RESEARCH

Open Access



Immunomodulatory insights of monoterpene glycosides in endometriosis: immune infiltration and target pathways analysis

Jin Zhan^{1*} and Jiajie Wu²

Abstract

Endometriosis is a complex gynecological condition characterized by abnormal immune responses. This study aims to explore the immunomodulatory effects of monoterpene glycosides from Paeonia lactiflora on endometriosis. Using the ssGSEA algorithm, we assessed immune cell infiltration levels between normal and endometriosis groups. Key targets were identified through differential expression analysis of the GSE51981 dataset. Potential immunomodulatory targets of Paeonia lactiflora compounds were identified through Venn diagram analysis, followed by enrichment and machine learning analyses. A nomogram was developed for predicting endometriosis, while molecular docking explored compound-target interactions. Significant differences in immune cell infiltration were observed, with increased CD8 T cells, cytotoxic cells, and others in endometriosis. Differential expression analysis identified 43 potential targets. Enrichment analysis highlighted pathways involved in immune and inflammatory responses. Machine learning identified SSTR5, CASP3, FABP2, and SYK as critical targets, contributing to a nomogram that demonstrated good predictive performance for endometriosis risk. Molecular docking revealed strong interactions between Paeoniflorigenone and CASP3. Our findings suggest that monoterpene glycosides have therapeutic effects on endometriosis by modulating key immune-related targets and pathways, providing a basis for further investigation into Paeonia lactiflora's potential as a treatment for this condition.

Keywords Endometriosis, Paeonia lactiflora, Monoterpene glycosides, Immune cell infiltration, Immunomodulatory targets, Machine learning, CASP3 interaction

*Correspondence: Jin Zhan

2024n249@zcmu.edu.cn

¹Department of Gynaecology, Ningbo Municipal Hospital of Traditional Chinese Medicine, Affiliated Hospital of Zhejiang Chinese Medical University, No.819, Liyuan North Road, Haishu District, Ningbo, Zhejiang Province 315010, China

²Emergency Department, Ningbo Municipal Hospital of Traditional Chinese Medicine, Affiliated Hospital of Zhejiang Chinese Medical University, Ningbo, Zhejiang Province, China

Introduction

Endometriosis is a prevalent gynecological disorder affecting approximately 10% of women of reproductive age, characterized by the presence of endometrial-like tissue outside the uterus, causing chronic pelvic pain and infertility [4, 5]. Despite extensive research, the precise etiology and pathogenesis of endometriosis remain elusive, largely due to its multifactorial nature involving genetic, hormonal, and immunological factors [17, 19]. Notably, the immune system plays a pivotal role in the development and progression of endometriosis, as evidenced by altered immune cell populations and cytokine profiles in affected individuals [6, 24, 44]. Understanding



© The Author(s) 2024. 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 Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

the immunological aspects of endometriosis is crucial for developing effective therapeutic strategies. Previous studies have highlighted the significance of immune cell infiltration in endometriosis, with increased levels of CD8 T cells and NK cells observed in ectopic lesions [15, 18, 31]. However, the exact mechanisms through which these immune cells contribute to disease pathology are not fully understood. Furthermore, current treatments for endometriosis, including hormonal therapies and surgical interventions, often provide only temporary relief and are associated with adverse side effects. This necessitates the exploration of novel therapeutic agents that can modulate the immune response and offer long-term benefits.

Paeonia lactiflora, a traditional Chinese medicinal herb, has garnered attention for its potential immunomodulatory properties [41]. Monoterpene glycosides, the active compounds derived from Paeonia lactiflora, have demonstrated anti-inflammatory and immunoregulatory effects in various disease models [26, 34, 43]. However, their specific role in endometriosis and the underlying molecular mechanisms remain largely unexplored. While previous research has established the immunomodulatory potential of these compounds, there remains a gap in understanding how they may specifically address the immune dysregulation observed in endometriosis. Recent studies from have begun to explore the intersection of machine learning approaches and endometriosis research, which could provide novel insights into this area [1, 12, 36]. Given the critical involvement of the immune system in endometriosis and the promising immunomodulatory effects of monoterpene glycosides from Paeonia lactiflora, this study aims to investigate the potential therapeutic effects of these compounds on endometriosis. By integrating bioinformatics and machine learning approaches, we aim to provide more precise insights into the immune-related pathways and key targets modulated by monoterpene glycosides, thereby advancing the current understanding of their mechanisms in the context of endometriosis.

The primary objective of this study is to explore the immunomodulatory effects of monoterpene glycosides from Paeonia lactiflora on endometriosis. This study employs a comprehensive approach combining bioinformatics and computational analyses. We utilize the ssG-SEA algorithm to evaluate immune cell infiltration levels and differential expression analysis to identify key targets from the GSE51981 dataset. Potential immunomodulatory targets are identified through Venn diagram analysis, followed by enrichment and machine learning analyses. A nomogram is constructed to predict endometriosis risk, and molecular docking is performed to explore interactions between active compounds and key targets. Figure S1 presents the flow chart of the study. This research provides novel insights into the immunomodulatory mechanisms of monoterpene glycosides from Paeonia lactiflora in endometriosis. By identifying key immunerelated targets and pathways, we offer a potential therapeutic strategy for managing endometriosis, addressing the limitations of current treatments, and paving the way for future clinical applications.

