

A five-gene risk model based on carnitine-related genes for predicting prognosis in ovarian cancer

Lan Liu

Xizang University

Xinfang Liu

Xizang University

Ping Luo

Houbei Hospital of Lhasa

Xinyan Fu

Xizang University

Yanhui Ding

Xizang University

Fanfan Li

Xizang University

Yue Wu

Xizang University

Yongcang Zhang

zyc@xzu.edu.cn

Xizang University

Research Article

Keywords: Ovarian cancer, Carnitine, Biomarkers, Risk model

Posted Date: October 29th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10891496/v1>

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

[Read Full License](#)

Additional Declarations: No competing interests reported.

A five-gene risk model based on carnitine-related genes for predicting prognosis in ovarian cancer

Lan Liu^{1,2}, Xinfang Liu^{1,2}, Ping Luo³, Xinyan Fu^{1,2}, Yanhui Ding^{1,2}, Fanfan Li^{1,2}, Yue Wu^{1,2}, Yongcang Zhang^{1,2*}

1. Medical College, Xizang University, Lhasa, Xizang, People's Republic of China;
2. Lhasa Key Laboratory of Enabling Technologies for Geo-Authentic Herbal Medicine Testing Instrumentation, Xizang University, Lhasa, Xizang, People's Republic of China;
3. Houbai Hospital of Lhasa, Lhasa, Xizang, People's Republic of China

*** Correspondence:**

Corresponding Author : Yongcang Zhang, Medical College, Xizang University, Lhasa, Xizang, People's Republic of China, zyc@xzu.edu.cn

Abstract

Background: Ovarian cancer (OV) remains a highly lethal gynecological malignancy and is associated with unfavorable long-term outcomes. Carnitine participates in metabolic adaptation of tumor cells by modulating fatty acid oxidation and redox balance. This study sought to identify carnitine-associated prognostic biomarkers in OV and to characterize the molecular features linked to these markers.

Methods: Publicly available datasets were used throughout the study. Candidate biomarkers were progressively selected through weighted gene co-expression network analysis (WGCNA), differential-expression analysis, regression-based screening, and machine-learning procedures. The selected genes were incorporated into a prognostic risk model, after which patients with OV were assigned to a high-risk group (HRG) or low-risk group (LRG) according to their calculated risk scores. Functional enrichment, immune-infiltration, and drug-sensitivity analyses were subsequently conducted to characterize biological differences associated with risk status. Biomarker expression was further examined using single-cell transcriptomic data to prioritize relevant cell populations.

Results: The multistage selection procedure yielded five biomarkers: ALPL, RAB38, SLC4A8, ZMAT4, and IGLON5. Using these genes, the risk model separated patients

into the HRG and LRG, with inferior survival observed in the HRG. Across the five biomarkers, 14 shared signaling pathways were identified, among them "Apoptosis", "Hedgehog signaling pathway", and "Ribosome". Scores for 21 immune-cell populations also differed significantly between the two risk categories. Drug-response prediction identified significant between-group differences in the half-maximal inhibitory concentration values of 73 compounds; several agents, including GSK1904529A and ML323, had lower predicted IC₅₀ values in the HRG. Single-cell analysis resolved five cell types and prioritized epithelial cells as the candidate key population.

Conclusion: The five genes ALPL, RAB38, SLC4A8, ZMAT4, and IGLON5 were identified as candidate prognostic biomarkers for OV. Their expression was linked to carnitine-associated metabolic characteristics and immune features, supporting further experimental investigation and prospective evaluation of their prognostic value.

Keywords: Ovarian cancer, Carnitine, Biomarkers, Risk model

1. Introduction

OV is a biologically heterogeneous malignancy that may originate from ovarian, fallopian-tube, or peritoneal tissues. The National Cancer Institute (NCI) Dictionary of Cancer Terms broadly distinguishes epithelial ovarian carcinoma, arising from epithelial tissue, from malignant germ-cell tumors derived from oocytes. Among cancers affecting women, OV contributes disproportionately to mortality, and women older than 50 years constitute an important high-risk population [1]. Its lethality is partly attributable to the scarcity of distinctive symptoms at an early stage; consequently, almost 70% of patients present with advanced disease, often accompanied by ascites and metastatic dissemination. These clinical features mark disease progression and may also favor chemoresistance and subsequent recurrence [2]. Peritoneal dissemination is frequently accompanied by OV-associated ascites, a complex fluid containing malignant cells, immune populations, and extracellular-matrix components. Within this milieu, ascites can support tumor-cell implantation, survival, and invasion [3]. A second major route of spread is lymphatic dissemination,

which commonly reaches pelvic and para-aortic lymph nodes and can subsequently facilitate distant tumor-cell transport through the lymphatic circulation. Together, peritoneal and lymphatic metastases substantially worsen clinical outcome [4]. Current first-line management generally relies on maximal cytoreductive surgery followed by platinum-based chemotherapy, yet clinically effective strategies directed specifically at metastatic lesions remain limited. Conventional pathological assessment does not fully resolve the cellular heterogeneity of metastatic tissue, while systemic cytotoxic agents cannot selectively distinguish malignant cells from surrounding normal or immune cells, thereby contributing to treatment-related toxicity [5]. A more detailed molecular understanding of OV progression is therefore needed to identify actionable targets and ultimately improve patient outcome and prognosis.

Metabolic adaptation has increasingly been implicated in both metastatic behavior and resistance to platinum therapy in OV [6]. In advanced disease, tumor cells undergo extensive metabolic reprogramming and can become strongly reliant on mitochondrial fatty acid oxidation (FAO) to satisfy energetic demands associated with dissemination and treatment resistance [6]. The L-carnitine shuttle is integral to this process because it controls transport of fatty acids for mitochondrial oxidation and can thereby affect cellular energy balance, oxidative stress, tumor growth, and remodeling of the local microenvironment [6]. L-carnitine itself is a quaternary ammonium compound synthesized endogenously from methionine and lysine, while animal-derived foods, particularly meat, provide major dietary sources [7]. In mammalian cells, L-carnitine has been reported to limit reactive oxygen species (ROS), reduce mitochondrial injury caused by oxidative stress, and attenuate apoptosis mediated through mitochondrial pathways [8]. Its canonical metabolic role is to enable transfer of long-chain fatty acids from the cytosol into mitochondria for oxidative energy production, a function that has also been associated with protection of ovarian tissue [9]. Carnitine palmitoyltransferase 1A (CPT1A), which catalyzes the rate-limiting step of the L-carnitine transport system, is frequently overexpressed in OV and has been linked to tumor-promoting activity [10]. Mechanistically, CPT1A can regulate mitochondrial fission factor MFF and enhance mitochondrial fission, thereby promoting OV-cell

proliferation [11]. Existing research has therefore provided important information about individual components such as CPT1A, but it has not adequately defined how the wider set of carnitine-related genes relates to long-term survival or the immune microenvironment in OV. The contribution of these gene programs to disease progression, platinum resistance, and immune remodeling consequently remains insufficiently resolved and formed the basis for the present investigation. Recent evidence suggests that CPT1A-driven fatty-acid oxidation and MFF succinylation may support mitochondrial remodeling, proliferation and stemness maintenance in ovarian cancer [6, 11, 12].

Against this background, we assembled publicly available bulk-transcriptomic and single-cell transcriptomic datasets from OV. Carnitine-related gene sets were integrated with these data in a multistep analytical workflow to screen prognostically relevant biomarkers and derive corresponding risk scores. The selected biomarkers were then used to establish and externally assess a prognostic risk model. Beyond survival stratification, we examined how the biomarkers related to biological signaling programs and patterns of immune infiltration within the tumor microenvironment. Single-cell data were subsequently used to map biomarker expression across cellular populations, prioritize candidate key cells, and investigate predicted intercellular communication as well as differentiation trajectories. The purpose of this integrated analysis was to define reproducible associations between carnitine-related gene programs and OV prognosis without presuming causal mechanisms. By jointly evaluating survival, pathway activity, immune characteristics, and cell-state information, the study provides testable candidates and a structured basis for subsequent mechanistic experiments and prospective validation.

2. Materials and methods

2.1 Data acquisition

For training set 1, gene-expression profiles for TCGA-OV were obtained from The Cancer Genome Atlas (TCGA, <https://portal.gdc.cancer.gov/>, accessed May 6, 2026).

After applying the survival criterion, 353 OV tumor samples with overall survival longer than 30 days were included. Training set 2 was GSE66957 (GPL15048), retrieved from the Gene Expression Omnibus (GEO, <https://www.ncbi.nlm.nih.gov/geo/>), and comprised 57 OV tumor tissues together with 12 normal ovarian tissues. GSE63885 (GPL570) served as the validation cohort; screening retained 75 OV tumor samples with survival exceeding 30 days, and the remaining specimens were excluded. For single-cell transcriptomic analyses, GSE130000 (GPL18573) was used, with 4 primary ovarian tumor samples retained after removal of ineligible samples.

Carnitine-related gene sets were additionally queried in the Molecular Signatures Database (MSigDB, <https://www.gsea-msigdb.org/gsea/msigdb/>, accessed May 6, 2026). The search terms were GOBP_CARNITINE_BIOSYNTHETIC_PROCESS, GOBP_CARNITINE_METABOLIC_PROCESS, GOBP_CARNITINE_TRANSPORT, GOMF_CARNITINE_O_ACYLTRANSFERASE_ACTIVITY, GOMF_CARNITINE_TRANSMEMBRANE_TRANSPORTER, GOMF_QUATERNARY_AMMONIUM_GROUP_BINDING, HP_ABNORMAL_CIRCULATING_CARNITINE_CONCENTRATION and REACTOME_CARNITINE_SHUTTLE. The retrieved gene lists were combined and duplicate entries were removed, yielding 91 carnitine-related genes (CRGs) for downstream analysis (Supplementary Table 1).

2.2 Weighted gene co-expression network analysis (WGCNA)

CRG scores for samples in TCGA-OV were computed with the GSVA package (v 1.46.0) [13]. Patients were dichotomized into high- and low-CRG-score groups using the optimal cutoff. Intergroup survival disparities were assessed with the survminer package (v 0.5.0) [14], which also generated Kaplan–Meier (K-M) curves; statistical significance was determined by the log-rank test (threshold $P < 0.05$).

Using the WGCNA package (v 1.71) [15], we applied WGCNA to the TCGA-OV cohort to detect genes within CRG-associated co-expression modules. Hierarchical clustering based on Euclidean distances among gene-expression profiles was first

performed to inspect the samples and exclude potential outliers. The pickSoftThreshold function was then applied to select an appropriate soft-thresholding power for the retained samples so that the resulting gene-gene network approximated a scale-free topology. Genes were partitioned into co-expression modules via hierarchical clustering and dynamic tree cutting after network construction, provided that the scale-free topology index R^2 exceeded 0.8 and mean connectivity approached 0. The association between module eigengenes and CRG scores was evaluated with the cor function from the WGCNA package (v 1.71). The module with the strongest CRG-score association was chosen as the key module; genes within it meeting the criteria of $P < 0.05$ and $|\text{correlation coefficient (cor)}| > 0.30$ were kept as key module genes.

2.3 Acquisition of candidate genes

Bulk differentially expressed genes (DEGs) between OV and normal ovarian tissues in GSE66957 were detected with the limma package (v 3.56.2) [16], using $\text{adj.}P < 0.05$ and $|\log_2 \text{ fold change (log2FC)}| > 1.5$ as the filtering criteria. DEGs were ordered according to log2FC, and the 10 most upregulated together with the 10 most downregulated genes were selected for visualization.

We defined candidate genes as the overlap between bulk DEGs and key module genes using the ggvenn package (v 0.1.10) [17], then functionally annotated these intersecting genes with clusterProfiler (v 4.10.1) [18]. GO and KEGG enrichment results were considered significant at $P < 0.05$. For display, the five highest-count terms in each GO category and the five highest-count KEGG pathways were selected. Candidate gene products were concurrently fed into the STRING database (<https://string-db.org/>) to generate a protein-protein interaction (PPI) network, applying a cutoff of ≥ 0.15 for interaction scores.

