



# Multi-Omics Characterization of a Glycerolipid Metabolism-Related Gene Enrichment Score in Colon Cancer

Zhiyu Wang<sup>1†</sup>, Zhuoqi Zhang<sup>2†</sup>, Ke Zhang<sup>3</sup>, Qiaoxia Zhou<sup>4</sup>, Sidong Chen<sup>4</sup>, Hao Zheng<sup>3</sup>, Guoqiang Wang<sup>4</sup>, Shangli Cai<sup>4</sup>, Fujing Wang<sup>3\*</sup> and Shenglong Li<sup>3\*</sup>

## OPEN ACCESS

### Edited by:

Ashok Kumar Pandurangan,  
B. S. Abdur Rahman Crescent Institute  
of Science and Technology, India

### Reviewed by:

Wenqing Cao,  
NYU Grossman School of Medicine,  
United States  
Qiong Zhang,  
Affiliated Hospital of Nantong  
University, China

### \*Correspondence:

Shenglong Li  
Lishenglong\_1982@163.com  
Fujing Wang  
wangfujing-hyd@163.com

<sup>†</sup>These authors have contributed  
equally to this work

### Specialty section:

This article was submitted to  
Gastrointestinal Cancers:  
Colorectal Cancer,  
a section of the journal  
Frontiers in Oncology

Received: 23 February 2022

Accepted: 04 April 2022

Published: 05 May 2022

### Citation:

Wang Z, Zhang Z, Zhang K,  
Zhou Q, Chen S, Zheng H,  
Wang G, Cai S, Wang F and Li S  
(2022) Multi-Omics Characterization  
of a Glycerolipid Metabolism-  
Related Gene Enrichment  
Score in Colon Cancer.  
*Front. Oncol.* 12:881953.  
doi: 10.3389/fonc.2022.881953

<sup>1</sup> Department of Medical Oncology, Hebei Key Laboratory of Cancer Radiotherapy and Chemotherapy, Affiliated Hospital of Hebei University, Baoding, China, <sup>2</sup> Department of Gastrointestinal Surgery, Affiliated Hospital of Hebei University, Baoding, China, <sup>3</sup> General Surgery Department, The Second Affiliated Hospital of Harbin Medical University, Harbin, China, <sup>4</sup> Medical Department, Burning Rock Biotech, Guangzhou, China

**Background:** Glycerolipid metabolism is involved in the genesis and progression of colon cancer. The current study aims at exploring the prognostic value and potential molecular mechanism of glycerolipid metabolism-related genes in colon cancer from the perspective of multi-omics.

**Methods:** Clinical information and mRNA expression data of patients with colon cancer were obtained from The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) databases. Single-sample gene set enrichment analysis (ssGSEA) was applied to calculate the glycerolipid metabolism-related gene enrichment score (GLMS). Univariable and multivariable Cox regression analyses were used to study the prognostic value of GLMS in TCGA-COAD and GSE39582 cohorts. The molecular mechanism of the prognostic factor was investigated via immune cell infiltration estimation and correlation analysis of cancer hallmark pathways. Single-cell transcriptomic dataset GSE146771 was used to identify the cell populations which glycerolipid metabolism targeted on.

**Results:** The GLMS was found to be associated with tumor location and consensus molecular types (CMSs) of colon cancer in TCGA-COAD cohort ( $P < 0.05$ ). Patients in the low-GLMS group exhibited poorer overall survival (OS) in TCGA cohort ( $P = 0.03$ ; HR, 0.63; 95% CI, 0.42–0.94), which was further validated in the GSE39582 dataset ( $P < 0.001$ ; HR, 0.57; 95% CI, 0.43–0.76). The association between the GLMS and OS remained significant in the multivariable analysis (TCGA cohort:  $P = 0.04$ ; HR, 0.64; 95% CI, 0.42–0.98; GSE39582 cohort:  $P < 0.001$ ; HR, 0.60; 95% CI, 0.45–0.80). The GLMS was positively correlated with cancer hallmark pathways including bile acid metabolism, xenobiotic metabolism, and peroxisome and negatively correlated with pathways such as interferon gamma response, allograft rejection, apoptosis, and inflammatory response ( $P < 0.05$ ). Increased immune infiltration and upregulated expression of immune checkpoints were observed in patients with lower GLMS ( $P < 0.05$ ). Single-cell datasets verified the different

distribution of GLMS in cell subsets, with significant enrichment of GLMS in malignant cells and Tprolif cells.

**Conclusion:** We demonstrated that GLMS was a potential independent prognostic factor for colon cancer. The GLMS was also correlated with several cancer hallmark pathways, as well as immune microenvironment.

**Keywords:** glycerolipid metabolism, colon cancer, prognostic signature, overall survival, multi-omics characterization

## INTRODUCTION

Colon cancer is the third most commonly diagnosed malignant tumor all over the world, accounting for almost 10% of all the cancer-related deaths in 2020 (1). Despite the rapid development in advanced surgical techniques and therapeutic strategies, the 5-year survival rate of late-stage colon cancer remains less than 30% due to metastasis and post-operation recurrence (2). The prognosis of colon cancer varies among different patients due to its high heterogeneity (3), since multifarious risk factors and multiple gene alterations are involved in its pathogenesis and progression, leading to distinct clinical pathological features and diverse responses to treatments (4). Nowadays, the establishment of public databases such as The Cancer Genome Atlas (TCGA) project and Gene Expression Omnibus (GEO) has enabled numerous studies exploring prognostic and predictive gene markers for colon cancer management; however, there is still a long way off identifying universally applied robust models.

Aberrant activation of lipid metabolism is an important hallmark for cancer cells (5), resulting in alterations of the lipid composition profile in different tumor tissues including lung (6), breast (7), and liver cancers (5). The lipid synthesis provides essential substrate for energy metabolism and ingredients for cell membrane construction during the proliferation process, suggesting an indispensable role of lipid metabolism pathways in the genesis and development of cancer. A subsequent study verified an upregulated expression of the key enzymes for lipid synthesis in colon cancer cells (8), followed by attempts of utilizing lipid synthesis inhibitors in colon cancer treatment (9), which suggested the vital role of lipid metabolism in colon cancer development. Among the various metabolites, triacylglycerol has been identified as a prognostic biomarker for colon cancer (10), indicating a possible contribution of glycerolipid metabolism in colon cancer progression. Therefore, genes related to glycerolipid metabolism may serve as potential prognostic markers of colon cancer, and a better understanding of cancer-related genes involved in glycerolipid metabolism may provide a reference for the selection of treatment strategies in colon cancer patients. However, relevant studies regarding glycerolipid metabolism and colon cancer development are still limited.

Thus, we conducted this study based on public databases including TCGA and GEO, proposing a prognostic model of colon cancer based on glycerolipid metabolism genes, and further explored the molecular mechanism of the glycerolipid metabolism-related genes involved in the development of colon cancer from the perspective of multi-omics.

## METHODS AND MATERIALS

### Data Acquisition

Three datasets were involved in this study. Clinical information and survival data of COAD patients from The Cancer Genome Atlas (TCGA-COAD) were downloaded from the UCSC XENA database (<https://xenabrowser.net/>), as well as the gene expression matrix (HTSeq-FPKM), somatic mutation, copy number variation (CNV, gene-level), and methylation (Illumina Human Methylation 450) sequencing results. Clinical information and survival data of colon cancer patients in the GSE39582 dataset were retrieved from the Gene Expression Omnibus (GEO) dataset (<http://www.ncbi.nlm.nih.gov/geo>). Microarray RAW “CEL” data files of colon cancer patients in the GSE39582 dataset were also downloaded, and normalization was conducted through the Robust Multiarray Average (RMA) method in R package “Affy.” A single-cell transcriptomic dataset of COAD patients in GSE146771 (Smart-seq2) was also obtained from the GEO database. The KEGG\_GLYCEROLIPID\_METABOLISM gene set and 50 cancer hallmark gene sets were downloaded from the Molecular Signature Database v. 7.4 (MSigDB). A total of 49 glycerolipid metabolism-related genes were obtained from the KEGG\_GLYCEROLIPID\_METABOLISM pathway.

