Research Article

Key Markers Involved in the Anticolon Cancer Response of CD8+ T Cells through the Regulation of Cholesterol Metabolism

Liang Dong,[1] Xi Yang,[2] Yangyanqiu Wang,[3] Yin Jin,[4] Qing Zhou,[5] Gong Chen,[6] and Shuwen Han [7]

[1]*Department of Gastroenterology, Huzhou Central Hospital, Affiliated Central Hospital Huzhou University, Sanhuan North Road No. 1558, Wuxing District, Huzhou 313000, Zhejiang, China*

[2]*Department of Intervention and Radiotherapy, Huzhou Central Hospital, Affiliated Central Hospital Huzhou University, Huzhou 313000, Zhejiang, China*

[3]*Graduate School of Medical College of Zhejiang University, Kaixuan Road No. 268, Jianggan District, Hangzhou 310029, Zhejiang, China*

[4]*Department of Laboratory Medicine, Huzhou Central Hospital, Affiliated Central Hospital Huzhou University, Huzhou 313000, Zhejiang, China*

[5]*Department of Nursing, Huzhou Central Hospital, Affiliated Central Hospital Huzhou University, Sanhuan North Road No.1558, Wuxing District, Huzhou 313000, Zhejiang, China*

[6]*Undergraduate School of Clinic Medicine, Huzhou University, Huzhou 313000, Zhejiang, China*

[7]*Department of Oncology, Huzhou Central Hospital, Affiliated Central Hospital Huzhou University, Huzhou 313000, Zhejiang, China*

Correspondence should be addressed to Shuwen Han; shuwenhan985@163.com

Background. T cell-mediated antitumor immune response is the basis of colorectal cancer (CRC) immunotherapy. Cholesterol plays an important role in T cell signal transduction and function. Apolipoprotein *E* (APOE) plays a major role in cholesterol metabolism. *Objective.* To screen and analyze key markers involved in the anticolon cancer response of CD8+ T cells through the regulation of cholesterol metabolism. *Methods.* Based on the median cutoff of the expression value of *APOE* according to the data downloaded from The Cancer Genome Atlas and Gene Expression Omnibus database, patients were grouped into low and high expression groups. Differences in clinical factors were assessed, and survival analysis was performed. Differentially expressed genes (DEGs) in the high and low expression groups were screened, followed by the analysis of differences in tumor-infiltrating immune cells and weighted gene coexpression network analysis results. The closely related genes to *APOE* were identified, followed by enrichment analysis, protein–protein interaction (PPI) network analysis, and differential expression analysis. Immunohistochemical staining (IHC) was used to detect the expression of CD8 in CRC tissues. *Results.* There were significant differences in prognosis and pathologic_N between the APOE low and high expression groups. A total of 2,349 DEGs between the high and low expression groups were selected. A total of 967 genes were obtained from the blue and brown modules. The probability of distribution of CD8+ T cells differed significantly between the two groups, and 320 closely related DEGs of *APOE* were screened. Genes including the *HLA* gene family, *B2M*, *IRF4*, and *STAT5A* had a higher degree in the PPI network. GEO datasets verified the prognosis and the related DEGs of *APOE*. IHC staining verified the relationship between the distribution of CD8+ T cells and *APOE* expression. *Conclusion.* Genes including the *HLA* gene family, *B2M*, *IRF4*, and *STAT5A* might be the key genes involved in the anticolon cancer response of CD8+ T cells through the regulation of cholesterol metabolism.

1. Introduction

Colon cancer is a major health problem worldwide because it is the second leading cause of cancer-related deaths [1]. It is reported that by 2030, especially in developed countries, the number of new cases will increase to 2.2 million, and cases of death will increase to 1.1 million [2]. Although surgical resection combined with radiation therapy and chemotherapy is the main treatment for colon cancer, the 5-year survival rate is only approximately 28% [3]. Excessive dietary fat is associated with the induction/exacerbation of various diseases, including colon cancer [4]. In addition, numerous studies have illustrated that a "Western" style diet with high-fat content is an important factor in the increased incidence of colon cancer [5–7]. A high-fat diet is also widely involved in various pathological conditions, such as obesity and metabolic diseases, which may promote the development of colon cancer [8]. Hyperlipidemia is often divided into four types: hypercholesterolemia, hypertriglyceridemia, mixed hyperlipidemia, and low high-density lipoprotein cholesterol. A previous study indicated that hyperlipidemia is a risk factor for colon cancer [9].

T cell-mediated antitumor immune response is the basis of tumor immunotherapy, which is associated with a favorable prognosis [10]. However, some tumors evolve to acquire immunosuppressive properties and escape the attack by T cells through various mechanisms in the tumor microenvironment [11]. Thus, reactivating the cytotoxicity of T cells is of great clinical interest in cancer immunotherapy. Cholesterol plays an important role in T cell signal transduction and function. For instance, Yang et al. found that the antitumor response of mouse CD8+ T cells can be enhanced by regulating cholesterol metabolism [12]. Ma et al. showed that high cholesterol can facilitate T cell immune checkpoint expression, which makes it easier for T cells to lose their antitumor function [13]. However, these studies only focused on T cell gene changes and endoplasmic reticulum stress and did not examine the impact of tumor cells.

