Research Article

Exploring the Interplay between Metabolism and Tumor Microenvironment Based on Four Major Metabolism Pathways in Colon Adenocarcinoma

Xiaofang Qiao,[1] **Guangmei Zhang,**[2] **Yajie Xiao,**[3] **Xiaoli Cui,**[3] **Zhikun Zhao,**[3] **Dongfang Wu** [iD][3] **and Xuefei Liu** [iD][4]

[1]*Department of Gastrointestinal Surgery, The Second Hospital of Jilin University, Changchun 220100, China*
[2]*Department of Internal Medicine, Jilin Cancer Hospital, Jilin 220200, China*
[3]*Department of Medicine, YuceBio Technology Co., Ltd., Shenzhen 440300, China*
[4]*Department of Oncology, Northern Theater Command General Hospital, Shenyang 210100, China*

Correspondence should be addressed to Dongfang Wu; wudf@yucebio.com and Xuefei Liu; feiliu6@163.com

Academic Editor: Jinfang Zheng

Tumor metabolism plays a critical role in tumor progression. However, the interaction between metabolism and tumor microenvironment (TME) has not been comprehensively revealed in colon adenocarcinoma (COAD). We used unsupervised consensus clustering to establish three molecular subtypes (clusters) based on the enrichment score of four major metabolism pathways in TCGA-COAD dataset. GSE17536 was used as a validation dataset. Single-cell RNA sequencing data (GSE161277) was employed to further verify the reliability of subtyping and characterize the correlation between metabolism and TME. Three clusters were identified and they performed distinct prognosis and molecular features. Clust3 had the worst overall survival and the highest enrichment score of glycolysis. 86 differentially expressed genes (DEGs) were identified, in which 11 DEGs were associated with favorable prognosis and 75 DEGs were associated with poor prognosis. Striking correlations were observed between hypoxia and glycolysis, clust3 and hypoxia, and clust3 and angiogenesis ($P < 0.001$).We constructed a molecular subtyping system which was effective and reliable for predicting COAD prognosis. The 86 identified key DEGs may be greatly involved in COAD progression, and they provide new perspectives and directions for further understanding the mechanism of metabolism in promoting aggressive phenotype by interacting with TME.

1. Introduction

Colon adenocarcinoma (COAD) is one of the most diagnosed cancer types worldwide, which represented 6.0% of new cancer cases and was the cause of 5.8% of new cancer deaths in 2020 according to global cancer statistics [1]. Risk factors of COAD are various, which can be generally classified into unhealthy lifestyles (such as smoking, heavy alcohol, and unhealthy diet), disease-caused or drug-induced (such as long-term androgen deprivation therapy, diabetes mellitus and insulin resistance, acromegaly, and renal transplantation under long-term immunosuppression), and family history [2]. Most of COAD individuals are sporadic,

about 70%, and ages above 50 years are common [3]. Although screening programs such as colonoscopy, fecal immunochemical test (FIT), and computed tomography (CT) colonography are recommended and popularized for high-risk populations in the recent decades, efficient and personalized therapeutics are still current challenges especially for metastatic patients.

Immunotherapy or immune checkpoint blockade raises the hope for metastatic patients. Targeted therapies such as vascular endothelial growth factor (VEGF) inhibitors and epidermal growth factor receptor (EGFR) inhibitors are examined to effectively hinder angiogenesis and thus impede tumor growth [4]. Immunotherapy and targeted therapy

regulate tumor microenvironment (TME) and inhibit tumor progression through targeting key proteins associated with poor prognosis. Tumor metabolism is a pivotal factor in TME modulation, where four critical metabolism pathways, glycolysis, pentose phosphate pathway (PPP), fatty acid oxidation (FAO), and glutaminolysis, are reported to be strikingly involved in tumor metabolic reprogramming contributing to TME modulation and tumor progression [5–9]. Particularly, a number of glycolysis-related prognostic biomarkers have been identified in various cancer types [10–13]. Nevertheless, the mechanism of the interaction between TME and metabolism pathways has not been well understood.

An increasing number of bioinformatics tools accelerate the development of understanding potential mechanisms of cancer development and the novel prognostic biomarkers for predicting cancer prognosis [14–17]. Moreover, the development of single-cell RNA sequencing (scRNA-seq) technology allows more accurately parsing the interplay or crosstalk of various pathways in TME. In this study, we sought to reveal the interaction between metabolism pathways and TME based on expression data and single-cell data gleaned from public database. Three molecular subtypes (clusters) were established based on the expression of genes within metabolism pathways. Three clusters showed distinct overall survival and expression patterns of metabolism. We identified key differentially expressed genes (DEGs) between two groups with distinct prognosis. Dramatic correlation between TME hallmarks (angiogenesis and hypoxia) and metabolism pathways in malignant cells were illustrated by single-cell analysis. Our study lays a foundation for further research on the mechanism of tumor metabolism in colon cancer.

2. Materials and Methods

2.1. Data Acquisition. TCGA-COAD dataset including expression data, mutation data, and clinical information was obtained from The Cancer Genome Atlas (TCGA) database (https://portal.gdc.cancer.gov/). Expression profiles with fragments per kilobase million (FPKM) format were transformed into transcript per million (TPM) format. GSE17536 [18] and GSE161277 [19] datasets were downloaded from Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/). Four metabolism pathways (glycolysis, PPP, FAO, and glutaminolysis) were obtained from Molecular Signatures Database [20] (MSigDB, https://www.gsea-msigdb.org/gsea/msigdb/).

2.2. Data Preprocessing. For TCGA-COAD dataset, samples without survival information were excluded, and survival time within 30 days to 10 years was restricted for all samples. 410 samples with expression data remained in TCGA-COAD dataset. For GSE17536 dataset, probes in chip files were converted to gene symbols. Samples were screened according to the same conditions as TCGA-COAD dataset. 170 samples with expression data remained in GSE17536 dataset.

2.3. Unsupervised Consensus Clustering. Firstly, single sample gene set enrichment analysis (ssGSEA) in GSVA R package was applied to calculate ssGSEA score of four metabolism pathways for each sample in TCGA-COAD dataset [21]. Then unsupervised consensus clustering was used to cluster samples based on the ssGSEA score of four metabolism pathways through ConsensusClusterPlus R package (v1.54.0) [22]. Kmdist algorithm and "Pearson" distance were used to conduct 500 bootstraps with each bootstrap containing at least 80% samples of TCGA-COAD dataset. Cluster number k was set from 2 to 10, and optimal k was determined according to cumulative distribution function (CDF) and area under CDF curve. The clustering was validated in GSE17536 dataset.

2.4. Mutation Analysis. Copy number variations (CNVs) and single nucleotide variations (SNVs) in TCGA-COAD dataset were analyzed separately. For CNV data, Genomic Identification of Significant Targets in Cancer (GISTIC) 2.0 software was employed to detect arm-level and focal CNVs based on hg38 reference genome [23]. The number of copies greater than 1 was considered the threshold of copy amplification and that less than −1 was considered the threshold of copy loss. R package of maftools was used to assess SNVs [24].

2.5. Tumor Microenvironment Analysis. To evaluate TME features in different clusters, several methodologies were applied. MCPcounter R package and the Proportion of Immune and Cancer cells (EPIC) methodology were used to calculate the enrichment score of immune cells [25, 26]. Estimation of STromal and Immune cells in MAlignant Tumors using Expression data (ESTIMATE) tool was introduced to evaluate immune infiltration and stromal infiltration by calculating immune score, stromal score, and ESTIMATE score based on signatures [27]. Gene sets of macrophage-related pathways including toll-like receptor signaling, natural killer cell mediated cytotoxicity, and antigen processing and presentation were obtained from MSigDB. Signatures of IFN-γ and cytotoxicity were obtained from previous research [28,29].

2.6. Differential Analysis between Clust12 and Clust3. DEGs between clust12 and clust3 were detected by limma R package under conditions of |log2 (fold change)| > log2 (1.5) and false discovery rate (FDR) < 0.05 [30]. String (https://cn.string-db.org/) online tool was applied to identify interactions among DEGs, and the interaction network was visualized by CytoScape (3.8.0) [31]. Degree of each DEG was analyzed by Analyze Network tool in CytoScape. Higher degree represents greater importance of the gene in the network. Annotation of gene ontology (GO) terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways was outputted by WebGestaltR package [32]. FDR < 0.05 was determined to screen significant GO terms and $P < 0.05$ was determined to screen significantly enriched KEGG pathways.

2.7. Identifying Immune Cells in scRNA-Seq Data. GSE161277 dataset containing scRNA-seq data was included for validating the reliability of molecular subtyping based on four metabolism pathways. Within the dataset, 13 samples, 4 patients, and 5 tissues (adenoma, blood, carcinoma, normal, and paracancer tissues) were included. Seurat R package was implemented to process scRNA-seq data [33]. Firstly, single cells were screened under conditions of each gene expressing at least in three cells and each cell expressing at least 250 genes. Then PercentageFeatureSet function was used to calculate the percentages of mitochondria and rRNA. Single cells with mitochondria <35% and unique molecular identifiers (UMI) > 100 within one cell were further screened. Log-normalization was conducted for normalizing single-cell data. FindVariableFeatures function was performed to identify highly variable genes based on "vst." ScaleData function was conducted to scale the data and principle component analysis (PCA) was used to reduce the dimensionality. Then FindNeighbors and FindClusters functions were implemented to cluster cells when dim = 40 and resolution = 0.9. CD45 (PTPRC) was set as the marker to identify immune cells. Finally, 26188 immune cells remained.

2.8. Annotating Cell Types in Processed scRNA-seq Data. The expression data of 26188 immune cells was rescaled and highly variable genes were recalculated. Cells were clustered under dim = 35 and resolution = 0.3 by FindNeighbors and FindClusters functions in Seurat R package [33]. Then RunTSNE function was performed to reduce dimensionality for 26188 ells and identify cell subgroups with different samples, tissues, and patients. Classical immune markers were used to annotate cell subgroups. Differential genes were screened by using FindAllMarkers function with parameters of logfc = 0.5 and Minpct = 0.35 ($P < 0.05$). Only the top 5 differential genes were visualized. clusterProfiler R package was applied to annotate enriched KEGG pathways for differential genes [34].

2.9. Statistical Analysis. Statistical analysis was performed in R software (4.1.1). R packages and tools used in the study were indicated. Statistical methods were described in the corresponding sections. $P < 0.05$ was considered significant. "ns" indicates no significance. $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, and $^{****}P < 0.0001$.

