

RESEARCH ARTICLE

Integrative analysis from multi-center studies identifies a weighted gene co-expression network analysis-based Tregs signature in ovarian cancer

Yang Cao  | Ying-Lei Liu | Xiao-Yan Lu | Hai-Li Kai | **Yun Han** | Yan-Li Zheng

Department of Obstetrics and Gynecology, Affiliated Hospital 2 of Nantong University, Nantong First People's Hospital, Nantong, China

Correspondence

Yun Han and Yan-Li Zheng, Department of Obstetrics and Gynecology, Affiliated Hospital 2 of Nantong University, Nantong First People's Hospital, Nantong 226000, China. Email: hanyun_888@163.com and gaoshan1189@sina.com

Funding information

Jiangsu Provincial Health Commission Project, Grant/Award Number: Z2021078; Nantong Science and Technology Bureau Plan Project, Grant/Award Number: MS22022005; Scientific Research and Innovation Team Project of Kangda College of Nanjing Medical University, Grant/Award Number: KD2022KYCXTD010

Abstract

Ovarian cancer (OC) is a malignancy associated with poor prognosis and has been linked to regulatory T cells (Tregs) in the immune microenvironment. Nevertheless, the association between Tregs-related genes (TRGs) and OC prognosis remains incompletely understood. The xCell algorithm was used to analyze Tregs scores across multiple cohorts. Weighted gene co-expression network analysis (WGCNA) was utilized to identify potential TRGs and molecular subtypes. Furthermore, we used nine machine learning algorithms to create risk models with prognostic indicators for patients. Reverse transcription-quantitative polymerase chain reaction and immunofluorescence staining were used to demonstrate the immunosuppressive ability of Tregs and the expression of key TRGs in clinical samples. Our study found that higher Tregs scores were significantly correlated with poorer overall survival. Recurrent patients exhibited increased Tregs infiltration and reduced CD8⁺ T cell. Moreover, molecular subtyping using seven key TRGs revealed that subtype B exhibited higher enrichment of multiple oncogenic pathways and had a worse prognosis. Notably, subtype B exhibited high Tregs levels, suggesting immune suppression. In addition, we validated machine learning-derived prognostic models across multiple platform cohorts to better distinguish patient survival and predict immunotherapy efficacy. Finally, the differential expression of key TRGs was validated using clinical samples. Our study provides novel insights into the role of Tregs in the immune microenvironment of OC. We identified potential therapeutic targets derived from Tregs (CD24, FHL2, GPM6A, HOXD8, NAP1L5, REN, and TOX3) for personalized treatment and created a machine learning-based prognostic model for OC patients, which could be useful in clinical practice.

KEYWORDS

immune microenvironment, ovarian cancer, prognosis, Tregs, WGCNA

1 | INTRODUCTION

Ovarian cancer (OC) is a deadly gynecological malignancy with a high mortality rate that remains largely unchanged despite advances in

diagnosis and treatment.¹ The complex immune microenvironment of OC poses significant challenges in the development of effective therapies.^{2,3} Regulatory T cells (Tregs), a subset of CD4⁺ T cells, play a crucial role in suppressing immune response and promoting tumor growth and

metastasis.⁴ Nevertheless, the association between Tregs-related genes (TRGs) and OC prognosis remains incompletely understood.

Tregs are responsible for immune suppression through various mechanisms, such as the expression of immune checkpoint receptors (e.g., CTLA4), thereby interacting with and suppressing effector T-cells and antigen-presenting cells (APCs).⁵ Furthermore, Tregs release immunomodulatory cytokines, including interleukin-10 (IL-10),⁶ IL-35,⁷ and transforming growth factor-beta (TGF- β),⁸ to inhibit the immune response.⁹ Tregs produce cytotoxic molecules that trigger apoptosis in effector immune cells.¹⁰ In addition, Tregs manipulate the metabolism of effector immune cells,¹¹ thus influencing their functions. Importantly, Tregs are linked to unfavorable prognoses in solid cancers, such as prostate cancer,¹² OC,¹³ and gastric cancer,¹⁴ underscoring the need to target Tregs when treating cancer. Altering the proportion of Tregs or their signaling pathways within the tumor microenvironment (TME) and reversing the resultant immunosuppression may improve treatment outcomes with immune checkpoint inhibitors.⁵ However, the biological function of Tregs within the TME is intricate and their targets, particularly co-stimulatory molecules, are predominantly expressed on effector T-cells or natural killer (NK) cells. As a result, targeting these molecules may damage effector cells equally, heightening the need for current drugs targeting Tregs to not only enhance patient outcomes but also maintain tolerability with low autoimmune risks.¹⁵

Several challenges arise in understanding and managing OC. These challenges may include factors such as the complexity and heterogeneity of the disease, the limitations of available diagnostic tools or biomarkers, the influence of various genetic and environmental factors, and potential difficulties in accurately interpreting clinical data. Weighted gene co-expression network analysis (WGCNA) is a robust tool for identifying gene expression patterns and molecular markers across several diseases, including OC,¹⁶ uterine leiomyosarcoma,¹⁷ and renal cell carcinoma.¹⁸ Specifically, WGCNA allows for the identification of biologically relevant gene modules, which are clusters of co-expressed genes, as well as hub genes that play a central role in gene regulation.¹⁹ In addition, WGCNA can identify disease subtypes based on gene expression, leading to more precise diagnoses and personalized treatments.²⁰ Currently, there is a lack of large-sample multicenter studies analyzing Tregs using WGCNA in OC cohorts. Here, we integrate 14 multicenter studies with complete follow-up information for bioinformatics analysis. Our results revealed novel therapeutic targets derived from Tregs (CD24, FHL2, GPM6A, HOXD8, NAP1L5, REN, and TOX3) for personalized treatment and a machine learning-based prognostic model for OC patients that can be implemented in clinical practice.

2 | MATERIALS AND METHODS

2.1 | Clinical samples

Ten pairs of tissue samples, including OC tissues and normal ovarian epithelial tissues, were procured from patients who underwent surgery at the Department of Obstetrics and

Gynecology, Affiliated Hospital 2 of Nantong University, between January 2022 and April 2023. The Ethical Committee of Affiliated Hospital 2 of Nantong University approved the collection of these tissues (2022KT142). Immunofluorescence staining was performed on two formalin-fixed paraffin-embedded (FFPE) tissue blocks (primary and recurrent tumors) obtained from our hospital using the services of Servicebio (Wuhan, Hubei, China). According to the company, detailed methods are available in a previous reference.¹⁶

2.2 | Reverse transcription-quantitative polymerase chain reaction

RNA extraction from tissues was performed using TRIzol (15 596 018, Thermo), according to the manufacturer's instructions. Subsequently, PrimeScriptT-MRT Kit (R232-01, Vazyme) was used for cDNA synthesis. Real-time polymerase chain reaction (RT-PCR) was undertaken utilizing SYBR Green Master Mix (Q111-02, Vazyme). The mRNA expression levels were normalized to GAPDH mRNA level and calculated using the $2^{-\Delta\Delta Ct}$ method. The primer sequences were available in the previous references, including CD24,²¹ FHL2,²² GPM6A,²³ HOXD8,²⁴ NAP1L5,²⁵ REN,²⁶ and TOX3.²⁷

2.3 | Collection of transcriptome data

Samples that lacked complete follow-up information, had zero days of survival, or were repetitive sequencing of the same patient were excluded. RNA sequencing (RNA-seq) data in transcripts per kilobase per million mapped reads (FPKM) format were downloaded from The Cancer Genome Atlas (TCGA) database for 374 samples. The raw data was subjected to $\log_2[(TPM) + 1]$ transformation. Somatic mutation data sourced from TCGA database and multiple databases from Gene Expression Omnibus (GEO) were integrated, utilizing various platforms, namely GPL570 (GSE19829, GSE18520, GSE9891, GSE26193, GSE30161, and GSE63885; $n = 597$), GPL96 (GSE3149, GSE23554, GSE276712, and GSE14764; $n = 409$), GPL7759 (GSE13876, $n = 415$), GPL14951 platform (GSE140082, $n = 380$), and GPL2986 (GSE49997, $n = 204$).¹⁶ Batch effect elimination was achieved using the “sva” package.

Our transcriptome-related bioinformatics analysis included 2379 patients with complete follow-up information gleaned from the integration of 14 multicenter studies.