Results

Immune cell infiltration

The immune cell infiltration levels between the control (Con) and endometriosis (EM) groups were assessed using the ssGSEA algorithm (Fig. 1A). Our analysis revealed significant differences in several immune cell types. Notably, CD8 T cells, cytotoxic cells, mast cells, neutrophils, NK CD56bright cells, NK cells, pDC, and Th17 cells showed increased infiltration in the EM group compared to the normal group (*p<0.05, **p<0.01, ***p < 0.001). Conversely, a significant decrease was observed in macrophages, T helper cells and Th2 cells in the EM group compared to the normal group (*p<0.05, **p < 0.01, ***p < 0.001). These findings suggest differential immune cell infiltration profiles between the normal and endometriosis groups. The ssGSEA algorithm was utilized to evaluate the immune-related pathway scores between the normal and endometriosis groups (Fig. 1B). Significant alterations were observed in several immune pathways. The antimicrobial pathway, chemokine receptors, cytokine receptors, cytokines, interferons, interleukins, and TGFb family member exhibited significantly higher scores in the EM group (*p < 0.05, **p < 0.01, ***p<0.001). Conversely, TGFb family member receptor pathway showed significant decrease in score in the EM group (***p < 0.001). These results indicate that the immune-related pathways are more actively present in endometriosis patients, potentially contributing to the disease pathology.

Identification the potential immunomodulatory targets of monoterpene glycosides from Paeonia lactiflora in treating endometriosis.

Differential expression analysis between the Con and EM groups was conducted using the endometriosis-related dataset (GSE51981) (Fig. 2A). This analysis identified 9872 upregulated and 3210 downregulated differentially expressed genes (DEGs) in the EM group compared to the normal group. Four active compounds from Paeonia lactiflora (Paeoniflorigenone, Paeoniflorin, Benzoyl paeoniflorin, and Albiflorin) were selected for the study. Using three online databases (Super-PRED, PharmMapper, and SwissTargetPrediction), we identified 478 potential targets. Additionally, we retrieved 2437 immune-related genes (IRGs) from the ImmPort database. A Venny diagram analysis using the Venny tool intersected the DEGs, compound targets, and IRGs,





Fig. 1 Comparison of immune cell infiltration levels and immune-related pathway scores. Immune cell infiltration levels and immune-related pathway scores in the Con and EM groups assessed using the ssGSEA algorithm. Statistical significance is indicated as follows: *p < 0.05, **p < 0.01, ***p < 0.001



Fig. 2 Exploration of potential immunomodulatory targets of monoterpene glycosides for the treatment of endometriosis. (A) Differential expression analysis based on the GSE51981 comparing Con and EM groups. (B) Venny diagram illustrating the intersection of DEGs, active ingredient targets from four monoterpene glycosides, and IRGs. A total of 43 common targets were identified. (C) PPI network analysis of the 43 common targets. (D) Heatmap displaying the expression profiles of the 43 common targets based on the GSE51981 dataset

resulting in the identification of 43 common targets. These 43 targets represent the potential immunomodulatory targets of the active compounds for the treatment of endometriosis (Fig. 2B). The PPI network displayed extensive interactions among the targets, suggesting key proteins that could be modulated by the active compounds in the context of endometriosis (Fig. 2C). The expression profiles of the 43 potential targets were visualized in a heatmap using the GSE51981 (Fig. 2D). The heatmap exhibits differential expression patterns of these targets between the normal and EM groups, further supporting their potential roles as therapeutic targets. These findings collectively propose that the monoterpene glycosides from Paeonia lactiflora might exert therapeutic

effects against endometriosis through modulating these identified targets.

Enrichment analysis of potential targets

The 43 potential targets identified for the monoterpene glycosides from Paeonia lactiflora against endometriosis were subjected to enrichment analysis. Figure 3A reveals the top enriched pathways and biological processes. The analysis indicates significant involvement in immune and inflammatory responses, including negative regulation of defense response, negative regulation of inflammatory response, and myeloid leukocyte activation. Other critical pathways include relaxin signaling pathway, erbB signaling pathway, EGFR tyrosine kinase



Fig. 3 Enrichment analysis of the potential targets of monoterpene glycosides against endometriosis. (A) The clustering tree diagram displays significantly enriched pathways and biological processes, with the dot color representing the adjusted p-value and the dot size indicating the number of genes involved. (B) Visualization of the enrichment analysis results related to immune and inflammatory pathways. The network diagram shows the connection between the identified targets and enriched pathways. Each line represents the association between a target gene and an enriched pathway, with the color intensity and size of nodes corresponding to the adjusted p-value and the number of genes involved, respectively

inhibitor resistance, hepatitis C, etc. Figure 3B visualizes the enriched pathways particularly related to immune and inflammatory processes. The network diagram highlights the connection between the identified targets and relevant pathways such as myeloid leukocyte activation, response to lipopolysaccharides, cellular response to biotic stimulus, and the B cell receptor signaling pathway. Pathways like the negative regulation of inflammatory response and defense response further emphasize the significant role of these targets in modulating inflammatory processes in endometriosis. These comprehensive enrichment analyses demonstrate that the monoterpene glycosides from Paeonia lactiflora target key pathways involved in immune regulation and inflammation, providing potential therapeutic avenues for managing endometriosis through these mechanisms.