2.4 Acquisition of biomarkers and construction of the risk model

Within TCGA-OV, candidate prognostic biomarkers were initially screened by univariate Cox regression using the survival package (v 3.8-3) [19], with $P < 0.05$ and $HR \neq 1$ required for selection. We then tested these candidates under the proportional hazards (PH) assumption using cox.zph and retained genes with PH-test $P > 0.05$ as key biomarkers. Subsequently, glmnet (v 4.1-8) [20] performed LASSO regression with

5-fold cross-validation, defining as biomarkers those genes with nonzero coefficients. These biomarkers were incorporated into the risk model, and the risk score was calculated as follows:

$$\text{risk score} = \sum_{(i=1)}^n \text{Coef}(i) \times \text{expr}(i)$$

Here, coef represents the risk coefficient assigned to each biomarker and expr denotes its corresponding expression value. The 353 TCGA-OV samples were separated into a high-risk group (HRG) and low-risk group (LRG) according to the optimal risk-score threshold. Group-specific survival was evaluated using the survminer package (v 0.5.0), and K-M curves were generated with a log-rank significance criterion of $P < 0.05$. Risk-score distributions and survival-status plots were also produced to show patient distributions across risk levels. Time-dependent receiver operating characteristic (ROC) curves at 3, 5, and 7 years were generated with the survivalROC package (v 1.0.3.1) [21], from which the area under the curve (AUC) was obtained; $\text{AUC} > 0.6$ was used as the evaluation criterion. Model validation was then performed in GSE63885 by applying the same analytical procedure.

2.5 Gene Set Enrichment Analysis (GSEA)

To profile pathway-level disparities between HRG and LRG, we first ran differential analysis on TCGA-OV expression data using DESeq2 (v 1.40.2) [22]. Genes were sorted from highest to lowest $\log_2\text{FC}$ to produce a ranked list representing risk-group disparities. GSEA was subsequently performed with clusterProfiler (v 4.10.1) using the MSigDB reference c2.cp.kegg_legacy.v2025.1.Hs.symbols; pathways achieving $|\text{normalized enrichment score (NES)}| > 1$ and $P < 0.05$ were considered significant, and the five with the largest $|\text{NES}|$ values were selected for visualization.

To explore pathways linked to individual biomarkers, we used the MSigDB reference file c2.cp.kegg_legacy.v2025.1.Hs.symbols.gmt. For each biomarker, we computed correlations with all TCGA-OV genes via the psych package (v 2.2.9) [23] and then sorted genes by descending correlation coefficients. These ranked lists were analyzed by GSEA with clusterProfiler (v 4.10.1), retaining pathways at $|\text{NES}| > 1$ and $P < 0.05$.

The five pathways with the greatest $|\text{NES}|$ values were displayed for each biomarker. Finally, ggvenn (v 0.1.10) was used to determine pathways shared among the biomarkers, and these common pathways were summarized in a heatmap.

2.6 Immune microenvironment analysis

Sample-level scores for 28 immune-cell types were calculated with the GSVA package (v 1.46.0) and the matching immune-cell signatures to quantify immune infiltration in TCGA-OV [24]. We compared scores between HRG and LRG via the Wilcoxon rank-sum test, retaining immune-cell types with $P < 0.05$. Then, using psych (v 2.2.9), we computed Spearman correlations to evaluate relationships among the selected immune populations and between these cells and the biomarkers. Correlations meeting $P < 0.05$ and $|\text{cor}| > 0.30$ were considered significant.

To examine predicted immune-evasion characteristics, Tumor Immune Dysfunction and Exclusion (TIDE) scores were obtained from the TIDE database (<http://tide.dfci.harvard.edu/>) for the two risk groups. TIDE provides a quantitative estimate related to immune dysfunction and exclusion in the tumor microenvironment (TME). TIDE-score distributions in the HRG and LRG were compared via the Wilcoxon test, with $P < 0.05$ deemed statistically significant.

Expression of the 36 immune checkpoint genes listed in Supplementary Table 2 was compared between the HRG and LRG in TCGA-OV. The Wilcoxon test was used for each comparison, and $P < 0.05$ was taken as the significance threshold.

2.7 Tumor mutational burden (TMB) analysis and drug sensitivity analysis

For TCGA-OV samples with available somatic-mutation information, mutation profiles in the HRG and LRG were summarized with the maftools package (v 2.18.0) [25]. The 20 genes with the highest mutation frequencies were shown in a waterfall plot. TMB scores were subsequently compared between the risk groups using the Wilcoxon test, where $P < 0.05$ was considered statistically significant.

Drug-response prediction was carried out in TCGA-OV for 198 conventional compounds from the GDSC database (<https://www.cancerrxgene.org/>). Their IC₅₀ values were estimated with the pRRophetic package (v 0.5) [26]. HRG and LRG were compared using the Wilcoxon test at $P < 0.05$, and the 10 drugs yielding the lowest P

values were chosen for display.

2.8 Construction of a GraphBAN model

Potential interactions between compounds and the proteins encoded by the biomarkers were screened with the Compound-Protein Interaction (CPI) prediction framework implemented in GraphBAN. Model training integrated three CPI resources: BindingDB (<https://www.bindingdb.org>), Kinase Inhibitor Binding Assay (KIBA), and Biomedical Network Dataset Collection (BioSNAP, <https://snap.stanford.edu/biodata/>). Compound structures in MOL2 format were retrieved from DrugBank (<https://www.drugbank.ca/>). Their molecular graph representations were generated from ChemBERTa features together with a Graph Convolutional Network (GCN). In parallel, amino-acid sequences for biomarker-encoded proteins were collected from UniProt (<https://www.uniprot.org>), and protein embeddings were produced using a One-Dimensional Convolutional Neural Network (1D-CNN) and Evolutionary Scale Modeling (ESM). The compound and protein features were jointly entered into GraphBAN to obtain CPI probabilities ranging from 0 to 1; values > 0.5 were interpreted as indicating potential binding. Only interaction pairs with probabilities > 0.5 in each of the three datasets, corresponding to the intersection across all datasets, were retained as candidate compounds.

Candidate compounds were next screened for druggability with ADMETlab3 (<https://admetlab3.scbdd.com/server/screening>). Molecular Weight (MW), the numbers of Hydrogen Bond Acceptors (HBA) and Hydrogen Bond Donors (HBD), and the Logarithm of the Partition Coefficient (LogP) were calculated and evaluated according to Lipinski's Rule of 5. In this analysis, $MW < 500$, $HBA < 5$, $HBD \leq 5$, and $LogP \leq 5$ were used as the stated criteria, with fewer violations interpreted as more favorable druggability. Topological Polar Surface Area (TPSA) was also assessed because of its relationship with membrane permeability; $TPSA < 140 \text{ \AA}^2$ was considered compatible with favorable cell-membrane permeability and oral absorption. Compounds passing these filters were designated key compounds. The compound with the highest predicted probability of interaction with the biomarkers was subsequently selected for molecular docking.

2.9 Molecular docking and molecular dynamics simulation

Three-dimensional structures of the key compounds were retrieved from PubChem. UniProt entries corresponding to each biomarker were queried, and available three-dimensional protein structures were obtained for docking. CB-Dock2 was used to perform molecular docking. On the basis of binding-energy rankings, the two highest-ranking biomarker-key compound pairs were advanced to molecular dynamics simulation.

Molecular dynamics (MD) simulations were run in GROMACS 2024.4 under the AMBER14SB force field. A 3-point water model was placed in a cubic simulation box, with the protein maintained 1 nm from the box boundary; ions were added to achieve overall electrical neutrality. The system underwent energy minimization by the steepest-descent algorithm, followed by sequential equilibration—first under NVT (constant particle number, volume and temperature), then under NPT (constant particle number, pressure and temperature). V-rescale temperature coupling was used during both equilibration stages with a reference temperature of 300 K. Simulations used a 2-femtosecond (fs) time step and were continued for a total of 100 nanoseconds (ns). Root-mean-square deviation (RMSD), root-mean-square fluctuation (RMSF), and total energy were calculated from the resulting trajectories.

2.10 Single-cell analysis

Quality control of GSE130000 was performed with the Seurat package (v 5.30) [23]. Cells were retained when nFeature_RNA, defined as the number of detected genes per cell, ranged from 200 to 3,000 and nCount_RNA, representing total gene counts per cell, was < 10,000. NormalizeData was used for data normalization, followed by FindVariableFeatures to select the 2,000 most highly variable genes (HVGs). The 10 leading HVGs were labeled for visualization with LabelPoints. ScaleData was then applied to scale the samples, and RunPCA was used to conduct principal component analysis (PCA) on the HVGs. Principal components (PCs) were selected with the ElbowPlot function in Seurat (v 5.30) under the stated criterion of $P < 0.05$. FindNeighbors and FindClusters were used for cell clustering, after which UMAP was performed for dimensionality reduction at a resolution of 0.4. Cell identities were

assigned using marker genes reported in the literature [27]. UMAP was used to display the annotated populations, and marker-gene expression across cell types was summarized in a bubble plot.

2.11 Identification of key cells and cell-cell communication analysis

The relative abundance of each annotated cell type in tumor samples was summarized with a histogram, and biomarker expression was examined across the different cell populations. Cell types that were both abundant and showed broad expression of multiple biomarkers were preferentially considered candidate key populations.

Cell-cell communication across all annotated populations in GSE130000 was inferred with the CellChat package (v 1.6.1) [28]. For the population designated as the key cell type, ligand-receptor pairs were further extracted when those cells acted as signal senders.

2.12 Exploration of cellular functions from multiple perspectives

Functional pathway activity across the cell types in GSE130000 was assessed using ReactomeGSA (v 1.16.1) [29]. For each pathway, the lowest enrichment score was taken as the reference, and differences between each cell-type score and this reference value were calculated. Pathways were ranked by these differences, and the 10 showing the largest differences were displayed. Cellular metabolic activity was examined separately with scMetabolism (v 0.2.1) [30], with the top 10 metabolic pathways selected for visualization.

CopyKAT (v 1.16.1) [31] was applied to GSE130000 to classify malignant cells within the candidate key-cell population. The malignant-cell assignments were visualized with UMAP. In addition, RunUMAP was used to generate a joint low-dimensional representation of malignant cells and the remaining annotated cell types.

Trajectory analysis of the key cells was conducted with Monocle (v 2.26.0) [32]. Biomarker-expression patterns were evaluated concurrently across the inferred stages of cellular differentiation.

2.13 Expression level analysis

Within GSE66957, biomarker expression was compared between tumor and normal ovarian samples using the Wilcoxon test, with $P < 0.05$ regarded as statistically

significant.

2.14 Statistical analysis

All bioinformatics analyses were conducted in R (v 4.3.3). For two-group comparisons, the Wilcoxon test was used, and statistical significance was defined as $P < 0.05$.

3. Results

3.1 Acquisition of 561 key module genes

Survival differed significantly between patients with high and low CRG scores when the optimal cutoff was set to 1.494254. The high-CRG-score group showed less favorable survival than the low-score group ($P = 0.0068$) (Fig. 1A).

Hierarchical sample clustering did not identify obvious outliers (Fig. 1B). A soft-thresholding power of 5 produced mean connectivity close to 0 (Fig. 1C). After exclusion of the unclassified grey module, 10 co-expression modules remained (Fig. 1D). Among them, the red module had the strongest inverse relationship with CRG scores ($\text{cor} = -0.42$, $P < 0.001$) and comprised 561 key module genes (Fig. 1E).

