


Identification of metabolism-associated molecular subtype in ovarian cancer

Xiaona Liu¹ | Aoshen Wu¹ | Xing Wang¹ | Yunhe Liu¹ | Yiang Xu^{2,3,4} | Gang Liu¹ | Lei Liu^{1,2} 

¹Institutes of Biomedical Sciences, Fudan University, Shanghai, China

²School of Basic Medical Sciences, Fudan University, Shanghai, China

³State Key Laboratory of Neuroscience, Institute of Neuroscience, Center for Excellence in Brain Science and Intelligence Technology, Chinese Academy of Sciences, Shanghai, China

⁴University of Chinese Academy of Sciences, Beijing, China

Correspondence

Lei Liu and Gang Liu, Shanghai Medical College Road 138, Shanghai 200032, China.

Emails: liulei_sibs@163.com (L.L.); liugang@fudan.edu.cn (G.L.)

Funding information

National Natural Science Foundation of China, Grant/Award Number: #91846302

Abstract

Ovarian cancer (OC) is the most lethal gynaecological cancer with genomic complexity and extensive heterogeneity. This study aimed to characterize the molecular features of OC based on the gene expression profile of 2752 previously characterized metabolism-relevant genes and provide new strategies to improve the clinical status of patients with OC. Finally, three molecular subtypes (C1, C2 and C3) were identified. The C2 subtype displayed the worst prognosis, upregulated immune-cell infiltration status and expression level of immune checkpoint genes, lower burden of copy number gains and losses and suboptimal response to targeted drug bevacizumab. The C1 subtype showed downregulated immune-cell infiltration status and expression level of immune checkpoint genes, the lowest incidence of BRCA mutation and optimal response to targeted drug bevacizumab. The C3 subtype had an intermediate immune status, the highest incidence of BRCA mutation and a secondary optimal response to bevacizumab. Gene signatures of C1 and C2 subtypes with an opposite expression level were mainly enriched in proteolysis and immune-related biological process. The C3 subtype was mainly enriched in the T cell-related biological process. The prognostic and immune status of subtypes were validated in the Gene Expression Omnibus (GEO) dataset, which was predicted with a 45-gene classifier. These findings might improve the understanding of the diversity and therapeutic strategies for OC.

KEYWORDS

immune, metabolic genes, ovarian cancer, prognosis, subtype

1 | INTRODUCTION

Ovarian cancer (OC) is the most lethal gynaecological cancer with 313,959 new cases, and 207,252 deaths reported worldwide in 2020; also, it represents the eighth most commonly diagnosed cancer among women in the world.^{1,2} About 75% of patients are diagnosed

at an advanced stage because of the asymptomatic nature of early-staged OC.¹ Advanced-stage patients have a 5-year relative survival rate of 29%, while this rate was 92% for early-stage disease.³ Despite improvements in surgery, chemotherapy, radiotherapy, targeted therapy and endocrine therapy, the current treatment effect of OC is still not satisfactory. Therefore, further understanding of

Liu, Wu and Wang contributed equally to this work.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *Journal of Cellular and Molecular Medicine* published by Foundation for Cellular and Molecular Medicine and John Wiley & Sons Ltd.

the biology of OC and effective treatment measures is required to improve the clinical outcome of patients with OC.

The anomalous metabolism in cancer was gradually found and regarded as a potential therapeutic approach. Metabolic reprogramming was an emerging hallmark of cancer in 2011⁴ and has been extensively investigated over the past decade. Cancer cells can reprogram their glucose metabolism even in the presence of oxygen, leading to a state termed as 'aerobic glycolysis'.⁴ Besides, the lipid-related pathway is a common and important mechanism in cancer cells for supplying energetic expenditure and membrane synthesis.⁵ This is also applicable to OC cells, and most studies have reported the abnormal lipid metabolism in OC.⁶⁻⁸ Although lipids and some subclasses of nutritional lipids may be associated with epithelial ovarian carcinoma (EOC) risk, the metabolism of lysophosphatidic acid (LPA) and arachidonic acid (AA) emerges as an important signaling network in EOC.⁸ Emerging evidence suggests that tumour cells have much more complex metabolic requirements than previously reported, and metabolic reprogramming is an important strategy of cancer therapy.⁹ Therefore, the molecular subtype of OC based on metabolomic signatures may have great clinical significance.

In this study, we used the expression data of metabolism-relevant genes from the Cancer Genome Atlas (TCGA) database and identified three molecular subtypes of OC, namely C1, C2 and C3. International Cancer Genome Consortium (ICGC) and GEO datasets were used for validation. We also evaluated the prognostic value, transcriptome features, metabolic signatures, immune-cell infiltration and drug sensitivity of the OC subtypes. Finally, a 45-gene classifier was developed and validated.

2 | MATERIALS AND METHODS

2.1 | Data source and processing

Three data repositories, including TCGA (<https://portal.gdc.cancer.gov/>), ICGC (<https://daco.icgc.org/>) and GEO (<https://www.ncbi.nlm.nih.gov/geo/>), were used to obtain required data for OC. RNA sequencing data (raw counts and Fragments per Kilobase Million [FPKM]) of 373 OC human samples with prognostic data were retrieved from the TCGA-OV cohort using R package 'TCGAbiolink'. RNAseq data (FPKM) of 81 OC samples with clinical information were downloaded from the ICGC database. The FPKM data were transformed into transcripts per kilobase million (TPM) values and then log₂ transformed for subsequent analysis. Additional raw microarray data of 260 OC samples from GSE32062 (GPL6480) were downloaded and processed by a robust multi-array average (RMA) algorithm using R package 'GEOquery'. Gene somatic mutation data (MAF files) and copy number data of the TCGA cohort were downloaded from UCSC Xena (<https://xena.ucsc.edu/>). The GSE140082 dataset including the bevacizumab response results was included in the analysis to investigate targeted drug response.