Identification of key anti-endometriosis targets using machine learning

A feature importance analysis was conducted using the Random Forest (RF) algorithm, ranking the importance of the 43 targets based on the Mean Decrease Gini index (Fig. 4A). SSTR5, KCNH2, CASP3, PTGER1, PPARD, and TYMP emerged as the top-ranked targets, indicating their potential significance in the treatment of endometriosis. The LASSO regression analysis identified optimal targets by minimizing the binomial deviance (Fig. 4B). The analysis selected 7 key targets with a lambda value optimized at the point of minimum deviance. Feature selection using SVM-REF revealed the number of features that maximized the 10-fold cross-validation (CV) accuracy (Fig. 4C). The model achieved the highest accuracy with 39 features, underscoring their relevance



Fig. 4 Identification of critical targets for treating endometriosis through the utilization of machine learning techniques. (A) Feature importance ranking of the 43 potential targets using the RF algorithm. (B) LASSO regression analysis results for the 43 potential targets. (C) SVM-REF analysis identifying the number of features that maximize the 10-fold cross-validation accuracy. (D) UpSet diagram integrating the results from RF, LASSO, and SVM-REF analyses

in distinguishing between normal and endometriosis conditions. Integrating the results from RF, LASSO, and SVM-REF, a UpSet diagram was constructed to identify the overlapping key targets (Fig. 4D). Four common key targets (SSTR5, CASP3, FABP2, and SYK) were identified across all three machine learning methods, suggesting their crucial role as therapeutic targets for the anti-endometriosis effects of monoterpene glycosides from Paeonia lactiflora.

Construction and evaluation of a nomogram for predicting endometriosis risk

The expression profiles of the four key anti-endometriosis targets (SSTR5, CASP3, FABP2, and SYK) were analyzed. Figure 5A illustrates significant upregulation of SSTR5 and FABP2, and downregulation of CASP3 and SYK in EM patients compared to the Con group. Based on the gene expression levels of the four identified targets, a nomogram was constructed to predict the risk of EM (Fig. 5B and Table S1). The nomogram assigns scores to each gene expression level, and the total score corresponds to a predicted risk of EM. The ROC curve (Fig. 5C) demonstrates the model's discriminative ability, with an Area Under the Curve (AUC) of 0.808, indicating good predictive performance. The calibration plot shows that the nomogram predictions closely match the observed probabilities, indicating good calibration (Fig. 5D). The nomogram yields a higher net benefit across a range of risk thresholds compared to treating all or none, emphasizing its potential clinical value (Fig. 5E). A higher net benefit indicates that the model reduces the number of false positives and false negatives, thus optimizing patient management by improving the accuracy of EM risk predictions. Clinically, this means that the nomogram can help identify high-risk patients more



Fig. 5 Construction and evaluation of a nomogram. (**A**) Gene expression profiles of the four key targets (SSTR5, CASP3, FABP2, and SYK) in Con and EM groups. Statistical significance is indicated as follows: *p < 0.05, ***p < 0.001. (**B**) Nomogram constructed based on the expression levels of the four key targets to predict the risk of endometriosis. (**C**) ROC curve evaluating the discriminative ability of the nomogram. (**D**) Calibration curve assessing the agreement between predicted probabilities and actual outcomes. (**E**) Decision curve analysis illustrating the net benefit of using the nomogram across a range of risk thresholds

accurately, enabling more targeted interventions and reducing unnecessary treatments for those at low risk.

Correlation analysis of key anti-endometriosis targets with immune cell infiltration

The correlation between gene expression of the four key anti-endometriosis targets (SSTR5, CASP3, FABP2, and SYK) and immune cell infiltration was analyzed. Figure 6A shows the correlation heatmap for various immune cell types. Notably, CASP3 expression exhibited significant negative correlations with several immune cell types including CD8 T cells, cytotoxic cells, DC, iDC, NK CD56bright cells, NKCD56dim cells, NK cells, and Th17 cells (p < 0.001). In contrast, SSTR5 and FABP2 expression showed significant positive correlations with several immune cell types, such as D8 T cells, cytotoxic cells, DC, NK CD56bright cells, NKCD56dim cells, NK cells, and Th17 cells (p < 0.001). Figure 6B illustrates the correlation heatmap between the gene expression of the four key targets and various immune-related pathways. CASP3 showed strong negative correlations with chemokine receptors, chemokines, cytokine receptors, cytokines, TGFb family member, and TNF family members receptors pathways (p < 0.001). SSTR5 and FABP2 exhibited significant positive correlations with several immune pathways, including the antimicrobials, chemokine receptors, cytokine receptors, cytokines, interferons, interleukins, and TGFb family member pathways (p < 0.001). These correlation analyses underscore the significant involvement of the key anti-endometriosis

A

Page 8 of 13

targets (SSTR5, CASP3, FABP2, and SYK) in modulating immune cell infiltration and immune-related pathways. This highlights the potential immunomodulatory mechanisms by which the monoterpene glycosides from Paeonia lactiflora exert their therapeutic effects against EM.

