



OPEN ACCESS

EDITED BY

Hao Peng,
Renmin Hospital of Wuhan University,
China

REVIEWED BY

Qun Xue,
Affiliated Hospital of Nantong University,
China

Bin Wang,
Sichuan University, China

Jinwei Li,
Department of Neurosurgery, West China
Hospital, Sichuan University, China

*CORRESPONDENCE

Lian Zou,
✉ jzxfckzl@163.com

RECEIVED 16 July 2023

ACCEPTED 05 September 2023

PUBLISHED 03 October 2023

CITATION

Zou L, Meng L, Xu Y, Wang K and Zhang J (2023), Revealing the diagnostic value and immune infiltration of senescence-related genes in endometriosis: a combined single-cell and machine learning analysis.
Front. Pharmacol. 14:1259467.
doi: 10.3389/fphar.2023.1259467

COPYRIGHT

© 2023 Zou, Meng, Xu, Wang and Zhang. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Revealing the diagnostic value and immune infiltration of senescence-related genes in endometriosis: a combined single-cell and machine learning analysis

Lian Zou^{1*}, Lou Meng¹, Yan Xu¹, Kana Wang² and Jiawen Zhang²

¹Chongqing Emergency Medical Center, Department of Obstetrics and Gynecology in Chongqing University Central Hospital, Chongqing, China, ²Department of Gynecology, West China Second Hospital of Sichuan University, Chengdu, China

Introduction: Endometriosis is a prevalent and recurrent medical condition associated with symptoms such as pelvic discomfort, dysmenorrhea, and reproductive challenges. Furthermore, it has the potential to progress into a malignant state, significantly impacting the quality of life for affected individuals. Despite its significance, there is currently a lack of precise and non-invasive diagnostic techniques for this condition.

Methods: In this study, we leveraged microarray datasets and employed a multifaceted approach. We conducted differential gene analysis, implemented weighted gene co-expression network analysis (WGCNA), and utilized machine learning algorithms, including random forest, support vector machine, and LASSO analysis, to comprehensively explore senescence-related genes (SRGs) associated with endometriosis.

Discussion: Our comprehensive analysis, which also encompassed profiling of immune cell infiltration and single-cell analysis, highlights the therapeutic potential of this gene assemblage as promising targets for alleviating endometriosis. Furthermore, the integration of these biomarkers into diagnostic protocols promises to enhance diagnostic precision, offering a more effective diagnostic journey for future endometriosis patients in clinical settings.

Results: Our meticulous investigation led to the identification of a cluster of genes, namely BAK1, LMNA, and FLT1, which emerged as potential discerning biomarkers for endometriosis. These biomarkers were subsequently utilized to construct an artificial neural network classifier model and were graphically represented in the form of a Nomogram.

KEYWORDS

machine learning, immune infiltration, endometriosis, senescence-related genes, aging, integrative bioinformatics, senescence-associated molecular

1 Introduction

Endometriosis (EM) is one of the most prevalent chronic inflammatory gynecological maladies, affecting nearly 10% of women in the stage of childbearing maturation. Its effects extend to the quality of life and overall wellbeing of a significant number of women (Liang and Yao, 2016; Becker et al., 2022). Its principal indications are pelvic discomfort, dysmenorrhea, and uterine hemorrhaging. Such symptoms affect nearly half of patients, leading to infertility and other complications (Eskenazi and Warner, 1997; Giudice and Kao, 2004; Meuleman et al., 2009). EM has emerged as a public health concern, putting substantial pressure upon women and their families. Nonetheless, there are insufficient precise diagnostic and therapeutic interventions. The pelvic discomfort characteristic of EM can be erroneously identified as abdominal pain originating from the menstrual cycle, thereby confounding the diagnostic process. Moreover, there is a consistent period of latency spanning 7–11 years, from the onset of symptomatic presentation to the eventual moment of definitive diagnosis (Greene et al., 2009; Rogers et al., 2013). Laparoscopic surgery is a benchmark for the diagnosis of EM. However, it is not widely supported as a diagnostic modality due to its considerable financial burden and potentially adverse effects on the physiological state of women (Giudice and Kao, 2004; Li et al., 2021).

A complete understanding of the pathogenesis underlying EM remains elusive. Several factors, including retrograde menstruation, cytokine involvement, estrogen influence, inflammatory reactions, and ectopic implantation, are considered potential etiological contributors. Among these, the process of epithelial–mesenchymal transition (EMT) is gaining recognition in research (Zondervan et al., 2018). Although the mechanism of the senescence genes in EM is not fully comprehended, endometrial epithelial cells are often characterized by senescence in lesions (Velarde and Menon, 2016; Brighton et al., 2017). By applying bioinformatic analyses, such as enrichment evaluations, a significant nexus between EM and the cellular cycle has been established. Notably, the affirmative enrichment of the cellular cycle within the context of EMT concurs with the cycle being intrinsically linked to genes associated with senescence. Furthermore, it has been found that the expression of SRGs is attenuated in the context of EM (Wang et al., 2022). Telomere correlation analysis also revealed that the average telomere length is shorter in women with EM (Gleason et al., 2022). Cellular senescence (CS) is a stress-induced response to various biological signals that results in stable cell-cycle arrest (Salama et al., 2014). CS is often associated with DNA damage, and DNA damage kinase is associated with telomere shortening (d’Adda di Fagagna, 2008; Harley et al., 1990).

Given such interrelations, SRGs show considerable promise as candidates for EM diagnosis. This investigation made a strategic choice to amalgamate the dataset of SRGs with the EM dataset obtained from both single-cell sequencing repositories and diverse transcriptomic compilations culled from public databases. Such integration facilitated comprehensive bioinformatic analyses, the formulation of diagnostic models, and the execution of single-cell scrutiny. We thereby corroborate the diagnostic potential of SRGs within the context of EM.

2 Materials and methods

2.1 Collection of gene microarray data and data pre-processing for endometriosis

The Gene Expression Omnibus (GEO) database, hosted on NCBI (<https://www.ncbi.nlm.nih.gov/geo/>), supplied three gene expression microarray datasets related to EM: GSE12768-GPL7304 (Borghese et al., 2008), GSE11691-GPL96 (Hull et al., 2008), and GSE7305-GPL570 (Hever et al., 2007). For comprehensive details concerning these datasets, see [Supplementary Table S1](#). We included a cohort of patients who had been definitively diagnosed with EM from the GEO database and transcribed and analyzed their normal and diseased endometrial human tissue. In addition, cohorts with complete patient follow-up information and complete clinical information, and cohorts with complete micro-matrix data, were required to ensure data quality for subsequent bioinformatics analysis. To minimize the influence of other gynecological conditions, we excluded subjects who may have concurrently suffered from other important gynecologic conditions to maintain the cohort’s integrity. To facilitate agreement with official platform annotations, probe IDs were transformed into gene symbols, wherein probes with multiple associated gene symbols and those without were systematically excluded. For instance, when multiple gene symbols were ascribed to a single gene, the median value was computed to represent the unique expression value. Datasets GSE12768 and GSE11691 were thus merged, and, to mitigate the influence of batch disparities within the microarray expression data, the ComBat function encompassed within the R package “sva” was used for batch effect rectification (Roder et al., 2019). The resulting processed microarrays were subsequently employed as the intrinsic training dataset, while GSE7305 functioned as the extraneous validation dataset. For this research, 307 SRGs (SRGs) were extracted from the Human Senescence Genome Resource (HAGR, <http://genomics.senescence.info/genes/>) (Tacutu et al., 2018). The complete enumeration of these genes is provided in [Supplementary Table S2](#).

2.2 Identification of differentially expressed SRGs

The gene expression microarrays specific to EM were intersected with the set of 307 SRGs, revealing their shared genes. This common gene pool was subsequently subjected to differential analysis. To facilitate this, the “limma” R package was employed, adopting established thresholds (p -value < 0.05 and $\text{Log}_2[\text{FC}] > 0.5$) to discriminate genes displaying differential expression (Song et al., 2022a; Zhao et al., 2022; Song et al., 2022b). A selection of 45 genes, termed differentially expressed SRGs (DE-SRGs), was thus culled from this analysis.

The “ggplot2” and “pheatmap” R packages were used to present these outcomes (Chi et al., 2022a; Shen et al., 2022; Huang et al., 2023). The former created a visually informative volcano plot, while the latter constructed a heatmap depicting the expression patterns of the DE-SRGs. Additionally, deviation maps were formulated to underscore the divergences in gene expression tendencies between the EM and healthy cohorts. This was achieved using

the “ggpubr” R package (version 4.0.2) (<https://github.com/kassambara/ggpubr>). These deviation maps delineate the trajectory of gene expression profiles within both the EM and healthy groups.

2.3 Functional enrichment analysis of differentially expressed SRGs

Utilizing the clusterProfiler R package (Jin et al., 2022; Chi et al., 2023a; Ren et al., 2023) (<https://bioconductor.org/packages/release/bioc/html/clusterProfiler.html>), a comprehensive enrichment analysis was executed for the DE-SRGs in the context of EM. This analysis interrogated the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways, Gene Ontology (GO) terms, and Disease Ontology (DO). The objective was to identify statistically significant biological functionalities and enrichment pathways, with strict criteria set at p -values <0.05 and Q -values <0.05 to ensure robustness. Prior to enrichment analysis, preprocessing was undertaken to ensure uniform gene annotation by transforming all gene symbols into Entrez IDs, which was facilitated by the human genome-wide annotation R package “org.Hs.e.g.,.db.” GO comprises three distinct components: biological process (BP), cellular component (CC), and molecular function (MF).