It has been reported that high cholesterol levels can promote the proliferation of stem cells, thereby increasing the growth rate of intestinal tumors by 100-fold [14]. RORα/hdac [15], CD36 [16], ACAT1 [17], and so on are involved in the regulation of the antitumour response of CD8+ T cells by modulating cholesterol metabolism. However, the key markers involved are still unclear. Thus, this study aimed to screen and analyze the key markers involved in the anticolon cancer response of CD8+ T cells through the regulation of cholesterol metabolism. The workflow of this study is shown in Figure 1.

2. Materials and Methods

2.1. Data Sources and Data Preprocessing.
Gene expression RNAseq data [$\log_2$(fpkm+1)] and clinical data of colon adenocarcinoma were obtained from The Cancer Genome Atlas database [18] (https://xenabrowser.net/). After sequencing data and clinical information matching, a total of 389 tumor samples were obtained (data version: 07-19-2019). RNA-Seq was annotated based on the annotation file
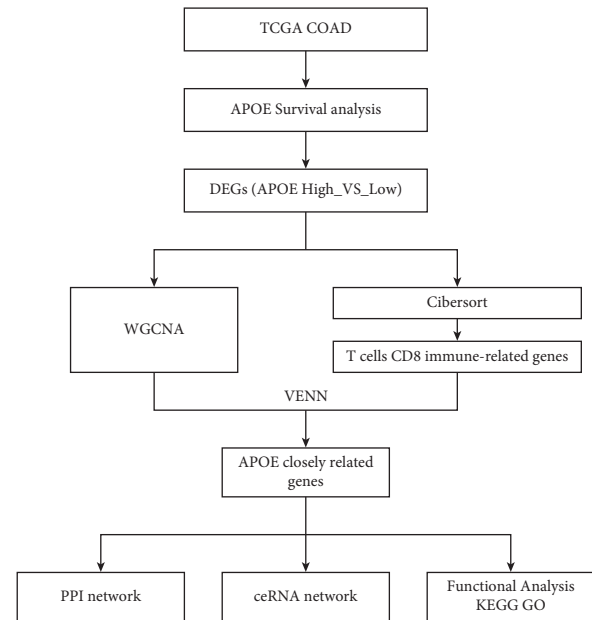

FIGURE 1: The workflow of this study.

of the Gencode database [19] (V23, https://www.gencodegenes.org/). In the expression profile analysis of Ensembl_ID, the mapping probe was used to calculate the gene expression value (obtained from the annotation files of the chip platform and microarray dataset) to Symbol_ID. The average value was taken as the level of Ensembl_ID expression when multiple probes matched one Symbol_ID. Genes with "protein_coding" annotations were then extracted as mRNAs.

2.2. APOE Expression-Related Clinical and Survival Analysis.
Apolipoprotein E (APOE) plays a major role in cholesterol metabolism [20]. Thus, in this study, to screen the key markers involved in the anticolon cancer response of CD8+ T cells via cholesterol metabolism regulation, the expression level of APOE was calculated using the Survminer package of R (version: 0.4.8, https://CRAN.Rproject.org/package=survminer), followed by the median cutoff of the expression value. Then, based on the median cutoff of the expression value, patients were grouped into low and high expression groups. In addition, the Kaplan–Meier survival curve analysis and log-rank test were performed to compare the prognosis between the two groups based on the survival information of the samples. Clinical factors (age, sex, TNM stage, and pathologic_stage) were compared between the low and high expression groups using the chi-squared test; differences were considered significant at a threshold of $P < 0.05$.

2.3. Differential Expression Analysis.
The typical Bayesian method in the limma package [21] (version 3.40.6) was used to analyze differentially expressed genes (DEGs) between high and low expression groups. The Benjamini and Hochberg method was used to adjust the P values. Adjusted

P value < 0.05 and $|logFC| > 0.263$ were used as the cutoff criteria for screening DEGs. Finally, the ggscatter function of the ggpubr package [22] in R (version: 0.2.2) was used to draw a volcano plot.

2.4. Screening of APOE Expression-Related Modules and Genes.

The weighted gene coexpression network analysis (WGCNA) algorithm was used to screen the modules and genes related to APOE expression based on the expression values of DEGs. The network construction and module screening procedures included dataset consistency analysis and gene coexpression correlation matrix, adjacent function, phase difference between nodes, and correlation analyses between network modules and diseases. A heat map was used to visualize the correlations of each module.