3. Results

3.1. Constructing Molecular Subtypes Based on Metabolism-Related Pathways. The workflow of this study is shown in Supplementary Figure S1. For each sample in TCGA-COAD dataset, we calculated the enrichment score of four metabolism pathways, glycolysis, PPP, FAO, and glutaminolysis, by using ssGSEA (Supplementary Table S1). Then unsupervised consensus clustering was applied to cluster 410 samples based on the enrichment score of four metabolism pathways. Cluster number k was set from 2 to 10 in CDF analysis. Cluster number $k = 3$ was confirmed as optimal

according to CDF and area under CDF curve (Figures 1(a) and 1(b)). Consensus matrix showed that samples were clearly divided into three clusters or molecular subtypes (Figure 1(c)). Three clusters showed different patterns of the enrichment of four metabolism pathways (Figure 1(d)). Glycolysis was the most enriched in clust3, while clust1 had the highest enrichment of glutaminolysis. The enrichment scores of PPP, FAO, and glutaminolysis were significantly higher in clust1 than in clust3. Over half of genes within four pathways (44 of 74) were differentially expressed among three clusters (Supplementary Figure S2).

Kaplan-Meier survival analysis revealed that three clusters had differential overall survival ($P = 0.0077$, Figure 1(e)). Using the same methodology in GSE17536 dataset, 170 samples were divided into three clusters with distinct overall survival ($P = 0.0049$, Figure 1(f)). Clust1 had the worst prognosis, while clust3 had the longest overall survival in both datasets (Figures 1(e) and 1(f)). PCA presented the different distribution of three clusters in two datasets (Supplementary Figure S3). Of the distribution of three clusters in different clinical features, we observed no significant difference in ages, genders, and stages (Supplementary Figure S4).

3.2. Differential Genomic Features among Three Subtypes. To know the genomic features of three subtypes, we applied gistic2 software to analyze the CNV data and visualized CNVs of 22 chromosomes (Figures 2(a)–2(c)). Three clusters showed obvious gain and loss of CNVs where gain of CNVs occurred largely in chromosomes 7, 8, 13, and 20, and loss of CNVs occurred largely in chromosome 18. Although similar CNV patterns were shown in three clusters, still a number of significantly differential CNVs were identified among them (Supplementary Figure S5). Clust2 had the most number of significantly amplified CNVs especially in chromosomes 17 and 20, compared to clust1 and clust3. Relatively lower counts of loss of CNVs were found in clust1 and clust2, but clust3 had more loss of CNVs than gain of CNVs itself (Supplementary Figure S5). In addition, we screened the top 15 mutated genes within four metabolism pathways based on SNV data (Figure 2(d)). The mutation frequencies ranged from 2% to 5%, and missense mutations contributed the majority of mutations. *GLUD2* was the top mutated gene, where missense mutations contributed the most of SNVs.

3.3. Differential TME among Three Clusters. Next we used MCPcounter to evaluate the estimated proportion of 22 immune-related cells in three clusters in TCGA-COAD dataset. 9 of 22 immune-related cells showed a significant difference on the proportion among three clusters (Figure 3(a)). In particular, resting memory CD4 T cells and M0 macrophages consisted of a relatively high proportion among 22 immune-related cells. Clust1 had the highest proportion of resting memory CD4 T cells, while clust3 had the highest proportion of M0 macrophages ($P < 0.01$, Figure 3(a)). EPIC analysis on seven cell types showed that clust3 had significantly higher proportions of cancer-

(a)

(b)

(c)
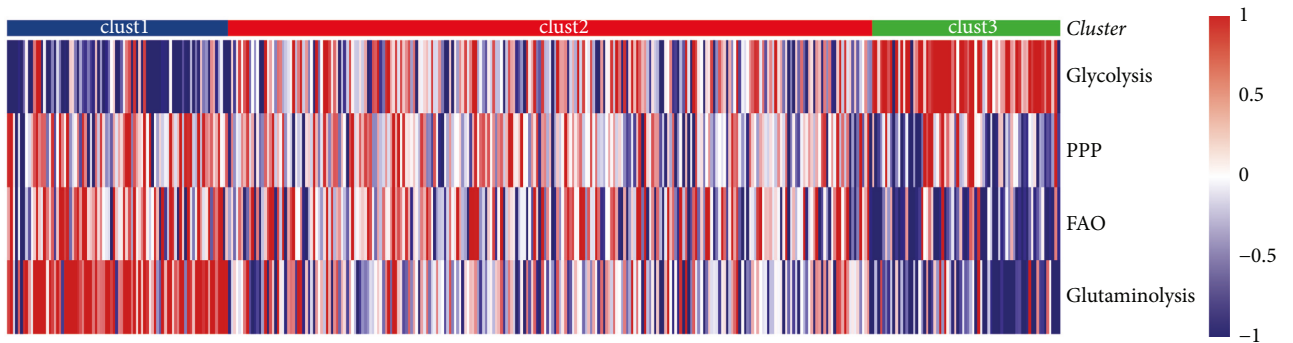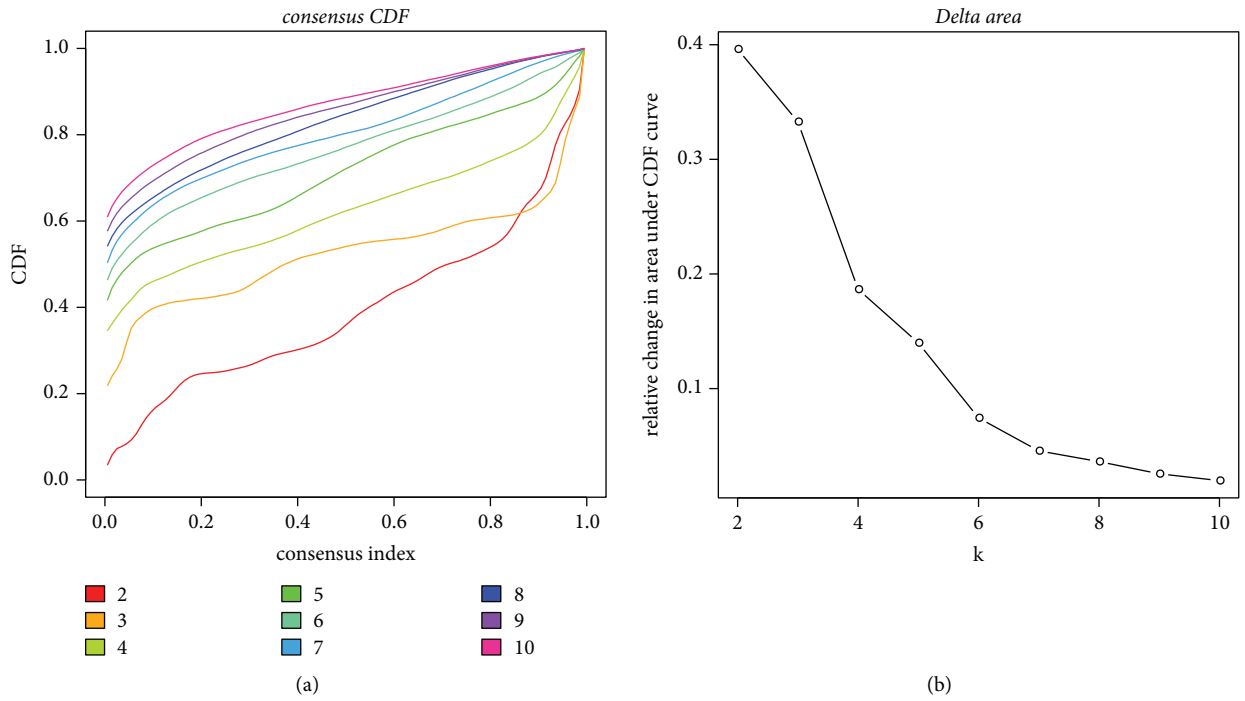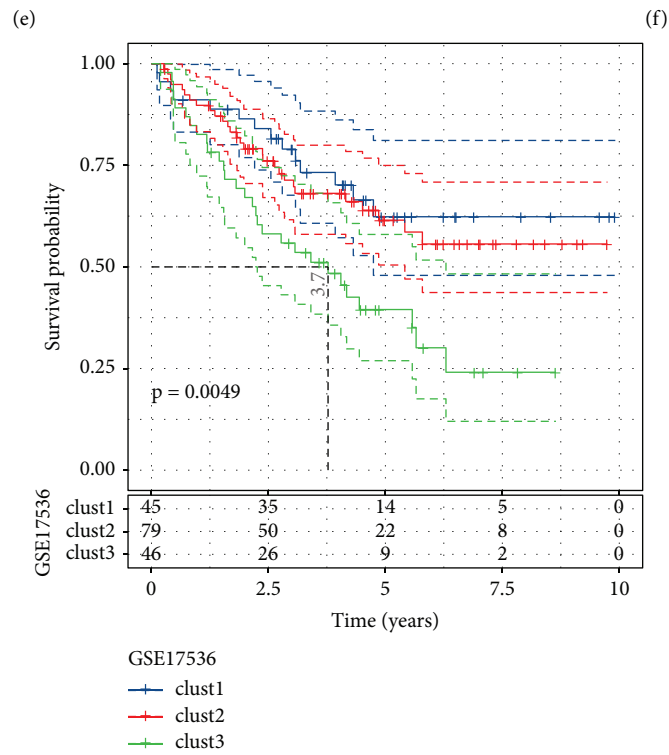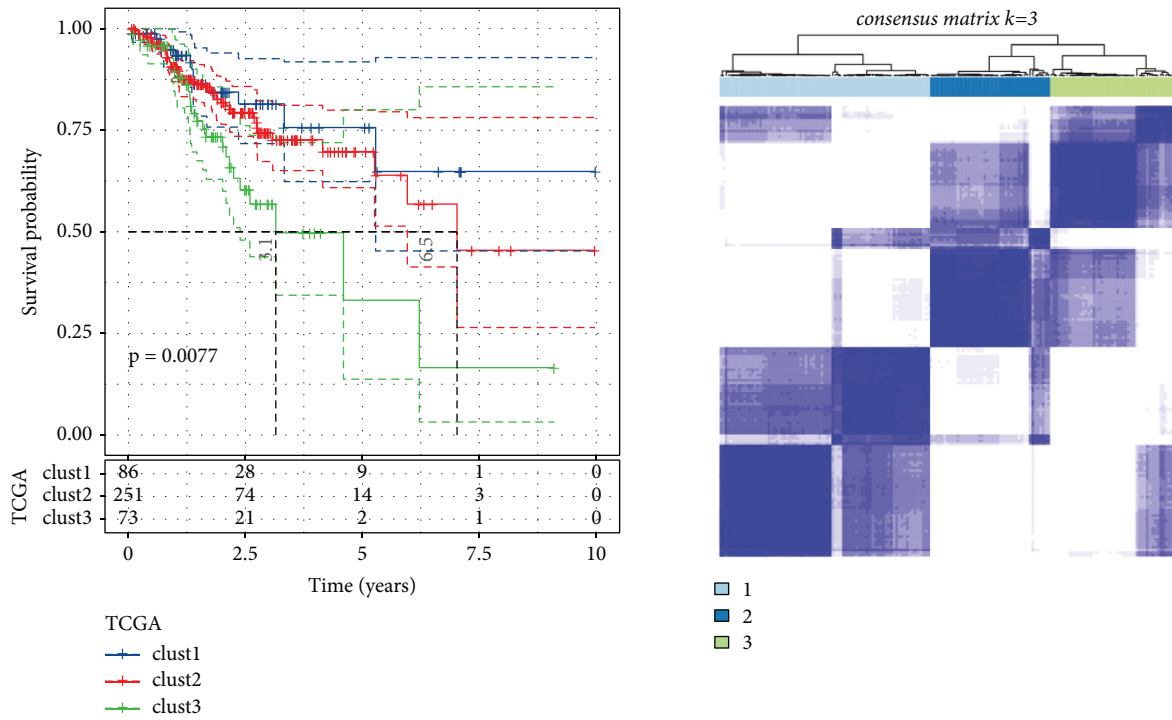
(d)

Figure 1: Continued.