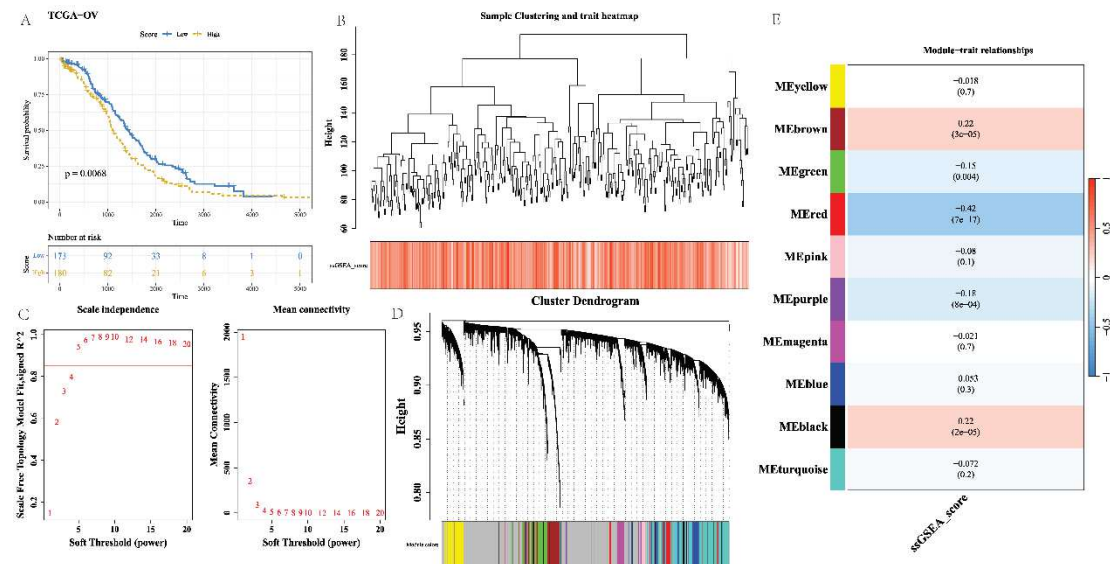


Fig. 1 Acquisition of 561 key module genes. (A) K-M survival analysis of high- and low-CRG score groups in the TCGA-OV dataset (optimal cutoff = 1.494254) ($P = 0.0068$). (B) Sample hierarchical clustering. (C) Soft threshold selection at 5. (D) Gene clustering into 10 modules. (E) Module correlation with CRG scores ($P < 0.05$ and $|\text{cor}| > 0.30$).

3.2 Acquisition of 118 candidate genes

Differential-expression analysis of GSE66957 detected 4,905 bulk DEGs, including 3,514 genes with higher expression and 1,391 with lower expression in tumor tissues (Fig. 2A-B). Intersecting these DEGs with the key module genes yielded 118 candidate genes (Fig. 2C). Functional enrichment identified 240 significant GO terms and 20 KEGG pathways. Representative GO terms included "epithelial tube morphogenesis" and "regulation of neuron projection development" (Fig. 2D, Supplementary Table 3), whereas prominent KEGG pathways included the "PI3K-Akt signaling pathway" and "ECM-receptor interaction" (Fig. 2D, Supplementary Table 4). After isolated nodes were removed from the PPI network, 112 proteins, including COL2A1 and COL4A5, remained connected (Fig. 2E).

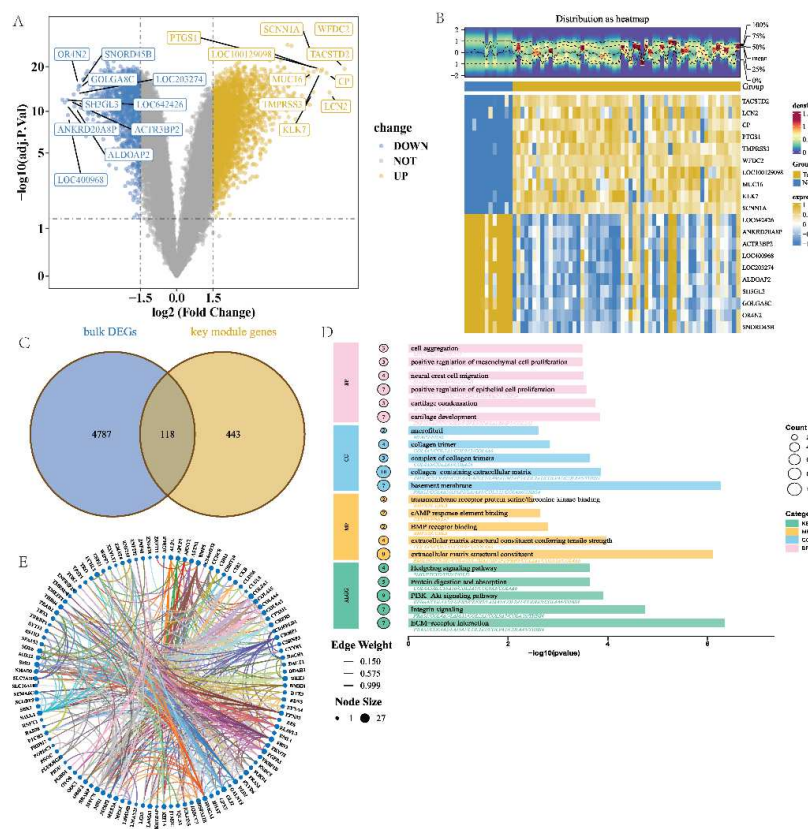


Fig. 2 Acquisition of 118 candidate genes. (A) Volcano plot of bulk DEGs (adj.P < 0.05 and |log2FC| > 1.5). (B) Heatmap of bulk DEGs (adj.P < 0.05 and |log2FC| > 1.5) (C) 118 candidate genes from the intersection of bulk DEGs and key module genes. (D)

GO and KEGG enrichment analysis of the 118 candidate genes ($P < 0.05$). (E) PPI network of proteins encoded by the 118 candidate genes (interaction score ≥ 0.15).

3.3 Development and validation of the prognostic risk model

Univariate Cox regression ($P < 0.05$; $HR \neq 1$), followed by the PH assumption test ($P > 0.05$), identified ALPL, RAB38, SLC4A8, ZMAT4, and IGLON5 as five candidate key biomarkers (Fig. 3A, Supplementary Table 5). LASSO regression retained the same five genes at $\lambda_{\min} = 0.00433297$ (Fig. 3B-C), and these biomarkers were used to construct the risk model (Supplementary Table 6). In TCGA-OV, the optimal risk-score cutoff was -0.6999897, assigning 217 samples to the HRG and 136 to the LRG. K-M analysis showed significantly poorer survival in the HRG ($P < 0.0001$) (Fig. 3F). Consistently, the risk-score and survival-status distributions showed more deaths and a higher mortality rate in the HRG than in the LRG (Fig. 3D-E). The 3-, 5-, and 7-year ROC AUCs were 0.61, 0.67, and 0.76, respectively, indicating modest discrimination at 3 years and improved discrimination at 7 years (Fig. 3G).

External evaluation in GSE63885 used an optimal risk-score cutoff of -0.4641285, which classified 56 OV samples as HRG and 19 as LRG. Applying the same analysis produced a pattern consistent with that of TCGA-OV. Survival was significantly worse for the HRG ($P = 0.0039$) (Fig. 4A), and the risk-score/survival-status distribution likewise showed greater mortality in this group (Fig. 4B). AUCs at 3, 5, and 7 years were 0.65, 0.76, and 0.73, respectively (Fig. 4C), supporting reproducibility of the risk-based stratification in the external cohort.

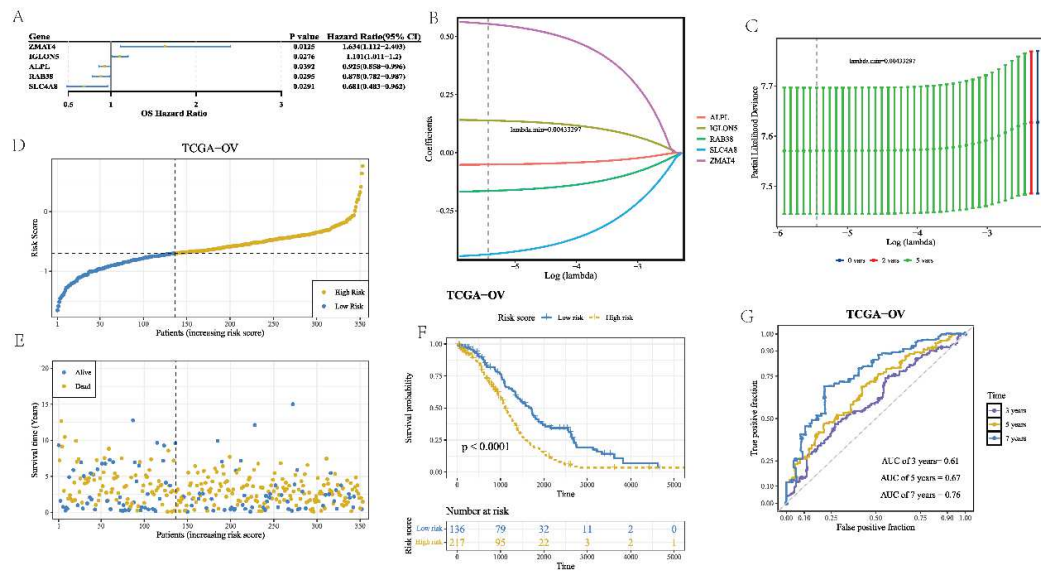


Fig. 3 Construction of the risk model. (A) Screening of candidate biomarkers via univariate Cox ($HR \neq 1$, $P < 0.05$). (B) LASSO coefficient profile for biomarker screening ($\lambda_{\min} = 0.00433297$). (C) Cross-validation plot for biomarker screening ($\lambda_{\min} = 0.00433297$). (D) Survival status of OV patients. (E) Risk curves of OV patients. (F) Comparison of survival differences between HRG and LRG ($P < 0.0001$). (G) 3-, 5-, and 7-year AUC curves of the risk model.

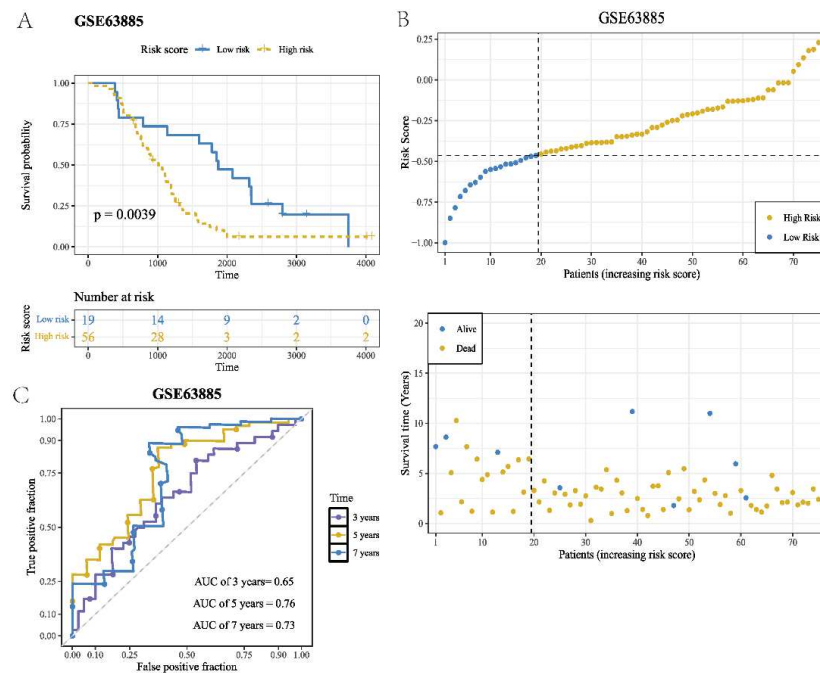


Fig. 4 Validation of the risk model in the GSE63885 dataset. (A) Comparison of survival differences between HRG and LRG ($P = 0.0039$). (B) Survival status and risk

curves of OV patients. (C) 3-, 5-, and 7-year AUC curves of the risk model for risk model validation.

3.4 Exploration of signaling pathways

Comparison of the two risk groups by GSEA identified 24 pathways enriched in common, including "Ribosome" and "Graft versus host disease" (Fig. 5A, Supplementary Table 7). The enrichment pattern indicates that these biological programs may be related to OV progression.