2.5. Screening of Immune-Related Genes.

The CIBERSORT algorithm [23], a useful method for obtaining high-throughput characteristics of 22 cell types in complex tissues, was used to analyze the abundance of tumor-infiltrating immune cells with the parameters of perm = 100 and QN = F. Subsequently, the abundance of all types of immune cells was tested in the high and low expression groups. Then, the Pearson correlation coefficient between DEGs and CD8+ T cells was calculated, and immune-related genes were obtained with a cutoff of $P < 0.05$ and $|r| > 0.3$. The module genes obtained from WGCNA were intersected with these immune-related genes, and the overlapping genes were redefined as closely related genes of *APOE*.

2.6. Enrichment Analysis.

Based on the closely related genes of *APOE*, the cluster Profiler package [24] in R was used to perform Gene Ontology (GO) [25] and Kyoto Encyclopedia of Genes and Genomes (KEGG) [26] pathway enrichment analyses with a cutoff value of $P < 0.05$ and count ≥2.

2.7. Protein-Protein Interaction (PPI) Network Analysis.

The STRING database [27] was used to evaluate the PPIs encoded by closely related genes of *APOE*. The PPI score was set at 0.9. Subsequently, the PPI network of T cell-related genes was analyzed using Cytoscape software [28].

2.8. Analysis of DEGs Related to APOE and CD8+ T Cells.

R software ggplot2 (version: 3.2.1, https://CRAN.R-project.org/package=ggplot2) and GGpubr ggpubr (version 0.2.2, t-test provided by https://CRAN.R-project.org/package=ggpubr) were used to analyze the differences of genes between the two groups. A box diagram was drawn for display with a threshold of $P < 0.05$.

2.9. Datasets Validation.

Two GEO datasets, GSE71187 and GSE39582, with prognostic information were selected from the GEO database to validate the result of APOE expression-related survival analysis. The series_matrix.txt data containing 99 tumors and 12 para-cancers was obtained after pretreatment of GSE71187. Furthermore, the series_matrix.txt

containing 566 tumors and 19 para-cancers was obtained after pretreatment of GSE39582. The expression of APOE in the two datasets was extracted. R Software Survminer (version 0.4.8) package calculated APOE expression to obtain optimal cutoff. The expression value > optimal cutoff was regarded as high expression, and the expression value ≤ optimal cutoff was regarded as low expression. Then, the prognosis of patients in the two groups was obtained. Combined with the survival information of the samples, the k-M curve was plotted, and log-rank was used to test its significance.

Patients in GSE71187 and GSE39582 were divided into two groups according to the median APOE expression. The expressions of transcription 5A (STAT5A), histocompatibility leukocyte antigen (HLA) gene family (HLA-E, HLA-C, HLA-B, etc.), beta-2 microglobulin (B2M), and interferon regulatory Factor 4 (IRF4) were tested in high and low APOE expression groups. R software GGplot2 (version 3.2.1, https://CRAN.R-project.org/package=ggplot2) and GGpubr ggpubr (version 0.2.2, the t-test provided by https://CRAN.R-project.org/package=ggpubr) were used to analyze the differences of genes between the two groups. A box diagram was drawn for display with a threshold of $P < 0.05$.

2.10. CRC Specimens and IHC Staining.

The study protocol was approved by the Human Ethics Committee of Huzhou Central Hospital (NO: 20210207). Forty CRC patients were recruited, and the basic information was presented in Supplementary Table 2. The expression of APOE in blood was detected by using immunoturbidimetry. APOE higher than 53.0 mg/L was included in the APOE-high group, and APOE lower than 29.0 mg/L was included in the APOE-low group. Further detection of CD8 expression was conducted by using immunohistochemistry. Immunohistochemistry test kit, antigen repair solution, and CD8 antibody were all purchased from Beijing Zhongshan Jinqiao Biotechnology Co., Ltd. The tissue sections were scanned at low magnification (X100) to select the areas with high CD8 positive cell density. Then, the positive rate of CD8 cells in five fields was determined at high magnification (X200) to calculate the average value. The positive rate of CD8 cells was expressed as mean ± standard deviation. Using SPSS 13.0 statistical software, the Pearson $\chi2$ test was performed.

