero coefficients were identified as selected signature genes. Next, we employed the sophisticated SVM-RFE algorithm with the e1071 package to construct another machine-learning model. These analyses provided a more nuanced understanding of these genes. Subsequently, we performed ROC analysis to construct models for assessing the predictive performance of the results and the diagnostic value of a factor. AUC values ≥ 0.5 and approaching 1 indicate high diagnostic accuracy. AUC values between 0.5 and 0.7 indicate low accuracy, 0.7 to 0.9 moderate accuracy, and ≥ 0.9 high accuracy. We used the R package pROC (v1.17.0.1) to perform the ROC analysis and obtain AUC values. Specifically, we obtained the disease status of the patient along with gene expression and evaluated the AUC and confidence intervals to clarify the diagnostic value of the genes in the disease.

Conclusion

The pathobiology of DED and POF involves an intricate tableau of multifactorial interactions encompassing a wide array of gene targets, pathways, signaling modalities, and regulatory frameworks. These components engage in a complex, often bidirectional interplay that exhibits both synergistic and antagonistic effects. Comorbid genes, including CDC20, BIRC5, PLK1, CDCA8, TOP2A, MCM5, MCM6, MCM7, MCM2, CENPA, FOXM1, GINS1, TIPIN, MAD2L1, and CDCA3, with processes encompassing DNA replication, the cell cycle, and the IL-17 signaling pathway, are central to this biological mosaic and are pivotal in catalyzing the biosynthesis of a spectrum of critical entities. The spotlight in this constellation of genes falls on CDC20, MAD2L1, and TIPIN, which is closely related to abnormal apoptosis and senescence. The functionality of these genes, whether in the activation or inhibition of metabolic pathways, is emblematic of the metabolic versatility integral to the pathogenesis of DED and POF.

Data availability

The datasets presented in this study can be found in GEO (<https://www.ncbi.nlm.nih.gov/geo/>). GSE44101, GSE39501, GSE171043, GSE48873 datasets were downloaded from the GEO database.

Received: 26 January 2024; Accepted: 9 July 2024

Published online: 13 July 2024

References

- Chon, S. J., Umair, Z. & Yoon, M. S. Premature ovarian insufficiency: Past, present, and future. *Front Cell Dev Biol.* **9**, 672890. <https://doi.org/10.3389/fcell.2021.672890> (2021).
- Torrealdy, S., Kodaman, P. & Pal, L. Premature Ovarian Insufficiency - an update on recent advances in understanding and management. *F1000Res.* **6**, 2069. <https://doi.org/10.12688/f1000research.11948.1> (2017).
- Kuchakzadeh, F., Ai, J. & Ebrahimi-Barough, S. Tissue engineering and stem cell-based therapeutic strategies for premature ovarian insufficiency. *Regen Ther.* **25**, 10–23 (2023).
- Rodgers, J. L. *et al.* Cardiovascular risks associated with gender and aging. *J. Cardiovasc. Dev. Dis.* **6**(2), 19 (2019).
- Stefanowski, B. *et al.* Cognitive decline and dementia in women after menopause: Prevention strategies. *Maturitas* **168**, 53–61 (2023).
- Nouri, N., Aghebati-Maleki, L. & Yousefi, M. Adipose-Derived mesenchymal stem cells: A promising tool in the treatment of premature ovarian failure. *J. Reprod Immunol.* **147**, 103363. <https://doi.org/10.1016/j.jri.2021.103363> (2021).
- Sheppard, J., Shen Lee, B. & Periman, L. M. Dry eye disease: identification and therapeutic strategies for primary care clinicians and clinical specialists. *Ann. Med.* **55**(1), 241–252 (2023).
- Pflugfelder, S. C. & Stern, M. E. Biological functions of tear film. *Exp. Eye Res.* **197**, 108115. <https://doi.org/10.1016/j.exer.2020.108115> (2020).
- Al-Saedi, Z. *et al.* Dry eye disease: Present challenges in the management and future trends. *Curr. Pharm Des.* **22**(28), 4470–4490 (2016).
- Rosenthal, P. Dry eye disease and the oculo-facial pain syndrome. *Int. Ophthalmol. Clin.* **57**(2), 41–45 (2017).
- Versura, P. & Campos, E. C. Menopause and dry eye. A possible relationship. *Gynecol. Endocrinol.* **20**(5), 289–298 (2005).
- Smith, J. A. *et al.* Dry eye signs and symptoms in women with premature ovarian failure. *Arch. Ophthalmol.* **122**(2), 151–156 (2004).
- Sumer, F., Subasi, S., Gurlek, B. & Ayazoglu, I. M. Meibography and tear function alterations in premature ovarian failure. *J. Fr. Ophthalmol.* **46**(10), 1169–1173 (2023).