### Clinical Dataset Preprocessing

The following preprocessing steps were applied to both TCGA-COAD and GSE39582 datasets. (1) Patients who had no clinical information were excluded from the datasets. (2) Patients whose survival time was 0 or survival status was unknown were excluded from the analysis. (3) The type of gene ID in the gene expression matrix of both datasets was synchronized to Gene Symbol. When multiple gene IDs were synchronized to one Gene Symbol, the median value was selected to represent the expression level.

### Multi-Omics Characterization of Glycerolipid Metabolism in Colon Cancer

The 49 glycerolipid metabolism-related genes were mapped to the gene expression matrix and gene methylation matrix. The gene expression level and methylation level of these genes were compared between tumor and normal tissues by the Mann-Whitney test, and  $P$ -value  $< 0.05$  was considered significant. The CNV amplification and deletion frequency, as well as the somatic single-nucleotide variation, of the glycerolipid metabolism-related genes were calculated from relative sequence results of tumor samples.

## Single-Sample Gene Set Enrichment Analysis

Single-sample gene set enrichment analysis (ssGSEA) is an unsupervised single-sample gene set enrichment analysis method (11). ssGSEA calculates the enrichment score of each sample by the cumulative distribution function difference between gene expression ranks inside the settled gene set and outside. The enrichment score was further normalized by the range of total gene expression matrix for all genes and samples (11). The ssGSEA analysis method included in R package “GSVA” was used to calculate the KEGG\_GLYCEROLIPID\_METABOLISM gene set enrichment score (GLMS) based on the gene expression matrix of TCGA-COAD and GSE39582 (12). Samples were categorized into high- and low-GLMS groups based on the median of GLMS for both TCGA-COAD and GSE39582 datasets.

## Association Between Clinical Characteristics and GLMS

The consensus molecular subtypes (CMSs) of TCGA-COAD datasets were estimated *via* the R package “CMScaller” (13). The clinical characteristics of TCGA-COAD datasets such as age, sex, BMI (classified into high and low groups based on median), stage, microsatellite instability, venous invasion, and tumor location were included for assessing the association between the GLMS and clinical characteristics, together with the CMS subtypes.

## Pathway Correlation Analysis

The ssGSEA analysis method was also applied to calculate the 50-cancer hallmark gene set enrichment score based on the gene expression matrix of TCGA-COAD dataset. Spearman correlation analysis was used to quantify the correlation between the GLMS and 50 cancer hallmark gene sets based on their ssGSEA scores of TCGA-COAD samples. The correlation between every single glycerolipid metabolism-related gene in the KEGG gene set and 50 cancer hallmark gene sets was also calculated *via* R package “WGCNA” corAndPvalue function based on gene expression and the hallmark ssGSEA score (14).

## Immune Microenvironment Analysis

The proportions of 64 cell types in the tumor microenvironment of TCGA-COAD tumor samples were estimated by R package “xCell” (15). The expressions of 14 immune checkpoints were compared between high- and low-GLMS groups including CD274, CD276, CD40, CTLA4, HAVCR2, ICOS, IDO1, LAG3, PDCD1, TIGIT, TNFRSF18, TNFRSF4, TNFRSF9, and VTCN1.

## Potential Drug Sensitivity Analysis

The pRRophetic algorithm was originally developed for drug response prediction based on gene expression microarray data and the half-maximal inhibitory concentration (IC50) of 138-drug response data in almost 700 cell lines from the Cancer Genome Project (CGP) database (16, 17). The drug sensitivity of those drugs in TCGA-COAD dataset, as indicated by IC50 prediction, was calculated based on the R package “pRRophetic” and gene expression matrix of these samples. The difference of estimated

drug sensitivity between high- and low-GLMS groups was evaluated for every drug *via* the Wilcoxon test, and *P*-value <0.05 was considered significant. Subsequently, the correlation between every single gene in the KEGG\_GLYCEROLIPID\_METABOLISM gene set and drug sensitivity was calculated *via* R package “WGCNA” corAndPvalue function based on expression and IC50 (14). Furthermore, potential drug response-related genes were selected according to *P*-value and correlation coefficient (*P*-value < 0.05, |cor| >0.5).

## Single-Cell Transcriptomic Analysis

GLMS of single-cell transcriptomic dataset GSE146771 was calculated by the ssGSEA method as well. The cell annotation information corresponding to this dataset was downloaded from the Tumor Immune Single-cell Hub (TISCH) database (<http://tisch.comp-genomics.org>). The GLMS of every single cell was shown by uniform manifold approximation and projection (UMAP) plots at two resolutions: “Celltype malignancy” and “Celltype major lineage.” The Kruskal–Wallis test was applied to evaluate whether GLMS varied among different cell types, and *P*-value <0.05 was considered significant. Finally, the cell subpopulations on which glycerolipid metabolism affected were estimated by the log2FoldChange (log2FC) value of each cell types *via* “one vs. rest” calculation.

## Statistical Analysis

Data preprocessing and R package analyses were completed in R (version 4.1). Statistical analysis and figures were completed in GraphPad Prism (version 9). Continuous variables were displayed as median (range), and categorical variables were displayed as number (percentage). The Wilcoxon test was applied to calculate the significance of the difference in continuous variables between two groups, and the chi-square test or Fisher’s exact test was used to calculate the significance of the difference in categorical variables between groups. Kaplan–Meier analysis was performed to estimate the overall survival (OS) rates, and the difference of survival curves between high- and low-GLMS groups was compared by the log-rank test through R package “survival.” Univariable and multivariable Cox proportional hazards regression analyses were conducted *via* R package “survival” coxph function to determine whether the GLMS could be an independent prognostic factor for colon cancer. Statistical significance was considered as *P*-value <0.05.

## RESULTS

### Epigenetic and Genomic Characteristics of Glycerolipid Metabolism-Related Genes in Colon Cancer

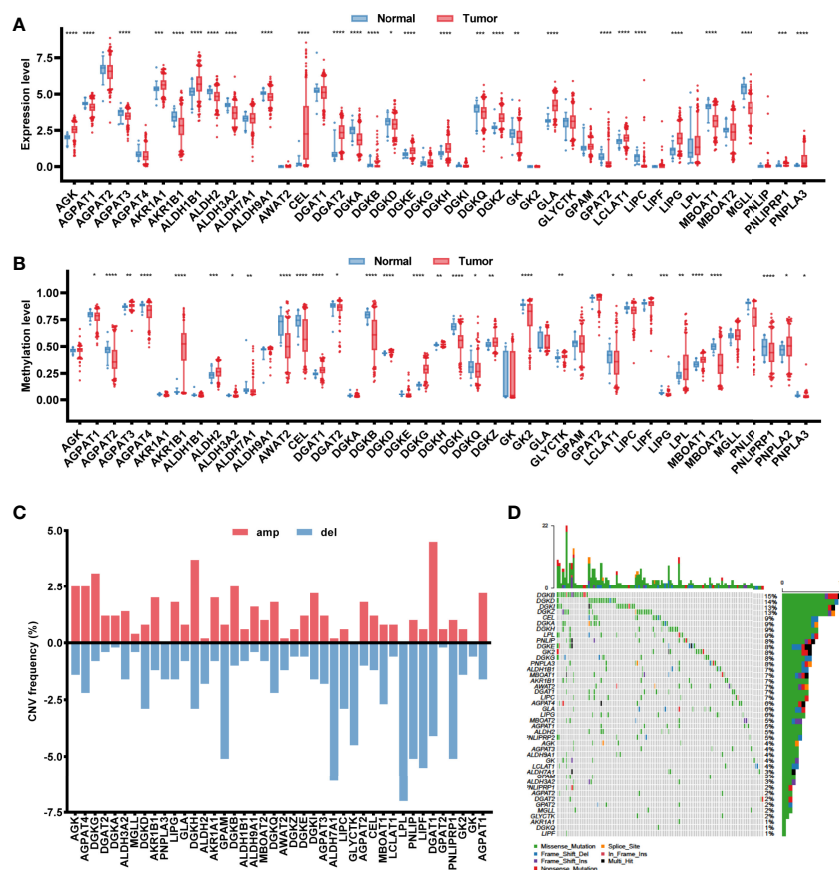
Based on the 430 cancer samples and 39 normal samples in TCGA-COAD dataset, the expression levels of 43 glycerolipid metabolism-related genes were available out of 49 genes from the KEGG\_GLYCEROLIPID\_METABOLISM gene set, while the methylation information was provided for 42 genes. The baseline characteristics of the samples in TCGA-COAD dataset