Biomarker-centered GSEA identified 64 enriched pathways for ALPL, including "Allograft rejection"; 77 for RAB38, including "Antigen processing and presentation"; 41 for SLC4A8, including "Ribosome"; 57 for ZMAT4, including "Antigen processing and presentation"; and 67 for IGLON5, including "Oxidative phosphorylation" (Fig. 5B-F, Supplementary Tables 8-12). Fourteen pathways were shared by all five biomarkers, among which were "Apoptosis", "Hedgehog signaling pathway", and "Ribosome" (Fig. 5G-H). These shared enrichment patterns suggest potential links between the biomarkers and signaling processes relevant to OV.

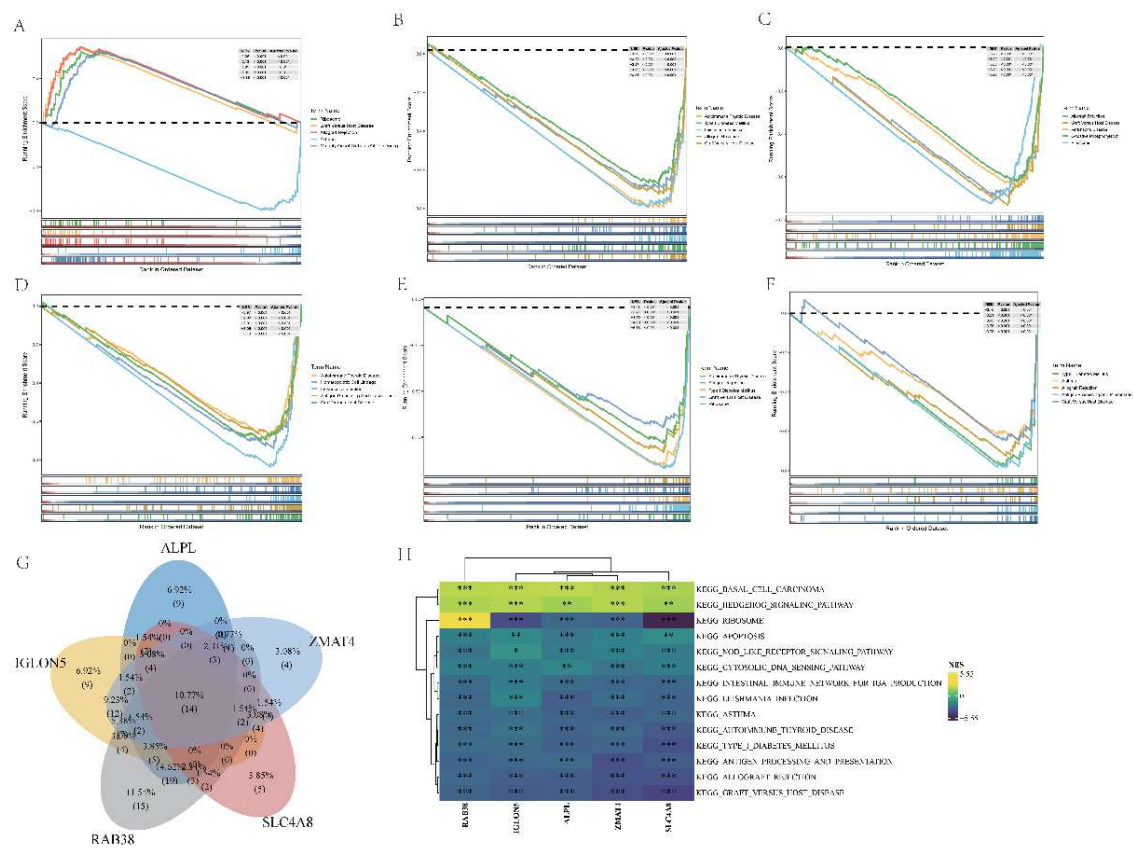


Fig. 5 Exploration of signaling pathways. (A) GSEA of HRG and LRG. (B) GSEA results of ALPL. (C) GSEA results of RAB38. (D) GSEA results of SLC4A8. (E) GSEA results of ZMAT4. (F) GSEA results of IGLON5. (G) Common enrichment of 5 biomarkers in 14 pathways. (H) Heatmap showing the activation or inhibition status of 14 different pathways. *** $P < 0.001$, ** $P < 0.01$, * $P < 0.05$

3.5 Significant association between biomarkers and immune microenvironment

Among the 28 evaluated immune-cell populations, 21 showed significantly different scores between the HRG and LRG, including activated B cells and activated CD8 T cells (Fig. 6A). The strongest positive correlations with myeloid-derived suppressor cells (MDSCs) were observed for regulatory T cells ($\text{cor} = 0.90$, $P < 0.001$) and type 1 T helper cells ($\text{cor} = 0.90$, $P < 0.001$) (Fig. 6B). Regarding biomarker-immune associations, ALPL correlated most strongly and negatively with effector memory CD8 T cells ($\text{cor} = -0.34$, $P < 0.001$), RAB38 showed its strongest positive correlation with activated dendritic cells ($\text{cor} = 0.39$, $P < 0.001$), and SLC4A8 was most positively correlated with type 1 T helper cells ($\text{cor} = 0.37$, $P < 0.001$) (Fig. 6C-E, Supplementary Table 13). These associations characterize differences in the OV immune microenvironment across the risk groups.

TIDE scores were also significantly higher in the HRG than in the LRG ($P < 0.05$) (Fig. 6F). This pattern was compatible with greater predicted immune dysfunction/exclusion in high-risk tumors and, consequently, a potentially lower predicted likelihood of benefiting from immune checkpoint blockade.

Eighteen immune checkpoint genes, including CD48, LAIR1, and CD86, were differentially expressed between the HRG and LRG (Fig. 6G). Several of these genes had higher expression in the HRG, further indicating distinct immune-regulatory profiles between the two risk categories.

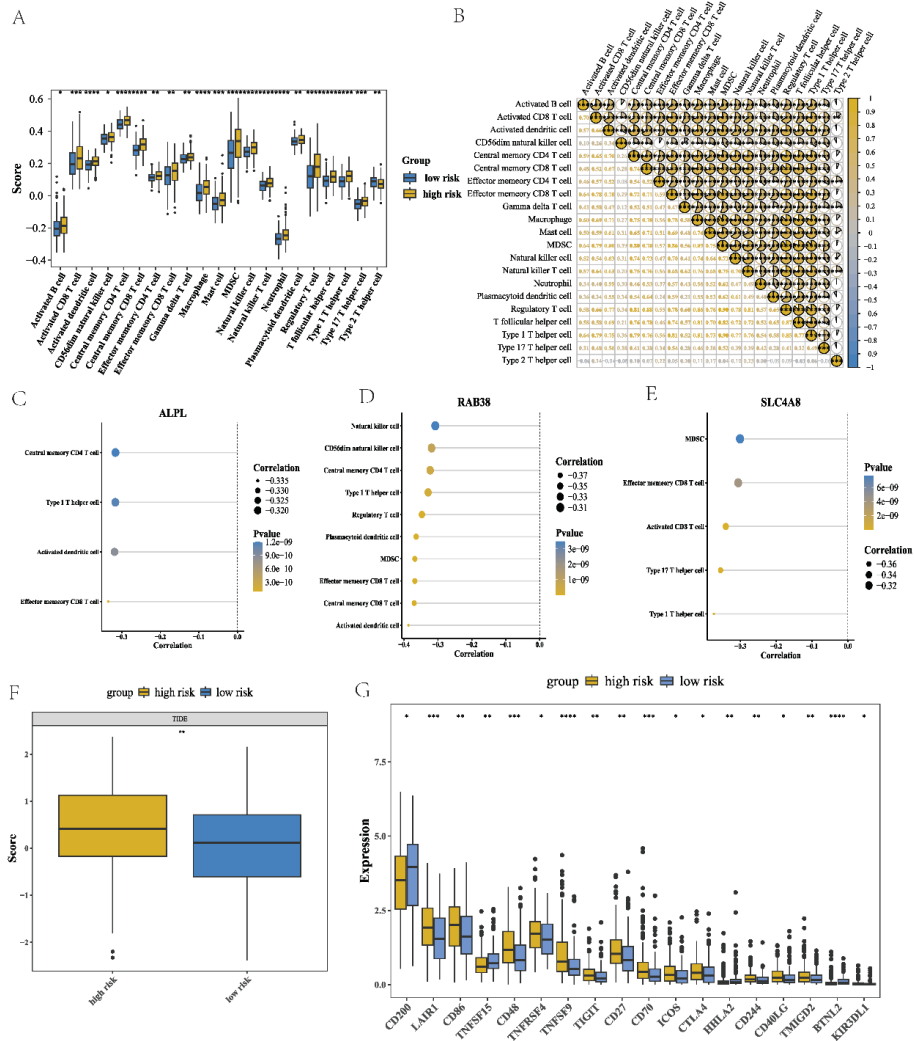
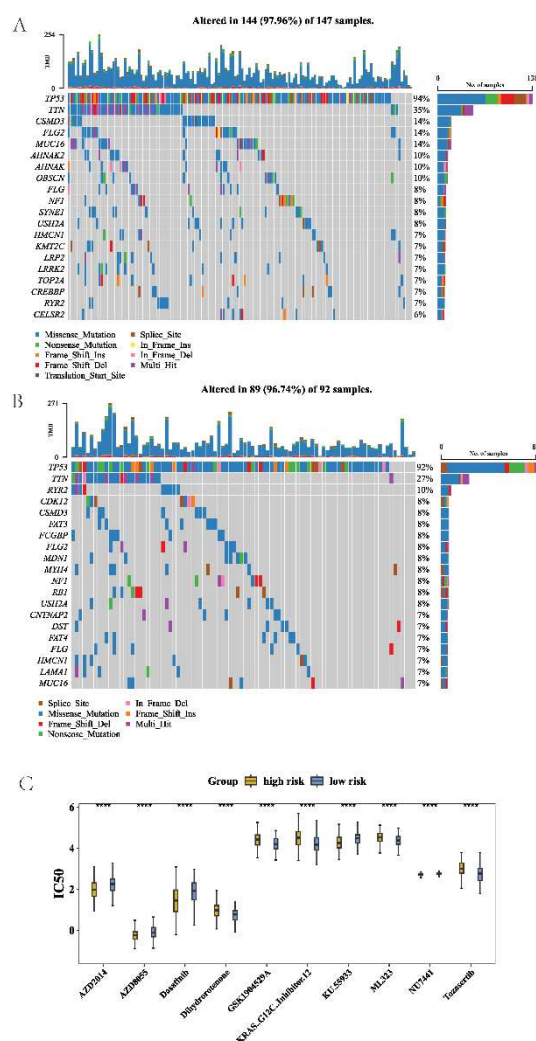


Fig. 6 Association between biomarkers and tumor immune microenvironment. (A) Box plots of scores of 28 immune cell types. (B) Significant correlations among differential immune cells ($|\text{cor}| > 0.30$ and $P < 0.05$). (C) Correlations of biomarker ALPL with differential immune cells ($|\text{cor}| > 0.30$, $P < 0.05$). (D) Correlations of biomarker RAB38 with differential immune cells ($|\text{cor}| > 0.30$, $P < 0.05$). (E) Correlations of biomarker SLC4A8 with differential immune cells ($|\text{cor}| > 0.30$, $P < 0.05$). (F) Differences in TIDE score between the HRG and LRG ($P < 0.05$). (G) Differences in immune checkpoint genes between the HRG and LRG ($P < 0.05$). **** $P < 0.0001$, *** $P < 0.001$, ** $P < 0.01$, * $P < 0.05$.

3.6 Results of TMB and drug sensitivity

Across both the HRG and LRG, TP53 was the most frequently mutated gene, and

missense mutation was the predominant mutation class (Fig. 7A-B). Although this pattern is compatible with TP53-associated genomic instability, its prognostic meaning should be interpreted cautiously.



