RESEARCH

Subtype-specific analysis of gene co-expression networks and immune cell profiling reveals high grade serous ovarian cancer subtype linkage to variable immune microenvironment

Kaylin M. Carey[1], Corey D. Young[2], Alexis J. Clark[1], Eric B. Dammer[3,4], Rajesh Singh[1] and James W. Lillard Jr.[1*]

Abstract

High-grade serous ovarian cancer (HGSOC) is marked by significant molecular diversity, presenting a major clinical challenge due to its aggressive nature and poor prognosis. This study aims to deepen the understanding of HGSOC by characterizing mRNA subtypes and examining their immune microenvironment (TIME) and its role in disease progression. Using transcriptomic data and an advanced computational pipeline, we investigated four mRNA subtypes: immunoreactive, differentiated, proliferative, and mesenchymal, each associated with distinct gene expression profiles and clinical behaviors. We performed differential expression analysis among mRNA subtypes using DESeq2 and conducted Weighted Gene Co-Expression Network Analysis (WGCNA) to identify co-expressed gene modules related to clinical traits, e.g., age, survival, and subtype classification. Gene Ontology (GO) analysis highlighted key pathways involved in tumor progression and immune evasion. Additionally, we utilized TIMER 2.0 to assess immune cell infiltration across different HGSOC subtypes, providing insights into the interplay between tumor immune microenvironment (TIME). Our findings show that the immunoreactive subtype, particularly the M3 module-associated network, was marked by high immune cell infiltration, including M1 ($p < 0.0001$) and M2 macrophages ($p < 0.01$), and Th1 cells ($p < 0.01$) along with LAIR-1 expression ($p = 1.63e-101$). The M18 module exhibited strong B cell signatures ($p = 6.24e-28$), along with significant FCRL5 (adj. $p = 3.09e-30$) and IRF4 (adj. $p = 3.09e-30$) coexpression. In contrast, the M5 module was significantly associated with the mesenchymal subtype, along with fibroblasts ($p < 0.0001$). The proliferative subtype was characterized by M15 module-driven cellular growth and proliferation gene expression signatures, along with significant ovarian stromal cell involvement ($p < 0.0001$). Our study reveals the complex interplay between mRNA subtypes and suggests genes contributing to molecular subtypes, underscoring the important clinical implications of mRNA subtyping in HGSOC.

Keywords High-grade serous ovarian cancer, Immune microenvironment, Transcriptomic subtypes, Gene expression, Patient outcomes

*Correspondence:
James W. Lillard Jr.
jlillard@msm.edu
Full list of author information is available at the end of the article

Carey et al. Journal of Ovarian Research   (2024) 17:240

Page 2 of 18

Background

High-grade serous ovarian cancer (HGSOC) is a highly heterogeneous disease and one of the most lethal gynecologic malignancies worldwide [1]. Ongoing research seeks to identify both diagnostic and prognostic biomarkers to better predict disease progression. Current first-line detection methods include transvaginal ultrasonography and CA-125 surveillance. Despite widespread use of CA-125 tests, its limitations are notable. Indeed, this biomarker is absent in about 20% of ovarian cancers and can be elevated in non-cancerous conditions like peritonitis, endometriosis, and even during the menstrual cycle or pregnancy [2, 3].

Molecular heterogeneity is particularly pronounced in serous and endometrioid tumors, highlighting the need for molecular subtyping to allow for more precise diagnosis and targeted treatment approaches [4, 5]. Previous studies identified four primary mRNA HGSOC subtypes, i.e., immunoreactive, differentiated, proliferative, and mesenchymal, each of which are associated with distinct biological pathways and clinical phenotypes [4–6]. Immunoreactive HGSOC subtype tumors show a strong immune response with higher numbers of tumor infiltrating lymphocytes and are associated with the best prognosis. Differentiated HGSOC subtype tumors exhibit features of epithelial differentiation, with moderate nuclear atypia and an intermediate prognosis. Proliferative tumors are characterized by rapid cell division and high expression of proliferation markers, leading to aggressive disease progression. Lastly, mesenchymal HCSOC tumors display a loss of epithelial characteristics and gain of mesenchymal features, often associated with increased invasion and the worst prognosis. These subtypes no doubt differ in their gene expression profiles, histological features, and clinical outcomes, highlighting the need for personalized treatment approaches [5, 7] While these subtypes have improved our understanding of HCSOC, the underlying molecular drivers of these signatures and their therapeutic implications, are not yet fully understood.

Immunotherapy has emerged as a promising treatment for various cancers [8]. Previous work from our laboratory was the first to demonstrate that CXCL13-CXCR5 and CCL25-CCR9 signaling promotes several cancer types including ovarian, breast, lung, and prostate cell growth and survival [9–11]. We also showed that this signaling pathway activates PI3K, AKT, ERK, and Jun, further driving cancer progression [12]. However, significant gaps remain in understanding how the gene networks in HGSOC influence immune cell activation, infiltration within the tumor microenvironment, and attributes of the four molecular subtypes of this cancer [13].

To advance our understanding of HGSOC, innovative computational methods are being applied to study its molecular landscape. Our study aimed to evaluate subtypes using transcriptomic data, focusing on the immune cell composition and its relationship to functional networks within the Tumor Immune Microenvironment (TIME). We identified 23 modules of co-expressed genes and explored their relationships with clinical traits, such as age of onset, survival, and mRNA subtype. While previous studies have established subtype-specific gene expression patterns, our findings reveal novel transcriptional network shifts and highlight specific immune cells driving these gene modules, which has important implications for understanding HGSOC biology.

Notably, the M3 module was strongly correlated with the immunoreactive subtype and exhibited heightened immune activity, driven primarily by M1 and M2 macrophages and Th1 cells. The differentiated subtype was associated with cellular differentiation, particularly through the M15 module, where DLK1—a key regulator of organ development and stem cell maintenance—plays an important role [14, 15]. The proliferative subtype showed strong correlation with the M6 module, which was associated with gene signatures involved in tumor growth and transcriptional regulation. Lastly, the mesenchymal subtype was significantly associated with the M5 module which displayed tumor invasion and extracellular matrix remodeling as well as cancer-associated fibroblasts (CAFs) molecular signatures. These insights suggest the potential for developing subtype-specific diagnostic and therapeutic markers to improve disease monitoring by targeting the unique molecular pathways within each HGSOC subtype.

Methods

Data collection and sample removal

The data used in this study is available at [The Cancer Genome Atlas (TCGA) via GDC (Genomic Data Commons)] at [https://portal.gdc.cancer.gov], last accessed on February 27, 2024. Primary tumor samples were collected from patients with ovarian serous cystadenocarcinoma. RNA sequencing STAR counts, processed using the mRNA pipeline [5], were obtained for 417 tumor samples and 60,660 genes. Recurrent tumor samples were removed, resulting in 398 primary tumor samples. Singleton batch cases were removed ($n = 2$), due to two unique site distributions, explained here (https://docs.gdc.cancer.gov/Encyclopedia/pages/TCGA_Barcode/). Data that contained 'NA' for mRNA subtypes were also removed ($n = 2$). Genes with row sums ≥ 32 normalized counts were removed to enhance the power and accuracy of differential expression detection resulting in 394 samples and 22,181 genes being used for DESeq2 analysis.

Carey *et al. Journal of Ovarian Research*   (2024) 17:240

Page 3 of 18

After variance stabilizing transformation (VST) normalization, samples with a bicor adjacency sample network connectivity that exceeded three standard deviations above the mean using the WGCNA R package, fundamentalNetworkConcepts function [16], in any of five rounds of successive outlier removal, were deemed outliers and removed from the cohort ($n=5$). The remaining 389 profiles were used to determine pairwise correlations between genes for the WGCNA connectivity network. PCA plots were compared for data before and after outlier removal steps. This normalized TCGA-OV cohort sample-gene transcript matrix was used to perform further downstream analysis, diagnostics, and visualization.

Differential expression (DESeq2) analysis

Raw counts from the above-described pipeline were used with DESeq2 (v. 1.42.0) in R to determine differential expression based on the negative binomial distribution method [17]. Filtering criteria were set with the smallestGroupSize as 50, and subset rows were removed with row means $> =32$, determined graphically by assessing the distribution of the dataset using a histogram. Our design formula (design $= \sim$ batch + condition) was used to explore gene expression changes of mRNA subtypes with batch effects factored as a confounding observation. Contrasts were specified using reference level ranking (immunoreactive, differentiated, proliferative, mesenchymal). The ranking was used to simulate 'best to worst' clinical survival outcomes [18], with the lowest rank (best prognosis) as the base level for comparison. Subsequently, we re-leveled the mRNA subtypes while preserving the original reference level (denominator) as much as possible. In this context, the denominator represents the baseline mRNA subtype that the other subtypes are compared. This strategy ensured that each mRNA subtype was properly accounted for in the denominator, facilitating the validation of clinical findings and enhancing our understanding of differential expression.

We used the Wald test to reveal genes with differential expression between the conditions with a student's p-value threshold of ≤ 0.05 and an adjusted p-value using a weighted Benjamini-Hochberg (BH) adjustment for multiple comparisons. Independent Hypothesis Weighting (IHW) weights were considered to control the false discovery rate [19]. Traditional methods, such as the BH method, assume independence amongst tests, which could lead to inflated FDR scores. However, the IHW method provides more accuracy for large-scale testing with many hypotheses. Log$_2$ fold-change (LFC) shrinkage estimates were determined for the immunoreactive base reference level using the apeglm method to improve on previous estimators for effect size shrinkage for visualization and ranking of genes, with little differences shown between standard IHW methods [20]. A LFC threshold of > -1 (upregulation) and LFC < -1 (downregulation) was set for both methods and all comparisons.

Weighted gene co-expression network construction and module detection

We employed the WGCNA R package v. 1.72–5 [21] clustering method to construct gene co-expression networks from transcriptomic data obtained from high-grade serous ovarian adenocarcinoma (HGS-OvCa) samples ($n=389$). WGCNA distinguishes gene clusters by calculating the 1-topology overlap matrix and a gene dissimilarity matrix across samples and creates gene clusters with similar expression patterns into modules. We performed a biweight mid correlation (bicor) matrix calculation before proceeding to build the network. The correlation matrix was transformed to create an adjacency matrix utilizing a soft-threshold power of 8 (at which there was a scale-free fit R2of 0.907 and mean connectivity of 153), and the R2 approach towards an asymptote was confirmed graphically. The blockwiseModules function of the WGCNA R package further leveraged the following parameters: power$=8$; deepSplit$=4$; mergeCutHeight$=0.15$; minModuleSize$=100$; TOMdenom$=$"mean"; corType$=$"bicor"; networkType$=$"signed"; pamStage$=$TRUE; pamRespectsDendro$=$TRUE; verbose$=3$; saveTOMs$=$FALSE; maxBlockSize$=30,000$; reassignThresh$=0.05$. After adjacency calculation, the adjacency matrix is transformed into Topological Overlap Matrix (TOM), and then 1-TOM, which represents measures of protein pair dissimilarity, upon which hierarchical clustering with dynamic tree cutting is performed, grouping highly interconnected genes into modules ($n=23$) with a minimum module size set by the minModuleSize parameter. After dendrogram-based module assignments of all gene products are completed, module eigengenes (ME) are initially calculated as the first principal component of the genes within a module, and then bicor correlation of each gene product to each ME (kME) is calculated to determine if any gene product should be reassigned to a module within which it has significantly higher kME are to be merged according to a p value for significance of correlation $p < 1e$-6. After reassignments, MEs (Additional Table 1; ME values with module size, number and patient statistics for each ME) and kMEs were finalized (Additional Table 3), as a record of module membership and granular relatedness. There were 7,810 genes that failed to remain in a module of minimum size ($n=100$), relegated to the grey module of proteins for which ME calculation and kMEs are meaningless.