are described in **Supplementary Table**. Differences in expression and methylation levels were compared between tumor and normal tissues (**Figure 1A**), which identified nine genes with a negative correlation between expression and methylation levels, including *CEL*, *DGAT2*, *LCLAT1*, and *LIPG* (hypomethylated and upregulated in tumor tissues), and *AGPAT3*, *AKR1B1*, *ALDH2*, *DGKD*, and *MBOAT1* (hypermethylated and downregulated in tumor tissues), indicating an inhibitory function of methylation in the mRNA expression of those genes. However, the methylation of six genes was positively correlated with expression levels, including *DGKH* and *DGKZ* (hypermethylated and upregulated in tumors), and *AGPAT1*, *DGKB*, *DGKQ*, and *LIPC* (hypomethylated and downregulated in tumors), suggesting the potential regulatory mechanisms of mRNA expression in those genes beyond methylation. The copy number variations and gene mutations were also analyzed based on cancer samples. Copy number amplification appeared in *AGPAT1*, *AKR1B1*, *DGKB*, *DGAT2*, and *DGKH*, while *AGPAT3*, *ALDH2*, *DGKD*, *MBOAT1*, and *LIPC* exhibited significant copy number deletions (**Figure 1**). A total of 120 (27.9%) tumor samples harbored mutations in glycerolipid metabolism-related genes, with two genes identified

as high frequently mutated, including *DGKB* (15%) and *DGKD* (14%) (**Figure 1**).

## Glycerolipid Metabolism Score as a Prognostic Indicator for Colon Cancer

Each sample in TCGA-COAD dataset was assigned with a glycerolipid metabolism-related gene enrichment score (GLMS) according to the single-sample gene set enrichment analysis (ssGSEA) algorithm and divided into high- and low-GLMS groups by the cutoff of the median value. We explored whether the GLMS could predict the prognosis of colon cancer in TCGA-COAD dataset. Kaplan–Meier analysis showed that the patients with higher GLMS exhibited improved overall survival (OS) compared with those with lower GLMS (log-rank  $P = 0.03$ ; HR, 0.63; 95% CI, 0.42–0.94, **Figure 2**). Univariable analysis was also conducted in other clinical factors, including age, sex, stage, tumor location, MSI status, and CMS, among which only clinical stage and GLMS exhibited statistical significance ( $P < 0.05$ , **Figure 2**). These clinical factors along with GLMS were included in the multivariable Cox regression, and the association between GLMS and OS remained significant ( $P = 0.04$ ; HR, 0.64; 95% CI, 0.42–



**FIGURE 1** | Characterization of the glycerolipid metabolism genes in colon cancer patients. **(A, B)** The mRNA expression **(A)** and methylation **(B)** of the glycerolipid metabolism-related genes between tumor and normal tissues from TCGA database. **(C, D)** Copy number variation (CNV) frequency showing gene amplification (amp) and deletion (del) **(C)** and mutation oncoprint **(D)** of 42 glycerolipid metabolism-related genes in 430 tumor tissues from TCGA-COAD cohorts. \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ ; and \*\*\*\* $P < 0.0001$  by the Mann–Whitney test.

0.98, **Figure 2**), suggesting that the GLMS was a potential independent prognostic factor of colon cancer.

To further verify the robustness of the GLMS in predicting the prognosis of colon cancer, another dataset (GSE39582) was used as a validation cohort. A total of 562 tumor samples were included in the GSE39582 dataset, with the baselined characteristics described in **Supplementary Table**. With the same algorithm of GLMS, patients in the high-GLMS group also had a better OS (log-rank  $P < 0.001$ ; HR, 0.57, 95% CI, 0.43–0.76, **Figure 2**). Similar results were also observed in the univariable (**Figure 2**) and multivariable Cox regression (**Figure 2**). Univariable analysis has identified tumor stage and GLMS that were associated with OS (**Figure 2**). The GLMS was further confirmed to be an independent prognostic factor of OS in the multivariable analysis ( $P < 0.001$ ; HR, 0.60; 95% CI, 0.45–0.80, **Figure 2**). Taken together, these results suggested that the GLMS was an independent prognostic factor for colon cancer.

## Association Between GLMS and Clinical Features

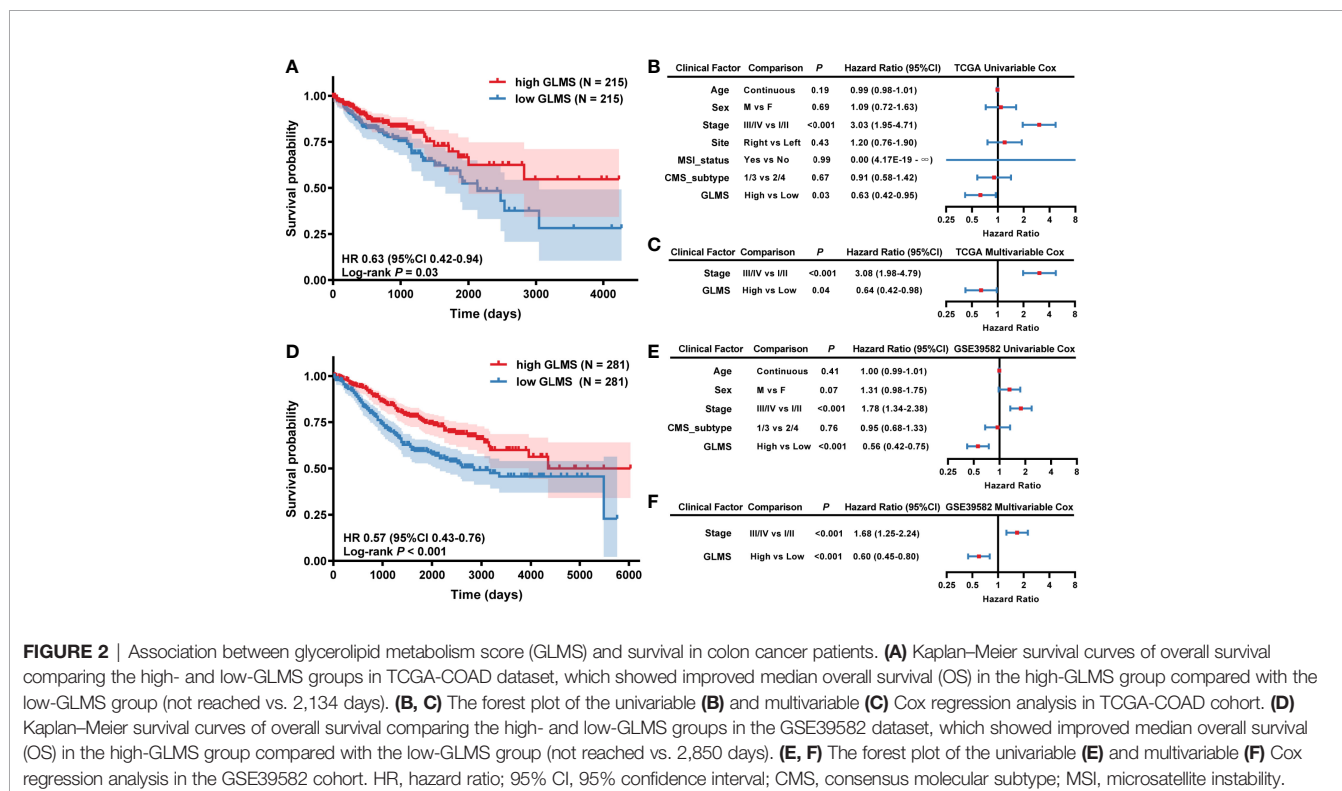
The potential association between GLMS and clinical features was further analyzed including age, sex, body mass index (BMI), tumor stage by American Joint Committee on Cancer (AJCC) TNM Classification, microsatellite status, vascular invasion status, tumor location, and consensus molecular subtypes (CMS) (**Figure 3**). Results showed that left colonic carcinomas exhibited significantly higher GLMS compared with right ones ( $P = 0.008$ , **Figure 3**), suggesting the capacity of the GLMS in differentiating the tumor location. Besides, CMS subtypes among tumors have revealed significant differences in GLMS distribution ( $P < 0.05$ , **Figure 3**),