3. Results

3.1. Comparison of Clinical Factors and Analysis of DEGs between the High and Low Expression Groups.

The Kaplan–Meier survival curve analysis and log-rank test were performed to compare the prognosis between the two groups based on the survival information of the samples. The results showed that there were significant differences in prognosis between the low and high expression groups ($P = 0.018$) (Figure 2(a)). In addition, a significant difference in pathologic N was found between the two groups ($P = 0.029$) (Figure 2(b)). In total, 2,349 DEGs, of which 1,949 and 400 were upregulated and downregulated, respectively, between the high and low expression groups were selected (Figure 2(c)).
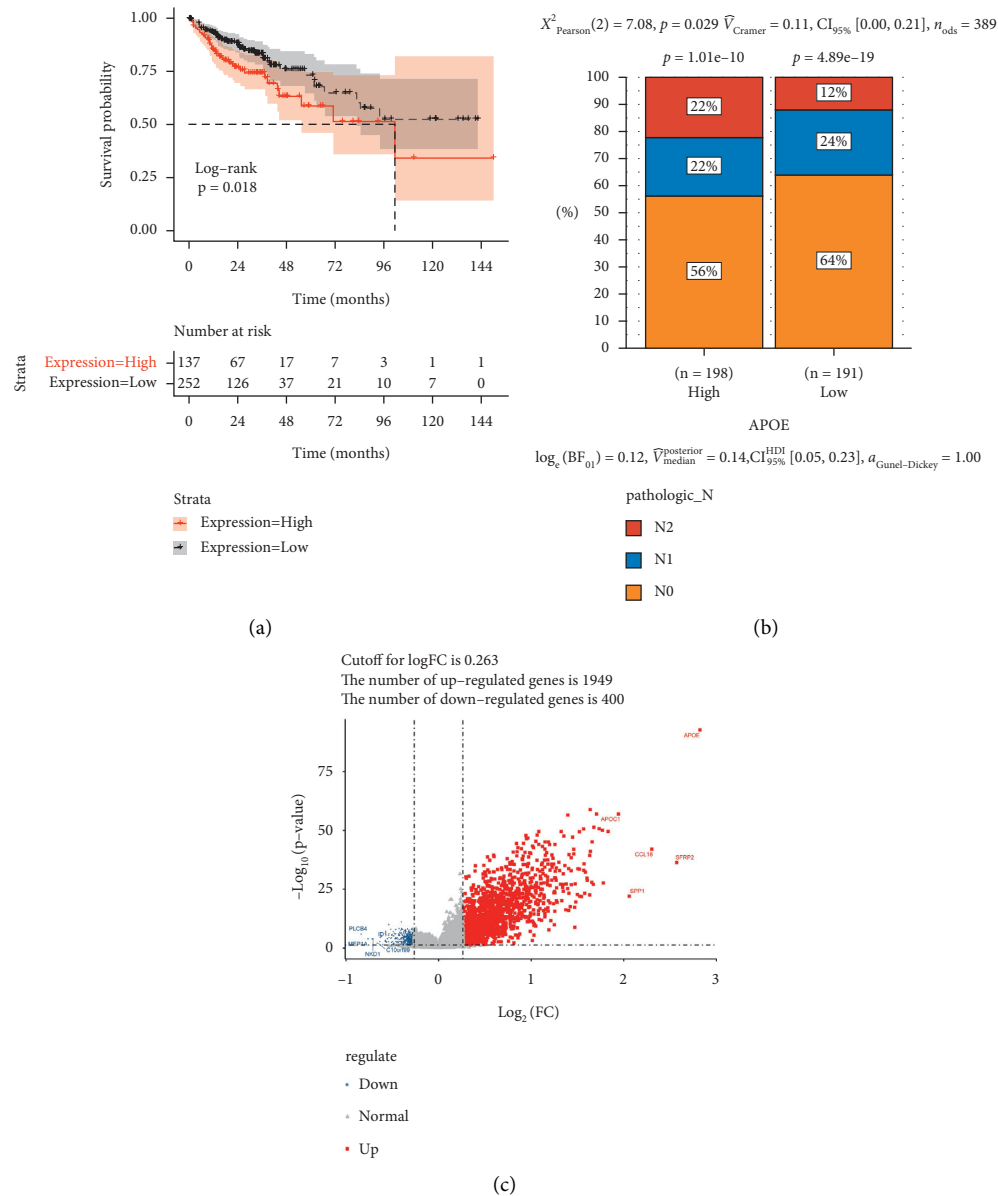
(a)

(b)

(c)

FIGURE 2: Comparison of clinical factors, survival analysis, analysis of DEGs between the high and low expression groups. (a) Survival analysis between low and high expression groups. (b) Comparison of clinical factors (age, sex, TNM stage, pathologic_stage) between the low and high expression groups. (c) DEGs between low and high expression groups; the red nodes represent upregulated DEGs, and the blue nodes represent downregulated DEGs.

3.2. Disease-Related Modules and Genes. A total of 2,349 DEGs were used for WGCNA analysis. The soft threshold power for matrix transformation was determined to be 8, with the square of the correlation coefficient between log2k and log2p (*k*) being 0.85 (Figure 3(a)). The minimum number of genes was set to 30 for each module, and the pruning height was cutHeight = 0.2. A total of five different disease-related modules were screened (Figure 3(b)), including blue, brown, green, red, and yellow modules. To ensure the reliability of the screening results of key network modules, the disease-related key network modules were analyzed again by calculating the absolute value of gene significance (GS) within the module, and the results showed that the blue and brown modules had a higher GS

(Figure 3(c)) and were significantly correlated with the disease (Figure D). Finally, a total of 967 genes were obtained from the blue and brown modules.