Following module definition, MEs (hereafter, modules) were correlated to each trait of interest encoded numerically as a continuous or binary variable to gauge

Carey *et al. Journal of Ovarian Research*   (2024) 17:240

Page 4 of 18

Table 1 Ontology enrichment analysis of WGCNA modules of interest

Module	ANOVA *p*	Top Biological processes	GO- Elite FET FDR
Red (M6)	7.53 e -62	Animal Organ Morphogenesis	0.000108
		Nervous System Process	0.000119
		Action Potential	0.000171
Pink (M8)	5.10 e -46	System Development	3.05e-28
		Multicellular Organism Development	2.32e-23
		Anatomical Structure Development	3.85e-21
Brown (M3)	2.20 e -77	Immune System Process	2.01e-223
		Immune Response	9.80e-210
		Defense Response	6.50e-166
Green (M5)	1.88 e -78	Multicellular Organismal Process	7.32e-69
		System Development	9.31e-67
		Circulatory System Development	2.77e-63
Tan (M12)	8.89 e -20	No significant biological processes were identified	NA
Midnight blue (M15)	2.75 e -13	Multicellular Organismal Process	0.00128
		Regulation Of Hormone Levels	0.00731
		Developmental Process	0.01086
Grey70 (M17)	1.73 e -12	No significant biological processes were identified	NA
Light green (M18)	6.24 e -28	Adaptive Immune Response	1.70 e-163
		Immune Response	1.04e-121
		Immune System Process	4.70 e-107

A Fisher Exact Test (one-tailed) FDC correction was used to detect overlap significance amongst GO lists of gene symbols and module membership for enrichment *p* < 0.05 and 5 genes minimum per ontology

the relevance of modules to these traits and vice versa. Bicor correlation was used instead of Pearson correlation, as defined by [22] to provide robustness to outliers; bicor rho and associated Student *p* values are ideal for determining robust correlation. Statistical significance module correlation to clinical traits (mRNA subtype, age, and survival) was determined by Student's *p*-value < 0.05.

Functional enrichment analysis and network visualization
We used an open-source R function GOparallel (https://www.github.com/edammer/GOparallel/) to find ontologies (among those in the following categories: [1] biological processes, and [2] molecular functions), which were significantly enriched in co-expression module gene product lists [23]. Briefly, this function leverages the R piano package to perform a one-tailed Fisher's exact test, which we modified to output signed Z score for either enrichment or depletion, as well as *p* value and Benjamini–Hochberg FDR for the enrichment significance. Further, the function scrapes the website of the Bader lab to download the above categories of ontologies which are updated on a monthly basis [24] for the species of interest (human).

Intramodular core connectivity was visualized with iGraphs using with the open source buildIGraphs function (https://www.github.com/edammer/netOps). The function further leverages the BioGRID database to highlight protein–protein interactions as emphasized edges among up to the top 100 nodes of a module sorted by decreasing intramodular kME. The top 10 genes ranked by intramodular kME are placed in the iGraph as centralized hub nodes. STRING v12.0 [24] was used to depict multiple integrated databases housing information on the protein interactions, generalizable coexpression, and other connectivities already established between the top hubs of selected modules.

TIME cell composition estimation and proportion regulation
We used cell estimation methods to estimate levels of non-cancerous, tumor immune and associated cell types present in the TIME. EPIC, Estimated Proportion of Immune Cell (https://epic.gfellerlab.org) (Reference the paper introducing this resource), method was employed to estimate cell types present in the bulk sample mRNA data (*n* = 389). Compared to this method, the one-tailed Fisher Exact Test was used to determine the possibility of observing an overlap between genes present in the WGCNA modules and genes associated with immune cell types, defined by immune cell markers. Our gene list consisted of 16 immune cell types, i.e., dendritic cells, B cells, M1 and M2 macrophages, fibroblasts, ovarian

Carey *et al. Journal of Ovarian Research*   (2024) 17:240

Page 5 of 18

cancer cells, ovarian stromal cells, endothelial cells, lymphatic endothelial cells, and cancer-associated adipocytes (Additional Table 5). TIMER 2.0 (http://timer.cistrome.org) analysis explored the association between mRNA expression and immune and stromal cell populations in HGSOC for ovarian serous cystadenocarcinoma patients ($n = 303$), based on adapted methods [25].

Results

Many studies have demonstrated a significant impact of mRNA subtypes on the clinical prognosis of HGSOC [5, 26, 27]. Our approach uniquely integrates weighted gene co-expression network analysis (WGCNA) and computational immune cell profiling to uncover the underlying network connectivity and immune microenvironment dynamics, providing a more comprehensive view of disease progression and potential therapeutic targets. The profiling success and accuracy in the immune profiling methods rely on precise statistical modeling of gene expression (TIMER's deconvolution using linear regression) and the rigorous association testing of immune cell types with HGSOC subtypes (Fisher's exact test using contingency tables and hypergeometric distribution). Our findings reveal distinct gene expression patterns and immune-driven profiles across subtypes, offering

potential implications for disease monitoring and therapeutic interventions.

Subtype-specific gene signatures revealed by differential expression analysis

We performed DESeq2 analysis on 22,181 genes from 394 primary tumor samples to identify differentially expressed genes across four mRNA subtypes: differentiated ($n = 104$), immunoreactive ($n = 82$), mesenchymal ($i = 95$), and proliferative ($n = 113$). The comparison followed a reference rank based on best to worst clinical prognosis, i.e., 4–5 years for immunoreactive, 3–4 years for differentiated, 2–3 years for proliferative, and ~ 2 years mesenchymal overall survival respectively [18]. Comparisons were made between the immunoreactive subtype and each of the other subtypes (Fig. 1), highlighting log fold change (LFC) and significance. Genes with an absolute value log2 fold change greater than 1 and FDR < 0.05 were marked in red.

We observed 83 genes showed higher expression levels in the differentiated subtype than the immunoreactive subtype, while 225 genes were downregulated. The top five significant genes in the differentiated subtype included SRGAP3-AS2 (upregulated, lfc = 3.60, p.adj = 5.68e-13, weight = 1.67) and four downregulated genes, including
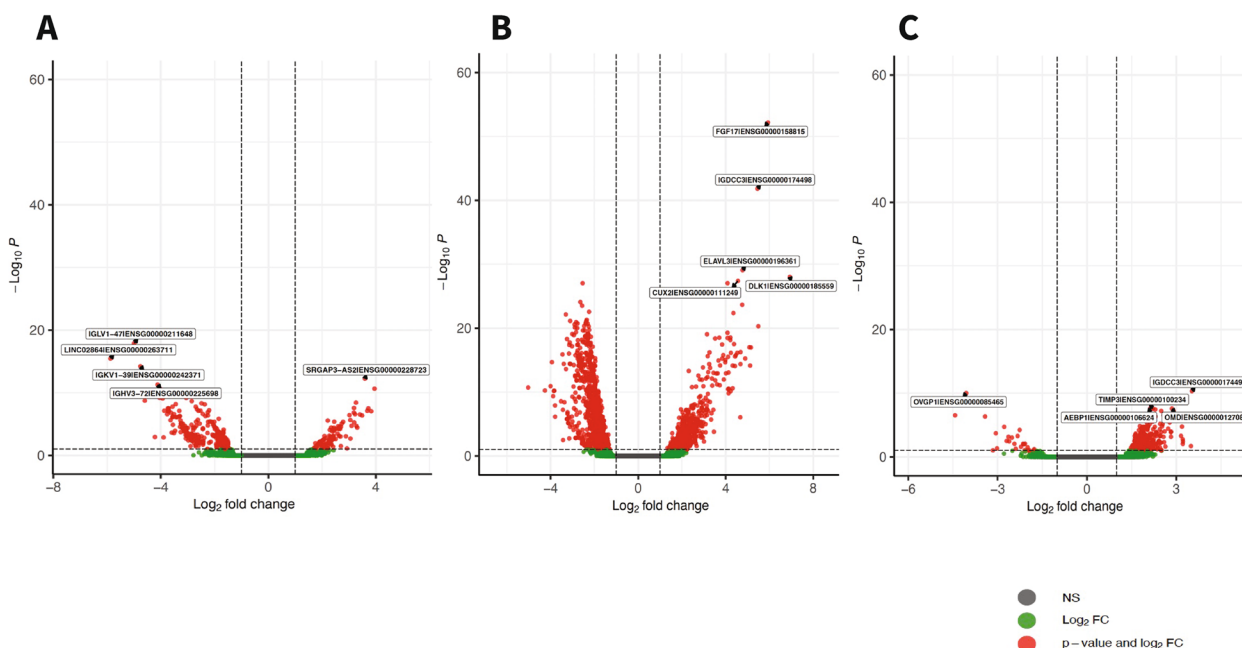

Fig. 1 Volcano Plots of DESEQ2 Comparisons ordered by best to worse clinical prognosis. These are estimates of the log2 fold changes (LFC) between conditions, (**A**) Differentiated vs. Immunoreactive (**B**) Proliferative vs. Immunoreactive and (**C**) Mesenchymal vs. Immunoreactive, for each gene with the top five genes within the criteria box. The analysis includes a total of 394 samples (Differentiated: $n = 104$, Immunoreactive: $n = 82$, Mesenchymal: $n = 95$, Proliferative: $n = 113$). The volcano plot showed gene expression changes' magnitude (log2 fold change) and statistical significance (-log10 adjusted *p*-value). Genes that are significantly differentially expressed having a log2 fold change greater than 1 and FDR < 0.05, BH corrected colored in red. Genes beyond the log fold change restriction but not FDR significant are in green, and grey genes are not significantly changed

Carey *et al. Journal of Ovarian Research*   (2024) 17:240

Page 6 of 18

IGLV1-47 (lfc = -4.99, p.adj = 1.28e-18, weight = 0.263) and LINC02864 (lfc = -5.88, p.adj = 3.49e-16, weight = 2.46). The differentiated subtype demonstrated fewer immune-related activities than the immunoreactive subtype. The immunoreactive subtype was marked by significant upregulation of immune activation, particularly of immunoglobulin genes (Additional Table 2). There were 455 upregulated and 661 downregulated genes observed in the proliferative subtype, as compared to the immunoreactive subtype. Notably, the proliferative subtype showed the most significant gene dysregulation. FGF17 (lfc = 5.93, p.adj = 6.67e-53, weight = 1.60) and DLK1 (lfc = 6.93, p.adj = 1.00e-28, weight = 0.77) were among the top upregulated genes, highlighting the proliferative subtype's emphasis on maintaining an undifferentiated state with the capacity for rapid tumor expansion. We observed 245 upregulated and 18 downregulated genes in the mesenchymal subtype, when compared to the immunoreactive subtype. Top upregulated genes included IGDCC3 (lfc = 3.52, p.adj = 4.68e-11, weight = 1.44) and AEBP1 (lfc = 2.15, p.adj = 2.59e-8, weight = 0.37), Downregulated genes included OVGP1 (lfc = -4.06, p.adj = 9.53e-11, weight = 0.58).

