



Analysis of Bulk RNA Sequencing Data Reveals Novel Transcription Factors Associated With Immune Infiltration Among Multiple Cancers

OPEN ACCESS

Edited by:

Zlatko Trajanoski,
Innsbruck Medical University, Austria

Reviewed by:

Hubert Hackl,
Medical University of Innsbruck,
Austria

Fatima Sanchez-Cabo,
Spanish National Centre for
Cardiovascular Research, Spain

***Correspondence:**

Luguo Sun
sunlg388@nenu.edu.cn
Xianling Cong
congxl888@hotmail.com
Yi Zhang
yizhang@ablife.cc

[†]These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Cancer Immunity
and Immunotherapy,
a section of the journal
Frontiers in Immunology

Received: 21 December 2020

Accepted: 30 July 2021

Published: 20 August 2021

Citation:

Liu L, Zhao Q, Cheng C, Yi J, Sun H,
Wang Q, Quan W, Xue Y, Sun L,
Cong X and Zhang Y (2021) Analysis of
Bulk RNA Sequencing Data Reveals
Novel Transcription Factors
Associated With Immune Infiltration
Among Multiple Cancers.
Front. Immunol. 12:644350.
doi: 10.3389/fimmu.2021.644350

Lei Liu^{1†}, Qiuchen Zhao^{2,3†}, Chao Cheng², Jingwen Yi⁴, Hongyan Sun⁵, Qi Wang²,
Weili Quan², Yaqiang Xue², Luguo Sun^{1,4*}, Xianling Cong^{5*} and Yi Zhang^{2*}

¹ National Engineering Laboratory for Druggable Gene and Protein Screening, Northeast Normal University, Changchun, China,

² Center of Genome Analysis, ABLife BioBigData Institute, Wuhan, China, ³ College of Life Sciences, Wuhan University, Wuhan, China, ⁴ Research Center of Agriculture and Medicine Gene Engineering of Ministry of Education, Northeast Normal University, Changchun, China, ⁵ Tissue Bank, China-Japan Union Hospital, Jilin University, Changchun, China

Tumor-infiltrating immune cells shape the tumor microenvironment and are closely related to clinical outcomes. Several transcription factors (TFs) have also been reported to regulate the antitumor activity and immune cell infiltration. This study aimed to quantify the populations of different immune cells infiltrated in tumor samples based on the bulk RNA sequencing data obtained from 50 cancer patients using the CIBERSORT and the EPIC algorithm. Weighted gene coexpression network analysis (WGCNA) identified eigengene modules strongly associated with tumorigenesis and the activation of CD4+ memory T cells, dendritic cells, and macrophages. TF genes *FOXM1*, *MYBL2*, *TAL1*, and *ERG* are central in the subnetworks of the eigengene modules associated with immune-related genes. The analysis of The Cancer Genome Atlas (TCGA) cancer data confirmed these findings and further showed that the expression of these potential TF genes regulating immune infiltration, and the immune-related genes that they regulated, was associated with the survival of patients within multiple cancers. Exome-seq was performed on 24 paired samples that also had RNA-seq data. The expression quantitative trait loci (eQTL) analysis showed that mutations were significantly more frequent in the regions flanking the TF genes compared with those of non-TF genes, suggesting a driver role of these TF genes regulating immune infiltration. Taken together, this study presented a practical method for identifying genes that regulate immune infiltration. These genes could be potential biomarkers for cancer prognosis and possible therapeutic targets.

Keywords: immune infiltration, pan-cancers, transcription factor, immune cell types, *FOXM1*, *MYBL2*, *TAL1*, *ERG*

INTRODUCTION

Cancer immunotherapy involves the use of drugs to either relieve the immune suppression of the tumor microenvironment or strengthen the immune system to eliminate cancer cells; it indeed provides enormous durable clinical benefits to patients with late-stage cancer across many tumor types (1). Tumor-infiltrating immune cells can shape the tumor microenvironment and are closely linked to immunotherapy. Compositions, localizations, and even orientations of tumor-infiltrating immune cells can influence the efficacy of anticancer immune responses (2, 3). The quantitation of the immune contexture is crucial to not only the prognosis (4) but also the checkpoint-blocker-based monotherapy or combination therapy (5). Many computation methods have been developed to study the immune cell composition dynamics during tumorigenesis, such as EPIC (6), TIMER (7), and CIBERSORT (8). Based on the gene expression profiles in the bulk tumor tissue and the specific gene expression profiles of various immune cells, these methods allow the quantification of immune cell composition by the conventional gene profiling methods, including bulk RNA-seq.

Several transcription factors (TFs) have been reported to suppress the antitumor immune response in various solid tumors. The expression pattern of forkhead box protein P3 (FOXP3) in tumor-infiltrating lymphocytes of primary cutaneous melanoma (PCM) suggests that FOXP3-expressing lymphocytes may suppress the local anti-PCM immune response in the microenvironment, thus favoring melanoma progression (9). A recent study in human breast cancer identified FOXP1 as an important negative regulator of antitumor immune responses *via* its control of the chemokine expression (10). Signal transducer and activator of transcription 3 (STAT3) is constitutively activated in both tumor and immune cells, representing a promising target for cancer therapy. STAT3 directly regulates the expression of oncogenes and triggers tumor progression and also induces tumor-induced immunosuppression that indirectly promotes human cancer growth (11). TF c-Maf controls the immunosuppressive macrophage polarization and function in cancer *via* the transcriptional regulation of M2-related genes, serving as a metabolic checkpoint, overexpressing in tumor-associated macrophages (TAMs), and regulating TAM immunosuppressive function. The knockout of c-Maf in myeloid cells contributes to decreased tumor burden with improved antitumor T-cell immunity (12). Solid tumors adapt to hypoxia by upregulating TF HIF-1 α . Meanwhile, tumor-infiltrating natural killer (NK) cells are frequently dysfunctional in killing tumor cells (13). A recent single-cell RNA-seq study has shown that the depletion of HIF-1 α in mouse NK cells elevates antitumor activity and inhibits tumor growth. Another study in a mouse model and clinical samples established that HIF-1 α is a potent inhibitor of nuclear factor kappa B (NF- κ B) signaling driven by interleukin (IL)-18 and the antitumor activity of tumor-infiltrating NK cells and, therefore, represents a potential immunotherapy target (14).

A few TFs are capable of motivating immunity cells to enhance their immune response, such as T-box expressed in T cells (T-bet) and forkhead box M1 (FoxM1). T-bet has been

shown as one of the crucial TFs responsible for controlling the fate of both innate and adaptive immune cells, and its expression in many immune cells on mucosal surfaces can increase immunity (15). After treatment with Tet-derivative doxycycline to induce FOXM1, transgenic mice exhibited hepatic infiltration of macrophages (16). Despite the importance of TFs in regulating immune cell infiltration and immune response, a paucity of large-scale studies investigating the association of TFs with immune infiltration in pan-cancers remains.

In this study, we used the CIBERSORT technology on 100 paired bulk RNA-seq data from nine different patients with solid tumors to systematically explore the tumor-associated changes in immune cell composition. Moreover, we used weighted gene coexpression network analysis (WGCNA) to identify novel TFs strongly associated with tumor-associated immune cell infiltration. We found that TFs, including FOXM1, MYBL2, ERG, and TAL1, together with many immune-related genes, form several TF-immune-related gene expression networks (TF-iGENs). The analysis of The Cancer Genome Atlas (TCGA) data showed that the expression level of these genes involved in the same TF-iGEN were consistently associated with tumorigenesis and the survival of patients within multiple cancers. Public datasets from TCGA also validated the connection between TF-iGEN networks and immune infiltration. The expression quantitative trait locus (eQTL) analysis of exome-seq data showed that mutations were significantly more frequent in the regions flanking the TF genes than those with the non-TF genes. Our research presented a practical method for identifying TFs regulating immune infiltration, which could be potential new prognosis indicators and possible therapeutic targets in pan-cancers.

MATERIALS AND METHODS

Tumor and Normal Dataset

We previously performed RNA-seq sequencing on 50 patients, including 100 RNA-seq data of the human tumor and tumor adjacent normal tissues covering nine cancer types, including six cervical squamous cell carcinoma (CSCC), six esophageal squamous cell carcinoma (ESCC), six gastric adenocarcinoma (GAC), six hepatocellular carcinoma (HCC), six lung adenocarcinoma (LUAD), six lung squamous cell carcinoma (LUSC), five papillary thyroid carcinoma (PTC), six small-cell lung carcinoma (SCLC), and three gastric signet-ring cell carcinoma (SRCC) (GEO accession: GSE87410) (Table S1).

RNA-seq Data Processing and Analysis of Differentially Expressed Genes

For RNA-seq data, adaptors and low-quality bases were trimmed from raw sequencing reads using the FASTX Toolkit (Version 0.0.13). Reads shorter than 16 nt were discarded. Clean reads were aligned to the human GRCh38 genome by Tophat2 (17) allowing four mismatches. Uniquely mapped reads were ultimately used to calculate read number and reads per kilobase of exon per million fragments mapped (FPKM) for

each gene. The software edgeR (18), which is specifically used to analyze differentially expressed genes (DEGs), was applied to screen the RNA-seq data for DEGs. The results were analyzed based on fold change ($FC \geq 2$ or ≤ 0.5) and false discovery rate ($FDR \leq 0.05$) to determine whether a gene was differentially expressed.