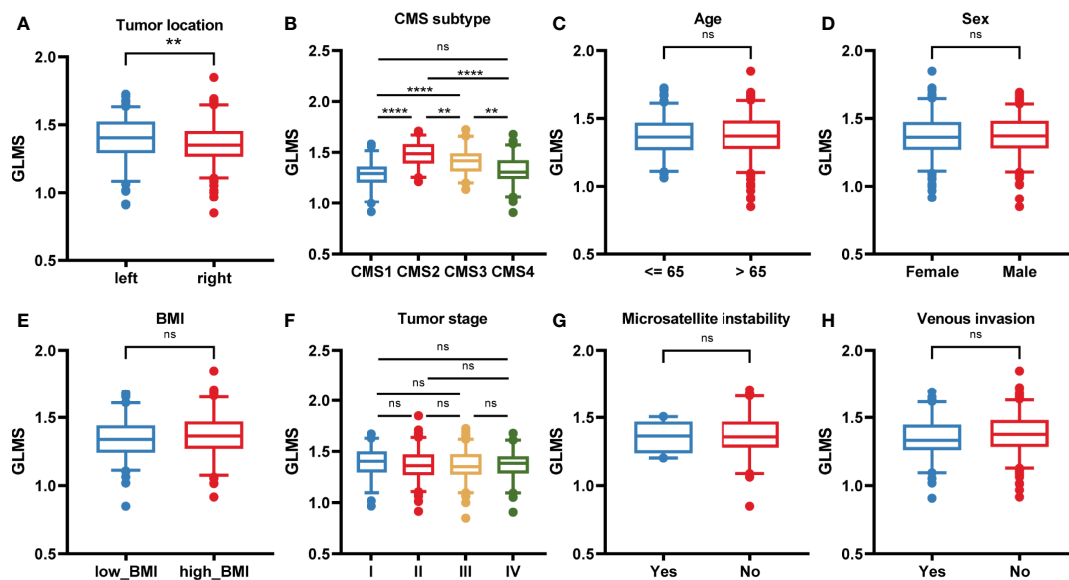
while patients in different groups by age, sex, BMI, tumor stage, microsatellite instability status, and venous invasion status were observed to have a similar distribution of the GLMS (**Figure 3C**).

## Underlying Mechanism of Glycerolipid Metabolism

To explore the potential pathways glycerolipid metabolism was involved in, the correlation between the GLMS and cancer hallmarks was analyzed. A significant positive association was observed between GLMS and pathways including bile acid metabolism, xenobiotic metabolism, and peroxisome, while pathways such as interferon gamma response, allograft rejection, apoptosis, and inflammatory response were found to be negatively correlated with the GLMS (**Figure 4**). The 42 genes involved in the GLMS were analyzed as well. Samples with an upregulated expression of AGPAT4, AKR1B1, and DGKI showed positive enrichment in pathways including cholesterol homeostasis, bile acid metabolism, inflammatory response, and NOTCH signaling (**Figure 4**), while expressions of AGPAT4, AKR1B1, DGKI, and MGLL were negatively associated with pathways such as protein secretion and IL2-STAT5 signaling (**Figure 4**), suggesting the core genes and pathways involved in glycerolipid metabolism.

The immune microenvironment was revealed by utilizing R package “xCell” to estimate the infiltration of the proportion of different immune cells, which was further compared between high- and low-GLMS groups. As presented in **Figure 4**, a total of 33 immune cells or stromal cells exhibited a significant infiltration difference between samples with high- and low-GLMS. Most of the immune-promoting cells, including B cells,





**FIGURE 3** | The association between the glycerolipid metabolism score and clinical features. (A–H) The glycerolipid metabolism score (GLMS) distribution in patients with different clinicopathologic features including tumor location (A), and consensus molecular subtype (CMS) (B), age (C), sex (D), body mass index (BMI) (E), tumor stage (F), microsatellite instability status (G), and venous invasion status (H). ns (non-significant)  $P > 0.05$ ,  $**P < 0.01$  and  $****P < 0.0001$  by the Wilcoxon test.

macrophages, activated dendritic cells (aDC), classical dendritic cells (cDC), and immature dendritic cells (iDC), accounted for higher proportions in the low-GLMS group, while cells negatively regulating immune response such as NKT cells were significantly infiltrated in high-GLMS group samples.

Furthermore, the expression of 14 genes involved in immune checkpoints was compared between the two groups. As a result, a total of 12 immune checkpoint molecules were upregulated in low-GLMS group samples (Figure 4), suggesting a potential benefit of immune therapies in patients with low GLMS.

## Drug Sensitivity

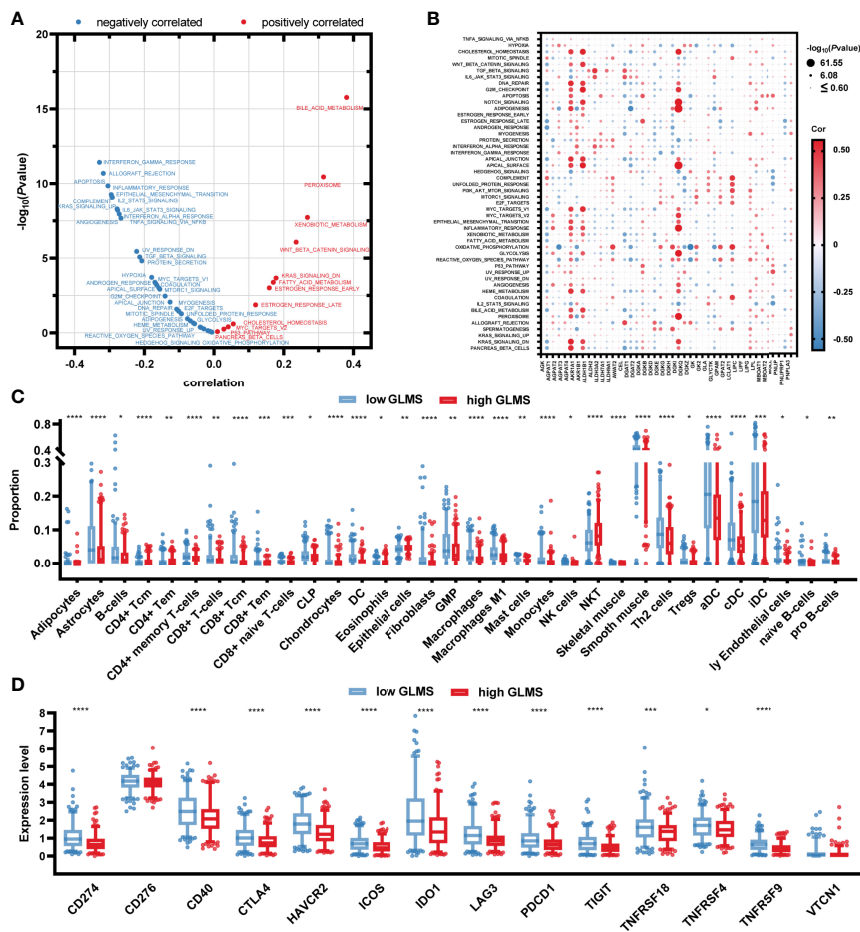
Considering that lower GLMS was associated with poor prognosis, we conducted further analysis exploring the relationship between GLMS and drug sensitivity. Based on the Genomics of Drug Sensitivity in Cancer (GDSC) database, the half-maximal inhibitory concentration (IC<sub>50</sub>) information was acquired to predict the treatment response. Among the 96 drugs which exhibited a significant response difference between groups with low and high GLMS, those that were currently used in the treatment of colon cancer were paid extra attention. Low-GLMS group samples were observed to be more sensitive to chemotherapeutics including cisplatin (Figure 5), gemcitabine (Figure 5), and camptothecin (Figure 5), which had analogues of oxaliplatin, capecitabine, and irinotecan, respectively. Besides, ponatinib (Figure 5) as an anti-angiogenesis inhibitor targeting vascular endothelial growth factor receptors (VEGFRs) and epidermal growth factor receptor (EGFR) tyrosine kinase inhibitors (TKIs) including gefitinib (Figure 5) and afatinib (Figure 5) also exhibited significant increased sensitivity associated with low-GLMS group samples, suggesting potential