2.4 | Weighted gene co-expression network analysis

The “IOBR” package's xCell algorithm was utilized to calculate Tregs abundance, which was then used as phenotypic data for subsequent weighted gene co-expression network analysis (WGCNA).²⁰ In

particular, the top 25% variants in the GPL570 and TCGA-OV cohorts were employed as input to select soft thresholds based on our previous protocol.¹⁶ The resulting adjacency matrix was clustered to identify the hub module, from which potential Tregs-related genes (TRGs) were identified. This was achieved by calculating the Pearson correlation coefficients between the modules and Treg scores.

2.5 | Consensus clustering

We performed consensus clustering analysis utilizing key TRGs. Principal component analysis (PCA) was conducted to determine the independence of each subtype. The number of clusters was determined using the “ConensusClusterPlus” package and validated using rank tests with 1000 replicates along with $pltem = 0.8$ parameter settings to evaluate subtype stability.²⁸ Differential gene analysis between subtypes was performed using the “limma” package, with parameters set at adjusted $p < .05$ and $|\log FC| > 1$. The prognostic value of each differential expression genes (DEG) was assessed by means of one-way Cox regression analysis ($p < .05$). Based on the prognostic DEGs obtained, a second round of consensus clustering analysis was performed.

2.6 | Immune-related analysis

To estimate immune cell abundance in different samples, we employed a range of methods including ssGSEA, TIMER, CIBERSORT, QUANTISEQ, MCP-counter, XCELL, and EPIC. Moreover, the ESTIMATE algorithm was utilized to calculate the immune score and stromal score, serving as indicators of the overall microenvironmental status.²⁹ Moreover, we downloaded the IMvigor-210 cohort from existing reference which had undergone immune checkpoint treatment to generate risk scores. This enabled us to validate the efficacy of our risk scores in the immunotherapy cohort.

2.7 | Enrichment analysis

We employed the clusterProfiler package to perform Kyoto Encyclopedia of Genes and Genomes (KEGG) based on all DEGs,³⁰ with significant enrichment pathways determined by p -value $< .05$ and q -value $< .05$. Furthermore, using $FDR < 0.05$ as the threshold value, we utilized Gene Set Variation Analysis (GSVA) to evaluate differences in biological pathways between subtypes.³¹

2.8 | Machine learning-derived signature

To facilitate machine learning modeling, we utilized the TCGA-OV cohort for modeling purposes, with validation being conducted on GSE18520, GSE19829, GSE26193, GSE30161, GSE63885, and GSE9891. Each dataset was normalized separately, with the gene

mean/SD being normalized to 1. Variable screening, which required a minimum threshold of four variables, was performed using Lasso, Cox-Boost, random survival forest (RSF), StepCox [both], and StepCox [backward] in the TCGA-OV cohort.²⁹ Subsequently, a combination of 10 machine learning algorithms (lasso, RSF, GBM, Survival-SVM, SuperPC, ridge regression, plsRcox, CoxBoost, StepCox, and enet) were used to create a composite model, with RSF being identified as the best prognostic model, based on the average C-index within all cohorts.

2.9 | Statistical analysis

We conducted all statistical analyses using version 4.1.2 of the R software. For detailed information regarding the statistical methods implemented, please refer to the previous sections and cited references. We set the level of statistical significance at $p < .05$.

3 | RESULTS

3.1 | Tregs could be considered prognostic marker for OC

We employed the xCell algorithm to predict the Tregs infiltration scores, utilizing data from the TCGA-OV, GPL96, GPL14951, GPL570, GPL2986, and GPL7759 cohorts. We classified all samples into either a high or low Tregs score group based on the best cutoff value of the scores. Intriguingly, except for the GPL2986 cohort, our findings revealed a significant association between higher Tregs scores and poorer overall survival (OS) (Figure 1A). $CD8^+$ T cells, a specific subset of leukocytes, secrete various cytokines involved in immune actions.¹³ Hence, we investigated the correlation between FOXP3, a specific marker of Tregs, and $CD8^+$ T cells in TCGA-OV cohort. Our results indicated that as the expression of FOXP3 increased, $CD8^+$ T cells decreased (Figure 1B). Moreover, the immunofluorescence results indicated a decrease in Tregs within the primary tumor compared to the metastasis tumor, with a subsequent increase in $CD8^+$ T cells content around Tregs (Figure 1C,D).

3.2 | Weighted gene co-expression network analysis of Tregs

In the subsequent analysis, we conducted WGCNA on the TCGA-OV and the GPL570 cohorts. We utilized the xCell-calculated Tregs score as the phenotypic data for WGCNA. For the TCGA-OV cohort, nine co-expression modules were clustered, and the turquoise color module exhibited the strongest positive correlation with the xCell score ($Cor = 0.31$, $P = 8e-10$) (Figure 2A). Similarly, for the GPL570 cohort, nine co-expression modules were clustered, where the pink color module demonstrated the strongest positive correlation with the Tregs score ($Cor = 0.15$, $P = 1e-04$) (Figure 2B). Ultimately, seven shared genes within the turquoise color module and pink color