2.4 Feature gene identification by integrating three machine learning algorithms

The cohort of 45 DE-SRGs underwent a meticulous screening process wherein three distinct machine learning algorithms were harmoniously amalgamated: support vector machine-recursive feature elimination (SVM-RFE), random forest (RF), and the least absolute shrinkage and selection operator (LASSO) (Chi et al., 2023b; Song et al., 2023; Zhang et al., 2023). SVM-RFE, an advance upon the sequential backward selection algorithm rooted in the SVM’s tenet of maximal margin, delivers superior and more proficient classification performance, particularly for high-dimensional datasets. The realization of SVM-RFE was facilitated using the “e1071” R package, a tool that identifies relevant attributes while discarding redundant ones (Mundra and Rajapakse, 2010). The RF algorithm embodies a high-precision and sensitivity-oriented integration technique that integrates multiple decision trees and amalgamates the outcomes of diverse classifiers into a unified decision, achieved through the utilization of the “randomForest” R package (Strobl et al., 2007). This investigation prioritized the relative significance among the 45(DE-SRGs, ultimately selecting candidate genes whose relative importance surpassed the threshold of 0.1. Due to its reputation for proficient dimensionality reduction, LASSO emerged as an exceptional tool. To optimize its performance, a rigorous assessment was undertaken involving ten-fold cross-validations, orchestrated through the “glmnet” R package, to discern the optimal tuning parameter (λ) (Friedman et al., 2008). The optimal configuration of λ was discerned as the point of minimal magnitude. In comparison with alternative algorithms, LASSO was very efficient at evaluating high-

dimensional datasets. The common genes extracted through the amalgamation of these three machine-learning algorithms were subsequently harnessed as biomarkers for the prediction of endometriosis.

2.5 Building an artificial neural network disease classification model and nomogram

Artificial neural networks (ANNs) are important in the field of deep learning within artificial intelligence systems, serving as consequential analytical instruments. Typically comprising three strata, ANN models encompass an input layer for information reception, one or more hidden layers for information processing, and an output layer for result computation. In this investigation, the “neuralnet” (Beck, 2018) R package was employed to construct ANN models founded upon three SRGs. These were orchestrated with a configuration comprising five hidden layers, and the gene weight data were harnessed to formulate a classification model primed for EM disease prediction. Similar parameters and processes were extended to an external dataset for model validation. ROC curves were computed employing the “pROC” R package to elicit AUC values, which are pivotal in evaluating the predictive efficacy of the ANN model. In parallel, a nomogram based on the three SRGs was meticulously formulated through the “rms” R package. Calibration curves and decision curve analysis (DCA) were used to visualize the discriminatory capacity of the model.

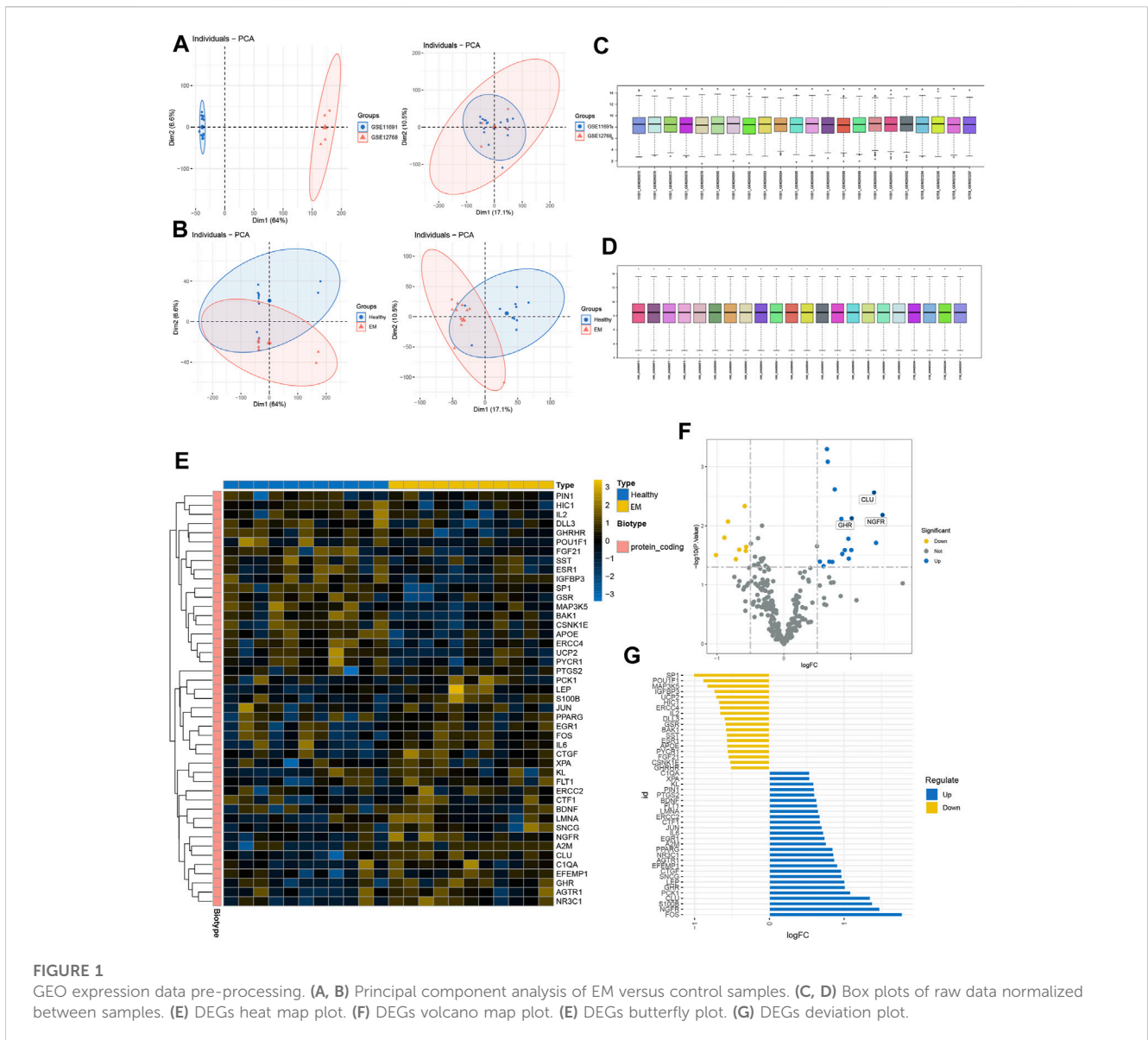
2.6 Network construction and functional enrichment analysis of WGCNA co-expression

The WGCNA algorithm clusters highly correlated genes into gene set modules, guided by a scale-free topological criterion and reveals the intricate interplay between these modules and disease traits. For this research, the “WGCNA” R package (Langfelder and Horvath, 2008) was pivotal, independently making connections between phenotypes (EM or health) and individual modules. The module exhibiting the highest correlation coefficient coupled with the most statistically significant p -value emerged as the principal module associated with endometriosis.

Subsequently, the genes nested within the pivotal modules were subjected to enrichment analysis which utilized Metascape data (<http://metascape.org>) (Zhou et al., 2019). The outcomes of this analysis highlighted the top 100 results from the enrichment, and a network diagram was generated that employed the top 20 clustering outcomes. In this diagram, the thickness of connecting lines served as an indicator of the similarity score, thus visually representing the intricate relationships established.

2.7 Immune cell infiltration abundance analysis

The R package IOBR (Zeng et al., 2021) (Immuno-Oncology Biology Research) amalgamated the four algorithms, TIMER,