benefits utilizing these therapies in patients with lower GLMS who had poorer survival.

In addition, further analysis also explored the association between IC<sub>50</sub> of the drugs and 42-glycerolipid metabolism-related gene expression. According to  $P$ -value  $< 0.05$  and the absolute value of correlation  $> 0.5$ , a total of 14 genes were identified including AKR1B1, DGKI, AGPAT2, GPAM, LCLAT, DGKZ, MGLL, DGKQ, AGPAT4, DGAT1, AKR1A1, DGKD, DGKH, and AGK (Figure 5), suggesting that these genes might play a crucial role in drug response.

## Origination of Glycerolipid Metabolic Disorders

The single-cell sequencing dataset GSE146771 was introduced in this current study to investigate the origination of glycerolipid metabolic disorders. Annotations of cell samples were obtained from the TISCH database to identify the subtype of the cells, which were presented by a rough classification of “malignant-immune-stromal cells” (Figure 6) and detailed subsets (Figure 6). Each cell sample was assigned with a GLMS according to the same ssGSEA algorithm (Figure 6). Results showed that the scores were generally higher in malignant cells and stromal cells, as well as in CD4Tconv, CD8Tex, and plasma cells when it came to the sophisticated categories. Further statistical analysis revealed a significant difference of the score distribution among various cell types ( $P < 2.2 \times 10^{-16}$ ) (Figure 6). Enrichment of the scores was analyzed based on the log<sub>2</sub>FC value of the scores in each cell subset, which showed significant enrichment in malignant cells and Tprolif cells (Figure 6), suggesting potential targets and originations of the glycerolipid metabolic disorders.



**FIGURE 4 |** Potential molecular mechanisms of glycerolipid metabolism. **(A)** Correlation between glycerolipid metabolism score (GLMS) and ssGSEA enrichment scores of 50 cancer hallmark pathways in TCGA-COAD cohort. **(B)** Correlation between the expression level of 42 glycerolipid metabolism-related genes and ssGSEA enrichment scores of 50 cancer hallmark pathways. **(C)** The abundance of each tumor microenvironment (TME)-infiltrating cell between high- and low-GLMS groups in TCGA-COAD cohort. **(D)** The difference of 14-immune checkpoint mRNA expression between high- and low-GLMS groups in TCGA-COAD cohort. \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$  and \*\*\*\* $P < 0.0001$  by the Mann–Whitney test.

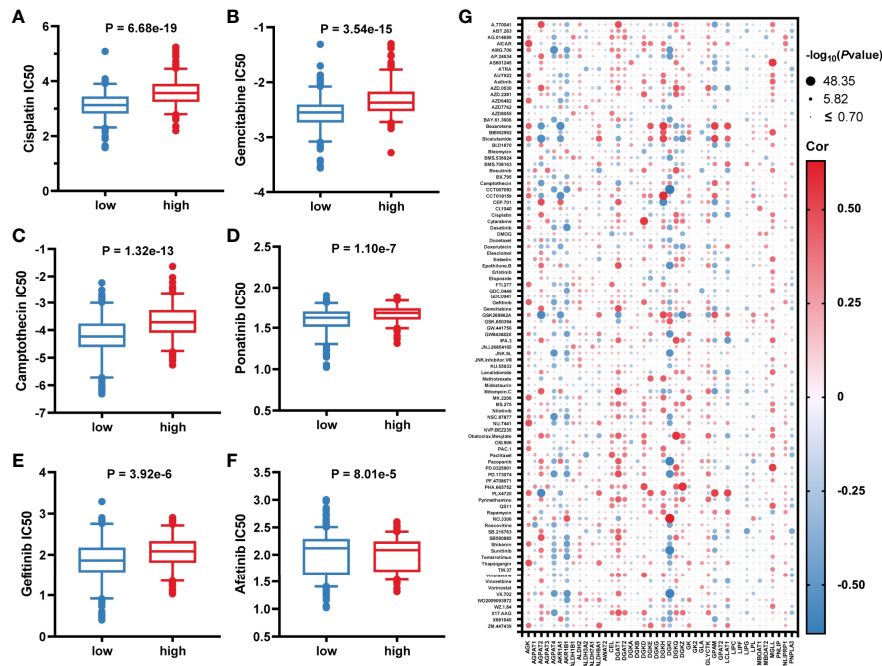
## DISCUSSION

Over the past decade or so, researchers have been extensively working on cancer metabolism as it has been considered as one of the cancer hallmarks (18, 19). Lipid synthesis and metabolism are known to get involved in cell proliferation (19–22). Additionally, it is reported that lipidomic profiling is significantly different between tumor tissues and healthy tissues (5–7). However, even though colon cancer patients with obesity and metabolic dysfunction have worse survival (23), there is still debate on whether lipid metabolism and related genes contribute to the disease progression of colorectal cancer (CRC) patients (24–26).

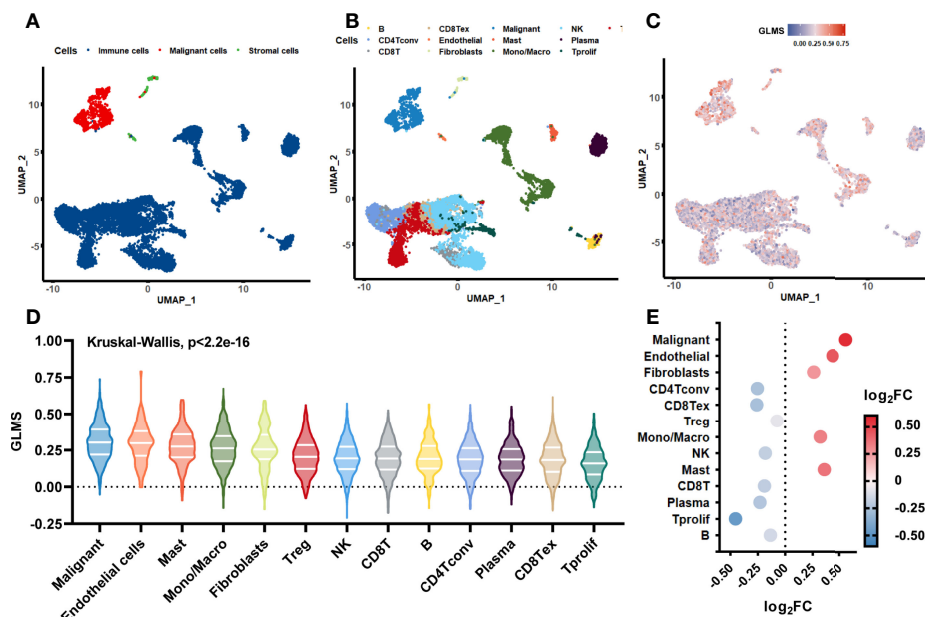
Glycerolipid is one of the most important lipid molecules, but little is known about glycerolipid metabolism in colon cancer (27). To address these questions, we developed a prognostic model GLMS for colon cancer based on glycerolipid metabolism-related genes *via* the ssGSEA algorithm and proved

that the GLMS was an independent clinical prognostic factor. We further investigated the molecular mechanism behind this model by evaluating the cross talk between GLMS and cancer hallmark pathways, as well as the immune microenvironment. Surprisingly, we observed that the glycerolipid metabolism pathway was highly correlated with several cancer hallmark pathways and the immune microenvironment. Last but not least, we identified that malignant and Tprolif cells were the most possible cell populations which glycerolipid metabolism targeted on.