Assessment of Tumor-Infiltrating Immune Cells

The CIBERSORT algorithm (8) (v1.03) was used with the default parameter for estimating immune cell fractions using FPKM values of each expressed gene. A total of 21 human immune cell phenotypes were analyzed in the study, including seven T cell types [CD8 T cells, naive CD4 T cells, memory CD4 resting T cells, memory CD4 activated T cells, T follicular helper cells, and regulatory T cells (Tregs)]; naive and memory B cells; plasma cells; resting and activated NK cells; monocytes; macrophages M0, M1, and M2; resting and activated dendritic cells; resting and activated mast cells; eosinophils; and neutrophils. An R package, ImmuneDeconv (19) (v2.0.2), that provided a unified interface to seven deconvolution methods was used for estimating immune cell fractions, while EPIC (20) was applied for estimating immune cell fractions.

WGCNA Analysis

We applied WGCNA to fully understand the gene expression pattern in pan-cancers (21) to cluster genes having similar expression pattern with default parameters. All DEGs between the tumor and adjacent normal samples for each cancer type were used as input data. Eigengenes for each clustering module were used as the representative expression pattern of genes in each module. Module–trait associations were also investigated using WGCNA.

TF-iGENs Analysis

Eigengene modules were selected to build a TF-iGEN. The expression levels of TF genes, LM22 marker genes, and immune-related genes (<https://www.immport.org/shared/genelists>) from the ImmPort database for each module were retained. Then, the Spearman correlation for each gene pair was calculated. Gene pairs with Spearman correlation coefficients >0.6 (or less than -0.6) and a corresponding p value (Benjamini–Hochberg corrected) <0.01 were considered to be significantly correlated. Then, TF–immune-related gene pairs or TF–LM22 gene pairs were retained when the TF–gene pair was found in the Encode TF target database (<https://maayanlab.cloud/Harmonizome/dataset/ENCODE+Transcription+Factor+Targets>) or TRRUST database (<https://www.grnpedia.org/trrust/>). The TF-iGEN was constructed using Cytoscape software.

TCGA Dataset Analysis

Expression data for 33 cancer types from TCGA were analyzed using GEPIA2 (22), a web-based tool that compares gene expression between tumor and normal tissues from TCGA. Furthermore, GEPIA was used to comprehensively analyze the association of gene expression with overall survival (OS) in various types of cancer. GEPIA uses the Mantel–Cox test for hypothesis test. $p < 0.05$ was labeled as significant.

Exome Sequencing

Exome sequencing of tumor and adjacent normal tissues was performed on 12 of the 50 patients having RNA-seq data (NCBI BIOPROJECT: PRJNA686225). Genomic DNA was extracted by the phenol chloroform extraction of nuclear pellets or using a Qiagen DNeasy Blood and Tissue kit (Qiagen). Purified genomic DNA was sheared to fragments of 100–500 base pairs, and 500 ng fragmented DNA was used for pair-end library preparation with a Truseq DNA library preparation kit (Illumina). After end-repair and 3' dA overhanging, fragmented DNA was ligated to Truseq adaptors (Illumina) and amplified for 10 cycles. Liquid-phase sequence capture was performed using a NimbleGen SeqCap kit (Roche). The Truseq DNA libraries were denatured into single-stranded DNA and hybridized to SeqCap oligo pools. The bound DNA fragments were eluted and PCR amplified for another 10 cycles. Fragments corresponding to 200–400 bp were purified with AMPure Xp beads and stored at -80°C until use for sequencing. Enriched libraries were sequenced on an Illumina Nextseq 500 system (ABLife Inc.).

Exome-seq Data Analysis

Adaptors were removed from raw reads using cutadapt (version 1.7.1) first, and then, reads were processed with the FASTX Toolkit (version 0.0.14) for trimming low-quality bases (qualities <20) and removing low-quality reads ($<70\%$ of read length with qualities <20). Then, N-containing reads were trimmed from N base. High-quality reads longer than 16 nt were aligned to the human genome (GRCh38) using BWA-MEM v 0.7.10-r789 (23) with default parameters. The resulting alignment was sorted by coordinates and further converted into binary alignment map (BAM) format using samtools v 1.6. The rmdup module of samtools was used to remove the duplicates from the data. The Genome Analysis Tool Kit (GATK) v3.5-0-g36282e4 (24) modules RealignerTargetCreator, Indel Re-aligner, and Base Re-calibrator were used to preprocess the alignments. During base quality recalibration, dbSNP variants and 1,000 genome variants were used as known sites. Target-capture efficiency metrics were determined using Picard HsMetrics. The realigned and recalibrated BAM file was used as an input to GATK HaplotypeCaller using the following parameters: `genotyping_mode DISCOVERY -stand_emit_conf 10 -stand_call_conf 30`. Finally, raw variant calls were soft filtered using GATK Variant Filtration based on the following parameters: `Low Qual (30 < Q < 50)`. Variants were annotated using ANNOVAR (25).

Detection of Somatic Mutations

Somatic mutations were identified using GATK Mutect2 in matched tumor and normal samples with default parameters. Candidate somatic mutations were further filtered based on gene annotations to identify those occurring in exon regions. The mutation landscape and visualization were created using the MafTools (v. 2.6.0) (26) in R software.

eQTL Analysis

We performed additional eQTL analysis in 22 exome samples to validate the correlation between expression and mutation. We used Matrix eQTL (v2.3) (27) to test *cis*-eQTL associations, and parameters

“useModel = modelLINEAR,” “errorCovariance = numeric(),” and “cisDist=1000000” were applied to assess the statistical significance between gene expression and single-nucleotide polymorphism (SNP) genotypes.

Functional Enrichment Analysis

To sort out functional categories of DEGs, Gene Ontology (GO) terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways were identified using KOBAS v2.0 (28). Hypergeometric test and Benjamini–Hochberg false discovery rate (FDR) controlling procedure were used to define the enrichment of each term. To explore the biological processes and signal pathways related to immune system process pathway, enrichment analysis of genes in each WGCNA modules was performed by Gene Ontology and pathway analysis in Metascape (29).

Statistical Analysis

All data presented were reproduced in at least three independent experiments. Statistical analysis was performed using the R software (<https://www.r-project.org/>) unless otherwise stated. Significance of differences was evaluated with either the Student’s *t*-test when only two groups were compared or the hypergeometric test for functional term enrichment analysis. No statistical methods were used to predetermine sample size (* $p \leq 0.05$, ** $p \leq 0.01$, and *** $p \leq 0.001$).

RESULTS

Weighted Gene Coexpression Analysis Obtained Tumor-Enriched Modules Associated With the Estimated Proportion of Immune Cell Type Traits

Tumor-infiltrating immune cells can shape the tumor microenvironment and are closely linked to immunotherapy. Using the CIBERSORT algorithm, we estimated the fraction of 21 subpopulations of immune cell types in 50 patients with CSCC, ESCC, GAC, HCC, LUAD, LUSC, PTC, SCLC, and SRCC (Table S2). Consistent with previous reports, the proportions of immune cells varied between different tumor types and their adjacent normal tissues (Figure S1A). Interestingly, the proportions of regulatory T cells (Tregs), M0 macrophages, T follicular helper cells, M1 macrophages, memory CD4 activated T cells, and dendritic cells in pan-cancer tissues were significantly ($p < 0.05$) higher than those in adjacent normal tissues, indicating their infiltration during tumorigenesis. On the contrary, the proportions of memory CD4 resting T cells and CD8 T cells were significantly ($p < 0.05$) lower (Figure 1A). The fluctuations in immune infiltration in tumor and adjacent normal samples across different cancer types were also exhibited using box plots (Figure S1B). We also applied EPIC, another tool, to estimate the proportions of different cell types from bulk gene expression data. We found that CD4+ T cells showed a lower proportion in tumors than in normal tissues,

while cancer-associated fibroblasts showed a higher proportion in tumors than in normal tissues (Figures S1C–E).

Then, WGCNA was used to identify gene coexpression modules correlated with tumor and immune cell-type proportion. We identified 55 coexpressed gene modules using all DEGs between tumor and adjacent normal tissues of all nine cancer types (Figure 1B). The modules were named by different colors. Among them, 3/55, including MEblue, MEdarkred, and MEmediumpurple3 module, significantly negatively correlated ($p < 0.001$), and 2/55, including MEyellow and MEwhite module, significantly positively correlated ($p < 0.001$) with the tumor group (Figures 1B, S2A, and Table S3).