Molecular docking analysis of key targets with active compounds

Figure 7A displays the docking scores represented as Vina scores. Among the four targets, CASP3 showed the highest binding affinity with Paeoniflorigenone, with a Vina score of -11.1, indicating a strong interaction. The other compounds also exhibited notable binding affinities with the targets; Paeoniflorin displayed significant binding with FABP2 (Vina score: -10.3) and SYK (Vina score: -9.3). The varying Vina scores across different compounds and targets highlight the differential binding propensities, suggesting specific interactions relevant to the therapeutic effects of the compounds. Figure 7B provides a detailed visualization of the molecular docking between CASP3 and Paeoniflorigenone. The docking illustration shows the binding site and interactions, including hydrogen bonds and hydrophobic interactions. Key residues involved in the interaction include R126, M21, Y70, etc. These interactions suggest potential mechanisms thorugh which Paeoniflorigenone could modulate CASP3 activity, thereby exerting therapeutic effects against endometriosis. These docking results support the hypothesis that monoterpene glycosides from Paeonia lactiflora can effectively interact with key targets involved



В

Fig. 6 Correlation analysis of key anti-endometriosis targets with immune cell infiltration and immune pathways. (A) Correlation heatmap of the four key anti-endometriosis targets with immune cell infiltration. (B) Correlation heatmap of the four key targets with immune-related pathways



Fig. 7 Molecular docking analysis of key anti-endometriosis targets with monoterpene Glycosides. (A) Heatmap of Vina docking scores showing the binding affinities between the four key anti-endometriosis targets and the four monoterpene glycosides from Paeonia lactiflora. (B) Molecular docking visualization of CASP3 with Paeoniflorigenone

in endometriosis, potentially modulating their activity and contributing to the therapeutic effects observed.

Discussion

Paeonia lactiflora, commonly known as white peony, has long been used in traditional Chinese medicine for its wide range of pharmacological effects, including anti-inflammatory, analgesic, and immunomodulatory properties [26, 40]. Recent studies have suggested that its active constituents, including monoterpene glycosides, may hold therapeutic potential for gynecological disorders such as endometriosis [25, 42]. Endometriosis is characterized by ectopic endometrial tissue growth, often accompanied by dysregulated immune responses and chronic inflammation [16, 27, 33]. Our study provides novel insights into the immunomodulatory effects of monoterpene glycosides from Paeonia lactiflora on endometriosis, highlighting their potential as a therapeutic agent.

In this study, we observed significant alterations in immune cell infiltration between normal and endometriosis tissues, with an increase in CD8 T cells and cytotoxic cells in the latter. These findings are consistent with previous reports that suggest a role for immune dysregulation in the pathogenesis of endometriosis [9, 10, 30]. Our differential expression analysis identified 43 potential targets, and enrichment analysis revealed pathways predominantly involved in immune and inflammatory responses. Notably, machine learning approaches identified key targets such as SSTR5, CASP3, FABP2, and SYK, which are implicated in immune modulation and apoptosis. Further exploration of the clinical implications of these targets, raises questions about their feasibility for translation into clinical applications. For instance, exploring the development of targeted therapies or biomarkers based on these targets could provide valuable insights into future treatment strategies for endometriosis. The identification of CASP3 as a key target is particularly noteworthy, as caspase-3 is a critical executor of apoptosis, and its dysregulation has been linked to the survival of ectopic endometrial cells [23]. Our molecular docking analysis revealed strong interactions between Paeoniflorigenone, a monoterpene glycoside, and CASP3, suggesting that Paeonia lactiflora may induce apoptosis in ectopic endometrial cells by modulating caspase-3 activity. This aligns with prior research indicating that Paeonia lactiflora exerts anti-apoptotic effects in various cell types [20, 21, 38, 39]. However, it is essential to acknowledge potential side effects and limitations of herbal therapies, as they may vary among individuals. A balanced perspective on these aspects will enhance the understanding of Paeonia lactiflora's therapeutic potential. The construction of a nomogram based on the identified targets provided a reliable tool for predicting endometriosis risk, with an AUC of 0.808, indicating good predictive performance. This model's predictive capacity adds to the growing body of evidence supporting the role of immune-related targets in the etiology and progression of endometriosis [45].

Our findings contribute to the current understanding of endometriosis by elucidating the immunomodulatory mechanisms of monoterpene glycosides from Paeonia lactiflora. The integration of immune infiltration analysis with machine learning and molecular docking represents a novel approach in endometriosis research, offering a comprehensive understanding of how these compounds may exert therapeutic effects. Nonetheless, it is important to consider potential limitations of our computational findings, such as biases inherent in the GSE51981 dataset and concerns regarding the predictive accuracy of machine learning models. Addressing these limitations will enhance the robustness of our conclusions. While our study provides compelling evidence for the potential of Paeonia lactiflora in endometriosis therapy, further experimental validation and clinical studies are warranted to fully elucidate its therapeutic efficacy and underlying mechanisms.