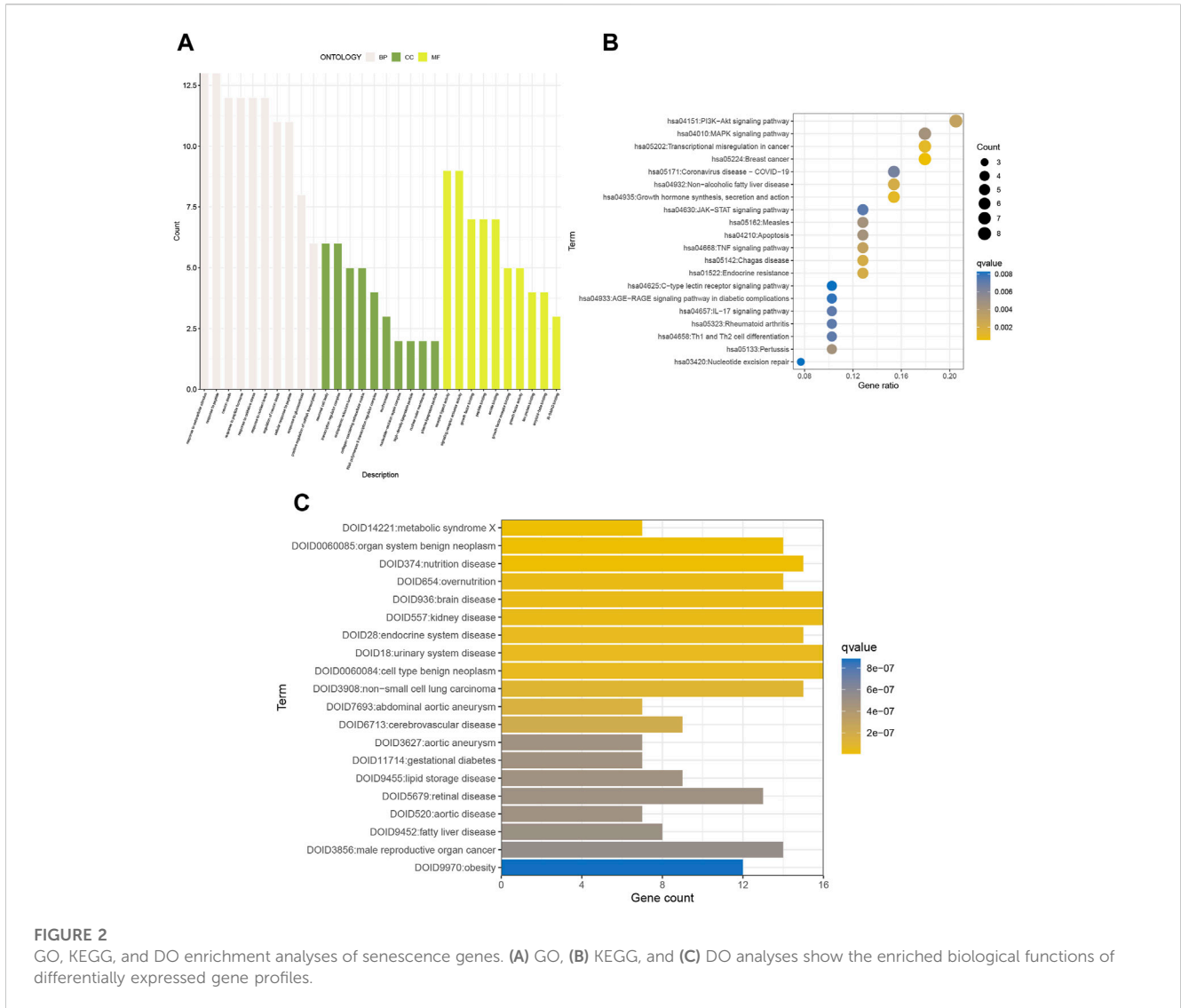


ESTIMATE, xCell, and CIBERSORT, which served as pivotal tools for quantifying the presence of immune cell infiltrations. In the context of this inquiry, the CIBERSORT algorithm emerged as the preferred choice, delineating the relative composition of 22 immune cell types. This characterization was achieved by discerning gene expression patterns in samples, ensuring that the cumulative presence of each immune cell type within a singular sample equated to 1. Spearman correlation was harnessed to unravel the intricate associations, enabling an exploration of the linkages between the three SRGs and the various constituents of the immune system.

2.8 Data processing and analysis of scRNA-seq

The 10x scRNA-seq data were processed as follows: 1) 10x scRNA-seq data were converted to Seurat objects using the R

software Seurat package (Macosko et al., 2015). 2) Counts were quality-controlled (quality control (QC)) by excluding low-quality cells based on mitochondrial or ribosomal gene percentages. 3) Screening of the top 2000 high variable genes after QC using the “FindVariableFeatures” function. 4) Principal component analysis (PCA) and uniform manifold approximation and projection (UMAP) were based on the 2000 genes for downscaling and cluster identification (Becht et al., 2018). 5) The “FindAllMarkers” function identified significant marker genes in different clusters by setting log₂FC to 0.3 and minimum pct to 0.25. 6) Our analysis of cluster annotations was performed using the SingleR package in the R software (Aran et al., 2019). Next, we performed Fisher precision tests to identify potentially important cell types. We calculated FC values for each cell type in tumor and normal samples and identified cell types with FC > 4 or FC < 0.25, with p < 0.05 being the critical cell type.



3 Results

3.1 GEO expression data pre-processing and analysis of variance

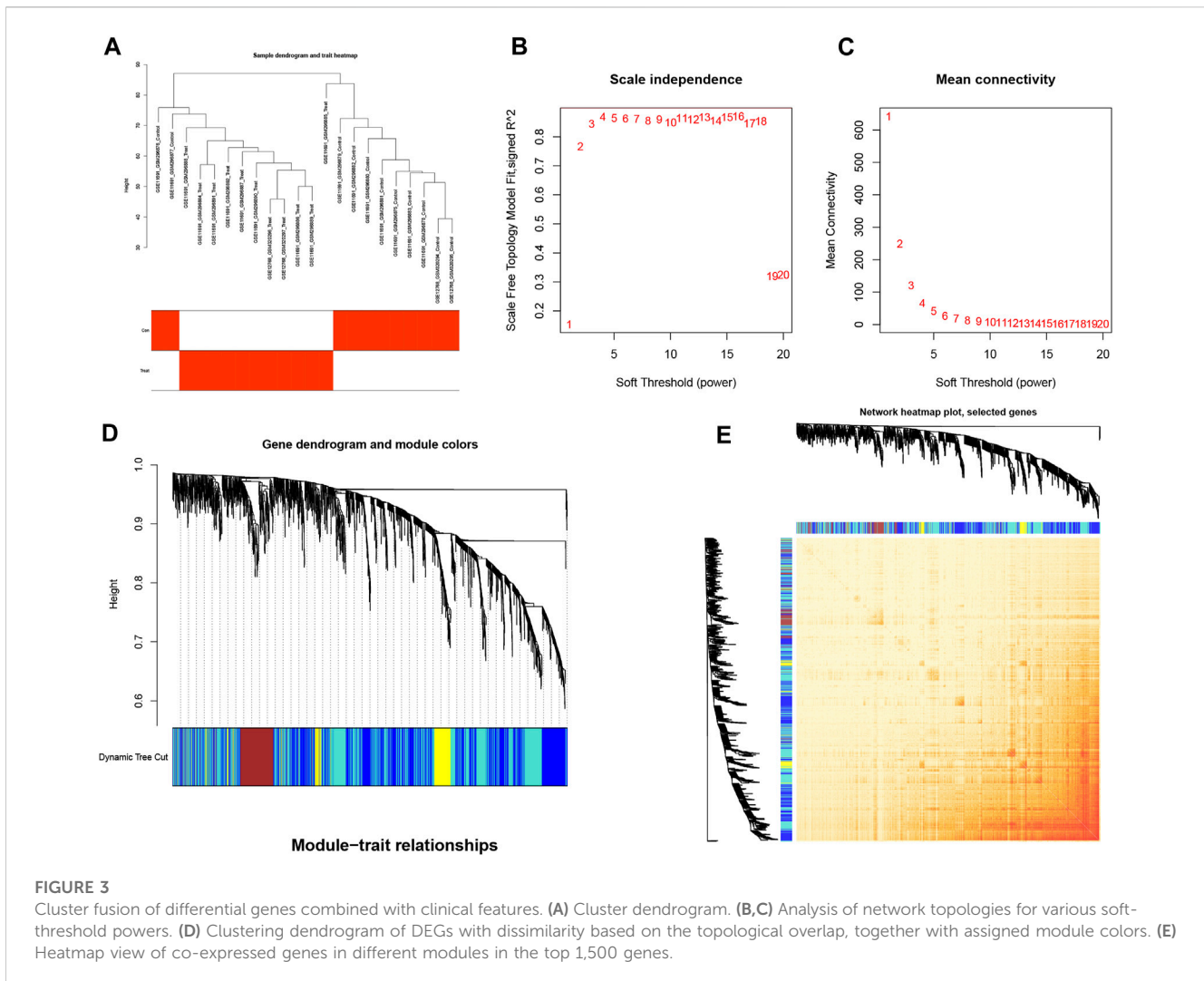
We first normalized the gene expression data of GSE11691 and GSE12768 with quantile normalization and merged the datasets with the R package “sva” to remove batch effects. Figure 1A shows a PCA of the two datasets before and after processing; as shown in the figure, the two datasets were initially separated without any intersection; after processing, their intersection can be used as a batch for subsequent analysis. Figure 1B shows the PCA results for the healthy and EM groups in the dataset, and Figures 1C, D show box plots of the two datasets after normalization—the different colors represent different samples and the columns indicate the gene expression values in the samples. We then performed disease differential gene analysis on the combined datasets and took the intersection with the collected senescence genes.

As shown in the figure, the intersection of the two datasets can be used as a batch of data for subsequent analysis. A total of 45 genes were identified as DEGs with 18 upregulated and 27 downregulated

genes under the criteria of p -adjustment < 0.05 and $\log_2 | \text{fold-change (FC)} | > 0.5$. Figure 1F shows the volcano map of the DEGs and the heatmap of the top 50 genes (Figure 1E). Finally, significantly different DEGs are displayed in the deviation plot (Figure 1G).

3.2 GO, KEGG, and DO enrichment analyses of senescence genes

We performed KEGG pathway analysis, GO functional analysis, and DO disease analysis to further evaluate the biological function of the differential genes. Consequently, we used p -value < 0.01 and Q -value < 0.05 as thresholds to identify items that were significantly enriched. Biological process (BP) includes response to extracellular stimulus, response to peptides, and the like. Cellular component (CC) includes neuronal cell body and transcriptional regulatory complex. Molecular function (MF) includes receptor ligand activity and signaling receptor activator activity (Figure 2A). KEGG analysis revealed enrichment in breast cancer, growth hormone synthesis, secretion, and action,



transcriptional misregulation in cancer, and other pathways associated with gynecological cancers (Figure 2B). DO analysis found an endocrine system disease—urinary system disease (Figure 2C). In conclusion, the differential genes have a great impact on cellular receptive signaling-related functions and pathways and presumably may play a major role in disease mechanisms.

3.3 Cluster fusion of differential genes combined with clinical features

We used WGCNA to analyze associations with clinical features of EM and differential genes, where the samples were clustered according to clinical features of the disease; the cohort was processed using the correlation coefficient method to obtain a sample clustering tree (Figure 3A). The samples were then analyzed by hierarchical clustering to construct a gene co-expression network with a soft threshold of 18 ($R^2 = 0.8$) and higher average connectivity (Figures 3B, C). By clustering and fusing the characteristic genes into the differential genes and merging the highly similar modules, five modules were ultimately obtained (Figure 3D). By showing the

relative independence of gene expression within these modules, the analysis of the combined results did not reveal any significant differences between the different modules (Figure 3E).