3.7 Good binding affinity of key compounds to biomarkers

GraphBAN screening initially identified 413 candidate compounds, of which 74 remained after filtering by physicochemical properties. The five leading key compounds were (1,2,6,7-³H)-Testosterone, 17 α -methyl-4-hydroxynandrolone, Methyltestosterone, Corticosterone and Boldione (Table 1). For molecular docking, Octasulfur was paired with ZMAT4, Perfluorodecalin with SLC4A8, and Perfluorooctane with ALPL. The corresponding binding affinities were -2.6 kcal/mol for Octasulfur-ZMAT4, -9.3 kcal/mol for Perfluorodecalin-SLC4A8, and -6.7 kcal/mol for Perfluorooctane-ALPL (Fig. 8A-C). These in silico findings indicate potential biomarker-compound binding and provide candidates for further experimental evaluation in OV.

Table 1 Top 5 key compounds

SMILES	Compound	MW	nH A	nH D	TPSA	logP	Gene
<chem>[3H]C1C2=CC(=O)C([3H])C([3H])[C@]2(C)[C@H]2CC[C@]3(C)[C@@H](O)CC[C@H]3[C@@H]2C1[3H]</chem>	(1,2,6,7- ³ H)Testosterone	296.24	2	1	37.3	3.19 9018 889	ALPL
<chem>C[C@]1(O)CC[C@H]2[C@@H]3CCCC4=C(O)C(=O)CC[C@@H]4[C@H]3CC[C@@]21C</chem>	17 α -methyl-4-hydroxynandrolone	304.2	3	2	57.53	2.92 7975 401	ALPL
<chem>C[C@]1(O)CC[C@H]2[C@@H]3CCCC4=CC(=O)CC[C@@H]4[C@H]3CC[C@@]21C</chem>	Methyltestosterone	288.21	2	1	37.3	3.41 4176 497	ALPL
<chem>C[C@]12C[C@H](O)[C@H]3[C@@H](C)CC4=CC(=O)CC[C@@]43C[C@@H]1C[C@@H]2C(=O)CO</chem>	Corticosterone	346.21	4	2	74.6	1.88 1014 061	ALPL
<chem>C[C@]12C=CC(=O)C=C1CC[C@@H]1[C@@H]2CC[C@]2(C)C(=O)CC[C@@H]12</chem>	Boldione	284.18	2	0	34.14	2.31 1551 812	ALPL

Because they had comparatively favorable docking energies, the Perfluorodecalin-SLC4A8 and Perfluorooctane-ALPL complexes were selected for molecular dynamics analysis. Across the 0-100 ns simulation interval, RMSD values for the two complexes ranged from 0 to 6 nm, consistent with overall structural stability during the simulation

(Fig. 8D). RMSF profiles indicated acceptable residue flexibility while maintaining the two complexes (Fig. 8E). Total-energy trajectories showed only limited fluctuations over time (Fig. 8F). Taken together, these simulation metrics supported relatively stable binding states for both complexes.

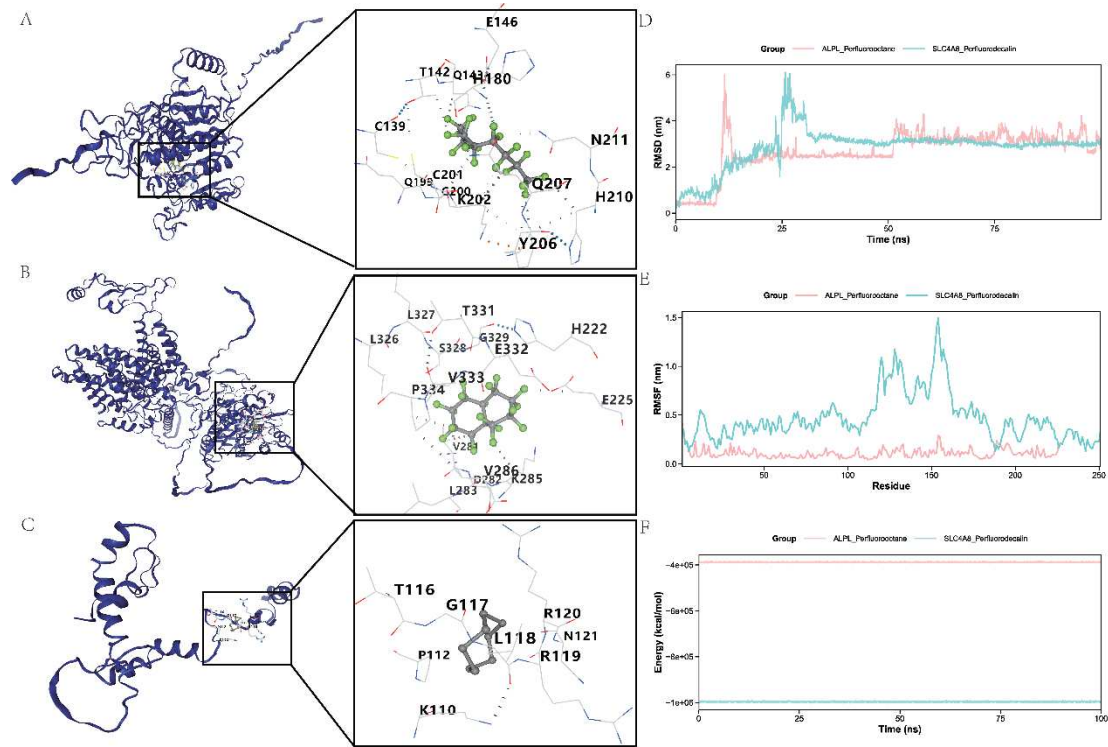


Fig. 8 Good binding affinity of key compounds to biomarkers. (A) Molecular docking conformation diagrams of biomarkers bound to Octasulfur-ZMAT4. (B) Molecular docking conformation diagrams of biomarkers bound to Perfluorodecalin-SLC4A8. (C) Molecular docking conformation diagrams of biomarkers bound to Perfluorooctane-ALPL. (D) RMSD of the 2 pairs of complexes: Perfluorodecalin-SLC4A8 and Perfluorooctane-ALPL. (E) RMSF of the 2 pairs of complexes. (F) Total energy of the 2 pairs of complexes.

3.8 Annotation of 5 cell types in OV

Before quality filtering, GSE130000 contained 13,447 cells and 18,777 genes; after quality control, 12,777 cells and the same 18,777 genes remained (Supplementary Fig. 1a). A total of 2,000 HVGs were identified, and the 10 most variable genes, including

CCL20, were labeled (Supplementary Fig. 1b). PCA was then performed, with the first 30 PCs retained for subsequent analyses (Supplementary Fig. 1c). Clustering separated the cells into 13 clusters (Fig. 9A), which were annotated into five major cell types: T cells, macrophages, cancer-associated fibroblasts, epithelial cells, and endothelial cells (Fig. 9B). Marker-gene expression across these populations was summarized with a bubble plot (Fig. 9C).

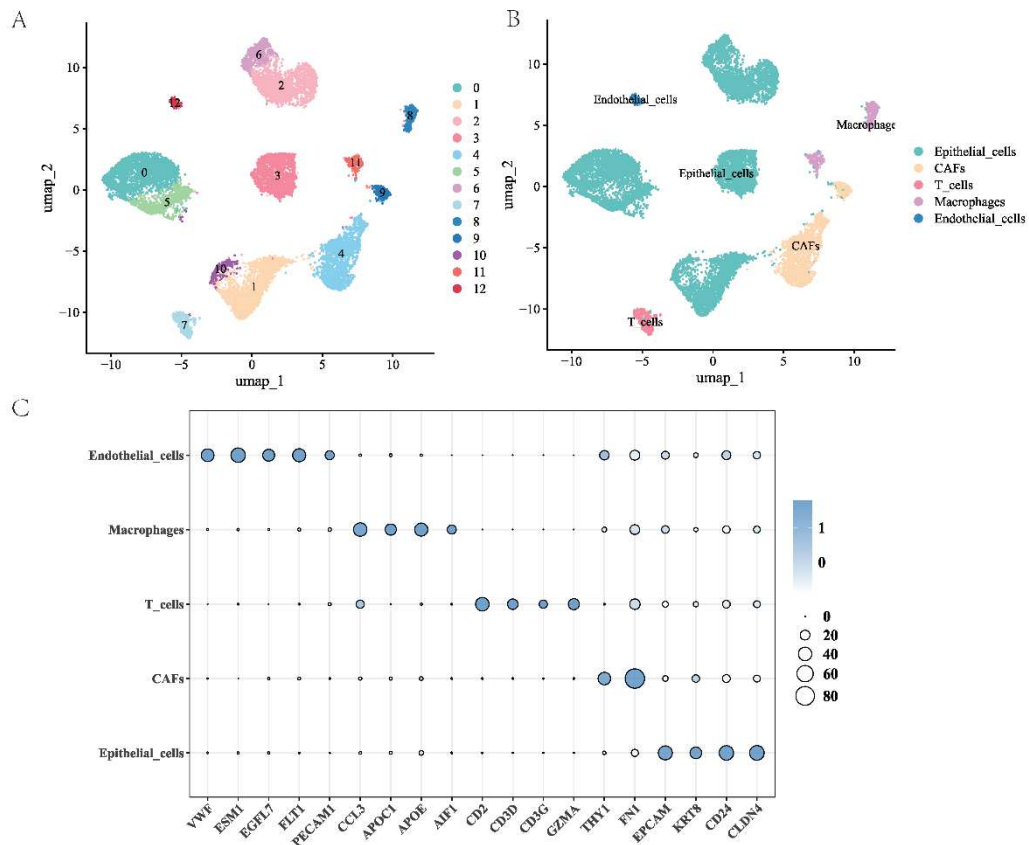


Fig. 9 Annotation of 5 cell types in OV. (A) 13 clusters of cells via clustering. (B) Annotation of cells into 5 cell types. (C) Expression of marker genes in different cells.

3.9 Epithelial cells as key cells and the cell-cell communication of key cells

Epithelial cells represented the largest cellular fraction in tumor samples (76.58%), whereas endothelial cells accounted for the smallest fraction (1.23%) (Fig. 10A). Both ALPL and RAB38 showed high expression in the epithelial-cell compartment (Fig. 10B). On the basis of abundance together with biomarker expression, epithelial cells

were therefore prioritized as the candidate key population.

CellChat predicted prominent communication between epithelial and endothelial cells in the tumor samples (Fig. 10C-D). When epithelial cells were specified as signal senders, high communication probabilities were also observed toward T cells, macrophages, cancer-associated fibroblasts, and endothelial cells, with MDK-NCL appearing as a shared ligand-receptor pair across these interactions (Fig. 10E). The predicted network therefore positions epithelial cells as a potential communication hub in the analyzed tumor microenvironment and nominates the MDK-NCL axis as a candidate mediator of crosstalk with stromal and immune populations.

