

## Research Article

# Bioinformatics Analysis of GFAP as a Potential Key Regulator in Different Immune Phenotypes of Prostate Cancer

Wencheng Yao,<sup>1,2</sup> Xiang Li,<sup>1,2</sup> Zhankui Jia ,<sup>1,2</sup> Chaohui Gu,<sup>1,2</sup> Zhibo Jin,<sup>1,2</sup> Jun Wang,<sup>1,2</sup> Bo Yuan,<sup>1,2</sup> and Jinjian Yang <sup>1,2</sup>

<sup>1</sup>Department of Urology, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan, China

<sup>2</sup>Henan Institute of Urology, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan, China

Correspondence should be addressed to Jinjian Yang; fccyangjj@zzu.edu.cn

Received 15 April 2021; Revised 26 May 2021; Accepted 5 June 2021; Published 17 June 2021

Academic Editor: Kyoung-Ho Pyo

Copyright © 2021 Wencheng Yao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Tumor immune escape plays an essential role in both cancer progression and immunotherapy responses. For prostate cancer (PC), however, the molecular mechanisms that drive its different immune phenotypes have yet to be fully elucidated. Patient gene expression data were analyzed from The Cancer Genome Atlas-prostate adenocarcinoma (TCGA-PRAD) and the International Cancer Genome Consortium (ICGC) databases. We used a Cell-type Identification by Estimating Relative Subsets of RNA Transcripts (CIBERSORT) analysis and an unsupervised clustering analysis to identify patient subgroups with distinct immune phenotypes. These distinct phenotypes were then explored for associations for differentially expressed genes (DEGs) and both epigenetic and genetic landscapes. Finally, we used a protein-protein interaction analysis to identify key hub genes. We identified two patient subgroups with independent immune phenotypes associated with the expression of Programmed death-ligand 1 (PD-L1). Patient samples in Cluster 1 (C1) had higher scores for immune-cell subsets compared to Cluster 2 (C2), and C2 samples had higher specific somatic mutations, MHC mutations, and genomic copy number variations compared to C1. We also found additional cluster phenotype differences for DNA methylation, microRNA (miRNA) expression, and long noncoding RNA (lncRNA) expression. Furthermore, we established a 4-gene model to distinguish between clusters by integrating analyses for DEGs, lncRNAs, miRNAs, and methylation. Notably, we found that glial fibrillary acidic protein (GFAP) might serve as a key hub gene within the genetic and epigenetic regulatory networks. These results improve our understanding of the molecular mechanisms underlying tumor immune phenotypes that are associated with tumor immune escape. In addition, GFAP may be a potential biomarker for both PC diagnosis and prognosis.

## 1. Introduction

Prostate cancer (PC) is the most frequently diagnosed type of cancer and ranks second among the leading causes of cancer-related deaths in men [1]. Although early detection has improved significantly, and the 5-year survival rate for early-stage cancer is excellent, the rate of false-positives is high for the prostate-specific antigen (PSA) screening test, and prostate metastatic disease is associated with both poor outcomes and deceptively low serum PSA levels that can lead to false-negative results [2]. Thus, it is critical to identify novel prognostic biomarkers and therapeutic strategies for improving overall survival (OS) for patients with PC.

Several studies have demonstrated that the tumor micro-environment has both tumor-antagonizing and tumor-promoting roles. Tumor-antagonizing immune cells mainly consist of M1-polarized macrophages, CD8<sup>+</sup> cytotoxic T cells, effector CD4<sup>+</sup>T cells, natural killer (NK) cells, dendritic cells (DCs), and N1-polarized neutrophils which are able to either present tumor cells, kill them directly, or secrete cytokines that interact with malignant cells [3, 4]. Tumor-promoting immune cells mainly include regulatory T cells (Tregs) and myeloid-derived suppressor cells (MDSCs) [3]. Notably, cancer cells can eventually evade immune surveillance through a variety of mechanisms and resist the cytotoxic effects of immune cells. As a new hallmark of cancer,

this ability of immune escape provides new opportunities for cancer treatment strategies. Recently, four groups of microenvironment immune types have been categorized using genomic analysis of 14 types of solid cancers. However, considering their complexity and the dynamic nature of immune cells, it is evident that a more in-depth evaluation of the associations between immune cells and PC is needed.

The programmed cell death-1 receptor (PD-1) is a representative immune checkpoint inhibitor expressed in activated T cells, B cells, and macrophages. Programmed death-ligand 1 (PD-L1), a ligand of PD-1, is often found on both tumor cells and antigen-presenting cells and provides potent inhibitory interactions within the tumor microenvironment [5]. The major function of PD-L1 limits tumor cells from evading immunity, but unfortunately, this has become a main mechanism for immune resistance in the immune microenvironment. In addition, the expression of PD-L1 in tumors is often regarded as a negative prognostic factor, but it is clearly a positive factor for PD-1/PD-L1 treatments [6].

In this PC study, we established two patient-sample clusters with different immunophenotypes using an integrative multiomic approach to better understand the molecular mechanisms of tumor immune escape and found the key regulatory node within the genetic and epigenetic regulatory networks. Furthermore, this key regulatory node may be a novel tumor immune escape predictor and a promising target for PC therapy.

## 2. Materials and Methods

**2.1. Patients and Gene Expression Profiles.** Clinical and RNA-expression patient data in pan-cancer were collected from the prostate adenocarcinoma dataset of The Cancer Genome Atlas (TCGA-PRAD; <https://portal.gdc.cancer.gov/>) and TCGA-PRAD serves as training datasets. The cohorts used for validation studies were downloaded from the International Cancer Genome Consortium (ICGC; <https://dcc.icgc.org/>). All patient data that were used in the present study had complete clinical information, including age, sex, grade, TNM stage, and overall survival time. The PRAD protein-expression data were obtained from <https://www.tcaportal.org/tcpa>. Functional lncRNA and microRNA data were downloaded from the TCGA data portal (<https://portal.gdc.cancer.gov/>) [7].

**2.2. Cell-Type Identification by Estimating Relative Subsets of RNA Transcripts (CIBERSORT) Analyses.** The CIBERSORT computational approach is commonly used to predict the infiltration of 22 types of immune cells from the gene expression profiles of complex tissues (<http://cibersort.stanford.edu>) [8]. The 22 types of immune cells include 7 types of T cells, 3 types of macrophages, naïve B cells, memory B cells, activated NK cells, plasma cells, monocytes, resting natural killer (NK) cells, resting mast cells, activated mast cells, resting dendritic cells, activated dendritic cells, neutrophils, and eosinophils.

In the present study, CIBERSORT is used to calculate the absolute immune-cell fraction in the PC samples

(model = absolute, permutation = 1,000, disable quantile normalization for RNA-Seq data as recommended) using the method of estimating the relative subset of RNA transcript for cell type identification.

**2.3. Unsupervised Hierarchical Cluster Analysis.** Immune-cell expression values were clustered using an unsupervised hierarchical clustering method (Euclidean distance and ward linkage) provided by the dendextend R software package [9]. The ClustVis tool (<http://biit.cs.ut.ee/clustvis/>) was used to process and modify these data for the final plotting of the principal component analysis and the heat map [10].

**2.3.1. Comparing Somatic Mutation Frequencies and Somatic Copy Number Variants (CNVs).** To assess the relative frequency of somatic mutation between the different immunophenotypes, the somatic mutation sequencing data from TCGA-PRAD clinical samples were acquired from cBioPortal (<http://www.cbioportal.org/>) [11] and were analyzed using the Complex Heatmap R software package [12]. Total gene mutations frequencies  $\geq 5\%$  for DNA damage responses and repair factors were identified. In addition, the characteristic molecular profiles of the cohorts were determined.

