Identification of a gene set correlated with immune status in ovarian cancer by transcriptome-wide data mining

Lili Fan

Central South University

Han Lei

Central South University

Ying Lin

Central South University

Tianxiang Zhang

Central South University

Zhengwei Zhou

Central South University

Guang Shu

Central South University

Zhipeng Yan

Central South University

Haotian Chen

Central South University

Gang Yin (✉ gangyin@csu.edu.cn)

Central South University Xiangya School of Medicine

Research

Keywords: ovarian cancer, TMB, immune cells infiltration, survival prognosis.

Posted Date: June 18th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-35960/v1>

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

[Read Full License](#)

Abstract

Background: Checkpoint block-mediated tumor immunotherapy has obtained huge success in treating various malignancies. However, the efficacy of immunotherapies against ovarian cancer (OC) was low largely due to limited information on the tumor immune microenvironment (TIME) of OC. This results in lacking reliable biomarkers to select the right patients for immunotherapy and prognosis prediction. Published studies have reported that high tumor mutation burden (TMB) can generate many neoantigens resulting in a higher degree of an anti-tumor immune response. However, the correlation between TMB and TIME of OC remains controversial let along a collection of TMB information is time and resource consuming. In this study, we considered the TMB and TIME of OC comprehensively to categorize patients based on their tumor local immune status and provide information that could guide further OC immunotherapy trials as well as the prognosis prediction.

Method: OC RNAseq data were downloaded from The Cancer Genome Atlas (TCGA) and divided into TMB^{high} and TMB^{low} subgroups according to TMB scores for comparison. OC cases were also grouped into immunity^{high} or immunity^{low} based on their profiles of tumor-infiltrating leukocytes (TILs) analyzed by ssGSEA (Single Sample Gene Set Enrichment Analysis). Besides, the function of genes enriched in both TMB^{high} and immunity^{high} was analyzed by Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes pathway (KEGG), and protein-protein interaction (PPI) network. The correlation of each gene with overall survival was also analyzed.

Results: Positive association was observed between the OC grade and the TMB. Both TMB^{high} and immunity^{high} were significantly correlated with a better prognosis. However, different from the findings of other studies, TMB of OC didn't correlated with preferable TIME in our analysis. By comparing the up-regulated signature genes in TMB^{high} and immunity^{high} cases, we found 14 overlapped genes that were mainly involved in immune response-related pathways. We further analyzed the prognostic value of these 14 genes and found the upregulation of 4 of them, CXCL13, FCRLA, PLA2G2D, and MS4A1 are significantly associated with better survival. With available data collected form a melanoma cohort, we also found that these four genes are positively associated with better response to immune checkpoint blockade-based immunotherapy.

Conclusion: By comparing signature genes enriched in TMB^{high} and immunity^{high} subgroups, four genes, CXCL13, FCRLA, PLA2G2D, and MS4A1, were found positively correlated with both OC immune infiltration and better prognosis. Additionally, these four genes predicted the response to checkpoint blockade-based immunotherapy in a melanoma cohort indicating that they can also be used as biomarkers for further immunotherapy clinical trials of OC.

Introduction

Ovarian cancer (OC) is the most fatal gynecological tumor around the world. Due to the lack of specific symptoms or effective screening strategies[1], more than 70% of patients are in the middle and late

stages at the time of diagnosis resulting in an increased risk of recurrence and poor prognosis[2]. Despite having various treatments, the five-year survival rate of late-stage patients was no more than 35%[3]. Traditional therapies to unsuitable patients, including chemotherapy and targeted therapy, up to 70% of patients with ovarian cancer still relapse after 12 to 18 months[4], highlighting the importance of exploring novel therapies.

As a rising novel treatment, PD-1/PD-L1 blockade based immunotherapy has demonstrated promising efficacy in treating various types of solid tumors[5]. Ample of clinical data indicate that pre-existing favorable tumor immune microenvironment (TIME), containing antigen-presenting cells, effector T cells, Th1 type cytokines and chemokines, and expression of PD-L1 are prerequisites for PD-1/PD-L1 blocking antibody-mediated immunotherapy. Current clinical trials on OC has shown that this immunotherapy can only benefit a small part of the patient population while a large number of patients have either limited or no response[6, 7]. Those nonresponsive cases are largely due to a far less understanding of the TIME of OCs that led to 1) missing targetable immune checkpoint molecules that are highly expressed in OC cells and inhibit anti-tumor immune response, or 2) lacking specific biomarkers to select the right patient population. Thus, it's critical to analyze the TIME of OC extensively to categorize patients based on their tumor immune cell profiles and to identify potential therapeutic targets and/or molecular biomarkers to facilitate the design of novel immunotherapies.

Tumor mutation burden (TMB) has been demonstrated as a potential biomarker to predict responses to immunotherapy in several types of solid tumors[8-10]. It was defined as the total number of replacement and insert/deletion mutations per mega-base in the exon coding region of the genome examined in tumor samples. It's proposed that tumors with high TMB may produce more neoantigens that can be recognized by the immune system and in turn elicit broader anti-tumor immune response[11]. Consistently, melanoma patients with higher TMB tumors manifested better clinical response to immune checkpoint inhibitors and long-term survival[12]. However, the clinical application of TMB as a prognostic marker was both time and financial consuming and the predictive significance of TMB has been questioned by some other studies[13, 14].

TIME has been considered to have essential impacts on tumor immunity[15-18]. It is composed of cancer cells and a variety of stromal cells, cytokines, chemokines, etc. Favorable TIME, containing antigen-presenting cells, effector T cells, Th1 type cytokines, and chemokines, is a positive factor for immunotherapy[19, 20]. However, the roles of the TMB combined with immunoactivity-related gene sets in predicting the biomarkers for prognosis of OC was unclear. Thus considering the correlation between the TMB, the immune cell types and the efficacy of cancer treatment, so the TMB and the immunoactivity-related gene sets were used to predict the biomarkers associated with the prognosis of OC was the focus of our studies. Therefore, our findings will accelerate the validation and the realization of TMB and TIME as the significant predictive markers for OC's clinical and prognosis. Altogether, understanding the relationship between the combination of TMB and immune microenvironment and the survival prognosis and clinical outcomes of certain malignancies was very favorable and promising for future

immunotherapy and could provide a mechanistic explanation and treatment direction for the survival and prognosis of clinical research.