14. İlgen, O., Yilmaz, F. & Kurt, S. Protective and/or therapeutic effects of berberine in a model of premature ovarian failure induced by cyclophosphamide in rats. *Turk. J. Obstet. Gynecol.* **20**(3), 227–233 (2023).
15. García-Catalán, M. R., Jerez-Olivera, E. & Benítez-Del-Castillo-Sánchez, J. M. Ojo seco y calidad de vida [Dry eye and quality of life]. *Arch. Soc. Esp. Oftalmol.* **84**(9), 451–458 (2009).
16. Craig, J. P. *et al.* TFOS DEWS II definition and classification report. *Ocul. Surf.* **15**(3), 276–283 (2017).
17. Volonte, D., Sedorovitz, M. & Galbiati, F. Impaired Cdc20 signaling promotes senescence in normal cells and apoptosis in non-small cell lung cancer cells. *J. Biol. Chem.* **298**(10), 102405. <https://doi.org/10.1016/j.jbc.2022.102405> (2022).
18. Sun, C. C. *et al.* Cisd2 plays an essential role in corneal epithelial regeneration. *EBio Med.* **7**, 103654. <https://doi.org/10.1016/j.ebiom.2021.103654> (2021).
19. Bai, Y., Zhang, K., Cao, X. & Chen, P. Aquaporins in lacrimal glands and their role in dry eye disease. *Exp. Eye Res.* **236**, 109676. <https://doi.org/10.1016/j.exer.2023.109676> (2023).
20. Hat, K., Kaštelan, S., Planinić, A., Muller, D. & Ježek, D. Pathohistological features of the aging human lacrimal gland. *Croat. Med. J.* **64**(5), 307–319 (2023).
21. Going, J. J. *et al.* Aberrant expression of minichromosome maintenance proteins 2 and 5, and Ki-67 in dysplastic squamous oesophageal epithelium and Barrett's mucosa. *Gut.* **50**(3), 373–377 (2002).
22. Zhang, Y. *et al.* The novel peptide PFAP1 promotes primordial follicle activation by binding to MCM5. *FASEB J.* **37**(5), e22931. <https://doi.org/10.1096/fj.202201495RR> (2023).
23. Swiech, L., Kisiel, K., Czolowska, R., Zientarski, M. & Borsuk, E. Accumulation and dynamics of proteins of the MCM family during mouse oogenesis and the first embryonic cell cycle. *Int. J. Dev. Biol.* **51**(4), 283–295 (2007).
24. Castro, A., Bernis, C., Vigneron, S., Labbé, J. C. & Lorca, T. The anaphase-promoting complex: A key factor in the regulation of cell cycle. *Oncogene.* **24**(3), 314–325 (2005).
25. Pflieger, C. M., Lee, E. & Kirschner, M. W. Substrate recognition by the Cdc20 and Cdh1 components of the anaphase-promoting complex. *Genes Dev.* **15**(18), 2396–2407 (2001).
26. Liu, M. *et al.* Evaluation of the antitumor efficacy of RNAi-mediated inhibition of CDC20 and heparanase in an orthotopic liver tumor model. *Cancer Biother. Radiopharm.* **30**(6), 233–239 (2015).
27. Wang, C. *et al.* Mutations in CCNB3 affect its location thus causing a multiplicity of phenotypes in human oocytes maturation by aberrant CDK1 activity and APC/C activity at different stages. *J. Ovar. Res.* **16**(1), 178 (2023).
28. Guo, C. *et al.* CDC20 inhibitor Apcin inhibits embryo implantation in vivo and in vitro. *Cell Biochem. Funct.* **38**(6), 810–816 (2020).
29. Liu, D. *et al.* Identification of transcriptome characteristics of granulosa cells and the possible role of UBE2C in the pathogenesis of premature ovarian insufficiency. *J. Ovarian Res.* **16**(1), 203 (2023).
30. Carmena, M., Wheelock, M., Funabiki, H. & Earnshaw, W. C. The chromosomal passenger complex (CPC): From easy rider to the godfather of mitosis. *Nat. Rev. Mol. Cell Biol.* **13**(12), 789–803 (2012).
31. Shuai, Y. *et al.* CDCA8 as an independent predictor for a poor prognosis in liver cancer. *Cancer Cell Int.* **21**(1), 159 (2021).
32. Cui, X. H. *et al.* Cell division cycle associated 8: A novel diagnostic and prognostic biomarker for hepatocellular carcinoma. *J. Cell Mol. Med.* **25**(24), 11097–11112. <https://doi.org/10.1111/jcmm.17032> (2021).
33. Jiang, D. *et al.* 2022 Cell division cycle-associated genes are potential immune regulators in nasopharyngeal carcinoma. *Front Oncol.* **12**, 779175. <https://doi.org/10.3389/fonc.2022.779175> (2022).
34. Ayad, N. G. *et al.* Tome-1, a trigger of mitotic entry, is degraded during G1 via the APC. *Cell.* **113**(1), 101–113 (2003).
35. Liu, X. *et al.* methylated in hepatocellular carcinoma, is a starvation stress sensor that induces apoptosis and autophagy through the AMPK-mTOR signaling cascade. *Canc. Lett.* **395**, 63–71 (2017).