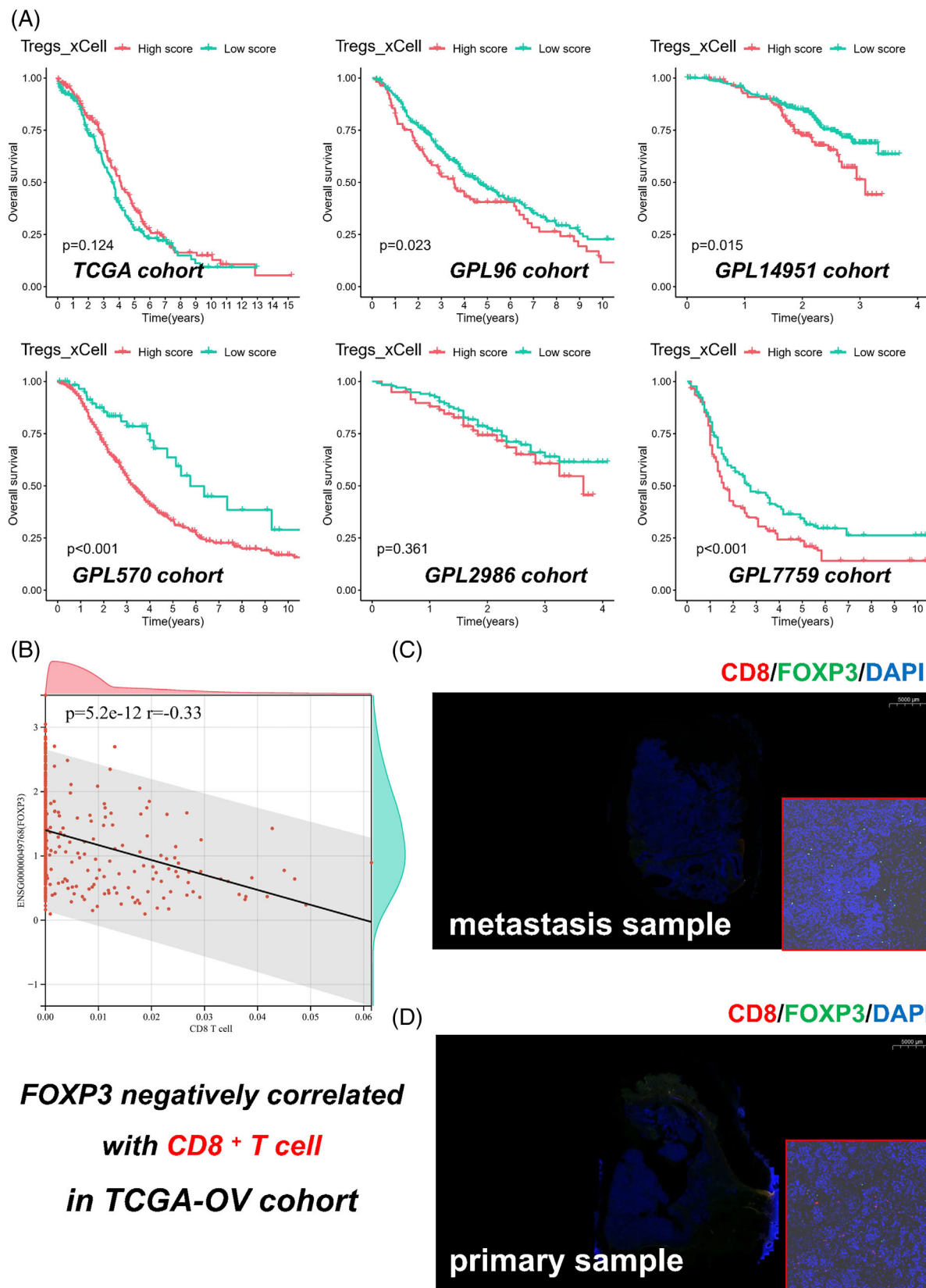
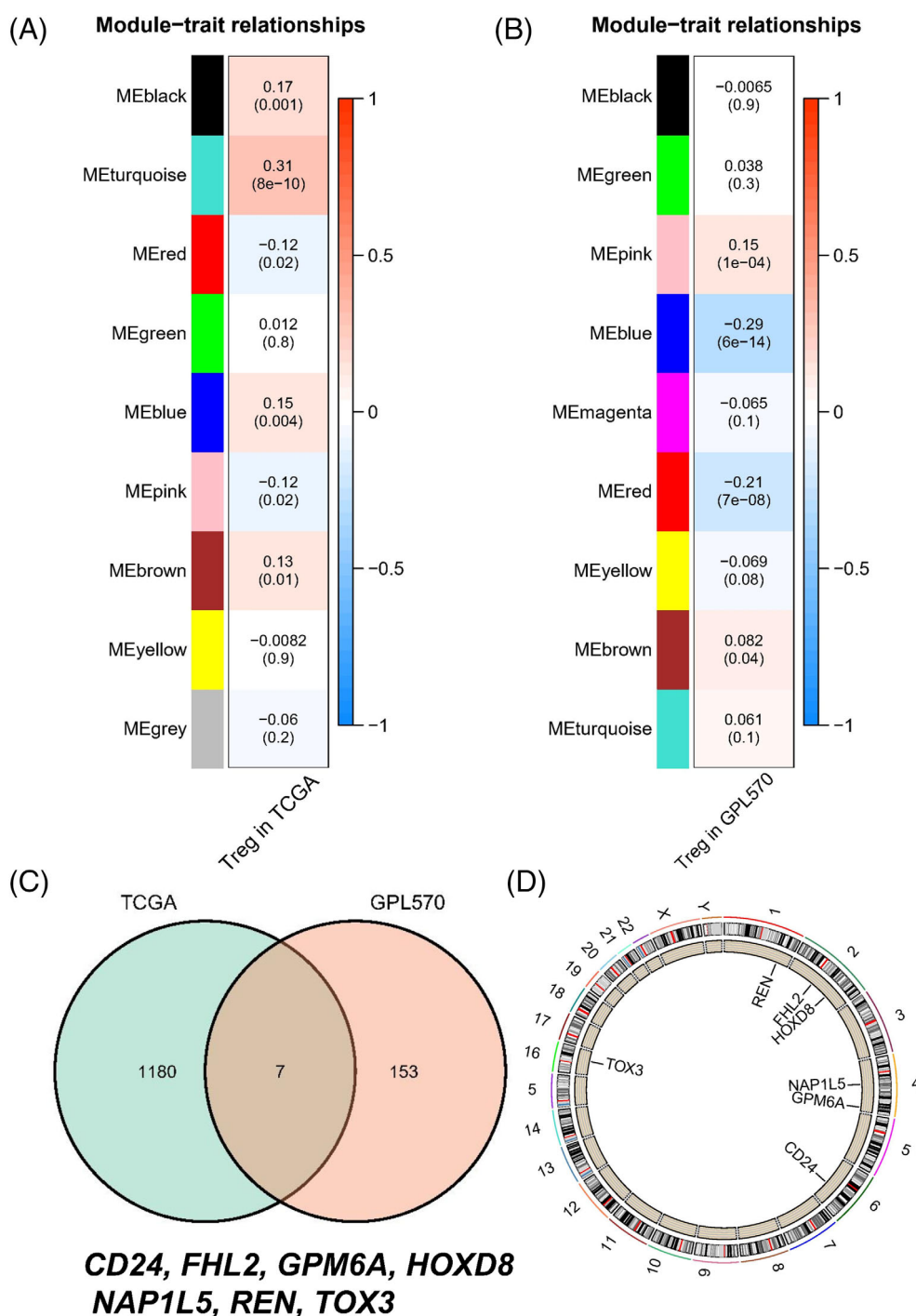


FIGURE 1 Tregs Infiltration Scores Predict Prognosis and Correlate with CD8⁺ T Cells in Ovarian Cancer Samples. (A) Kaplan–Meier curves illustrating the overall survival (OS) of ovarian cancer patients in high and low Tregs score groups, as determined by the xCell algorithm. Significance was determined using the log-rank test, with p -values indicated on the graph. (B) Scatter plot depicting the negative correlation between FOXP3 expression and score of CD8⁺ T cells in the TCGA-OV cohort. (C) Representative images of immunofluorescence staining showing changes in Tregs and CD8⁺ T cells from metastasis tumors. (D) Representative images of immunofluorescence staining showing changes in Tregs and CD8⁺ T cells from primary tumors.

FIGURE 2 Weighted Gene Co-Expression Network Analysis of Tregs in Ovarian Cancer. (A) weighted gene co-expression network analysis (WGCNA) on the TCGA-OV cohort revealed nine co-expression modules. The turquoise color module displayed the strongest positive correlation with the xCell score, suggesting a potential association with Tregs. (B) WGCNA on the GPL570 cohort also identified nine co-expression modules, with the pink color module exhibiting the strongest positive correlation with the Tregs score. (C) Seven shared genes within the turquoise color module and pink color module were identified as key Tregs-related genes (TRGs) in ovarian cancer. (D) Chromosomal location of the seven TRGs: CD24, FHL2, GPM6A, HOXD8, NAP1L5, REN, and TOX3.



module were identified as key Tregs related genes (TRGs) in OC (Figure 2C). Figure 2D illustrated the chromosomal location of the aforementioned 7 TRGs (CD24, FHL2, GPM6A, HOXD8, NAP1L5, REN, and TOX3).

3.3 | Key TRGs-derived molecular subtypes

The unsupervised classification of OC patients based on the expression of key TRGs in meta cohorts proved optimal with a K value of

2. We identified two distinct molecular subtypes, with subtype A consisting of 484 cases and subtype B consisting of 487 cases (Figure 3A). The PCA displayed a relative dispersion between the two subtypes (Figure 3B). Furthermore, the heatmap exhibited the distribution of clinical features across different subtypes, with most TRGs being notably over-expressed in subtype A (Figure 3C). To investigate the underlying cause for their varying prognosis, we conducted GSVA. Remarkably, subtype B exhibited significant enrichment of pathways linked to multiple carcinogenesis compared to subtype A (Figure 3D, E). Additionally, we analyzed mutation profiles of different subtypes