3.4 Identifying high-relevance modules and analyzing related functions

In an endeavor to quantify the interrelation between each module and the two clinical attributes, heatmaps were created to illustrate the intricate network of interconnections within the module features (Figure 4A). These heatmaps display the correlations between trait genes and the traits associated with obesity. Within this context, an exhaustive co-expression analysis of DEGs was undertaken, which identified the turquoise module as the epicenter of particularly robust correlation. This assertion was reinforced by scatter plots which yielded 106 hub genes that displayed an exceptionally high level of connectivity within the turquoise module ($cor = 0.78$, $p < 1e-200$) (Figure 4B). KEGG and GO enrichment analyses determined the functional significance of these 106 hub genes by employing the Metascape library. This identified over 100 prime genes with profound enrichment in lipid-related pathways and functions.

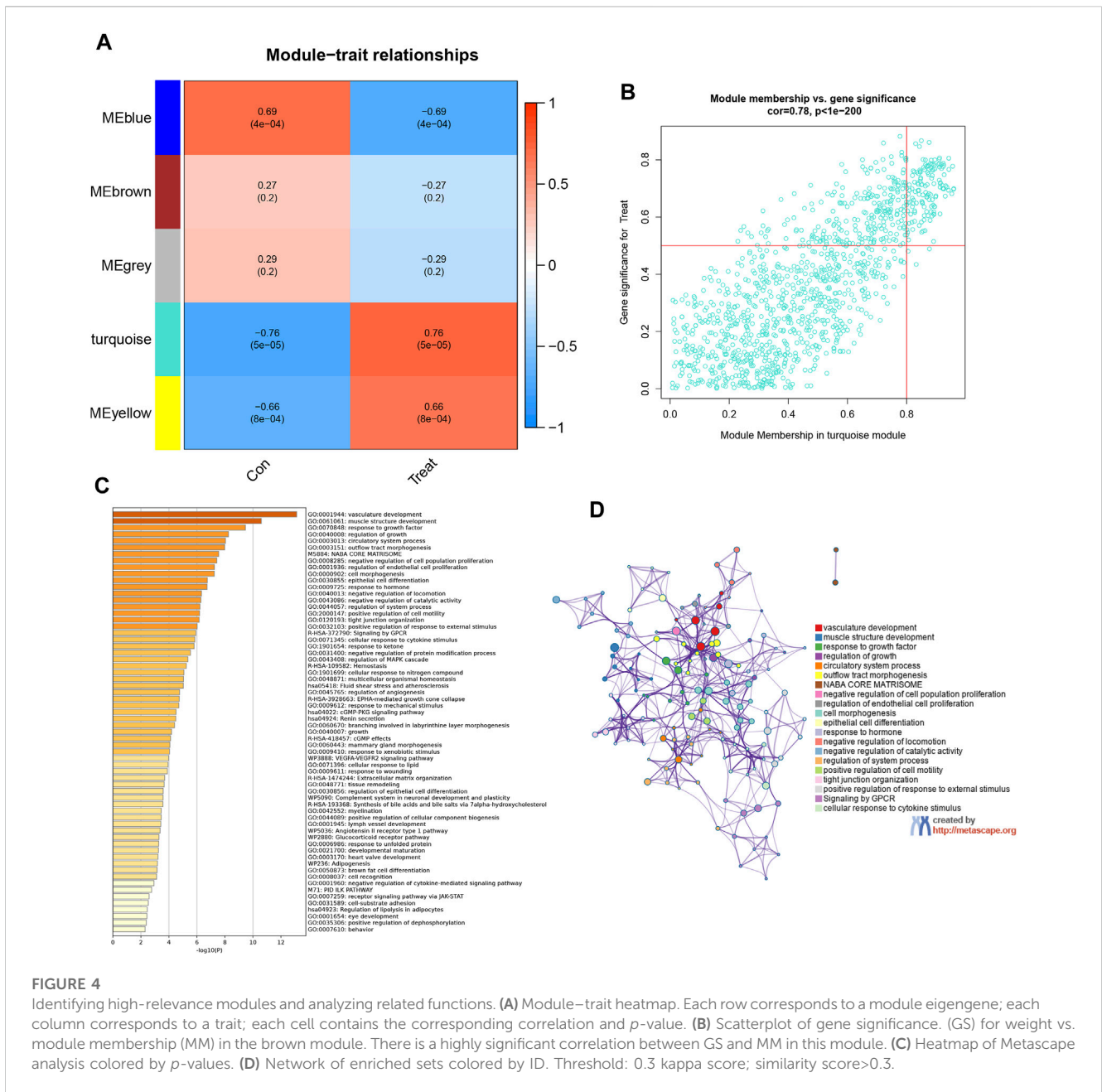


FIGURE 4

Identifying high-relevance modules and analyzing related functions. (A) Module-trait heatmap. Each row corresponds to a module eigengene; each column corresponds to a trait; each cell contains the corresponding correlation and p-value. (B) Scatterplot of gene significance. (GS) for weight vs. module membership (MM) in the brown module. There is a highly significant correlation between GS and MM in this module. (C) Heatmap of Metascape analysis colored by p-values. (D) Network of enriched sets colored by ID. Threshold: 0.3 kappa score; similarity score>0.3.

These genes also demonstrated close associations with blood vessel establishment and morphological structures (Figures 4C, D). Such interplay strongly suggests that the hub genes may have a substantial impact on EM within hormone synthesis and action.

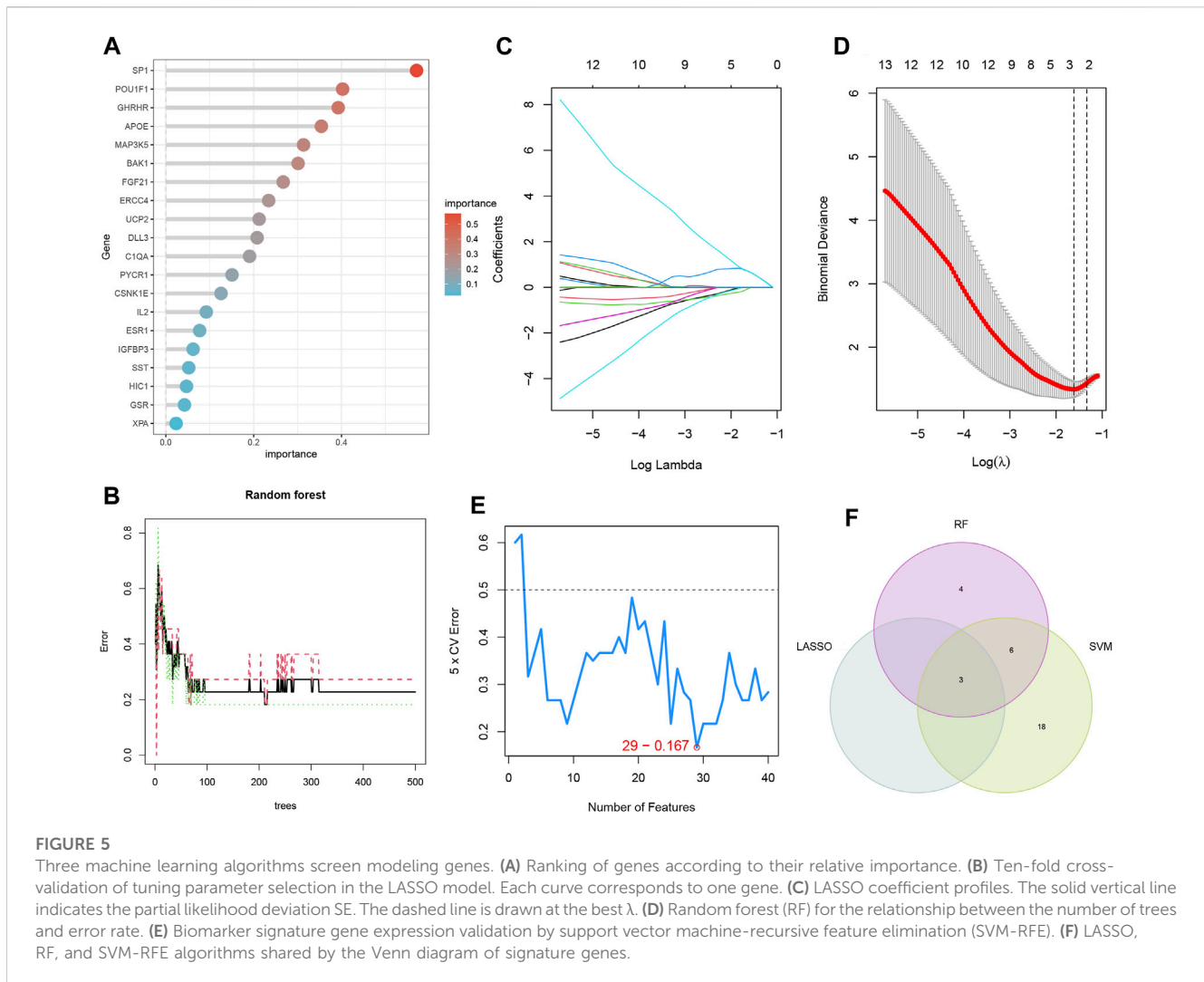
3.5 Three machine learning algorithms to screen for modeled genes

The LASSO, RF, and SVM-RFE algorithms were used to screen feature genes among the differentially expressed genes associated with key EM progression and senescence processes. For the LASSO algorithm, we selected after cross-validation the minimum criteria for constructing the LASSO classifier to identify three feature genes (Figures 5C, D); the error was minimized when the number of

features was 29 (Figure 5E). Thus, 29 relevant feature genes were obtained, and the top 13 with importance greater than 0.1 were selected from the classification tree results (Figures 5A, B) in combination with RF feature selection. Through crossover, the three feature genes shared by LASSO, RF, and SVM-RFE algorithms were finally identified, namely, BAK1, FLT1, and LMNA, and their relationships were represented by the Venn diagram (Figure 5F).