In comparison to other studies that have investigated differentially expressed genes (DEGs) in HGSOC, several of the genes that were highlighted in our analysis –such as FGF17, DLK1 and SRGAP3-AS2–have been previously reported, validating their role in HGSOC progression and immune evasion [28]. However, LINC02864 (lfc = -5.88, p.adj = 3.49e-16, weight = 2.46) stands out as a novel finding, not previously associated with HGSOC mRNA subtypes, showcasing a unique role in the biology of the differentiated subtype. These results provide both confirmation of established DEGs and offer new insights into molecular signatures that could influence future targeted strategies.

Co-expression network analysis reveals modules associated with mRNA subtypes

We employed WGCNA to evaluate highly correlated gene clusters across clinical conditions in a signed bicor network. This analysis identified 23 distinct modules ($p < 0.05$), associated with different mRNA subtypes. The eigengene network highlights distinct module-trait correlations (Fig. 2). The M3 module was strongly positively correlated with the immunoreactive subtype ($p = 1e-28$) and negatively correlated with the proliferative subtype ($p = 2e-54$), indicating a highly coordinated gene expression profile distinct from other subtypes. Hub genes in the M3 module—such as *LCP2* (kME = 0.965), *CD53* (kME = 0.956), and *LAIR1* (0.945)—were significantly enriched for immune response pathways, including "Immune System Process" and "Immune Response" shown in Table 1 (GO:0002376, GO:0006955).

Interestingly, M18 module followed a similar immune-related expression signature as M3, illustrated by the close node branching (Fig. 2). However, the differentiated subtype displayed negative co-expression within M18 ($p = 4e-08$), suggesting subtype-specific gene regulation. GO analysis of M18 revealed enrichment in pathways such as "Adaptive Immune Response" and "Immune Response" (GO:0002250, GO:0006955), further emphasizing its immune-related nature. The differentiated subtype ($p = 1e-07$) was positively correlated with the M17 module, which had the strongest association with diagnostic age in women aged 45–55 ($n = 108$) and moderate correlation in women under 45 ($n = 32$).

The M6 module was strongly correlated with the proliferative subtype ($p = 2e-56$) and included hub genes *SALL2* (kME = 0.729) and *TCF7L1* (kME = 0.700), both of which regulate cell growth and proliferation. GO analysis of M6 identified significant enrichment in "Wnt signaling" and "cell cycle regulation" (GO:0016055, GO:0007049), supporting its role in driving the proliferative subtype's tumorigenic potential. Notably, upregulated genes such as *FGF17*, *IGDCC3*, and *ELAVL3* in the proliferative vs. immunoreactive subtype comparison also exhibited membership in M6, contributing to the proliferative HCSOC subtype. The M6 module was also positively associated with age at diagnosis (> 55 years, $p = 0.002$), suggesting an age-related influence on this gene network. M15 was also noteworthy, exhibiting a moderate but significant co-expression with the proliferative subtype ($p = 3e-11$). GO analysis implicated

(See figure on next page.)

Fig. 2 Co-expression network of module eigengenes and clinical traits within TCGA-OV cohort. Twenty-three network coexpression module eigengenes (MEs) (ranking from largest number of module members to smallest) were identified: M1 to M23 for 389 patient samples. **A** Module Trait Relationships Heatmap; The correlation between the summary of gene expression profiles, module eigengenes (rows) and clinical traits (columns) is shown. The cells in the heatmap contain the bicor correlation coefficient with red denoting a positive correlation and blue a negative correlation. Strength of association is measured in color intensity. *P* values relate to significance of correlation between module eigengenes and clinical traits. **B** Hierarchical clustering of module eigengenes depicts branches within the dendrogram, grouping eigengenes with positive correlations. Row and columns are paired within the heatmap, correlating to one module eigengene labeled by color. The blue represents low adjacency, vindictive of a negative correlation, while the red shows high adjacency or a positive correlation. The meta-modules are depicted through squares of red color placed diagonally
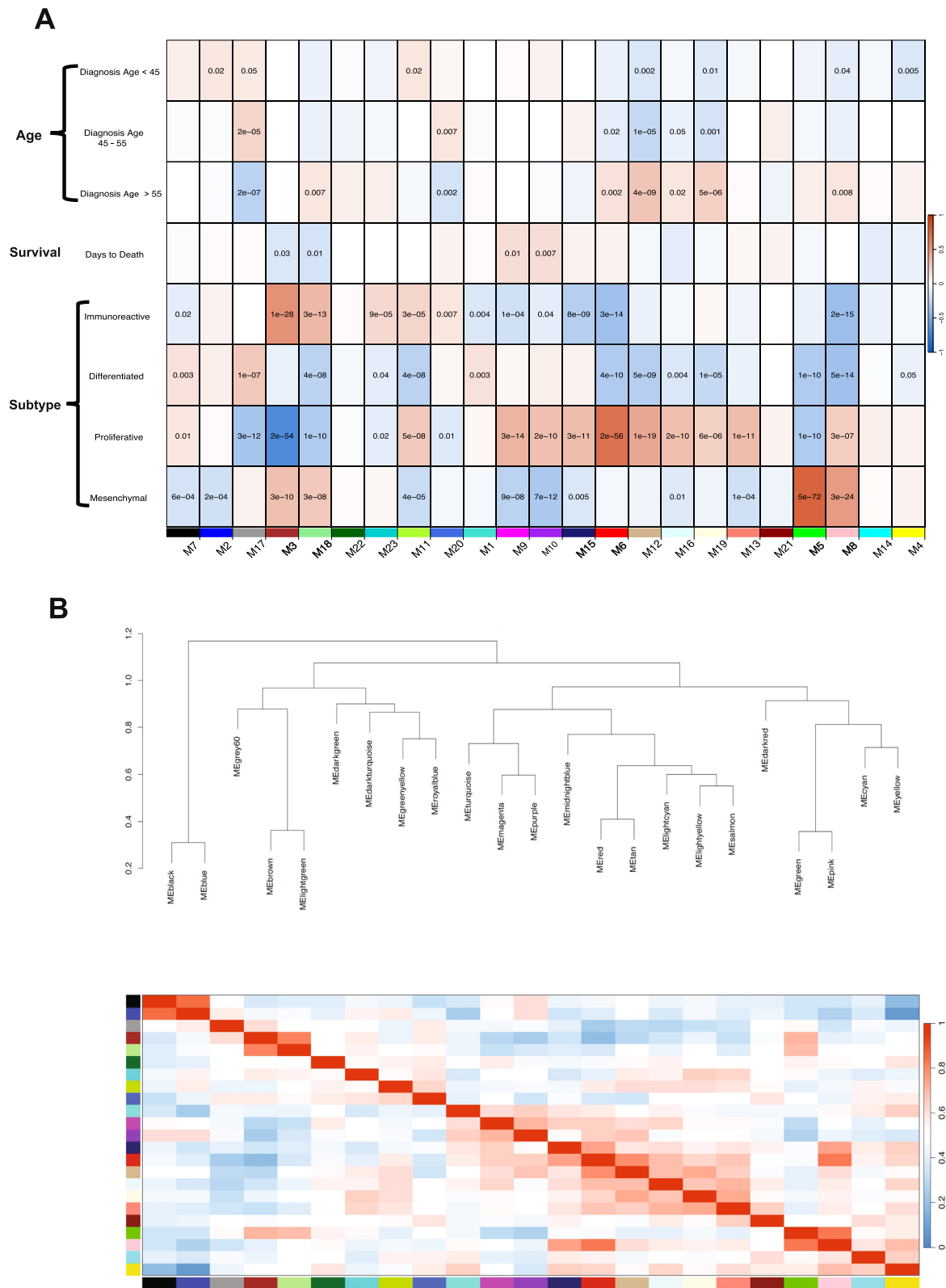
Fig. 2 (See legend on previous page.)

Carey *et al. Journal of Ovarian Research*   (2024) 17:240

Page 8 of 18

processes related to multicellular organismal processes and regulation of hormone levels.

The M5 module, highly associated with the mesenchymal subtype ($p = 5e-72$), includes hub genes critical for extracellular matrix (ECM) dynamics, such as *FBN1* (kME = 0.942), *COL3A1* (kME = 0.938), and *COL1A1* (kME = 0.928). GO analysis of M5 revealed significant enrichment in pathways related to "Multicellular Organismal Process" and "System Development" (GO:0032501, GO:0048731). This reflects the role of the mesenchymal subtype in promoting tumor invasion and metastasis through ECM remodeling. Furthermore, *AEBP1*, *TIMP3*, and *OMD*, previously identified as upregulated in the mesenchymal vs. immunoreactive comparison, are key contributors to this ECM-rich module (Fig. 3). Similarly, the M8 module positively correlates with the proliferative subtype, further reinforcing the distinct gene expression patterns associated with these subtypes. GO analysis for M8 indicated enrichment in "System Development" (GO:0048731) and "Multicellular Organism Development" (GO:0032501), suggesting its involvement in tumor growth and self-renewal. Overall, our analysis highlights distinct molecular landscapes of the four mRNA subtypes and reveals key biological pathways and gene functions driving the molecular differences between subtype expression profiles.

Immune cell profiling and association with HGSOC gene modules

We used the Fisher Exact Test (FET) to investigate the overlap between gene modules and specific immune cell types, aiming to elucidate the relationship between gene co-expression networks and immune cell infiltration. Initially, we categorized genes into modules using WGCNA, which organized the genes based on co-expression patterns across samples. We identified immune cell marker genes associated with various immune cell types (Fig. 4), sourced from literature and databases. To evaluate the association between gene modules and immune cells, we constructed a contingency table that cross-tabulated gene assignments to modules against the presence or absence of each immune cell type, based on the expression levels of their marker genes. Given the shared nature of some marker genes across cell types, duplicates were permitted.

Fisher's exact test was used to determine if the distribution of genes across modules significantly differed between samples with and without each immune cell type. Significant p-values indicated non-random associations between module assignments and immune cell presence, suggesting potential interactions or regulatory relationships. Multiple testing was adjusted using the Benjamini–Hochberg method. The null hypothesis posited no significant association between modules and immune cell markers.

We failed to reject the null hypothesis for the following modules: Brown (M3), Green (M5), Midnight Blue (M15), and Light Green (M18). However, the heatmap shows the Brown (M3) module exhibited a strong association with M1 macrophages ($p < 0.0001$) and additional correlations with M2 macrophages and Th1 cells ($p < 0.05$). The Green (M5) module showed the highest immune marker significance with fibroblasts ($p < 0.0001$), while the Midnight Blue module (M15) was significantly associated with ovarian stromal cells ($p < 0.0001$) and the Light Green module (M18) was significantly associated with B cells ($p < 0.001$).

Immune cell fraction analysis and Estimated Proportions of Immune Cells (EPIC) in Bulk mRNA

To further investigate the immune landscape, EPIC analysis focused on *cellFractions*, representing the possible proportions of different cell types given in the sample. This approach offers an accurate comparison against bulk sample proportions, reflecting the fraction of each cell type present. Seven cell types were quantified: B Cells, CD4 + T Cells, CD8 + T cells, NK cells, macrophages, endothelial cells, and other (uncharacterized) cells [29]. Uncharacterized cells exhibited the highest fraction across the dataset. However, the subtype-specific analysis revealed the mesenchymal subtype had the highest proportion of cancer-associated fibroblasts (CAFs), supporting previous findings in FET analysis. These findings suggest a distinctive cellular composition in the

(See figure on next page.)