We examined the correlation between the expression of module eigengenes from MEyellow, MEwhite, MEblue, MEmediumpurple3, and MEdarkred, and the fraction of 21 immune cell populations to explore the association of immune cell composition dynamics with gene coexpression networks (Figure S2B and Table S4). The expression of these modules correlated with the infiltration of specific immune cell types (Figures 1C–F), indicating that we could use the gene expression from these modules to predict the fluctuation of the immune cell population in the tumor microenvironment. The significant correlation coefficients suggested the increased expression of the MEyellow module accompanied by the relatively high percentage of nonpolarized macrophages, M0 (Figure 1C); MEwhite module was most positively associated with memory CD4 activated T cells (Figure 1D). Similarly, improved infiltration of activated dendritic cells could be linked to the upregulated expression of MEblue module (Figure 1E). Nonetheless, the expression of ME mediumpurple3 and MEdarkred modules showed a low association with infiltrating immune cells (Figures 1F and S2C). We also observed that MEsalmon4 eigengene expression was positively associated with the populations of Tregs, follicular helper T cells, and CD8 T cells (Figure S2D), and these eigengenes were strikingly overexpressed in HCC tumor tissues (Figure S2E). These results indicated that some gene coexpression networks strongly correlated with the proportion of specific immune cell types, contributing to the development of the specific immune cell population in the tumor microenvironment.

Construction of TF-iGENs in MEyellow and MEblue Modules

TFs can control the expression of critical genes and thus play a prominent role in controlling the infiltration and functionality of immune cells (30, 31). We decided to identify the key TF regulators in MEyellow and MEblue modules that were more globally associated with tumor-infiltrating immune cell dynamics (see below). For tumor-associated genes in the MEyellow module, the KEGG analysis showed that their most enriched pathways were related to cell growth, including cell cycle, chromosome segregation, DNA replication, and DNA repair (Figure S3A and Table S5).

Next, Metascape analysis was performed to cluster the tumor-associated genes in the MEyellow module; the parent GO term of the immune system process pathway was enriched, and its child

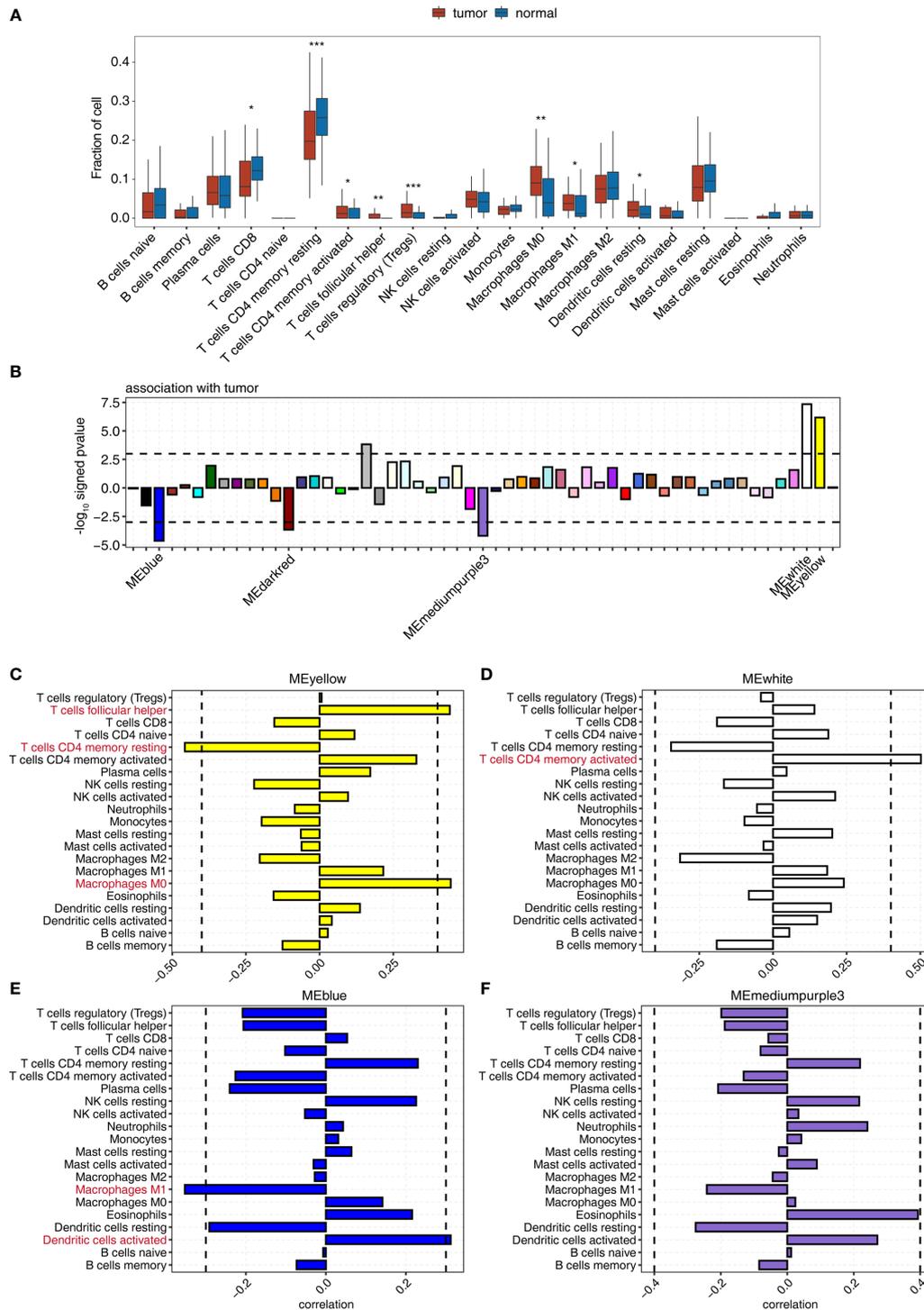


FIGURE 1 | WGCNA determination of the correlation between the expression of module eigengenes whose expression was tumor deregulated and the change in the proportion of tumor-associated immune cells. **(A)** Boxplot showing the fraction of each immune cell type in tumor or normal samples; the significant difference in the immune cell fractions between these two groups was calculated using the Student's t-test. *pvalue \leq 0.05; **pvalue \leq 0.01; ***pvalue \leq 0.001 **(B)** Signed association of the module eigengene expression with tumor. Positive values indicate modules with increased expression in tumor samples. Negative values indicate modules with decreased expression in tumor samples. Dashed lines signify tumor-associated modules. **(C–F)** Correlation between immune cell population and the expression of module eigengenes from **(C)** MEyellow, **(D)** MEwhite, **(E)** MEblue, and **(F)** MEMediumpurple3. Dashed lines signify tumor-associated modules.

GO terms were extracted and presented (**Figure 2A** and **Table S6**). These child GO pathways were mainly enriched in the differentiation and activation of immune cells, including T and B cells. These results were consistent with the positive correlation between MEyellow gene expression and follicular helper T cells and memory CD4 activated T cells, and with the negative correlation of MEyellow gene expression with CD4 memory resting T cells.

The aforementioned results encouraged us to analyze TF and the immune-related genes in the MEyellow module. To this end, a TF-iGEN was constructed based on both the correlation of gene expression between TFs and immune-related genes and also the binding potential of TFs on the correlated immune-related genes. Specifically, we first screened TFs, immune-related genes, and LM22 marker genes from the MEyellow module. Then, we calculated the Spearman correlation for each gene pair. Gene pairs with Spearman correlation coefficients >0.6 (or less than -0.6) and a corresponding p value (Benjamini-Hochberg corrected) <0.01 were considered significantly correlated. Then, TF-immune gene pair or TF-LM22 gene pair was filtered when the TF-gene pair was found in the Encode TF target database or TRRUST database (**Table S7**). The TF-iGEN was visualized in Cytoscape. Finally, we found that two hub TF genes, FOXM1 and MYBL2, were significantly associated with immune-related genes including *BIRC5*, *CACYBP*, *CDK4*, and *HDGF* (**Figure 2B**). TP53, another TF gene, was slightly associated with other immune-related genes (**Figure 2B**). We further demonstrated that the genes involved in TF-iGENs showed significantly higher expression levels in the tumor than in the adjacent normal tissues (**Figure 2C**).

A similar approach was then applied to analyze genes in the MEblue module; the most enriched KEGG pathways were different from the MEyellow module, mainly involving blood vessel development, extracellular structure organization, and cell adhesion (**Figure S3B** and **Table S8**). Child GO terms of the immune system process were mainly enriched in the activation of immune cells, such as myeloid leukocytes, neutrophils, macrophages, microglial cells, and T cells (**Figure 2D** and **Table S9**). We constructed TF-iGENs for the MEblue module (**Figure 2E**). Two hub TFs, TAL1 and ERG, were identified, which were significantly associated with immune-related genes, including *SLC11A1*, *EDN1*, *ICAM2*, *ICAM1*, *RARA*, *IFNGR1*, and *ARRB1*. Another TF gene, *RARA*, which was also involved in immune response pathways, was slightly associated with other immune-related genes. Different from that from the MEyellow module, the gene expression profile of TF-iGENs from the MEblue module showed a higher expression level in the adjacent normal samples than in the tumor (**Figure 2F**). These results indicated that the TF-iGENs derived from the two modules might carry different functionality.