Genome-wide gene expression profiling provides an extremely valuable tool to unwind complicated biological processes. In this study, we sought to identify prognostic biomarkers for OC by an integrated TMB and TIME analysis using the OC expression profile data and mutation annotation files downloaded from The Cancer Genome Atlas (TCGA) database. All OC cases were classed into TMB^{high} group and TMB^{low} group based on the OC's somatic mutation data. Additionally, we also used ssGSEA (Single Sample Gene Set Enrichment Analysis) scores of immune cell types to conduct unsupervised clustering of OC patients by the profiles of tumor-infiltrating leukocytes (TILs) and obtained three different immune infiltrating clusters, here some of them were defined as the low immunity group and the high immunity group. We further compared upregulated genes enriched in either TMB^{high} or immunity^{high} patients and found 14 genes were shared by both of these two subgroups. We further sorted the function of these genes by Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes pathway (KEGG), protein-protein interaction (PPI) networks, and found that they were involved in various immune-related pathways. Immune infiltration and survival analysis indicated that four of these 14 genes, CXCL13, FCRLA, PLA2G2D, and MS4A1, were significantly correlated with better prognosis. Data of a published melanoma cohort also suggested that these four genes were associated with better response to immune checkpoint blockade-based immunotherapy. Altogether, our analysis strongly indicated that those four signature genes can be applied to predict prognosis and select OC patients for anti-PD-1/PD-L1 treatment (Figure1).

Methods

Database

RNAseq and mutation annotation files for OC patients were gained from the TCGA database portal (<https://portal.gdc.cancer.gov/>). Clinical data such as staging type, survival status and time were also downloaded from the TCGA database.

Implementation of ssGSEA

We obtained the marker gene sets for immune cell types from the TCGA database. Then we used the ssGSEA program to derive the enrichment scores of each immune-related term. The ssGSEA applies gene signatures expressed by immune cell populations to individual cancer samples. The computational approach used in our study included immune cell types that are involved in innate immunity and adaptive immunity. Tumors with qualitatively different immune cell infiltration patterns were grouped using hierarchical clustering.

Heatmaps and clustering analysis

Heatmaps and clustering were generated using the R package.

Identification of differentially expressed genes (DEGs)

TMB group: Top 200 upregulated genes of OC patients were selected by R package limma, $FDR < 0.05$, and $|\log_2 \text{fold change (FC)}| \geq 1$. A total of 427 DEGs consisting of all upregulated genes of OC patients were selected by R package limma, $FDR < 0.05$, and $|\log_2 \text{fold change (FC)}| \geq 1$.

Functional annotation and pathway analysis of DEGs

The online tool, Database for Annotation, Visualization and Integrated Discovery (DAVID, <https://david.ncifcrf.gov/>) provided full-scale information for a list of genes by GO and KEGG pathway analysis. The ClusterProfiler and enrich plot packages to predict the function of DEGs. GO and KEGG pathway analysis with the standard of $p < 0.05$, the results were downloaded in text format. Gene set variation analysis (GSVA) is an open-source software package for R and can be downloaded on the website, then condensing information from Up-DEGs and Down-DEGs expression profiles respectively into KEGG pathway. Compared with single-gene analysis, the advantages of this method include noise and dimension reduction[21]. Moreover, Unsupervised hierarchical clustering of these top KEGG pathways was able to distinguish between normal ovaries group (GTEx datasets) and OC (TCGA datasets).

PPI Networks Construction and Module Analysis

In order to analyze the links between these proteins, DEGs were uploaded to the Search Tool for the Retrieval of Interacting Genes (STRING, <https://string-db.org/>). Only single networks with more than ten nodes were included for further analysis.

Survival Analysis of Hub Genes

Download the datum matrix containing all ovarian patients' prognosis information from the GEO database. The data was processed by GraphPad Prism 8 and then exported the results. Kaplan-Meier plots were drawn to demonstrate the relationship between patients' overall survival and gene expression levels of DEGs. The relationship was tested by the log-rank test.

Kaplan-Meier plotter [Ovarian Cancer]

The online Kaplan-Meier plotter (<http://www.kmplot.com>) was used to assess the effect of four selected characteristic genes on ovarian cancer prognostic. The ovarian cancer patients were split into high and low expression groups by mRNA expression of auto-selected best cutoff. The Kaplan-Meier survival plot showed Hazard ratio (HR), 95% confidence intervals (CI) log-rank p -value for each corresponding gene, and the number at risk below each plot.

TIMER (Tumor Immune Estimation Resource)

TIMER (<https://cistrome.shinyapps.io/timer/>) is an interactive web server that can systematically analyze immune infiltrates in cancer with its own algorithm. We analyzed the correlation between four selected characteristic genes and the abundance of immune infiltrates in OC, including purity, B cells, CD4+ T cells,

CD8+ T cells, by gene modules. The correlation analysis is the Spearman's correlation statistical significance with $p < 0.05$ was indicated with *, $p < 0.01$ was indicated with **, $p < 0.001$ were indicated with *** and $p < 0.0001$ were indicated with ****.

Statistical and computational analyses

SPSS 20.0 software was used for statistical analysis. $p < 0.05$ was considered statistically significant in all statistical tests. All the statistical and computational analyses were performed using the R statistical software environment.

Results

TMB scores and the immunoactivity-related gene sets are conjoined with OC's overall survival

To answer if a high load of somatic mutation is associated with enhanced tumor immune cell infiltration and result in prolonged overall survival, datasets comprising RNAseq from 379 TCGA tumors and mutation annotation files from 436 TCGA tumors (272 which overlapped with RNAseq data) were obtained from the TCGA Data Portal in March 2019 (<https://portal.gdc.cancer.gov/>). We evaluated all OC cases with complete gene expression data and clinical information in TCGA. For mutation annotation files, we presented the Word Cloud of gene symbols identified in OC which the mutation sites > 5 based on the WordArt (<https://wordart.com>). We found that *TP53* was the most significant (90% cases) mutated gene[22] (**Figure S1A**). We then screened the top 25 genes and employed the DISCOVER (Discrete Independence Statistic Controlling for Observations with Varying Event Rates)[23] algorithm to examine co-occurrence and mutual exclusion among these genes and found that *TTN* and *LRP2* are significantly mutual exclusive ($p < 0.001$) (**Figure S1B**). In order to better visualized the mutually exclusive genes pairs, we selected 420 (96.33%) of 436 samples and showed in a waterfall plot that the most frequently mutated genes were *TP53* and *TTN* ($p < 0.05$, FDR < 0.01) (**Figure S1C**). These analyses of OC samples' mutation can be used to guide further TMB analysis.