Conclusions

In conclusion, our study identifies monoterpene glycosides from Paeonia lactiflora as promising candidates for the treatment of endometriosis, potentially through the modulation of immune-related targets and pathways. To further validate these findings, we recommend conducting experimental studies or clinical trials to test the therapeutic efficacy of Paeonia lactiflora in endometriosis. These findings not only enhance our understanding of the immunopathogenesis of endometriosis but also pave the way for the development of novel therapeutic strategies based on traditional medicine.

Methods

Active ingredient target screening

To identify potential active ingredients of Paeonia lactiflora for the treatment of endometriosis, we employed a systematic screening approach using the Traditional Chinese Medicine Systems Pharmacology Database and Analysis Platform (TCMSP) [28]. The screening criteria were set to an oral bioavailability (OB) \geq 30% and druglikeness (DL) \geq 0.1. Based on these parameters and extensive literature review, four compounds from Paeonia lactiflora were selected: Paeoniflorigenone, Paeoniflorin, Benzoyl Paeoniflorin, and Albiflorin. The canonical SMILES of these compounds were retrieved from the PubChem database. Using these SMILES, we predicted the potential targets of the compounds through three distinct databases: Super-PRED, PharmMapper, and SwissTargetPrediction [7, 11, 35]. After obtaining the predicted targets from each database, duplicate targets were removed to consolidate the results.

Disease target screening

We conducted a search for datasets related to endometriosis in the GEO database, utilizing "endometriosis" as the search term. The inclusion criteria for the datasets were as follows: (1) the organism must be Homo sapiens; (2) the sample size for both the disease and control cohorts should be no less than 40. Following our screening process, the GSE51981 dataset, which comprises 71 Con samples and 78 EM samples, fulfilled the criteria. To identify differentially expressed genes (DEGs) associated with endometriosis, we utilized the publicly available dataset GSE51981 from the Gene Expression Omnibus (GEO) database. The raw matrix files of the GSE51981 dataset were extracted and preprocessed using the affy package (v.3.8). This step involved background correction, normalization, and summarization of probe-level data to ensure accurate and reliable expression measurements. This dataset was annotated using the GPL570 platform annotation file to convert probe IDs to gene symbols. For cases where multiple probes corresponded to a single gene symbol, the average expression value of the probes was used to represent the gene's expression level. The limma package (v.3.22.7) was employed to perform differential expression analysis between the normal and disease groups within the dataset. DEGs were identified based on an adjusted p-value threshold of < 0.05. The ggplot2 package (v.3.3.6) was used to visualize the results of the differential expression analysis.

Monoterpene glycosides anti-endometriosis target screening

Immune-related genes (IRGs) were retrieved from the ImmPort database (http://www.immport.org). This dataset provided a comprehensive list of genes involved in immune responses [2]. DEGs identified from the GSE51981 dataset, IRGs from the ImmPort database, and predicted targets of the Monoterpene glycosides (Paeoniflorigenone, Paeoniflorin, Benzoyl Paeoniflorin, Albiflorin) were intersected using the Venny tool. The intersected genes were considered as the Monoterpene glycosides anti-endometriosis targets.

Construction of protein-protein Interaction (PPI) network

The common targets obtained from the intersection analysis were submitted to the STRING database (https://stri ng-db.org) [32]. The interaction data for the common targets were extracted from STRING with a high confidence score (interaction score>0.7) to ensure the reliability of the interactions included in the network. The extracted PPI network data were imported into Cytoscape software (version 3.7.2) for visualization and further analysis.

Enrichment analysis of monoterpene glycosides antiendometriosis targets

To explore the biological functions and pathways associated with the common targets of monoterpene glycosides in the treatment of endometriosis, we conducted Gene Ontology (GO) functional and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses. The common targets identified were subjected to enrichment analysis using ClusterProfiler package (version 2.2.7). For both GO and KEGG enrichment analyses, a significance threshold of p < 0.05 was applied to ensure the identification of statistically significant biological functions and pathways.

Machine learning identification of core monoterpene glycosides anti-endometriosis targets

To identify the core targets of monoterpene glycosides for the treatment of endometriosis, we utilized three distinct machine learning algorithms based on the gene expression profiles of the common targets from the GSE51981 dataset. The selection of these specific machine learning models was based on their proven effectiveness in handling high-dimensional biological data and their ability to provide interpretable results. The Random Forest algorithm was employed using the randomForest package in R. The importance of each gene was ranked based on the mean decrease in accuracy, which allows for robust feature selection and helps in understanding the contribution of each gene to the model's predictive power [14]. LASSO regression was conducted using the glmnet package in R. This method applies a penalty to the absolute size of the regression coefficients, effectively shrinking some coefficients to zero and thereby performing variable selection and regularization. LASSO is particularly useful in scenarios with many predictors, as it helps to prevent overfitting and enhances model interpretability [22]. The optimal lambda value was determined through crossvalidation to identify significant predictors among the common targets. SVM-RFE was implemented using the e1071 package in R. This method ranks features by recursively eliminating the least significant ones and re-fitting the model until the optimal set of features is achieved. SVM-RFE is advantageous for its ability to handle nonlinear relationships and its effectiveness in identifying the most relevant features in complex datasets [29]. The results from the RF, LASSO, and SVM-RFE analyses were integrated using the Venny tool. Common targets identified by all three methods were considered core Monoterpene glycosides anti-endometriosis targets [8, 37].