3.6 Development and validation of ANN models and nomogram based on three SRGs

We ultimately selected three pivotal SRGs to construct a hidden layer number 5 for the ANN model for clinical EM diagnosis



(Figure 6A). The satisfactory reliability and accuracy of the prediction model were validated by the ROC curves, and it was found that the AUC values were greater than 0.85 in both the training (AUC = 0.822) and the validation cohorts (AUC = 0.895) (Figures 6E, F). To further promote the clinical application of the three SRG biomarkers in this study, nomogram plots were also drawn (Figure 6B) to calculate patient scores based on the expression data of the three biomarkers to infer the probability of prevalence in EM patients. In addition, both the DCA and calibration curves indicated a higher net benefit and accuracy of using the nomogram for diagnostic prediction in EM patients (Figures 6C, D).

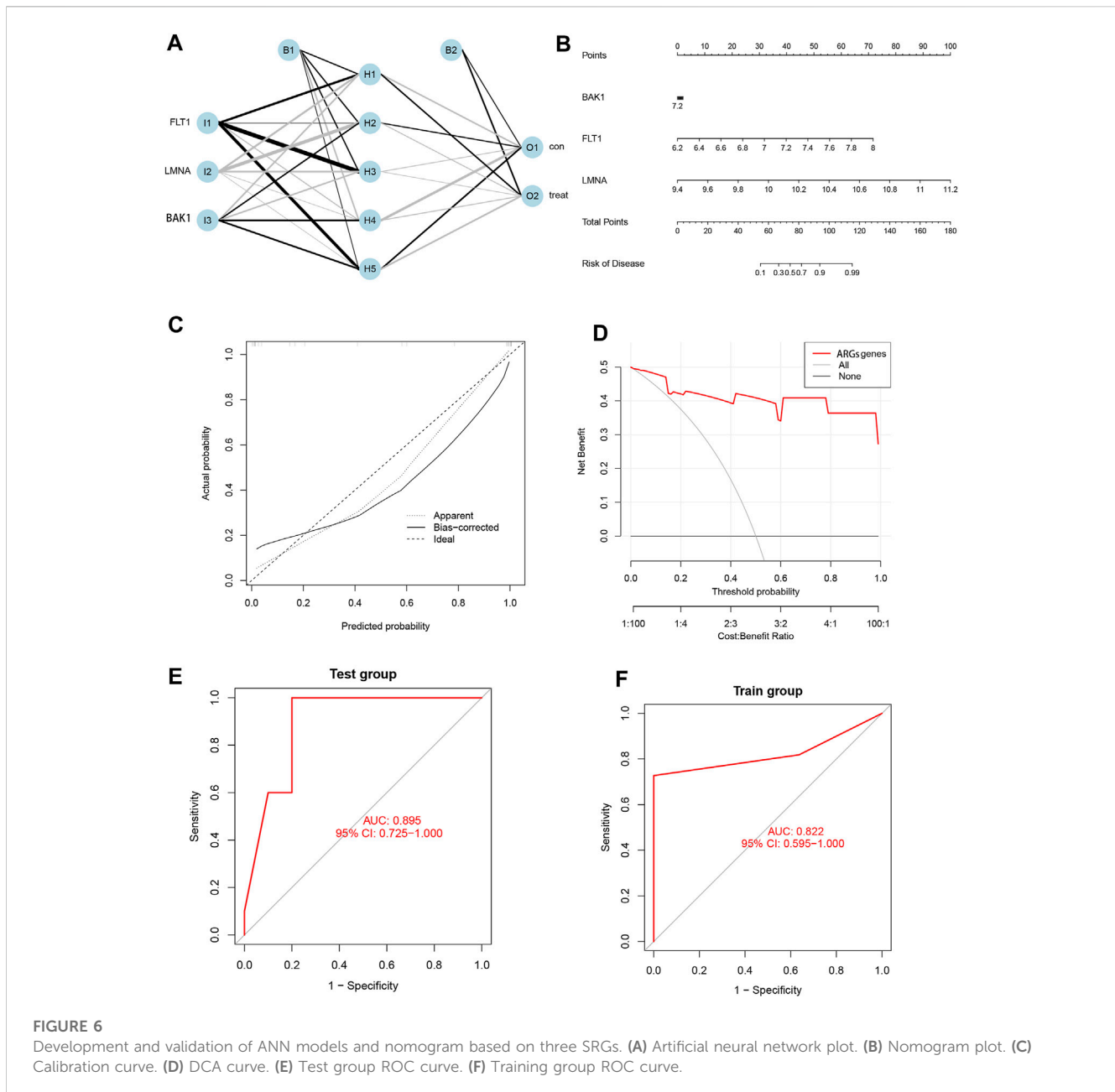
3.7 Detection of immunological features in EM and normal samples

Immune cells play an essential role in endometrial shedding, tissue repair, and preventing infection (Shen et al., 2021), although a disproportionate number of immune cells in an abnormal immune environment is thought to play a key role in EM pathogenesis (Wu et al., 2017). The CIBERSORT algorithm in IOBR was used to

calculate the proportion of 22 immune cell infiltrates in EM tissues. A stacked plot was used to fully visualize the distribution of the immune infiltrate in the EM tissue (Figure 7A), with the sum of all immune cell proportions being 1. A significant increase in B cells naive and NK cells activated in the patient's uterine tissue was observed, while the proportion of immune cells such as macrophages m1 and m2 decreased significantly (Figure 7B). In addition, the correlation between the three SRGs and immune cells was explored based on the permanent method (Figure 7C), where FLT1 and Macrophages_M1 had the highest positive correlation while LMNA and NK_cells_activated had the highest negative correlation; their correlation scatter plots are shown in Figures 7D, E.

3.8 Single-cell analysis

Following quality control procedures applied to the single-cell dataset, a crucial phase was PCA-based dimensionality reduction, where the determination of the optimal number of selected PCs was substantiated by insights gleaned from the JackStrawPlot and

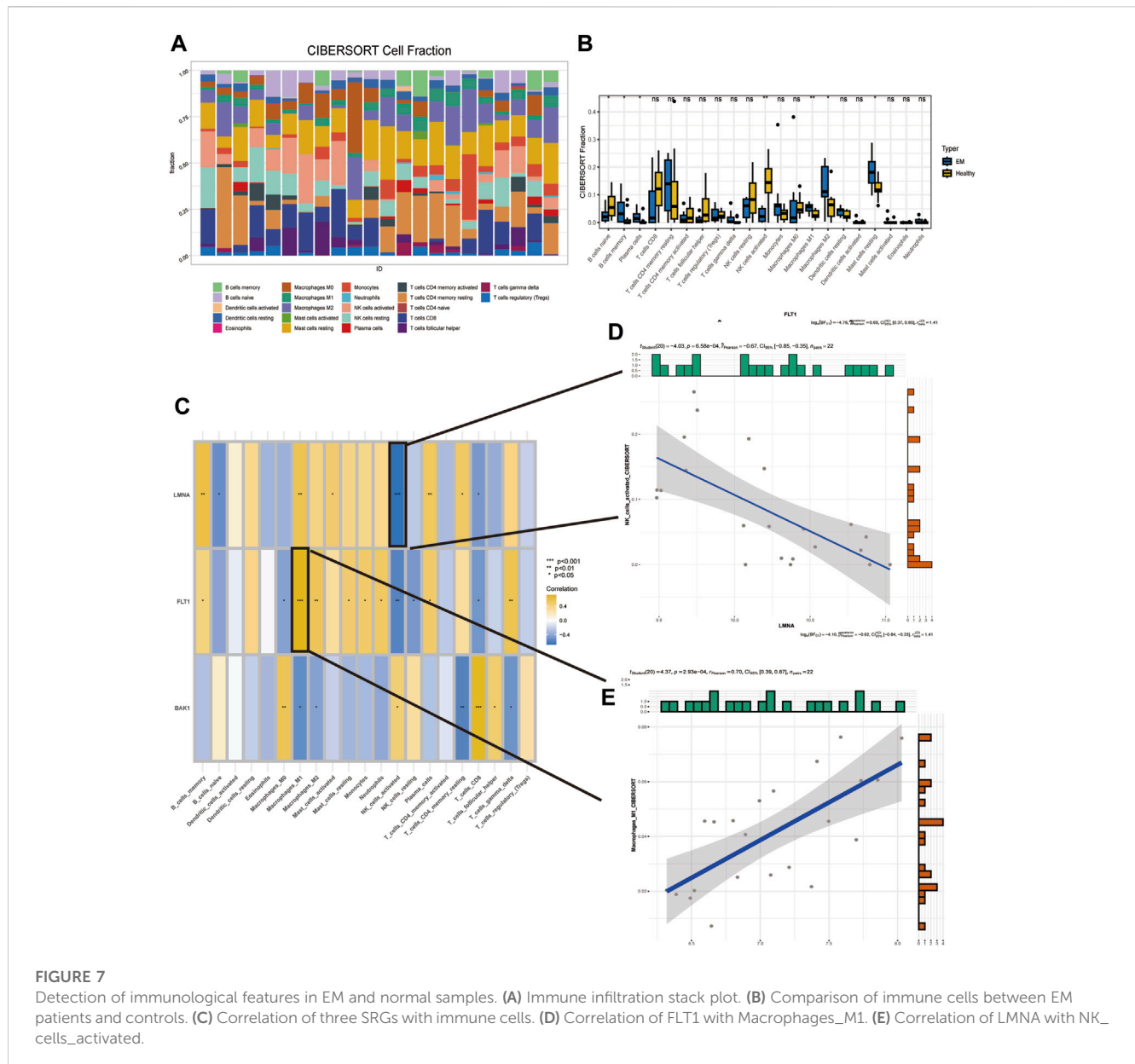


ElbowPlot plot analyses (Figures 8A, B). Subsequently, the “FindClusters” function harnessed from the “Seurat” package was deployed to cluster the individual cells, which were further annotated using SingleR. The resulting visualization was realized through t-SNE and UMAP downscaling techniques, thereby culminating in the identification and characterization of six distinct cell types (Figures 8C, D). The exploration extended to evaluating the expression occupancy of the three modeled genes—BAK1, FLT1, and LMNA—within these six cell types. Intriguingly, the results revealed that LMNA exhibited heightened expression across nearly all cell types, whereas FLT1 demonstrated pronounced expression solely within epithelial cells (Figures 8E, F). We then explored deeper into ligand–receptor networks and the delineation of specific pathways, thus facilitating the inference of cell–cell

communication networks. Based on these findings, MHC-I and TGFb emerged as prominent drivers of signaling communication within NK cells and epithelial cells, respectively (Figures 8G, H).