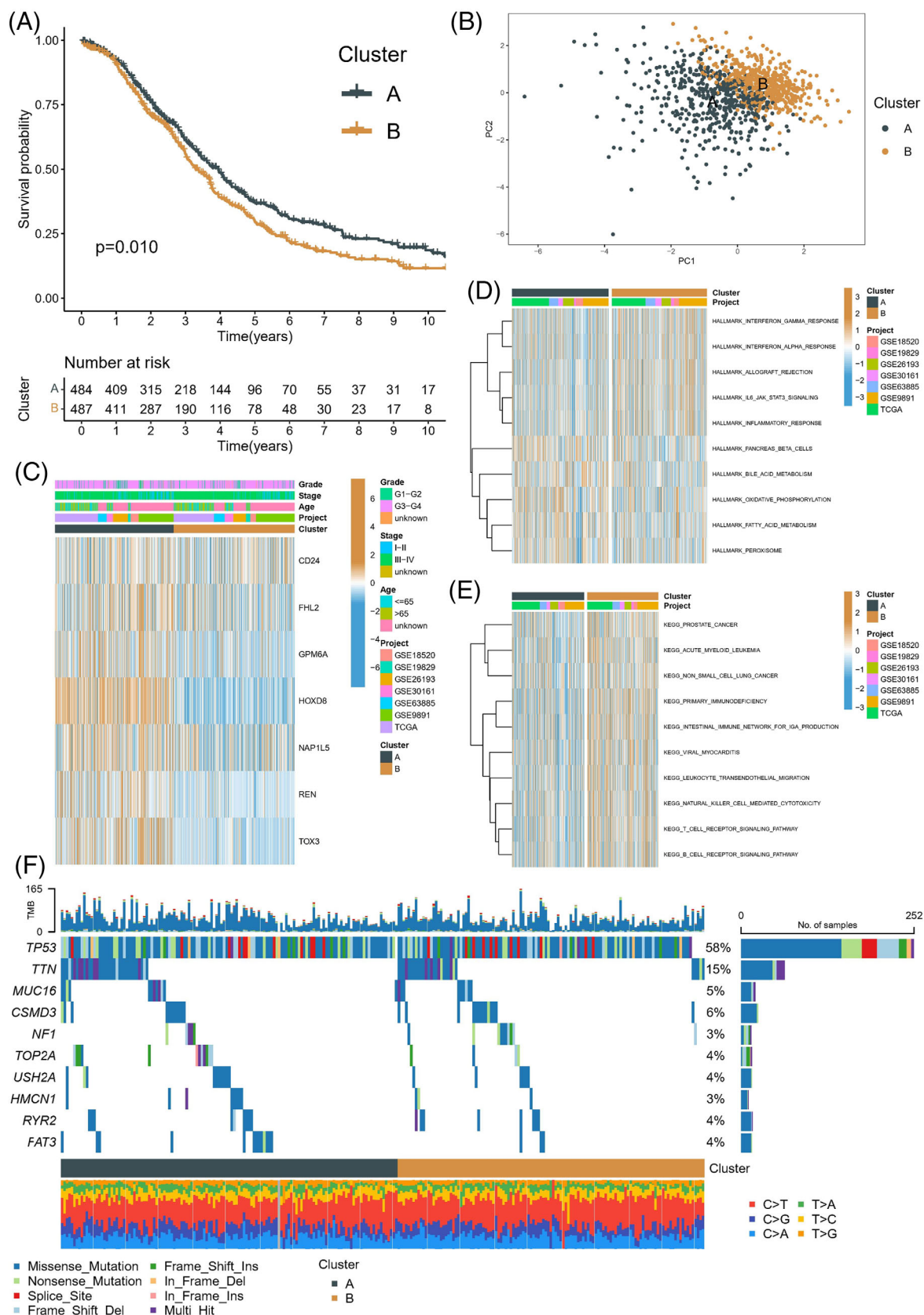


FIGURE 3 Key Tregs-related genes (TRGs)-derived Molecular Subtypes in Ovarian Cancer. (A) Unsupervised classification of ovarian cancer (OC) patients based on the expression of key TRGs in meta cohorts, resulting in two distinct molecular subtypes: subtype A (484 cases) and subtype B (487 cases). (B) Principal component analysis (PCA) displaying relative dispersion between the two subtypes. (C) Heatmap showing the distribution of clinical features across different subtypes, with the majority of TRGs over-expressed in subtype A. (D, E) Gene Set Variation Analysis (GSVA) revealing significant enrichment of pathways linked to multiple carcinogenesis in subtype B compared to subtype A. (F) Mutation profiles of different subtypes, with significantly higher TTN mutations observed in subtype A samples compared with subtype B.

and observed significantly higher TTN mutations in subtype A samples compared with subtype B (Figure 3F).

3.4 | Characteristics of immune microenvironment in TRGs-derived molecular subtypes

Remarkably, subtype B exhibited higher mRNA expression in most Human leukocyte antigen (HLA) and immune checkpoint-related mRNA (Figure 4A,B). Furthermore, the ESTIMATE algorithm reaffirmed that subtype B had a higher immune score (Figure 4C). The box plot displayed the TME status in various molecular subtypes, demonstrating significantly upregulated Treg cells in subtype B. This may help to explain the poor prognosis observed in some patients with a “hot tumor” status (Figure 4D). In addition, we predicted IC50 values of various common drugs and found that subtype B had better sensitivity for most of them (Figure 4E). Furthermore, we compared the differences between the two subtypes and identified 58 DEGs (Figure 5A). The main enrichment pathways of these genes may be related to biological processes such as the Wnt signaling pathway (Figure 5B). These results provide evidence to support the prognosis of the different subtypes. Using one-way cox regression analysis, we identified 11 prognostic DEGs (TIMP3, SFRP2, AGR2, FN1, POSTN, FAP, COL11A1, NDP, COMP, GREM1, TFF3). Patients were categorized according to the expression of these prognostic DEGs, achieving the best classification with a K value of 3. Ultimately, three different regulatory clusters were identified (Figure 5C), of which cluster C had the best prognosis. PCA revealed relative dispersion of the three molecular clusters (Figure 5D). Interestingly, the heat map showed that the expression of key TRGs also differed across the different clusters (Figure 5E).

3.5 | Machine learning-derived Tregs signature

The TCGA-OV cohort was implemented for modeling in the machine learning model. Validation was performed on the GSE18520, GSE19829, GSE26193, GSE30161, GSE63885, and GSE9891 cohorts. RSF was identified as the best prognostic model based on the mean C-index. The model incorporated the 11 previously identified prognostic DEGs (Figure 6A). To validate the model, we implemented it in multiple cohorts. Notably, the model better-differentiated patient survival in most cohorts (Figure 6B–F).

3.6 | Heterogeneity of cohorts distinguished by Tregs signature

The Sankey plot effectively displayed the relationship between risk subtypes and molecular subtypes, indicating that there is a higher risk of death in the high-risk group (Figure 7A). Moreover, molecular subtypes with poorer prognosis have a higher risk score (Figure 7B,C). Notably, the proportion of wound healing (C1) subtypes was

significantly higher in the high-risk subtypes and had a higher risk score (Figure 7D,E). For immune cell analysis, multiple algorithms were used simultaneously to estimate immune cell infiltration in various samples. The heat map revealed that the low risk group had a more active TME (Figure 7F). Moreover, mutated genes based on whole-exome data were analyzed, and no significant differences were observed between high- and low-risk patients (Figure 7G). In the immunotherapy cohort, the RSF model also differentiated survival risk (Figure 7H), with the risk score higher in the SD/PD group (Figure 7I).

3.7 | Validation of key Tregs-related genes in clinical samples

In order to better articulate the expression levels of various TRGs in clinical samples, we conducted an investigation into mRNA expression using RNA-seq technology among normal and OC samples. Our findings indicated that of the TRGs examined, only CD24 exhibited high expression in tumor samples, while GPM6A, HOXD8, NAP1L5, REN, and TOX showed elevated expression in normal samples. Conversely, FHL3 demonstrated no differential expression across various sample types (Figure 8A). Subsequently, we further explored the protein expression patterns of these TRGs. Representative results demonstrated a notable concordance between CD24 and NAP1L5 protein and mRNA expression trends (Figure 8B). However, there were no significant differences observed in the expression of GPM6A, TOX, REN (FHL3, HOXD8 without corresponding IHC samples). In addition, we analyzed the GSE151214 single-cell dataset across nine clusters (Figure 8C) to examine the expression of risk score. We found that epithelial and malignant cells exhibited particularly high-risk score, suggesting a critical role for TRGs in the development and progression of OC (Figure 8D). To further confirm our findings, we conducted qRT-PCR experiments based on clinical samples obtained from our hospital. Our results confirmed the differential expression of CD24, REN, TOX3, and GPM6A, while the remaining data for TRGs was statistically insignificant and therefore not presented (Figure 8E). It is worth noting that the use of clinical samples from our hospital provided additional support for the significance of these findings and reinforces the potential importance of these TRGs in OC.

4 | DISCUSSION

OC is one of the most lethal gynecologic malignancies worldwide, with a high mortality rate.¹ The current standard of care for OC includes surgical debulking followed by chemotherapy. However, despite advances in treatment, OC remains difficult to treat due to its heterogeneity and lack of reliable biomarkers for diagnosis and prognosis.³² There has been a growing interest in the role of tumor-infiltrating lymphocytes (TILs) in the pathogenesis of OC, and various studies have shown that TILs can be used as potential prognostic markers for this disease.^{13,33} Tregs, a subpopulation of CD4⁺ T cells, are known to play a key role in suppressing immune responses