Over the past few years, many prognostic models have been established based on gene signature or other biological molecules from different signaling pathways of colon cancer (28). A majority part of these prognostic models was based on immune-relevant signaling pathways. For example, Zhou et al. constructed a tumor microenvironment risk score (TMRS) with 100 tumor microenvironment-related genes based on the GEO database using the Lasso Cox regression model. This model was an effective tool for survival prediction of not only colon cancer



**FIGURE 5** | Association between glycerolipid metabolism score and drug sensitivity. (A–F) The correlation between glycerolipid metabolism scores (GLMS) of cell samples and estimated half-maximal inhibitory concentration (IC50) value of drugs including cisplatin (A), gemcitabine (B), camptothecin (C), ponatinib (D), gefitinib (E), and afatinib (F). (G) The correlation between glycerolipid metabolism-related genes and estimated IC50 value of 96 drugs.



**FIGURE 6** | Glycerolipid metabolism score in single-cell datasets. (A, B) The distribution of the cell samples by uniform manifold approximation and projection (UMAP) presented in rough classification of “malignant-immune-stromal cells” (A) and detailed subsets (B). (C) The UMAP distribution of the cells presented with glycerolipid metabolism scores (GLMS) of each sample. (D) Distribution difference of the GLMS in 13 subsets of cells with statistical analysis by the Kruskal–Wallis test. (E) GLMS enrichment by log2 fold change (log2FC) value in 13 subsets of cells.



but also gastric cancer (29). Xu et al. developed a prognostic model for colon cancer with 11 immune-related genes based on single-cell sequencing data (30). Intriguingly, one recent study identified that one single immune-related gene, CXCL11, could be an effective independent prognostic biomarker for colon cancer patients and upregulation of CXCL11 expression was correlated with a PD-L1 high expression level (31). For miRNA-constructed models, Zhang et al. used Lasso Cox regression to construct a classifier of 6 miRNAs which could predict recurrence in patients with stage II colon cancer and whether patients would benefit from postoperative adjuvant therapy (32).

Surprisingly, not many prognostic models based on lipid synthesis- and metabolism-related genes have been developed so far. Jiang et al. constructed a prognostic model of 8 lipid metabolism-related genes for colorectal cancer based on TCGA and GEO databases (33). A latest research investigated the significance of lipid metabolism for the prognostic of CRC based on multi-omics. In their research, a Cox regression analysis of lipid metabolites was utilized to identify the key lipid metabolites and construct the prognostic model. Furthermore, they distinguished the glycerophospholipid metabolism-related hub genes and developed a prognostic signature through Cox regression and LASSO regression analysis (34). Moreover, a research team developed and validated a robust prognostic model for colorectal cancer based on a lipid species molecule signature from lipidomics quantification analysis instead of gene or miRNA (10). More importantly, one major component of this robust CRC-specific lipid signature model was glycerolipid which gave us a hint that glycerolipid metabolism might be very important for colon cancer prognostic prediction. To date, our study provided the first colon cancer prognostic factor based on the glycerolipid metabolism-related gene enrichment score.

Pathway correlation analysis results suggested that glycerolipid metabolism was significantly positively correlated with bile acid metabolism pathway, peroxisome pathway, and xenobiotic metabolism pathway. In terms of the peroxisome pathway, peroxisome proliferator-activated receptor  $\gamma$  (PPAR $\gamma$ ) is a nuclear receptor that regulates the expression of genes related to lipid metabolism and known to have anticancer activity (35). It is reported that tumor tissues of colon cancer patients expressing PPAR $\gamma$  have better prognosis, which supports our prognostic factor and correlation analysis (36). A high plasma bile acid level is reported to be positively correlated with colon cancer risk. This result indicated that a high load of bile acid due to inadequate metabolism might be promotive for colon cancer tumorigenesis and disease progression which was consistent with our analyses as well (37). On the other hand, analysis results indicated that the interferon- $\gamma$  (IFN- $\gamma$ ) response pathway was negatively correlated with glycerolipid metabolism in colon cancer. The IFN- $\gamma$  signaling pathway plays a critical role in anticancer immunity mainly through anti-proliferation (38), apoptosis (39), necrosis (40), and ferroptosis (41); however, it is also reported that IFN- $\gamma$  induces the expression of immune checkpoints which may lead to an immune-suppression microenvironment, and this was consistent with our further analysis about the immune microenvironment (42–44).

Taken together, our results suggested that the IFN- $\gamma$  response pathway might mainly contribute to the immune-suppression microenvironment in glycerolipid metabolism-activated colon cancer, other than antitumor effect.

The tumor immune microenvironment is critical to tumorigenesis, disease progression, and treatment outcomes in colon cancer (45, 46). Researchers generated an Immunoscore (IS) based on the density of CD3+ and CD8+ T-cell effectors within the tumor, as well as its invasive margin, and developed a colon cancer recurrence risk model accordingly (47, 48). However, due to the complexity of the immune system, all different types of immune cells could contribute to the immune microenvironment. Hence, in our study, we evaluated the infiltration of 64 cell types and compared their proportion between high- and low-GLMS groups. We observed that the infiltration proportions of B cells, macrophages, aDC, cDC, and iDC were significantly higher in the low-GLMS group; in contrast, the infiltration proportion of NKT cells was significantly higher in the high-GLMS group. It is previously reported that fatty acid metabolism, which is an important step of glycerolipid metabolism, controls the immune suppressive phenotype of tumor-associated macrophages in the colon cancer cell line as well as CRC patients (49). Dendritic cells are a group of antigen-presenting cells. Researchers have reported that accumulation of lipids, especially triacylglycerol, affects the function of dendritic cells, which leads to ineffective tumor-associated antigens presenting in the colon cancer mouse model (50). Together with this result, our analyses give us a hint that elevated DC infiltration in the low-GLMS group may be associated with glycerolipid accumulation, which might also affect the subsequent antigen-presenting process. Last but not least, the expression level of different checkpoints in the low- and high-GLMS groups showed that the expression of 12 immune checkpoints was significantly increased in the low-GLMS group which may lead to an immune-suppression microenvironment. These results supported the notion that lipid metabolism not only plays a role in the tumorigenesis and disease progression of cancer cells but also is involved in the responses of tumor recruited immune, stromal cell, and regulating immune microenvironment (21). More importantly, our single-cell transcriptomics analyses indicated that glycerolipid metabolism may affect the biological activity of malignant cells and Tprolif cells in colon cancer. It is also reported that, in hepatocellular carcinoma (HCC), palmitic acyl-containing glycerophospholipids inhibit HCC cell growth and metastatic abilities, which was consistent with our analyses (51).

There are some limitations of the study. Firstly, we developed this model based on the gene set enrichment score of the total 49 KEGG\_GLYCEROLIPID\_METABOLISM genes, which may cause over-interpretation about the results. Secondly, the conclusions of our study are mainly based on *in silico* bioinformatics analyses of two retrospective cohorts. The retrospective nature and sample sizes may limit the interpretation of the conclusion of the present study. Further *in vitro* or *in vivo* validation studies are needed and prospective studies are warranted.