4 Discussion

Despite affecting approximately 10% of women within the reproductive age bracket, the intricate connection between EM and various health ramifications, including early natural menopause (ENM), remains enigmatic. A substantial corpus of clinically documented cases underscores the notion that EM is intrinsically related to ovarian senescence, thereby exerting a tangible influence on the temporal span leading to menopause

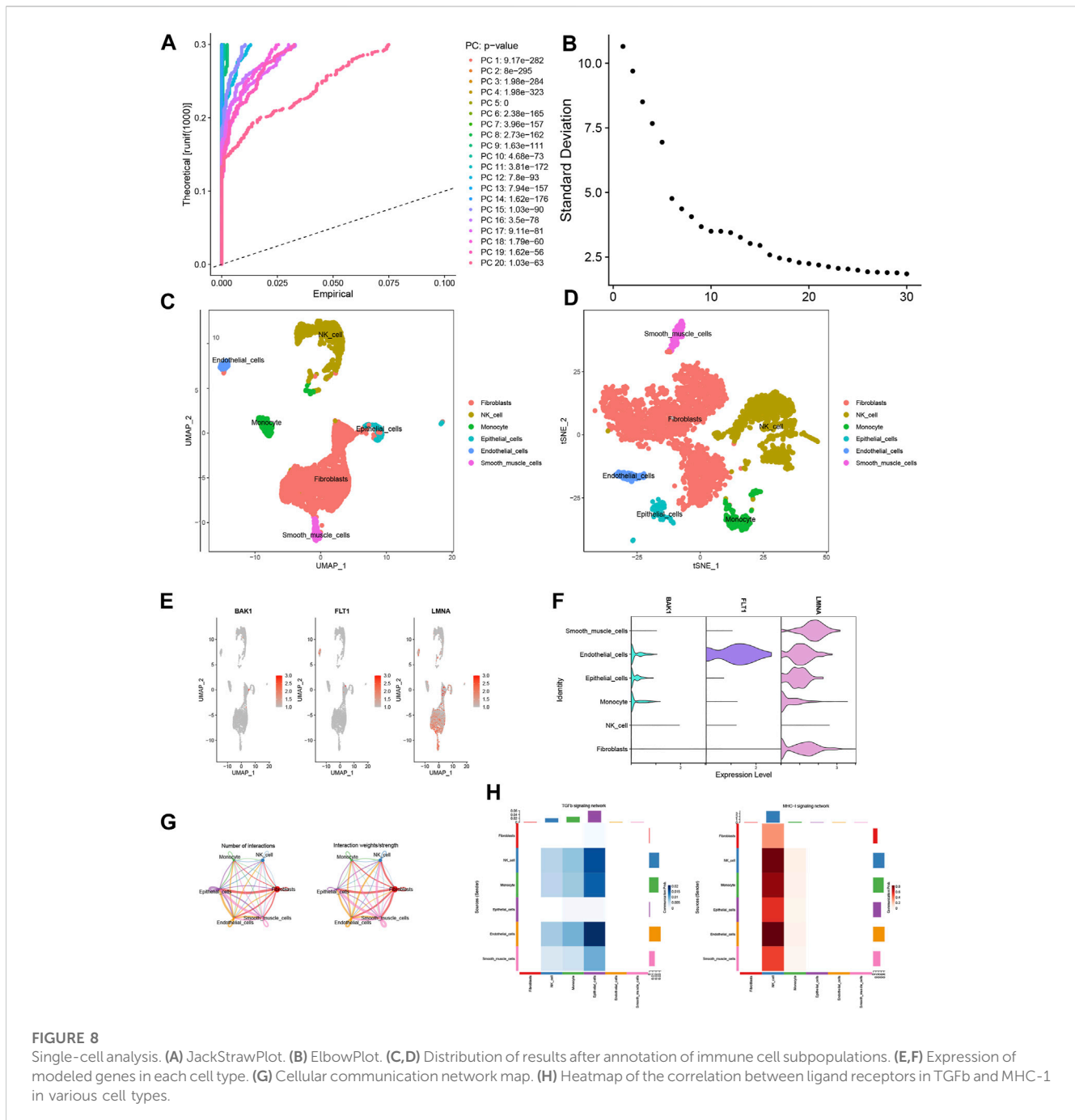


(Pokoradi et al., 2011). Nonetheless, it is crucial to note that these linkages exhibit non-linear characteristics and manifest inconsistency across distinct markers. Moreover, factors that introduce perturbations to ovarian functionality, encompassing instances such as ovarian surgery, autoimmune variations, smoking habits, and the utilization of androgenic exogenous hormones, have the potential to reverberate across the ovarian reserve. This resonance is characterized by diminished follicle counts and compromised oocyte quality and quantity. Consequently, this is closely aligned with an earlier onset of menopause and the initiation of ovarian senescence (Seyhan et al., 2015; Romanski et al., 2019).

Consequently, the computation of senescence-related genes (DEGs) is intricately linked with EM. This was accomplished through a co-expression network analysis of genes differentially expressed within the context of EM. Subsequently, leveraging the prowess of LASSO, SVM-RFE, and RF identified three

candidate attributes: FLT1, LMNA, and BAK1. These attributes revealed insights from the gene selection process. Within this framework, the neural network model assumed a defining role, identifying the predictive weights attributed to the associated genes. This led to the establishment of a classification model score for EM. Finally, the correlation between the modeled genes and the immune characteristics of EM patients was explored.

Angiogenesis emerges as a plausible player in the intricate pathogenic fabric of EM, wherein FLT1—a prominent receptor for vascular endothelial growth factor (VEGF)—is pivotal as a major regulator of pro-angiogenic factors. Differential expression of FLT1 characterizes all subtypes of EM in comparison to their normal tissue counterparts. Furthermore, the emergence of high-risk red lesions may be attributed to elevated levels of VEGF and its receptors (FLT1 and FLT2), instigating an augmentation in the subperitoneal vascular network. This then fosters the propagation of



malignant retroperitoneal gap implantation (Donnez et al., 1998; Aydin et al., 2021).

Changes in the hypermethylation status of the CpG islets of the laminin A/C (LMNA) gene are associated with insulin resistance in patients with polycystic ovary syndrome (PCOS), suggesting that this gene may be involved in the regulation of PCOS-associated insulin resistance (Ting et al., 2013). In addition, insulin resistance is an important pathological mechanism in endometrial cancer and EM (Wei and Li, 2020), and mutations in LMNA are thought to be a causal factor in non-valvular atrial fibrillation (NVAf) (Saj et al., 2012) and Hutchinson–Gilford premature senescence syndrome (HGFS) (Auld et al., 2010).

The Bcl-2 family gene BCL2 Antagonist/Killer 1 (BAK1) was found expressed at high levels in the apoptotic tissues of several clinical diseases (Zhou et al., 2008; Xu et al., 2011; Magee et al., 2014), and EM has been shown by some scholars to be a pathological consequence of inflammation and apoptosis (Laganà et al., 2019); thus, the high expression of BAK1 is involved in the pathomechanical process that regulates inflammation and apoptosis in the endometrium of EM patients.

The peritoneal immune milieu plays a pivotal role in the intricate EM pathophysiology (Vallvé-Juanico et al., 2019). This environment encompasses compromised natural killer cells, perturbed T-cell differentiation, and autoantibodies stemming

from activated B cells (Chi et al., 2022b; Zhao et al., 2022; Chi et al., 2023c; Xiong et al., 2023). Consequently, it is critical to scrutinize the interplay that links senescence-associated prognostic genes with the diverse immune attributes exhibited by EM patients. This exploration is driven by the aim of formulating potential immunotherapeutic regimens or pharmaceutical agents for clinical implementation. Moreover, it seeks to enhance prevailing clinical treatment protocols, thereby heightening their overall efficacy and engendering superior therapeutic outcomes.

This investigation seeks to leverage machine learning algorithms to sift through the genetic signatures of EM to create a novel diagnostic framework grounded in artificial neural networks. Its ultimate objective is tangible benefits for clinical patients. However, it is important to acknowledge certain inherent limitations within this study. First is the relatively modest sample size within the cohort, which may not well capture the broader population dynamics, consequently influencing the applicability and generalization of the diagnostic model. Additionally, it is important to recognize that the diagnostic model is predicated on preliminary findings and lacks robust experimental validation essential to substantiate its clinical viability and dependability. Given these constraints, the diagnostic model necessitates further in-depth scrutiny and exploration to ascertain its potential utility in shaping clinical decision-making processes.