Based on the somatic mutation data across all the OC cases in this study, TMB scores range from 0.0263 to 6.5260. Interestingly, the TMB scores of OC cases at Grade 3 & Grade 4 were significantly higher than those of Grade 1 & Grade 2 OC cases (**Figure 2A**, $n=417$, except for the blank values, $p=0.048$). And then based on the ssGSEA scores of immune cell types, we divided the immunoactivity-related gene sets into three immune infiltration clusters by unsupervised clustering, and some of which were defined as immunity^{high} clusters ($n=193$) and immunity^{low} clusters ($n=149$) (**Figure 2B**). Besides, the tumor purity in immunity^{high} group was significantly lower than that in immunity^{low} group validating the reliability of our classification (**Figure S1D**). Moreover, the PD-L1 expression level of the immunity^{high} group is significantly higher than that of immunity^{low} group (**Figure 2B**). We further accessed if the TMB or immune infiltration profile could affect the overall survival. Kaplan-Meier survival curves showed that the overall survival time of TMB^{high} or immunity^{high} group was significantly longer than TMB^{low} or immunity^{low} group (**Figure 2C** 2.58 years vs. 1.92 years, $p=0.0295$ in the log-rank test. **Figure 2D** 2.50

years vs. 2.45 years, $p=0.00082$ in log-rank test). We then asked if there is any association between the TMB group and the immunity group. However, we did not see any correlation between these two factors across 379 OC cases (**Figure 2E**).

We further studied the relationship between the immunity group and the immune cell infiltration in OC. So We further studied the abundance of 22 different immune cells subpopulations in the high&low immunity group or the TMB group of cancers. However, only memory B cells and NK cells were increased in TMB^{high} group while there are seven immune cell subpopulations were enriched in high immune scores group, that was most of these immune cells have a higher infiltration degree in immunity^{high} group than in immunity^{low} group of OC (**Figure S2A and 2B**).

KEGG pathway and GO term analysis of DEGs associated with OC.

Next, as to reveal the correlation between the TMB, the immunoactivity-related gene sets, and global gene expression profiles, and wonder if individual genes that are upregulated in both TMB^{high} and immunity^{high} groups are involved in anti-tumor related biological pathways. To this end, we overlapped TMB^{high} gene expression signature (top 200 genes) with immunity^{high} gene expression signature (427 genes) by Venn diagrams and finally obtained 14 differentially expressed genes (DEGs) that may be closely predicted the response to checkpoint blockade-based OC immunotherapy (**Figure 3A**). Heatmaps of the expression levels of 14 up-regulated DEGs in the TMB group and immunity group were shown in **Figure S2C**. Next, we conducted unsupervised hierarchical clustering of KEGG pathways related to Up-DEGs (**Figure 3B**) and found that most of these KEGG pathways are immune-related, such as “Cytokine cytokine receptor interaction”, “Chemokine signaling pathway”, as well as “Primary immunodeficiency”. Moreover, in order to summarize the potential function of the DEGs, we performed functional enrichment analysis of the 14 up-regulated genes between TMB^{high} and immunity^{high} group. GO terms and KEGG pathway analysis also demonstrate that the Up-DEGs function in the immune and inflammatory response, cytokine activities and T cell proliferation (**Figure 3C and 3D**). These results indicated that the gene sets shared by both TMB^{high} and immunity^{high} groups have active immune functions.

Protein-protein interactions among genes associated with prognosis in OC

In order to understand the interactions between the identified DEGs, we performed a PPI networking study using the Search Tool for the Retrieval of Interacting Gene (STRING) tool. All 14 DEGs were uploaded to STRING as a whole Up-DEGs module (**Figure 4**) to construct PPI networks. A network consisting of 9 nodes and 10 edges were generated when the minimum required interaction score was set 1.5 and disconnected nodes were hidden in the network in the STRING website (**Figure 4A**). By counting the number of connections (**Figure 4B**), it shows that CXCL13 has the most key link with other members of the module in the network. Besides CXCL13, several key genes associated with T cell proliferation, inflammation, and immune response were located at the center of the module, such as PLA2G2D and FCRLA. These results indicated that PLA2G2D, MS4A1, ADAMDEC1, FCRLA, CXCL13 interact more with

other molecules, suggesting that these genes are more likely to be key genes in regulating tumor restricting functions in OCs.

Correlation of expression of individual DEGs in overall survival

Since both TMB^{high} group and immunity^{high} group were correlated with better survival, we reasoned that there could be potential prognosis predicting biomarkers among the 14 shared DEGs defined above. To this end, we analyzed the association of individual DEGs with overall survival time. Among them, four genes, including PLA2G2D, CXCL13, FCRLA, and MS4A1, were validated to be significantly relevant to good survival and prognosis outcomes (log-rank test, $p < 0.05$) (**Figure 5A-D**). To validate our results, we downloaded and analyzed gene expression data from a different cohort from the GEO database. Meta-analysis further indicated that OC cases that have high expression of these four genes showed a significant survival benefit and all the four genes are independent predictors for the prognosis of OCs (**Figure 5E-H**). These results indicated that CXCL13, FCRLA, PLA2G2D, and MS4A1 could be a useful biomarkers for predicting responses to immunotherapy of OC.

Correlation of four genes expression with immune infiltration level in OC and pathway analysis.