2.2 | Identification of OC subtypes

Previously published 2752 metabolism-relevant genes encoding known human metabolic enzymes and transporters¹⁰ were used as candidate genes for identifying subtypes based on non-negative matrix factorization (NMF) clustering.¹¹ A filtering step was performed on the TCGA cohort before NMF clustering. Genes with a high median absolute deviation (MAD) value (MAD >0.5) across all samples and significant prognostic value ($p < 0.05$, Cox regression using R packages 'survival') were used for NMF clustering. Subsequently, an unsupervised NMF clustering method was used to cluster samples of TCGA and ICGC cohorts using the R package 'NMF'.¹² The value of k where the magnitude of the cophenetic correlation coefficient began to fall was chosen as the optimal number of clusters.¹³ A class mapping (SubMap) analysis (Gene Pattern, <https://www.genepattern.org/>) was used to evaluate the similarity of the subtypes identified in the two aforementioned datasets based on their expression profiles.¹⁴ T-distributed stochastic neighbour embedding (t-SNE)¹⁵ (R packages 'Rtsne') was also used to validate the subtype assignments with the mRNA expression data of the aforementioned metabolic genes.

2.3 | Differential analysis and function enrichment analysis

The differentially expressed genes (DEGs) among OC subtypes were identified using the *limma* package in R on the TCGA-OV raw count data. Absolute log₂FC >1.5 and adjusted $p < 0.01$ were thresholds for DEGs. Subsequently, Gene Set Enrichment Analysis (GSEA) software¹⁶ was used for KEGG pathway enrichment analysis with the expression profile of TCGA-OV. DAVAD^{17,18} website (<https://david.ncifcrf.gov/home.jsp>) was used for GO BP (Gene Ontology, Biological Process) enrichment analysis using the expression profiles of DEGs.

2.4 | Metabolism-relevant pathway gene signature score using gene set variation analysis

Gene set variation analysis (GSVA), a nonparametric and unsupervised gene set enrichment method, can estimate the score of a certain pathway or signature based on transcriptomic data.¹⁹ The 114 metabolism-relevant pathway gene signatures were acquired from previously published studies.²⁰ Each sample received 114 scores corresponding to 114 metabolism pathway signatures using the R package 'GSVA'. Subsequently, differential analysis was conducted based on 114 metabolism scores using the 'limma' package in the R software, and adjusted $p < 0.05$ was used for identifying differentially signatures.

2.5 | Estimation of immune infiltration

Immune scores and stromal scores were calculated using the ESTIMATE algorithm (R package 'estimate'), which could infer the fraction of stromal and immune cells in tumour samples.²¹ We evaluated the abundance of two immune cell populations (T cells and myeloid dendritic cells) and two nonimmune stromal cell population (cancer-associated fibroblasts and endothelial cells) using a micro-environment cell population-counter (MCP-counter)²² in the Timer website. Besides, a single-sample GSEA (ssGSEA) algorithm in the GSVA R package was used to estimate the immune cell infiltration of 28 immune cell populations.²³

2.6 | Generation of classifier and performance validation

The top 15 DEGs with the largest log₂FC value (only DEGs with log₂FC >0 were chosen) in each subtype were selected to develop the prediction model, and a 45-gene classifier was generated. The subtype prediction was conducted based on the 45-gene classifier on the GSE32062 dataset using nearest template prediction (NTP) algorithm (*CMScaller* R package). Subsequently, the prognosis and immune-cell infiltration features of predicted subtypes were analysed to evaluate the performance of classifier and stability of subtypes.

2.7 | Response prediction of each subtype for targeted therapy

The response of bevacizumab to each subtype of the TCGA cohort was predicted based on data derived from the GSE140082 dataset using the SubMap algorithm.

2.8 | Other statistical analyses

R software (4.0.2, <https://www.r-project.org/>) was used for statistical analysis. Log-rank test and Kaplan-Meier methods were used to analyse the prognostic differences in different subtypes. The unpaired Student's *t* test was used to compare the two groups with normally distributed variables, while the Mann-Whitney *U* test was used to compare two groups with the non-normally distributed variables. All scripts were placed on GitHub (https://github.com/Nanahaha/Article_script).

3 | RESULTS

3.1 | NMF identified three molecular subtypes in OC

The study workflow is shown in Figure 1. Previously reported 2752 metabolism-relevant genes were chosen as the basis of NMF analysis. After screening, the expression profile of 109 metabolism-relevant

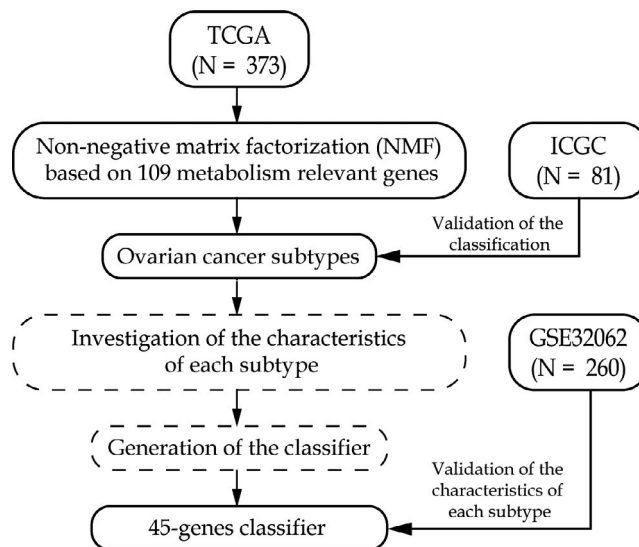


FIGURE 1 The workflow of the study

genes was used to identify molecular subtypes using NMF consensus clustering. As shown in Figure 2A, the optimal number of clusters was $k = 3$ according to the cophenetic correlation coefficients. When $k = 3$, the consensus matrix heatmap still had clear boundaries, suggesting stable and robust clustering for the samples (Figure S1A). T-SNE was performed on the expression profile of the 109 metabolism-relevant genes to validate the assignment, and the three subtypes were clustered separately in a two-dimensional t-SNE distribution patterns (Figure 2B). In addition, the ICGC dataset with 81 OC samples was used to validate the classification based on 109 metabolism-relevant genes and NMF consensus clustering algorithm. Consistently, the results of the ICGC dataset also revealed that of the number of optimized subtypes was also 3, suggesting the robustness and stability of the three subtypes (Figure 2C and Figure S1B). A SubMap analysis was conducted to analyse the correlation of subtypes between TCGA and ICGC datasets. The result showed that C1, C2 and C3 subtypes in the TCGA dataset highly correlated with corresponding subtypes in the ICGC dataset, suggesting the robustness of subtyping across cohorts (Figure S2).

The differences in survival among subtypes were analysed by the Kaplan-Meier method (Figure 2D,E). Significant differences in overall survival among subtypes were found in both the TCGA dataset ($p < 0.0001$) and the ICGC dataset ($p = 0.0055$). The overall survival of the C2 subtype was the worst in both datasets.