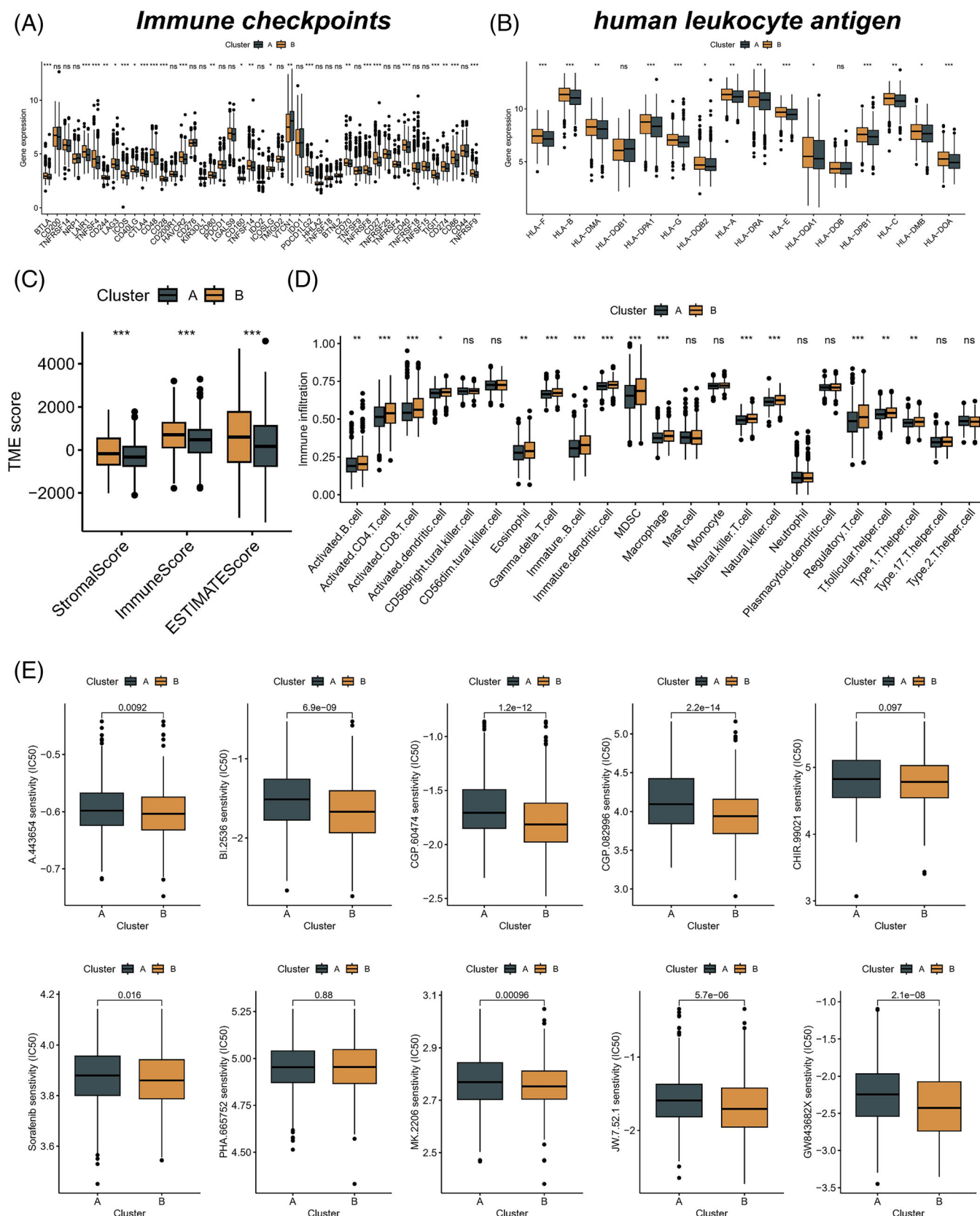


FIGURE 4 Immune-related characteristics and drug sensitivity of key Tregs-related genes (TRGs)-derived molecular subtypes in ovarian cancer. (A) Box plot showing the mRNA expression of most immune checkpoint-related mRNA for different subtypes, with subtype B exhibiting higher expression in these genes. (B) Box plot showing the mRNA expression of most Human leukocyte antigen (HLA) for different subtypes, with subtype B exhibiting higher expression in these genes. (C) ESTIMATE algorithm re-affirming the higher immune score for subtype B compared to subtype A. (D) Box plot demonstrating the differences in tumor microenvironment (TME) status in various molecular subtypes, showing significantly upregulated Treg cells in subtype B. (E) Predicted IC50 values of various common drugs, with subtype B exhibiting better sensitivity for most of the drugs.

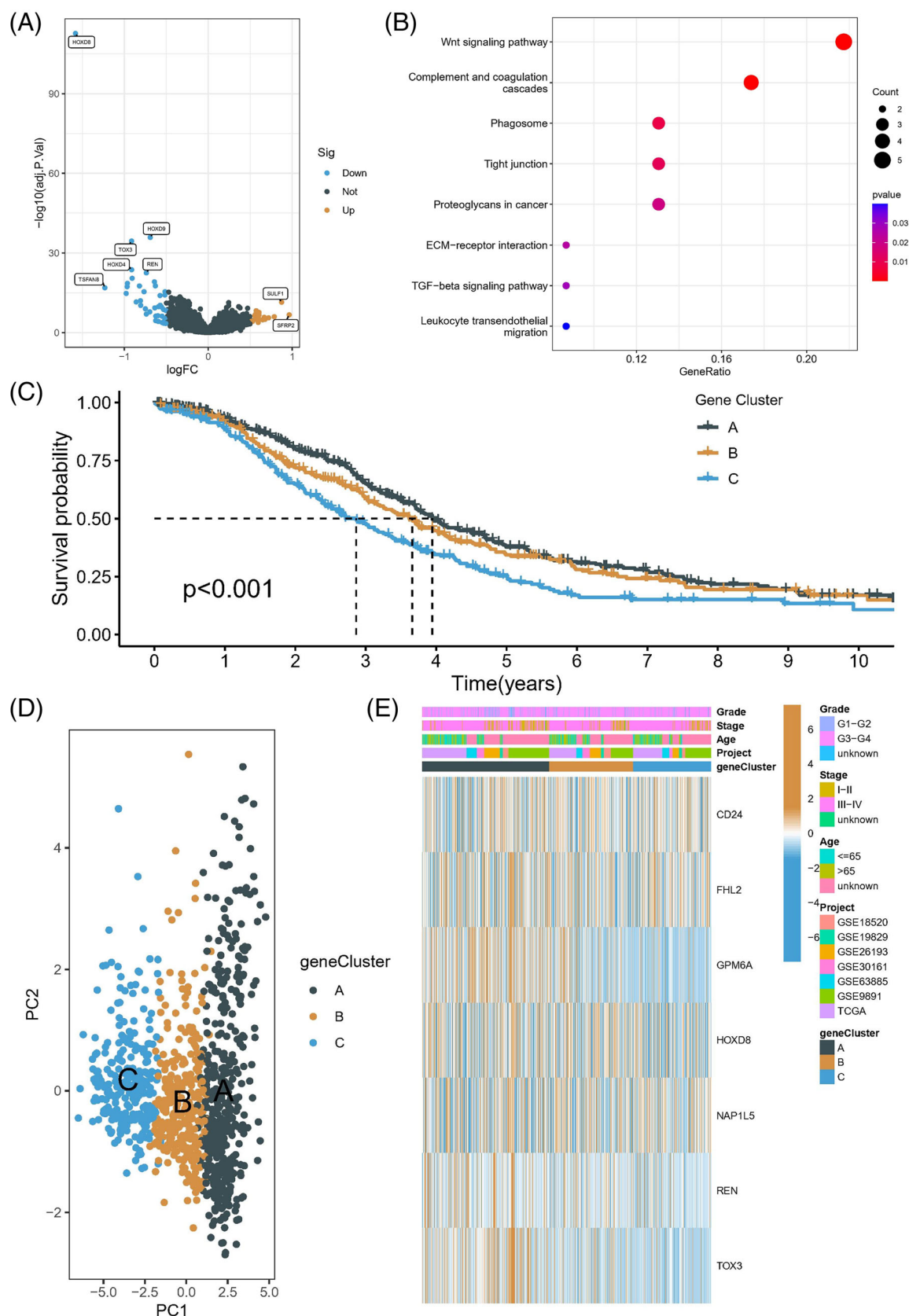


FIGURE 5 Identification of differential expression genes (DEGs) and Prognostic Analysis of Key Tregs-related genes (TRGs)-derived molecular subtypes in ovarian cancer. (A) Volcano plot displaying the differential expression of 58 DEGs between the two subtypes, indicating significant differences in gene expression profiles. (B) Pathway enrichment analysis indicating that the main enrichment pathways of these genes are related to biological processes such as the Wnt signaling pathway. (C) Kaplan-Meier curves illustrating three different regulatory clusters (Clusters A, B, and C), with Cluster C having the best prognosis. (D) Principal component analysis showing relative dispersion of the three molecular clusters. (E) Heatmap demonstrating differences in expression of key TRGs across the three clusters.