(e)

(f)

(g)

FIGURE 1: Constructing molecular subtypes based on metabolism pathways. ((a) and (b)) CDF curve and area under CDF curve when cluster number $k = 2$ to 10 analyzed by unsupervised consensus clustering. (c) Consensus matrix when $k = 3$ in TCGA-COAD dataset. (d) A heatmap of the enrichment of four metabolism pathways grouped by clusters. Blue and red indicate relatively low and high enrichment scores, respectively. (e) Kaplan-Meier survival curves of three clusters in TCGA-COAD dataset. (f) Consensus matrix when $k = 3$ in GSE17536 dataset. (g) Kaplan-Meier survival curves of three clusters in GSE17536 dataset. Log-rank test was conducted in (e) and (g).

associated fibroblasts and endothelial cells (Supplementary Figure S6). Interestingly, clust1 with the longest overall survival had the least immune infiltration, and clust3 had the highest score of stromal and immune score ($P < 0.0001$, Figure 3(b)). We then assessed the enrichment of 10 oncogenic pathways [35] and observed that 6 of 10 oncogenic pathways were differentially enriched among three clusters ($P < 0.01$, Figure 3(c)). Hippo, Notch, RAS, and Wnt signaling pathways were more activated in clust3, while NRF1 and TP53 were more activated in clust1.

As M0 macrophages were identified to be significantly differentially enriched among three clusters, we further assessed the immune regulation related to macrophages. We selected three pathways related to macrophages from MSigDB including antigen processing and presentation, toll-like receptor signaling pathway, and natural killer (NK) cell mediated cytotoxicity. GSEA revealed that the three pathways were all most activated in clust3 ($P < 0.01$, Figures 3(d)–3(f)), which was accordant to the highest proportion of macrophages in clust3. In addition, clust3 also displayed high enrichment of interferon-γ (IFN-γ) and T cell cytotoxicity (CYT) ($P < 0.05$, Figures 3(g) and 3(h)). In GSE17536 dataset, similar results were presented (Supplementary Figure S7). Overall, clust3 showed more significantly active immune response than clust1 and clust2. Differential TME among three clusters indicated that metabolism pathways were possibly involved in complicated TME modulation.

3.4. Identifying Differentially Expressed Genes among Clusters.
Given that clust1 and clust2 had superior prognosis than clust3, we combined the data of clust1 and clust2 named clust12. By comparing expression profiles between clust12 and clust3 in TCGA-COAD and GSE17536 datasets, we identified a number of DEGs between two groups under conditions of |log2 fold change (FC)| > log2 (1.5) and FDR <0.05 (Figures 4(a) and 4(b)). 11 upregulated and 75 downregulated genes (clust12 versus clust3) were screened in both datasets (Figure 4(c)). Furthermore, based on 86 genes, we used String (https://cn.string-db.org/) to screen their interactions and found 69 interacted genes visualized by CytoScape (Figure 4(d)). Most of the genes were downregulated and only two genes (ADH1C and UGT2A3) within 69 interacted genes were upregulated. Genes with degree >10 were listed (Supplementary Table S2). Subsequently, we conducted GO and KEGG analysis on 69 genes. Within GO analysis, 92 terms of biological process, 20 terms of molecular function, and 13 terms of cellular component were annotated (Supplementary Table S3, FDR < 0.05). The top 10 enriched terms were visualized (Supplementary Figures S8(a)–S8(c)). In addition, 9 KEGG pathways were significantly enriched such as ECM-receptor interaction, focal adhesion, and PI3K-Akt signaling pathway (FDR < 0.05, Supplementary Figure S8(d)).

### 3.5. Identifying Eight Cell Types Based on scRNA-Seq Data.
To further verify the reliability of our molecular subtyping based on four metabolism pathways, we introduced scRNA-seq data of COAD and sought to distinguish malignant and nonmalignant cells. Single-cell data was preliminarily processed and screened to meet the standards of each gene expressed at least in three cells, each cell expressing at least 250 genes, each cell containing less than 35% mitochondria, and UMI > 100 for each cell (Supplementary Figures S9(a)–S9(c)). 45892 single cells remained. Then PCA was applied to diminish the dimensionality of the data. Cells were clustered into 34 clusters (resolution = 0.9) and CD45 marker was used to identify immune cells with a total number of 26188 (Supplementary Figures S9(d) and S9(e)). Subsequently, 26188 immune cells were further clustered under resolution = 0.3. 12 subgroups were identified, and t-SNE plots grouped by different samples, patients, and tissues were shown (Figures 5(a)–5(d)). By using immune markers from CellMarker (http://biocc.hrbmu.edu.cn/CellMarker/), 8 cell types were annotated, where T cells and B cells represented the majority (Figures 5(e)–5(g)). The top 5 DEGs among eight cell types were identified (Figure 5(f)), indicating the distinct expression features of them. KEGG analysis on 8 cell types manifested that 29 pathways were significantly enriched and many of them were related to immunity (Figure 5(h)).

3.6. Malignant Cells Were More Enriched in Clust3.
As the complicity of tumor tissues that may contain normal cells, we employed CopyKat to estimate genomic copy number profiles for more accurately distinguishing malignant and nonmalignant cells (Supplementary Figure S10) [36]. As a result, 8055 malignant cells and 18133 normal cells were distinguished. In the previous section, we identified 11 upregulated genes and 75 downregulated genes by comparing clust12 with clust3 (Figure 4(c)). We used ssGSEA to calculate their enrichment score of upregulated and downregulated genes in malignant and nonmalignant cells, respectively. Significantly differential enrichment of these genes was exhibited in two groups, with the 11 DEGs higher expression in nonmalignant cells and the 75 DEGs higher expression in malignant cells ($P < 0.0001$, Figure 6). The result demonstrated that the molecular subtyping based on metabolism pathways was reliable, and these dysregulated genes may be highly associated with the regulation of four metabolism pathways.