Tumor purity was usually taken into account when obtaining immunotherapy expression markers from transcriptome data. Some studies have shown that tumor samples with low purity tend to have more immune cells and a higher mutation load, because inflammatory responses induced by immune cells can increase the mutation rate of tumor cells, and the effect of immunotherapy may be better[24, 25]. Therefore, to study if the four signature genes can predict tumor immune infiltration, we used TIMER (Tumor Immune Estimation Resource, <https://cistrome.shinyapps.io/timer/>), an online database, to analyze the infiltration of different immune cells in tumor tissues based on available RNA-seq data. Our results demonstrated that CXCL13, FCRLA, MS4A1, and PLA2G2D have a negative correlation with tumor purity in OC (**Figure 6A-D**). Besides CXCL13 expression showed a very weak relationship with B cell infiltration level in OC (**Figures 6A**) while FCRLA, MS4A1, PLA2G2D have no significant correlations with B cell infiltration level in OC (**Figure 6B-D**). Moreover, there were moderate to strong positive relationships between the expression levels of PLA2G2D, FCRLA, MS4A1, CXCL13, and infiltration level of CD4⁺ T cells, as well as have a prominent positive correlation between expression level and infiltration level of CD8⁺ T cells in OC, especially CXCL13, MS4A1, PLA2G2D, FCRLA (**Figures 6A-D, Table S1**). In conclusion, CXCL13, FCRLA, MS4A1, and PLA2G2D might have multiple and closely function in immune infiltration.

Next, we conducted gene set enrichment analysis (GSEA) to analyze the four genes and found that all of four genes were related to immune cell signaling pathways. They were all associated with B cell receptor, chemokine, cytokine-cytokine receptor interaction, natural killer cell-mediated cytotoxicity and T cell receptor signaling pathway (**Figure 6E-H**). Interestingly, CXCL13, FCRLA, and PLA2G2D were also associated with primary immunodeficiency (**Figure 6E-H**). MS4A1 was associated with autoimmune thyroid disease (**Figure 6G**). All of these findings revealed that CXCL13, FCRLA, MS4A1 and PLA2G2D's expression was closely related to immune function were enriched for hallmarks of OC. More information

can be got from **Table S2**. These results indicated that CXCL13, FCRLA, PLA2G2D, and MS4A1 were validated as prognostic biomarkers for the immunotherapy of OC. These results indicate that CXCL13, FCRLA, PLA2G2D, and MS4A1 are positively correlated with immune cell infiltration, and these genes are enriched to be associated with immune pathways by GSEA enrichment method.

Prognostic evaluation of four genes in melanoma patients treated with immuncheckpoint inhibitors.

Immune infiltration and active inflammations are prerequisites of anti-tumor immune reaction and better response to immunotherapy. We have shown that CXCL13, FCRLA, PLA2G2D, and MS4A1 are upregulated in the immune^{high} group and involved in inflammation and T cell function. We then asked if these four genes can be used as biomarkers to predict responses to immune checkpoint blockade-based immunotherapy. Since the data of OC response to immunotherapy are not available, to address that, we explored an online database, TIDE (Tumor Immune Dysfunction and Exclusion, <http://tide.dfci.harvard.edu/query/>) and analyzed the data collected from melanoma. As shown in **Figure 7**, our results showed that among the melanoma patients treated with anti-PD-L1 or anti-CTLA4, the CXCL13, MS4A1, FCRLA, and PLA2G2D were positively correlated with the number of CTL infiltration. The two-sided t-test *p* values for correlations in CXCL13, MS4A1, FCRLA, PLA2G2D and mean are 1.43×10^{-10} , 8.12×10^{-2} , 6.6×10^{-10} and 2.1×10^{-4} , respectively. And the high expression levels of these genes were associated with the better response of melanoma patients.

Furthermore, we also investigated the protein expression of CXCL13, FCRLA, and MS4A1 in OC and normal ovarian tissues based on the IHC (immunohistochemistry) staining in The Human Protein Atlas (<https://www.proteinatlas.org>). We found that the staining intensity of CXCL13 and MS4A1 was slightly different in normal ovarian tissues and OC tissues, the protein expression levels in cancer were slightly lower than in adjacent tissues, but the difference was not significant. And the level of FCRLA protein was up-regulated in OC tissues compared to the normal tissues (**Figure S3**). There are no PLA2G2D IHC data available in the database currently. Considering the positive association between PLA2G2D mRNA level and prognosis data in OC, it is worth developing antibodies for future studies.

Discussion

OC has been one of the major threats to women's health for decades[26, 27]. Current treatments for late-stage OC patients are far less satisfying. Immune checkpoint blockade based immunotherapy, especially PD-1/PD-L1 blocking antibody, has been approved for clinical treatment to multiple solid tumors[28-30]. Accumulating clinical data indicated that inflammatory tumors with the high expression level of PD-L1 are more likely to respond[31]. However, the PD-L1 pathway blocking antibody has not been approved for treating OC. The reasons are likely due to scarce information of the TIME of OC resulting in missing targetable immune-modulatory molecules as well as promising biomarkers for choosing the right patient population for immunotherapy.

TMB represents the overall amount of somatic mutations that could generate neoantigens to elicit immune response[32-34]. Consistently, tumors with high TMB scores contain more TILs[35] and are associated with better response to immunotherapy[36], such as melanoma[37] and NSCLC[38], with high TMB have shown better response and treatment results in immune checkpoint blockade treatment. In our analysis, we found that late-stage OCs tend to have higher TMB, which has not been described before. We also found that a high TMB score is positively associated with a better prognosis (**Figure 2**).

The extent of tumor immune cell infiltration has been evaluated in many different ways. We analyzed the TIME of OCs by the ssGSEA and assigned each case an immune score. Consistent with other studies, high immune scores significantly associated with better prognosis. However, we did not find any correlation between TMB and immune scores (**Figure 2**). This may be due to genetic differences between tumor cells, and the inconsistency between TMB and immune infiltration has been reported in other studies, such as YardenaSamuels and EytanRuppin's research team found that in melanin, intra-tumor heterogeneity (ITH) played a decisive role in tumor immune infiltration compared with other factors, like TMB and CNV[39]. Additionally, we also studied the correlation between TMB and the infiltration of specific immune cells and found TMB scores only associated with the infiltration of memory B cells and NK cells (**Figure S2**). These results indicate that OC has other layers of regulation upon the innate immune system, such as antigen uptake and presentation, that inhibit the activation of T cells. Alternatively, other mechanisms that inhibit immune cell accumulating might present in the microenvironment of OCs. Further studies are warranted to explore these possibilities.