Nomogram construction and evaluation

Using the rms package in R, we integrated the core target genes to construct a nomogram. The predictive ability of the nomogram for endometriosis was assessed through Receiver Operating Characteristic (ROC) curve analysis. To assess the accuracy of the nomogram, calibration curve analysis was performed. Decision Curve Analysis was conducted to evaluate the clinical utility and net benefit of the nomogram across different threshold probabilities.

Immune characterization analysis

Gene sets used for evaluating the enrichment levels of various immune cell subsets were obtained from a previous study [3]. Additionally, gene sets related to immune response pathways were downloaded from the ImmPort database (http://www.immport.org) [2]. The enrichment levels of immune cell subsets and immune response pathways in endometriosis and normal samples were assessed using single-sample gene set enrichment analysis (ssG-SEA) [13]. The results of the ssGSEA were visualized using the ggplot2 package. To explore the associations between core target genes and immune cell components as well as immune response pathways, pearson correlation analysis was performed. The correlation coefficients were calculated to determine the strength of the relationships. The results of the correlation analysis were also visualized using the ggplot2 package.

Molecular docking analysis

To investigate the interaction between small molecule compounds and core target proteins associated with endometriosis, we conducted a molecular docking analysis. Small molecule compounds were obtained from the PubChem database in the form of Structure Data Files (SDF). These files were processed and converted into the PDBQT file for docking analysis using AutoDock Tools (version 1.5.6). The three-dimensional structures of the core target proteins were downloaded from the Protein Data Bank (PDB) (https://www.rcsb.org). These structur es were prepared for docking by removing water molecules, adding hydrogen atoms, and assigning appropriate charges using AutoDock Tools. Molecular docking was performed using AutoDock to predict the binding affinity and interaction modes between the small molecule compounds and core target proteins. The docking results were analyzed to identify the best binding poses based on the binding energy scores. The interactions between the ligands and the target proteins were visualized using PyMOL software (version 1.0.0).

Abbreviations

TCMSP	Traditional Chinese Medicine Systems Pharmacology Database and
	Analysis Platform
OB	Oral bioavailability
DL	Drug-likeness
DEGs	Differentially expressed genes
GEO	Gene Expression Omnibus
IRGs	Immune-related genes
PPI	Protein-Protein Interaction

GO	Gene Ontology
KEGG	Kyoto Encyclopedia of Genes and Genomes
ROC	Receiver Operating Characteristic
ssGSEA	Single-sample gene set enrichment analysis
SDF	Structure Data Files
PDB	Protein Data Bank
Con	Control
FМ	Endometriosis

Supplementary Information

The online version contains supplementary material available at https://doi.or g/10.1186/s41065-024-00354-8.

Supplementary Material 1

Acknowledgements

Not applicable.

Author contributions

Jin Zhan wrote the manuscript and designed the study. Jiajie Wu analyzed the data and produced the figures. All authors reviewed and edited the manuscript. All authors reviewed the manuscript.

Funding

This work was supported by the Science and Technology Project of Traditional Chinese Medicine in Zhejiang Province (2024ZR158).

Data availability

The data used in our study are available from the GEO (https://www.ncbi.nlm. nih.gov/geo/) database.

Declarations

Ethics approval and consent to participate Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 4 November 2024 / Accepted: 8 December 2024 Published online: 03 January 2025

References

- Balogh DB, Hudelist G, Blizņuks D, Raghothama J, Becker CM, Horace R, Krentel H, Horne AW, Bourdel N, Marki G, et al. FEMaLe: the use of machine learning for early diagnosis of endometriosis based on patient self-reported data-study protocol of a multicenter trial. PLoS ONE. 2024;19(5):e0300186. eng.
- 2 Bhattacharya S, Dunn P, Thomas CG, Smith B, Schaefer H, Chen J, Hu Z, Zalocusky KA, Shankar RD, Shen-Orr SS et al. 2018. ImmPort, toward repurposing of open access immunological assay data for translational and clinical research. Scientific data. 5:180015. eng.
- 3 Bindea G, Mlecnik B, Tosolini M, Kirilovsky A, Waldner M, Obenauf AC, Angell H, Fredriksen T, Lafontaine L, Berger A, et al. Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. Immunity. 2013;39(4):782–95. eng.
- 4 Bulun SE, Yilmaz BD, Sison C, Miyazaki K, Bernardi L, Liu S, Kohlmeier A, Yin P, Milad M, Wei J. Endometr Endocr Reviews. 2019;40(4):1048–79. eng.
- 5 Burney RO, Giudice LC. Pathogenesis and pathophysiology of endometriosis. Fertil Steril. 2012;98(3):511–9. eng.
- 6 Chen S, Liu Y, Zhong Z, Wei C, Liu Y, Zhu X. 2023. Peritoneal immune microenvironment of endometriosis: Role and therapeutic perspectives. Frontiers in immunology. 14:1134663. eng.