3.3. CD8+ T Cell-Related DEGs. Immune infiltration in the high and low expression groups was analyzed based on the CIBERSORT algorithm. As illustrated in Figures 4(a) and 4(b), 22 tumor-immune cell proportions in the high and low expression groups were analyzed. To observe the probability of the distribution of different tumor-immune cells between the two groups, a violin plot was drawn. As shown in Figure 4(c), the probability of the distribution of CD8+ T cells was the most significant difference between the two
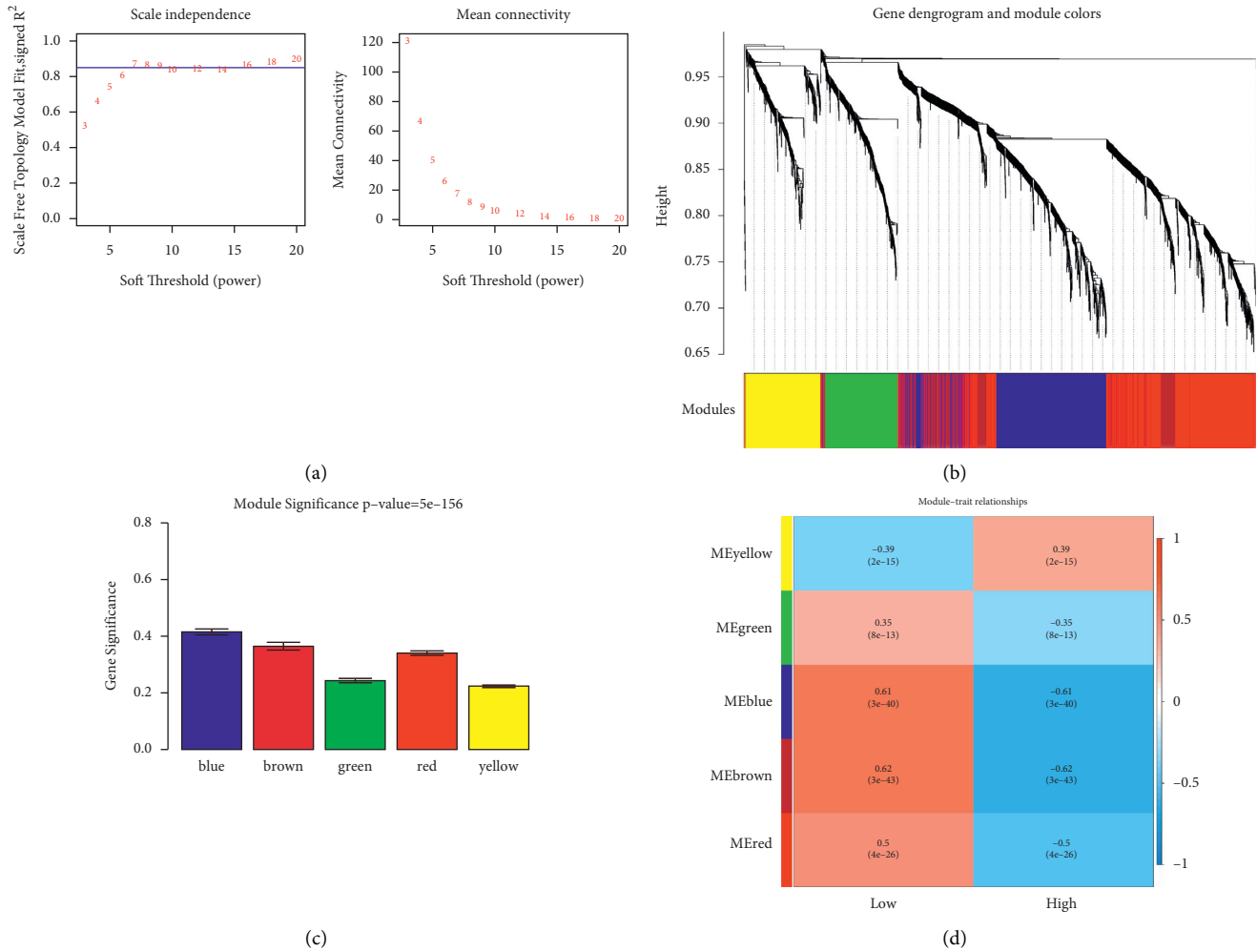
(a)

(b)

(c)

(d)

FIGURE 3: Disease-related modules and genes. (a) Determination of the soft threshold for the adjacency matrix. The horizontal axis represents the soft threshold power, and the vertical axis represents the square of the correlation coefficient between $\log_2 k$ and $\log_2 p$ (k). The blue line indicates a correlation coefficient of 0.85 and a corresponding soft threshold power of 8. (b) Gene dendrogram derived from hierarchical clustering. Different modules are indicated by the colors below the dendrogram. (c) Gene significance (GS) of five modules. (d) Module–disease relationships. The modules are represented by different colors along the horizontal axis. Correlation coefficients are indicated as numbers in corresponding positions, and p values are shown in brackets along the coefficients. Red indicates higher correlation coefficients.

groups ($P < 0.01$). The correlation coefficient between DEGs and CD8+ T cell-related genes was then calculated, and a total of 612 CD8+ T cell-related DEGs were identified. In addition, the module genes obtained from WGCNA intersected with these CD8+ T cell-related genes, and a total of 320 closely related DEGs of *APOE* was obtained (Figure 4(d), Supplementary Table 1).