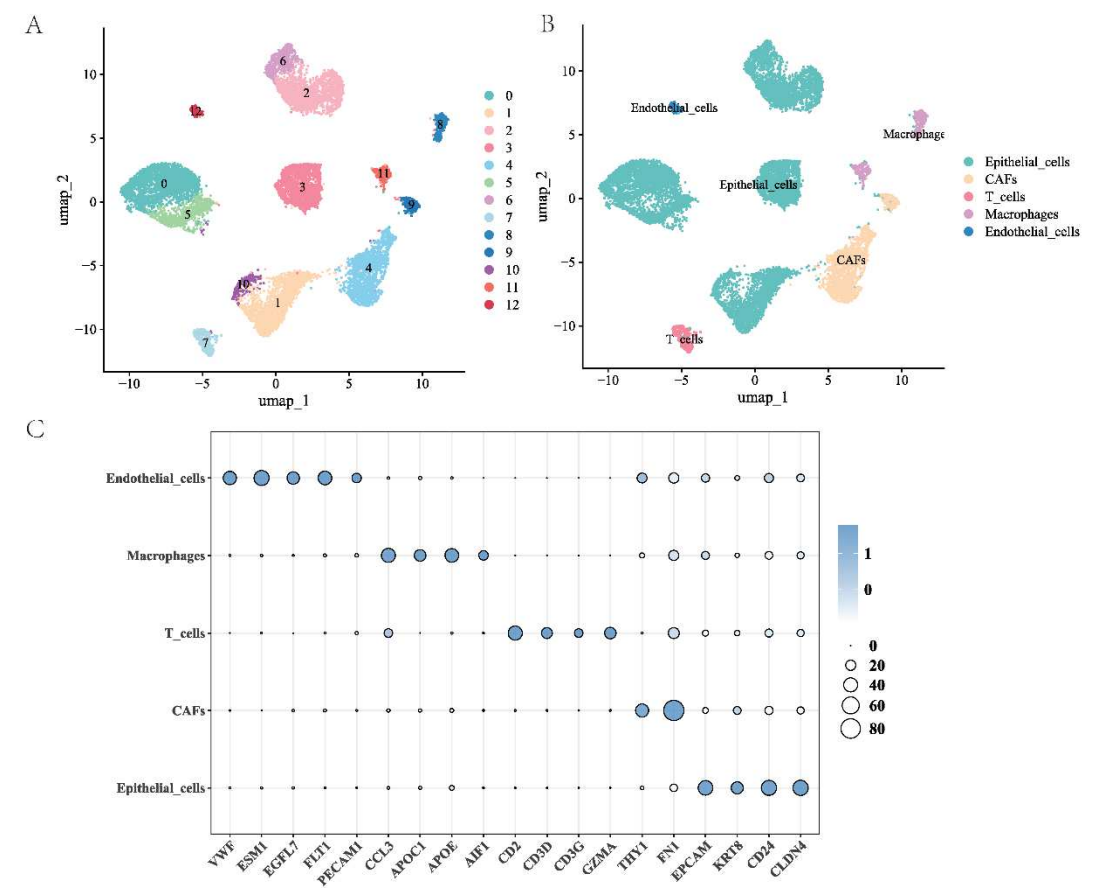


Fig. 10 Epithelial cells as key cells and the cell-cell communication of key cells. (A) Proportions of 5 cell types in the tumor group. (B) Expression of biomarkers in 5 cell types. (C) Network diagrams of intercellular communication count in tumor group. (D) Network diagrams of intercellular communication weight in tumor group. (E) Profiles of ligand-receptor pairs between the key cells and other cells.

3.10 Different functions of cells

ReactomeGSA enrichment showed that epithelial cells had high scores for pathways such as "Orexin and neuropeptides FF and QRFP bind to their respective receptors" (Fig. 11A). Metabolic analysis additionally indicated increased activity of "pyruvate metabolism", "glycolysis/gluconeogenesis", and the "citrate cycle (TCA cycle)" in the candidate key cells (Fig. 11B). These findings describe pathway and metabolic features associated with the epithelial-cell state in this dataset.

CopyKAT classified 1,363 cells within the epithelial-cell population as malignant (Fig. 11C). Their distribution relative to the remaining annotated cell types was subsequently displayed in a UMAP representation (Fig. 11D).

Pseudotime ordering indicated a trajectory of the key cells from dark-blue toward light-blue states. Malignant cells constituted a larger relative fraction at later positions along the trajectory than at the early stage. Five developmental states were defined by the color-coded trajectory, with State 1 located at the early differentiation stage (Fig. 11E). SLC4A8 expression was low early in pseudotime and increased during the middle and late phases, whereas ALPL, RAB38, ZMAT4, and IGLON5 showed comparatively higher expression early and decreased at intermediate and later stages (Fig. 11F). These patterns add a dynamic view of epithelial-cell state transitions during OV progression to the static cell-state analysis.

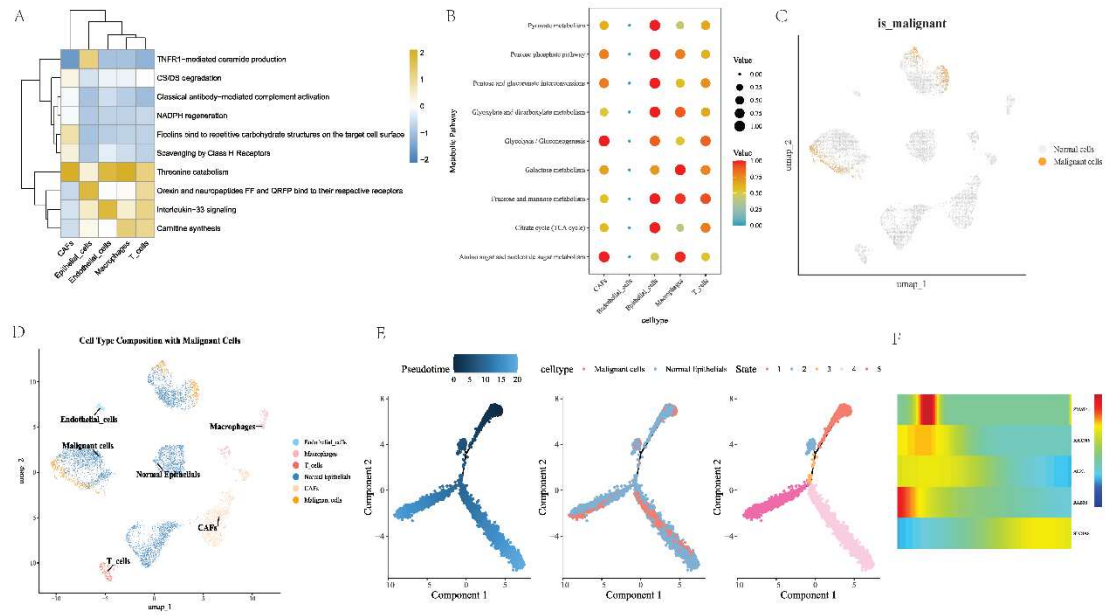
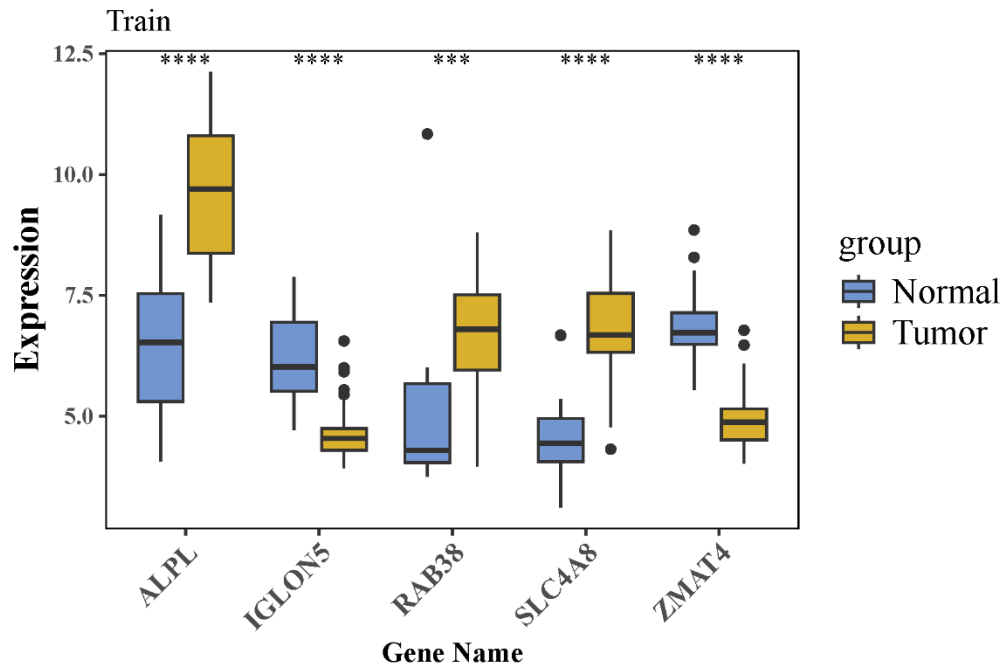


Fig. 11 Different functions of cells. (A) Functional enrichment analysis across 5 distinct cell types. (B) Activity of different metabolic pathways in 5 cell types. (C) Identification of malignant cells. (D) UMAP visualization of malignant versus other cell types. (E) Trajectory plots of differences in cell differentiation time, distribution of malignant and normal epithelial cells along cell differentiation trajectory, and state diagrams of cell differentiation showing cell differentiation states. (F) Differences in the expression of biomarkers between different differentiation stages.

3.11 Expression level of biomarkers

Relative to normal ovarian samples, tumor samples showed significantly higher expression of ALPL, RAB38, and SLC4A8, whereas ZMAT4 and IGLON5 were significantly lower ($P < 0.001$) (Fig. 12).



12

Fig. 12 Expression of the 5 Biomarkers in the GSE66957 dataset. Significant upregulation of ALPL, RAB38, SLC4A8 and downregulation of ZMAT4 and IGLON5 in tumor tissues compared with normal tissues ($P < 0.001$). **** $P < 0.0001$, *** $P < 0.001$.

Supplementary Materials

Supplementary Table 1 List of 91 CRGs.

Supplementary Table 2 List of 36 immune checkpoint genes.

Supplementary Table 3 Detailed enrichment results of GO.

Supplementary Table 4 Detailed enrichment results of KEGG.

Supplementary Table 5 Results of PH assumption test for univariate Cox regression.

Supplementary Table 6 The LASSO coefficients of 5 biomarkers.

Supplementary Table 7 GSEA enrichment result file for HRG and LRG.

Supplementary Table 8 GSEA enrichment result file for ALPL.

Supplementary Table 9 GSEA enrichment result file for RAB38.

Supplementary Table 10 GSEA enrichment result file for SLC4A8.

Supplementary Table 11 GSEA enrichment result file for ZMAT4.

Supplementary Table 12 GSEA enrichment result file for IGLON5.

Supplementary Table 13 Correlation coefficient table between biomarkers and differential immune cells.

Supplementary Fig. 1 Single-cell analysis. (A) nFeature_RNA and nCount_RNA before and after quality control. (B) Acquisition of HVGs. (C) Selection of usable PCs.

4. Discussion

L-carnitine contributes to mitochondrial fatty-acid transport and redox homeostasis; in an experimental ovarian-injury model, co-administration with spirulina was associated with SIRT1-linked antioxidant, anti-inflammatory, and anti-apoptotic effects [6, 11, 33]. Building on this metabolic context, our integrative analysis identified five prognostic biomarkers—*ALPL*, *RAB38*, *SLC4A8*, *ZMAT4*, and *IGLON5* [27, 34, 35]. The resulting model stratified overall survival in OV patients. Pathway, immune, drug-sensitivity, and single-cell analyses linked the signature to apoptosis, Hedgehog signaling, ribosome-related processes, immune-cell variation, candidate compounds, and epithelial-cell states in OV progression. Collectively, these results provide hypothesis-generating evidence for carnitine-associated biology in OV [36, 37].

Solute carrier (SLC) transporters participate in diverse cancer-associated metabolic reprogramming events to meet the elevated nutritional and energy demands of proliferating tumor cells [38]. The SLC4 superfamily contains ten members that primarily govern intracellular pH homeostasis and mediate coupled transmembrane transport of chloride/bicarbonate and sodium/bicarbonate [39]. Within the ovarian tumor microenvironment, OV cells undergo characteristic metabolic remodeling, which upregulates hydrogen ion transporters and drives extracellular acidification to boost tumor invasion and metastasis [40]. In addition, previous studies have identified genetic variants in *SLC4A8* among OV patients that are significantly correlated with OV prognosis [34], which aligns with our present results. Organic cation/carnitine transporter 2 (OCTN2, encoded by *SLC22A5*) can be secreted via exosomes, retaining Na⁺-dependent, pH-sensitive carnitine transport activity identical to membrane-bound OCTN2 [41]. Tumors rely on carnitine-dependent FAO to maintain redox homeostasis,

with OCTN2 acting as a critical transport hub [42]. In the present study, *SLC4A8* was more highly expressed in OV samples than in normal ovarian tissues, while univariate Cox regression associated it with a lower hazard of death ($HR < 1$). These results indicate that *SLC4A8* expression is associated with OV prognosis; whether it reflects a compensatory metabolic response requires direct functional and metabolomic validation. The discordance between higher tumor expression and a favorable hazard ratio may indicate context-dependent or compensatory metabolic adaptation rather than a uniformly oncogenic role [6, 36].