Validation of the TF-iGEN in TCGA Datasets

We downloaded and profiled the expression of these candidate genes in 33 TCGA cancer datasets to further study the clinical relevance of the TF-iGEN in tumors (**Figures 3A–D**). FOXM1

and MYBL2 were significantly higher in tumors than in adjacent normal tissues in 70% (23/33) of cancers (**Figures 3A, B**). In contrast, TAL1 was significantly underexpressed in 52% (17/33) of tumors compared with normal tissues (**Figure 3C**), and ERG was significantly underexpressed in 33% (11/33) of tumors (**Figure 3D**). Another TF in the TF-iGEN from the MEyellow module, TP53, showed higher expression in pan-cancer compared with normal tissues (12/33) (**Figure S4A**). The differential expression of these six TFs was either never shown to be opposite or only shown to be opposite in one or a few cancer types (**Figures 3A–D** and **S4A, B**), suggesting a general role for these immune-related genes and TFs in many cancers. On the contrary, RARA was highly expressed in 6/33 cancers and lowly expressed in 5/33, demonstrating the controversial expression pattern (**Figure S4B**).

Since the expression of these networks closely correlated with the fluctuation in the immune cell population, we speculated that TF-iGENs had some prognostic value for patients with cancer. GEPIA2 analysis was further used to perform survival analyses for genes in the TF-iGENs. For each group, we conducted a heatmap to exhibit the survival analysis results across multiple cancer types (**Figures 3E, F**). In the heatmap, the red blocks denoted significantly high risk, and the blue ones denoted significantly low risk associated with higher gene expression. For the TF-iGENs from the MEyellow module, the high expression meant high risk. Especially in cancers such as ACC, KIRP, LGG, LIHC, LUAD, MESO, SARC, and SKCM, at least five genes in TF-iGENs showed consistent effects on the prognosis (**Figure 3E**). On the contrary, for the TF-iGENs from the MEblue module, high expression was more indicative of low risk (**Figure 3F**), suggesting that the survival time negatively correlated with the expression of these genes. These results demonstrated that some of the genes from TF-iGENs could serve as important indicators for the prognostic analysis of patients with cancers.

Besides expression analysis, two datasets from TCGA, LUAD, and LUSC were also used to validate the connection between TF-iGEN networks and immune infiltration. In our previous studies, compared with the elevated expression of genes from the MEyellow module, the fraction of some kinds of cells, such as CD4 memory activated T cells and M1 macrophages, increased in cancer tissues, while other cells such as CD4 memory resting T cells followed a different trend (**Figure S1B**). In LUAD and LUSC datasets, we observed a similar alteration of these immune cells from normal to cancer tissues (**Figures 4A** and **S5A**). Furthermore, Pearson's correlation analysis was performed to test whether the expression of genes in TF-iGEN networks strongly correlated with the infiltration of immune cells. The expression levels of genes from the MEyellow module exhibited a positive correlation with CD4 memory activated T cells, Tregs, M1 macrophage M1, and so forth, and a negative correlation with monocytes, CD4 memory resting T cells, and so forth (**Figures 4B** and **S5B**). The expression of genes from the MEblue module also strongly correlated with specific immune cell fractions (**Figures 4C** and **S5C**). Together with our previous findings (**Figures 1C, E**), these results confirmed the potential usage of gene expression of TF-iGEN networks to

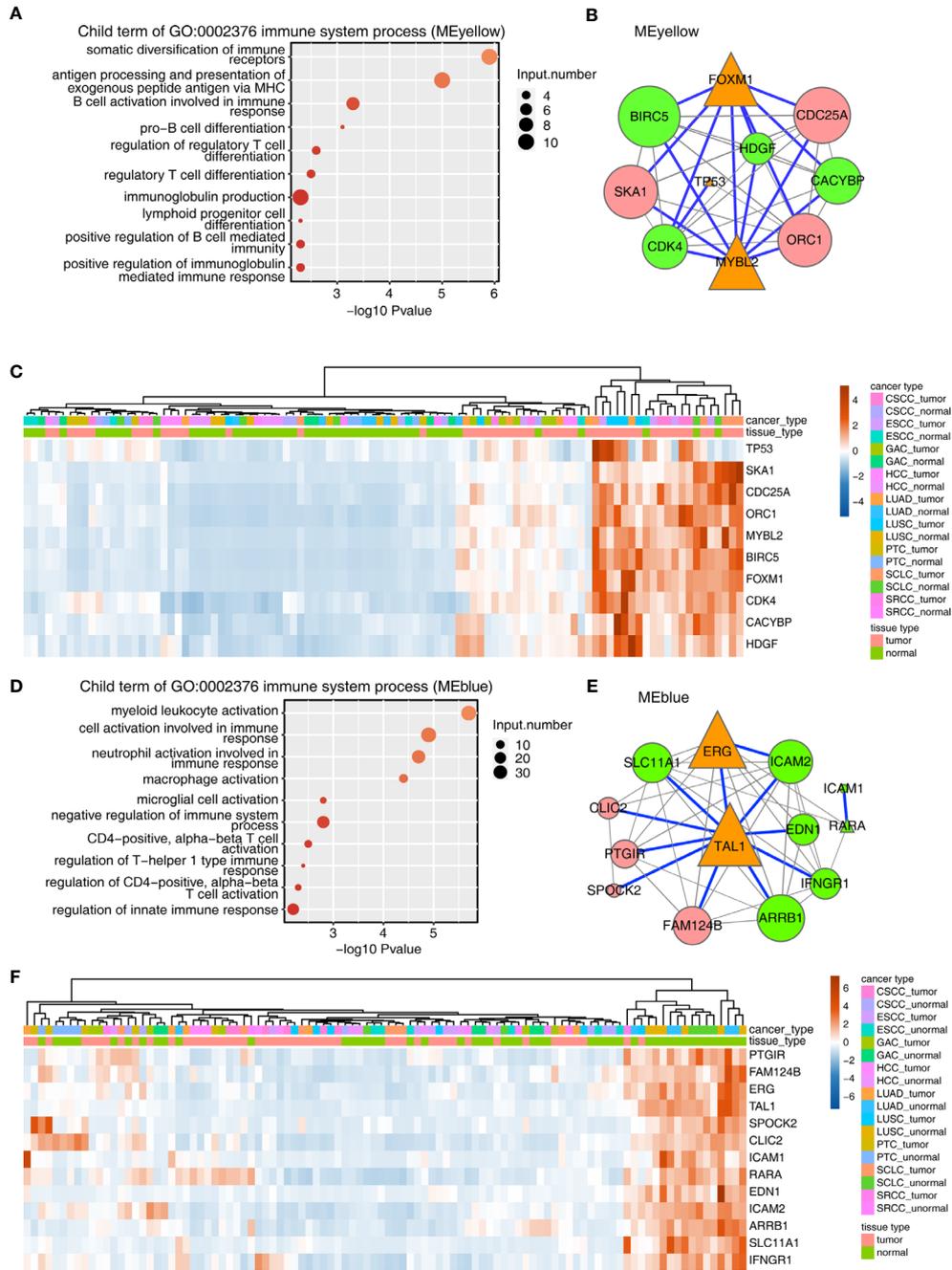
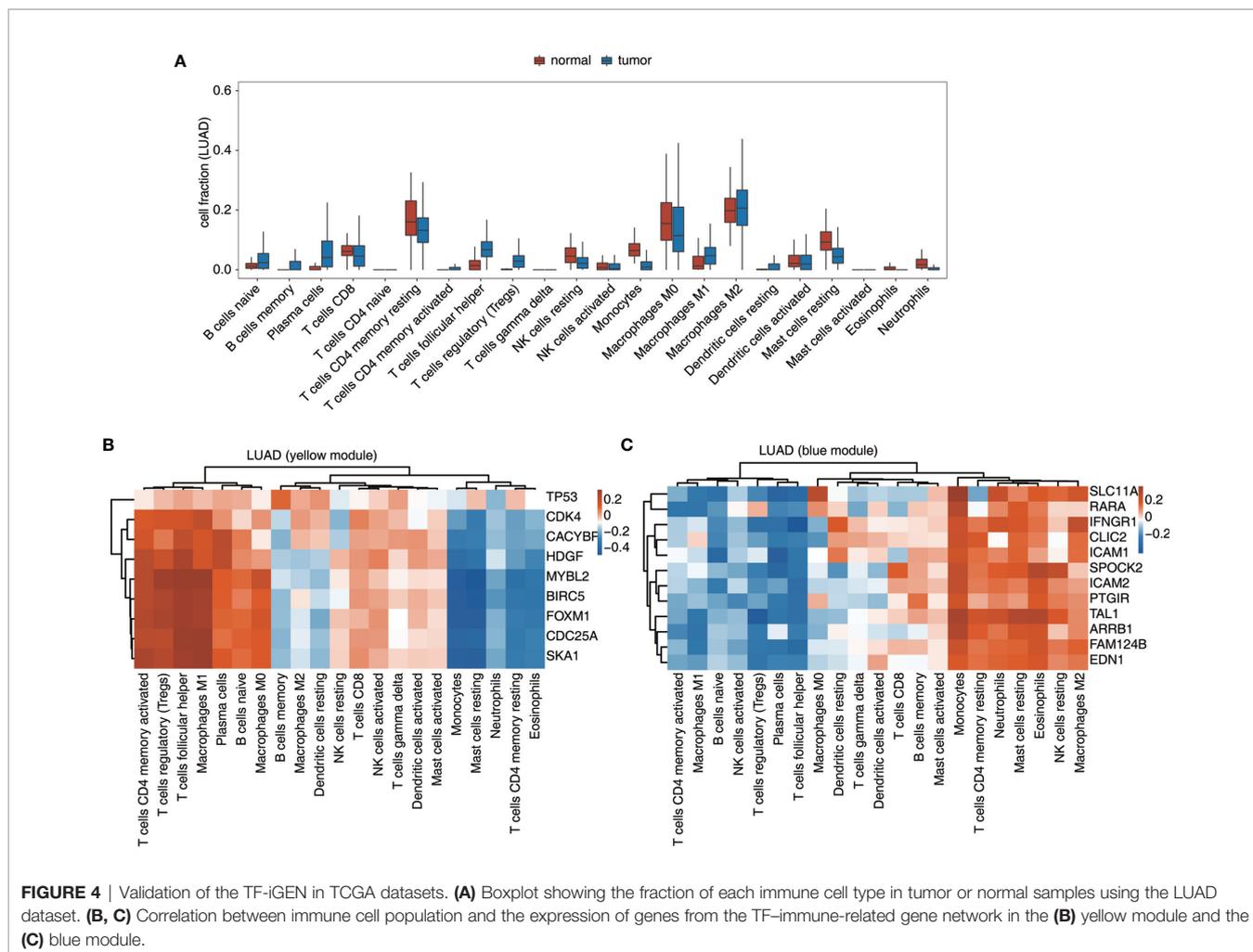