3.4. Enrichment Analysis and PPI Network Analysis of Closely Related DEGs of APOE.
Enrichment analysis showed that the closely related upregulated DEGs were mainly involved in 544 GO-biological processes (BPs; e.g., T cell activation; regulation of lymphocyte activation; and regulation of T cell activation) and 54 KEGG pathways (e.g., hsa05169: Epstein–Barr virus (EBV) infection and hsa04514: cell adhesion molecules) (Figures 5(a) and 5(b)), whereas the closely

related downregulated DEGs were not involved in the GO-BP and KEGG pathways. PPI network analysis of 320 closely related DEGs of *APOE* revealed 191 nodes and 1,153 interaction pairs (Figure 5(c)). In this PPI network, genes including the histocompatibility leukocyte antigen (HLA) gene family (HLA-E, HLA-C, HLA-B, etc.), beta-2-microglobulin (*B2M*), interferon regulatory factor 4 (*IRF4*), and STAT5A had a higher degree, which could lead to their classification as key genes (Table 1).

3.5. Analysis of DEGs Related to APOE and CD8+ T Cells.
As shown in Figure 5(d), histocompatibility leukocyte antigen (HLA) gene family (HLA-E, HLA-C, HLA-B, etc.), beta-2-microglobulin (B2M), interferon regulatory factor 4 (IRF4), and STAT5A were significantly differentially expressed in high and low APOE groups ($P < 0.001$).
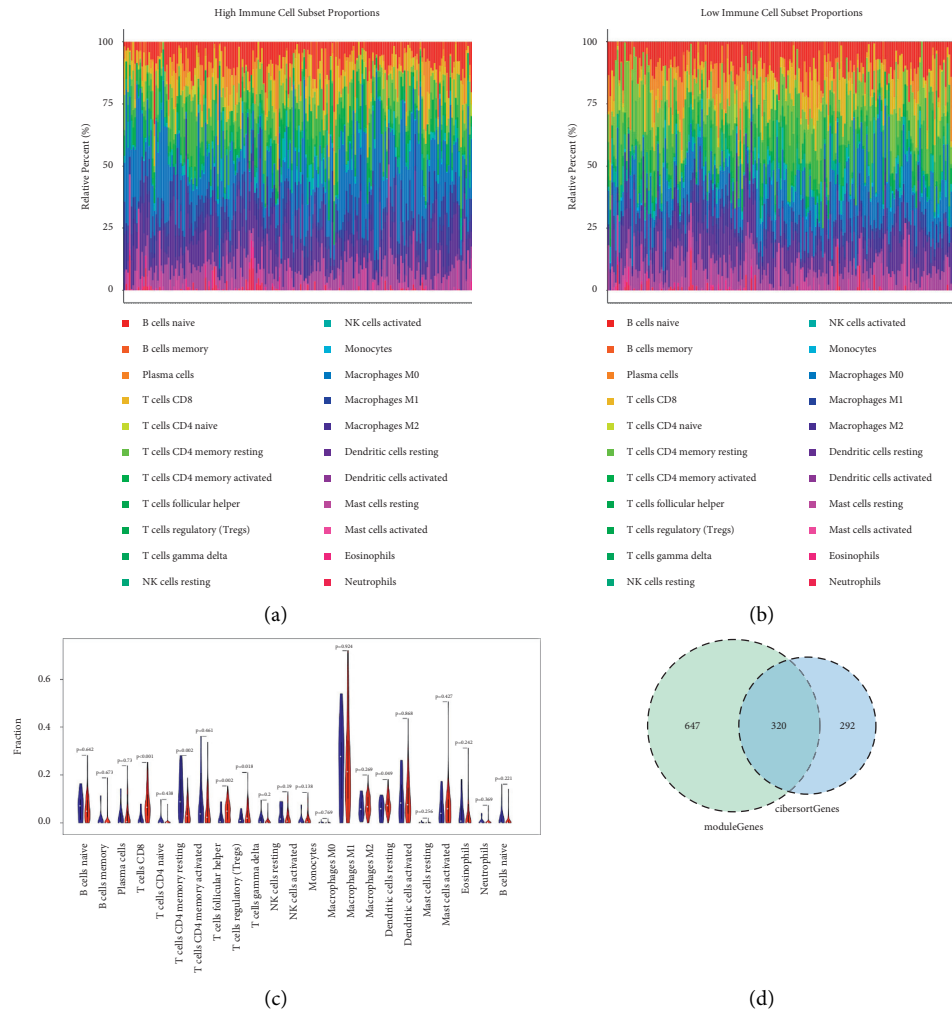
(a)

(b)

(c)

(d)

FIGURE 4: Screening of CD8+ T cell-related DEGs. (a) The 22 tumor-immune cell proportions of high expression groups. (b) The 22 tumor-immune cell proportions of low expression groups. (c) The violin plot of different tumor-immune cells between high and low expression groups; blue represents the high expression group, and red represents the control group. (d) The Venn diagram of closely related DEGs of *APOE*. WGCNA: weighted gene coexpression network analysis; APOE: apolipoprotein E.