- 7 Daina A, Michielin O, Zoete V. 2019. SwissTargetPrediction: updated data and new features for efficient prediction of protein targets of small molecules. Nucleic acids research. 47(W1):W357-w364. eng.
- 8 Dong G, Gao H, Chen Y, Yang H. 2023. Machine learning and bioinformatics analysis to identify autophagy-related biomarkers in peripheral blood for rheumatoid arthritis. Frontiers in genetics. 14:1238407. eng.
- 9 Dymanowska-Dyjak I, Terpiłowska B, Morawska-Michalska I, Michalski A, Polak G, Terpiłowski M, Rahnama-Hezavah M, Grywalska E. 2024. Immune Dysregulation in Endometriomas: Implications for Inflammation. International journal of molecular sciences. 25(9). eng.
- 10 Fan D, Wang X, Shi Z, Jiang Y, Zheng B, Xu L, Zhou S. Understanding endometriosis from an immunomicroenvironmental perspective. Chin Med J. 2023;136(16):1897–909. eng.
- 11 Gallo K, Goede A, Preissner R, Gohlke BO. 2022. SuperPred 3.0: drug classification and target prediction-a machine learning approach. Nucleic acids research. 50(W1):W726-w731. eng.
- 12 Goldstein A, Cohen S. Self-report symptom-based endometriosis prediction using machine learning. Sci Rep. 2023;13(1):5499. eng.
- 13 Hänzelmann S, Castelo R, Guinney J. 2013. GSVA: gene set variation analysis for microarray and RNA-seq data. BMC bioinformatics. 14:7. eng.
- 14 Hu J, Szymczak S. A review on longitudinal data analysis with random forest. Brief Bioinform. 2023;24(2):eng.
- 15 Huang ZX, Lin DC, Zhang HY, Yang MJ, Chen JH, Ding XY, Dai SJ, Hong YH, Liang GS, Li QY, et al. The dysfunction of CD8(+) T cells triggered by endometriotic stromal cells promotes the immune survival of endometriosis. Immunology. 2024;172(3):469–85. eng.
- 16 Jeljeli M, Riccio LGC, Chouzenoux S, Moresi F, Toullec L, Doridot L, Nicco C, Bourdon M, Marcellin L, Santulli P et al. 2020. Macrophage Immune Memory Controls Endometriosis in Mice and Humans. Cell reports. 33(5):108325. eng.
- 17 Kapoor R, Stratopoulou CA, Dolmans MM. 2021. Pathogenesis of Endometriosis: New Insights into Prospective Therapies. International journal of molecular sciences. 22(21). eng.
- 18 Kisovar A, Becker CM, Granne I, Southcombe JH. 2023. The role of CD8+T cells in endometriosis: a systematic review. Frontiers in immunology. 14:1225639. eng.
- 19 Koninckx PR, Ussia A, Adamyan L, Wattiez A, Gomel V, Martin DC. Pathogenesis of endometriosis: the genetic/epigenetic theory. Fertil Steril. 2019;111(2):327–40. enq.
- 20 Li LY, Ma L, Dong WL. Total glucosides of paeony (Paeonia lactiflora) alleviates blood-brain barrier disruption and cerebral ischemia/reperfusion injury in rats via suppressing inflammation and apoptosis. Pharmazie. 2020;75(5):208– 12. eng.
- 21 Li Y, Deng X, Hu Q, Chen Y, Zhang W, Qin X, Wei F, Lu X, Ma X, Zeng J, et al. Paeonia lactiflora pall. Ameliorates acetaminophen-induced oxidative stress and apoptosis via inhibiting the PKC-ERK pathway. J Ethnopharmacol. 2024;329:118107. eng.
- 22 Liang C, Wanling L, Maofeng W. 2024. LASSO-derived model for the prediction of bleeding in aspirin users. Scientific reports. 14(1):12507. eng.
- 23 Mao Y, Qiao JD, Chen S, Zhou X, Wang Z, Cai S, Li L, Luo Y. 2021. Kallistatin inhibits Anoikis Resistance and Metastasis of Ectopic Endometrium cells by modulating MnSOD and Caspase 3 signaling. Reprod Sci. 2021;28(4):1012– 1019. eng.
- 24 Omidvar-Mehrabadi A, Ebrahimi F, Shahbazi M, Mohammadnia-Afrouzi M. 2024. Cytokine and chemokine profiles in women with endometriosis, polycystic ovary syndrome, and unexplained infertility. Cytokine. 178:156588. eng.
- 25 Park MJ, Han SE, Kim HJ, Heo JD, Choi HJ, Ha KT, Yang SW, Lee KS, Kim SC, Kim CW et al. 2020. Paeonia lactiflora improves ovarian function and oocyte quality in aged female mice. Animal reproduction. 17(2):e20200013. eng.
- 26 Parker S, May B, Zhang C, Zhang AL, Lu C, Xue CC. A pharmacological review of Bioactive constituents of Paeonia lactiflora Pallas and Paeonia Veitchii Lynch. Phytother Res. 2016;30(9):1445–73. eng.
- 27 Peng L, Luo X, Cao B, Wang X. 2024. Exploring the link: Systemic immuneinflammation index as a marker in endometriosis-Insights from the NHANES 2001–2006 cross-sectional study. PloS one. 19(6):e0304591. eng.