3.2 | Transcriptomic characteristics and metabolism-related pathway characteristics of OC subtypes

Subtype-specific genes (DEGs) were identified with the thresholds of the false discovery rate (FDR) <0.01 and $|\log_2\text{FC}| >1.5$. A total of 520, 292 and 43 DEGs were identified in C1, C2 and C3 (Table S1), respectively. Subsequently, the GSEA pathway and GO

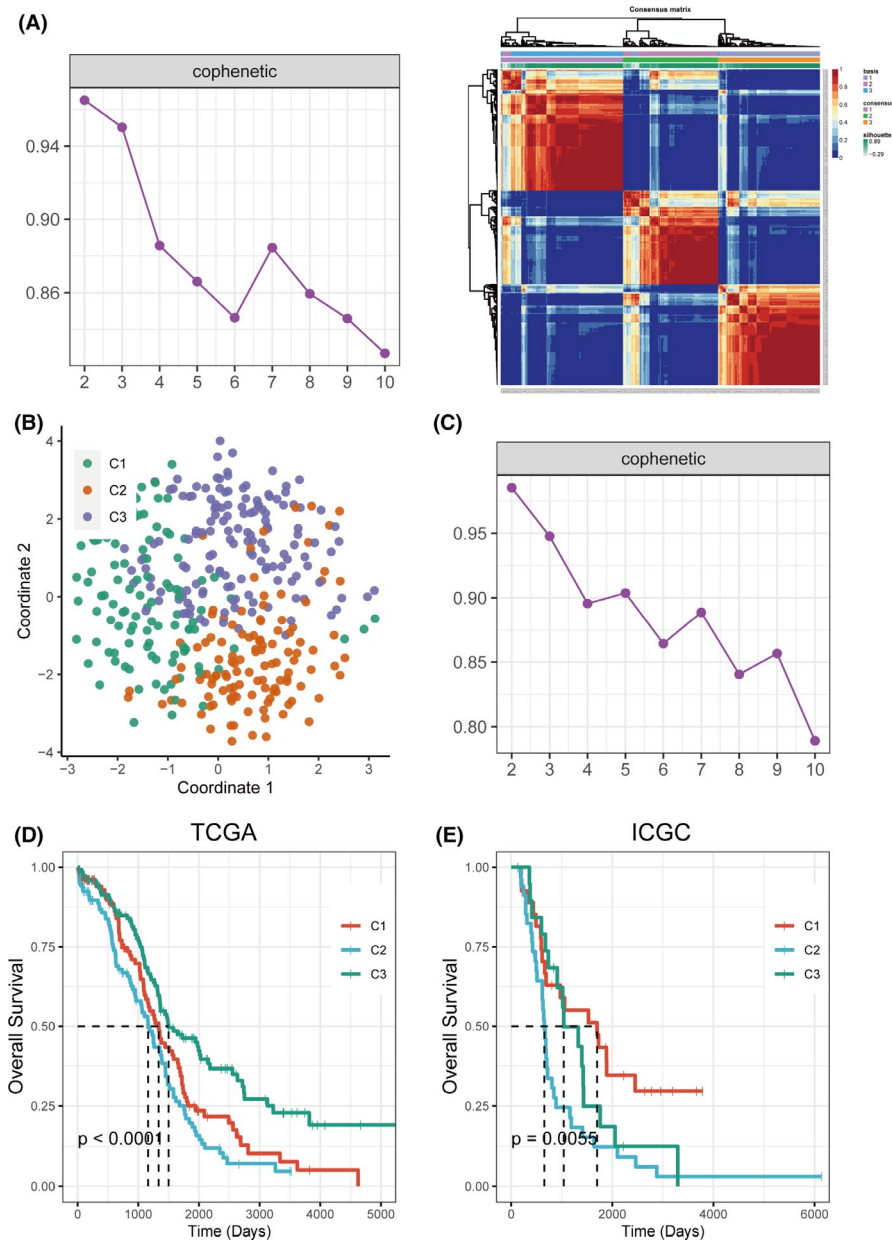


FIGURE 2 Identification of subtypes using NMF consensus clustering based on the metabolism-associated genes. (A) NMF clustering using 109 metabolism-associated genes in TCGA cohort. Cophenetic correlation coefficient for $k = 2-10$ was shown. The consensus matrix heatmap for $k = 3$ was shown. (B) t-SNE analysis scatter plot based on 109 metabolism-associated genes. (C) Cophenetic correlation coefficient of NMF clustering for $k = 2-10$ in ICGC cohort was shown. (D) KM curves showed prognostic relationship of three subtypes in TCGA and ICGC cohort, respectively; the p -value was calculated using the log-rank test, by comparing the overall survival of three subtypes

BP enrichment analyses were performed (Figure S3, Figure S4 and Table S2). Most DEGs of the C1 subtype were downregulated, while those of the C2 subtype were upregulated. Both C1 and C2 subtypes were significantly enriched in the 'cytokine cytokine receptor interaction' and 'hematopoietic cell lineage' pathway, but had the opposite enrichment status. Consistent with this, the GO BP enrichment analysis showed a similar pattern. Gene signatures of C1 and C2 subtypes were mainly enriched in proteolysis, immune-related biological process and Fc receptor-related process, but the regulatory status was also opposite depending on the relative expression status of the genes in GO terms. The C3 subtype was mainly enriched in the phosphorylation pathway and T cell-related biological process but with no statistical significance.

Since the subtyping was based on metabolism-relevant genes, the correlation between OC subtypes and metabolism pathways

was investigated. The metabolic pathways were retrieved from a previous published study.²⁰ The GSVA algorithm was used to score each sample targeted to 114 metabolism pathways. Then, the difference in metabolic pathways between one subtype and the other subtypes was analysed using the *limma* package. The FDR <math>< 0.05</math> was used as the threshold for identifying significantly different metabolic pathways of each subtype. The heatmap of metabolism pathway signatures of the C2 subtype is shown in Figure S5A. The C2 subtype had significantly increased glycan biosynthesis-related pathways and several carbohydrate metabolism- and lipid metabolism-related pathways (eg glycerophospholipid, glycosphingolipid and sphingolipid metabolism) (Figure S5B). Aldosterone, ketone, steroid hormone and testosterone metabolism had the lowest scores in the C2 subtype (Figure S5B).