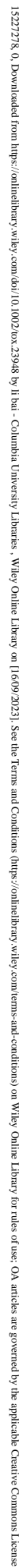


FIGURE 6 Machine Learning-derived Tregs signature for prognosis of ovarian cancer patients. (A) The machine learning model incorporated the 11 previously identified prognostic DEGs and was trained using the TCGA-OV cohort. Random survival forest (RSF) was identified as the best prognostic model based on the mean C-index. (B) Validation of the model in multiple cohorts (GSE18520, GSE19829, GSE26193, GSE30161, GSE63885, and GSE9891), demonstrating that the model better differentiated patient survival in most cohorts.

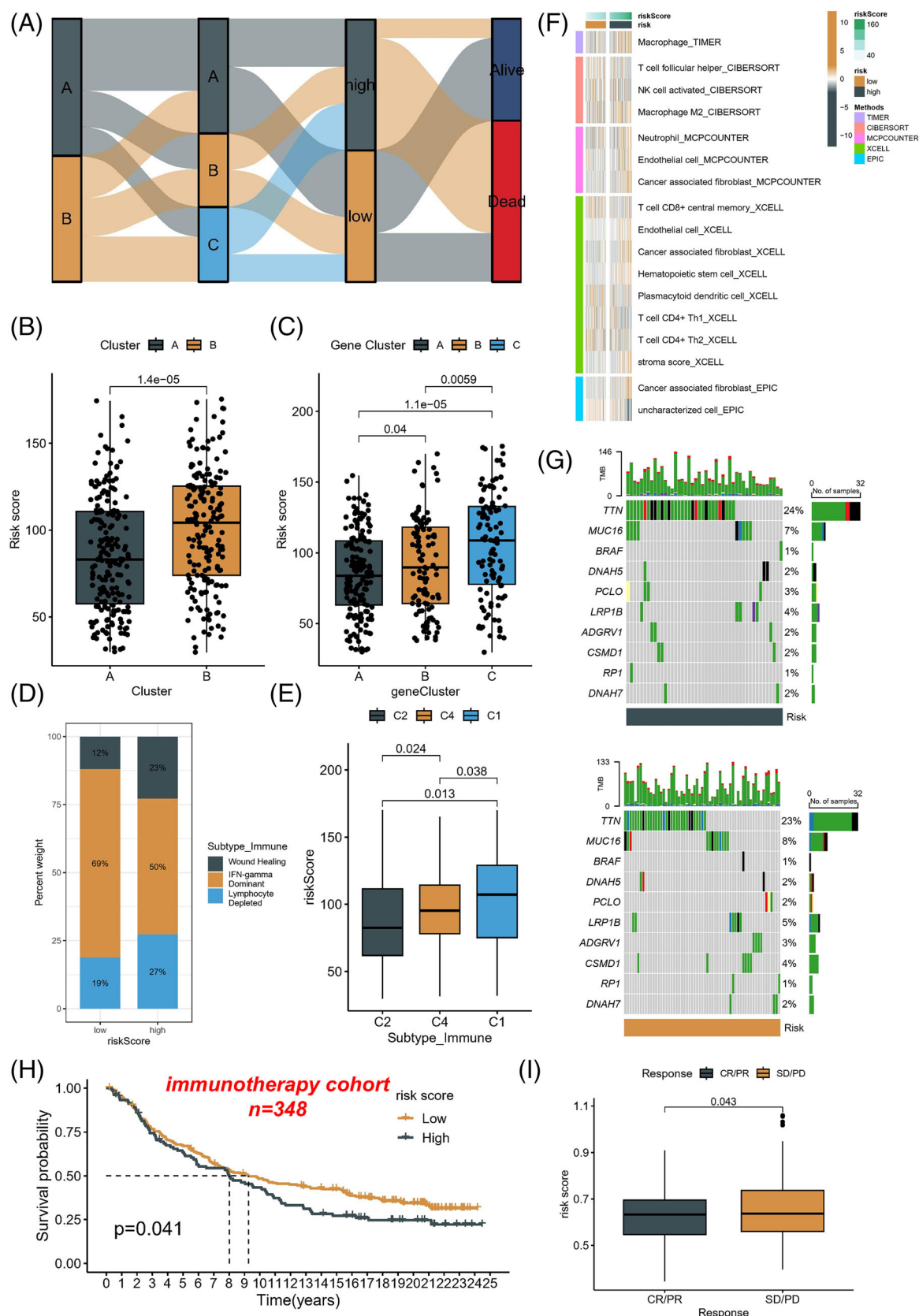


FIGURE 7 Heterogeneity of cohorts distinguished by Tregs signature in ovarian cancer. (A) Sankey plot demonstrating the relationship between risk subtypes and molecular subtypes, with a higher risk of death observed in the high-risk group. (B, C) Molecular subtypes (Cluster B and regulator Cluster C) with poorer prognosis have a higher risk score, as demonstrated by the boxplots. (D, E) Proportions of wound healing (C1) subtypes are significantly higher in the high-risk subtypes and have a higher risk score. (F) Heatmap showing that the low-risk group has a more active tumor microenvironment (TME), based on immune cell analysis using multiple algorithms. (G) Analysis of mutated genes based on whole-exome data, showing no significant differences between high and low-risk patients. (H) The random survival forest (RSF) model differentiated survival risk in the immunotherapy cohort, as demonstrated by the Kaplan-Meier plots. (I) Box plot showing risk scores were higher in the SD/PD group.