3.6. Validation of APOE-related survival analysis and DEGs related to APOE and CD8+ T cells. As shown in the K-M curve in Figures 6(a) and 6(b), the survival analysis results of GSE3958 and GSE71187 were verified that the high expression group had a poor prognosis ($P < 0.05$). Moreover, as shown in the box diagram in Figures 6(c) and 6(d), most of the concerned genes in the GSE39582 and GSE71187 datasets were highly expressed ($P < 0.05$). The results were consistent with the analysis results of the TCGA database.

3.7. The Infiltration of CD8+T Cells Correlated with APOE in Tissue Specimen Sections. The basic information of recruited CRC patients was present in Supplementary Table 2. IHC staining detection of CD8 proteins showed that CD8 was upregulated expression in the APOE-high expression group (Figure 7(a)), while CD8 was downregulated expression in the APOE-low expression group (Figure 7(b)) ($P = 0.001$). The statistical results are presented in Table 2.

4. Discussion

In this study, there were significant differences in prognosis and pathologic_N between the low and high expression groups. A total of 2,349 DEGs between the high and low expression groups were selected. In addition, the probability of CD8+ T cell distribution was the most significant difference between the two groups. A total of 320 closely related DEGs of *APOE* were identified. Genes including the *HLA* gene family, *B2M*, and *IRF4* had a higher degree in the PPI network, and *STAT5A* was regulated by numerous miRNAs in the ceRNA network.

Cholesterol has been reported to be associated with colon cancer. Eggs low in fat contain large amounts of cholesterol, and egg consumption has been associated with an increased risk of colon cancer [29]. In addition, Jacobs et al. showed that high cholesterol levels seem to increase the risk of colon cancer [30]. In line with our data, in this study, the survival analysis results showed that the high cholesterol group was related to poor survival, which suggested that
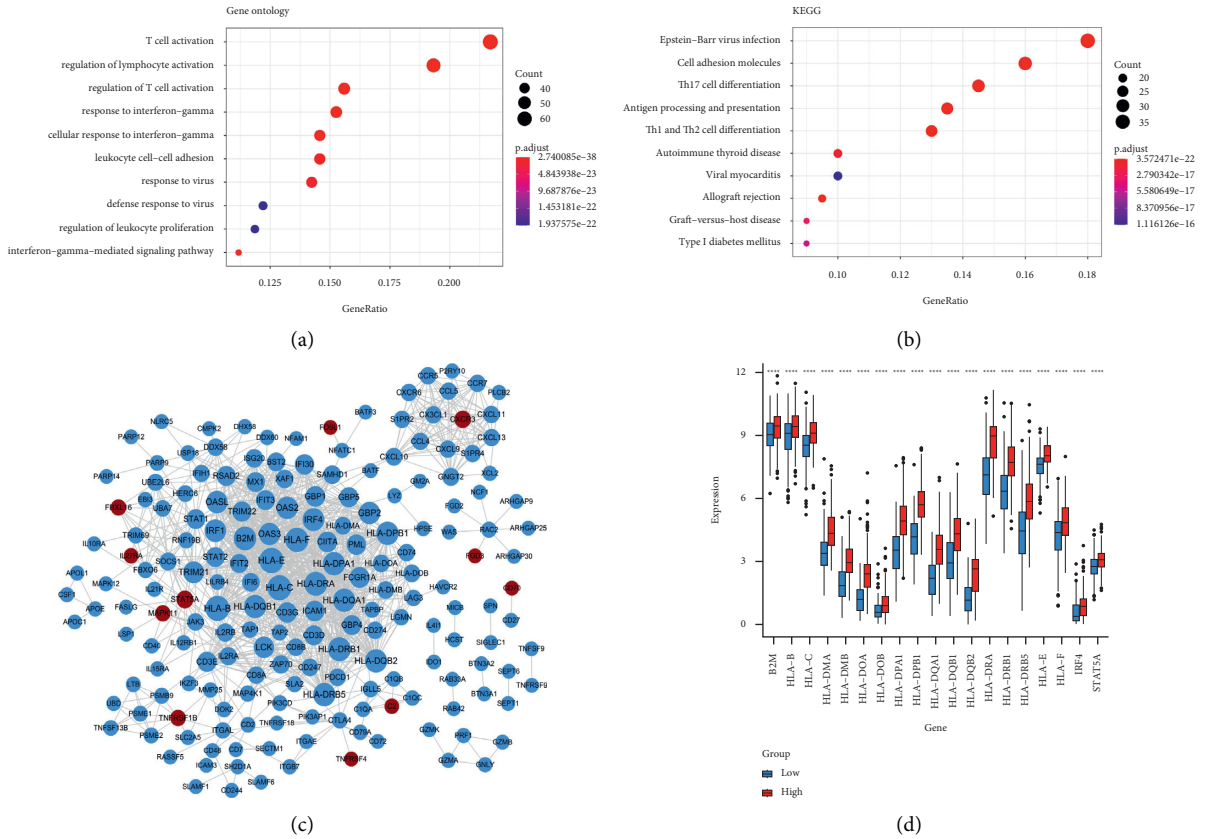
(a)