Transcriptome studies provide genome-wide gene expression data. By performing the differential gene expression analysis, we found that most genes that are upregulated in the immune^{high} group are associated with an immune or inflammatory signaling pathway, which validates our analysis (**Figure 3**). However, we did not see a significant difference in gene expression between TMB^{high} and TMB^{low} groups. Our analysis did not suggest any group-specific or enriched mutation, which partially explains the similar transcriptome profiles between these two groups (**Figure S2**).

Given TMB^{high} and Immune^{high} are both correlated with better survival, we focused on 14 genes that are upregulated in immune^{high} group and also enriched in TMB^{high} group. These 14 genes are involved in many immune regulatory pathways, as shown by GO and KEGG terms analysis (**Figure 3**). Consistently, PPI network analysis also suggested that several key genes related to immune response were located in the center of the module (**Figure 4**). Nevertheless, these predictions could be biased by previous knowledge[40, 41]. The real biology of these genes in regulating the anti-tumor immune response in OCs need more investigations.

We further performed an overall survival analysis upon each of the 14 DEGs, and identified four genes, PLA2G2D, CXCL13, FCRLA, and MS4A1, were associated with better prognosis in OC patients. We also validated the four genes related to the prognosis of OC patients from the GEO database (**Figure 5**). In addition, multiple reports have shown that these four genes were involved in immune reactions, such as CXCL13 has recently been linked with TFH cell infiltration and improved survival[42]. FCRLA has been

shown to be a potential target gene in immunotherapy for B-cell lymphoma[43]. MS4A1 played a vital role in the apoptosis of B-cell lymphoma Ramos cells[44]. And PLA2G2D involved in inflammation and immune response and may be used to treat inflammatory diseases[45]. The association of these four genes and immune cell infiltration was further validated by an independent algorithm presented in the TIMER database (**Figure 6**). However, positive association does not mean these genes are causative factors that affect the tumor inflammatory status, which takes more studies.

To evaluate if these four genes could help us to choose the right patients for future immunotherapy, we investigated the correlation between the expression of CXCL13, MS4A1, FCRLA, PLA2G2D and the response rate of immunotherapy. Due to the lack of data on OC's response to immunotherapy, we analyzed the data collected from a melanoma cohort. It was found that in melanoma patients treated with anti-PD-L1 or anti-CTLA4, these four genes were positively correlated with the number of CTL infiltration, and their high expression was associated with a better prognosis of patients (**Figure 7**). Further effort will focus on validating antibodies or developing other diagnostic methods to detecting the expression of these genes in the tumor microenvironment, which allows us to pick the right patients for immune checkpoint blockade based immune therapy against OCs.

In summary, by a comprehensive study of TMB and TIME, we found some clues that provide additional data in decoding the complex interaction of TMB and TIME in OC. We also identified four genes signature that could be used as biomarkers to guide future immunotherapy on OC.

Conclusions

In our study, TMB was combined with TIME of OC to classify and analyze patients according to their local tumor immune status that can guide for further immunotherapy trials with OC. We found that higher TMB can promote immune cell infiltration and increases immune activity and have a better prognosis of OC. In summary, we found that CXCL13, FCRLA, PLA2G2D, and MS4A1 may be biomarkers for OC immunotherapy. Finally, through further study of these genes that may shed new light on the potential relationship between tumor microenvironment and OC prognosis. Therefore, our researches provided a new perspective on the relationship between TMB and tumor immune response in OCs.

Abbreviations

TMB: Tumor Mutational Burden; Tumor Immune Microenvironment: TIME; OC: Ovarian Cancer; TCGA: The Cancer Genome Atlas; ssGSEA: Single-sample Gene Set Enrichment Analysis; Up-DEGs: Up-regulated Differential Expression Genes; GO: Gene Ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes pathway; PPI: Protein-Protein Interaction; DAVID: Database for Annotation, Visualization, and Integrated Discovery; GSVA: Gene Set Variation Analysis; STRING: Search Tool for the Retrieval of Interacting Genes; DISCOVER: Discrete Independence Statistic Controlling for Observations with Varying Event Rates; GEO: Group on Earth Observations; OS: Overall Survival; TIMER: Tumor Immune Estimation Resource; GSEA:

Gene Set Enrichment Analysis; ssGSEA: Single Sample Gene Set Enrichment Analysis; TIDE: Tumor Immune Dysfunction and Exclusion; IHC: immunohistochemistry.

Declarations

Acknowledgments

The authors gratefully acknowledge contributions from the TCGA network, the GTEx network, and the CCLE Network. This study was supported by grants from the National Natural Science Foundation of China (No. 81572900). The Fundamental Research Funds for the Central Universities of Central South University (No. 502221804).

Authors' contributions

Lili Fan conceived and designed this study, also crafted figures and tables, and was responsible for the critical reading of the manuscript. Also, she contributed to the data collection and analysis and the critical reading of the manuscript. Han Lei, Ying Lin, and Tianxiang Zhang were responsible for the writing and helped with the figures and tables. Han Lei, Zhengwei Zhou, Zhipeng Yan, Guang Shu, and Gang Yin supervised and contributed to the critical reading of the manuscript.

Funding

This work was supported by the National Natural Science Foundation of China [No. 81572900]; The Fundamental Research Funds for the Central Universities of Central South University [No. 502221804]; National Key R&D Program of China, Stem Cell, and Translation Research [No. 2016YFA0102000]

Availability of data and materials

All data generated or analyzed during the study period were included in this published article and its supplementary information.

Ethics approval and consent to participate

Not Applicable

Consent for publication

We have obtained consent to publish this paper from all the participants of this research.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Department of Pathology, Xiangya Hospital, School of Basic Medical Sciences, Central South University, Changsha, Hunan Province, China. ²Department of Immunology, School of Basic Medical Sciences, Central South University, Changsha, Hunan Province, China. ³Department of Immunobiology, Yale University School of Medicine, New Haven, CT, USA. ⁴School of Basic Medical Sciences, Central South University, Changsha, Hunan Province, China. ⁵Hunan Cancer Hospital/the affiliated Cancer Hospital of Xiangya School of Medicine, Central South University, Changsha, Hunan Province, China.