Fig. 3 iGRAPH Module Network and STRING Analysis of Hub Genes in HGSOC Molecular Subtypes. **A**, **C**, **E** iGraph network plots displaying the hub genes from three immune-mediated co-expression modules (M3, M5, M18) in HGSOC for 389 patient samples. The size of each node represents the top 10% of hub genes based on kME values, and edge thickness reflects the strength of co-expression between genes. The network highlights central genes driving each module's function, revealing distinct molecular signatures across subtypes. **B**, **D**, **F** STRING analysis of each module's top 10 hub genes illustrates predicted protein–protein interactions. The nodes represent individual proteins, while the edges reflect the confidence of interactions (line thickness indicates interaction confidence). Enrichment analysis within STRING revealed key pathways related to immune response, cellular differentiation, and extracellular matrix remodeling, which are linked to immunoreactive and mesenchymal subtypes
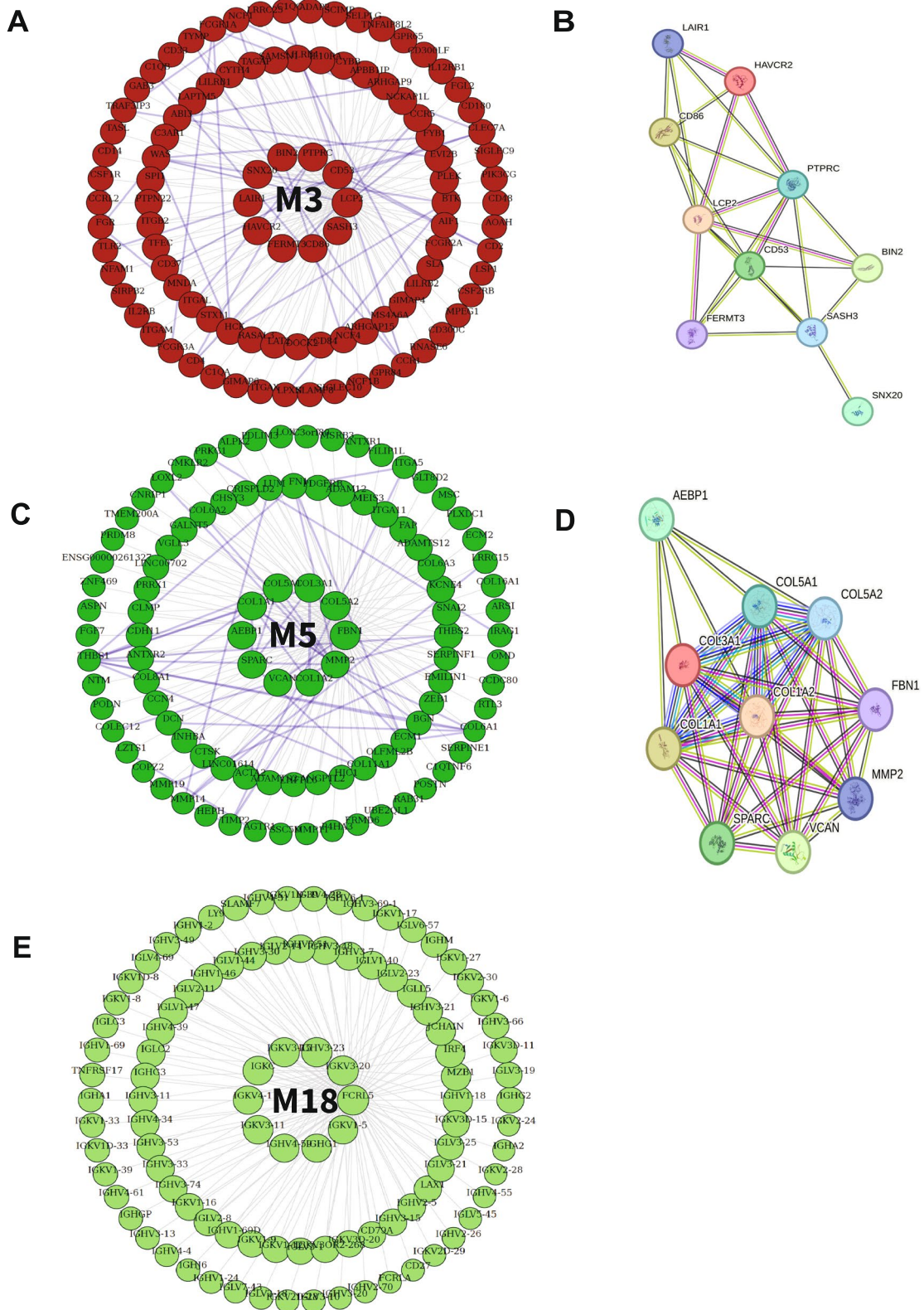
Carey *et al. Journal of Ovarian Research*   (2024) 17:240

Page 9 of 18

Fig. 3 (See legend on previous page.)

mesenchymal subtype, potentially contributing to its more aggressive clinical behavior.

TIMER 2.0 analysis

To further assess the association between gene expression and immune infiltration levels of relevant cell types for each module identified by our FET analysis, we benchmarked alternative immune cell deconvolution methods using the TIMER 2.0 platform. We focused on the top hub genes, which are defined as the most central module hubs in each WGCNA module. A functional heatmap was generated to show the association between gene expression and immune infiltration level of a given cell type, estimated by several algorithms, including EPIC, CIBERSORT, TIMER, QUANTISEQ, XCELL, and TIDE [30].

The correlation of gene expression, tumor purity and infiltration level of specific cell types is depicted in the heatmap generated using TIMER 2.0, adjusted *p*-values were used to account for multiple comparisons, while the scatterplot reports raw *p*-values for the correlation between gene expression and immune infiltration (Fig. 5). The full list of correlations between gene expression profiles and different cell types across multiple deconvolution methods (Additional Table 6). The presence of macrophages and T helper cells in bulk samples were correlated with the most central hub genes of the Brown (M3) module. The most significant hub gene, LAIR-1, exhibited a robust positive correlation with macrophage/monocyte infiltration levels ($\rho = 0.920$, adjusted $p = 1.63\text{e-}101$), as estimated by MCPCOUNTER. Tumor purity, defined as the proportion of cancer cells in the sample, showed a significant negative correlation with LAIR-1 expression levels ($\rho = -0.658$), adjusted $p = 1.96\text{e-}32$).

The significant association of the Green (M5) module with fibroblasts, as indicated by our Fisher Exact Test (FET) analysis, prompted us to investigate the presence of CAFs in relation to the expression of centralized hub genes from the M5 module. There were strong positive correlations between collagen-producing genes—COL1A1, COL1A2, COL3A1, COL5A1, and COL5A2—and CAF infiltration levels across the EPIC, MCP-COUNTER, and TIDE algorithms ($\rho > 0.9$).

Additionally, COL3A1 demonstrated a robust positive correlation with CAF infiltration ($\rho = 0.973$, adjusted $p = 8.38\text{E-}158$), while tumor purity showed a significant negative correlation with COL3A1 expression ($\rho = -0.541$, adjusted $p = 1.91\text{e-}20$). These findings suggest a potential role for COL3A1 in CAF activation.

Central hub genes in the Light Green module (M18) showed a significant association with B cells ($p < 0.001$), with two hub genes available for TIMER 2.0 analysis. FCRL5 ($\rho = 0.667$, adjusted $p = 1.44\text{e-}32$) and IRF4 ($\rho = 0.649$, adjusted $p = 3.09\text{e-}30$) exhibited moderate significant positive correlations with B cell infiltration levels estimated by EPIC. Additionally, tumor purity values correlated negatively with FCRL5 ($\rho = -0.507$, adjusted $p = 1.05\text{e-}17$) and IRF4 ($\rho = -0.542$, adjusted $p = 1.7\text{e-}20$) gene expression adjusted, suggesting that both genes may be expressed by B cells infiltrating the tumor and could serve as biomarkers for immune infiltration.

Discussion

This study presents an integrative analysis of immune cell infiltration and gene co-expression networks in HGSOC, revealing distinct cellular landscapes and their potential impact on tumor biology. By using differential expression among mRNA subtypes, applying WGCNA to identify key modules, and leveraging multiple deconvolution methods to estimate immune cell proportions in bulk mRNA samples, we discovered specific immune cell fractions—particularly macrophages, Th1 cells, B cells, ovarian stromal cells, and CAFs are closely associated with distinct gene networks. These findings highlight the significant role of immune infiltration in modulating gene expression within the tumor microenvironment, thereby influencing tumor progression.

Identifying hub genes from WGCNA provided insights into the functional roles of gene clusters in HSCOC subtypes. Hub genes, identified through high kME scores, are central to the biological processes represented by module eigengenes and often drive the module's behavior [21]. Notably, the M3 module is influenced by M1 and M2 macrophages and Th1 cells, suggesting potential therapeutic targets within this immune-driven network. This module highlights the interplay between immune

(See figure on next page.)