Level 3 CNV data from TCGA were retrieved from the firebrowse data portal (<http://firebrowse.org/>) using a threshold of 0.2 for segmented mean amplification values and  $-0.2$  for deletion values. Fisher's exact test was used for determining statistical significance (<http://convaq.combio.sdu.dk/>), using the average number of deletions and duplications per sample as background data, and the number of samples with deletions and amplifications for the loci [13]. The CNV summary diagram was generated by IGV-2.2.19, and the Circos figure was analyzed using the RCircos R software package.

**2.4. The Single-Sample Gene Set Enrichment Analysis (ssGSEA).** As described previously, ssGSEA was used to score each sample based on the list of screened genes in the TCGA-PRAD cohort. To determine any correlations between the immune phenotypes and the screened genes and to verify grouping accuracy, we use the ssGSEA algorithm to calculate the ssGSEA enrichment score of immune genes in the ICGC-PRAD-CA cohort based on the published immune-related gene set genes (doi:10.1172/JCI91190).

**2.5. Target-Genes Predictions for MicroRNAs (miRNAs) and Long Noncoding RNAs (lncRNAs).** To study miRNA functionality, the target mRNAs for identified miRNAs were predicted using DIANA (<http://diana.imis.athena-innovation.gr/DianaTools/index.php>). The screening threshold used was a miTG score of  $\geq 0.95$ . Target-gene predictions for identified lncRNAs were made using lncRNA2Target software (<http://123.59.132.21/lncrna2target/download.jsp>).

**2.6. DNA Methylation Analysis.** The DNA methylation analyses were performed using Illumina 450K chip-based information from the TCGA-PRAD database (<http://gdc.xenahubs.net/download/TCGA-PRAD.methylation450.tsv.gz/>). The methylation data were standardized using the watermelon package in R software, and the probes for

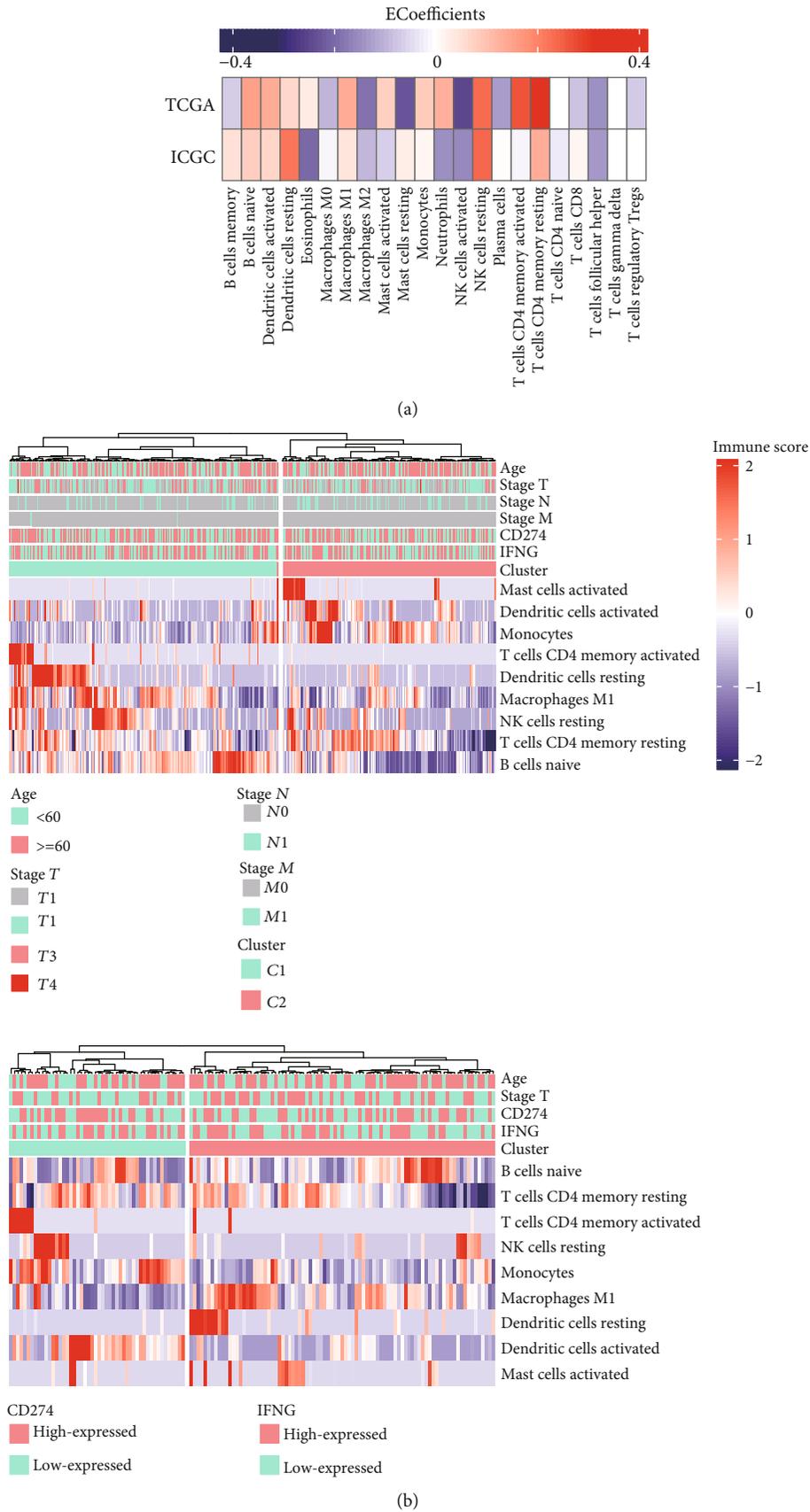


FIGURE 1: Continued.

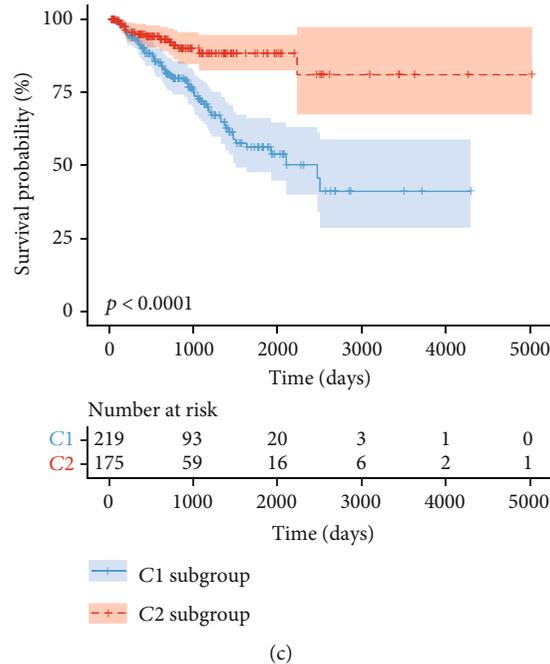


FIGURE 1: The immune-related subtypes for prostate cancer. (a) Correlations between the expression of PD-L1 and immune-cell infiltration ratios in TCGA-PRAD and ICGC-PRAD cohorts. (b) The distribution of immune-related subtypes and associated clinical characteristics of TCGA (top) and ICGC (bottom) cohorts. (c) Kaplan-Meier analysis of groups C1 and C2 in TCGA-PRAD cohort.

differential methylations were detected using the minfi package in R software.

**2.7. Analysis of Protein-Protein Interactions (PPIs).** The PPI network was analyzed using the STRING Consortium online database (<https://string-db.org/>). In Cytoscape (version 3.6.1) software, the PPI network was analyzed using Cytoscape plugin-MCODE (for clustering of the PPI network) and cytoHubba (for identifying hub proteins), as described previously [14].

**2.8. Statistical Analyses.** All statistical analyses were performed using SPSS 25 (IBM, Chicago, IL, USA) and GraphPad Prism 7.0 (GraphPad Software, La Jolla, CA, USA). The versions of R software used were more recent than v.3.5.1. For the correlation analyses, Student's *t*-tests were used. A univariate analysis was performed to analyze possible prognostic factors. All statistical results with  $p < 0.05$  were considered statistically significant.