The Hedgehog (HH) pathway is normally constrained in most adult tissues but can be reactivated in cancer, where it supports stemness, proliferation, and tissue remodeling [11, 43-45]. Persistent HH activation has been reported in OV and associated with tumor-cell proliferation [46]. Expression of SHH, DHH, PTCH, SMO, and GLI1 has been reported to increase across benign, borderline, and malignant ovarian lesions, supporting pathway activation during malignant progression [47]. Experimental studies further suggest that SHH signaling promotes OV-cell proliferation and migration and interacts with autophagy-related processes [48]. Because lipid modifications are integral to HH signaling, pathway activation may intersect with carnitine transport and fatty-acid oxidation; however, the proposed GLI-PPAR α -OCTN2 loop requires direct validation in OV [49]. GSEA showed that *SLC4A8* expression was associated with HH-related signatures. This correlation does not establish regulatory direction; whether *SLC4A8* modifies HH activity or carnitine flux requires functional validation.

At the cellular level, type 1 helper T cells coordinate cellular immunity. They activate macrophages and support B-cell and CD8⁺ T-cell responses [50-52]. In cisplatin-resistant ovarian cancer, altered helper T-cell differentiation and reduced Th1 cytokine secretion have been associated with immune evasion and platinum resistance [53]. HH signaling can modulate T-cell receptor (TCR) signaling and downstream calcium, AP-1, NF- κ B, and ERK activation, thereby influencing T-cell activation [54]. HH activation has also been associated with a shift toward type 2 helper T-cell states, whereas pathway inhibition may enhance type 1 helper T-cell responses [44]. In our analysis, *SLC4A8* expression correlated positively with type 1 T-helper-cell scores ($cor = 0.37$, $P < 0.001$)

[44, 45].

Integrating the present multi-omics results with published evidence suggests a testable model in which *SLC4A8* may link OV progression with carnitine-associated metabolism, HH signaling, and Th1-related immunity. However, the observed expression and enrichment patterns are correlative. Whether *SLC4A8* alters intracellular pH, OCTN2-dependent carnitine uptake, or FAO requires direct metabolic assays. Likewise, the proposed effects of *SLC4A8* on HH activity and carnitine-related metabolism should be regarded as hypotheses rather than established regulatory relationships. Future studies should combine gene perturbation, metabolite tracing, and immune co-culture models to test whether *SLC4A8* has context-dependent effects on Th1-cell states and tumor progression. These experiments will be essential before therapeutic targeting is considered.

In this study, within the TCGA-OV training set, the prognosis of patients in the high-risk group was significantly worse than that of the low-risk group. The 3-, 5-, and 7-year AUCs were 0.61, 0.67, and 0.76, respectively, indicating modest discrimination at 3 years and stronger performance at later time points. External validation yielded 3-, 5-, and 7-year AUCs of 0.65, 0.76, and 0.73, supporting reproducibility but not yet establishing clinical generalizability. Direct comparisons with published models should be interpreted cautiously because cohorts, endpoints, and validation strategies differ. A previously constructed risk model for OV (PMID: 40375037) [55] reported AUCs of approximately 0.6 across short-term horizons. Another risk model for OV (PMID: 39809219) [56] reported 1-, 3-, and 5-year AUCs of approximately 0.6–0.7 in TCGA-OV. These estimates are broadly comparable but do not demonstrate superiority. Longer-term assessment may capture prognostic differences not apparent at early time points. The inclusion of 7-year survival provides a longer-term perspective. Nevertheless, the cohort-specific cutoffs and absence of prospective calibration or decision-curve analysis limit clinical interpretation. Accordingly, the model should be considered a research stratification tool; integration with clinicopathological variables and prospective validation are needed before clinical use.

ML323 is a small-molecule inhibitor of the USP1–UAF1 deubiquitinase complex

involved in DNA-damage responses [57, 58]. In ovarian cancer models, ML323 inhibited proliferation and induced S-phase arrest [58]. In this study, the HRG showed a lower predicted IC₅₀ for ML323, identifying it as a computationally prioritized compound rather than confirming treatment sensitivity. Its efficacy, selectivity, and safety in molecularly defined OV models require experimental validation. Future work could test whether ML323 responses depend on *SLC4A8*, carnitine metabolism, or HH signaling and whether treatment alters the immune microenvironment.

High-grade serous ovarian carcinoma is now understood to frequently arise from fallopian-tube secretory epithelium rather than exclusively from ovarian surface epithelium [59-61]. Single-cell studies have implicated *MYBL2* in malignant epithelial programs associated with OV progression [62]. EMT-like plasticity and multicellular spheroid formation facilitate peritoneal dissemination in OV [63]. Malignant epithelial cells also communicate with stromal and immune compartments through cytokines and matrix-remodeling signals [60]. In endometrial carcinoma, the MDK–NCL axis has been linked to epithelial–endothelial communication and an immunosuppressive microenvironment [61]. In our dataset, epithelial cells were abundant, biomarkers were expressed in this compartment, and MDK–NCL appeared among predicted epithelial–endothelial ligand–receptor pairs. These observations support, but do not prove, a role for MDK–NCL in OV progression. Validation should quantify MDK and NCL across risk groups and use spatial or perturbation experiments to determine the direction and functional consequence of this interaction.

Overall, this study identified five prognostic biomarkers—*ALPL*, *RAB38*, *SLC4A8*, *ZMAT4*, and *IGLON5*—and generated a risk model associated with survival, immune features, predicted drug response, and epithelial-cell states. These results provide a hypothesis-generating framework for carnitine-related biology in OV rather than definitive evidence of mechanism or clinical utility.

The study has several limitations. It relies on retrospective public cohorts, lacks protein-level and experimental validation, and infers regulatory relationships from transcriptomic correlations. In addition, metabolomic data were unavailable, epithelial OV subtypes were not analyzed separately, and no prospective cohort assessed

calibration or clinical utility. Therefore, the relationships between *SLC4A8* and HH/Th1 biology, as well as predicted drug responses, require confirmation using gene perturbation, metabolic-flux assays, spatial profiling, and in vivo models. Future studies should also evaluate subtype-specific performance and prospectively validate the model alongside established clinicopathological factors.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not adapted.

Availability of data and materials

The dataset (ICD -10 C56) analysed during the current study are available in the [IARC] repository, [<https://gco.iarc.who.int/en>].

Conflict of Interest

The authors declare that they have no competing interests.

Funding

This work was supported by the Lhasa Science and Technology Bureau(grant numbers LSKJ202465, LSKJ202413), and the Science and Technology Department of Xizang Autonomous Region (grant numbers XZ202401ZY0051).

Author Contributions

Lan Liu: Investigation, Methodology, Data curation, Formal analysis, Visualization, Writing - original draft preparation, Funding acquisition. **Xinfang Liu:** Investigation, Methodology, Data curation, Formal analysis, Visualization. **Ping Luo:** Investigation, Data curation, Formal analysis, Visualization. **Xinyan Fu:** Investigation, Data curation, Formal analysis, Visualization. **Yanhui Ding:** Investigation, Methodology, Data curation, Formal analysis, Visualization. **Fanfan Li:** Formal analysis, Visualization.

Yue Wu: Investigation, Formal analysis. **Yongcang Zhang:** Conceptualization, Project administration, Resources, Supervision, Writing - review & editing, Funding acquisition. All authors contributed to the article and approved the submitted version.

References

1. Thi, H.V., A.D. Ngo, and D.T. Chu, *Epigenetic regulation in ovarian cancer*. Int Rev Cell Mol Biol, 2024. **387**: p. 77-98.
2. O'Shea, A.S., *Clinical Staging of Ovarian Cancer*. Methods Mol Biol, 2022. **2424**: p. 3-10.
3. Rickard, B.P., et al., *Malignant Ascites in Ovarian Cancer: Cellular, Acellular, and Biophysical Determinants of Molecular Characteristics and Therapy Response*. Cancers (Basel), 2021. **13**(17).
4. Pal, S., et al., *Lymphatic vasculature in ovarian cancer*. Biochim Biophys Acta Rev Cancer, 2023. **1878**(5): p. 188950.
5. Smolarz, B., et al., *Ovarian Cancer-Epidemiology, Classification, Pathogenesis, Treatment, and Estrogen Receptors' Molecular Backgrounds*. Int J Mol Sci, 2025. **26**(10).
6. Liu, C., et al., *Metabolic Reprogramming Shapes the Progression and Therapeutic Landscape of Ovarian Cancer*. Cancer Manag Res, 2025. **17**: p. 1707-1722.
7. Adeva-Andany, M.M., et al., *Significance of l-carnitine for human health*. IUBMB Life, 2017. **69**(8): p. 578-594.
8. Placidi, M., et al., *Carnitines as Mitochondrial Modulators of Oocyte and Embryo Bioenergetics*. Antioxidants (Basel), 2022. **11**(4).
9. Vašková, J., et al., *The Importance of Natural Antioxidants in Female Reproduction*. Antioxidants (Basel), 2023. **12**(4).
10. Schlaepfer, I.R. and M. Joshi, *CPT1A-mediated Fat Oxidation, Mechanisms, and Therapeutic Potential*. Endocrinology, 2020. **161**(2).
11. Zhu, Y., et al., *Carnitine palmitoyltransferase 1A promotes mitochondrial fission by enhancing MFF succinylation in ovarian cancer*. Commun Biol, 2023. **6**(1): p. 618.
12. Zhu, Y., et al., *CPT1A-mediated MFF succinylation promotes stemness maintenance in ovarian cancer stem cells*. Commun Biol, 2025. **8**(1): p. 250.
13. Hänzelmann, S., R. Castelo, and J. Guinney, *GSVA: gene set variation analysis for microarray and RNA-seq data*. BMC Bioinformatics, 2013. **14**: p. 7.
14. Jia, W., et al., *Identifying immune infiltration by deep learning to assess the prognosis of patients with hepatocellular carcinoma*. J Cancer Res Clin Oncol, 2023. **149**(14): p. 12621-12635.
15. Langfelder, P. and S. Horvath, *WGCNA: an R package for weighted correlation network analysis*. BMC Bioinformatics, 2008. **9**: p. 559.
16. Ritchie, M.E., et al., *limma powers differential expression analyses for RNA-sequencing and microarray studies*. Nucleic Acids Res, 2015. **43**(7): p. e47.
17. Zheng, Y., et al., *Ferroptosis and Autophagy-Related Genes in the Pathogenesis of Ischemic Cardiomyopathy*. Front Cardiovasc Med, 2022. **9**: p. 906753.
18. Chen, H. and P.C. Boutros, *VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R*. BMC Bioinformatics, 2011. **12**: p. 35.