In conclusion, EM imposes an immeasurable burden beyond its localized symptoms, with far-reaching effects on all aspects of overall health. The chronic pain, reproductive challenges, and potential complications associated with it not only severely impact the physical health of patients but also their mental and emotional wellbeing, placing a heavy economic burden on society. We have identified a better diagnostic model than previously available (Ji et al., 2022), and these data highlight the need for further research to uncover senescence-related mechanisms in EM. It has laid some foundation and direction for more in-depth research afterward.

5 Conclusion

Within this investigation, a trio of pivotal senescence hub genes—BAK1, LMNA, and FLT1—were meticulously chosen. This selection was grounded in comprehensive bioinformatics analysis, and these hub genes were leveraged to create a “classifier” and forge a nomogram plot dedicated to endometriosis (EM). Furthermore, the study revealed the existence of two discernible senescence-associated subtypes, coupled with the discernment of pivotal regulatory pathways and the intricate landscape of the immune microenvironment.

References

- Aran, D., Looney, A. P., Liu, L., Wu, E., Fong, V., Hsu, A., et al. (2019). Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. *Nat. Immunol.* 20 (2), 163–172. doi:10.1038/s41590-018-0276-y
- Auld, D., Shen, M., and Thomas, C. (2010). “A high-throughput screen for pre-mRNA splicing modulators,” in *Probe reports from the NIH molecular libraries program* (Bethesda (MD): National Center for Biotechnology Information US).

This research may deliver fresh perspectives concerning the future of drug interventions and the molecular underpinnings of EM.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

Author contributions

LZ: writing—original draft. LM: writing—original draft. YX: writing—original draft. KW: writing—original draft. JZ: writing—original draft.

Funding

The authors declare that no financial support was received for the research, authorship, and/or publication of this article.

Conflict of interest

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

The reviewer JL declared a shared parent affiliation with the author(s) KW & JZ to the handling editor at the time of review.

Publisher’s note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fphar.2023.1259467/full#supplementary-material>

- Aydin, G. A., Ayyaci, H., Koc, N., Tarhan, N., and Demirci, O. (2021). The relationship between decorin and VEGF in endometriosis. *J. Coll. Physicians Surg. Pak* 31 (11), 1285–1290. doi:10.29271/jcsp.2021.11.1285

- Becht, E., McInnes, L., Healy, J., Dutertre, C. A., and Kwok, I. W. H. (2018). Dimensionality reduction for visualizing single-cell data using UMAP. *Nat. Biotechnol.* 37, 38–44. doi:10.1038/nbt.4314

- Beck, M. W. (2018). NeuralNetTools: Visualization and analysis tools for neural networks. *J. Stat. Softw.* 85 (11), 1–20. doi:10.18637/jss.v085.i11
- Becker, C. M., Bokor, A., Heikinheimo, O., Horne, A., Jansen, F., Kiesel, L., et al. (2022). ESHRE guideline: Endometriosis. *Hum. Reprod. Open* 2022 (2), hoac009. doi:10.1093/hropen/hoac009
- Borghese, B., Mondon, F., Noël, J. C., Fayt, I., Mignot, T. M., Vaiman, D., et al. (2008). Gene expression profile for ectopic versus eutopic endometrium provides new insights into endometriosis oncogenic potential. *Mol. Endocrinol.* 22 (11), 2557–2562. doi:10.1210/me.2008-0322
- Brighton, P. J., Maruyama, Y., Fishwick, K., Vrljicak, P., Tewary, S., Fujihara, R., et al. (2017). Clearance of senescent decidual cells by uterine natural killer cells in cycling human endometrium. *Elife* 6, e31274. doi:10.7554/eLife.31274
- Chi, H., Chen, H., Wang, R., Zhang, J., Jiang, L., Zhang, S., et al. (2023b). Proposing new early detection indicators for pancreatic cancer: Combining machine learning and neural networks for serum miRNA-based diagnostic model. *Front. Oncol.* 13, 1244578. doi:10.3389/fonc.2023.1244578
- Chi, H., Gao, X., Xia, Z., Yu, W., Yin, X., Pan, Y., et al. (2023a). FAM family gene prediction model reveals heterogeneity, stemness and immune microenvironment of UCEC. *Front. Mol. Biosci.* 10, 1200335. doi:10.3389/fmolb.2023.1200335
- Chi, H., Jiang, P., Xu, K., Zhao, Y., Song, B., Peng, G., et al. (2022a). A novel anoiikis-related gene signature predicts prognosis in patients with head and neck squamous cell carcinoma and reveals immune infiltration. *Front. Genet.* 13, 984273. doi:10.3389/fgene.2022.984273
- Chi, H., Xie, X., Yan, Y., Peng, G., Strohmayer, D. F., Lai, G., et al. (2022b). Natural killer cell-related prognosis signature characterizes immune landscape and predicts prognosis of HNSCC. *Front. Immunol.* 13, 1018685. doi:10.3389/fimmu.2022.1018685
- Chi, H., Zhao, S., Yang, J., Gao, X., Peng, G., Zhang, J., et al. (2023c). T-cell exhaustion signatures characterize the immune landscape and predict HCC prognosis via integrating single-cell RNA-seq and bulk RNA-sequencing. *Front. Immunol.* 14, 1137025. doi:10.3389/fimmu.2023.1137025
- d'Adda di Fagnana, F. (2008). Living on a break: Cellular senescence as a DNA-damage response. *Nat. Rev. Cancer* 8 (7), 512–522. doi:10.1038/nrc2440
- Donnez, J., Smoes, P., Gillerot, S., Casanas-Roux, F., and Nisolle, M. (1998). Vascular endothelial growth factor (VEGF) in endometriosis. *Hum. Reprod.* 13 (6), 1686–1690. doi:10.1093/humrep/13.6.1686
- Eskenazi, B., and Warner, M. L. (1997). Epidemiology of endometriosis. *Obstet. Gynecol. Clin. North Am.* 24 (2), 235–258. doi:10.1016/s0889-8545(05)70302-8
- Friedman, J., Hastie, T., and Tibshirani, R. (2008). Sparse inverse covariance estimation with the graphical lasso. *Biostatistics* 9 (3), 432–441. doi:10.1093/biostatistics/kxm045
- Giudice, L. C., and Kao, L. C. (2004). Endometriosis. *Lancet* 364 (9447), 1789–1799. doi:10.1016/S0140-6736(04)17403-5
- Gleason, J. L., Thoma, M. E., Zukerman Willinger, N., and Shenassa, E. D. (2022). Endometriosis and uterine fibroids and their associations with elevated C-reactive protein and leukocyte telomere length among a representative sample of U.S. Women: Data from the national health and nutrition examination survey, 1999–2002. *J. Womens Health (Larchmt)* 31 (7), 1020–1028. doi:10.1089/jwh.2021.0044
- Greene, R., Stratton, P., Cleary, S. D., Ballweg, M. L., and Sinaii, N. (2009). Diagnostic experience among 4,334 women reporting surgically diagnosed endometriosis. *Fertil. Steril.* 91 (1), 32–39. doi:10.1016/j.fertnstert.2007.11.020
- Harley, C. B., Futcher, A. B., and Greider, C. W. (1990). Telomeres shorten during ageing of human fibroblasts. *Nature* 345 (6274), 458–460. doi:10.1038/345458a0
- Hever, A., Roth, R. B., Hevezi, P., Marin, M. E., Acosta, J. A., Acosta, H., et al. (2007). Human endometriosis is associated with plasma cells and overexpression of B lymphocyte stimulator. *Proc. Natl. Acad. Sci. U. S. A.* 104 (30), 12451–12456. doi:10.1073/pnas.0703451104
- Huang, X., Chi, H., Gou, S., Guo, X., Li, L., Peng, G., et al. (2023). An aggrephagy-related lncRNA signature for the prognosis of pancreatic adenocarcinoma. *Genes (Basel)* 14 (1), 124. doi:10.3390/genes14010124
- Hull, M. L., Escareno, C. R., Godtsland, J. M., Doig, J. R., Johnson, C. M., Phillips, S. C., et al. (2008). Endometrial-peritoneal interactions during endometriotic lesion establishment. *Am. J. Pathol.* 173 (3), 700–715. doi:10.2353/ajpath.2008.071128
- Ji, X., Huang, C., Mao, H., Zhang, Z., Zhang, X., Yue, B., et al. (2022). Identification of immune- and autophagy-related genes and effective diagnostic biomarkers in endometriosis: A bioinformatics analysis. *Ann. Transl. Med.* 10 (24), 1397. doi:10.21037/atm-22-5979
- Jin, W., Yang, Q., Chi, H., Wei, K., Zhang, P., Zhao, G., et al. (2022). Ensemble deep learning enhanced with self-attention for predicting immunotherapeutic responses to cancers. *Front. Immunol.* 13, 1025330. doi:10.3389/fimmu.2022.1025330
- Laganà, A. S., Garzon, S., Götte, M., Viganò, P., Franchi, M., Ghezzi, F., et al. (2019). The pathogenesis of endometriosis: Molecular and cell Biology insights. *Int. J. Mol. Sci.* 20 (22), 5615. doi:10.3390/ijms20225615
- Langfelder, P., and Horvath, S. (2008). Wgcna: an R package for weighted correlation network analysis. *BMC Bioinforma.* 9, 559. doi:10.1186/1471-2105-9-559
- Li, B., Duan, H., Wang, S., and Li, Y. (2021). Ferroptosis resistance mechanisms in endometriosis for diagnostic model establishment. *Reprod. Biomed. Online* 43 (1), 127–138. doi:10.1016/j.rbmo.2021.04.002
- Liang, Y., and Yao, S. (2016). Potential role of estrogen in maintaining the imbalanced sympathetic and sensory innervation in endometriosis. *Mol. Cell. Endocrinol.* 424, 42–49. doi:10.1016/j.mce.2016.01.012
- Macosko, E. Z., Basu, A., Satija, R., Nemes, J., Shekhar, K., Goldman, M., et al. (2015). Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* 161 (5), 1202–1214. doi:10.1016/j.cell.2015.05.002
- Magee, T. R., Ross, M. G., Wedekind, L., Desai, M., Kjos, S., and Belkacemi, L. (2014). Gestational diabetes mellitus alters apoptotic and inflammatory gene expression of trophoblasts from human term placenta. *J. Diabetes Complicat.* 28 (4), 448–459. doi:10.1016/j.jdiacomp.2014.03.010
- Meuleman, C., Vandenabeele, B., Fieuws, S., Spiessens, C., Timmerman, D., and D'Hooghe, T. (2009). High prevalence of endometriosis in infertile women with normal ovulation and normospermic partners. *Fertil. Steril.* 92 (1), 68–74. doi:10.1016/j.fertnstert.2008.04.056
- Mundra, P. A., and Rajapakse, J. C. (2010). SVM-RFE with MRMR filter for gene selection. *IEEE Trans. Nanobioscience* 9 (1), 31–37. doi:10.1109/TNB.2009.2035284
- Pokoradi, A. J., Iversen, L., and Hannaford, P. C. (2011). Factors associated with age of onset and type of menopause in a cohort of UK women. *Am. J. Obstet. Gynecol.* 205 (1), 34.e1–13. doi:10.1016/j.ajog.2011.02.059
- Ren, Q., Zhang, P., Lin, H., Feng, Y., Chi, H., Zhang, X., et al. (2023). A novel signature predicts prognosis and immunotherapy in lung adenocarcinoma based on cancer-associated fibroblasts. *Front. Immunol.* 14, 1201573. doi:10.3389/fimmu.2023.1201573
- Roder, J., Oliveira, C., Net, L., Tsy-pin, M., Linstid, B., and Roder, H. (2019). A dropout-regularized classifier development approach optimized for precision medicine test discovery from omics data. *BMC Bioinforma.* 20 (1), 325. doi:10.1186/s12859-019-2922-2
- Rogers, P. A., D'Hooghe, T. M., Fazleabas, A., Giudice, L. C., Montgomery, G. W., Petraglia, F., et al. (2013). Defining future directions for endometriosis research: Workshop report from the 2011 world congress of endometriosis in montpellier, France. *Reprod. Sci.* 20 (5), 483–499. doi:10.1177/1933719113477495
- Romanski, P. A., Brady, P. C., Farland, L. V., Thomas, A. M., and Hornstein, M. D. (2019). The effect of endometriosis on the antimüllerian hormone level in the infertile population. *J. Assist. Reprod. Genet.* 36 (6), 1179–1184. doi:10.1007/s10815-019-01450-9
- Saj, M., Dabrowski, R., Labib, S., Jankowska, A., Szperl, M., Broda, G., et al. (2012). Variants of the lamin A/C (LMNA) gene in non-valvular atrial fibrillation patients: A possible pathogenic role of the Thr528Met mutation. *Mol. Diagn. Ther.* 16 (2), 99–107. doi:10.1007/BF03256434
- Salama, R., Sadaie, M., Hoare, M., and Narita, M. (2014). Cellular senescence and its effector programs. *Genes. Dev.* 28 (2), 99–114. doi:10.1101/gad.235184.113
- Seyhan, A., Ata, B., and Uncu, G. (2015). The impact of endometriosis and its treatment on ovarian reserve. *Semin. Reprod. Med.* 33 (6), 422–428. doi:10.1055/s-0035-1567820
- Shen, H. H., Zhang, T., Yang, H. L., Lai, Z. Z., Zhou, W. J., Mei, J., et al. (2021). Ovarian hormones-autophagy-immunity axis in menstruation and endometriosis. *Theranostics* 11 (7), 3512–3526. doi:10.7150/tno.55241
- Shen, Y., Chi, H., Xu, K., Li, Y., Yin, X., Chen, S., et al. (2022). A novel classification model for lower-grade glioma patients based on pyroptosis-related genes. *Brain Sci.* 12 (6), 700. doi:10.3390/brainsci12060700
- Song, B., Chi, H., Peng, G., Song, Y., Cui, Z., Zhu, Y., et al. (2022b). Characterization of coagulation-related gene signature to predict prognosis and tumor immune microenvironment in skin cutaneous melanoma. *Front. Oncol.* 12, 975255. doi:10.3389/fonc.2022.975255
- Song, B., Wu, P., Liang, Z., Wang, J., Zheng, Y., Wang, Y., et al. (2022a). A novel necroptosis-related gene signature in skin cutaneous melanoma prognosis and tumor microenvironment. *Front. Genet.* 13, 917007. doi:10.3389/fgene.2022.917007
- Song, G., Peng, G., Zhang, J., Song, B., Yang, J., Xie, X., et al. (2023). Uncovering the potential role of oxidative stress in the development of periodontitis and establishing a stable diagnostic model via combining single-cell and machine learning analysis. *Front. Immunol.* 14, 1181467. doi:10.3389/fimmu.2023.1181467
- Strobl, C., Boulesteix, A. L., Zeileis, A., and Hothorn, T. (2007). Bias in random forest variable importance measures: Illustrations, sources and a solution. *BMC Bioinforma.* 8, 25. doi:10.1186/1471-2105-8-25
- Tacutu, R., Thornton, D., Johnson, E., Budovsky, A., Barardo, D., Craig, T., et al. (2018). Human ageing genomic resources: New and updated databases. *Nucleic Acids Res.* 46 (D1), D1083–d1090. doi:10.1093/nar/gkx1042
- Ting, W., Yanyan, Q., Jian, H., Keqin, H., and Duan, M. (2013). The relationship between insulin resistance and CpG island methylation of LMNA gene in polycystic ovary syndrome. *Cell. Biochem. Biophys.* 67 (3), 1041–1047. doi:10.1007/s12013-013-9602-z
- Vallvé-Juanico, J., Houshdaran, S., and Giudice, L. C. (2019). The endometrial immune environment of women with endometriosis. *Hum. Reprod. Update* 25 (5), 564–591. doi:10.1093/humupd/dmz018