### 3. Results

**3.1. Identification of the Immune-Cell Subsets Related to PD-L1 Expression in PC Samples.** To evaluate the infiltrations of different immune-cell subtypes, a training set (TCGA-PRAD,  $n = 519$ ) and a validation cohort (ICGC-PRAD-CA,  $n = 144$ ) were used for the CIBERSORT analysis. Overall, nine immune-cell subtypes (activated mast cells, monocytes, M1 macrophages, activated dendritic cells, naive B cells, activated memory T cells-CD4, resting dendritic cells, resting NK cells, and resting memory T cells-CD4) were identified as being positively correlated with PD-L1 transcript levels

(Figure 1(a)). These nine cell types were then further characterized.

**3.2. Classification of PC Samples into Subgroups Based on Immune-Cell Subsets.** To better describe the relationship between tumor-infiltrating immune cells and tumor-cell immune escape, we classified the PC sample data into two major clusters (C1 and C2) based on the above nine immune-cell subsets (Figure 1(b)). Patient samples in group C1 had higher scores for selected immune-cell subsets compared to C2 samples (Figure 1(b)), so C1 patient samples were defined as having a high cell-cytotoxic immune phenotype, and C2 patient samples were defined as having a low cell-cytotoxic immune phenotype. The same findings were also verified using the ICGC-PRAD-CA validation cohort. A Kaplan-Meier analysis demonstrated that C1 patients also had shorter overall survive times than C2 patients according to TCGA-PRAD dataset (Figure 1(c)).

**3.3. Differences in MHC Class I Gene Expression Related to Immune Phenotypes.** MHC class I (MHC-I) tumor antigens expressed on tumor-cell surfaces determine the capacity of cytotoxic T lymphocytes (CTLs) to recognize and eliminate tumor cells. The lack of MHC-I antigens on CTLs is an important mechanism leading to tumor-cell immune evasion. The expression of MHC-I antigens on cancer cells also affects tumor responses to immunotherapy.

We therefore evaluated the transcript levels of  $\beta 2$ -microglobulin (B2M) and human leukocyte antigen (HLA) genes encoding MHC class I protein in the independent TCGA-PRAD validation cohort. The results showed that C2 expressions were lower than C1 expressions (Figure 2(a)), indicating

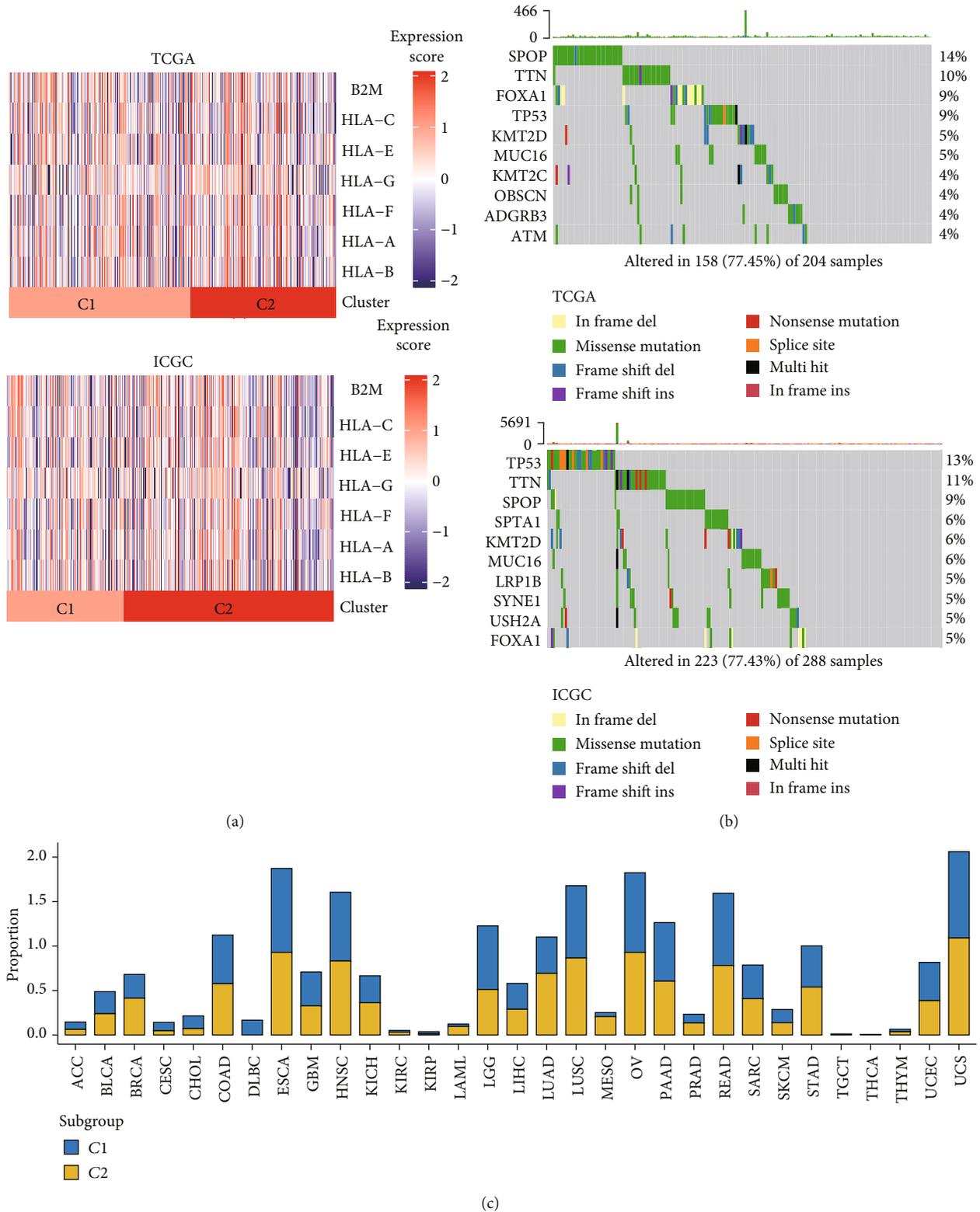


FIGURE 2: Differences in MHC class I gene expression and somatic mutations between the two immune phenotypes. (a) The transcription levels for B2M and for human leukocyte antigen (HLA) genes encoding MHC class I protein in TCGA-PRAD (top) and ICGC-PRAD cohorts (bottom). (b) The top-10 mutated genes in groups C1 (top) and C2 (bottom) in the PRAD database. (c) In TCGA-PRAD cohort, the mutation frequency of TP53 in group C2 was relatively higher than in group C1 for most types of cancer.