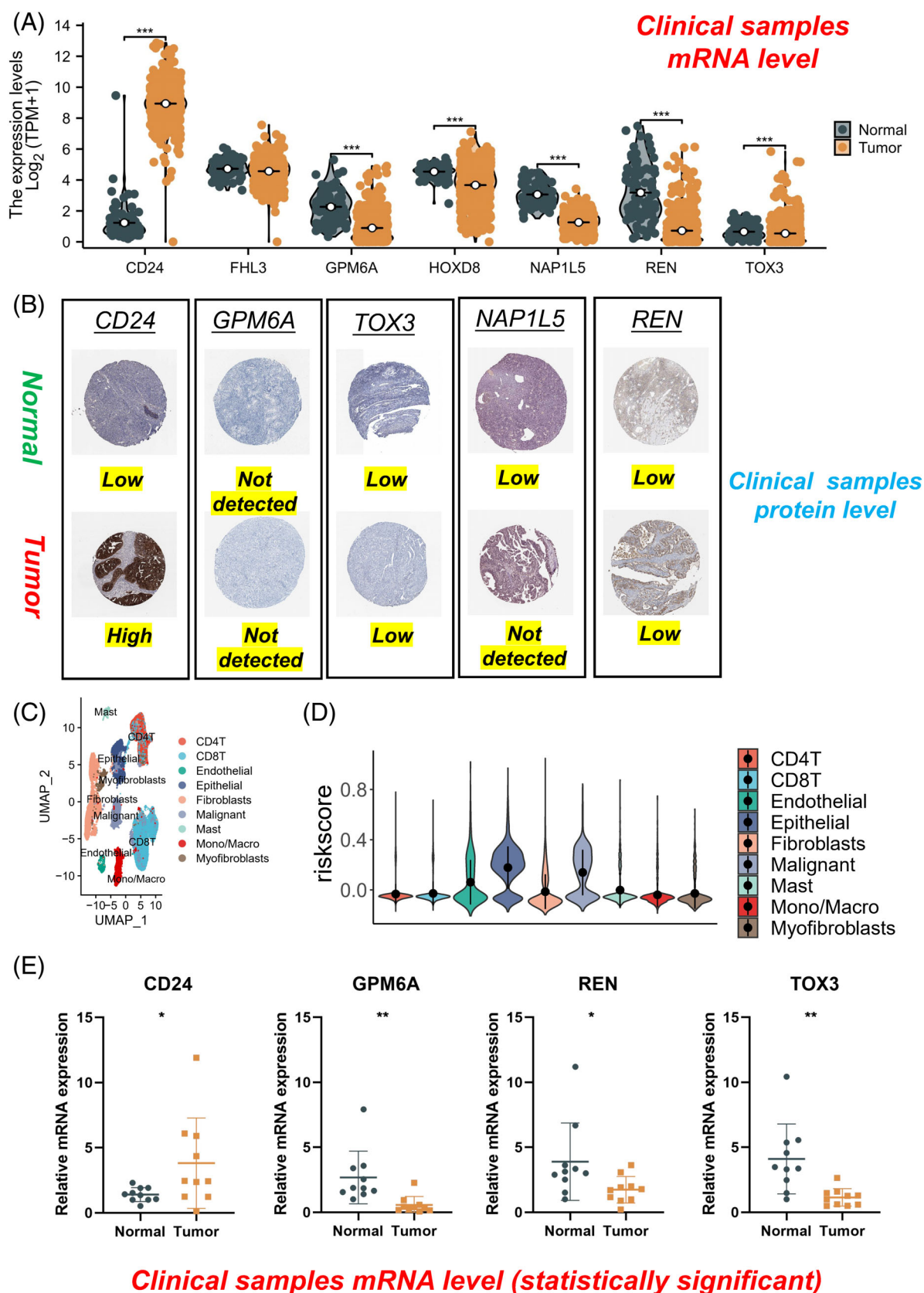


FIGURE 8 Validation of Key Tregs-related genes (TRGs) in clinical samples of ovarian cancer. (A) mRNA expression levels of various TRGs examined in GTEx database indicate that only CD24 was highly expressed in tumor samples, while GPM6A, HOXD8, NAP1L5, REN, and TOX showed elevated expression in normal samples. (B) Concordance between CD24 and NAP1L5 protein and mRNA expression trends were observed, but no significant differences were present in the expression of GPM6A, TOX, REN, HOXD8, and FHL3 without corresponding IHC samples. (C) Single-cell dataset GSE151214 was analyzed across nine cluster. (D) Epithelial and malignant cells exhibited particularly high-risk scores. (E) qRT-PCR experiments conducted based on clinical samples obtained from our hospital confirmed the differential expression of CD24, REN, TOX3, and GPM6A.

and promoting immune tolerance.¹⁵ Tregs have long been recognized as potent immunosuppressive cells that regulate immune responses, and recent studies have shown that Tregs play a role in promoting tumor growth and metastasis in various types of cancer, including OC [5; 14; 15].

We identified a potential Tregs-related signature using an integrative bioinformatics approach. Firstly, we found that several Tregs-related genes were differentially expressed between OC tissues and normal tissues. We also found that the expression of these genes was associated with patient survival, suggesting that these genes may be useful prognostic markers for OC. Moreover, the expression of FOXP3 increased, CD8⁺ T cells decreased. The results indicated a decrease in Tregs within the primary tumor compared to the metastasis tumor, with a subsequent increase in CD8⁺ T cells content around Tregs. These findings are consistent with previous studies that have shown a correlation between TILs and patient prognosis in OC.¹³

The increasing prevalence of machine learning techniques for predicting patient survival can be attributed to several reasons. These methods have gained prominence due to their ability to handle complex and high-dimensional data, which is often encountered in healthcare settings. By employing advanced algorithms, machine learning models can effectively extract meaningful patterns and relationships from diverse datasets comprising clinical, genomic, and other patient-specific information. Hence, we used machine learning to develop a prognostic model based on the Tregs-related signature. The model incorporated 11 previously identified prognostic DEGs and was validated in multiple cohorts. We found that the model better differentiated patient survival in most cohorts, suggesting that it may serve as a reliable prognostic marker for OC. CD24, FHL2, GPM6A, HOXD8, NAP1L5, REN, and TOX3 are seven Tregs-related genes (TRGs) that have been identified as potential prognostic markers for various types of cancer. Clinicians can use the signature to assess the likelihood of disease progression, recurrence, or response to specific treatments. This information can aid in determining appropriate follow-up intervals, monitoring disease response, and facilitating early intervention if necessary. CD24 is a cell surface glycoprotein that has been implicated in various biological processes, including cell adhesion, migration, and differentiation.³⁴ High levels of CD24 expression have been associated with poor prognosis in several types of cancer, including breast cancer, pancreatic cancer, and colorectal cancer (CRC).³⁴ FHL2 is a member of the four-and-a-half LIM domain protein family that plays a role in regulating gene expression and cell proliferation.³⁵ On the one hand, FHL2 functions as a tumor suppressor by down-regulating in tumor tissue and impeding cell proliferation to inhibit tumor growth. On the other hand, FHL2 serves as an oncoprotein by up-regulating in tumor tissue and binding to various transcription factors to suppress cell apoptosis, promote cell proliferation and migration, and facilitate tumor progression.³⁵ GPM6A is a glycoprotein that is involved in regulating cell adhesion and migration.³⁶ HOXD8 is a homeobox transcription factor that plays a role in regulating embryonic development and cell differentiation, and overexpression of HOXD8 inhibits tumor growth both in vitro and in vivo by suppressing AKT and mTOR phosphorylation and inducing apoptosis.³⁷ Another

study demonstrated that stable expression of HOXD8 inhibits cell proliferation, anchorage-independent growth, and invasion, and can induce apoptotic events in CRC cells by upregulating executioner caspases 6 and 7 and cleaved PARP.³⁸ NAP1L5 is a nucleosome assembly protein that is involved in regulating chromatin structure and gene expression.²⁵ The study by Zhao and colleagues examined the function of NAP1L5 in hepatocellular carcinoma (HCC), and low expression of NAP1L5 was found in HCC, and its downregulation was associated with shorter survival and disease-free survival.³⁹ In addition, the study found that NAP1L5 inhibits the PI3K/AKT/mTOR signaling pathway in HCC by regulating MYH9, suggesting that NAP1L5 may be a potential target.³⁹ REN is a renal peptide hormone that plays a role in regulating blood pressure and fluid balance.²⁶ TOX3 is a transcription factor that plays a role in regulating gene expression and cell proliferation.⁴⁰ Yang et al. investigated the expression and role of TOX high mobility group box family member 3 in CRC.⁴¹ Notably, the authors found that messenger RNA and protein expression levels of TOX3 were markedly upregulated in CRC tissues and cell lines. High TOX3 expression was associated with high T stage, nodal invasion, and advanced tumor stage.⁴¹

Importantly, the Sankey plot revealed a higher risk of death in the high-risk group and a correlation between molecular subtypes with poorer prognosis and a higher risk score, supporting the prognostic relevance of the Tregs-related signature. These findings are consistent with previous studies that have shown the prognostic value of molecular subtyping in OC.^{28,42} The results of this study suggest that Tregs and their related genes may play a critical role in OC progression and prognosis. The identification of a Tregs-related signature that can be used as a prognostic marker for OC has important clinical implications, as it may enable clinicians to identify patients at high risk of recurrence or poor survival outcomes and tailor treatment accordingly.

One limitation of this study is that it relied on bioinformatics analysis and machine learning algorithms, which may be subject to bias. We recognize that patient demographics can vary across different cohorts and geographic regions. Factors such as age, gender, ethnicity, and socioeconomic status may influence disease presentation, progression, and treatment outcomes. While our study includes a diverse multicenter cohort, we acknowledge that there might be inherent biases due to these demographic differences. Therefore, future studies should validate the prognostic value of the Tregs-related signature using experimental methods, such as flow cytometry, to confirm the expression levels of TILs and their related genes in OC tissues. Another limitation is that this study focused on public cohorts of patients with OC. Future studies should investigate the prognostic value of the Tregs-related signature in larger and more diverse patient populations to determine its applicability across different subtypes of OC.

5 | CONCLUSIONS

In conclusion, the results of this study provide compelling evidence that a Tregs-related signature can serve as a reliable prognostic marker for OC. These findings have important implications for the

development of personalized treatment strategies for patients with OC and may enable clinicians to identify patients at high risk of recurrence or poor survival outcomes. Further research is needed to validate these findings and determine the clinical utility of Tregs-related gene signatures in OC prognosis and treatment.

AUTHOR CONTRIBUTIONS

Yang Cao contributed in conception and design of the study. Yang Cao and Ying-lei Liu drafted the manuscript. Xiao-Yan Lu, Hai-Li Kai, and Yun Han critically revised the manuscript for important intellectual content. Yan-Li Zheng approved the final draft.

FUNDING INFORMATION

This work was supported by the Jiangsu Provincial Health Commission Project (grant Z2021078), Nantong Science and Technology Bureau Plan Project (grant MS2022005), and Scientific Research and Innovation Team Project of Kangda College of Nanjing Medical University (KD2022KYCXTD010).