Fig. 4 Immune Cell Enrichment and Composition in HGSOC. **A** Heatmap depicting the enrichment of immune cell markers across HGSOC subtypes (*n* = 389; additional details in Additional File 5). Rows represent module membership, and columns represent specific immune cell types. Colors indicate the significance of the relative enrichment scores present on the scale, with purple denoting higher enrichment and yellow indicating lower enrichment. Box plots representing the fractions of eight different immune cell types across HGSOC subtypes within the bulk tumor expression (**B**) across the whole dataset and (**C**) subtype-specific, as calculated by EPIC. Each box plot shows each cell fraction's median, interquartile range, and outliers for each cell fraction. Statistical significance is indicated by asterisks using the following scale (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$)
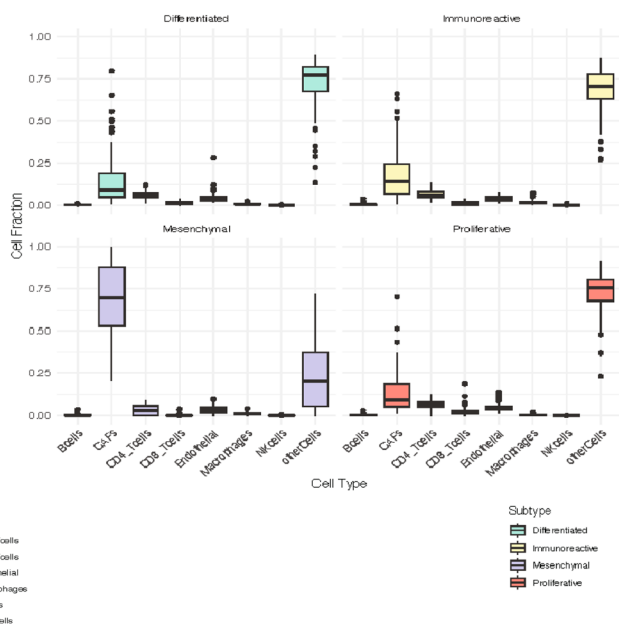
A
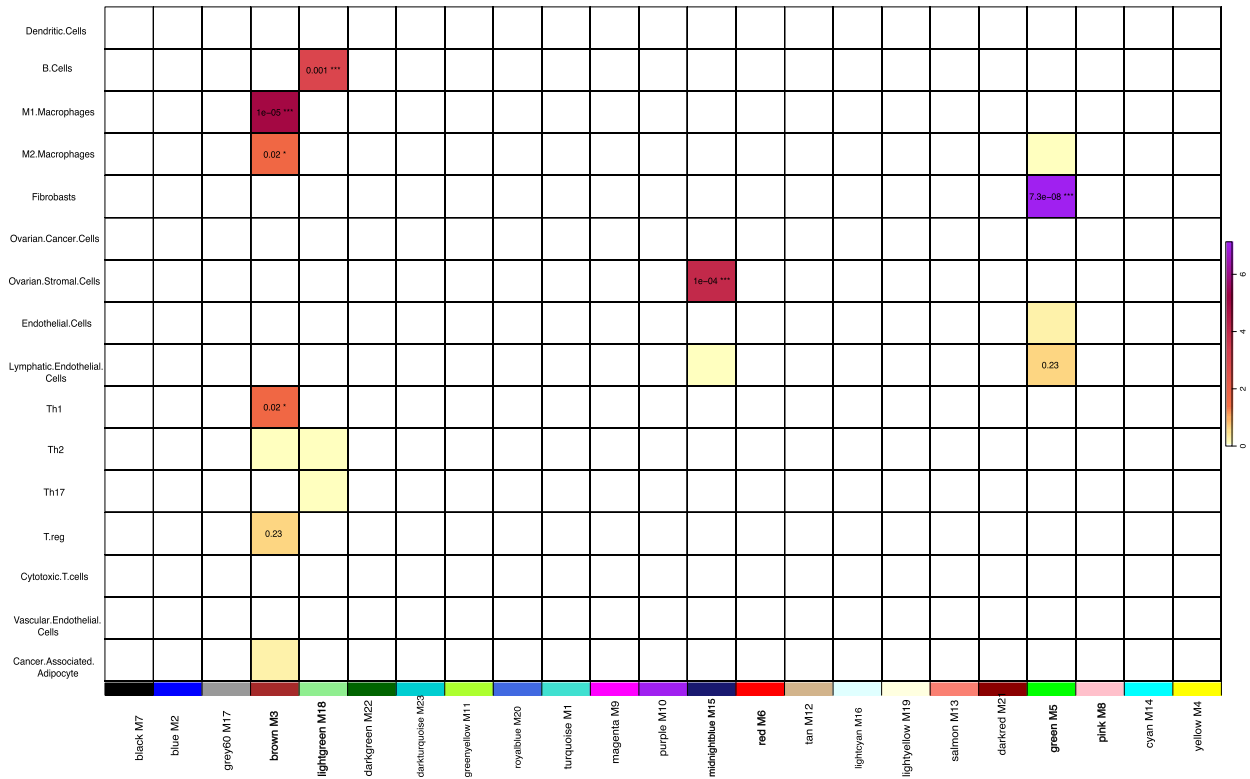
B
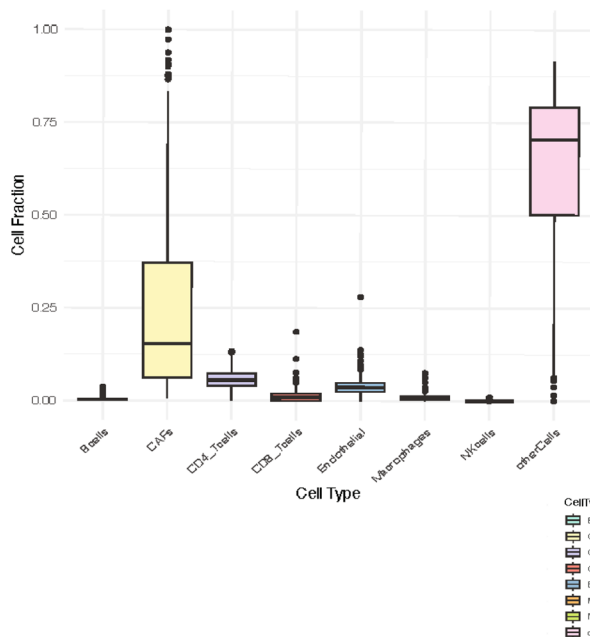
C

Fig. 4 (See legend on previous page.)

cell activity and gene regulation in the tumor microenvironment, emphasizing its relevance for future therapeutic strategies.

Macrophages play a crucial role in maintaining homeostasis and are implicated in both autoimmune diseases and malignancies [31, 32]. M1 macrophages exhibit a pro-inflammatory phenotype and are essential for anti-infectious, anti-viral, and anti-tumor immunity [32]. In contrast, M2 macrophages are characterized by anti-inflammatory functions and primarily involved in tissue repair mechanisms, wound healing, and promoting angiogenesis [33, 34]. As a result, M2 macrophages are often associated with poor cancer outcomes [35]. Tumor-associated macrophages (TAMs) are present in significantly greater numbers in HGSOC, than compared to low grade serous ovarian tumors [36]. Of the HGSOC TAM infiltrates, significantly more M2 macrophages have been observed than compared to M1 macrophages. However, our study shows a higher level of M1 ($p < 0.0001$) than compared to M2 ($p < 0.02$) macrophage molecular signatures, which suggest M1 macrophages play a greater role in the M3 module gene-network than M2 macrophages and a significant association with the immunoreactive subtype. Higher M1/M2 macrophages have been linked to better clinical outcomes in HGSOC [36, 37]. Consequently, the immunoreactive subtype has been found to have the best clinical prognosis [27]. Future studies could explore the underlying mechanism present in the M3 gene network and the immunoreactive HGSOC subtype, focusing on M2 / M1 repolarization and its impact on outcomes.

Notably, LAIR-1 or leukocyte-associated immunoglobulin-like receptor 1 gene expression was positively correlated with both M1 and M2 macrophages across several deconvolution methods. The only negative correlation observed was with the M2 macrophage estimate derived from the TIDE method ($\rho = -0.603$, adjusted $p = 1.33\text{e-}25$). TIDE focuses specifically on immune evasion and predicting immunotherapy response by measuring a tumor's potential of evading an immune checkpoint blockade [38]. While this negative correlation could reflect minor differences in the underlying algorithmic

approaches of TIDE's method, it may also have biological implications. LAIR-1, an immune inhibitory receptor in various immune cell types, has been implicated in macrophage polarization [39].

Our results for tumor purity suggest that LAIR-1 is likely influencing M2 macrophages infiltrating the tumor. Tumor purity refers to the proportion of cancer cells present in a tumor, thus, a negative correlation, or low tumor purity, suggests that there is a higher infiltration of immune cells [30, 40]. Therefore, the observed negative correlation between LAIR-1 and M2 macrophage infiltration, in the context of low tumor purity, suggests that LAIR-1 may influence the balance between M1 and M2 macrophage states, potentially driving M3 expression and promoting immune suppression and tissue remodeling functions within the TIME.

Recent studies have demonstrated that LAIR-1 is expressed by macrophages and regulates cancer cell recruitment [41], as well as downregulates immune activation [42]. It has also been implicated in epithelial ovarian cancer, correlating with tumor grade [43], and in oral squamous cell carcinoma, where it mediates immune suppression [44]. Furthermore, LAIR-1 has been shown to suppress cell growth through the PI3K-AKT-mTOR mediated pathway present in various ovarian cancer cell lines [45]. Given LAIR-1's role in immune regulation and expression in macrophages, it may significantly contribute to tumor-promoting functions within the TIME.

TAMs work with other tumor-infiltrating lymphocytes (TILs) to modulate tumor responses [46]. In the context of HGSOC, TILs are known to influence survival and enhance T cell immune surveillance. However, TILs are also associated with expressing multiple inhibitory-mediated receptors, such as LAG-3, PD-1, and TIM-3 [47], indicative of a partially exhausted state [48]. The complex interplay between LAIR-1 expression, its influence on the TIME, and the presence of exhausted TILs may contribute to the lack of immunotherapy options in HGSOC. The transition from M1/M2 macrophage- and Th1-related genes in M3 to B cell-related genes in M18 indicates a shift in the TIME. Central hub genes in M18, FCRL5 (Fc Receptor-Like 5) and IRF4 (Interferon

(See figure on next page.)

Fig. 5 Immune Cell Infiltration Analysis in High-Grade Serous Ovarian Cancer (HGSOC) using TIMER 2.0. This figure illustrates the immune cell infiltration profiles for three cell types —Macrophages, B cells, and Cancer-Associated Fibroblasts (CAFs)—using multiple deconvolution methods in TIMER 2.0, with $n = 303$ (**A**, **C**, **E**). Heatmaps illustrating expression intensity of Macrophages (**A**), B cells (**C**), and CAFs (**E**) across different HGSOC samples based on several deconvolution methods: EPIC, TIMER, CIBERSORT, MCPCOUNTER, TIDE, QUANTISEQ and xCell. The color gradient found in the heatmap represents Spearman correlation values, with red indicating positive correlations and blue representing negative ones. Statistically significant correlations ($p < 0.05$) are marked. Significant variations in infiltration profiles are seen across samples and methods. **B**, **D**, **F** Scatter plots of gene expression levels (LAIR-1, COL3A1, COL1A1, FCRL5 and IRF4) vs. tumor purity and infiltration levels for each of the three immune cell types. Tumor purity is negatively correlated with immune infiltration, and these panels visualize the relationship between gene expression and estimated cell infiltration
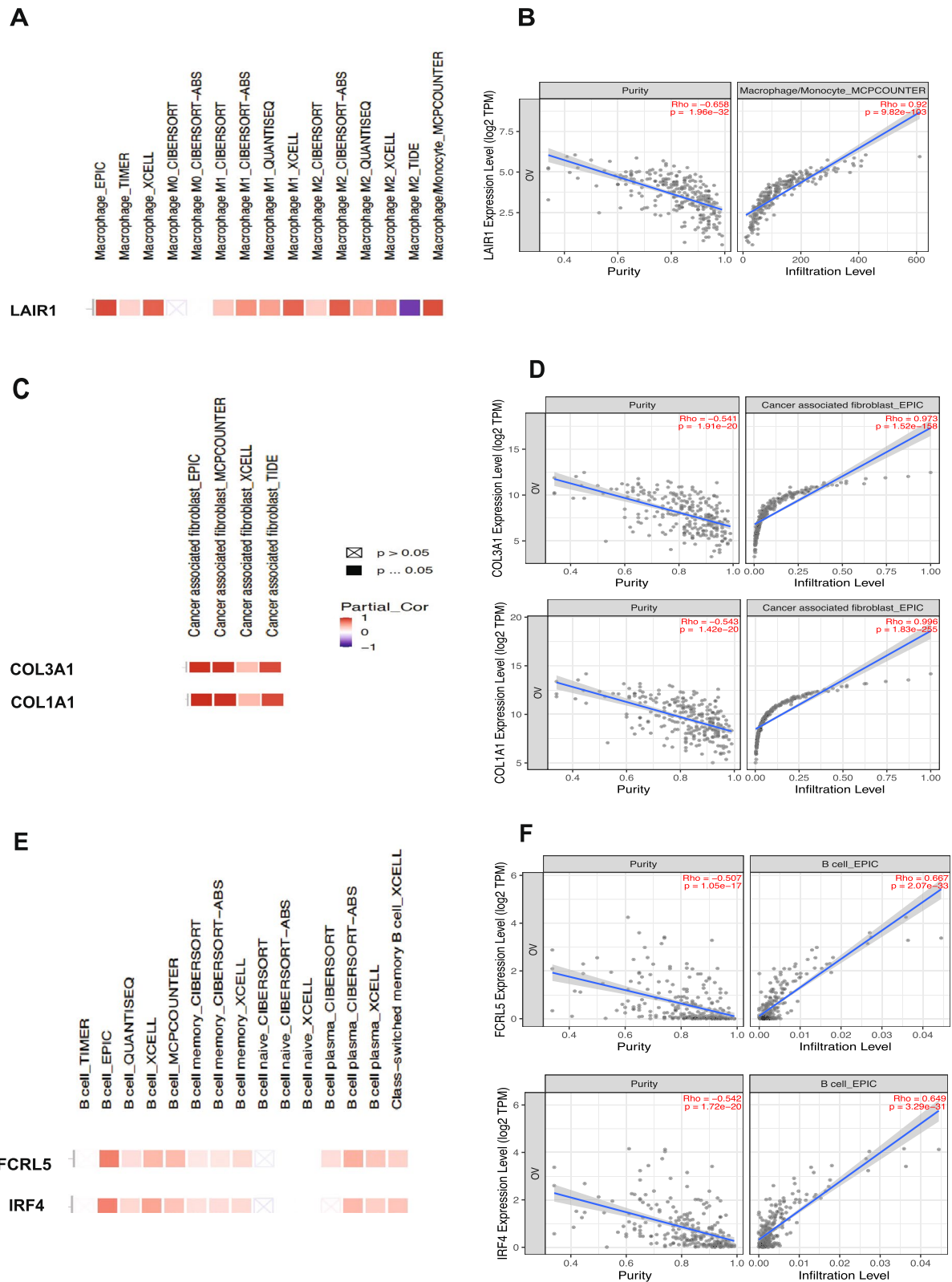
Carey *et al. Journal of Ovarian Research*   (2024) 17:240

Page 13 of 18

Fig. 5 (See legend on previous page.)