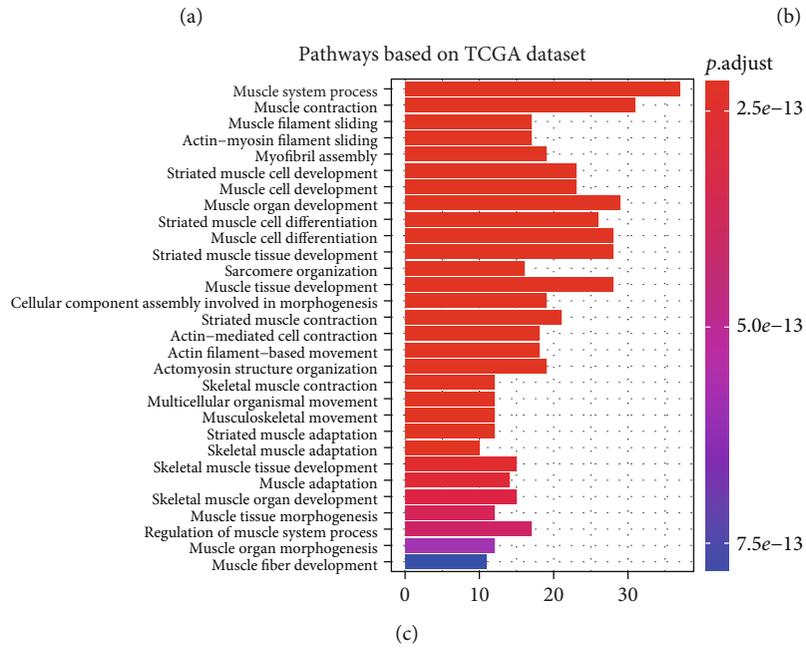
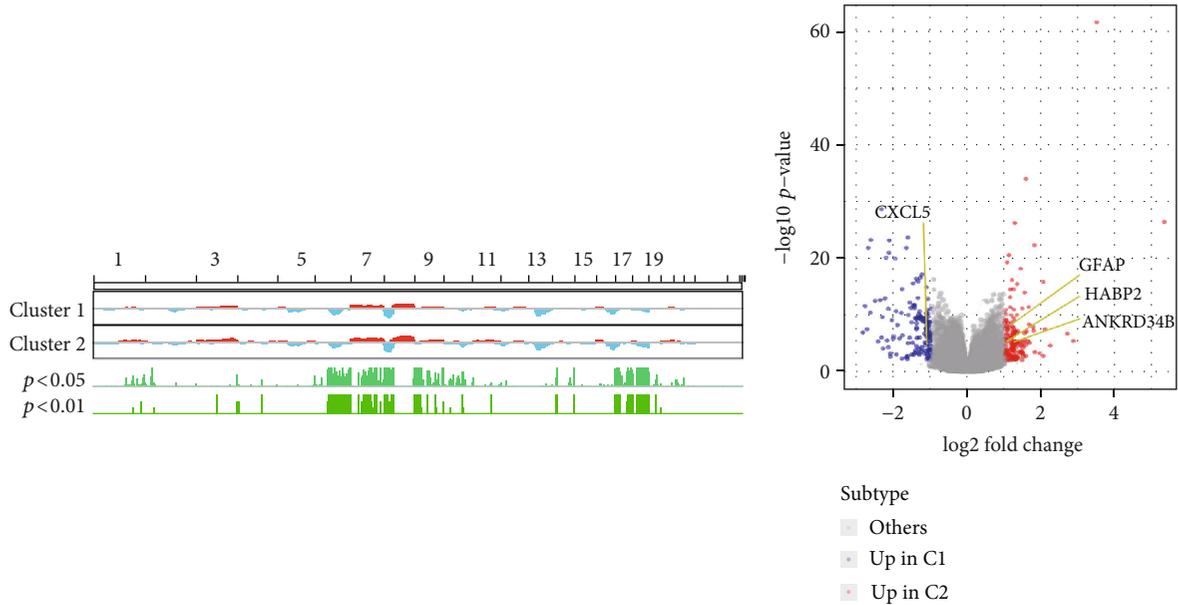


FIGURE 3: Continued.