### 3.7. Metabolism of Malignant Cells Was Associated with TME.
To understand if TME made a difference on the metabolism of malignant cells, we introduced hypoxia and angiogenesis as indicators. Here hypoxia score was calculated by ssGSEA based on genes within hypoxia pathway, and angiogenesis score was calculated based on the signature from Masiero et al. [37]. Pearson correlation analysis revealed positive correlations between either two of the four metabolism pathways in malignant cells, especially between glycolysis and PPP (Figure 7(a)). In addition, we applied Mantel test to evaluate the relation of angiogenesis and hypoxia with four metabolism pathways. We observed that angiogenesis score was negatively correlated with four pathways, while positive correlation was shown between hypoxia and four pathways
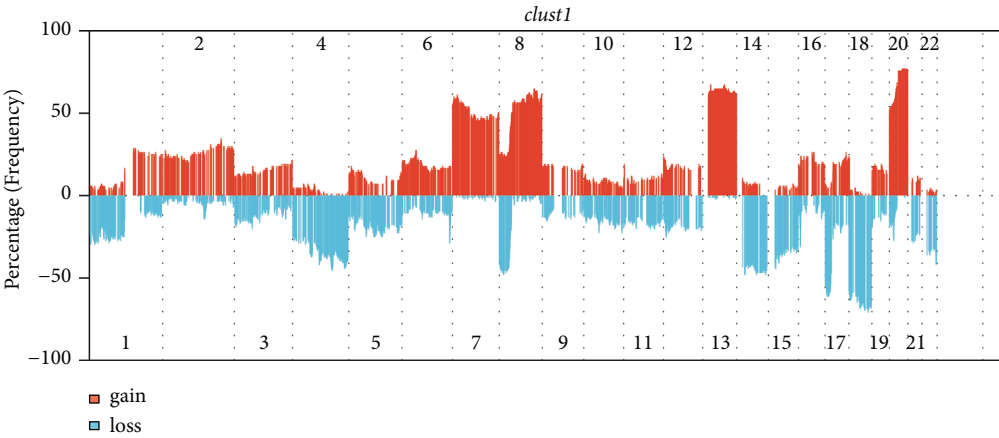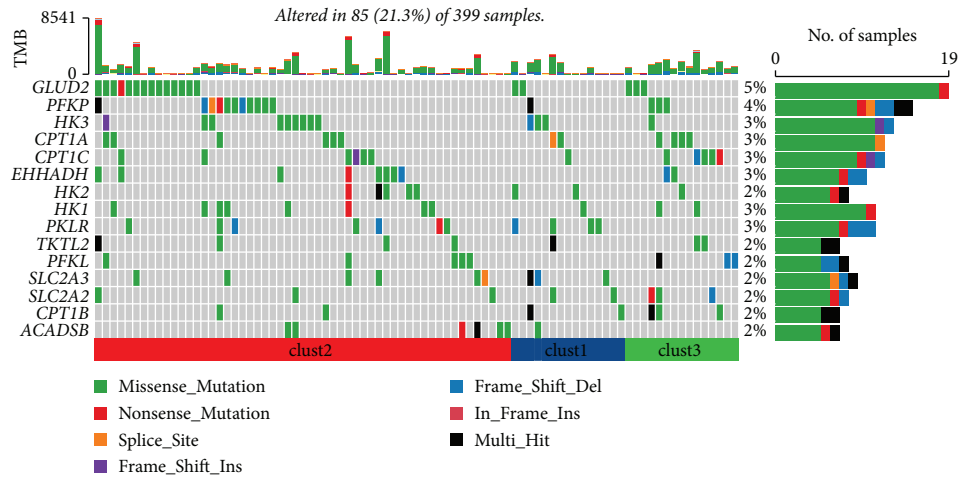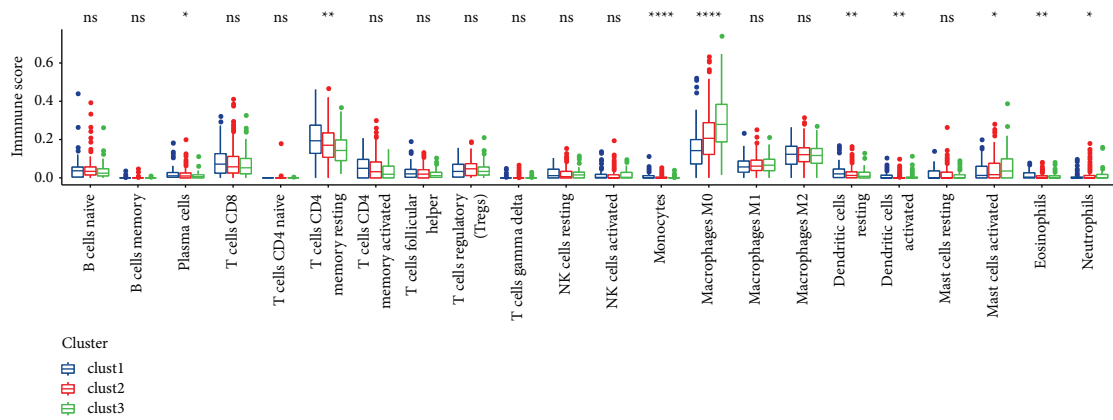
(a)
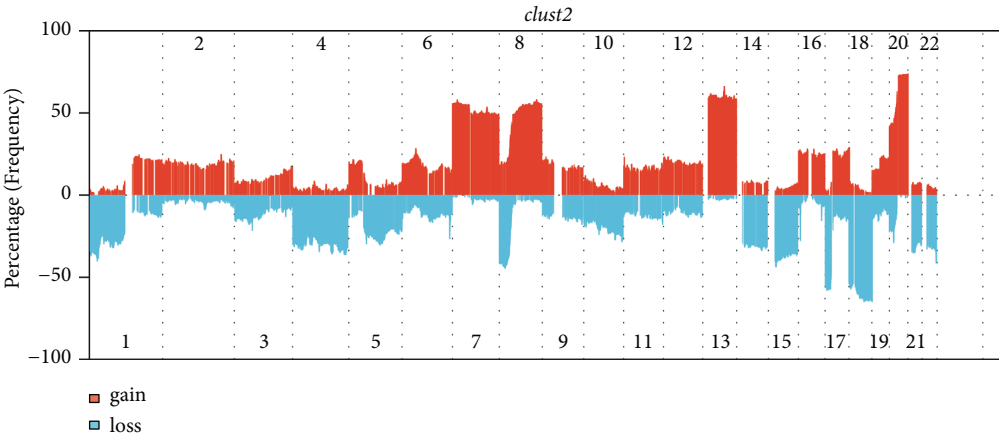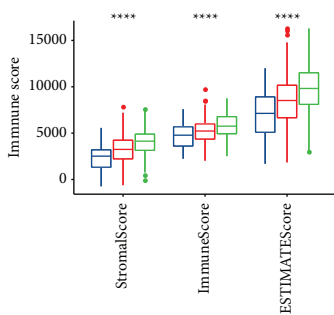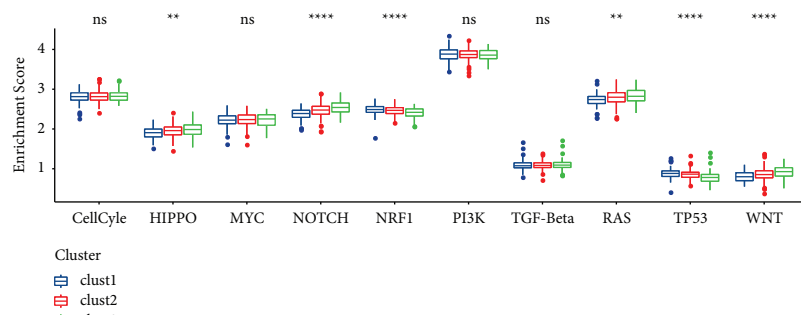

(b)
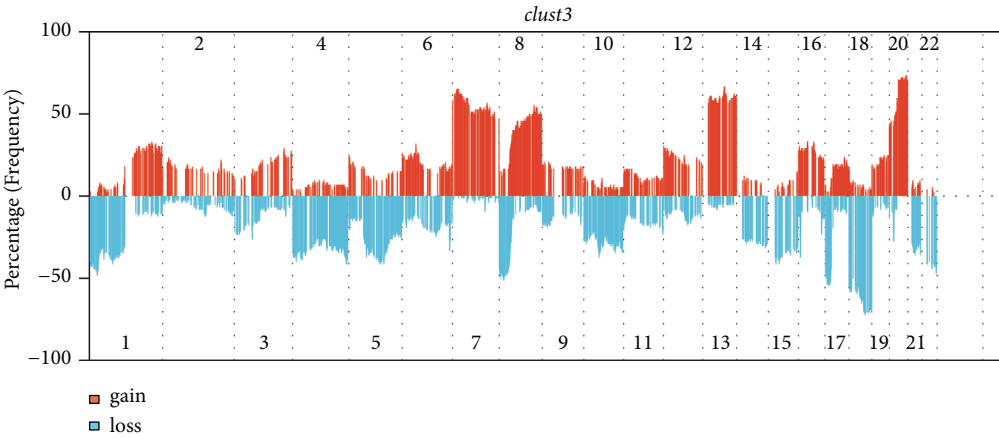
(c)

FIGURE 2: Continued.

(d)

FIGURE 2: Genomic features of three clusters in TCGA-COAD dataset. ((a)–(c)) The proportion of CNVs in 22 chromosomes grouped by clust1, clust2, and clust3. (d) The top 10 mutated genes within metabolism pathways (Fisher's test, $P < 0.05$).

(a)

(b)

(c)

FIGURE 3: Continued.
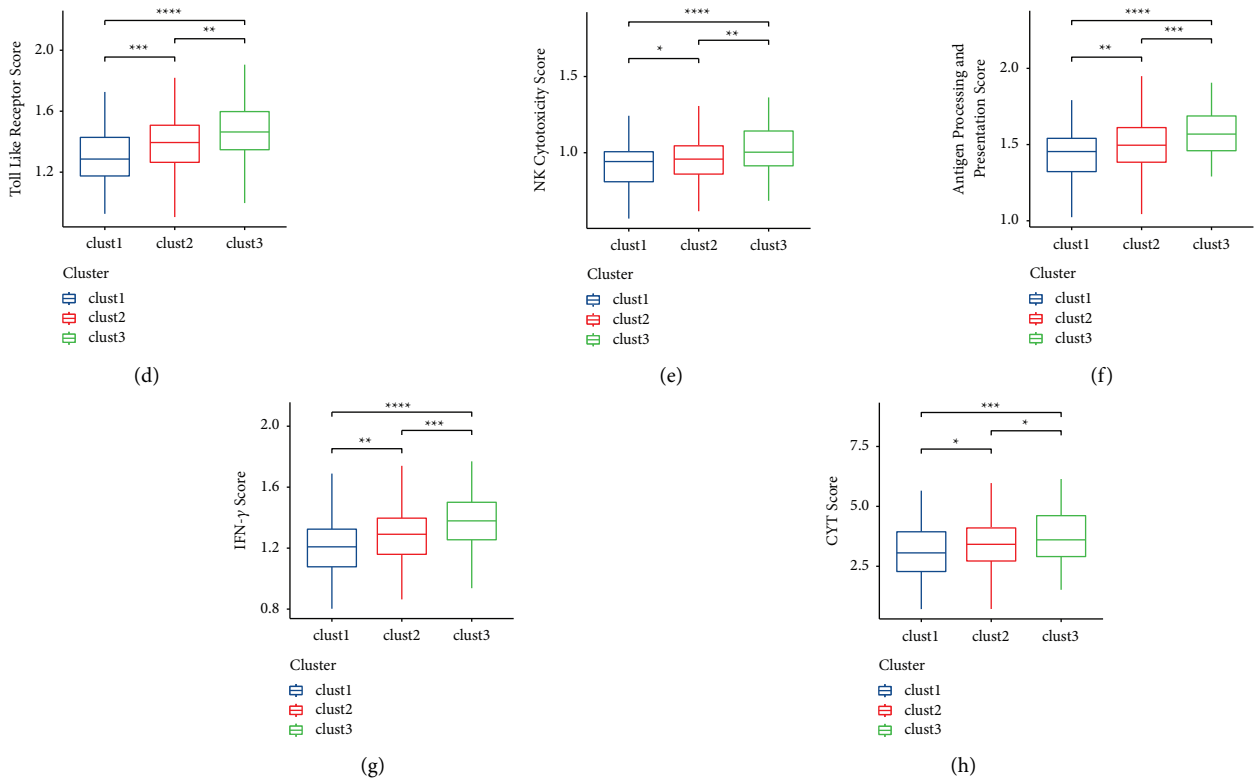
(d)

(e)

(f)

(g)

(h)