Regulatory Factor 4), play significant roles in regulating B cell responses, potentially contributing to immune suppression and tolerance within the tumor. In essence, FCRL5 might act as a regulator of B cell function by modulating the effects of IRF4's transcriptional control.

The modules that were identified in WGCNA reflect the distinct associations of the three mRNA subtypes of HGSOC (i.e., immunoreactive, proliferative, mesenchymal), showcasing subtype-specific gene networks. The differentiated subtype, exhibiting overlapping gene expression profiles across modules, suggests a mediatory role between subtypes, potentially acting as a microenvironment stabilizing subtype. Genes such as GATA4, FOXL2, and MYOCD indicate prioritization of differentiation and structural integrity prioritization, distinguishing it from the proliferative subtype. This suggests that while the differentiated HGSOC subtype shares some common features with other subtypes, its overall gene expression signature prioritizes cell differentiation and structural integrity, as shown in previous studies [7, 27, 49].

Interestingly, LINC02864, a long non-coding RNA (lncRNA), was identified in our differential expression analysis as one of the top five downregulated genes in the differentiated subtype, while being significantly upregulated in the immunoreactive subtype. Moreover, this contrasting pattern of expression suggests LINC02864 plays a role in immune regulation, since its elevated levels in the immunoreactive subtype further correlate with better clinical outcomes [18]. lncRNAs regulate cellular function by modulating transcriptional activity and biological signaling [50], with some also influencing tumor suppression [51]. Previous studies have shown that LINC02864 is prevalent in head and neck cancers and gastric cancers and interacts with several miRNAs (miR-3155a, miR-3678-3p, miR-875-3p, miR-758-3p, miR-3155b, miR-3612, miR-6873-5p, miR-4443, miR-484) [52]. However, it's direct mechanism is largely unknown. To this end, miR-875-3p has been implicated in immune suppression in gastric cancer [53] potentially linking LINC02864 activity in the immunoreactive subtype. Though LINC02864 is part of the grey module in WGCNA, which often contains genes that are not clustered in functional networks, its differential expression remains noteworthy.

The M6 module, focusing on transcriptional and cell cycle signaling, shows a positive correlation with the proliferative subtype, implicating cell cycle progression in tumor development. SALL2 and TCF7L1 play critical roles in regulating cell growth and proliferation, with SALL2 more focused on stress responses and apoptosis [54], and TCF7L1 involved in Wnt signaling and stem cell maintenance [55]. However, there was a negative correlation with the differentiated subtype for the M6 module which may influence cell growth, affecting tumor development or cancer stem cell maintenance in HGSOC.

Although our differential expression analysis showed that DLK1 was upregulated in the proliferative versus immunoreactive HGSOC subtypes, it belongs to the M17 module, which is positively correlated with the differentiated subtype and negatively correlated with the proliferative subtype. This discrepancy suggests DLK1 plays a dual role, promoting differentiation in certain contexts while being elevated in proliferative environments. Interestingly, IGDCC3 was found to be upregulated in both proliferative and mesenchymal HGSOC subtypes. IGDCC3 (Immunoglobulin Superfamily DCC Subclass Member 3), known for its role in nervous system development, contributes to cell adhesion and migration involved in tumor progression [56]; however, more research is needed to confirm this theory. Thus, the additional hub genes presented in the modulated subtypes with either cell growth (proliferative) or division or tissue remodeling and invasion (mesenchymal), highlighting the plasticity present in the HGSOC TIME. This suggests the shared upregulation of a gene across different subtypes may trigger distinct molecular switches in each subtype due to the unique gene networks that influence pathway activation.

GO analysis of the M15 module revealed significant enrichment in processes related to hormone regulation (GO: 0010817), suggesting a strong stromal influence on tumor biology, particularly in connection with the proliferative subtype. This finding aligns with previous studies, highlighting the stroma as a key modulator of the TIME [57, 58]. The proliferative subtype, associated with more aggressive clinical outcomes, exhibited the highest degree of gene dysregulation in our differential analysis. Hence, the M15 module, driven by ovarian stromal cells, underscores the critical role of the tumor stroma in HGSOC progression.

The mesenchymal subtype is linked to the most aggressive disease and poorest prognosis. Differential expression analysis between mesenchymal and immunoreactive subtypes revealed upregulation of TIMP3 (Tissue Inhibitor of Metalloproteinases 3), OMD (Osteomodulin), and AEBP1 (Adipocyte Enhancer-Binding Protein 1). These hub genes were also observed in the mesenchymal-dominant M5 module. AEBP1, a key transcriptional repressor, plays a crucial role in collagen synthesis, contributing to tumor stiffness and fibrosis, which promotes tumor progression [59]. Several studies have shown that several collagen-related genes such as AEBP1, COL5A1 and TIMP3 are associated with poorer survival in HGSOC [60]. Others have shown COL1A1 is elevated in ascites from patients with epithelial ovarian cancer, confirming that

it has mainly originated from fibroblasts [61]. COL1A1 promotes metastatic disease by activating the ITGB1/AKT pathway, highlighting a potential target for disease progression.

The downregulation of OVGP1 in the mesenchymal subtype and role in reproductive support and ECM regulation, suggest its loss of expression may diminish normal cellular functions and increase tumorigenesis. However, as OVGP1 was assigned to the grey module in our analysis, further investigation is needed to clarify its role in this context. Studies have shown that OVGP1 is more apparent in earliest-stage serous carcinomas and tumors defined as borderline and mucinous-associated carcinomas [62], suggesting it is an important modulator for earlier and less aggressive serous phenotypes.

Therapeutic targets identified in our analysis include several key genes associated with distinct modules. For instance, LAIR-1, found in the M3 module was characterized by immunoreactive and M1/M2/Th1 cell interactions, is a target for therapeutic intervention in acute myeloid leukemia (AML) [63]. Additionally, COL3A1 and COL1A1 were part of the M5 module and associated with mesenchymal and fibroblast characteristics. COL3A1 serves as a target for Ehlers-Danlos syndrome [64], while COL1A1 is prognostic for breast cancer outcomes [65]. Moreover, FCRL5 and IRF4, located within the M18 module and link to B cells, represent promising targets in hematological malignancies. Indeed, alvocidib, an IRF4 inhibitor, has shown potential in treating adult T-cell leukemia/lymphoma [66], whereas CAR-T cells
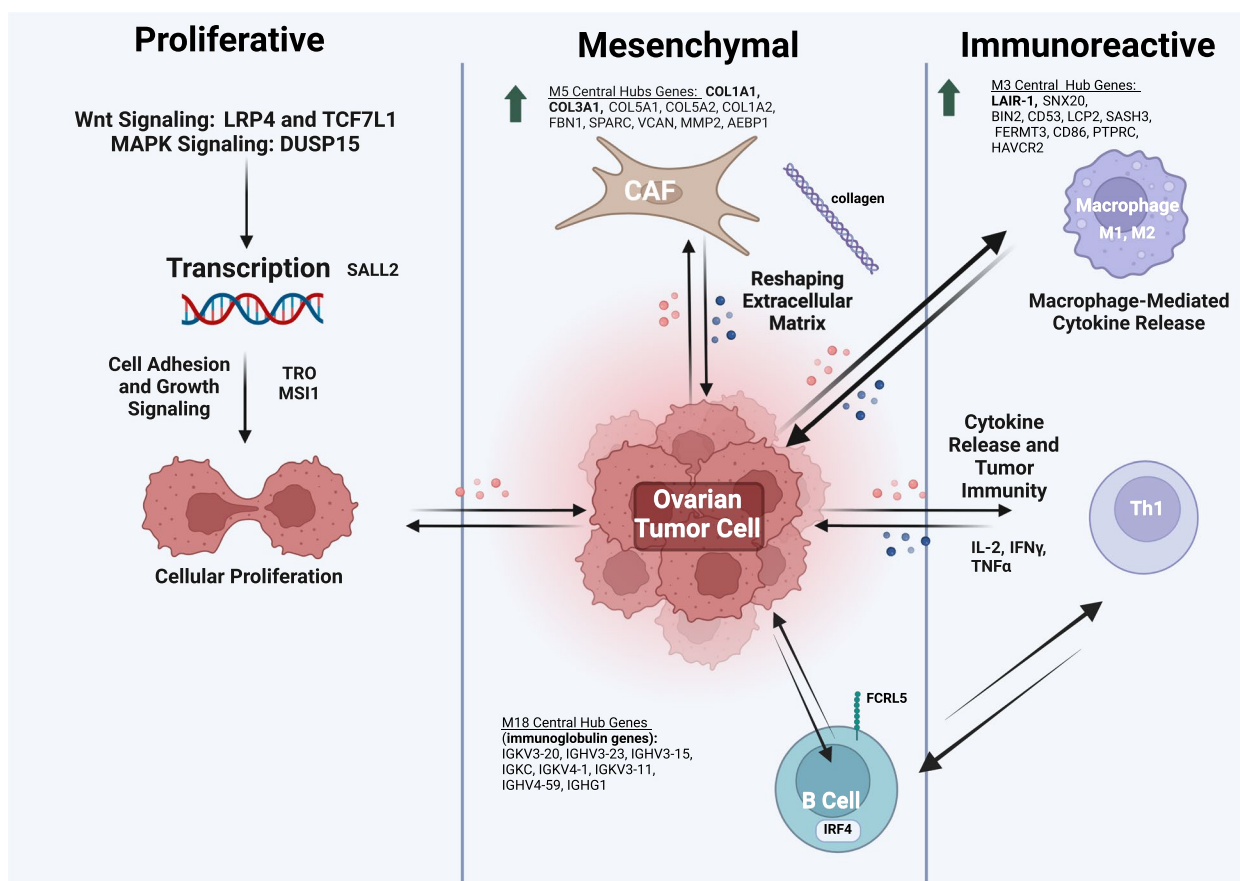

Fig. 6 mRNA Subtypes and Immune Cell Interactions in the Tumor Immune Microenvironment (TIME) present in HGSOC. This schematic represents the interactions between distinct mRNA subtypes (proliferative, mesenchymal, and immunoreactive) and the immune microenvironment in HGSOC tumors. Key immune cell populations, including B cells, macrophages, Th1 cells, and Cancer-Associated Fibroblasts (CAFs), are depicted in relation to their transcriptional signatures and subtype-specific gene expression. The proliferative subtype shows elevated expression of genes involved in cell cycle regulation, driving tumor growth. CAFs and macrophages in this subtype promote extracellular matrix remodeling and tumor invasion, with proliferative gene signatures influencing immune suppression within the TIME. The mesenchymal subtype is characterized by high infiltration of CAFs and M2 macrophages, supporting aggressive invasion and matrix remodeling. This subtype is linked to poor prognosis, with CAF-driven pathways enhancing immune evasion. The immunoreactive subtype features high levels of B and Th1 cells, correlating with immune activation. M1 macrophages, present in this subtype, promote tumor immune responses, which are reflected in the strong immune-related gene expression signature. This figure showcases the complex immune dynamics and transcriptional networks in the TIME. BioRender.com/q00f536

that are FCRL5-directed show promise in treating multiple myeloma [67].