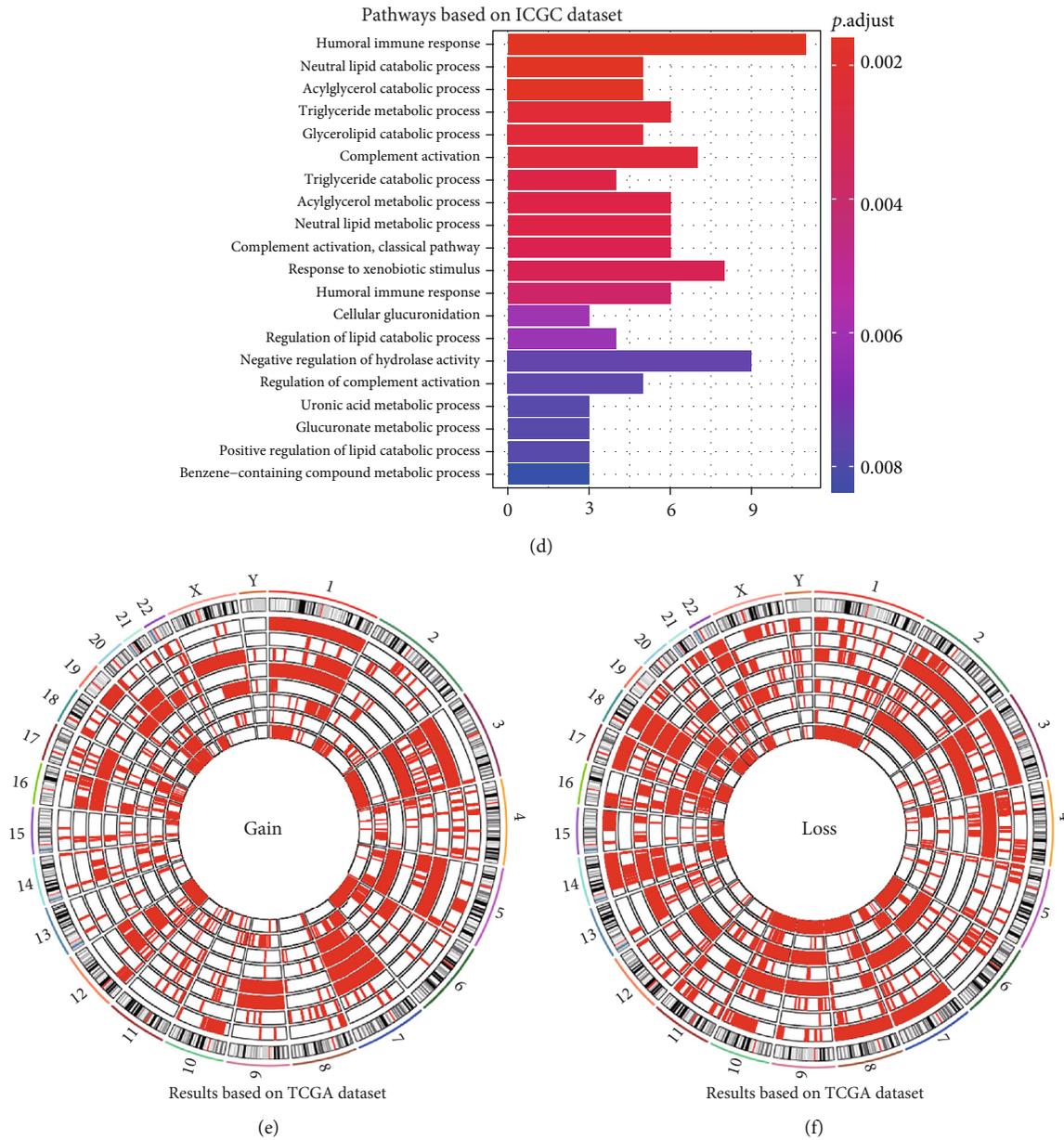


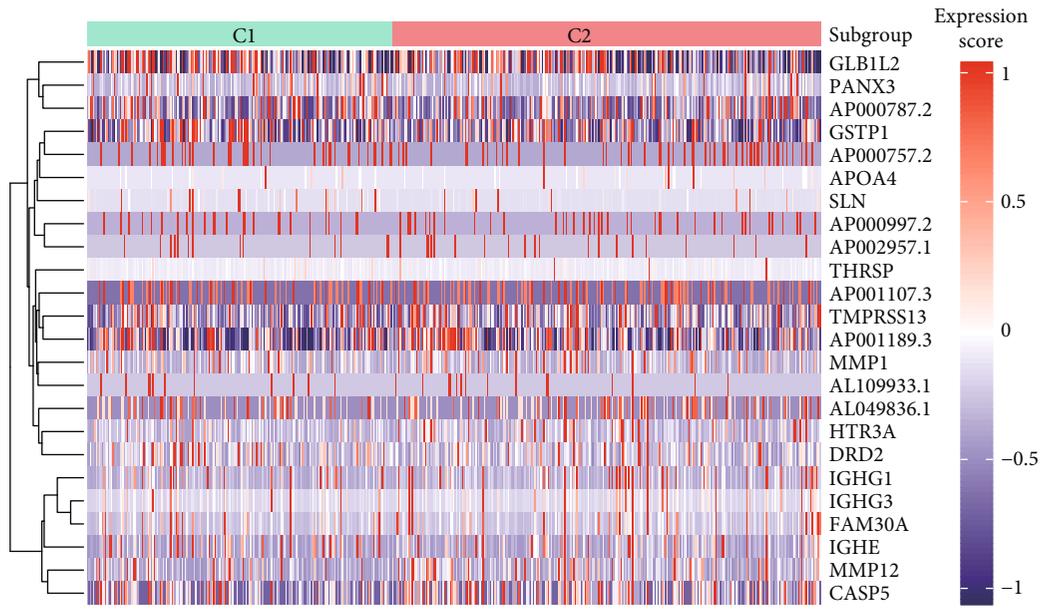
FIGURE 3: Differences in genomic copy number variations (CNVs) related to the two immune phenotypes. (a) CNV statistics between groups C1 and C2 in TCGA-PRAD cohort. (b) The distribution of differentially expressed genes (DEGs) between groups C1 and C2. (c) GO enrichment analysis of DEGs in group C1. (d) GO enrichment analysis of DEGs in group C2. (e and f) CNVs compared among six other tumor types in TCGA cohorts. Cohorts from outside to inside diameters: PRAD, GBM, KIRP, LGG, PAAD, SARC, and TGCT.

impaired antigen presentation for C2 tumor cells and an escape mechanism for C2 immune surveillance. The same results were obtained using the independent ICGC-PRAD-CA validation cohort (Figure 2(a)).

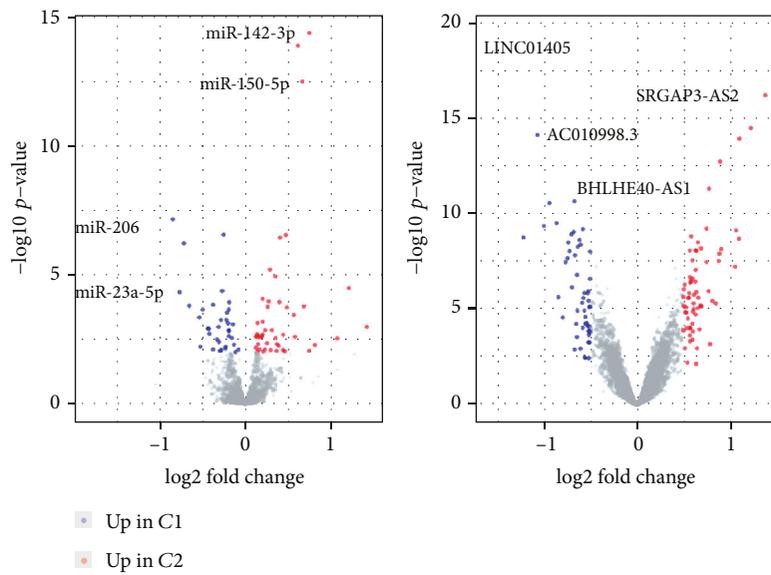
**3.4. Somatic Mutation Differences between Immune Phenotypes.** To understand changes at the gene level, we analyzed the number and quality of somatic mutations in the C1 and C2 subgroups of the TCGA-PRAD cohort. We found that the TP53 mutation frequency in group C2 was higher than that in group C1, and the 10 most-mutated genes are shown in Figure 2(b). Similar to the TCGA-PRAD results,

the mutation frequency of TP53 in C2 was relatively high in most types of TCGA cohort cancers (Figure 2(c)).

**3.5. Differences in Genomic CNVs between Immune Phenotypes.** Studies have found that genomic CNVs are closely associated with immune system evasion [15, 16]. Analyses of changes in the genomes of TCGA-PRAD cohort showed that group C2 had significant amplifications (chromosomes 1q, 4q, 14q, and 17q) and significant deletions (chromosomes 7q, 8p, 14q, and 19q) in several hotspot regions compared to group C1 (Figure 3(a)). To assess whether CNVs affected gene expression, both C1 and C2



(a)



(b)

FIGURE 4: Continued.

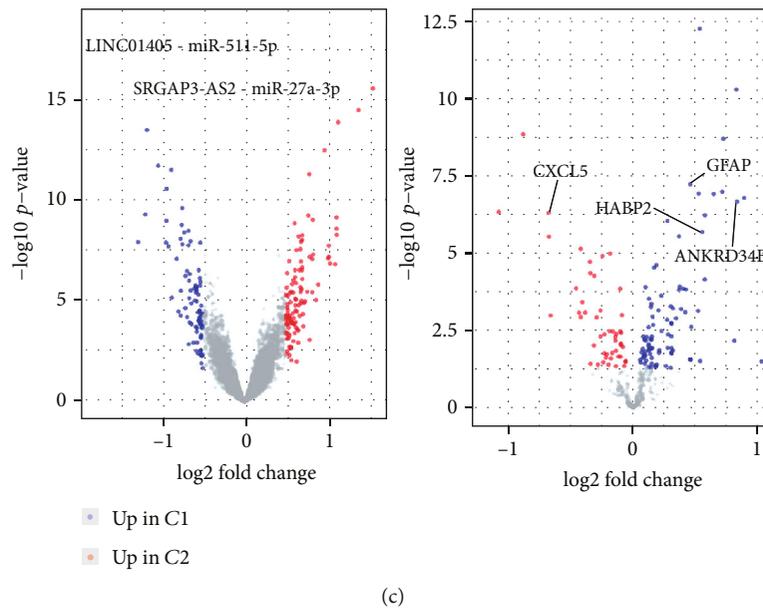


FIGURE 4: The relationship between miRNAs, lncRNAs, and differentially expressed genes (DEGs) between the two subtypes. (a) Analysis of differentially-methylated gene expression between groups C1 and C2. (b) Differentially expressed miRNAs (top left) and differentially expressed lncRNAs (top right) between the two subgroups. (c) Differentially expressed lncRNA-miRNA links (bottom left) and DEGs (bottom right) between the two subgroups.

were screened for DEGs, and the results showed that 134 genes were upregulated in group C1 and 135 genes were upregulated in group C2 (Figure 3(b)). GO enrichment analysis determined that the upregulated C1 genes were involved in immune-cell adhesion and movement, functions related to the activation of immune cells (Figure 3(c)). Upregulated C2 genes were involved in functions related to keratinization, development, and differentiation (Figure 3(d)). Similar to the PRAD dataset results, several TCGA cohorts (glioblastoma multiforme (GBM), kidney renal papillary cell carcinoma (KIRP), brain lower-grade glioma (LGG), liver hepatocellular carcinoma (LIHC), pancreatic adenocarcinoma (PAAD), sarcoma (SARC), testicular germ cell tumor (TGCT), and prostate adenocarcinoma (PRAD)) shared significantly enriched deletions in chromosomes 6p, 10p, and 18p, and increased copy numbers of chromosome 1q in group C2 compared to group C1 (Figures 3(e) and 3(f)). Collectively, the results indicate that group C2 cancer cells adapt to the presence of tumor-infiltrating lymphocytes by acquiring specific somatic mutations, MHC mutations, and genomic CNVs, and group C1 escapes immune surveillance by changing the activation and proliferation of immune cells in the immune microenvironment.

**3.6. Differences in DNA Methylation and miRNA Expression between Immune Phenotypes.** Aberrant DNA methylation is associated with both cellular identity and tumor immune surveillance [17]. To explore how DNA methylation affected tumor immunophenotype, we analyzed genome-wide methylation data in the TCGA-PRAD cohort. In total, we found 43 differentially methylated regions between groups C1 and C2 ( $FDR < 0.05$ ), and a total of 33 gene sequences were represented by these regions. The regions with higher beta values

for group C1 compared to group C2 had relatively low gene expression levels (Figure 4(a)).