FIGURE 3: TME features of three clusters in TCGA-COAD dataset. (a) Estimated proportion of 22 immune cells. (b) Stromal score, immune score, and ESTIMATE score calculated by ESTIMATE. (c) Enrichment score of 10 oncogenic pathways. ((d)–(h)) Enrichment score of toll-like receptor, NK cytotoxicity, antigen processing and presentation, IFN-γ, and CYT. ANOVA was conducted. ns, no significance. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, and $^{****}P < 0.0001$.
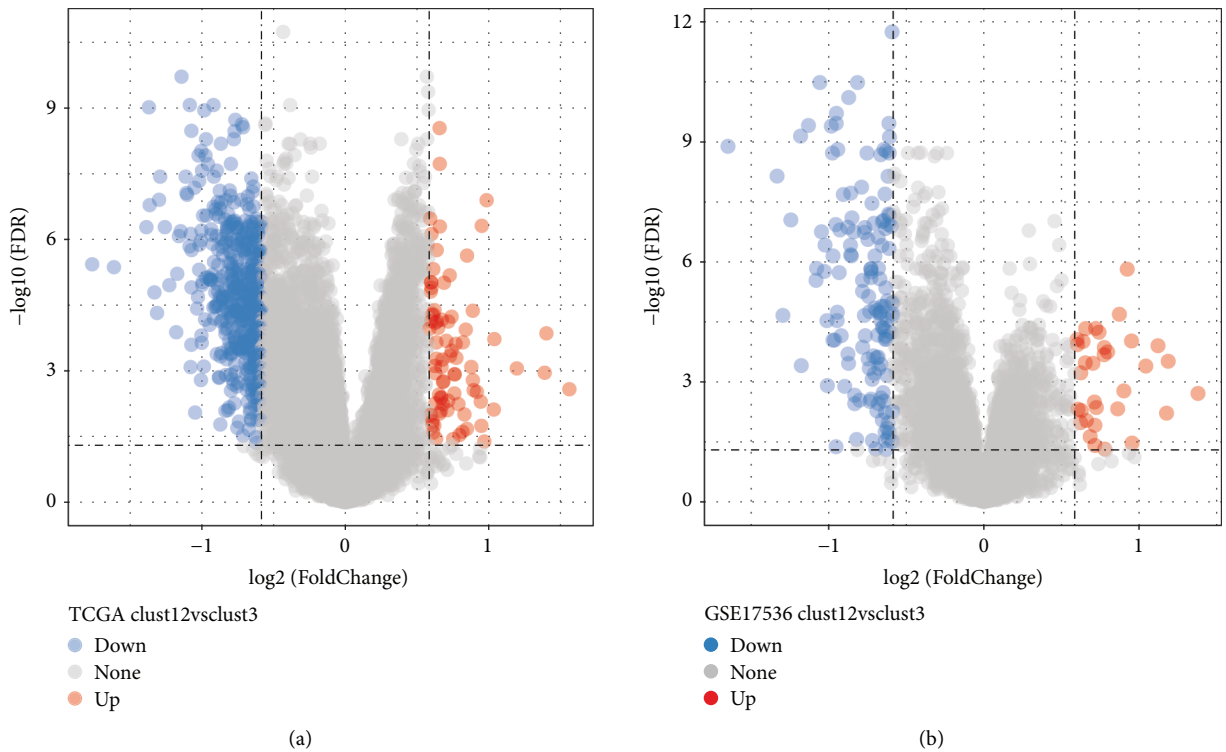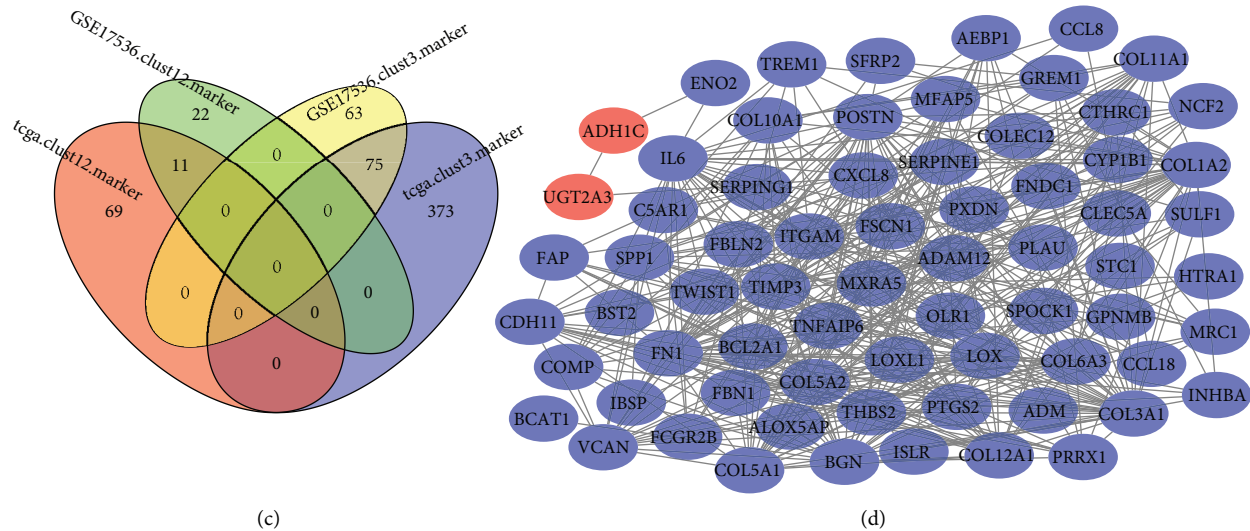

(a)

(b)

FIGURE 4: Continued.

(c)

(d)

Figure 4: Identifying dysregulated genes related to prognosis. ((a) and (b)) DEGs between clust12 and clust3 in TCGA-COAD and GSE17536 datasets. Red and blue indicate upregulated and downregulated genes, respectively. (c) Venn plot of DEGs between clust12 and clust3 in two datasets. (d) Protein-protein interaction (PPI) analysis for 86 dysregulated genes. Blue indicates downregulated genes and red indicates upregulated genes.

(Figure 7(a)). Notably, significantly high correlation coefficient was observed between hypoxia and glycolysis score ($R = 0.62$), indicating the strong interaction between two pathways.

Malignant cells mostly consisted of B cells, macrophages, and T cells (Supplementary Table S4). Notably, all detected cell types were more enriched in clust12 compared with clust3 ($P < 0.0001$, Figure 7(b)). Furthermore, we found that clust12 score was negatively correlated with hypoxia ($R = -0.46$), glycolysis ($R = -0.58$), PPP ($R = -0.51$), and FAO ($R = -0.35$) particularly (Figure 7(c)). Conversely, clust3 score was positively correlated with hypoxia and angiogenesis especially ($R = 0.50$ and 0.40, respectively, Figure 7(c)). In addition, we found that the enrichment score of hypoxia was significantly differential among three subtypes, where clust3 had the highest enrichment score (Supplementary Figure S11), which was consistent with the above result. Hypoxia and angiogenesis are important TME hallmarks in cancer. The results suggested that metabolism regulation of malignant cells was possibly regulated by hypoxia and angiogenesis, as well as TME. Additionally, the significantly differential distribution of immune cells between clust12 and clust3 proved that 86 markers (DEGs) were effective and reliable for distinguishing two groups in single-cell data.

4. Discussion

Tumor metabolism has a major impact in TME modulation [38,39], but the link between tumor metabolism and TME has not been clearly understood in COAD so far. In this study, we dug out key genes interacting with both four major metabolism pathways and TME by analyzing expression data and single-cell data. Initially, three clusters were established through unsupervised consensus clustering

based on the enrichment score of four metabolism pathways. Given that clust1 and clust2 had similar overall survival, we combined the expression data of two clusters. Then 86 important DEGs were identified between clust12 and clust3, and 69 of 86 DEGs presented close interactions through PPI analysis.

We estimated that these DEGs may play a critical role in tumor metabolism and have a pronounced impact on COAD prognosis. As demonstrated in the single-cell analysis, malignant cells and nonmalignant cells had distinct expression levels of 86 DEGs (11 as clust12 markers and 75 as clust3 markers). Malignant cells had extremely high expression level of 75 DEGs associated with poor prognosis in clust3, which illustrated the important role of 75 DEGs in tumor progression, as well as the effectiveness and reasonability of subtyping based on four metabolism pathways.

Hypoxia and angiogenesis are two important characteristics in TME of solid tumor, which contributes to poor prognosis and drug resistance or inferior efficiency of immunotherapy [40–43]. Numerous studies have demonstrated that hypoxia can induce glycolytic flux through hypoxia-inducible factor-1α (HIF-1α). HIF-independent reprogramming is activated, which promotes expression of glucose transporters and enzymes involved in glucose pathway [44]. Immune cell metabolism can be altered HIF-1α activated glycolysis, where various immune cells are affected such as lymphocytes, dendritic cells, and macrophages [45,46]. Immunosuppressive cells including tumor-associated macrophages (TAMs) and myeloid-derived suppressor cells (MDSCs) are increased via HIF-1α signaling [47]. High activity of glycolysis also inhibits T cell function but favors infiltration of M2 macrophages which can suppress antitumor immune response and promote tumor growth [48]. In this study, striking correlation was observed between hypoxia and glycolysis ($R = 0.62$), between clust3
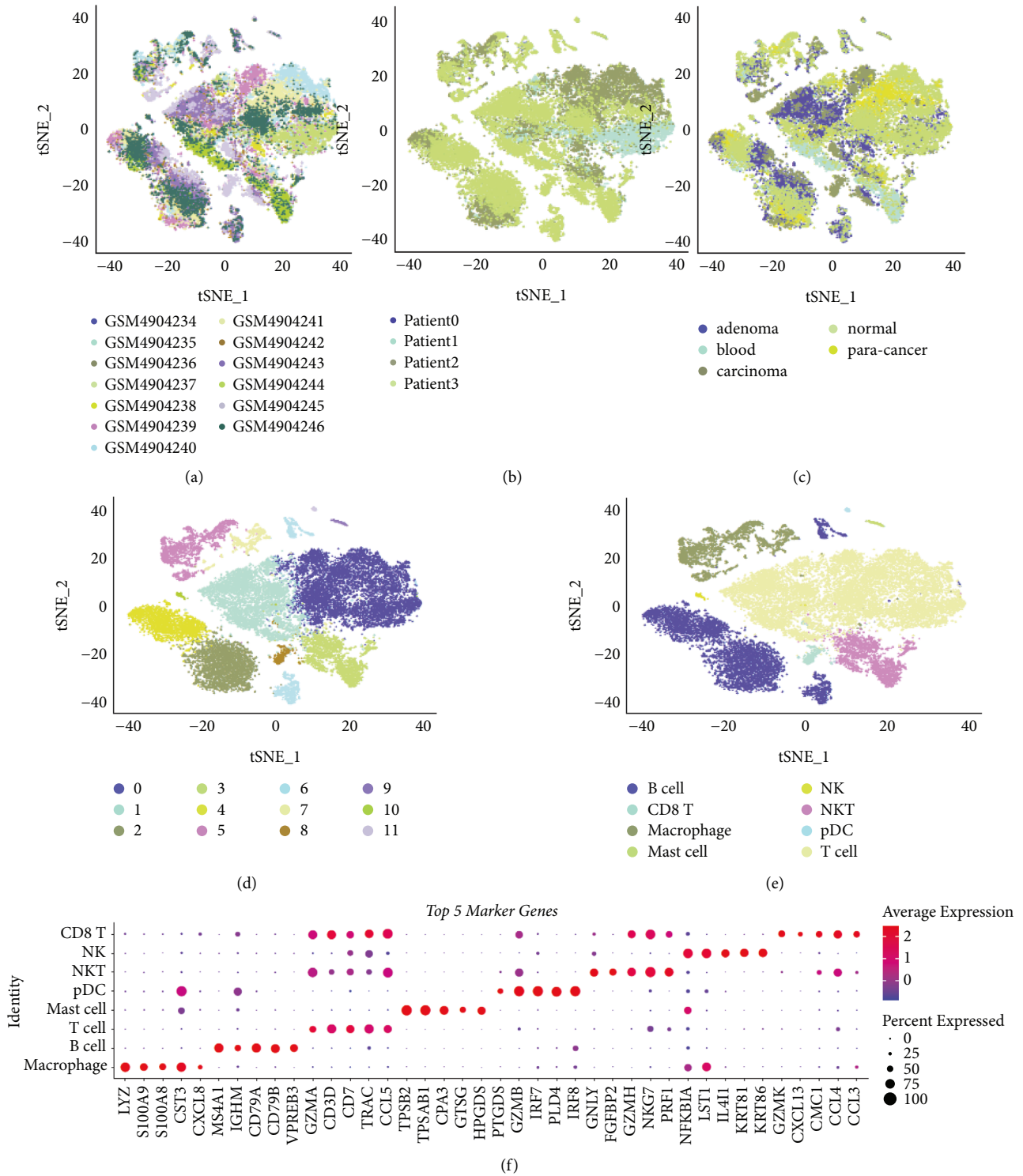
(a)