FIGURE 2 | Construction of the TF-immune-related gene expression network in MEyellow and MEblue modules. **(A)** Metascape analysis was performed to cluster the tumor-associated genes in the MEyellow module. The parent GO term of the immune system process pathway was enriched, and its child GO terms were extracted and presented. **(B)** TF-immune-related gene expression network of the tumor-associated DEGs clustered in the MEyellow module. The expression levels of TF genes, LM22 marker genes, and immune-related genes were filtered. Then, we calculated the Spearman correlation for each gene pair. Gene pairs with Spearman correlation coefficients >0.6 (or less than -0.6) and a corresponding p value (Benjamini-Hochberg corrected) <0.01 were considered significantly correlated. Then, the TF-immune gene pair or TF-LM22 gene pair was filtered when the TF-gene pair was found in the Encode TF target database or the TRRUST database. The TF-immune-related gene expression network was built with Cytoscape in which the triangle represented TF, the green represented immune genes, and the red represented LM22 marker genes. Specifically, the connection of TF-immune-related genes with database support were illustrated with blue lines. Orange triangles indicated TF genes, green circles indicated immune-related genes, and pink circles indicated LM22 genes. Node size indicated the number of correlated genes in the gene network. **(C)** Expression profile of genes involved in TF-immune-related genes in **(B)**. **(D-F)** Same as **(A-C)**, except that tumor-associated DEGs from the MEblue module were used.



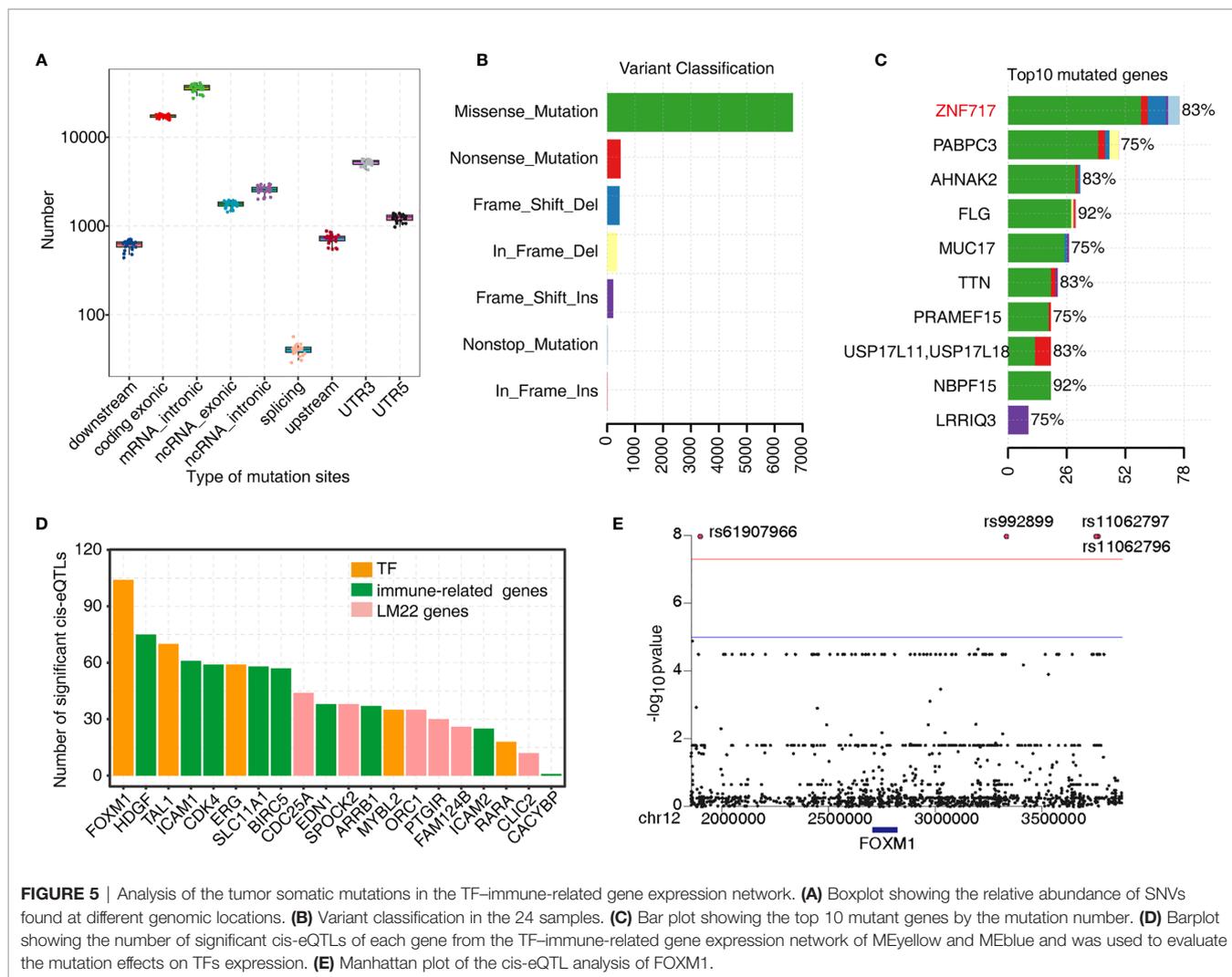
predict immune infiltration. Furthermore, the expression pattern of all genes in two modules was validated using two datasets from TCGA (**Figures S6A, B**), and the results were similar to our findings.

TF Genes in TF-iGENs Showed More Frequent *cis*-eQTL Compared With the Non-TF Genes

As mentioned earlier, TF-iGENs strongly correlated with immune infiltration in the tumor microenvironment. Therefore, investigating the dynamic change in gene expression of TF-iGENs helped explain the altered immune cell population during tumorigenesis. TF-binding sites are frequently somatically mutated in cancer, leading to oncogenic activation (32, 33). To further explore whether the expression alteration of TF-iGENs was associated with somatic mutations, we performed exome sequencing on the paired samples from 24 of the 50 patients; RNA-seq data from all the 50 patients were available and analyzed in this study earlier (**Table S10**). Data were analyzed with R, and the results were visualized with

“maftools” package. The distribution of somatic mutations in different genomic regions is shown (**Figure 5A**). We further classified these mutations according to different categories. As shown in **Figure 5B**, missense mutation was the most common type of variant classification. The bar plot showed the top 10 mutant genes by mutation number (**Figure 5C**), including *ZNF717* (83%), *PABPC3* (75%), *AHNAK2* (83%), *FLG* (92%), *MUC17* (75%), *TTN* (83%), *PRAMEF15* (75%), *USP17L11* (83%), *NBPF15* (92%), and *LRRIQ3* (75%). The waterfall chart was used to demonstrate the frequency mutation profile of the top 20 mutant genes (**Figure S6A**). The functional enrichment analysis on genes with mutations in more than 5/12 patients revealed that the genes were mainly involved in cell differentiation, proliferation, and apoptotic process (**Figures S6B, C**).