Next, 83 differentially expressed miRNAs (significance =  $-1 < \text{fold change (FC)} < 1$  or  $\text{FC} > 1$ , false-discovery rate ( $FDR < 0.01$ ) were assessed between groups C1 and C2 based on TCGA-PRAD cohort (Figure 4(b)). Considering that identifying the target genes for these miRNAs would be useful for understanding miRNA regulatory functions, we identified 5185 target genes and 10437 gene links. Finally, we also identified 166 links for DEGs ( $-1.41 < \text{FC} < 1.41$ ,  $FDR < 0.01$ ), of which 101 were upregulated in group C1 and 65 were upregulated in group C2 (Figure 4(b)).

**3.7. Differences in lncRNA Expression between Immune Phenotypes.** lncRNAs compete with miRNAs by occupying shared binding sequences, thereby sequestering miRNAs and changing the expression of their downstream target genes. Interaction networks between lncRNAs, miRNAs, and mRNAs have been documented for a variety of biological processes in many diseases [18]. Here, we found 118 differentially expressed lncRNAs (significance =  $-0.5 < \log_2 \text{FC} < 0.5$ ,  $FDR < 0.05$ ) between groups C1 and C2 using TCGA-PRAD cohort, with 53 of them highly expressed in group C1, and 65 highly expressed in group C2 (Figure 4(c)). In addition, our results revealed 2752820 lncRNA-miRNA links, including 2588 miRNAs and 14666 lncRNAs from the miRcode database. By combining the differentially expressed lncRNA analysis with the lncRNA-miRNA links information from the miRcode database, we documented 46 differentially expressed lncRNAs that interacted with miRNAs and 746 lncRNA-miRNA links. We next used

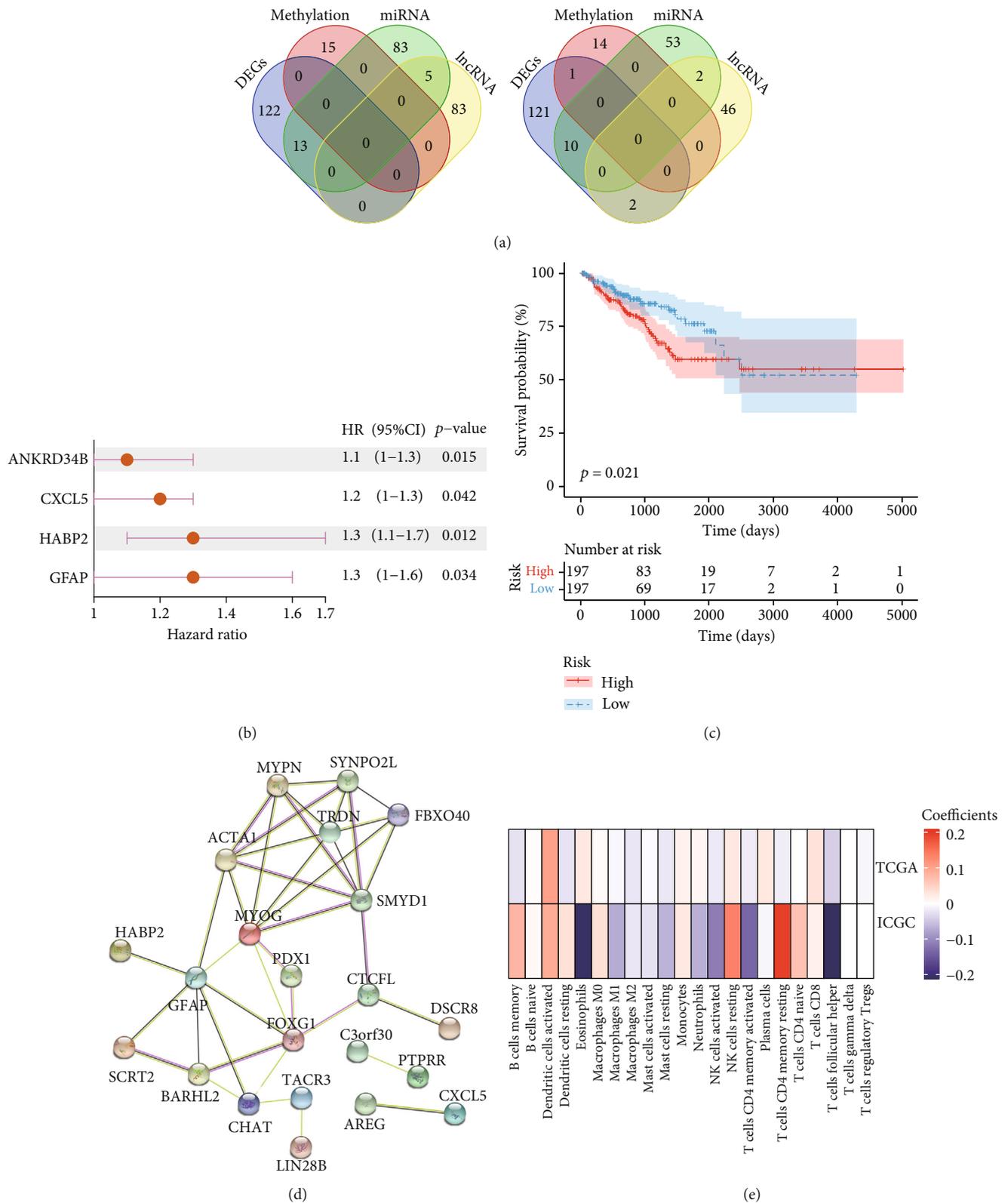


FIGURE 5: Identification of a key node within the genetic and epigenetic regulatory networks. (a) The integration of DEG, lncRNA, miRNA, and methylation analyses to determine upregulated differentially expressed genes in groups C1 and C2. (b) The four genes positively associated with prognosis were confirmed using a random forest blot in TCGA-PRAD cohort. (c) A Kaplan-Meier analysis of the 4-gene model using TGCA-PRAD cohort. (d) GFAP was identified as a key hub gene within the network using a protein-protein interaction analysis of the STRING Consortium database. (e) Spearman's correlation analysis was performed to analyze correlations between GFAP transcription levels and immune-cell subsets for TCGA-PRAD cohort.

the lncRNA2Target database to predict the target genes for these lncRNAs. A total of 138 DEGs were found to interact with these lncRNAs. Among them, 88 targeted DEGs were upregulated in group C1, and 50 targeted DEGs were upregulated in group C2 (Figure 4(c)).

**3.8. Glial Fibrillary Acidic Protein (GFAP) as a Key Node within the Genetic and Epigenetic Regulatory Networks.** Cancer is the result of complicated interactions between genetic and epigenetic variations [19], and we considered that these two tumor immunophenotypes might be driven by both genetic and epigenetic regulators. We identified 13 DEGs that were increased in group C1 and 13 DEGs were increased in group C2 by integrating the analyses of DEGs, lncRNAs, miRNAs, and methylation (Figure 5(a)). We used a single-factor Cox regression for these 26 genes and identified four genes that were significantly associated with PC prognosis. Three of these four genes were highly expressed in group C1, and one of the four was highly expressed in group C2. As expected, the ssGSEA scores based on these genes were confirmed for both the C1 and C2 groups using TCGA-PRAD cohort and the independent validation ICGC-PRAD cohort (Supplementary Figure 1). In addition, the ssGSEA scores were further confirmed for both the C1 and C2 groups for other tumor types in TCGA cohort (Supplementary Figure 2).

These four identified genes were also confirmed using a random forest blot from TCGA-PRAD cohort data (Figure 5(b)). After calculating the risk scores for the four gene sets using a multivariate Cox regression analysis, they were used for a prognosis analysis of data from the TCGA-PRAD cohort. These four genes were found to have excellent predictive ability for PC prognoses ( $p < 0.05$ , Figure 5(c)).