(b)

(c)

(d)

(e)

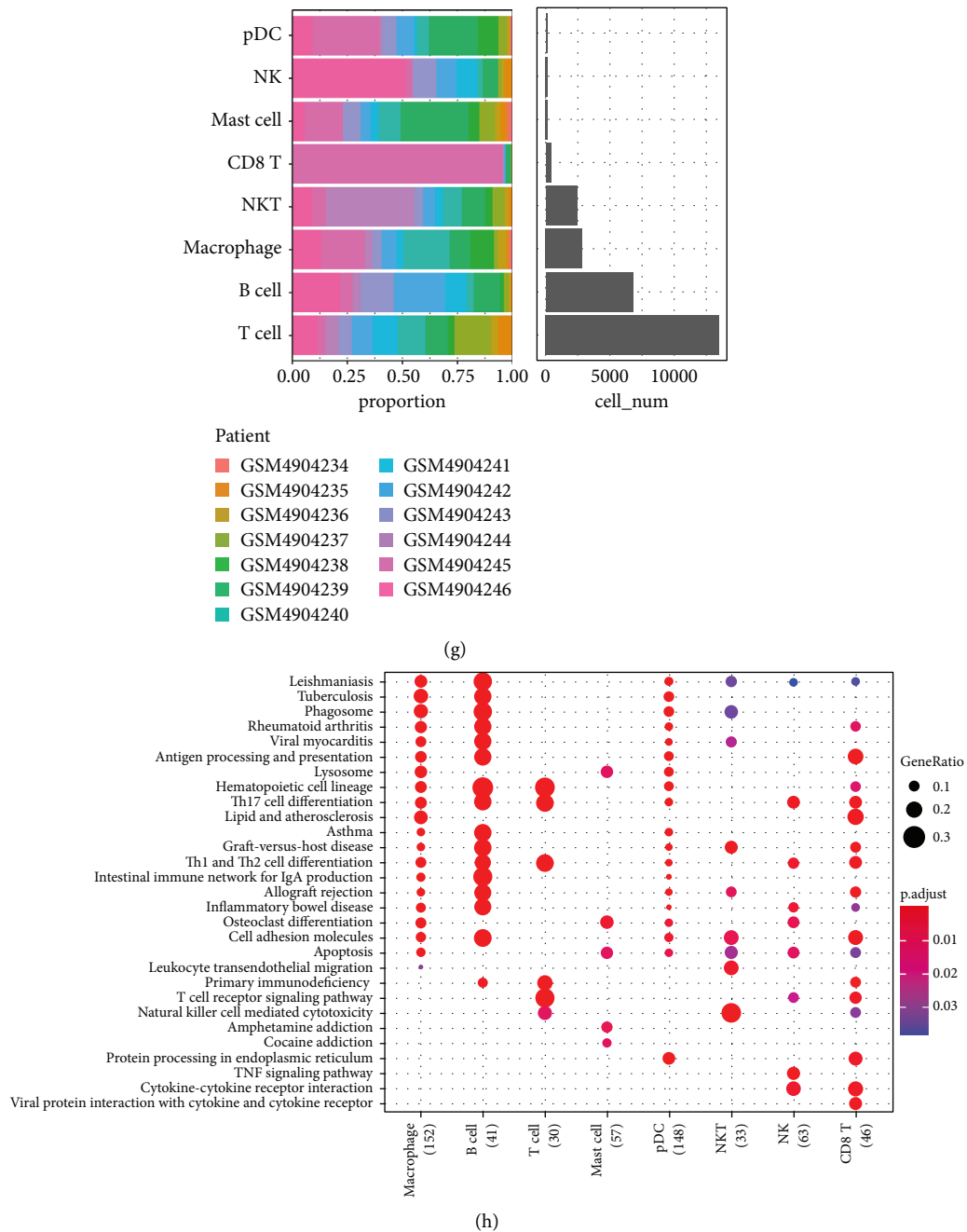Top 5 Marker Genes

(f)

FIGURE 5: Continued.

(g)

(h)

Figure 5: Single-cell analysis of GSE161277 dataset. ((a)–(c)) T-SNE plots of 26188 single cells grouped by samples (a), patients (b), and tissues (c). (d) Subgrouping of single cells. (e) Identification of eight cell types. (f) The top 5 DEGs (markers) of eight cell types. (g) The cell counts of different cell types and the distribution of samples. (h) Significantly enriched pathways of eight cell types.

score and hypoxia ($R = 0.50$), between clust12 score and hypoxia ($R = -0.46$), and between clust12 score and glycolysis ($R = -0.58$). The correlation between clust3 score and glycolysis was relatively slight ($R = 0.21$), indicating that 75 markers were more involved in the signaling of hypoxia and thus leading to activated glycolysis in tumor tissues related to poor prognosis.

Besides glycolysis, PPP was also strongly correlated with hypoxia ($R = 0.42$), and clust12 score was shown to be negatively associated with PPP ($R = -0.51$). PPP is one of

major pathway for glucose catabolism, and it is reprogrammed accompanied with the activation of glycolysis in tumor cells [49]. In the human colon cancer cell line (HT29), enzymes within PPP pathway are enhanced during cell cycle progression [50], suggesting that PPP is a potential target for inhibiting tumor cell proliferation [51]. High expression of 11 marker genes in clust12 may play an important role for alleviating PPP activation.

Angiogenesis is more activated in clust3 revealed from a correlation coefficient of 0.40 between clust3 score

(a)

(b)

(c)
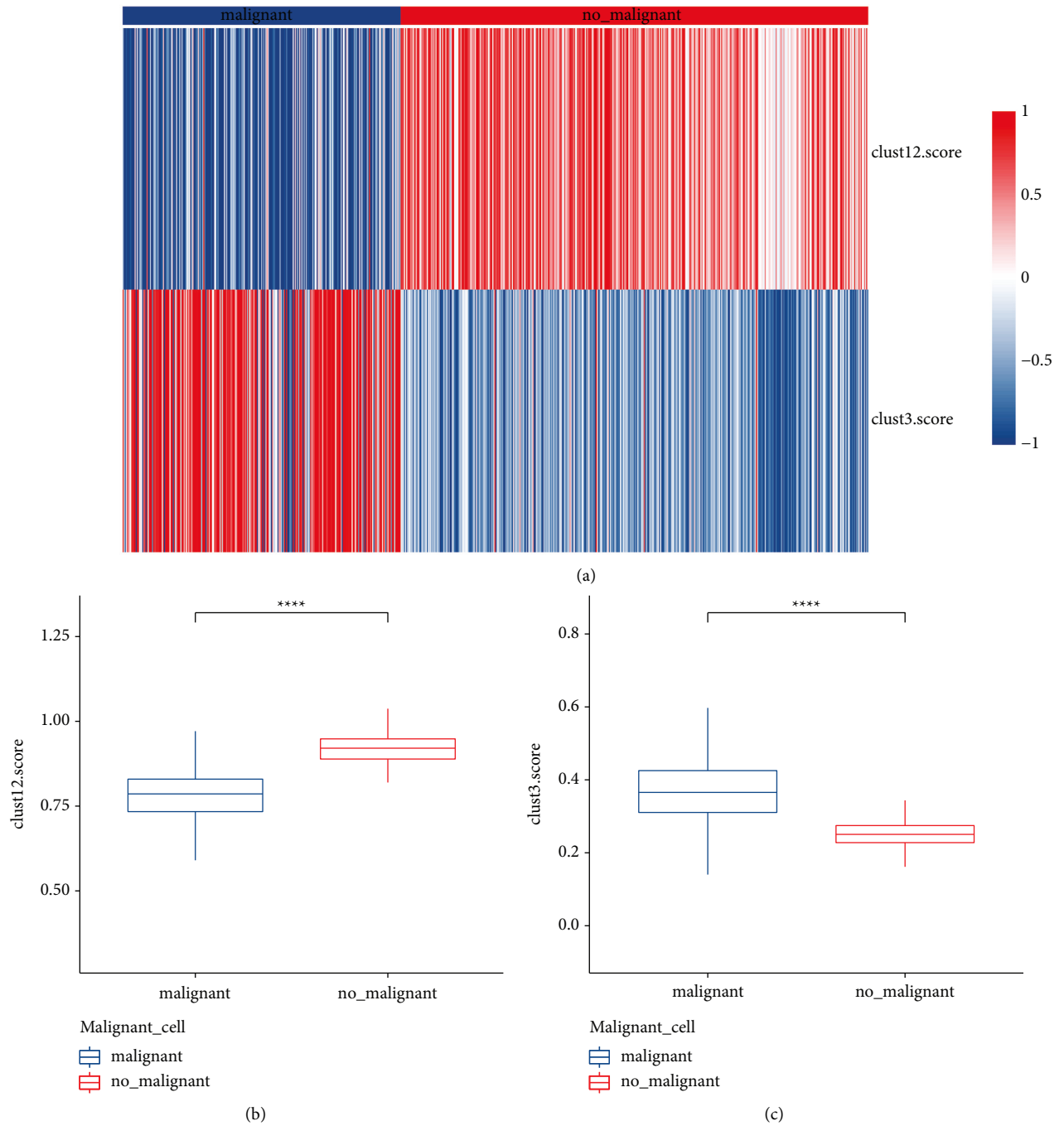
FIGURE 6: The distribution of clust12 score and clust3 score in malignant and nonmalignant cells. Wilcoxon test was performed. ****$P < 0.0001$.

and angiogenesis. Several targeted drugs display anti-angiogenesis effect in colon cancer, such as regorafenib [52], anti-interleukin-6 receptor antibody [53], and SARI [54]. Angiogenesis is regulated by innate immune cells such as TAMs, MDSCs, tumor-associated neutrophils