3.3 | Correlation of OC subtypes with the immune status

Immune therapy is a hotspot of cancer therapy at present. Thus, the immune characteristics of the three subtypes were explored. Firstly, the tumour purity of each subtype was evaluated using the ESTIMATE algorithm (Figure 3A,B), which is used to estimate stromal and immune cells in malignant tumour tissues with expression data. The results showed that the C2 subtype had significantly increased immune and stromal scores than C1 and C3 subtypes, while the C1 subtype had the lowest scores. These results suggested that the subtypes of OC might have different immune responses.

Subsequently, the immune-cell infiltration score was analysed using the MCP-counter algorithm and the ssGSEA algorithm. The infiltration scores of 32 immune cell types were evaluated (Figure 3C, Figure S6). Consistent with previous results, the C2 subtype had a significantly increased immune cell infiltration level (except activated CD4 T cells, activated CD8 T cells, CD56bright natural killer cells, CD56dim natural killer cells and type 2 T helper cells), while the C1 subtype had a significantly decreased immune cell infiltration (except memory B cells). Further, the C3 subtype had a moderate immune cell infiltration (except activated CD4 T cells, activated CD8 T cells, CD56bright natural killer cells, CD56dim natural killer cells, type 2 T helper cells and memory B cells; Figure 3C, Figure S6).

The immune checkpoint is the critical target for immune therapy. The expression of 18 immune checkpoints (PDCD1, CD274, PDCD1LG2, CCL2, CTLA4, CXCR4, IL1A, IL6, LAG3, TGFB1, TNFRSF4, TNFRSF9, TNFSF4, CD4, CD86, CD80, CD276 and VTCN1) was analysed among the three subtypes (Figure 3D). The results showed that the C2 subtype had significantly increased expression of most checkpoint genes except CXCR4, LAG3 and VTCN1, while the C1 subtype had significantly decreased expression except CXCR4 and CD276 (Figure 3D).

Collectively, the three subtypes of OC showed significantly differences in the immune status.

3.4 | Mutations and copy number aberrations of OC subtypes

BRCA is a tumour suppressor gene that plays an important role in regulating cell replication, repair of DNA damage and normal cell growth.²⁴ BRCA mutation leads to the loss of function of tumour growth inhibition.²⁵ We analysed the relationship between BRCA1 and BRCA2 mutations and the three subtypes (Figure 4A). The results showed that the mutation rate of C3 BRCA1 was significantly higher than that of C1 (Fisher's test, $p = 0.025$). The mutation rate of C3 BRCA2 was higher than C1/C2, but with no statistical significance. In terms of somatic copy number aberrations, C2 showed a lower burden of both gains and losses compared with other subtypes, while no statistically significant difference in gain and loss was found between C1 and C3 (Figure 4B).

3.5 | Correlation of OC subtypes with the sensitivity of targeted drugs

Targeted drugs are an effective approach to cancer therapy. Bevacizumab, a monoclonal antibody, targets the vascular endothelial growth factor (VEGF) pathway and inhibits tumour blood vessel growth. Thus, this study evaluated the correlation between bevacizumab responses and OC subtypes. An array expression profile including the bevacizumab response results in OC was downloaded from GEO (GSE140082). A SubMap analysis was conducted to analyse the correlation between the three subtypes and the bevacizumab response. The results showed that the C1 subtype in the TCGA dataset highly correlated with an optimal response to bevacizumab in the GSE140082 dataset, the C2 subtype was highly correlated with a suboptimal response to bevacizumab, while the C3 subtype was secondarily correlated with the optimal response to bevacizumab (Figure 4C). These results suggested that bevacizumab therapeutic strategies might have different responses to subtypes.

3.6 | Forty-five-gene classifier and performance validation

The GSE32062 dataset including 260 samples was downloaded from the GEO database to validate the three subtypes further. Considering clinical application potential, the top informative subtype-associated signature genes including top 15 genes with the largest log₂FC value (>0) in each subtype were selected. The NTP algorithm was used to predict the subtype of the GSE32062 dataset using a 45-gene classifier (Figure 5A, Table S3). Further, 63 samples were classified into the C1 subtype, 86 samples into the C2 subtype and 112 samples into the C3 subtype. We first analysed the difference in overall survival among three predicted subtypes in the GEO dataset and found that the C2 subtype had the worst survival (Figure S7). Next, we analysed the tumour purity of the predicted three subtypes. The results showed that the immune score and stromal scores were significantly higher in the C2 subtype than in other subtypes, and the immune scores were significantly lower in the C1 subtype than in other subtypes (Figure 5B). Besides, the expression distribution of 18 immune checkpoint genes was also analysed among the three subtypes. High expression of most immune checkpoint genes in the C2 subtype (Figure 5D) was observed. The results in the GSE32062 dataset were consistent with those in the TCGA dataset.

4 | DISCUSSION

The studies on the molecular subtyping of OC were mainly based on the overall expression profile. Despite comprehensive transcriptomic information, redundancy and noise were also employed. The importance of metabolism in cancer was gradually realized.⁴ To investigate the association between OC subtypes and metabolic