References

1. Chornokur G, Amankwah EK, Schildkraut JM, Phelan CM: **Global ovarian cancer health disparities.** *Gynecol Oncol* 2013, **129**:258-264.
2. Nash Z, Menon U: **Ovarian cancer screening: Current status and future directions.** *Best Pract Res Clin Obstet Gynaecol* 2020, **65**:32-45.
3. Siegel RL, Miller KD, Jemal A: **Cancer statistics, 2016.** *CA Cancer J Clin* 2016, **66**:7-30.
4. Lee JY, Kim S, Kim YT, Lim MC, Lee B, Jung KW, Kim JW, Park SY, Won YJ: **Changes in ovarian cancer survival during the 20 years before the era of targeted therapy.** *BMC Cancer* 2018, **18**:601.
5. Hoos A: **Development of immuno-oncology drugs - from CTLA4 to PD1 to the next generations.** *Nat Rev Drug Discov* 2016, **15**:235-247.
6. Hamanishi J, Mandai M, Ikeda T, Minami M, Kawaguchi A, Murayama T, Kanai M, Mori Y, Matsumoto S, Chikuma S, et al: **Safety and Antitumor Activity of Anti-PD-1 Antibody, Nivolumab, in Patients With Platinum-Resistant Ovarian Cancer.** *J Clin Oncol* 2015, **33**:4015-4022.
7. Varga A, Piha-Paul S, Ott PA, Mehnert JM, Berton-Rigaud D, Morosky A, Yang P, Ruman J, Matei D: **Pembrolizumab in patients with programmed death ligand 1-positive advanced ovarian cancer: Analysis of KEYNOTE-028.** *Gynecol Oncol* 2019, **152**:243-250.
8. Goodman AM, Kato S, Bazhenova L, Patel SP, Frampton GM, Miller V, Stephens PJ, Daniels GA, Kurzrock R: **Tumor Mutational Burden as an Independent Predictor of Response to Immunotherapy in Diverse Cancers.** *Mol Cancer Ther* 2017, **16**:2598-2608.
9. Birkbak NJ, Kochupurakkal B, Izarzugaza JM, Eklund AC, Li Y, Liu J, Szallasi Z, Matulonis UA, Richardson AL, Iglehart JD, Wang ZC: **Tumor mutation burden forecasts outcome in ovarian cancer with BRCA1 or BRCA2 mutations.** *PLoS One* 2013, **8**:e80023.
10. Duffy MJ, Crown J: **Biomarkers for Predicting Response to Immunotherapy with Immune Checkpoint Inhibitors in Cancer Patients.** *Clin Chem* 2019, **65**:1228-1238.
11. McGranahan N, Furness AJ, Rosenthal R, Ramskov S, Lyngaa R, Saini SK, Jamal-Hanjani M, Wilson GA, Birkbak NJ, Hiley CT, et al: **Clonal neoantigens elicit T cell immunoreactivity and sensitivity to immune checkpoint blockade.** *Science* 2016, **351**:1463-1469.
12. Samstein RM, Lee CH, Shoushtari AN, Hellmann MD, Shen R, Janjigian YY, Barron DA, Zehir A, Jordan EJ, Omuro A, et al: **Tumor mutational load predicts survival after immunotherapy across multiple cancer types.** *Nat Genet* 2019, **51**:202-206.

13. Garassino M, Rodriguez-Abreu D, Gadgeel S, Esteban E, Felip E, Speranza G, Reck M, Hui R, Boyer M, Cristescu R, et al: **Evaluation of TMB in KEYNOTE-189: Pembrolizumab Plus Chemotherapy vs Placebo Plus Chemotherapy for Nonsquamous NSCLC.** *Journal of Thoracic Oncology* 2019, **14**:S216-S217.
14. Langer C, Gadgeel S, Borghaei H, Patnaik A, Powell S, Gentzler R, Yang JC, Gubens M, Sequist L, Awad M, et al: **KEYNOTE-021: TMB and Outcomes for Carboplatin and Pemetrexed With or Without Pembrolizumab for Nonsquamous NSCLC.** *Journal of Thoracic Oncology* 2019, **14**:S216-S216.
15. Manji GA, Olive KP, Saenger YM, Oberstein P: **Current and Emerging Therapies in Metastatic Pancreatic Cancer.** *Clin Cancer Res* 2017, **23**:1670-1678.
16. Nishida N, Kudo M: **Immunological Microenvironment of Hepatocellular Carcinoma and Its Clinical Implication.** *Oncology* 2017, **92 Suppl 1**:40-49.
17. Şenbabaoğlu Y, Gejman RS, Winer AG, Liu M, Van Allen EM, de Velasco G, Miao D, Ostrovnaya I, Drill E, Luna A, et al: **Tumor immune microenvironment characterization in clear cell renal cell carcinoma identifies prognostic and immunotherapeutically relevant messenger RNA signatures.** *Genome Biol* 2016, **17**:231.
18. Tsai MS, Chen WC, Lu CH, Chen MF: **The prognosis of head and neck squamous cell carcinoma related to immunosuppressive tumor microenvironment regulated by IL-6 signaling.** *Oral Oncol* 2019, **91**:47-55.
19. Hanahan D, Weinberg RA: **The hallmarks of cancer.** *Cell* 2000, **100**:57-70.
20. Hanahan D, Coussens LM: **Accessories to the crime: functions of cells recruited to the tumor microenvironment.** *Cancer Cell* 2012, **21**:309-322.
21. Hänzelmann S, Castelo R, Guinney J: **GSVA: gene set variation analysis for microarray and RNA-seq data.** *BMC Bioinformatics* 2013, **14**:7.
22. Donehower LA, Soussi T, Korkut A, Liu Y, Schultz A, Cardenas M, Li X, Babur O, Hsu TK, Lichtarge O, et al: **Integrated Analysis of TP53 Gene and Pathway Alterations in The Cancer Genome Atlas.** *Cell Rep* 2019, **28**:1370-1384.e1375.
23. Canisius S, Martens JW, Wessels LF: **A novel independence test for somatic alterations in cancer shows that biology drives mutual exclusivity but chance explains most co-occurrence.** *Genome Biol* 2016, **17**:261.
24. Aran D, Sirota M, Butte AJ: **Systematic pan-cancer analysis of tumour purity.** *Nat Commun* 2015, **6**:8971.
25. Mao Y, Feng Q, Zheng P, Yang L, Liu T, Xu Y, Zhu D, Chang W, Ji M, Ren L, et al: **Low tumor purity is associated with poor prognosis, heavy mutation burden, and intense immune phenotype in colon cancer.** *Cancer Manag Res* 2018, **10**:3569-3577.
26. Siegel RL, Miller KD, Jemal A: **Cancer statistics, 2019.** *CA Cancer J Clin* 2019, **69**:7-34.
27. Bast RC, Jr., Hennessy B, Mills GB: **The biology of ovarian cancer: new opportunities for translation.** *Nat Rev Cancer* 2009, **9**:415-428.