In conclusion, we constructed an independent colon cancer prognostic factor based on the glycerolipid metabolism gene enrichment score. Our prognostic factor, GLMS, was correlated

with important signaling pathways in colon cancer, as well as the immune microenvironment. In addition, GLMS mainly targeted on malignant cells in colon cancer. Thus, our findings provide a new prognostic factor for colon cancer and the possible molecular mechanism behind it.

## DATA AVAILABILITY STATEMENT

The datasets generated and analyzed in the present study are available in the public data repository, TCGA-COAD: <https://portal.gdc.cancer.gov/repository>, and GEO: <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE39582>.

## REFERENCES

- Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: Globocan Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* (2021) 71(3):209–49. doi: 10.3322/caac.21660
- Kennedy RD, Bylesjo M, Kerr P, Davison T, Black JM, Kay EW, et al. Development and Independent Validation of a Prognostic Assay for Stage II Colon Cancer Using Formalin-Fixed Paraffin-Embedded Tissue. *J Clin Oncol* (2011) 29(35):4620–6. doi: 10.1200/JCO.2011.35.4498
- Sagaert X, Vanstapel A, Verbeek S. Tumor Heterogeneity in Colorectal Cancer: What Do We Know So Far? *Pathobiology* (2018) 85(1-2):72–84. doi: 10.1159/000486721
- Sasaki N, Clevers H. Studying Cellular Heterogeneity and Drug Sensitivity in Colorectal Cancer Using Organoid Technology. *Curr Opin Genet Dev* (2018) 52:117–22. doi: 10.1016/j.gde.2018.09.001
- Vriens K, Christen S, Parik S, Broekaert D, Yoshinaga K, Talebi A, et al. Evidence for an Alternative Fatty Acid Desaturation Pathway Increasing Cancer Plasticity. *Nature* (2019) 566(7744):403–6. doi: 10.1038/s41586-019-0904-1
- Cifkova E, Holcapek M, Lisa M, Vrana D, Gatek J, Melichar B. Determination of Lipidomic Differences Between Human Breast Cancer and Surrounding Normal Tissues Using HILIC-HPLC/ESI-MS and Multivariate Data Analysis. *Anal Bioanal Chem* (2015) 407(3):991–1002. doi: 10.1007/s00216-014-8272-z
- Eggers LF, Muller J, Marella C, Scholz V, Watz H, Kugler C, et al. Lipidomes of Lung Cancer and Tumour-Free Lung Tissues Reveal Distinct Molecular Signatures for Cancer Differentiation, Age, Inflammation, and Pulmonary Emphysema. *Sci Rep* (2017) 7(1):11087. doi: 10.1038/s41598-017-11339-1
- Zaytseva YY, Harris JW, Mitov MI, Kim JT, Butterfield DA, Lee EY, et al. Increased Expression of Fatty Acid Synthase Provides a Survival Advantage to Colorectal Cancer Cells Via Upregulation of Cellular Respiration. *Oncotarget* (2015) 6(22):18891–904. doi: 10.18632/oncotarget.3783
- Peck B, Schug ZT, Zhang Q, Dankworth B, Jones DT, Smethurst E, et al. Inhibition of Fatty Acid Desaturation Is Detrimental to Cancer Cell Survival in Metabolically Compromised Environments. *Cancer Metab* (2016) 4:6. doi: 10.1186/s40170-016-0146-8
- Ecker J, Benedetti E, Kindt ASD, Horing M, Perl M, Machmuller AC, et al. The Colorectal Cancer Lipidome: Identification of a Robust Tumor-Specific Lipid Species Signature. *Gastroenterology* (2021) 161(3):910–23.e19. doi: 10.1053/j.gastro.2021.05.009
- Barbie DA, Tamayo P, Boehm JS, Kim SY, Moody SE, Dunn IF, et al. Systematic RNA Interference Reveals That Oncogenic Kras-Driven Cancers Require Tbk1. *Nature* (2009) 462(7269):108–12. doi: 10.1038/nature08460
- Hänzelmann S, Castelo R, Guinney J. Gsva: Gene Set Variation Analysis for Microarray and RNA-Seq Data. *BMC Bioinf* (2013) 14(1):7. doi: 10.1186/1471-2105-14-7
- Eide PW, Bruun J, Lothe RA, Sveen A. Cmscaller: An R Package for Consensus Molecular Subtyping of Colorectal Cancer Pre-Clinical Models. *Sci Rep* (2017) 7(1):16618. doi: 10.1038/s41598-017-16747-x
- 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

## AUTHOR CONTRIBUTIONS

SL and FW designed this work. ZW and ZZ integrated and analyzed the data. KZ, QZ, and SdC wrote this manuscript. HZ, GW, and SIC edited and revised the manuscript. All authors contributed to the article and approved the submitted version.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fonc.2022.881953/full#supplementary-material>

- Aran D, Hu Z, Butte AJ. Xcell: Digitally Portraying the Tissue Cellular Heterogeneity Landscape. *Genome Biol* (2017) 18(1):220. doi: 10.1186/s13059-017-1349-1
- Geeleher P, Cox NJ, Huang RS. Clinical Drug Response Can Be Predicted Using Baseline Gene Expression Levels and *In Vitro* Drug Sensitivity in Cell Lines. *Genome Biol* (2014) 15(3):R47. doi: 10.1186/gb-2014-15-3-r47
- Geeleher P, Cox N, Huang RS. Prorhetic: An R Package for Prediction of Clinical Chemotherapeutic Response From Tumor Gene Expression Levels. *PLoS One* (2014) 9(9):e107468. doi: 10.1371/journal.pone.0107468
- Hanahan D, Weinberg RA. Hallmarks of Cancer: The Next Generation. *Cell* (2011) 144(5):646–74. doi: 10.1016/j.cell.2011.02.013
- Zhu J, Thompson CB. Metabolic Regulation of Cell Growth and Proliferation. *Nat Rev Mol Cell Biol* (2019) 20(7):436–50. doi: 10.1038/s41580-019-0123-5
- Hoy AJ, Nagarajan SR, Butler LM. Tumour Fatty Acid Metabolism in the Context of Therapy Resistance and Obesity. *Nat Rev Cancer* (2021) 21(12):753–66. doi: 10.1038/s41568-021-00388-4
- Corn KC, Windham MA, Rafat M. Lipids in the Tumor Microenvironment: From Cancer Progression to Treatment. *Prog Lipid Res* (2020) 80:101055. doi: 10.1016/j.plipres.2020.101055
- Santos CR, Schulze A. Lipid Metabolism in Cancer. *FEBS J* (2012) 279(15):2610–23. doi: 10.1111/j.1742-4658.2012.08644.x
- Cespedes Feliciano EM, Kroenke CH, Meyerhardt JA, Prado CM, Bradshaw PT, Dannenberg AJ, et al. Metabolic Dysfunction, Obesity, and Survival Among Patients With Early-Stage Colorectal Cancer. *J Clin Oncol* (2016) 34(30):3664–71. doi: 10.1200/JCO.2016.67.4473
- O’Keefe SJ. Diet, Microorganisms and Their Metabolites, and Colon Cancer. *Nat Rev Gastroenterol Hepatol* (2016) 13(12):691–706. doi: 10.1038/nrgastro.2016.165
- Jia W, Xie G, Jia W. Bile Acid-Microbiota Crosstalk in Gastrointestinal Inflammation and Carcinogenesis. *Nat Rev Gastroenterol Hepatol* (2018) 15(2):111–28. doi: 10.1038/nrgastro.2017.119
- Butler LM, Perone Y, Dehairs J, Lupien LE, de Laat V, Talebi A, et al. Lipids and Cancer: Emerging Roles in Pathogenesis, Diagnosis and Therapeutic Intervention. *Adv Drug Deliv Rev* (2020) 159:245–93. doi: 10.1016/j.addr.2020.07.013
- Prentki M, Madiraju SR. Glycerolipid Metabolism and Signaling in Health and Disease. *Endocr Rev* (2008) 29(6):647–76. doi: 10.1210/er.2008-0007
- Ahluwalia P, Kolhe R, Gahlay GK. The Clinical Relevance of Gene Expression Based Prognostic Signatures in Colorectal Cancer. *Biochim Biophys Acta Rev Cancer* (2021) 1875(2):188513. doi: 10.1016/j.bbcan.2021.188513
- Zhou R, Zeng D, Zhang J, Sun H, Wu J, Li N, et al. A Robust Panel Based on Tumour Microenvironment Genes for Prognostic Prediction and Tailoring Therapies in Stage I-III Colon Cancer. *EBioMedicine* (2019) 42:420–30. doi: 10.1016/j.ebiom.2019.03.043
- Xu J, Dai S, Jiang K, Xiao Q, Yuan Y, Ding K. Combining Single-Cell Sequencing to Identify Key Immune Genes and Construct the Prognostic Evaluation Model for Colon Cancer Patients. *Clin Transl Med* (2021) 11(7):e465. doi: 10.1002/ctm2.465
- Cao Y, Jiao N, Sun T, Ma Y, Zhang X, Chen H, et al. Cxcl11 Correlates With Antitumor Immunity and an Improved Prognosis in Colon Cancer. *Front Cell Dev Biol* (2021) 9:646252. doi: 10.3389/fcell.2021.646252