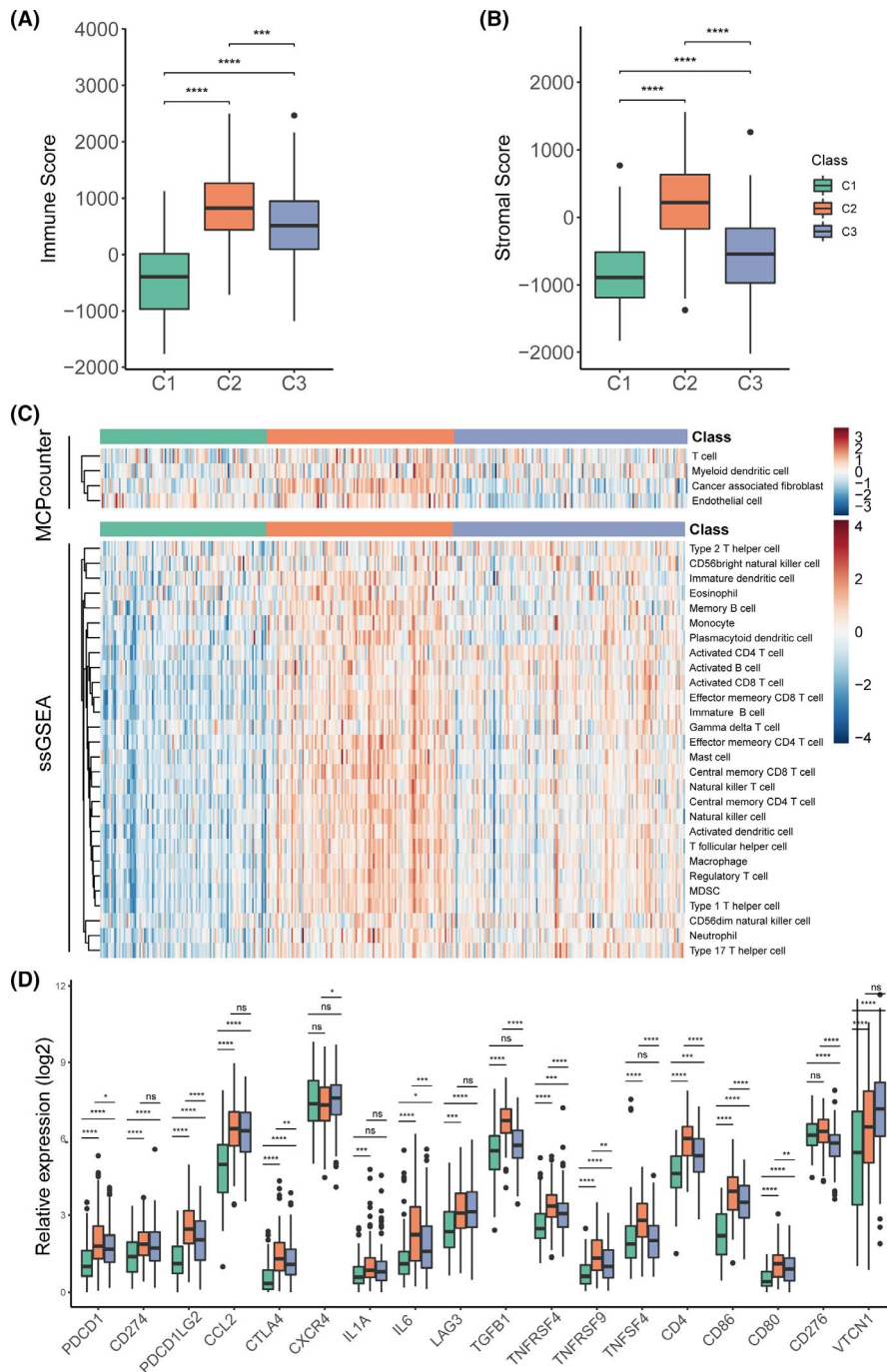


FIGURE 3 Immune characteristics of the three subtypes in the TCGA-OV cohort. Boxplot of immune score (A) and stromal score (B) from ESTIMATE of three subtypes. (C) Heatmap describing the abundance of immune and stromal cell population in C1, C2 and C3. (D) Expression level of 18 immune checkpoint genes in three OC subtypes. The statistical difference was compared by wilcox.test, and adjusted by 'holm' method (ns represents no significance, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$)

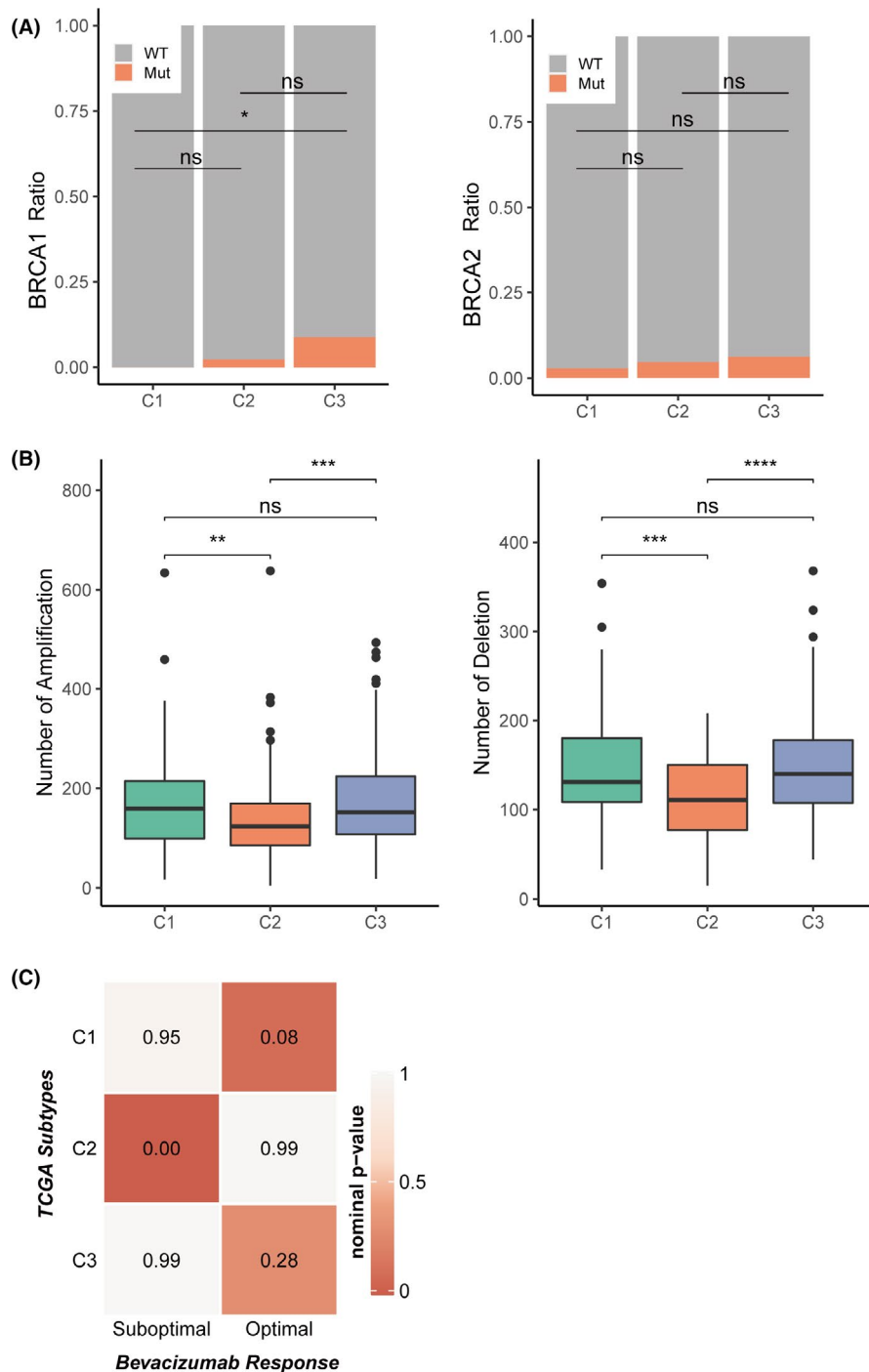
processes, OC subtypes were established in this study based on 2752 metabolic genes previously studied,¹⁰ which provided more details about the metabolomic landscape of OC.