28. Hirano F, Kaneko K, Tamura H, Dong H, Wang S, Ichikawa M, Rietz C, Flies DB, Lau JS, Zhu G, et al: **Blockade of B7-H1 and PD-1 by monoclonal antibodies potentiates cancer therapeutic immunity.** *Cancer Res* 2005, **65**:1089-1096.
29. Rizvi NA, Hellmann MD, Snyder A, Kvistborg P, Makarov V, Havel JJ, Lee W, Yuan J, Wong P, Ho TS, et al: **Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer.** *Science* 2015, **348**:124-128.
30. Bracarda S, Altavilla A, Hamzaj A, Sisani M, Marrocolo F, Del Buono S, Danielli R: **Immunologic checkpoints blockade in renal cell, prostate, and urothelial malignancies.** *Semin Oncol* 2015, **42**:495-505.
31. Radziewicz H, Ibegbu CC, Fernandez ML, Workowski KA, Obideen K, Wehbi M, Hanson HL, Steinberg JP, Masopust D, Wherry EJ, et al: **Liver-infiltrating lymphocytes in chronic human hepatitis C virus infection display an exhausted phenotype with high levels of PD-1 and low levels of CD127 expression.** *J Virol* 2007, **81**:2545-2553.
32. Hellmann MD, Nathanson T, Rizvi H, Creelan BC, Sanchez-Vega F, Ahuja A, Ni A, Novik JB, Mangarin LMB, Abu-Akeel M, et al: **Genomic Features of Response to Combination Immunotherapy in Patients with Advanced Non-Small-Cell Lung Cancer.** *Cancer Cell* 2018, **33**:843-852.e844.
33. Chan TA, Yarchoan M, Jaffee E, Swanton C, Quezada SA, Stenzinger A, Peters S: **Development of tumor mutation burden as an immunotherapy biomarker: utility for the oncology clinic.** *Ann Oncol* 2019, **30**:44-56.
34. Gubin MM, Zhang X, Schuster H, Caron E, Ward JP, Noguchi T, Ivanova Y, Hundal J, Arthur CD, Krebber WJ, et al: **Checkpoint blockade cancer immunotherapy targets tumour-specific mutant antigens.** *Nature* 2014, **515**:577-581.
35. Jiang P, Gu S, Pan D, Fu J, Sahu A, Hu X, Li Z, Traugh N, Bu X, Li B, et al: **Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response.** *Nat Med* 2018, **24**:1550-1558.
36. Klebanov N, Artomov M, Goggins WB, Daly E, Daly MJ, Tsao H: **Burden of unique and low prevalence somatic mutations correlates with cancer survival.** *Sci Rep* 2019, **9**:4848.
37. Larkin J, Hodi FS, Wolchok JD: **Combined Nivolumab and Ipilimumab or Monotherapy in Untreated Melanoma.** *N Engl J Med* 2015, **373**:1270-1271.
38. Herbst RS, Baas P, Kim DW, Felip E, Pérez-Gracia JL, Han JY, Molina J, Kim JH, Arvis CD, Ahn MJ, et al: **Pembrolizumab versus docetaxel for previously treated, PD-L1-positive, advanced non-small-cell lung cancer (KEYNOTE-010): a randomised controlled trial.** *Lancet* 2016, **387**:1540-1550.
39. Wolf Y, Bartok O, Patkar S, Eli GB, Cohen S, Litchfield K, Levy R, Jiménez-Sánchez A, Trabish S, Lee JS, et al: **UVB-Induced Tumor Heterogeneity Diminishes Immune Response in Melanoma.** *Cell* 2019, **179**:219-235.e221.
40. Forschner A, Battke F, Hadaschik D, Schulze M, Weißgraeber S, Han CT, Kopp M, Frick M, Klumpp B, Tietze N, et al: **Tumor mutation burden and circulating tumor DNA in combined CTLA-4 and PD-1 antibody therapy in metastatic melanoma - results of a prospective biomarker study.** *J Immunother Cancer* 2019, **7**:180.

41. Hugo W, Zaretsky JM, Sun L, Song C, Moreno BH, Hu-Lieskovan S, Berent-Maoz B, Pang J, Chmielowski B, Cherry G, et al: **Genomic and Transcriptomic Features of Response to Anti-PD-1 Therapy in Metastatic Melanoma.** *Cell* 2016, **165**:35-44.
42. Gu-Trantien C, Migliori E, Buisseret L, de Wind A, Brohée S, Garaud S, Noël G, Dang Chi VL, Lodewyckx JN, Naveaux C, et al: **CXCL13-producing TFH cells link immune suppression and adaptive memory in human breast cancer.** *JCI Insight* 2017, **2**.
43. Inozume T, Mitsui H, Okamoto T, Matsuzaki Y, Kawakami Y, Shibagaki N, Shimada S: **Dendritic cells transduced with autoantigen FCRLA induce cytotoxic lymphocytes and vaccinate against murine B-cell lymphoma.** *J Invest Dermatol* 2007, **127**:2818-2822.
44. Kawabata KC, Ehata S, Komuro A, Takeuchi K, Miyazono K: **TGF- β -induced apoptosis of B-cell lymphoma Ramos cells through reduction of MS4A1/CD20.** *Oncogene* 2013, **32**:2096-2106.
45. Miki Y, Yamamoto K, Taketomi Y, Sato H, Shimo K, Kobayashi T, Ishikawa Y, Ishii T, Nakanishi H, Ikeda K, et al: **Lymphoid tissue phospholipase A2 group IID resolves contact hypersensitivity by driving antiinflammatory lipid mediators.** *J Exp Med* 2013, **210**:1217-1234.