Importantly, we identified GFAP as a key hub gene within the network using PPIs from the STRING Consortium database (Figure 5(d)), and we used Spearman's correlation analysis to assess any correlations between GFAP transcription levels and immune-cell subsets in the TCGA-PRAD cohort. We found that GFAP expression was positively correlated with T cell CD4 memory resting and NK cells resting, while GFAP expression was negatively correlated with T cells follicular helper, eosinophils, NK cells activated, plasma cells, and macrophage M1 (Figure 5(e)). The results confirmed that high GFAP expression was positively correlated with immunosuppression.

#### 4. Discussion

Tumor-associated immune cells are an important part of the tumor microenvironment. A growing number of evidences show that immune cells in tumor microenvironment play a vital role in the initiation and progression in prostate cancer [20]. The infiltrating immune cells in the prostate tumor microenvironment mainly include macrophages, neutrophils, myeloid-derived suppressor cells (MDSCs), and T regulatory cells (Tregs) [21]. In the present study, the results of CIBERSORT analyses showed that 9 immune-cell subtypes (activated mast cells, monocytes, M1 macrophages, activated dendritic cells, naive B cells, activated memory T cells-CD4, resting dendritic cells, resting NK cells, and resting memory

T cells-CD4) were positively correlated with PD-L1 transcript levels. To better describe the relationship between tumor-infiltrating immune cells and tumor-cell immune escape, we classified the PC sample data into two major clusters (C1 and C2) based on the above nine immune-cell subsets.

Different immune subtypes have been identified among patients with cancer, a promising approach for understanding the effects of the immune microenvironment on tumors. For hepatocellular carcinoma, Sia et al. also identified two subgroups characterized by exhausted or adaptive immune responses to predict the therapeutic effects of PD-1, PD-L1, or transforming growth factor  $\beta$ 1 inhibitor [22]. Similarly, Li et al. reported both immune-reduced and immune-enhanced subtypes with differing immune-related and clinical characteristics to provide targets for new treatments [23]. For breast cancer, Zheng et al. developed an immune phenotype classifier for predicting both immune activity within the tumor microenvironment and prognoses for patients with triple-negative breast cancer [24]. In addition, numerous investigations have been conducted using immune subtypes to predict clinical prognoses and treatment guidance for bladder cancer [25], ovarian carcinoma [26], uveal melanoma [27], lung cancer [28, 29], and head and neck cancers [30]. Here, we have identified two immune system-related PC phenotypes with differing transcription level scores among immune-cell subsets using a CIBERSORT analysis and further characterized their differences using a multiomic approach. The results revealed that identifying immune-related cancer subtypes was meaningful for both a better understanding of tumor molecular mechanisms and for clinical prognoses.

Emerging evidence suggests that both genetic and epigenetic alterations in cancer cells drive malignancy. Here, we have highlighted the genetic mutations and epigenetic aberrations driving different PC immunophenotypes. Similarly, Feng et al. described the mutational and epigenetic landscape for head and neck cancers using immune-related phenotypes [30]. For glioblastomas, tumor drug tolerance would evolve by acquiring genetic and epigenetic alterations [31]. For ovarian cancers, both genetic and epigenetic factors likely contribute to shaping the immunosuppressive tumor microenvironment and to improving the response rate of ovarian cancer to immune checkpoint therapies [32]. Similarly, for acute myeloid leukemia, genetic abnormalities and aberrant epigenetic regulators both play essential roles that affect responses to therapy and prognoses [33].

By integrating the genetic and epigenetic alterations and their effects on DEGs between these two immune subgroups, we found that GFAP was the key hub gene in this regulation network. Previous research has demonstrated that the expression of GFAP was aberrant in astrocytoma tissue compared to normal brain tissue [34–38], with GFAP levels decreasing as with astrocytoma grading increased [39–41], so GFAP levels may serve as a novel cancer diagnostic marker.

#### 5. Conclusions

The present data provide further insight into the underlying molecular mechanisms for alterations to both the mutational landscape and the epigenome for these two

immunophenotypes. GFAP, a key node among mechanistic pathways, may be a potential biomarker both for PC diagnosis and for predicting prognoses.

## Data Availability

Publicly available datasets were analyzed in this study.

## Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

## Authors' Contributions

Jinjian Yang, Wencheng Yao, Xiang Li, and Zhankui Jia designed the research plan and analysis strategy. Wencheng Yao, Chaohui Gu, Zhibo Jin, Jun Wang, and Bo Yuan wrote the manuscript and prepared the figures. All authors agreed to be accountable for all aspects of the work and gave final approval of the version to be published.

## Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant number: 81570685).

## Supplementary Materials

Supplementary Figure 1: comparisons of ssGSEA scores from TCGA-PRAD cohort. (a) The ssGSEA scores from all screened genes between the two subgroups in the TCGA-PRAD cohort. (b) The ssGSEA scores from the two subgroups based on C1-upregulated genes in the TCGA-PRAD cohort. (c) The ssGSEA scores from the two subgroups based on C2-upregulated genes in the TCGA-PRAD cohort. Supplementary Figure 2: the ssGSEA scores based on C1- and C2-upregulated genes. (a) The ssGSEA scores for C1-upregulated genes between C1 and C2 subgroups were applied to total tumor entities in the TCGA cohort. (b) The ssGSEA scores based on C2-upregulated genes between the C1 and C2 subgroups applied to total tumor entities in the TCGA cohort. (*Supplementary Materials*)

## References

- [1] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2018," *CA: a Cancer Journal for Clinicians*, vol. 68, no. 1, pp. 7–30, 2018.
- [2] M. M. Center, A. Jemal, J. Lortet-Tieulent et al., "International variation in prostate cancer incidence and mortality rates," *European Urology*, vol. 61, no. 6, pp. 1079–1092, 2012.
- [3] X. Lei, Y. Lei, J. K. Li et al., "Immune cells within the tumor microenvironment: biological functions and roles in cancer immunotherapy," *Cancer Letters*, vol. 470, pp. 126–133, 2020.
- [4] M. J. Smyth, G. P. Dunn, and R. D. Schreiber, "Cancer immunosurveillance and immunoediting: the roles of immunity in suppressing tumor development and shaping tumor immunogenicity," *Advances in Immunology*, vol. 90, pp. 1–50, 2006.
- [5] F. Schütz, S. Stefanovic, L. Mayer, A. von Au, C. Domschke, and C. Sohn, "PD-1/PD-L1 pathway in breast cancer," *Oncol Res Treat.*, vol. 40, no. 5, pp. 294–297, 2017.
- [6] A. V. Balar and J. S. Weber, "PD-1 and PD-L1 antibodies in cancer: current status and future directions," *Cancer Immunology, Immunotherapy*, vol. 66, no. 5, pp. 551–564, 2017.
- [7] A. Chu, G. Robertson, D. Brooks et al., "Large-scale profiling of microRNAs for the Cancer Genome Atlas," *Nucleic Acids Research*, vol. 44, no. 1, article e3, 2016.
- [8] A. M. Newman, C. L. Liu, M. R. Green et al., "Robust enumeration of cell subsets from tissue expression profiles," *Nature Methods*, vol. 12, no. 5, pp. 453–457, 2015.
- [9] T. Galili, "Dendextend: an R package for visualizing, adjusting and comparing trees of hierarchical clustering," *Bioinformatics*, vol. 31, no. 22, pp. 3718–3720, 2015.
- [10] T. Metsalu and J. Vilo, "ClustVis: a web tool for visualizing clustering of multivariate data using principal component analysis and heatmap," *Nucleic Acids Research*, vol. 43, no. W1, pp. W566–W570, 2015.
- [11] E. Cerami, J. Gao, U. Dogrusoz et al., "The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data," *Cancer Discovery*, vol. 2, no. 5, pp. 401–404, 2012.
- [12] Z. Gu, R. Eils, and M. Schlesner, "Complex heatmaps reveal patterns and correlations in multidimensional genomic data," *Bioinformatics*, vol. 32, no. 18, pp. 2847–2849, 2016.
- [13] S. V. Laddha, S. Ganesan, C. S. Chan, and E. White, "Mutational landscape of the essential autophagy gene BECN1 in human cancers," *Molecular Cancer Research*, vol. 12, no. 4, pp. 485–490, 2014.
- [14] Y. Yang, D. Chen, H. Liu, and K. Yang, "Increased expression of lncRNA CASC9 promotes tumor progression by suppressing autophagy-mediated cell apoptosis via the AKT/mTOR pathway in oral squamous cell carcinoma," *Cell Death & Disease*, vol. 10, no. 2, p. 41, 2019.
- [15] A. Bhattacharya, R. D. Bense, C. G. Urzúa-Traslaviña, E. G. E. de Vries, M. van Vugt, and R. S. N. Fehrmann, "Transcriptional effects of copy number alterations in a large set of human cancers," *Nature Communications*, vol. 11, no. 1, p. 715, 2020.
- [16] C. S. Grasso, M. Giannakis, D. K. Wells et al., "Genetic mechanisms of immune evasion in colorectal cancer," *Cancer Discovery*, vol. 8, no. 6, pp. 730–749, 2018.
- [17] J. Li, Y. Li, W. Li et al., "Guide positioning sequencing identifies aberrant DNA methylation patterns that alter cell identity and tumor-immune surveillance networks," *Genome Research*, vol. 29, no. 2, pp. 270–280, 2019.
- [18] L. Wang, K. B. Cho, Y. Li, G. Tao, Z. Xie, and B. Guo, "Long noncoding RNA (lncRNA)-mediated competing endogenous RNA networks provide novel potential biomarkers and therapeutic targets for colorectal cancer," *International Journal of Molecular Sciences*, vol. 20, no. 22, p. 5758, 2019.
- [19] S. Zhou, A. E. Treloar, and M. Lupien, "Emergence of the non-coding cancer genome: a target of genetic and epigenetic alterations," *Cancer Discovery*, vol. 6, no. 11, pp. 1215–1229, 2016.
- [20] P. Hu, Y. Gao, Y. Huang et al., "Gene expression-based immune cell infiltration analyses of prostate cancer and their associations with survival outcome," *DNA and Cell Biology*, vol. 39, no. 7, pp. 1194–1204, 2020.
- [21] S. Q. Wu, H. Su, Y. H. Wang, and X. K. Zhao, "Role of tumor-associated immune cells in prostate cancer: angel or devil?," *Asian Journal of Andrology*, vol. 21, no. 5, pp. 433–437, 2019.