In this study, three subtypes (C1, C2 and C3) of OC were identified based on metabolism-relevant genes in the TCGA-OV cohort. The robustness of the subtype number was verified in the external dataset ICGC-OV. A classifier consisting of 45 genes was evaluated in the external dataset GSE32062. The transcriptome features, metabolic pathway signatures, prognosis, immune infiltration and targeted drug response of the three subtypes were explored. Overall, C1 and C2 subtypes showed significant differences in most molecular characteristics.

The mutation rate of BRCA1 was significantly higher in the C3 subtype compared with the other subtypes. BRCA1/2 absence or mutation increased the response sensitivity to DNA cross-linkers (eg cisplatin), and it has been shown that cisplatin and paclitaxel combination chemotherapy benefited patients with BRCA1 mutation more,²⁶ which was consistent with the fact that C3 subtype with the highest mutation rate of BRCA1 had a good prognosis in the present study. In terms of copy number aberrations, C2 patients showed a lower burden of gains and losses, suggesting the stability of the C2 genome.

Previous studies demonstrated that metabolic alterations were existed in OC,²⁷ however, the metabolic heterogeneity was rarely

FIGURE 4 The association between three subtypes and mutations, copy number aberrations and targeted drug response in TCGA-OV cohort. (A) BRCA1 and BRCA2 mutation type/wild-type ratio in three subtypes. (C) Boxplot of number of copy number aberrations. (D) Correlation between three subtypes and bevacizumab response. Between three subtypes. The statistical difference was compared by wilcox.test, and adjusted by 'holm' method (ns represents no significance, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$)



reported. The results of the present study indicated that each subtype had unique metabolism signatures. The C2 subtype was chiefly involved in glycan biosynthesis and metabolism-relevant signatures, carbohydrate metabolism and lipid metabolism processes. Among the signatures, glycans, which decorated all eukaryotic cell surfaces, underwent changes in structure with the onset of diseases such as cancer and inflammation. Several glycan-based vaccines are presently undergoing clinical evaluation with some encouraging preliminary results.²⁸ The expression of Sialyl-Tn (STn), a type of O-glycans, was increased in cancer, and the prognosis was poor in patients with cancer.²⁹ Our results

revealed that the C2 subtype had a high enrichment score in glycan biosynthesis and metabolism-relevant signatures compared with the other two subtypes, suggesting that the C2 subtype might be more sensitive to glycan-based vaccines. Lipid metabolism alteration is an important characteristic of OC.³⁰ Our results showed that several lipid metabolism-relevant signatures (eg cyclooxygenase arachidonic acid, glycerolipid, fatty acid biosynthesis, glycerophospholipid, glycosphingolipid, sphingolipid) had higher enrichment scores in the C2 subtype compared with the other subtypes. Aberrant lipid metabolic genes were found to be associated with worse clinical outcomes. For example, fatty acid