36. Guo, P. *et al.* TRIM31 is upregulated in hepatocellular carcinoma and promotes disease progression by inducing ubiquitination of TSC1-TSC2 complex. *Oncogene* **37**(4), 478–488 (2018).
37. Leman, A. & Noguchi, E. Local and global functions of timeless and Tipin in replication fork protection. *Cell Cycle* **11**(21), 3945–3955 (2012).
38. Prorok, P. & Wolf, E. Cardoso MC 2024 timeless-tipin interactions with MCM and RPA mediate DNA replication stress response. *Front Cell Dev. Biol.* **12**, 1346534. <https://doi.org/10.3389/fcell.2024.1346534> (2024).
39. Kim, S. J. *et al.* Mitochondrial catalase overexpressed transgenic mice are protected against lung fibrosis in part via preventing alveolar epithelial cell mitochondrial DNA damage. *Free Radic. Biol. Med.* **101**, 482–490 (2016).
40. Wang, L. *et al.* Inhibition of MAD2L1 mediates pulmonary fibrosis through impairment of mitochondrial function and induction of cell senescence. *Can. Respir. J.* <https://doi.org/10.1155/2022/9663354> (2022).
41. Kanehisa, M., Furumichi, M., Sato, Y., Kawashima, M. & Ishiguro-Watanabe, M. KEGG for taxonomy-based analysis of pathways and genomes. *Nucl. Acids Res.* **51**, D587–D592 (2023).
42. Li, Z. *et al.* Research progress on the premature ovarian failure caused by cisplatin therapy. *Front. Oncol.* **13**, 1276310. <https://doi.org/10.3389/fonc.2023.1276310> (2023).
43. Chen, Q. *et al.* Involvement of aberrant acinar cell proliferation in scopolamine-induced dry eye mice. *Exp. Eye Res.* **227**, 109391. <https://doi.org/10.1016/j.exer.2023.109391> (2023).
44. Gu, C., Wu, L. & Li, X. IL-17 family: Cytokines, receptors and signaling. *Cytokine.* **64**(2), 477–485 (2013).
45. Faustine, G., Prijanti, A. R. & Wibowo, H. Effect of Lutein on ocular goblet cell, IFN- γ , and IL-17 concentration in dry eye-induced mice model. *Cureus.* **15**(7), e42009. <https://doi.org/10.7759/cureus.42009> (2023).
46. Harrington, L. E. *et al.* Interleukin 17-producing CD4+ effector T cells develop via a lineage distinct from the T helper type 1 and 2 lineages. *Nat. Immunol.* **6**(11), 1123–1132 (2005).
47. Yang, B., Sun, L. & Liang, L. MiRNA-802 suppresses proliferation and migration of epithelial ovarian cancer cells by targeting YWHAZ. *J. Ovarian Res.* **12**(1), 100 (2019).
48. Jarych, D. *et al.* Differential microRNA expression analysis in patients with HPV-infected ovarian neoplasms. *Int. J. Mol. Sci.* **25**(2), 762 (2024).
49. Yu, H., Xu, Y., Zhang, D. & Liu, G. Long noncoding RNA LUCAT1 promotes malignancy of ovarian cancer through regulation of miR-612/HOXA13 pathway. *Biochem. Biophys. Res. Commun.* **503**(3), 2095–2100 (2018).
50. Ji, Y. J., Shao, Y., Zhang, J., Zhang, X. & Qiang, P. Bromodomain-containing protein 4 silencing by microRNA-765 produces anti-ovarian cancer cell activity. *Aging (Albany NY).* **13**(6), 8214–8227 (2021).
51. Raudszus, R. *et al.* Pharmacological inhibition of TRPV2 attenuates phagocytosis and lipopolysaccharide-induced migration of primary macrophages. *Br. J. Pharmacol.* **180**(21), 2736–2749 (2023).
52. Gao, Z. *et al.* Valdecoxib protects against cell apoptosis induced by endoplasmic reticulum stress via the inhibition of PERK-ATF4-CHOP pathway in experimental glaucoma. *Int. J. Mol. Sci.* **23**(21), 12983. <https://doi.org/10.3390/ijms232112983> (2022).
53. Kahm, Y. J. & Kim, R. K. BIRC5: A novel therapeutic target for lung cancer stem cells and glioma stem cells. *Biochem. Biophys. Res. Commun.* **682**, 141–147 (2023).
54. Fuhr, L., Rousseau, M., Plauth, A., Schroeder, F. C. & Sauer, S. Amorfrutins are natural PPAR γ agonists with potent anti-inflammatory properties. *J. Nat. Prod.* **78**(5), 1160–1164 (2015).
55. Zou, Y. *et al.* Investigating the active chemical constituents and pharmacology of *Nanocnide lobata* in the treatment of burn and scald injuries. *PLoS One.* **18**(6), e0287147. <https://doi.org/10.1371/journal.pone.0287147> (2023).
56. Jiang, L. *et al.* Upregulation of endothelial cell-derived exosomal microRNA-125b-5p protects from sepsis-induced acute lung injury by inhibiting topoisomerase II α . *Inflamm. Res.* **70**(2), 205–216 (2021).