- 28 Ru J, Li P, Wang J, Zhou W, Li B, Huang C, Li P, Guo Z, Tao W, Yang Y et al. 2014. TCMSP: a database of systems pharmacology for drug discovery from herbal medicines. Journal of cheminformatics. 6:13. eng.
- 29 Sanz H, Valim C, Vegas E, Oller JM, Reverter F. SVM-RFE: selection and visualization of the most relevant features through non-linear kernels. BMC Bioinformatics. 2018;19(1):432. eng.
- 30 Shen HH, Zhang T, Yang HL, Lai ZZ, Zhou WJ, Mei J, Shi JW, Zhu R, Xu FY, Li DJ, et al. Ovarian hormones-autophagy-immunity axis in menstruation and endometriosis. Theranostics. 2021;11(7):3512–26. eng.
- 31 Shin S, Chung YJ, Moon SW, Choi EJ, Kim MR, Chung YJ, Lee SH. Single-cell profiling identifies distinct hormonal, immunologic, and inflammatory signatures of endometriosis-constituting cells. J Pathol. 2023;261(3):323–34. eng.
- 32 Szklarczyk D, Kirsch R, Koutrouli M, Nastou K, Mehryary F, Hachilif R, Gable AL, Fang T, Doncheva NT, Pyysalo S et al. 2023. The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. Nucleic acids research. 51(D1):D638-d646. eng.
- 33 Vallvé-Juanico J, Houshdaran S, Giudice LC. The endometrial immune environment of women with endometriosis. Hum Reprod Update. 2019;25(5):564–91. eng.
- 34 Wang QS, Gao T, Cui YL, Gao LN, Jiang HL. Comparative studies of paeoniflorin and albiflorin from Paeonia lactiflora on anti-inflammatory activities. Pharm Biol. 2014;52(9):1189–95. eng.
- Wang X, Shen Y, Wang S, Li S, Zhang W, Liu X, Lai L, Pei J, Li H. PharmMapper 2017 update: a web server for potential drug target identification with a comprehensive target pharmacophore database. Nucleic Acids Res. 2017;45W1:W356–60. eng.
- Wang W, Zeng W, Yang S. 2024. A stacked machine learning-based classification model for endometriosis and adenomyosis: a retrospective cohort study utilizing peripheral blood and coagulation markers. Frontiers in digital health. 6:1463419. eng.
- 37 Wu Z, Chen H, Ke S, Mo L, Qiu M, Zhu G, Zhu W, Liu L. Identifying potential biomarkers of idiopathic pulmonary fibrosis through machine learning analysis. Sci Rep. 2023;13(1):16559. eng.
- 38 Xing D, Ma Y, Lu M, Liu W, Zhou H. 2023. Paeoniflorin alleviates hypoxia/reoxygenation injury in HK-2 cells by inhibiting apoptosis and repressing oxidative damage via Keap1/Nrf2/HO-1 pathway. BMC nephrology. 24(1):314. eng.
- 39 Yang L, Liu S, Mu S, Guo R, Zhou L, Fu Q. 2021. Paeoniflorin Attenuates Dexamethasone-Induced Apoptosis of Osteoblast Cells and Promotes Bone Formation via Regulating AKT/mTOR/Autophagy Signaling Pathway. Evidence-based complementary and alternative medicine: eCAM. 2021:6623464. eng.
- Zhang W, Dai SM. Mechanisms involved in the therapeutic effects of Paeonia lactiflora Pallas in rheumatoid arthritis. Int Immunopharmacol. 2012;14(1):27– 31. eng.
- 41 Zhang L, Wei W. Anti-inflammatory and immunoregulatory effects of paeoniflorin and total glucosides of paeony. Pharmacol Ther. 2020;207:107452. eng.
- Zhang J, Wang F, Wang H, Wang Y, Wu Y, Xu H, Su C. Paeoniflorin inhibits proliferation of endometrial cancer cells via activating MAPK and NF-kB signaling pathways. Experimental Therapeutic Med. 2017;14(6):5445–51. eng.
- Zhang L, Yu J, Wang C, Wei W. The effects of total glucosides of paeony (TGP) and paeoniflorin (Pae) on inflammatory-immune responses in rheumatoid arthritis (RA). Funct Plant Biology: FPB. 2019;46(2):107–17. eng.
- 44 Zhong S, Liang Y, Wu Z, Wei L. 2023. Association between polymorphisms of cytokine genes and endometriosis: A comprehensive systematic review and meta-analysis. Journal of reproductive immunology. 158:103969. eng.
- 45 Zhou C, Feng M, Chen Y, Lv S, Zhang Y, Chen J, Zhang R, Huang X. 2023. Unraveling immunotherapeutic targets for endometriosis: a transcriptomic and single-cell analysis. Frontiers in immunology. 14:1288263. eng.

Publisher's note

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