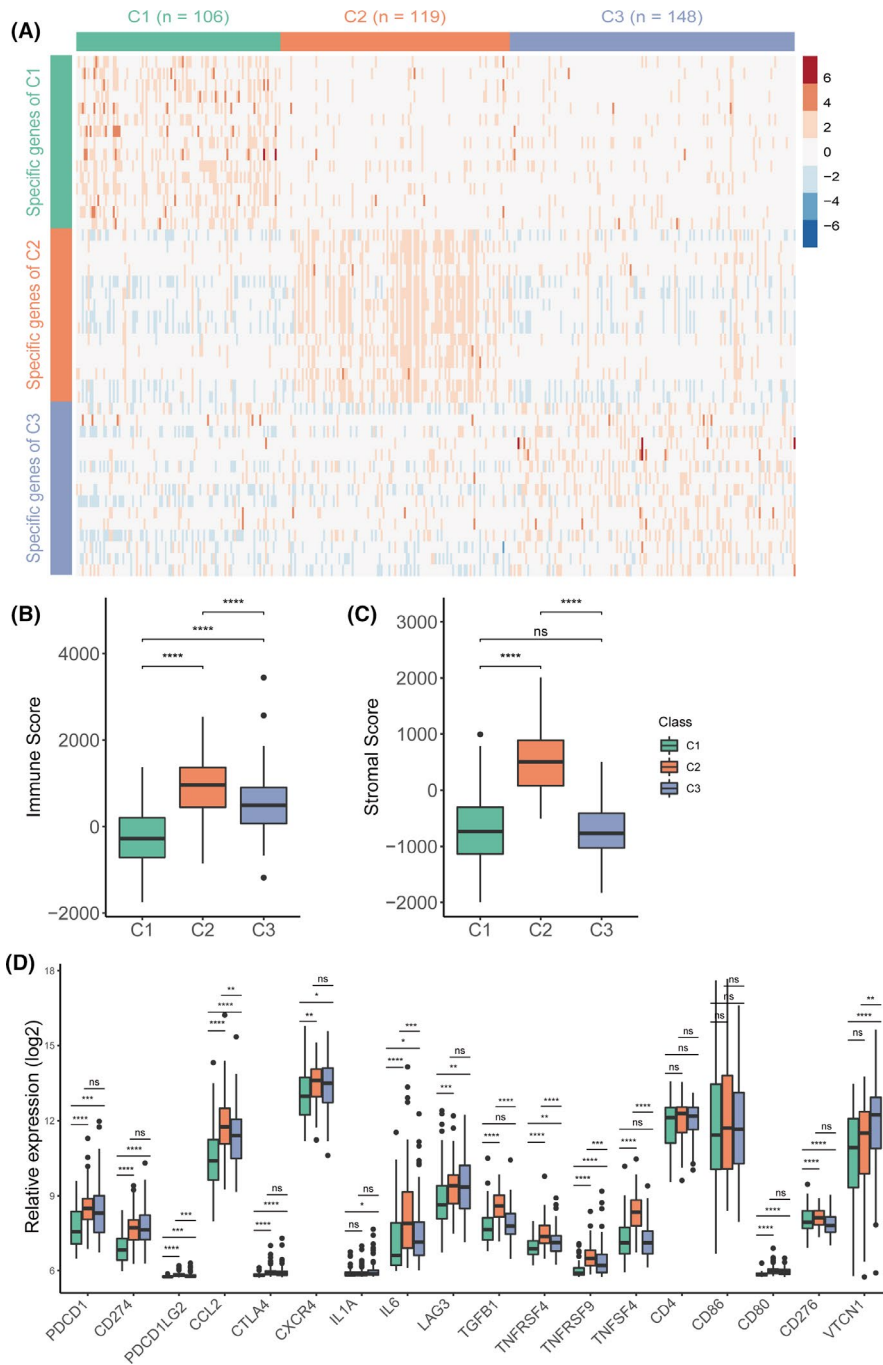


FIGURE 5 Validation of external dataset. (A) Heatmap of the expression level of the 45-classifier in TCGA-OV cohort. Boxplot of immune score (B) and stromal score (C) from ESTIMATE of three predicted subtypes in the validation set. (D) Expression distribution of 18 checkpoint genes in three subtypes in the validation set. The statistical difference was compared by wilcox.test, and adjusted by 'holm' method (ns represents no significance, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$)

synthase (FASN) is highly expressed in OC tissues and is associated with poor prognosis.³¹ The increased levels of ASAHI1 and sphingosine-1-phosphate (S1P), which participated in sphingolipid metabolism, were associated with the poor survival of OC.^{27,30} These reports might explain why the C2 subtype had the worst overall survival among the subtypes.

Previous studies suggested that immune checkpoints played a crucial role in the immune escape of cancer.³² The monoclonal antibodies developed for PD-1 and its ligands as well as CTLA4 have been approved for the clinical therapy of several cancers (eg melanoma and non-small-cell lung carcinoma).^{33,34} Thus, the difference in the expression of immune checkpoint genes among the subtypes

was evaluated. The C2 subtype had the highest level immune cell infiltration. An OC subtype classification based on immune genes reported the worst prognosis in one subtype with abundant immune infiltration,³⁵ which was consistent with the result of the present study. PD-1, which was upregulated in the C2 subtype in this study, inhibits excessive immune responses and protects normal cells from immune attack.³⁶ High PD-1+ expression of lymphocytes was associated with the poor survival of OC.³⁷ Regulatory CD4+ T (Treg) cells and myeloid-derived suppressor cells (MDSCs) were also significantly enriched in C2. Consistently, regulatory CD4+ T (Treg) cells mediated homeostatic peripheral tolerance by suppressing autoreactive T cells.³⁸ A previous study showed that

tumour Treg cells were associated with a high death hazard and reduced survival in OC.³⁸ Baert et al.³⁹ reported that MDSCs, which belonged to the innate immune system, were the key drivers of immunosuppression in OC. These findings might also explain why the C2 subtype had a high degree of immune cell infiltration but a poor prognosis. Although no immune therapies are currently approved for OC,⁴⁰ our subtyping results might provide some information for OV immune therapy.

This study first established the molecular subtypes of OC using a metabolism-relevant gene expression profile and investigated the characteristics of subtypes. However, the study had some limitations. First, the number of metabolism-relevant genes used to identify subtypes was small. Second, the number of genes in classifiers was also small. Third, no difference in prognosis was found in the GEO32062 validation set, but was consistent with the survival trends in the TCGA dataset. Fourth, the survival of subtypes in the TCGA and ICGC datasets was inconsistent. This might be due to the small sample size of the ICGC dataset or the inherently small difference in survival rates between subtypes. The worst survival rate for the C2 subtype was the most important result in all datasets.

5 | CONCLUSION

This study identified three subtypes in OC using gene expression profiles of metabolism-relevant genes through large databases of TCGA, ICGC and GEO. The three subtypes exhibited unique characteristics in terms of prognosis, transcriptome, immune and mutation statuses, and copy number. These findings may shed new light on individual therapeutic strategies for OC.

ACKNOWLEDGEMENTS

This research was supported by the National Natural Science Foundation of China (#91846302). Thanks to the Medical Research Data Center of Fudan University.

CONFLICT OF INTEREST

The authors confirm that there are no conflicts of interest.

AUTHOR CONTRIBUTIONS

Lei Liu: Funding acquisition (lead) and Supervision (lead). **Xiaona Liu:** Conceptualization (equal); Formal analysis (equal); Methodology (equal); and Writing—original draft (equal). **Aoshen Wu:** Formal analysis (equal); Software (equal); and Writing—original draft (equal). **Xing Wang:** Conceptualization (equal); Formal analysis (equal); and Writing—original draft (equal). **Yunhe Liu:** Methodology (equal). **Yiang Xu:** Investigation (equal). **Gang Liu:** Project administration (lead) and Writing—review & editing (lead).

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analysed in this study.