Our findings contribute to the complex interplay between tumor cells and the immune microenvironment. By demonstrating how immune cell infiltration shapes gene expression networks, we deepened our knowledge of tumor biology and the potential for developing immunotherapies targeting this interaction. Talhouk et al. [68] developed a gene expression-based predictor, PrOTYPE, that classifies subtypes of high-grade serous ovarian carcinoma, further emphasizing the importance of such gene expression analyses. Studies like INOVate (Individualized Ovarian Cancer Treatment Through Integration of Genomic Pathology into Multidisciplinary Care) [68] highlight efforts to refine patient classification based on tumor genomic profiling and histological subtyping underscores the relevance of the current study and its potential application in patient care.

Limitations

While our study provides valuable insights into the regulation and landscape of HGSOC TIME, we acknowledge several limitations. The reliance on bulk RNA-seq data may mask important cellular heterogeneity and our sample size limits the statistical power of our analyses. Future studies should incorporate single-cell RNA sequence analysis to better elucidate specific contributions of various immune cell could improve precision medicine approaches by offering a more comprehensive view of how the factors influencing subtype-specific immune responses impact patient outcomes. Davidson et al. [69] demonstrated that while the prevalence of HGSOC varied across self-identified racial groups (i.e., Black, White, Japanese), subtype-specific expression patterns were similar for survival, showcasing the potential for subtype-specific prognostic outcomes. However, further investigation is needed to understand the deeper role of genetic ancestry. While our analysis had limited power for the under-45 age group, prospective analyses could further examine how age may influence subtype prevalence, offering valuable insights into which gene networks are particularly relevant at different life stages. An additional key next step involves exploring potential therapeutic targets, particularly related to immune modulation, among the hub genes identified. Future studies are also needed to validate our computational findings for clinical implications, using immunostaining for the most centralized hub genes as candidate biomarkers. Additionally, RT-PCR could also be adapted to assess expression levels and prognostic persistence.

Conclusion

Our results pave the way for future investigations into the mechanistic roles of specific hub genes in TIME modulation in HGSOC and highlight distinct gene expression profiles across mRNA subtype groups in HGSOC patients (Fig. 6). Differential expression of genes involved in immune response and cell proliferation suggests potential differences in TIME and biological characteristics among subtype groups. Our study suggests that specific immune cell profiles, particularly the presence of M1 macrophages, could serve as biomarkers for patient stratification as well as LAIR-1 and IRF4 as therapeutic targets in HGSOC. Taken together, our findings provide valuable insights into the molecular mechanisms underlying HGSOC subtype-specific differences, which have implications for prognosis and therapeutic strategies.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13048-024-01556-4.

Additional file 1: Table 1. DESEQ2_master_table_padj_ALLsubtypecombinations. This table contains the comprehensive DESeq2 analysis results for all subtype combinations of high-grade serous ovarian cancer (HGSOC). It includes the adjusted *p*-values (p-adj) for differential gene expression across all mRNA subtype comparisons, using different subtypes as denominators for releveling. The subtypes evaluated are immunoreactive, differentiated, proliferative, and mesenchymal. The table captures all pairwise comparisons between subtypes, providing a detailed breakdown of significant gene expression changes and allowing for cross-subtype analysis of molecular signatures. This file is a critical resource for identifying subtype-specific gene expression patterns and their potential clinical relevance.

Additional file 2: Table 2. ME values with module size, number and ANOVA p value statistics. This table contains all the WGCNA module eigengene (ME) assignments including module size, color and ANOVA *p* values for M1-M23 for spreadsheet one. Spreadsheet two has all of the patients ME statistics.

Additional file 3: Table 3. WGCNAModuleAssignmentsAugust2024_Post-DESEQ2_WGCNA.ratio_power8_MergeHeight0.15_PAMstageTRUE.txt: This file contains all WGCNA module assignments and corresponding kME values (module membership scores) values for all genes in the dataset. It is based on a WGCNA analysis performed post-DESeq2, using a soft power threshold of 8, a merge height of 0.15, and PAM stage assignments set to TRUE.

Additional file 4: Table 4. ALL Modules for GO ANALYSIS_FDR.BHcorrected. txt: This table includes the results of all Gene Ontology (GO) results for all modules identified in the WGCNA analysis. The GO terms used a multiple testing method known as the FDR (False Discovery Rate) with Benjamini-Hochberg (BH) adjustment.

Additional file 5: Table 5. HGSOC-CellTypeMarkerListFET with annotations. This file includes all marker genes used for cell type classification in the Tumor Immune Microenvironment (TIME) for HGSOC. The table includes Entrez Gene IDs, PMIDs, and citations to the relevant publications.

Additional file 6: Table 6. TIMER 2.0 Results of OV samples (*n*=303). This spreadsheet data contains the results of TIMER 2.0 analysis for ovarian cancer (OV) samples (*n*=303). It includes adjusted*p*-values (padj) for selected hub genes, providing insights into immune cell infiltration and tumor purity correlations.

Carey *et al. Journal of Ovarian Research*   (2024) 17:240

Page 17 of 18

Data availability

No datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate

This study is aligned with TCGA's Ethics and Policies (https://www.cancer.gov/ccg/research/genome-sequencing/tcga/history/ethics-policies) for collecting and analyzing data. Informed consent was received from all patients in the study, informing and ensuring them of the purpose, risks and benefits associated with the study. All researchers are required to have a token, a secure authentication code, for controlled-access data to ensure data privacy and security. No conflicts of interests were identified.

Competing interests

The authors declare no competing interests.

Author details

[1] Department of Microbiology, Biochemistry, and Immunology, Morehouse School of Medicine, 720 Westview Dr SW, HG 341B, Atlanta, GA 30310, USA. [2] Department of Health and Human Services, Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA. [3] Center for Neurodegenerative Disease, Emory University School of Medicine, Atlanta, GA 30322, USA. [4] Department of Biochemistry, Emory University School of Medicine, Atlanta, GA 30322, USA.

References

1. Wessolly M, Mairinger E, Borchert S, Bankfalvi A, Mach P, Schmid KW, et al. CAF-Associated Paracrine Signaling Worsens Outcome and Potentially Contributes to Chemoresistance in Epithelial Ovarian Cancer. Front Oncol. 2022;12:Article 850847. https://doi.org/10.3389/fonc.2022.850847.
2. Buamah PK, Skillen AW. Serum CA 125 concentrations in patients with benign ovarian tumours. J Surg Oncol. 1994;56(2):71–4.
3. Buamah P. Benign conditions associated with raised serum CA-125 concentration. J Surg Oncol. 2000;75(4):264–5.
4. Tothill RW, Tinker AV, George J, Brown R, Fox SB, Lade S, et al. Novel molecular subtypes of serous and endometrioid ovarian cancer linked to clinical outcome. Clin Cancer Res. 2008;14(16):5198–208.
5. Cancer Genome Atlas Research Network. Integrated genomic analyses of ovarian carcinoma. Nature. 2011;474(7353):609–15.
6. Yang Z, Xu S, Jin P, Yang X, Li X, Wan D, et al. MARCKS contributes to stromal cancer-associated fibroblast activation and facilitates ovarian cancer metastasis. Oncotarget. 2016;7(25):37618–30.
7. Waldron L, Riester M, Birrer M. Molecular subtypes of high-grade serous ovarian cancer: the holy grail? J Natl Cancer Inst. 2014 Sep 30;106(10):dju297. https://doi.org/10.1093/jnci/dju297.
8. Giusti R, Porzio G, Maltoni M, Filetti M, Cuomo A, Bandieri E, et al. Association of opioid use with survival in patients with cancer treated with immune checkpoint inhibitors: it is time for evidence-based behaviors. Oncologist. 2024;29(1):1–10.
9. Singh R, Singh S, Grizzle WE, Lillard JW. CXCL13 and CXCR5 expression by non-small cell lung carcinoma. J Immunol. 2011;186(1):149.15.
10. El Haibi CP, Sharma PK, Singh R, Johnson PR, Suttles J, Singh S, et al. PI3Kp110-, Src-, FAK-dependent and DOCK2-independent migration and invasion of CXCL13-stimulated prostate cancer cells. Mol Cancer. 2010;9:85.
11. Singh R, Stockard CR, Grizzle WE, Lillard JW Jr, Singh S. Expression and histopathological correlation of CCR9 and CCL25 in ovarian cancer. Int J Oncol. 2011;39(2):373–81.
12. El-Haibi CP, Singh R, Sharma PK, Singh S, Lillard JW Jr. CXCL13 mediates prostate cancer cell proliferation through JNK signalling and invasion through ERK activation. Cell Prolif. 2011;44(4):311–9.
13. Bergsten TM, Levy SE, Zink KE, Lusk HJ, Pergande MR, Cologna SM, et al. Fallopian tube secreted protein affects ovarian metabolites in high grade serous ovarian cancer. Front Cell Dev Biol. 2022;10:1042734.
14. Grassi ES, Pietras A. Emerging roles of DLK1 in the Stem cell niche and cancer stemness. J Histochem Cytochem. 2022;70(1):17–28.
15. Pittaway JFH, Lipsos C, Mariniello K, Guasti L. The role of delta-like non-canonical Notch ligand 1 (DLK1) in cancer. Endocr Relat Cancer. 2021;28(12):R271–87.
16. Oldham MC, Konopka G, Iwamoto K, Langfelder P, Kato T, Horvath S, et al. Functional organization of the transcriptome in human brain. Nat Neurosci. 2008;11(11):1271–82.
17. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550.
18. Kassuhn W, et al. Classification of molecular subtypes of high-grade serous ovarian cancer by MALDI-imaging. Cancers (Basel). 2021 Mar 25;13(7):1512. https://doi.org/10.3390/cancers13071512.
19. Ignatiadis N, Klaus B, Zaugg JB, Huber W. Data-driven hypothesis weighting increases detection power in genome-scale multiple testing. Nat Methods. 2016;13(7):577–80. https://doi.org/10.1038/nmeth.3885.
20. Zhu A, Ibrahim JG, Love MI. Heavy-tailed prior distributions for sequence count data: removing the noise and preserving large differences. Bioinformatics. 2018;35(12):2084–92.
21. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008;9:559.
22. Langfelder P, Horvath S. Fast R functions for robust correlations and hierarchical clustering. J Stat Softw. 2012;46(11):i11. https://doi.org/10.18637/jss.v046.i11.
23. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. Gene Ontology Consortium Nat Genet. 2000;25(1):25–9.
24. Reimand J, Isserlin R, Voisin V, Kucera M, Tannus-Lopes C, Rostamianfar A, et al. Pathway enrichment analysis and visualization of omics data using g:Profiler, GSEA Cytoscape and EnrichmentMap. Nat Protoc. 2019;14(2):482–517.
25. Jiménez-Sánchez A, Cybulska P, Mager KL, Koplev S, Cast O, Couturier D-L, et al. Unraveling tumor–immune heterogeneity in advanced ovarian cancer uncovers immunogenic effect of chemotherapy. Nat Genet. 2020;52(6):582–93.
26. Seitz S, Dreyer TF, Stange C, Steiger K, Bräuer R, Scheutz L, et al. CXCL9 inhibits tumour growth and drives anti-PD-L1 therapy in ovarian cancer. Br J Cancer. 2022;126(10):1470–80.
27. Konecny GE, Wang C, Hamidi H, et al. Prognostic and therapeutic relevance of molecular subtypes in high-grade serous ovarian cancer. J Natl Cancer Inst. 2014;106(10):dju249. https://doi.org/10.1093/jnci/dju249.
28. Keathley R, Kocherginsky M, Davuluri R, Matei D. Integrated multiomic analysis reveals immunosuppressive phenotype associated with poor outcomes in high-grade serous ovarian cancer. Cancers. 2023;15(14):3649.
29. Racle J, Gfeller D. EPIC: a tool to estimate the proportions of different cell types from bulk gene expression data. Methods Mol Biol. 2020;2120:233–48.
30. Li T, Fan J, Wang B, Traugh N, Chen Q, Liu JS, et al. TIMER: a web server for comprehensive analysis of tumor-infiltrating immune cells. Cancer Res. 2017;77(21):e108–10.