Next, we investigated the genetic basis of gene expression variation in TF-iGENs. Matrix eQTL was used for identifying eQTL based on the significance criterion. We classified eQTL with peaks within 1 MB of the transcript position and $p < 0.01$ as *cis*-eQTLs (**Table S11**). The number of significant *cis*-eQTLs of each gene from TF-iGENs is illustrated in **Figure 5D**. The



number of *cis*-eQTLs helped evaluate the mutation effects on gene expression, including TF genes. Mutations were found to be more frequent in the regions flanking the TF genes than those with the non-TF genes (**Figure 5D**). A Manhattan plot of the *cis*-eQTL analysis of FOXM1 was further conducted (**Figure 5E**). Significant variants rs61907966, rs992899, rs11062797, and rs11062796 were identified. The enrichment of somatic mutations in tumor-deregulated TF genes than non-TF-immune-related genes supported the hypothesis that somatic mutation in key immune gene-related TFs might drive the tumor-specific immune cell infiltration.

DISCUSSION

In this study, we identified four key TFs significantly associated with tumor-infiltrating immune cells by analyzing 100 paired RNA-seq data obtained from 50 patients across nine cancer types. First, we estimated the fraction of 21 subpopulations of immune cells in each RNA-seq data using CIBERSORT and

EPIC. Our results showed that the proportions of different types of T cells and macrophages were similarly altered in pan-cancer samples. The WGCNA analysis allowed us to further identify tumor-deregulated gene modules strongly associated with the tumor-induced change in the proportion of TIICs. On further constructing the TF-iGEN in two such modules, FOXM1, MYBL2, TAL1, and ERG were identified as key TFs regulating the expression of genes specifically involved in the immune cell differentiation or activation. Previous studies reported that FOXM1, ERG, and MYBL2 played a role in tumor immune infiltration (16, 34, 35). The possible involvement of TAL1 in regulating tumor immune infiltration was first reported in the present study. We further verified the pan-cancer-regulated expression of these TF-iGENs in TCGA datasets derived from 33 cancer types. Additionally, the tumor-deregulated expression of these genes was strongly associated with the survival rate of pan-cancer patients. The eQTL analysis with exome-seq showed that mutations were significantly more frequent in the regions flanking the TF genes than those with the non-TF genes. These findings suggested that the TF-iGEN identified in this study

might play some roles in regulating immune infiltration and affecting immune therapy and could serve as prognostic biomarkers of pan-cancers (30–32, 36, 37).

A previous study showed that CD4 memory activated T cells resembled the CD62 ligand low (CD62L-low) memory subset, showing rapid activation kinetics and high proliferative capacity (38). The tissue-resident dendritic cells that can capture antigen in peripheral tissues migrate into the lymph node and present peptides to the already activated CD4+ T cells (39), which motivates CD4 memory T cells to function in the microenvironment. In addition, many CD45RA-Foxp3 non-suppressive Treg cells in human cancer could differentiate into memory effector CD4+ T cells (40). This was consistent with our observation that the proportion of CD4 memory activated T cells in pan-cancer tissues was significantly higher than that in adjacent normal tissues, while the proportion of CD4 memory resting T cells was the opposite. Moreover, we demonstrated that both M0 and M1 macrophages were more enriched in tumor samples. Macrophages can switch from an unpolarized (M0) to a polarized (M1) phenotype in response to various stimuli such as tumors (41, 42), and M1 macrophages can exert proinflammatory functions. In contrast, M2 macrophages exert anti-inflammatory functions and facilitate tissue repair (43, 44). Therefore, it is reasonable that tumor induces M0 and M1, but not M2. In our study, M2 macrophages were higher in the non-cancer tissues.

Several TFs have been reported to regulate immune cell infiltration in cancer, such as FOXP1, FOXP3, and c-Maf (10, 12, 45). However, studies investigating TFs associated with immune cell infiltration in pan-cancers are lacking globally. Inkeles et al. previously integrated WGCNA gene modules with cell-type-specific gene signatures to investigate the genes and pathways associated with immune cell types that contributed to host defense and tissue injury at the site of infection in the different subtypes of leprosy (46). Inspired by this work, we used CIBERSORT and WGCNA to explore the correlation between the tumor-deregulated gene expression and tumor-infiltrating immune cell dynamics and further filtered out the TF-iGENs. Four key TFs (FOXM1, MYBL2, TAL1, and ERG) were identified as reliable key regulators in the network.

Previous studies revealed that forkhead box M1 (FoxM1) was overexpressed in HCC, and the induction of FOXM1 led to the hepatic infiltration of macrophages in mice (16). FOXM1 also weakened the promotion of T cell proliferation and depleted IL-12 p70 in tumor-bearing mice (47). Consistently, we found that FOXM1 was overexpressed in the tumor tissues of many different cancer types, and the expression dynamics of the eigengene module containing FOXM1 was positively associated with the increase in the proportion of follicular helper T cells and M0 macrophages. Moreover, the positive correlation between FOXM1 overexpression and the survival rate of multiple cancers was demonstrated. These findings suggested a more general role of FOXM1 in regulating immune infiltration, confirming the efficacy of our methods. The interactions between other TFs and immune infiltration were not studied much and comprehensively. MYBL2 was known to overexpress and associate with poor patient outcomes in many cancer entities (48); its expression positively correlated with CD4+ T cell infiltration but

negatively correlated with B cell infiltration (35). In prostatic cancer, the expression of ERG was positively associated with CD204+ and CD3+ cell infiltration (34). The connection between TAL1 and the regulation of immune infiltration has not been reported due to insufficient data and hence deserves further investigation. Since the expression of these TFs strongly correlated with the infiltration of specific immune cell types, we believed that their important roles in regulating immune infiltration in the tumor microenvironment were revealed in our study.

The immune-related genes in the TF-iGEN are strongly associated with immune defense. Intercellular adhesion molecule 2 increases the antitumor immunity by accelerating the infiltration of immature myeloid dendritic cells in the tumor epithelium, followed by cellular immune responses, and promoting the susceptibility of the tumor cells to cytotoxic T-cell-mediated cytolysis in intraductal papillary mucinous adenoma (49). The solute carrier family 11a member 1 modulates macrophage activation by regulating immune-inflammation genes in macrophages (50); these include tumor necrosis factor- α (TNF- α), interferon gamma (IFN- γ), and IL-1 (51), and major histocompatibility complex (MHC) class II expression (52). The inhibitor of cyclin-dependent kinases 4 (CDK4), which has already been approved by the Food and Drug Administration (FDA) for the treatment of breast cancer (53, 54), markedly repressed the proliferation of Tregs (55). This finding was consistent with our research that the expression of CDK4 was positively associated with the population of Tregs. Therefore, the pan-cancer-associated TF-iGEN is functionally related to immune infiltration, which likely represents a mechanism underlying the altered immune cell contexture in the tumor microenvironment. Moreover, the expressions of many TFs and immune-related genes in the networks increase or decrease significantly during cancerization, suggesting that individually or combined, they can serve as indicators to predict the prognosis situation. It was remarkable that most of the genes in our TF-iGENs showed a significant prognostic value in brain lower-grade glioma, adrenocortical carcinoma, liver hepatocellular carcinoma, sarcoma, and kidney renal clear cell carcinoma, which demonstrated the clinical relevance of our results.

Mutations in both coding or non-coding areas (56) could affect transcriptional regulatory mechanisms. Many TFs have been found to have a high mutation frequency and are related to the occurrence and development of tumors. ATBF1 encodes a transcription factor that could inhibit cell proliferation. ATBF1 messenger RNA (mRNA) is abundant in normal prostates but more scarce in approximately half of prostate cancers tested. Frequent somatic mutations of the transcription factor ATBF1 in human prostate cancer were found, many of which impair ATBF1 function (57). RUNX1, another transcription factor mutated in breast cancer, was found as a key regulator of the ER+ luminal lineage whose disruption may contribute to the development of ER+ luminal breast cancer when under the background of either TP53 or RB1 loss (58). Mutations of TP53 (59) significantly correlated with the programmed death 1, programmed death-ligand 1, and programmed death-ligand 2

axis (60–62), which had a significant influence on immune infiltration. The eQTL analysis in the genes of our TF-iGEN showed that mutations were significantly more frequent in the regions flanking the TF genes than those with the non-TF genes, supporting a potential mechanism that somatic mutation in key immune gene-related TFs might drive the altered expression of immune infiltration during cancerization and consequently altered immune infiltration. This hypothesis requires further investigation.

DATA AVAILABILITY STATEMENT

The data reported in this paper have been deposited with the NCBI Gene Expression Omnibus (GEO, <https://www.ncbi.nlm.nih.gov/geo/>) under accession number GSE87410.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the ethics committee of the China–Japan Union Hospital of Jilin University. The patients/participants provided their written informed consent to participate in this study.