- [22] D. Sia, Y. Jiao, I. Martinez-Quetglas et al., "Identification of an immune-specific class of hepatocellular carcinoma, based on molecular features," *Gastroenterology*, vol. 153, no. 3, pp. 812–826, 2017.
- [23] W. Li, H. Wang, Z. Ma et al., "Multi-omics analysis of micro-environment characteristics and immune escape mechanisms of hepatocellular carcinoma," *Frontiers in Oncology*, vol. 9, p. 1019, 2019.
- [24] S. Zheng, Y. Zou, X. Xie et al., "Development and validation of a stromal immune phenotype classifier for predicting immune activity and prognosis in triple-negative breast cancer," *International Journal of Cancer*, vol. 147, no. 2, pp. 542–553, 2020.
- [25] B. N. Song, S. K. Kim, J. Y. Mun, Y. D. Choi, S. H. Leem, and I. S. Chu, "Identification of an immunotherapy-responsive molecular subtype of bladder cancer," *eBioMedicine*, vol. 50, pp. 238–245, 2019.
- [26] M. Zheng, Y. Hu, R. Gou et al., "Identification of immune-enhanced molecular subtype associated with BRCA1 mutations, immune checkpoints and clinical outcome in ovarian carcinoma," *Journal of Cellular and Molecular Medicine*, vol. 24, no. 5, pp. 2819–2831, 2020.
- [27] H. Pan, L. Lu, J. Cui, Y. Yang, Z. Wang, and X. Fan, "Immunological analyses reveal an immune subtype of uveal melanoma with a poor prognosis," *Aging (Albany NY)*, vol. 12, no. 2, pp. 1446–1464, 2020.
- [28] F. L. Qin, Z. Y. Xu, L. Q. Yuan et al., "Novel immune subtypes of lung adenocarcinoma identified through bioinformatic analysis," *FEBS Open Bio*, vol. 10, no. 9, pp. 1921–1933, 2020.
- [29] Y. Kurebayashi, K. Emoto, Y. Hayashi et al., "Comprehensive immune profiling of lung adenocarcinomas reveals four immunosubtypes with plasma cell subtype a negative indicator," *Cancer Immunology Research*, vol. 4, no. 3, pp. 234–247, 2016.
- [30] B. Feng, Y. Shen, X. Pastor Hostench et al., "Integrative analysis of multi-omics data identified EGFR and PTGS2 as key nodes in a gene regulatory network related to immune phenotypes in head and neck cancer," *Clinical Cancer Research*, vol. 26, no. 14, pp. 3616–3628, 2020.
- [31] C. E. Eyler, H. Matsunaga, V. Hovestadt, S. J. Vantine, P. van Galen, and B. E. Bernstein, "Single-cell lineage analysis reveals genetic and epigenetic interplay in glioblastoma drug resistance," *Genome Biology*, vol. 21, no. 1, p. 174, 2020.
- [32] Y. Cho, L. Milane, and M. M. Amiji, "Genetic and epigenetic strategies for advancing ovarian cancer immunotherapy," *Expert Opinion on Biological Therapy*, vol. 19, no. 6, pp. 547–560, 2019.
- [33] C. Bret, E. Viziteu, A. Kassambara, and J. Moreaux, "Identifying high-risk adult AML patients: epigenetic and genetic risk factors and their implications for therapy," *Expert Review of Hematology*, vol. 9, no. 4, pp. 351–360, 2016.
- [34] V. C. Chumbalkar, C. Subhashini, V. M. Dhople et al., "Differential protein expression in human gliomas and molecular insights," *Proteomics*, vol. 5, no. 4, pp. 1167–1177, 2005.
- [35] R. Laczko, K. M. Szauter, M. K. Jansen et al., "Active lysyl oxidase (LOX) correlates with focal adhesion kinase (FAK)/paxillin activation and migration in invasive astrocytes," *Neuropathology and Applied Neurobiology*, vol. 33, no. 6, pp. 631–643, 2007.
- [36] M. A. Sharpe and D. S. Baskin, "Monoamine oxidase B levels are highly expressed in human gliomas and are correlated with the expression of HiF-1 $\alpha$  and with transcription factors Sp1 and Sp3," *Oncotarget*, vol. 7, no. 3, pp. 3379–3393, 2016.
- [37] S. Bien-Möller, E. Balz, S. Herzog et al., "Association of glioblastoma multiforme stem cell characteristics, differentiation, and microglia marker genes with patient survival," *Stem Cells International*, vol. 2018, Article ID 9628289, 2018.
- [38] X. Su, X. Liu, L. Ni et al., "GFAP expression is regulated by Pax 3 in brain glioma stem cells," *Oncology Reports*, vol. 36, no. 3, pp. 1277–1284, 2016.
- [39] E. J. van Bodegraven, J. V. van Asperen, P. A. J. Robe, and E. M. Hol, "Importance of GFAP isoform-specific analyses in astrocytoma," *Glia*, vol. 67, no. 8, pp. 1417–1433, 2019.
- [40] J. Tichy, S. Spechtmeyer, M. Mittelbronn et al., "Prospective evaluation of serum glial fibrillary acidic protein (GFAP) as a diagnostic marker for glioblastoma," *Journal of Neuro-Oncology*, vol. 126, no. 2, pp. 361–369, 2016.
- [41] J. Gállego Pérez-Larraya, S. Paris, A. Idbaih et al., "Diagnostic and prognostic value of preoperative combined GFAP, IGFBP-2, and YKL-40 plasma levels in patients with glioblastoma," *Cancer*, vol. 120, no. 24, pp. 3972–3980, 2014.