Carey *et al. Journal of Ovarian Research*   (2024) 17:240

Page 18 of 18

31. Shapouri-Moghaddam A, Mohammadian S, Vazini H, Taghadosi M, Esmaeili SA, Mardani F, et al. Macrophage plasticity, polarization, and function in health and disease. J Cell Physiol. 2018;233(9):6425–40.

32. Kadomoto S, Izumi K, Mizokami A. Macrophage polarity and disease control. Int J Mol Sci. 2021;23(1):144. https://doi.org/10.3390/ijms23010144.

33. Zhu S, Yi M, Wu Y, Dong B, Wu K. Roles of tumor-associated macrophages in tumor progression: implications on therapeutic strategies. Exp Hematol Oncol. 2021;10(1):60. https://doi.org/10.1186/s40164-021-00252-z.

34. Zhu S, Yi M, Wu Y, Dong B, Wu K. Correction to: Roles of tumor-associated macrophages in tumor progression: implications on therapeutic strategies. Exp Hematol Oncol. 2022;11(1):4.

35. Yang Q, Guo N-n, Zhou Y, Chen J, Wei Q, Han M. The role of tumor-associated macrophages (TAMs) in tumor progression and relevant advance in targeted therapy. Acta Pharmaceutica Sinica B. 2020;10:2156–70.

36. Kepuladze S, Devadze R, Gvenetadze A, Burkadze G. Distribution of tumor-associated macrophages and M1/M2 polarization in different types and grades of ovarian tumors. Indian J Pathol Oncol. 2022;9:318–21. https://doi.org/10.18231/j.ijpo.2022.076.

37. Yin M, Shen J, Yu S, Fei J, Zhu X, Zhao J, et al. Tumor-Associated Macrophages (TAMs): a critical activator in ovarian cancer metastasis. Onco Targets Ther. 2019;12:8687–99.

38. Jiang P, Gu S, Pan D, Fu J, Sahu A, Hu X, et al. Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. Nat Med. 2018;24(10):1550–8.

39. Jin J, Wang Y, Ma Q, Wang N, Guo W, Jin B, et al. LAIR-1 activation inhibits inflammatory macrophage phenotype in vitro. Cell Immunol. 2018;331:78–84.

40. Deng Y, Song Z, Huang L, Guo Z, Tong B, Sun M, et al. Tumor purity as a prognosis and immunotherapy relevant feature in cervical cancer. Aging (Albany NY). 2021;13(22):24768–85.

41. Helou DG, Quach C, Hurrell BP, Li X, Li M, Akbari A, et al. LAIR-1 limits macrophage activation in acute inflammatory lung injury. Mucosal Immunol. 2023;16(6):788–800.

42. Tang X, Narayanan S, Coligan J, Borrego F. Characterization of the interaction between human LAIR-1 and collagens. J Immunol. 2007;178(Suppl 1):S202-S202. doi: 10.4049/jimmunol.178.Supp.101.8.

43. Cao Q, Fu A-L, Yang S, He X, Wang Y, Zhang X, et al. Leukocyte-associated immunoglobulin-like receptor-1 expressed in epithelial ovarian cancer cells and involved in cell proliferation and invasion. Biochem Biophys Res Commun. 2015;458(2):399–404.

44. Yang L, Zhang M-J, Wu L, Mao L, Chen L, Yu G-T, et al. LAIR-1 overexpression and correlation with advanced pathological grade and immune suppressive status in oral squamous cell carcinoma. Head Neck. 2018;41:1080–6.

45. Liu Y, Ma L, Shangguan F, Zhao X, Wang W, Gao Z, et al. LAIR-1 suppresses cell growth of ovarian cancer cell via the PI3K-AKT-mTOR pathway. Aging (Albany NY). 2020;12:16142–54.

46. Bercovici N, Guérin MV, Trautmann A, Donnadieu E. The remarkable plasticity of macrophages: a chance to fight cancer. Front Immunol. 2019;10:1563.

47. Sakuishi K, Apetoh L, Sullivan JM, Blazar BR, Kuchroo VK, Anderson AC. Targeting Tim-3 and PD-1 pathways to reverse T cell exhaustion and restore anti-tumor immunity. J Exp Med. 2010;207(10):2187–94.

48. Bergamini A, Tassi E, Wignall JM, Ruggiero E, Bocciolone L, Candiani M, et al. P169 Activated effector T cells co-expressing multiple inhibitory receptors are enriched in the tumor immune microenvironment of HGSOC expressing high and uniform WT-1 antigen. Int J Gynecol Cancer. 2019;29:A161–2.

49. Verhaak RG, Tamayo P, Yang JY, Hubbard D, Zhang H, Creighton CJ, et al. Prognostically relevant gene signatures of high-grade serous ovarian carcinoma. J Clin Invest. 2013;123(1):517–25.

50. Hu Z, Yuan L, Yang X, Yi C, Lu J. The roles of long non-coding RNAs in ovarian cancer: from functions to therapeutic implications. Front Oncol. 2024;14:1332528.

51. Gokulnath P, de Cristofaro T, Manipur I, Di Palma T, Soriano AA, Guarracino MR, et al. Long non-coding RNA HAND2-AS1 acts as a tumor suppressor in high-grade serous ovarian carcinoma. Int J Mol Sci. 2020;21(11):4059.

52. Ershov PV, Yablokov EO, Mezentsev YV, Ivanov AS. Long intergenic non-coding RNAs of human chromosome 18: focus on cancers. Biomedicines. 2024;12(3):544.

53. Gao S, Zhang Z, Wang X, Ma Y, Li C, Liu H, et al. hsa-miR-875-5p inhibits tumorigenesis and suppresses TGF-β signalling by targeting USF2 in gastric cancer. J Transl Med. 2022;20(1):115.

54. Hermosilla VE, Hepp MI, Escobar D, Farkas C, Riffo EN, Castro AF, et al. Developmental SALL2 transcription factor: a new player in cancer. Carcinogenesis. 2017;38(7):680–90.

55. Shy BR, Wu C, Khramtsova GF, Zhang JY, Olopade OI, Goss KH, et al. Regulation of Tcf7l1 DNA binding and protein stability as principal mechanisms of Wnt/β-catenin signaling. Cell Rep. 2013;4(1):1–9.

56. Salbaum JM. Punc, a novel mouse gene of the immunoglobulin superfamily, is expressed predominantly in the developing nervous system. Mech Dev. 1998;71(1–2):201–4.

57. Zhang S, et al. Stroma-associated master regulators of molecular subtypes predict patient prognosis in ovarian cancer. Sci Rep. 2015 Nov 4;5:16066. doi: 10.1038/srep16066.

58. Zhang L, Cascio S, Mellors JW, Buckanovich RJ, Osmanbeyoglu HU. Single-cell analysis reveals the stromal dynamics and tumor-specific characteristics in the microenvironment of ovarian cancer. Commun Biol. 2024 Jan 15;7(1):20. https://doi.org/10.1038/s42003-023-05733-x.

59. Majdalawieh AF, Massri M, Ro H-S. AEBP1 is a novel oncogene: mechanisms of action and signaling pathways. J Oncol. 2020;2020(1):8097872.

60. Cheon DJ, Tong Y, Sim MS, Dering J, Berel D, Cui X, Lester J, Beach JA, Tighiouart M, Walts AE, Karlan BY, Orsulic S. A collagen-remodeling gene signature regulated by TGF-β signaling is associated with metastasis and poor survival in serous ovarian cancer. Clin Cancer Res. 2014 Feb 1;20(3):711-23. doi: 10.1158/1078-0432.CCR-13-1256.

61. Li M, Wang J, Wang C, Xia L, Xu J, Xie X, Lu W. Microenvironment remodeled by tumor and stromal cells elevates fibroblast-derived COL1A1 and facilitates ovarian cancer metastasis. Exp Cell Res. 2020;394(1):112153. https://doi.org/10.1016/j.yexcr.2020.112153.

62. Maines-Bandiera SL, Woo MMM, Borugian MJ, Molday LL, Hii T, Gilks B, et al. Oviductal Glycoprotein (OVGP1, MUC9): a differentiation-based mucin present in serum of women with ovarian cancer. Int J Gynecol Cancer. 2009;20:16–22.

63. Xie J, Gui X, Deng M, Chen H, Chen Y, Liu X, Ku Z, Tan L, Huang R, He Y, Zhang B, Lewis C, Chen K, Xu L, Xu J, Huang T, Liao XC, Zhang N, An Z, Zhang CC. Blocking LAIR1 signaling in immune cells inhibits tumor development. Front Immunol. 2022;13:996026. https://doi.org/10.3389/fimmu.2022.996026.

64. Ruscitti F, Trevisan L, Rosti G, Gotta F, Cianflone A, Geroldi A, et al. A novel mutation in COL3A1 associates to vascular Ehlers-Danlos syndrome with predominant musculoskeletal involvement. Mol Genet Genomic Med. 2021;9(9):e1753.

65. Liu J, Shen J-X, Wu H-T, Li X, Wen X-F, Du C-W, et al. Collagen 1A1 (COL1A1) promotes metastasis of breast cancer and is a potential therapeutic target. Discov Med. 2018;25(139):211–23.

66. Sakamoto H, Ando K, Imaizumi Y, Mishima H, Kinoshita A, Kobayashi Y, et al. Alvocidib inhibits IRF4 expression via super-enhancer suppression and adult T-cell leukemia/lymphoma cell growth. Cancer Sci. 2022;113:4092–103.

67. Yu Z, Li H, Lu Q, et al. Fc receptor-like 5 (FCRL5)-directed CAR-T cells exhibit antitumor activity against multiple myeloma. Signal Transduct Target Ther. 2024;9:16. https://doi.org/10.1038/s41392-023-01702-2.

68. Talhouk A, George J, Wang C, Budden T, Tan TZ, Chiu DS, et al. Development and validation of the gene expression predictor of high-grade serous ovarian carcinoma molecular SubTYPE (PrOTYPE). Clin Cancer Res. 2020;26(20):5411–23.

69. Davidson NR, Barnard ME, Hippen AA, Campbell A, Johnson CE, Way GP, Dalley BK, Berchuck A, Salas LA, Peres LC, Marks JR, Schildkraut JM, Greene CS, Doherty JA. Molecular subtypes of high-grade serous ovarian cancer across racial groups and gene expression platforms. Cancer Epidemiol Biomarkers Prev. 2024;33(8):1114–25.

Publisher's Note