32. Zhang J-X, Song W, Chen Z-H, Wei J-H, Liao Y-J, Lei J, et al. Prognostic and Predictive Value of a MicroRNA Signature in Stage II Colon Cancer: A MicroRNA Expression Analysis. *Lancet Oncol* (2013) 14(13):1295–306. doi: 10.1016/s1470-2045(13)70491-1
33. Jiang C, Liu Y, Wen S, Xu C, Gu L. *In Silico* Development and Clinical Validation of Novel 8 Gene Signature Based on Lipid Metabolism Related Genes in Colon Adenocarcinoma. *Pharmacol Res* (2021) 169:105644. doi: 10.1016/j.phrs.2021.105644
34. Sun Y, Liu B, Chen Y, Xing Y, Zhang Y. Multi-Omics Prognostic Signatures Based on Lipid Metabolism for Colorectal Cancer. *Front Cell Dev Biol* (2021) 9:811957. doi: 10.3389/fcell.2021.811957
35. Koeffler HP. Peroxisome Proliferator-Activated Receptor  $\Gamma$  and Cancers. *Clin Cancer Res* (2003) 9(1):1.
36. Ogino S, Shima K, Baba Y, Noshio K, Irahara N, Kure S, et al. Colorectal Cancer Expression of Peroxisome Proliferator-Activated Receptor Gamma (Pparg, Ppargamma) Is Associated With Good Prognosis. *Gastroenterology* (2009) 136(4):1242–50. doi: 10.1053/j.gastro.2008.12.048
37. Kuhn T, Stepien M, Lopez-Nogueroles M, Damms-Machado A, Sookthai D, Johnson T, et al. Prediagnostic Plasma Bile Acid Levels and Colon Cancer Risk: A Prospective Study. *J Natl Cancer Inst* (2020) 112(5):516–24. doi: 10.1093/jnci/djz166
38. Wang L, Wang Y, Song Z, Chu J, Qu X. Deficiency of Interferon-Gamma or Its Receptor Promotes Colorectal Cancer Development. *J Interferon Cytokine Res Off J Int Soc Interferon Cytokine Res* (2015) 35(4):273–80. doi: 10.1089/jir.2014.0132
39. Chawla-Sarkar M, Lindner DJ, Liu YF, Williams BR, Sen GC, Silverman RH, et al. Apoptosis and Interferons: Role of Interferon-Stimulated Genes as Mediators of Apoptosis. *Apoptosis an Int J Programmed Cell Death* (2003) 8(3):237–49. doi: 10.1023/a:1023668705040
40. Thapa RJ, Basagoudanavar SH, Nogusa S, Irrinki K, Mallilankaraman K, Slikker MJ, et al. Nf-Kappab Protects Cells From Gamma Interferon-Induced Rip1-Dependent Necroptosis. *Mol Cell Biol* (2011) 31(14):2934–46. doi: 10.1128/mcb.05445-11
41. Wang W, Green M, Choi JE, Gijón M, Kennedy PD, Johnson JK, et al. Cd8(+) T Cells Regulate Tumour Ferroptosis During Cancer Immunotherapy. *Nature* (2019) 569(7755):270–4. doi: 10.1038/s41586-019-1170-y
42. Du W, Frankel TL, Green M, Zou W. Ifngamma Signaling Integrity in Colorectal Cancer Immunity and Immunotherapy. *Cell Mol Immunol* (2021) 19(1):23–32. doi: 10.1038/s41423-021-00735-3
43. Ivashkiv LB. Ifngamma: Signalling, Epigenetics and Roles in Immunity, Metabolism, Disease and Cancer Immunotherapy. *Nat Rev Immunol* (2018) 18(9):545–58. doi: 10.1038/s41577-018-0029-z
44. Zaidi MR, Merlino G. The Two Faces of Interferon- $\Gamma$  in Cancer. *Clin Cancer Res* (2011) 17(19):6118–24. doi: 10.1158/1078-0432.Ccr-11-0482
45. Schmitt M, Greten FR. The Inflammatory Pathogenesis of Colorectal Cancer. *Nat Rev Immunol* (2021) 21(10):653–67. doi: 10.1038/s41577-021-00534-x
46. Ganesh K, Stadler ZK, Cercek A, Mendelsohn RB, Shia J, Segal NH, et al. Immunotherapy in Colorectal Cancer: Rationale, Challenges and Potential. *Nat Rev Gastroenterol Hepatol* (2019) 16(6):361–75. doi: 10.1038/s41575-019-0126-x
47. Pagès F, Mlecnik B, Marliot F, Bindea G, Ou F-S, Bifulco C, et al. International Validation of the Consensus Immunoscore for the Classification of Colon Cancer: A Prognostic and Accuracy Study. *Lancet* (2018) 391(10135):2128–39. doi: 10.1016/s0140-6736(18)30789-x
48. Marliot F, Chen X, Kirilovsky A, Sbarrato T, El Sissy C, Batista L, et al. Analytical Validation of the Immunoscore and Its Associated Prognostic Value in Patients With Colon Cancer. *J Immunother Cancer* (2020) 8(1):e000272. doi: 10.1136/jitc-2019-000272
49. Wu H, Han Y, Rodriguez Sillke Y, Deng H, Siddiqui S, Treese C, et al. Lipid Droplet-Dependent Fatty Acid Metabolism Controls the Immune Suppressive Phenotype of Tumor-Associated Macrophages. *EMBO Mol Med* (2019) 11(11):e10698. doi: 10.15252/emmm.201910698
50. Herber DL, Cao W, Nefedova Y, Novitskiy SV, Nagaraj S, Tyurin VA, et al. Lipid Accumulation and Dendritic Cell Dysfunction in Cancer. *Nat Med* (2010) 16(8):880–6. doi: 10.1038/nm.2172
51. Lin L, Ding Y, Wang Y, Wang Z, Yin X, Yan G, et al. Functional Lipidomics: Palmitic Acid Impairs Hepatocellular Carcinoma Development by Modulating Membrane Fluidity and Glucose Metabolism. *Hepatology* (2017) 66(2):432–48. doi: 10.1002/hep.29033

**Conflict of Interest:** Authors QZ, SdC, GW, and SIC were employed by Burning Rock Biotech.

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 © 2022 Wang, Zhang, Zhang, Zhou, Chen, Zheng, Wang, Cai, Wang and Li. 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.