- Velarde, M. C., and Menon, R. (2016). Positive and negative effects of cellular senescence during female reproductive aging and pregnancy. *J. Endocrinol.* 230 (2), R59–R76. doi:10.1530/JOE-16-0018
- Wang, M., Wu, Y., He, Y., Liu, J., Chen, Y., Huang, J., et al. (2022). SIRT1 upregulation promotes epithelial-mesenchymal transition by inducing senescence escape in endometriosis. *Sci. Rep.* 12 (1), 12302. doi:10.1038/s41598-022-16629-x
- Wei, C., and Li, X. (2020). The role of photoactivated and non-photoactivated verteporfin on tumor. *Front. Pharmacol.* 11, 557429. doi:10.3389/fphar.2020.557429
- Wu, J., Xie, H., Yao, S., and Liang, Y. (2017). Macrophage and nerve interaction in endometriosis. *J. Neuroinflammation* 14 (1), 53. doi:10.1186/s12974-017-0828-3
- Xiong, J., Chi, H., Yang, G., Zhao, S., Zhang, J., Tran, L. J., et al. (2023). Revolutionizing anti-tumor therapy: Unleashing the potential of B cell-derived exosomes. *Front. Immunol.* 14, 1188760. doi:10.3389/fimmu.2023.1188760
- Xu, C., Hao, Y., Wei, B., Ma, J., Li, J., Huang, Q., et al. (2011). Apoptotic gene expression by human periodontal ligament cells following cyclic stretch. *J. Periodontal Res.* 46 (6), 742–748. doi:10.1111/j.1600-0765.2011.01397.x
- Zeng, D., Ye, Z., Shen, R., Yu, G., Wu, J., Xiong, Y., et al. (2021). Iobr: Multi-Omics immuno-oncology biological research to decode tumor microenvironment and signatures. *Front. Immunol.* 12, 687975. doi:10.3389/fimmu.2021.687975
- Zhang, J., Peng, G., Chi, H., Yang, J., Xie, X., Song, G., et al. (2023). CD8 + T-cell marker genes reveal different immune subtypes of oral lichen planus by integrating single-cell RNA-seq and bulk RNA-sequencing. *BMC Oral Health* 23 (1), 464. doi:10.1186/s12903-023-03138-0
- Zhao, S., Chi, H., Ji, W., He, Q., Lai, G., Peng, G., et al. (2022a). A bioinformatics-based analysis of an anoikis-related gene signature predicts the prognosis of patients with low-grade gliomas. *Brain Sci.* 12 (10), 1349. doi:10.3390/brainsci12101349
- Zhao, Y., Wei, K., Chi, H., Xia, Z., and Li, X. (2022b). IL-7: A promising adjuvant ensuring effective T cell responses and memory in combination with cancer vaccines? *Front. Immunol.* 13, 1022808. doi:10.3389/fimmu.2022.1022808
- Zhou, J., Liu, M., Zhai, Y., and Xie, W. (2008). The antiapoptotic role of pregnane X receptor in human colon cancer cells. *Mol. Endocrinol.* 22 (4), 868–880. doi:10.1210/me.2007-0197
- Zhou, Y., Zhou, B., Pache, L., Chang, M., Khodabakhshi, A. H., Tanaseichuk, O., et al. (2019). Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat. Commun.* 10 (1), 1523. doi:10.1038/s41467-019-09234-6
- Zondervan, K. T., Becker, C. M., Koga, K., Missmer, S. A., Taylor, R. N., and Viganò, P. (2018). Endometriosis. *Nat. Rev. Dis. Prim.* 4 (1), 9. doi:10.1038/s41572-018-0008-5