REFERENCES

1. Walk EE, Yohe SL, Beckman A, Schade A, Zutter MM, Pfeifer J, et al. The Cancer Immunotherapy Biomarker Testing Landscape. *Arch Pathol Lab Med* (2020) 144(6):706–24. doi: 10.5858/arpa.2018-0584-CP
2. Galluzzi L, Chan TA, Kroemer G, Wolchok JD, López-Soto A. The Hallmarks of Successful Anticancer Immunotherapy. *Sci Trans Med* (2018) 10(459): eaat7807. doi: 10.1126/scitranslmed.aat7807
3. Fridman WH, Zitvogel L, Sautès-Fridman C, Kroemer G. The Immune Contexture in Cancer Prognosis and Treatment. *Nat Rev Clin Oncol* (2017) 12:717. doi: 10.1038/nrclinonc.2017.101
4. Jérôme G, Anne C, Fatima S-C, Amos K, Bernhard M, Christine L-P, et al. Type, Density, and Location of Immune Cells Within Human Colorectal Tumors Predict Clinical Outcome. *Science* (2006) 313(5795):1960. doi: 10.1126/science.1129139
5. Galon J, Bruni D. Approaches to Treat Immune Hot, Altered and Cold Tumours With Combination Immunotherapies. *Nat Rev Drug Discov* (2019) 18(3):197–218. doi: 10.1038/s41573-018-0007-y
6. Racle J, de Jonge K, Baumgaertner P, Speiser DE, Gfeller D. Simultaneous Enumeration of Cancer and Immune Cell Types From Bulk Tumor Gene Expression Data. *ELife* (2017) 6:e26476. doi: 10.7554/eLife.26476
7. Li B, Severson E, Pignon J-C, Zhao H, Li T, Novak J, et al. Comprehensive Analyses of Tumor Immunity: Implications for Cancer Immunotherapy. *Genome Biol* (2016) 17(1):174. doi: 10.1186/s13059-016-1028-7
8. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust Enumeration of Cell Subsets From Tissue Expression Profiles. *Nat Methods* (2015) 12(5):453–7. doi: 10.1038/nmeth.3337
9. De Panfilis G, Campanini N, Santini M, Mori G, Tognetti E, Maestri R, et al. Phase- and Stage-Related Proportions of T Cells Bearing the Transcription Factor FOXP3 Infiltrate Primary Melanoma. *J Invest Dermatol* (2008) 128(3):676–84. doi: 10.1038/sj.jid.5701046
10. De Silva P, Garaud S, Solinas C, de Wind A, Van den Eyden G, Jose V, et al. FOXP1 Negatively Regulates Tumor Infiltrating Lymphocyte Migration in Human Breast Cancer. *EBioMedicine* (2019) 39:226–38. doi: 10.1016/j.ebiom.2018.11.066
11. Wang Y, Shen Y, Wang S, Shen Q, Zhou X. The Role of STAT3 in Leading the Crosstalk Between Human Cancers and the Immune System. *Cancer Lett* (2018) 415:117–28. doi: 10.1016/j.canlet.2017.12.003

AUTHOR CONTRIBUTIONS

LS, XC, and YZ conceived and supervised the study. LL, JY, and HS performed the experiments. CC, QW, and QZ performed data analyses. LL, QZ, and CC wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was funded by the Research Foundation of Jilin Provincial Science & Technology Committee (Nos. 20200201135JC and 20200404124YY) and the Scientific Research Foundation of Jilin Province (20200601010JC and 20190701061GH).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fimmu.2021.644350/full#supplementary-material>

12. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical Features of Patients Infected With 2019 Novel Coronavirus in Wuhan, China. *Lancet* (2020) 395(10223):497–506. doi: 10.1016/S0140-6736(20)30183-5
13. Raulet DH, Marcus A, Coscoy L. Dysregulated Cellular Functions and Cell Stress Pathways Provide Critical Cues for Activating and Targeting Natural Killer Cells to Transformed and Infected Cells. *Immunol Rev* (2017) 280(1):93–101. doi: 10.1111/imr.12600
14. Ni J, Wang X, Stojanovic A, Zhang Q, Wincher M, Bühler L, et al. Single-Cell RNA Sequencing of Tumor-Infiltrating NK Cells Reveals That Inhibition of Transcription Factor HIF-1 α Unleashes NK Cell Activity. *Immunity* (2020) 52(6):1075–87. doi: 10.1016/j.immuni.2020.05.001
15. Powell N, Canavan JB, MacDonald TT, Lord GM. Transcriptional Regulation of the Mucosal Immune System Mediated by T-Bet. *Mucosal Immunol* (2010) 3(6):567–77. doi: 10.1038/mi.2010.53
16. Kurahashi T, Yoshida Y, Ogura S, Egawa M, Furuta K, Hikita H, et al. Forkhead Box M1 Transcription Factor Drives Liver Inflammation Linking to Hepatocarcinogenesis in Mice. *Cell Mol Gastroenterol Hepatol* (2020) 9(3):425–46. doi: 10.1016/j.jcmgh.2019.10.008
17. Kim D, Pertege A, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: Accurate Alignment of Transcriptomes in the Presence of Insertions, Deletions and Gene Fusions. *Genome Biol* (2013) 14(4):R36. doi: 10.1186/gb-2013-14-4-r36
18. Robinson MD, McCarthy DJ, Smyth GK. EdgeR: A Bioconductor Package for Differential Expression Analysis of Digital Gene Expression Data. *Bioinformatics* (2010) 26(1):139–40. doi: 10.1093/bioinformatics/btp616
19. Sturm G, Finotello F, List M. Immunedeconv: An R Package for Unified Access to Computational Methods for Estimating Immune Cell Fractions From Bulk RNA-Sequencing Data. *Methods Mol Biol* (2020) 2120:223–32. doi: 10.1007/978-1-0716-0327-7_16
20. 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. doi: 10.1007/978-1-0716-0327-7_17
21. Langfelder P, Horvath S. WGCNA: An R Package for Weighted Correlation Network Analysis. *BMC Bioinf* (2008) 9:559. doi: 10.1186/1471-2105-9-559
22. Tang Z, Li C, Kang B, Gao G, Li C, Zhang Z. GEPIA: A Web Server for Cancer and Normal Gene Expression Profiling and Interactive Analyses. *Nucleic Acids Res* (2017) 45(W1):W98–102. doi: 10.1093/nar/gkx247

23. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map Format and SAMtools. *Bioinformatics* (2009) 25 (16):2078–9. doi: 10.1093/bioinformatics/btp352
24. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernysky A, et al. The Genome Analysis Toolkit: A MapReduce Framework for Analyzing Next-Generation DNA Sequencing Data. *Genome Res* (2010) 20(9):1297–303. doi: 10.1101/gr.107524.110
25. Wang K, Li M, Hakonarson H. ANNOVAR: Functional Annotation of Genetic Variants From High-Throughput Sequencing Data. *Nucleic Acids Res* (2010) 38(16):e164. doi: 10.1093/nar/gkq603
26. Mayakonda A, Lin DC, Assenov Y, Plass C, Koeffler HP. Maftools: Efficient and Comprehensive Analysis of Somatic Variants in Cancer. *Genome Res* (2018) 28(11):1747–56. doi: 10.1101/gr.239244.118
27. Shabalin AA. Matrix eQTL: Ultra Fast eQTL Analysis via Large Matrix Operations. *Bioinformatics* (2012) 28(10):1353–8. doi: 10.1093/bioinformatics/bts163
28. Xie C, Mao X, Huang J, Ding Y, Wu J, Dong S, et al. KOBAS 2.0: A Web Server for Annotation and Identification of Enriched Pathways and Diseases. *Nucleic Acids Res* (2011) 39(Web Server issue):316–22. doi: 10.1093/nar/gkr483
29. Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, et al. Metascape Provides a Biologist-Oriented Resource for the Analysis of Systems-Level Datasets. *Nat Commun* (2019) 10(1):1523. doi: 10.1038/s41467-019-09234-6
30. Stephan P, Lautreite R, Voisin A, Grinberg-Bleyer Y. Transcriptional Control of Regulatory T Cells in Cancer: Toward Therapeutic Targeting? *Cancers (Basel)* (2020) 12(11):3194. doi: 10.3390/cancers12113194
31. Wang H, Morse HC 3rd, Bolland S. Transcriptional Control of Mature B Cell Fates. *Trends Immunol* (2020) 41(7):601–13. doi: 10.1016/j.it.2020.04.011
32. Capasso M, Lasorsa VA, Cimmino F, Avitabile M, Cantalupo S, Montella A, et al. Transcription Factors Involved in Tumorigenesis Are Over-Represented in Mutated Active DNA-Binding Sites in Neuroblastoma. *Cancer Res* (2020) 80(3):382–93. doi: 10.1158/0008-5472.CAN-19-2883
33. Le Gallo M, Rudd ML, Urlick ME, Hansen NFP. National Institutes of Health Intramural Sequencing Center Comparative Sequencing, Merino MJ, et al. The FOXA2 Transcription Factor Is Frequently Somatic Mutated in Uterine Carcinosarcomas and Carcinomas. *Cancer* (2018) 124(1):65–73. doi: 10.1002/cncr.30971
34. Burdova A, Rulisek P, Bouchal J, Král M, Student V, Kolar Z. Infiltration of Prostate Cancer by CD204+ and CD3+ Cells Correlates With ERG Expression and TMPRSS2-ERG Gene Fusion. *Klinicka Onkologie Casopis Ceske Slovenske Onkologicke Spolecnosti* (2018) 31(6):421–8. doi: 10.14735/amko2018421
35. Gu H-Y, Zhang C, Guo J, Yang M, Zhong H-C, Jin W, et al. Risk Score Based on Expression of Five Novel Genes Predicts Survival in Soft Tissue Sarcoma. *Aging* (2020) 12(4):3807–27. doi: 10.18632/aging.102847
36. Ge P, Wang W, Li L, Zhang G, Gao Z, Tang Z, et al. Profiles of Immune Cell Infiltration and Immune-Related Genes in the Tumor Microenvironment of Colorectal Cancer. *BioMed Pharmacother* (2019) 118:109228. doi: 10.1016/j.biopha.2019.109228
37. Xiong Y, Wang K, Zhou H, Peng L, You W, Fu Z. Profiles of Immune Infiltration in Colorectal Cancer and Their Clinical Significant: A Gene Expression-Based Study. *Cancer Med* (2018) 7(9):4496–508. doi: 10.1002/cam4.1745
38. Ahmadzadeh M, Hussain SF, Farber DL. Heterogeneity of the Memory CD4 T Cell Response: Persisting Effectors and Resting Memory T Cells. *J Immunol (Baltimore Md 1950)* (2001) 166(2):926–35. doi: 10.4049/jimmunol.166.2.926
39. Palucka K, Banchereau J. Cancer Immunotherapy via Dendritic Cells. *Nat Rev Cancer* (2012) 12(4):265–77. doi: 10.1038/nrc3258
40. Yung-Chang L, Jayashri M, Jy-Ming C, Po-Jung S, Yu-Yi C, Hsin-Yi L, et al. Activated But Not Resting Regulatory T Cells Accumulated in Tumor Microenvironment and Correlated With Tumor Progression in Patients With Colorectal Cancer. *Int J Of Cancer* (2013) 132(6):1341–50. doi: 10.1002/ijc.27784
41. Yu X, Xiao-Tao H, Xin-Yue X, Bei-Min T, Ying A, Fa-Ming C. Exosomes Derived From M0, M1 and M2 Macrophages Exert Distinct Influences on the Proliferation and Differentiation of Mesenchymal Stem Cells. *PeerJ* (2020) 8:e8970–0. doi: 10.7717/peerj.8970
42. Mantovani A, Biswas SK, Galdiero MR, Sica A, Locati M. Macrophage Plasticity and Polarization in Tissue Repair and Remodelling. *J Pathol* (2013) 229(2):176–85. doi: 10.1002/path.4133
43. Murray PJ, Allen JE, Biswas SK, Fisher EA, Gilroy DW, Goerdts S, et al. Macrophage Activation and Polarization: Nomenclature and Experimental Guidelines. *Immunity* (2014) 41(1):14–20. doi: 10.1016/j.immuni.2014.06.008
44. Shapouri-Moghaddam A, Mohammadian S, Vazini H, Taghadosi M, Esmaeili S-A, Mardani F, et al. Macrophage Plasticity, Polarization, and Function in Health and Disease. *J Cell Physiol* (2018) 233(9):6425. doi: 10.1002/jcp.26429
45. West NR, Kost SE, Martin SD, Milne K, deLeeuw RJ, Nelson BH, et al. Tumour-Infiltrating FOXP3+ Lymphocytes Are Associated With Cytotoxic Immune Responses and Good Clinical Outcome in Oestrogen Receptor-Negative Breast Cancer. *Br J Cancer* (2012) 108(1):155–62. doi: 10.1038/bjc.2012.524
46. Inkeles MS, Teles RM, Poulard D, Andrade PR, Madigan CA, Lopez D, et al. Cell-Type Deconvolution With Immune Pathways Identifies Gene Networks of Host Defense and Immunopathology in Leprosy. *JCI Insight* (2016) 1(15):e88843. doi: 10.1172/jci.insight.88843
47. Zhongshi Z, Hongdan C, Rui X, Hongjie W, Senlin L, Qianqian X, et al. Epigenetically Modulated FOXM1 Suppresses Dendritic Cell Maturation in Pancreatic Cancer and Colon Cancer. *Mol Oncol* (2019) 13(4):873–93. doi: 10.1002/1878-0261.12443
48. Musa J, Aynaud M-M, Mirabeau O, Delattre O, Grünewald TGP. MYBL2 (B-Myb): A Central Regulator of Cell Proliferation, Cell Survival and Differentiation Involved in Tumorigenesis. *Cell Death Dis* (2017) 8(6):e2895–5. doi: 10.1038/cddis.2017.244
49. Hiraoka N, Yamazaki-Itoh R, Ino Y, Mizuguchi Y, Yamada T, Hirohashi S, et al. CXCL17 and ICAM2 Are Associated With a Potential Anti-Tumor Immune Response in Early Intraepithelial Stages of Human Pancreatic Carcinogenesis. *Gastroenterology* (2011) 140(1):310–21. doi: 10.1053/j.gastro.2010.10.009
50. Correa M, Canhamero T, Borrego A, Katz I, Jensen J, Guerra J, et al. Slc11a1 (Nramp-1) Gene Modulates Immune-Inflammation Genes in Macrophages During Pristane-Induced Arthritis in Mice. *Inflammation Res* (2017) 66(11):969–80. doi: 10.1007/s00011-017-1077-8
51. Lalmanach AC, Montagne A, Menanteau P, Lantier F. Effect of the Mouse Nramp1 Genotype on the Expression of IFN-Gamma Gene in Early Response to Salmonella Infection. *Microbes Infect* (2001) 3(8):639–44. doi: 10.1016/S1286-4579(01)01419-8
52. Wojciechowski W, DeSanctis J, Skamene E, Radziach D. Attenuation of MHC Class II Expression in Macrophages Infected With Mycobacterium Bovis Bacillus Calmette-Guérin Involves Class II Transactivator and Depends on the Nramp1 Gene. *J Immunol (Baltimore Md 1950)* (1999) 163(5):2688–96.
53. Yu Q, Sicinska E, Geng Y, Ahnström M, Zagodzón A, Kong Y, et al. Requirement for CDK4 Kinase Function in Breast Cancer. *Cancer Cell* (2006) 9(1):23–32. doi: 10.1016/j.ccr.2005.12.012
54. Zhang Q-F, Li J, Jiang K, Wang R, Ge J-L, Yang H, et al. CDK4/6 Inhibition Promotes Immune Infiltration in Ovarian Cancer and Synergizes With PD-1 Blockade in a B Cell-Dependent Manner. *Theranostics* (2020) 10(23):10619–33. doi: 10.7150/thno.44871
55. Goel S, DeCristo MJ, Watt AC, BrinJones H, Sceneay J, Li BB, et al. CDK4/6 Inhibition Triggers Anti-Tumour Immunity. *Nature* (2017) 548(7668):471–5. doi: 10.1038/nature23465
56. Maurano MT, Hao W, Kutayin T, Stamatoypoulos JA. Widespread Site-Dependent Buffering of Human Regulatory Polymorphism. *PLoS Genet* (2012) 8(3):1–12. doi: 10.1371/journal.pgen.1002599
57. Sun X, Frierson HF, Chen C, Li C, Ran Q, Otto KB, et al. Frequent Somatic Mutations of the Transcription Factor ATBF1 in Human Prostate Cancer. *Nat Genet* (2005) 37(4):407–12. doi: 10.1038/ng1528
58. van Bragt MP, Hu X, Xie Y, Li Z, RUNXI, a Transcription Factor Mutated in Breast Cancer, Controls the Fate of ER-Positive Mammary Luminal Cells. *Elife* (2014) 3:e03881. doi: 10.7554/eLife.03881
59. Bykov VJN, Eriksson SE, Bianchi J, Wiman KG. Targeting Mutant P53 for Efficient Cancer Therapy. *Nat Rev Cancer* (2018) 18(2):89–102. doi: 10.1038/nrc.2017.109
60. Xu G, Qi F, Li H, Yang Q, Wang H, Wang X, et al. The Differential Immune Responses to COVID-19 in Peripheral and Lung Revealed by Single-Cell RNA Sequencing. *Cell Discov* (2020) 6(1):73. doi: 10.1038/s41421-020-00225-2
61. Serra P, Petat A, Maury J-M, Thivolet-Bejui F, Chalabreysse L, Barritault M, et al. Programmed Cell Death-Ligand 1 (PD-L1) Expression Is Associated With RAS/TP53 Mutations in Lung Adenocarcinoma. *Lung Cancer* (2018) 118:62–8. doi: 10.1016/j.lungcan.2018.02.005
62. Wieser V, Gaugg I, Fleischer M, Shivalingiah G, Wenzel S, Sprung S, et al. BRCA1/2 and TP53 Mutation Status Associates With PD-1 and PD-L1 Expression in Ovarian Cancer. *Oncotarget* (2018) 9(25):17501–11. doi: 10.18632/oncotarget.24770

Conflict of Interest: QZ, CC, QW, WQ, YX, and YZ were employed by the company ABLife BioBigData Institute.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in

this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Liu, Zhao, Cheng, Yi, Sun, Wang, Quan, Xue, Sun, Cong and Zhang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.