Supplementary Materials

Figure S1. Database analysis.

Figure S2. Comparison of the infiltration levels of immune cells in the high and low immunity group.

Figure S3. The protein expressions of the three selected regulators.

Table S1. The correlation coefficients of four genes with immune infiltration in OC.

Table S2. The correlation coefficients of four genes with immune-related pathways in OC.

Figures

Figure 1

The workflow of the current work.

Figure 2

TMB scores and the immunoactivity-related gene sets are conjoined with OC's overall survival. (A) The violin plots showed the interrelationship between TMB and OC grades. It exhibited that there was a powerful relationship between OC grades and the level of TMB scores (n=417, except for the blank values, p=0.0480). (B) Up: Unsupervised clustering of OC patients from the TCGA cohort using ssGSEA scores

from immune cell types. Hierarchical clustering was performed with Euclidean distance and Ward linkage. Three distinct immune infiltration clusters, here we define some of them as immunitylow and immunityhigh. Down: the correlation analysis between the expression level of PD-L1 and the high & low immunity clusters obtained by ssGSEA score based on immune cell type (As displayed by Mann-Whitney test, $p < 0.0001$). (C) OC patients were classified into two groups based on their TMB scores: the samples with TMBhigh scores in the top 20% and the TMBlow samples with TMB scores in the bottom 20% of all OC samples. According to the Kaplan-Meier survival curve, the total survival time of TMBhigh group is longer than TMBlow group (2.58 years vs. 1.92 years, $n=430$, except for the blank values, as displayed by the long-rank test, $p=0.0295$). (D) OC patients were classified into two groups based on their immunity: those with immunityhigh group in 193 patients and those with immunitylow group in 149 patients. The total survival time of immunityhigh group is longer than immunitylow group (2.77 years vs. 2.67 years, $n=364$, except for the blank values, as displayed by the log-rank test, $p=0.0082$). (E) The correlation analysis between the TMB group and the immunity group (Linear regression test, $p=0.27$).

Figure 3

KEGG pathway and GO term analysis of DEGs associated with OC. (A) Venn diagrams displayed the amount of conjointly up-regulated DEGs in both the TMB group and the high immunoactivity-related gene sets group. (B) Unsupervised hierarchical clustering of KEGG pathways associated with Up-DEGs. Blue represents normal samples, data from GTEx ($n=88$), red represents treat samples, data from TCGA ($n=379$). The average linkage method and the Pearson distance measurement method were used to depict the Heatmaps. High expression genes are shown in red, low expression genes in green, and genes with the same level of expression are shown in black. False discovery rate (FDR) < 0.05 , fold change > 1 . (C, D) The overview of the GO and KEGG terms. The X-axis serves as the proportion of relevant genes, and the y-axis serves as the GO and KEGG terms. Each bubble means a term. The size of the bubble implies

the account of relevant genes. Lighter colors suggest the smaller P values. C: Enriched GO terms. D: Enriched KEGG terms. FDR of GO and KEGG pathway analysis was acquired from the DAVID functional annotation tool. $p < 0.05$.

Figure 4

PPI networks of Up-DEGs Modules. (A) PPI network established by the STRING database for the Up-DEGs module. Minimum required interaction score >0.04 , Disconnected nodes were hidden in the network. (B) Quantification diagram of the Up-DEGs module.

Figure 5

Correlation of expression of individual Up-DEGs in overall survival. (A-D) Kaplan-Meier survival curves were produced for selected DEGs extracted from the comparison of groups of high (red line) and low (blue line) gene expression. The elevated expression of CXCL13 (A), FCRLA (B), MS4A1 (C) and PLA2G2D (D) were related to better overall survival (OS) in OC. $p < 0.05$ in the Log-rank test. (E-H) Verification of the correlation between DEGs extracted from the TCGA database and overall survival time in the GEO cohort. Forest plots showed the association with the clinical benefits of the four genes in the survival prognosis study and can as a univariate predictor of OC.

Figure 6

Correlation of four genes expression with immune infiltration level in OC and pathway analysis. (A-D) Scatter diagram of the correlation between the expression levels of CXCL13, FCRLA, MS4A1, PLA2G2D and immune infiltration level in OC. (E-H) An evaluated enrichment degree within the four genes-expressing. KEGG pathway analysis of CXCL13, FCRLA, MS4A1, and PLA2G2D. The vertical bars on the x-

axis of the GSEA diagram represent the positions of genes in a given set in a sorted list. Positive GSEA curve means positive enrichment and the negative GSEA curve means anti-enrichment.

Figure 7

Prognostic evaluation of four genes in melanoma patients treated with immunotherapy. (A-D) Scatter diagram of the correlation between the expression levels of CXCL13 (A), FCRLA (B), PLA2G2D (C) and MS4A1 (D) the response rate of immune checkpoint inhibitors in melanoma. The expression correlations of top (> 0) and bottom (< 0) were obtained in patients with melanoma treated with anti-PD1 and anti-CTLA4. The x-axis shows the CTL level for each tumor (average expression level of CD8A, CD8B, GZMA, GZMB, and PRF1). The Pearson correlation (R) between the plotted values is shown in the upper of each plot. (E-F) Kaplan–Meier plots of overall survival (OS) for melanoma patients with top half and bottom half CXCL13 (E), FCRLA (F), PLA2G2D (G) and MS4A1 (H) expression levels, using the data from an anti-PD1 and anti-CTLA4 study. The p-value was calculated by testing the association between high/low expression and overall survival with the two-sided Wald test in a Cox-PH regression. The z-score of each gene is calculated by dividing the interaction coefficient by its standard error. The elevated expression of CXCL13 (E), FCRLA (F), PLA2G2D (G) and MS4A1 (H) were related to better overall survival (OS) in melanoma. $p < 0.05$ in the cox-PH regression.

Supplementary Files

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

- [SupplementaryMaterials.pdf](#)