57. Wang, T., Zhang, Y., Bai, J., Xue, Y. & Peng, Q. MMP1 and MMP9 are potential prognostic biomarkers and targets for uveal melanoma. *BMC Canc.* **21**(1), 1068 (2021).
58. Hu, J., Zhou, S. & Guo, W. Construction of the coexpression network involved in the pathogenesis of thyroid eye disease via bioinformatics analysis. *Hum. Genomics.* **16**(1), 38 (2022).
59. Qin, Y., Wen, C. & Wu, H. CXCL10-based gene cluster model serves as a potential diagnostic biomarker for premature ovarian failure. *PEER J.* **11**, e16659. <https://doi.org/10.7717/peerj.16659> (2023).
60. Liu, B. *et al.* Comprehensive analysis of lncRNA-miRNA-mRNA ceRNA network and key genes in granulosa cells of patients with biochemical primary ovarian insufficiency. *J. Assist. Reprod. Genet.* **41**(1), 15–29 (2024).
61. Nemeth, K., Bayraktar, R., Ferracin, M. & Calin, G. A. Non-coding RNAs in disease: From mechanisms to therapeutics. *Nat Rev Genet.* **25**(3), 211–232 (2023).
62. Hudder, A. & Novak, R. F. miRNAs: Effectors of environmental influences on gene expression and disease. *Toxicol. Sci.* **103**(2), 228–240 (2008).
63. Mattick, J. S. *et al.* Long non-coding RNAs: Definitions, functions, challenges and recommendations. *Nat. Rev. Mol. Cell Biol.* **24**(6), 430–447 (2023).
64. Tan, J. Y. *et al.* Extensive microRNA-mediated crosstalk between lncRNAs and mRNAs in mouse embryonic stem cells. *Genome Res.* **25**(5), 655–666 (2015).

Acknowledgements

We thank all research group members for the helpful discussion.

Author contributions

X.L. proposed the concept of this research and drafted the manuscript. Z.W. was in charge of data collection and analysis, part of writing—original draft. P.J., K.T. and P.L. were in charge of data collection and curation, investigation. Q.P. conceived and designed this article, and was in charge of project administration and writing—review and editing.

Funding

National Natural Science Foundation of China (81804150, 81574031); 2021 “Academician Liu Liang’s Workstation” guiding project (21YS002); Hunan Provincial Department of Education key project (21A0238); Hunan Provincial Natural Science Foundation Program (2023JJ40479).

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Q.P.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024