ORCID

Lei Liu  <https://orcid.org/0000-0002-9995-9080>

REFERENCES

- Lheureux S, Gourley C, Vergote I, Oza AM. Epithelial ovarian cancer. *Lancet*. 2019;393(10177):1240-1253.
- Horikawa N, Abiko K, Matsumura N, et al. Expression of vascular endothelial growth factor in ovarian cancer inhibits tumor immunity through the accumulation of myeloid-derived suppressor cells. *Clin Cancer Res*. 2017;23(2):587-599.
- Reid BM, Permuth JB, Sellers TA. Epidemiology of ovarian cancer: a review. *Cancer Biol Med*. 2017;14(1):9-32.
- Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell*. 2011;144(5):646-674.
- Buckley D, Duke G, Heuer TS, et al. Fatty acid synthase – modern tumor cell biology insights into a classical oncology target. *Pharmacol Ther*. 2017;177:23-31.
- Liu X, Liu G, Chen L, et al. Untargeted metabolomic characterization of ovarian tumors. *Cancers*. 2020;12(12):3642.
- Sans M, Gharpure K, Tibshirani R, et al. Metabolic markers and statistical prediction of serous ovarian cancer aggressiveness by ambient ionization mass spectrometry imaging. *Cancer Res*. 2017;77(11):2903-2913.
- Reinartz S, Finkernagel F, Adhikary T, et al. A transcriptome-based global map of signaling pathways in the ovarian cancer microenvironment associated with clinical outcome. *Genome Biol*. 2016;17(1):108.
- Boroughs LK, DeBerardinis RJ. Metabolic pathways promoting cancer cell survival and growth. *Nat Cell Biol*. 2015;17(4):351-359.
- Possemato R, Marks KM, Shaul YD, et al. Functional genomics reveal that the serine synthesis pathway is essential in breast cancer. *Nature*. 2011;476(7360):346-350.
- Lee DD, Seung HS. Learning the parts of objects by non-negative matrix factorization. *Nature*. 1999;401(6755):788-791.
- Gaujoux R, Seoighe C. A flexible R package for nonnegative matrix factorization. *BMC Bioinform*. 2010;11:367.
- Brunet JP, Tamayo P, Golub TR, Mesirov JP. Metagenes and molecular pattern discovery using matrix factorization. *Proc Natl Acad Sci USA*. 2004;101(12):4164-4169.
- Hoshida Y, Brunet JP, Tamayo P, Golub TR, Mesirov JP. Subclass mapping: identifying common subtypes in independent disease data sets. *PLoS One*. 2007;2(11):e1195.
- Hinton LvdMG. Visualizing data using t-SNE. *J Mach Learn Res*. 2008;9:2579-2605.
- Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA*. 2005;102(43):15545-15550.
- da Huang W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*. 2009;4(1):44-57.
- da Huang W, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res*. 2009;37(1):1-13.
- Hänzelmann S, Castelo R, Guinney J. GSEA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinform*. 2013;14:7.
- Rosario SR, Long MD, Affronti HC, Rowsam AM, Eng KH, Smiraglia DJ. Pan-cancer analysis of transcriptional metabolic dysregulation using the cancer genome atlas. *Nat Commun*. 2018;9(1):5330.
- Yoshihara K, Shahmoradgoli M, Martínez E, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat Commun*. 2013;4:2612.

22. Becht E, Giraldo NA, Lacroix L, et al. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome Biol.* 2016;17(1):218.
23. Jia Q, Wu W, Wang Y, et al. Local mutational diversity drives intratumoral immune heterogeneity in non-small cell lung cancer. *Nat Commun.* 2018;9(1):5361.
24. Mahdavi M, Nassiri M, Kooshyar MM, et al. Hereditary breast cancer; genetic penetrance and current status with BRCA. *J Cell Physiol.* 2019;234(5):5741-5750.
25. Bulanova DR, Helenius M, Sokolenko AP, Kuznetsov SG, Imyanitov EN. Response to: the GPRC5A frameshift variant c.183del is not associated with increased breast cancer risk in BRCA1 mutation carriers. *Int J Cancer.* 2019;144(7):1758-1760.
26. Zhang J, Lin Y, Sun XJ, et al. Biomarker assessment of the CBCSG006 trial: a randomized phase III trial of cisplatin plus gemcitabine compared with paclitaxel plus gemcitabine as first-line therapy for patients with metastatic triple-negative breast cancer. *Ann Oncol.* 2018;29(8):1741-1747.
27. Ryland LK, Fox TE, Liu X, Loughran TP, Kester M. Dysregulation of sphingolipid metabolism in cancer. *Cancer Biol Ther.* 2011;11(2):138-149.
28. Dube DH, Bertozzi CR. Glycans in cancer and inflammation-potential for therapeutics and diagnostics. *Nat Rev Drug Discov.* 2005;4(6):477-488.
29. Brockhausen I. Mucin-type O-glycans in human colon and breast cancer: glycodynamics and functions. *EMBO Rep.* 2006;7(6):599-604.
30. Braicu EI, Darb-Esfahani S, Schmitt WD, et al. High-grade ovarian serous carcinoma patients exhibit profound alterations in lipid metabolism. *Oncotarget.* 2017;8(61):102912-102922.
31. Cai Y, Wang J, Zhang L, et al. Expressions of fatty acid synthase and HER2 are correlated with poor prognosis of ovarian cancer. *Med Oncol.* 2015;32(1):391.
32. Marin-Acevedo JA, Dholaria B, Soyano AE, Knutson KL, Chumsri S, Lou Y. Next generation of immune checkpoint therapy in cancer: new developments and challenges. *J Hematol Oncol.* 2018;11(1):39.
33. Hegde PS, Chen DS. Top 10 challenges in cancer immunotherapy. *Immunity.* 2020;52(1):17-35.
34. Hirsch FR, Scagliotti GV, Mulshine JL, et al. Lung cancer: current therapies and new targeted treatments. *Lancet.* 2017;389(10066):299-311.
35. Zheng M, Hu Y, Gou R, et al. Identification of immune-enhanced molecular subtype associated with BRCA1 mutations, immune checkpoints and clinical outcome in ovarian carcinoma. *J Cell Mol Med.* 2020;24(5):2819-2831.
36. Tu L, Guan R, Yang H, et al. Assessment of the expression of the immune checkpoint molecules PD-1, CTLA4, TIM-3 and LAG-3 across different cancers in relation to treatment response, tumor-infiltrating immune cells and survival. *Int J Cancer.* 2020;147(2):423-439.
37. Chatterjee J, Dai W, Aziz NHA, et al. Clinical use of programmed cell death-1 and its ligand expression as discriminatory and predictive markers in ovarian cancer. *Clin Cancer Res.* 2017;23(13):3453-3460.
38. Curiel TJ, Coukos G, Zou L, et al. Specific recruitment of regulatory T cells in ovarian carcinoma fosters immune privilege and predicts reduced survival. *Nat Med.* 2004;10(9):942-949.
39. Baert T, Vankerckhoven A, Riva M, et al. Myeloid derived suppressor cells: key drivers of immunosuppression in ovarian cancer. *Front Immunol.* 2019;10:1273.
40. Odunsi K. Immunotherapy in ovarian cancer. *Ann Oncol.* 2017;28:viii1-viii7.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Liu X, Wu A, Wang X, et al.

Identification of metabolism-associated molecular subtype in ovarian cancer. *J Cell Mol Med.* 2021;25:9617-9626. <https://doi.org/10.1111/jcmm.16907>