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article and further enquiries can be directed to the corresponding author. The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Yang Cao  <https://orcid.org/0009-0008-7031-4116>

REFERENCES

- Menon U, Karpinskyj C, Gentry-Maharaj A. Ovarian cancer prevention and screening. *Obstet Gynecol*. 2018;131:909-927.
- Jimenez-Sanchez A, Cybulska P, Mager KL, et al. Unraveling tumor-immune heterogeneity in advanced ovarian cancer uncovers immunogenic effect of chemotherapy. *Nat Genet*. 2020;52:582-593.
- Xiao Y, Yu D. Tumor microenvironment as a therapeutic target in cancer. *Pharmacol Ther*. 2021;221:107753.
- Nishikawa H, Sakaguchi S. Regulatory T cells in tumor immunity. *Int J Cancer*. 2010;127:759-767.
- Tanaka A, Sakaguchi S. Regulatory T cells in cancer immunotherapy. *Cell Res*. 2017;27:109-118.
- Batchu RB, Gruzdyn OV, Kolli BK, et al. IL-10 signaling in the tumor microenvironment of ovarian cancer. *Adv Exp Med Biol*. 2021;1290:51-65.
- Ye C, Yano H, Workman CJ, Vignali DAA. Interleukin-35: structure, function and its impact on immune-related diseases. *J Interferon Cytokine Res*. 2021;41:391-406.
- Moreau JM, Velegaki M, Bolyard C, Rosenblum MD, Li Z. Transforming growth factor-beta1 in regulatory T cell biology. *Sci Immunol*. 2022;7:eabi4613.
- Horii M, Matsushita T. Regulatory B cells and T cell regulation in cancer. *J Mol Biol*. 2021;433:16685.
- Kalbasi A, Ribas A. Tumour-intrinsic resistance to immune checkpoint blockade. *Nat Rev Immunol*. 2020;20:25-39.
- Kurniawan H, Soriano-Baguet L, Brenner D. Regulatory T cell metabolism at the intersection between autoimmune diseases and cancer. *Eur J Immunol*. 2020;50:1626-1642.
- Ju M, Fan J, Zou Y, et al. Computational recognition of a regulatory T-cell-specific signature with potential implications in prognosis, immunotherapy, and therapeutic resistance of prostate cancer. *Front Immunol*. 2022;13:807840.
- Santoiemma PP, Powell DJ Jr. Tumor infiltrating lymphocytes in ovarian cancer. *Cancer Biol Ther*. 2015;16:807-820.
- Wang B, Zhang Z, Liu W, Tan B. Targeting regulatory T cells in gastric cancer: pathogenesis, immunotherapy, and prognosis. *Biomed Pharmacother*. 2023;158:114180.
- Shan F, Somasundaram A, Bruno TC, Workman CJ, Vignali DAA. Therapeutic targeting of regulatory T cells in cancer. *Trends Cancer*. 2022;8:944-961.
- Feng S, Xu Y, Dai Z, Yin H, Zhang K, Shen Y. Integrative analysis from multicenter studies identifies a WGCNA-derived cancer-associated fibroblast signature for ovarian cancer. *Front Immunol*. 2022;13:951582.
- Shen X, Yang Z, Feng S, Li Y. Identification of uterine leiomyosarcoma-associated hub genes and immune cell infiltration pattern using weighted co-expression network analysis and CIBERSORT algorithm. *World J Surg Oncol*. 2021;19:223.
- Zhang J, Zhang H, Wang Y, Wang Q. MCM2-7 in clear cell renal cell carcinoma: MCM7 promotes tumor cell proliferation. *Front Oncol*. 2021;11:782755.
- Zhao W, Langfelder P, Fuller T, Dong J, Li A, Hovarth S. Weighted gene coexpression network analysis: state of the art. *J Biopharm Stat*. 2010;20:281-300.
- Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*. 2008;9:559.
- Li B, Zhou CX, Pu YQ, Qiu L, Mei W, Xiong W. Expression of CD24 gene in human malignant pleural mesothelioma and its relationship with prognosis. *Zhonghua Lao Dong Wei Sheng Zhi Ye Bing Za Zhi*. 2023;41:168-176.
- Algaber A, Madhi R, Hawez A, Ronnow CF, Rahman M. Targeting FHL2-E-cadherin axis by miR-340-5p attenuates colon cancer cell migration and invasion. *Oncol Lett*. 2021;22:637.
- Zhang Q, Deng S, Li Q, Wang G, Guo Z, Zhu D. Glycoprotein M6A suppresses lung adenocarcinoma progression via inhibition of the PI3K/AKT pathway. *J Oncol*. 2022;2022:4601501.
- Ye W, Huang Y, Zhu G, et al. miR-30a inhibits the osteogenic differentiation of the tibia-derived MSCs in congenital pseudarthrosis via targeting HOXD8. *Regen Ther*. 2022;21:477-485.
- Wang B, Liu W, Sun F. Nucleosome assembly protein 1-like 5 alleviates Alzheimer's disease-like pathological characteristics in a model. *Front Mol Neurosci*. 2022;15:1034766.
- Mishima S, Mitsui T, Tani K, et al. Endothelin-1 production via placental (pro)renin receptor in a mouse model of preeclampsia. *Placenta*. 2023;138:44-50.
- Mohamad Shah NS, Sulong S, Wan Sulaiman WA, Halim AS. Two novel genes TOX3 and COL21A1 in large extended Malay families with nonsyndromic cleft lip and/or palate. *Mol genet. Genomic Med*. 2019;7:e635.
- Feng S, Xia T, Ge Y, et al. Computed tomography imaging-based Radiogenomics analysis reveals hypoxia patterns and immunological characteristics in ovarian cancer. *Front Immunol*. 2022;13:868067.
- Li Y, Niu JH, Wang Y. Machine learning-based neddylation landscape indicates different prognosis and immune microenvironment in endometrial cancer. *Front Oncol*. 2023;13:1084523.
- Kanehisa M, Goto S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res*. 2000;28:27-30.
- Hanzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics*. 2013;14:7.
- White KL, Schildkraut JM, Palmieri RT, et al. Ovarian cancer association, ovarian cancer risk associated with inherited inflammation-related variants. *Cancer Res*. 2012;72:1064-1069.
- Quail DF, Joyce JA. Microenvironmental regulation of tumor progression and metastasis. *Nat Med*. 2013;19:1423-1437.

34. Fang X, Zheng P, Tang J, Liu Y. CD24: from a to Z. *Cell Mol Immunol*. 2010;7:100-103.
35. Zhang J, Zeng Q, She M. The roles of FHL2 in cancer. *Clin Exp Med*. Published online April 27, 2023. doi:[10.1007/s10238-023-01076-3](https://doi.org/10.1007/s10238-023-01076-3)
36. Leon A, Aparicio GI, Scorticati C. Neuronal glycoprotein M6a: an emerging molecule in chemical synapse formation and dysfunction. *Front Synaptic Neurosci*. 2021;13:661681.
37. Zhang Y, Yu Y, Su X, Lu Y. HOXD8 inhibits the proliferation and migration of triple-negative breast cancer cells and induces apoptosis in them through regulation of AKT/mTOR pathway. *Reprod Biol*. 2021;21:100544.
38. Mansour MA, Senga T. HOXD8 exerts a tumor-suppressing role in colorectal cancer as an apoptotic inducer. *Int J Biochem Cell Biol*. 2017;88:1-13.
39. Zhao R, Ge Y, Gong Y, Li B, Xiao B, Zuo S. NAP1L5 targeting combined with MYH9 inhibit HCC progression through PI3K/AKT/mTOR signaling pathway. *Aging (Albany NY)*. 2022;14:9000-9019.
40. Wang Y, Wang Z, Shao C, et al. Melatonin may suppress lung adenocarcinoma progression via regulation of the circular noncoding RNA hsa_circ_0017109/miR-135b-3p/TOX3 axis. *J Pineal Res*. 2022;73:e12813.
41. Yang W, Wu W, Liang H, Chen J, Dong X. TOX3 regulates the proliferation and apoptosis of colorectal cancer by downregulating RhoB via the activation of the MAPK pathway. *Cell Biol Int*. 2022;46:1074-1088.
42. Feng S, Yin H, Zhang K, et al. Integrated clinical characteristics and omics analysis identifies a ferroptosis and iron-metabolism-related lncRNA signature for predicting prognosis and therapeutic responses in ovarian cancer. *J Ovarian Res*. 2022;15:10.

How to cite this article: Cao Y, Liu Y-L, Lu X-Y, Kai H-L, Han Y, Zheng Y-L. Integrative analysis from multi-center studies identifies a weighted gene co-expression network analysis-based Tregs signature in ovarian cancer. *Environmental Toxicology*. 2023;1-15. doi:[10.1002/tox.23948](https://doi.org/10.1002/tox.23948)