19. Wang, Y., et al., *Construction of an acute myeloid leukemia prognostic model based on m6A-related efferocytosis-related genes*. Front Immunol, 2023. **14**: p. 1268090.
20. Friedman, J., T. Hastie, and R. Tibshirani, *Regularization Paths for Generalized Linear Models via Coordinate Descent*. J Stat Softw, 2010. **33**(1): p. 1-22.
21. Zhang, S., et al., *Development and Validation of PET/CT-Based Nomogram for Preoperative Prediction of Lymph Node Status in Esophageal Squamous Cell Carcinoma*. Ann Surg Oncol, 2023. **30**(12): p. 7452-7460.
22. Love, M.I., W. Huber, and S. Anders, *Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2*. Genome Biol, 2014. **15**(12): p. 550.
23. Nagy, C., et al., *Single-nucleus transcriptomics of the prefrontal cortex in major depressive disorder implicates oligodendrocyte precursor cells and excitatory neurons*. Nat Neurosci, 2020. **23**(6): p. 771-781.
24. Lin, G., et al., *scRNA-seq revealed high stemness epithelial malignant cell clusters and prognostic models of lung adenocarcinoma*. Sci Rep, 2024. **14**(1): p. 3709.
25. Xu, Q., et al., *Landscape of Immune Microenvironment Under Immune Cell Infiltration Pattern in Breast Cancer*. Front Immunol, 2021. **12**: p. 711433.
26. Geeleher, P., N. Cox, and R.S. Huang, *pRRophetic: an R package for prediction of clinical chemotherapeutic response from tumor gene expression levels*. PLoS One, 2014. **9**(9): p. e107468.
27. Zhu, Y., et al., *Elucidating the role of liver enzymes as markers and regulators in ovarian cancer: a synergistic approach using Mendelian randomization, single-cell analysis, and clinical evidence*. Hum Genomics, 2024. **18**(1): p. 71.
28. Fang, Z., et al., *Integration of scRNA-Seq and Bulk RNA-Seq Reveals Molecular Characterization of the Immune Microenvironment in Acute Pancreatitis*. Biomolecules, 2022. **13**(1).
29. Wan, Z., et al., *Single-cell transcription analysis reveals the tumor origin and heterogeneity of human bilateral renal clear cell carcinoma*. Open Life Sci, 2023. **18**(1): p. 20220569.
30. Griss, J., et al., *ReactomeGSA - Efficient Multi-Omics Comparative Pathway Analysis*. Mol Cell Proteomics, 2020. **19**(12): p. 2115-2125.
31. Chen, H., et al., *Identification of metabolic signatures related to metastasis and immunotherapy resistance in oral squamous cell carcinoma*. Am J Transl Res, 2023. **15**(1): p. 373-391.
32. Du, J., et al., *Single-cell and spatial heterogeneity landscapes of mature epicardial cells*. J Pharm Anal, 2023. **13**(8): p. 894-907.
33. Embaby, E.M., et al., *Co-administration of Spirulina and L-carnitine preserves ovarian reserve in a rat model of premature ovarian insufficiency via SIRT1 regulation of oxidative stress, inflammation, and apoptosis*. Sci Rep, 2025. **15**(1): p. 32527.
34. Jiao, J., L. Jiang, and Y. Luo, *N6-Methyladenosine-Related RNA Signature Predicting the Prognosis of Ovarian Cancer*. Recent Pat Anticancer Drug Discov, 2021. **16**(3): p. 407-416.

35. Yadav, R., et al., *RNAseq-based meta-analyses revealed tumor suppressor-inducer fusion events in liver, oral, and ovarian cancer in the Indian population: a cancer cell surviving mechanism*. Nucleosides Nucleotides Nucleic Acids, 2026: p. 1-16.
36. Wang, H., et al., *Single nucleotide polymorphisms in ovarian cancer impacting lipid metabolism and prognosis: an integrated TCGA database analysis*. BMC Cancer, 2025. **25**(1): p. 462.
37. Zhang, H., et al., *A novel mitochondrial autophagy and aging-related gene signature for predicting ovarian cancer*. Front Immunol, 2025. **16**: p. 1594021.
38. Nordlund, P. and P. Reichard, *Ribonucleotide reductases*. Annu Rev Biochem, 2006. **75**: p. 681-706.
39. Qin, L., T. Li, and Y. Liu, *High SLC4A11 expression is an independent predictor for poor overall survival in grade 3/4 serous ovarian cancer*. PLoS One, 2017. **12**(11): p. e0187385.
40. Sanhueza, C., et al., *Sodium/proton exchanger isoform 1 regulates intracellular pH and cell proliferation in human ovarian cancer*. Biochim Biophys Acta Mol Basis Dis, 2017. **1863**(1): p. 81-91.
41. Console, L., et al., *Carnitine Traffic in Cells. Link With Cancer*. Front Cell Dev Biol, 2020. **8**: p. 583850.
42. Specks, J., et al., *A Single Conserved Residue Mediates Binding of the Ribonucleotide Reductase Catalytic Subunit RRM1 to RRM2 and Is Essential for Mouse Development*. Mol Cell Biol, 2015. **35**(17): p. 2910-7.
43. Jeng, K.S., C.F. Chang, and S.S. Lin, *Sonic Hedgehog Signaling in Organogenesis, Tumors, and Tumor Microenvironments*. Int J Mol Sci, 2020. **21**(3).
44. Wang, J., et al., *The emerging roles of Hedgehog signaling in tumor immune microenvironment*. Front Oncol, 2023. **13**: p. 1171418.
45. Giammona, A., E. Crivaro, and B. Stecca, *Emerging Roles of Hedgehog Signaling in Cancer Immunity*. Int J Mol Sci, 2023. **24**(2).
46. Bhattacharya, R., et al., *Role of hedgehog signaling in ovarian cancer*. Clin Cancer Res, 2008. **14**(23): p. 7659-66.
47. Chen, X., et al., *Hedgehog signal pathway is activated in ovarian carcinomas, correlating with cell proliferation: it's inhibition leads to growth suppression and apoptosis*. Cancer Sci, 2007. **98**(1): p. 68-76.
48. Pan, Y., et al., *The Sonic Hedgehog signaling pathway regulates autophagy and migration in ovarian cancer*. Cancer Med, 2021. **10**(13): p. 4510-4521.
49. Nguyen, T.D., M.E. Truong, and J.F. Reiter, *The Intimate Connection Between Lipids and Hedgehog Signaling*. Front Cell Dev Biol, 2022. **10**: p. 876815.
50. Kravtsov, D.S., et al., *Roles of CD4+ T cells as mediators of antitumor immunity*. Front Immunol, 2022. **13**: p. 972021.
51. Teixeira, F.F.C., et al., *Effects of Calcium Hydroxide Intracanal Medications on T Helper (Th1, Th2, Th9, Th17, and Tfh) and Regulatory T (Treg) Cell Cytokines in Apical Periodontitis: A CONSORT RCT*. J Endod, 2022. **48**(8): p. 975-984.
52. Ge, J., et al., *STK11 mutation affects tumor proliferation by impacting CD4(+)*

- T cell activity in lung adenocarcinoma*. Cent Eur J Immunol, 2024. **49**(3): p. 320-330.
53. Chen, P., et al., *[Lumican enhanced the therapeutic effect of cisplatin-resistant ovarian cancer by inhibiting the differentiation of Th1 cells]*. Xi Bao Yu Fen Zi Mian Yi Xue Za Zhi, 2024. **40**(8): p. 687-695.
 54. Furmanski, A.L., et al., *The transcriptional activator Gli2 modulates T-cell receptor signalling through attenuation of AP-1 and NFκB activity*. J Cell Sci, 2015. **128**(11): p. 2085-95.
 55. Lei, H., et al., *GJB2 as a novel prognostic biomarker associated with immune infiltration and cuproptosis in ovarian cancer*. Apoptosis, 2025. **30**(5-6): p. 1589-1613.
 56. Chen, X., et al., *Identification and validation of a novel defined stress granule-related gene signature for predicting the prognosis of ovarian cancer via bioinformatics analysis*. Medicine (Baltimore), 2024. **103**(47): p. e40608.
 57. Liang, Q., et al., *A selective USP1-UAF1 inhibitor links deubiquitination to DNA damage responses*. Nat Chem Biol, 2014. **10**(4): p. 298-304.
 58. Song, B., et al., *ML323 suppresses the progression of ovarian cancer via regulating USP1-mediated cell cycle*. Front Genet, 2022. **13**: p. 917481.
 59. Kurman, R.J. and M. Shih Ie, *The origin and pathogenesis of epithelial ovarian cancer: a proposed unifying theory*. Am J Surg Pathol, 2010. **34**(3): p. 433-43.
 60. Chai, C., et al., *Single-cell transcriptome analysis of epithelial, immune, and stromal signatures and interactions in human ovarian cancer*. Commun Biol, 2024. **7**(1): p. 131.
 61. Yu, X., et al., *Integrating single-cell RNA-seq and spatial transcriptomics reveals MDK-NCL dependent immunosuppressive environment in endometrial carcinoma*. Front Immunol, 2023. **14**: p. 1145300.
 62. Shao, W., et al., *Single-cell RNA sequencing reveals that MYBL2 in malignant epithelial cells is involved in the development and progression of ovarian cancer*. Front Immunol, 2024. **15**: p. 1438198.
 63. Klymenko, Y., O. Kim, and M.S. Stack, *Complex Determinants of Epithelial: Mesenchymal Phenotypic Plasticity in Ovarian Cancer*. Cancers (Basel), 2017. **9**(8).

Abbreviations

Abbreviation	Full term
1D-CNN	one-dimensional convolutional neural network
ADMET	absorption, distribution, metabolism, excretion, and toxicity
ALPL	alkaline phosphatase, biomineralization associated
AUC	area under the curve
CAFs	cancer-associated fibroblasts
CPI	compound–protein interaction
CPT1A	carnitine palmitoyltransferase 1A
CRGs	carnitine-related genes
DEGs	differentially expressed genes
ECM	extracellular matrix

EMT	epithelial–mesenchymal transition
ESM	evolutionary scale modeling
FAO	fatty acid oxidation
fs	femtoseconds
GCN	graph convolutional network
GDSC	Genomics of Drug Sensitivity in Cancer
GEO	Gene Expression Omnibus
GO	Gene Ontology
GraphBAN	Graph-Based Attention-Network
GSEA	gene set enrichment analysis
HBA	hydrogen-bond acceptors
HBD	hydrogen-bond donors
HH	Hedgehog
HR	hazard ratio
HRG	high-risk group
HVGs	highly variable genes
IC ₅₀	half-maximal inhibitory concentration
KEGG	Kyoto Encyclopedia of Genes and Genomes
KIBA	Kinase Inhibitor Binding Assay
K-M	Kaplan–Meier
LASSO	least absolute shrinkage and selection operator
LogP	logarithm of the partition coefficient
LRG	low-risk group
MD	molecular dynamics
MDSCs	myeloid-derived suppressor cells
MFF	mitochondrial fission factor
MSigDB	Molecular Signatures Database
MW	molecular weight
NCI	National Cancer Institute
NES	normalized enrichment score
NPT	constant particle number, pressure, and temperature ensemble
NVT	constant particle number, volume, and temperature ensemble
OCTN2	organic cation/carnitine transporter 2
OV	ovarian cancer
PCA	principal component analysis
PCs	principal components
PH	proportional hazards
PPI	protein–protein interaction
RMSD	root-mean-square deviation
RMSF	root-mean-square fluctuation
ROC	receiver operating characteristic
ROS	reactive oxygen species
scRNA-seq	single-cell RNA sequencing
SLC	solute carrier
TCGA	The Cancer Genome Atlas
TCR	T-cell receptor
TIDE	Tumor Immune Dysfunction and Exclusion
TMB	tumor mutational burden
TME	tumor microenvironment
TPSA	topological polar surface area
UMAP	uniform manifold approximation and projection
WGCNA	weighted gene co-expression network analysis

Data Availability Statement

The data that support the findings of this study are openly available in public repositories. The genomic and clinical profiles from TCGA-OV were obtained from The Cancer Genome Atlas (TCGA) (<https://portal.gdc.cancer.gov/>). The Gene Expression Omnibus (GEO) series datasets GSE66957, GSE63885, and GSE130000 are available at NCBI GEO (<https://www.ncbi.nlm.nih.gov/geo/>). Reference gene sets were obtained from the Molecular Signatures Database (MSigDB) . Interaction networks were analyzed using the STRING database (<https://string-db.org/>). Immune evasion profiles were derived from the TIDE database (<http://tide.dfci.harvard.edu/>). Drug-response data were obtained from the GDSC database (<https://www.cancerrxgene.org/>). Finally, chemical, protein, and pharmacokinetic resources including BindingDB, KIBA, BioSNAP, DrugBank, UniProt, PubChem, and ADMETlab3 were utilized.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryTables.xls](#)
- [SupplementaryFigure1.tiff](#)