(TANs), mast cells, and NK cells, which produce proangiogenic factors [55], and promote immunosuppressive TME [56]. Other two metabolism pathways, FAO and glutaminolysis, exhibited a relatively weak correlation with either hypoxia or angiogenesis,

(a)

(b)

(c)
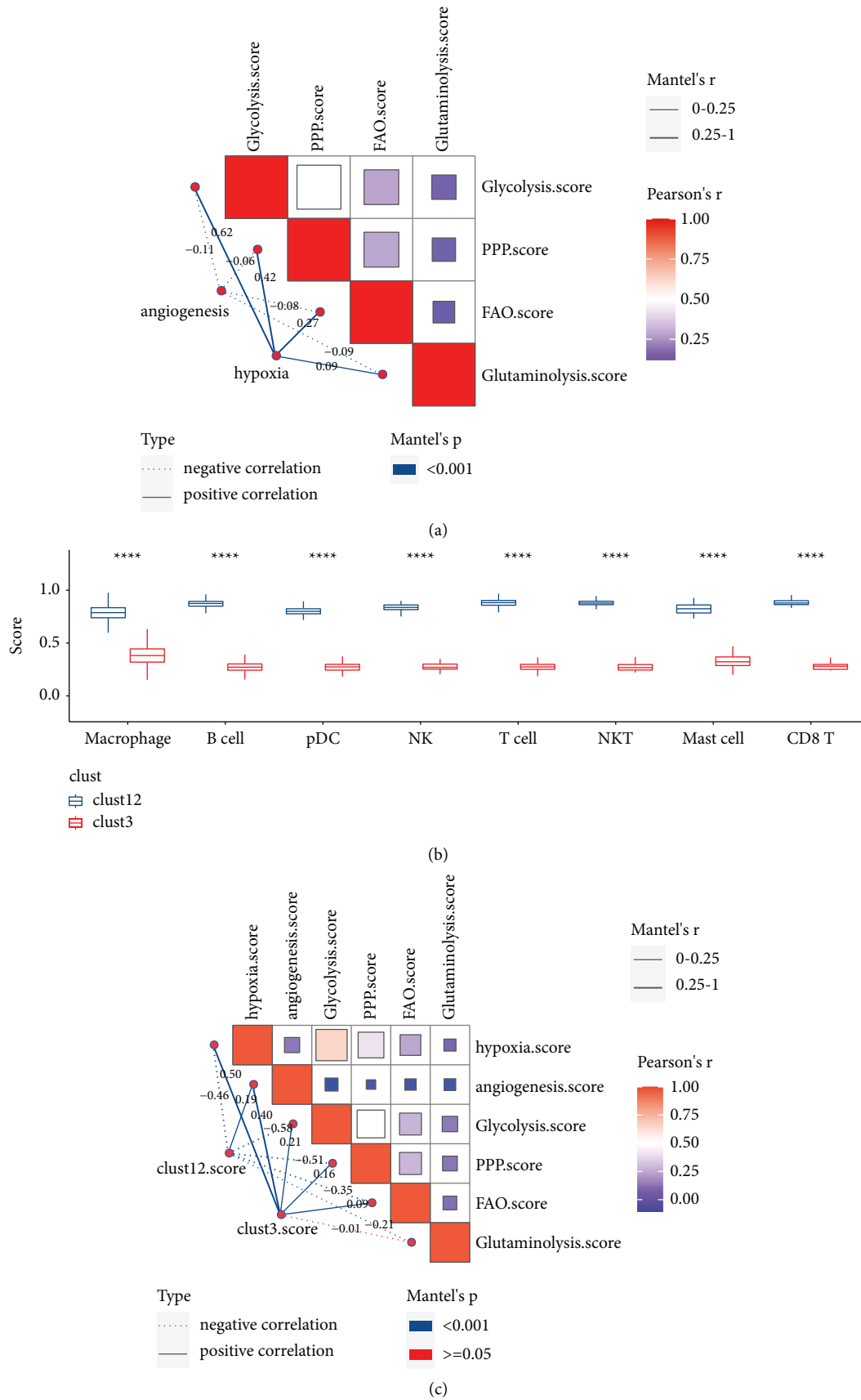
FIGURE 7: Analysis of the relation between TME and metabolism pathways. (a) Correlation analysis between hypoxia, angiogenesis, and four metabolism pathways. (b) Enrichment score of eight immune cell types in malignant and nonmalignant cells. (c) The correlation of clust12 score and clust3 score with hypoxia, angiogenesis, and four metabolism pathways. In (a) and (c), dashed lines indicate negative correlation and solid lines indicate positive correlation. Thicker lines indicate higher correlation. Pearson correlation analysis was conducted among pathways in top-right matrix. Mantel test was conducted between two matrixes. Wilcoxon test was performed in (b) and (c). ****$P < 0.0001$.

indicating that they may be less involved in dysregulated immunity in COAD.

5. Conclusions

In conclusion, this study depicted a crosstalk between tumor metabolism and TME features for COAD through integrated analysis of expression data and scRNA-seq data. Three molecular subtypes were constructed based on metabolism-related genes, and they performed significant differences in metabolism, genomic variations, immune infiltration, and enriched pathways, suggesting that the four metabolism pathways may regulate TME and thus affect COAD progression. Importantly, we identified 86 differentially expressed genes from three subtypes that were illustrated in scRNA-seq data to be highly associated with malignancy, hypoxia, and glycolysis. We envisioned that these key genes may provide new directions for further revealing the mechanism of metabolism in TME modulation.

Data Availability

The datasets analyzed in this study can be found in [GSE17536] at https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE17536 and in [GSE161277] at https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE161277.

Conflicts of Interest

The authors declare no potential conflicts of interest.

Authors' Contributions

Xiaofang Qiao, Guangmei Zhang, and Yajie Xiao share equal contribution.

Supplementary Materials

Supplementary Figure S1: the flowchart of this study. Supplementary Figure S2: expression of genes within four metabolism pathways in three clusters. Kruskal test was conducted. ns, no significance. $**P < 0.01$, $***P < 0.001$, and $****P < 0.0001$. Supplementary Figure S3: the distribution of three clusters in PCA plots. Supplementary Figure S4: the distribution of different clinical features in three clusters. Chi-square test was conducted. Supplementary Figure S5: significant CNVs detected by gistic2. Left line indicates chromosomes. Right line indicates significant CNVs (q value < 0.25). The bottom line indicates q values and the upper line indicates G score. Supplementary Figure S6: the enrichment score of immune cells in TCGA-COAD dataset analyzed by EPIC. Kruskal test was conducted. ns, no significant. $**P < 0.01$ and $****P < 0.001$. Supplementary Figure S7: TME features of three clusters in GSE17536 dataset. (a) Estimated proportion of 22 immune cells. (b) Stromal score, immune score, and ESTIMATE score calculated by ESTIMATE. (c) Enrichment score of 10 oncogenic pathways. ((d)–(h)) Enrichment scores of toll-like receptor, NK cytotoxicity, antigen processing and presentation, IFN-γ,

and CYT. ANOVA was conducted. ns, no significance. $*P < 0.05$, $**P < 0.01$, $***P < 0.01$, and $****P < 0.0001$. Supplementary Figure S8: functional analysis on 69 genes within PPI network in TCGA-COAD dataset. ((a)–(c)) The top 10 significantly enriched terms in biological process (a), cellular component (b), and molecular function (c). (d) Nine significantly enriched KEGG pathways. Supplementary Figure S9: quality control and preprocessing of single-cell data. Supplementary Figure S10: the CNV heatmap of CopyKat for distinguishing malignant and nonmalignant cells. Orange indicates malignant cells and green indicates nonmalignant cells. Supplementary Figure S11: the ssGSEA enrichment score of hypoxia in three subtypes in TCGA-COAD dataset. Wilcoxon test was conducted. $****P < 0.0001$. Supplementary Table S1: a gene list of four metabolism pathways. Supplementary Table S2: screened genes with degree >10. Supplementary Table S3: the significantly enriched GO terms and KEGG pathways in TCGA-COAD dataset. Supplementary Table S4: the composed cell types of malignant cells. (*Supplementary Materials*)

References

[1] H. Sung, J. Ferlay, R. L. Siegel et al., "Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA: A Cancer Journal for Clinicians*, vol. 71, no. 3, pp. 209–249, 2021.

[2] M. Ahmed, "Colon cancer: a clinician's perspective in 2019," *Gastroenterology Research*, vol. 13, no. 1, pp. 1–10, 2020.

[3] A. Hadjipetrou, D. Anyfantakis, C. G. Galanakis, M. Kastanakis, and S. Kastanakis, "Colorectal cancer screening and primary care: a mini literature review," *World Journal of Gastroenterology*, vol. 23, no. 33, pp. 6049–6058, 2017.

[4] Y. Ohhara, N. Fukuda, S. Takeuchi et al., "Role of targeted therapy in metastatic colorectal cancer," *World Journal of Gastrointestinal Oncology*, vol. 8, no. 9, p. 642, 2016.

[5] C. Li, G. Zhang, L. Zhao, Z. Ma, and H. Chen, "Metabolic reprogramming in cancer cells: glycolysis glutaminolysis and Bcl-2 proteins as novel therapeutic targets for cancer," *World Journal of Surgical Oncology*, vol. 14, no. 1, p. 15, 2016.

[6] L. Yang, S. Venneti, and D. Nagrath, "Glutaminolysis: a hallmark of cancer metabolism," *Annual Review of Biomedical Engineering*, vol. 19, no. 1, pp. 163–194, 2017.

[7] Y. Ma, S. M. Temkin, A. M. Hawkridge et al., "Fatty acid oxidation: an emerging facet of metabolic transformation in cancer," *Cancer Letters*, vol. 435, pp. 92–100, 2018.

[8] E. S. Cho, Y. H. Cha, H. S. Kim, N. H. Kim, and J. I. Yook, "The pentose phosphate pathway as a potential target for cancer therapy," *Biomolecules & therapeutics*, vol. 26, no. 1, pp. 29–38, 2018.

[9] A. V. Orang, J. Petersen, R. A. McKinnon, and M. Z. Michael, "Micromanaging aerobic respiration and glycolysis in cancer cells," *Molecular Metabolism*, vol. 23, pp. 98–126, 2019.

[10] Z. H. Wang, Y. Z. Zhang, Y. S. Wang, and X. X. Ma, "Identification of novel cell glycolysis related gene signature predicting survival in patients with endometrial cancer," *Cancer Cell International*, vol. 19, no. 1, p. 296, 2019.

[11] J. Bi, F. Bi, X. Pan, and Q. Yang, "Establishment of a novel glycolysis-related prognostic gene signature for ovarian cancer and its relationships with immune infiltration of the

tumor microenvironment," *Journal of Translational Medicine*, vol. 19, no. 1, p. 382, 2021.

[12] L. Zhang, Z. Zhang, and Z. Yu, "Identification of a novel glycolysis-related gene signature for predicting metastasis and survival in patients with lung adenocarcinoma," *Journal of Translational Medicine*, vol. 17, no. 1, p. 423, 2019.

[13] J. Zhu, S. Wang, H. Bai et al., "Identification of five glycolysis-related gene signature and risk score model for colorectal cancer," *Frontiers in Oncology*, vol. 11, 2021.

[14] W. Shi, Y. Luo, D. Zhao, H. Huang, and W. Pang, "Evaluation of the benefit of post-mastectomy radiotherapy in patients with early-stage breast cancer: a propensity score matching study," *Oncology Letters*, vol. 17, 2019.

[15] W. Shi, D. Hu, S. Lin, and R. Zhuo, "Five-mRNA signature for the prognosis of breast cancer based on the ceRNA network," *BioMed Research International*, vol. 2020, Article ID 9081852, 17 pages, 2020.

[16] Z. Chen, M. Wang, R. L. De Wilde et al., "A machine learning model to predict the triple negative breast cancer immune subtype," *Frontiers in Immunology*, vol. 12, 2021.

[17] W. Sun, H. Shi, Z. Yuan et al., "Prognostic value of genes and immune infiltration in prostate tumor microenvironment," *Frontiers in Oncology*, vol. 10, 2020.

[18] J. J. Smith, N. G. Deane, F. Wu et al., "Experimentally derived metastasis gene expression profile predicts recurrence and death in patients with colon cancer," *Gastroenterology*, vol. 138, no. 3, pp. 958–968, 2010.

[19] X. Zheng, J. Song, C. Yu et al., "Single-cell transcriptomic profiling unravels the adenoma-initiation role of protein tyrosine kinases during colorectal tumorigenesis," *Signal Transduction and Targeted Therapy*, vol. 7, no. 1, p. 60, 2022.

[20] A. Liberzon, C. Birger, H. Thorvaldsdottir, M. Ghandi, J. Mesirov, and P. Tamayo, "The molecular signatures database hallmark gene set collection," *Cell systems*, vol. 1, no. 6, pp. 417–425, 2015.

[21] S. Hänzelmann, R. Castelo, and J. Guinney, "GSVA: gene set variation analysis for microarray and RNA-seq data," *BMC Bioinformatics*, vol. 14, no. 1, p. 7, 2013.

[22] M. D. Wilkerson and D. N. Hayes, "ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking," *Bioinformatics*, vol. 26, no. 12, 2010.

[23] C. H. Mermel, S. E. Schumacher, B. Hill, M. L. Meyerson, R. Beroukhim, and G. Getz, "GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers," *Genome Biology*, vol. 12, no. 4, 2011.

[24] A. Mayakonda, D. C. Lin, Y. Assenov, C. Plass, and H. P. Koeffler, "Maftools: efficient and comprehensive analysis of somatic variants in cancer," *Genome Research*, vol. 28, no. 11, pp. 1747–1756, 2018.

[25] E. Becht, N. A. Giraldo, L. Lacroix et al., "Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression," *Genome Biology*, vol. 17, no. 1, p. 218, 2016.

[26] J. Racle, K. de Jonge, P. Baumgaertner, D. E. Speiser, and D. Gfeller, "Simultaneous enumeration of cancer and immune cell types from bulk tumor gene expression data," *Elife*, vol. 6, 2017.

[27] K. Yoshihara, M. Shahmoradgoli, E. Martinez et al., "Inferring tumour purity and stromal and immune cell admixture from expression data," *Nature Communications*, vol. 4, no. 1, 2013.

[28] L. Danilova, W. J. Ho, Q. Zhu et al., "Programmed cell death ligand-1 (PD-L1) and CD8 expression profiling identify an immunologic subtype of pancreatic ductal adenocarcinomas

with favorable survival," *Cancer Immunology Research*, vol. 7, no. 6, pp. 886–895, 2019.

[29] M. S. Rooney, S. A. Shukla, C. J. Wu, G. Getz, and N. Hacohen, "Molecular and genetic properties of tumors associated with local immune cytolytic activity," *Cell*, vol. 160, 2015.

[30] M. E. Ritchie, B. Phipson, D. Wu et al., "Limma powers differential expression analyses for RNA-sequencing and microarray studies," *Nucleic Acids Research*, vol. 43, 2015.

[31] G. Su, J. H. Morris, B. Demchak, and G. D. Bader, "Biological network exploration with Cytoscape 3," *Current protocols in bioinformatics*, vol. 47, 2014.

[32] Y. Liao, J. Wang, E. J. Jaehnig, Z. Shi, and B. Zhang, "WebGestalt 2019: gene set analysis toolkit with revamped UIs and APIs," *Nucleic Acids Research*, vol. 47, 2019.

[33] A. Gribov, M. Sill, S. Luck et al., "SEURAT: visual analytics for the integrated analysis of microarray data," *BMC Medical Genomics*, vol. 3, no. 1, p. 21, 2010.

[34] G. Yu, L. G. Wang, Y. Han, and Q. Y. He, "clusterProfiler: an R package for comparing biological themes among gene clusters," *OMICS: A Journal of Integrative Biology*, vol. 16, no. 5, pp. 284–287, 2012.

[35] F. Sanchez-Vega, M. Mina, J. Armenia et al., "Oncogenic signaling pathways in the cancer genome Atlas," *Cell*, vol. 173, 2018.

[36] R. Gao, S. Bai, Y. C. Henderson et al., "Delineating copy number and clonal substructure in human tumors from single-cell transcriptomes," *Nature Biotechnology*, vol. 39, no. 5, pp. 599–608, 2021.

[37] M. Masiero, F. Simoes, H. Han et al., "A core human primary tumor angiogenesis signature identifies the endothelial orphan receptor ELTD1 as a key regulator of angiogenesis," *Cancer Cell*, vol. 24, no. 2, pp. 229–241, 2013.

[38] M. Reina-Campos, J. Moscat, and M. Diaz-Meco, "Metabolism shapes the tumor microenvironment," *Current Opinion in Cell Biology*, vol. 48, pp. 47–53, 2017.

[39] C. A. Lyssiotis and A. C. Kimmelman, "Metabolic interactions in the tumor microenvironment," *Trends in Cell Biology*, vol. 27, no. 11, pp. 863–875, 2017.

[40] A. Vito, N. El-Sayes, and K. Mossman, "Hypoxia-driven immune escape in the tumor microenvironment," *Cells*, vol. 9, no. 4, p. 992, 2020.

[41] X. Jing, F. Yang, C. Shao et al., "Role of hypoxia in cancer therapy by regulating the tumor microenvironment," *Molecular Cancer*, vol. 18, no. 1, p. 157, 2019.

[42] M. De Palma, D. Biziato, and T. V. Petrova, "Microenvironmental regulation of tumour angiogenesis," *Nature Reviews Cancer*, vol. 17, no. 8, pp. 457–474, 2017.

[43] R. I. Teleanu, C. Chircov, A. M. Grumezescu, and D. M. Teleanu, "Tumor angiogenesis and anti-angiogenic strategies for cancer treatment," *Journal of Clinical Medicine*, vol. 9, no. 1, p. 84, 2019.

[44] I. Papandreou, R. A. Cairns, L. Fontana, A. L. Lim, and N. C. Denko, "HIF-1 mediates adaptation to hypoxia by actively downregulating mitochondrial oxygen consumption," *Cell Metabolism*, vol. 3, no. 3, pp. 187–197, 2006.

[45] K. Ganeshan and A. Chawla, "Metabolic regulation of immune responses," *Annual Review of Immunology*, vol. 32, no. 1, pp. 609–634, 2014.

[46] S. P. Colgan, G. T. Furuta, and C. T. Taylor, "Hypoxia and innate immunity: keeping up with the HIFsters," *Annual Review of Immunology*, vol. 38, no. 1, pp. 341–363, 2020.

[47] C. A. Corzo, T. Condamine, L. Lu et al., "HIF-1α regulates function and differentiation of myeloid-derived suppressor

cells in the tumor microenvironment," *Journal of Experimental Medicine*, vol. 207, no. 11, pp. 2439–2453, 2010.

[48] C. H. Chang, J. Qiu, D. O'Sullivan et al., "Metabolic competition in the tumor microenvironment is a driver of cancer progression," *Cell*, vol. 162, no. 6, pp. 1229–1241, 2015.

[49] P. Jiang, W. Du, and M. Wu, "Regulation of the pentose phosphate pathway in cancer," *Protein & Cell*, vol. 5, no. 8, pp. 592–602, 2014.

[50] P. Vizán, G. Alcarraz-Vizan, S. Diaz-Moralli, O. N. Solovjeva, W. M. Frederiks, and M. Cascante, "Modulation of pentose phosphate pathway during cell cycle progression in human colon adenocarcinoma cell line HT29," *International Journal of Cancer*, vol. 124, no. 12, pp. 2789–2796, 2009.

[51] W. Gao, Y. Xu, T. Chen et al., "Targeting oxidative pentose phosphate pathway prevents recurrence in mutant Kras colorectal carcinomas," *PLoS Biology*, vol. 17, no. 8, 2019.

[52] L. Abou-Elkacem, S. Arns, G. Brix et al., "Regorafenib inhibits growth, angiogenesis and metastasis in a highly aggressive orthotopic colon cancer model," *Molecular Cancer Therapeutics*, vol. 12, no. 7, pp. 1322–1331, 2013.

[53] T. Nagasaki, M. Hara, H. NakanisHi, H. TakaHasHi, M. Sato, and H. Takeyama, "Interleukin-6 released by colon cancer-associated fibroblasts is critical for tumour angiogenesis: anti-interleukin-6 receptor antibody suppressed angiogenesis and inhibited tumour-stroma interaction," *British Journal of Cancer*, vol. 110, no. 2, pp. 469–478, 2014.

[54] L. Dai, X. Cui, X. Zhang et al., "SARI inhibits angiogenesis and tumour growth of human colon cancer through directly targeting ceruloplasmin," *Nature Communications*, vol. 7, no. 1, 2016.

[55] A. Albini, A. Bruno, D. M. Noonan, and L. Mortara, "Contribution to tumor angiogenesis from innate immune cells within the tumor microenvironment: implications for immunotherapy," *Frontiers in Immunology*, vol. 9, p. 527, 2018.

[56] G. T. Motz and G. Coukos, "The parallel lives of angiogenesis and immunosuppression: cancer and other tales," *Nature Reviews Immunology*, vol. 11, no. 10, pp. 702–711, 2011.