(b)

(c)

(d)

Figure 5: Enrichment analysis, PPI network analysis, and DEGs related to APOE and CD8+ T cells. (a) GO-BP. (b) KEGG pathway. The size of the ball represents the number of genes enriched in each term. The color of the ball represents the value of the *P* value. (c) PPI network; blue nodes represent the genes in the blue module, and brown nodes represent genes in the brown module. (d) Analyses of DEGs related to APOE and CD8+ T cells.

Table 1: The degree of top 20 nodes in protein-protein interaction (PPI) network.

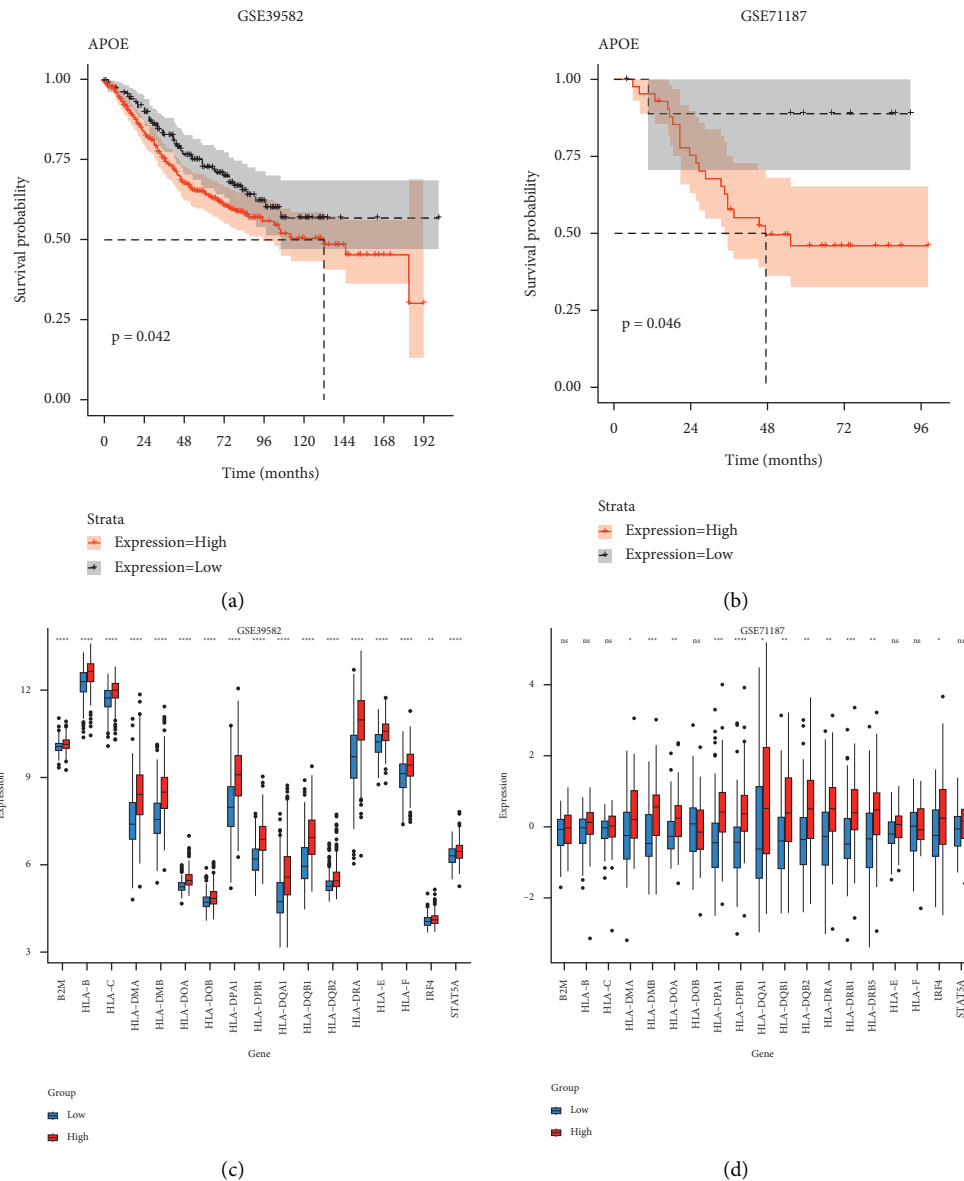| Gene	Degree
HLA-E	52
HLA-C	51
HLA-B	51
HLA-F	47
HLA-DRB1	45
HLA-DRA	45
B2M	44
HLA-DQA1	42
HLA-DQB1	42
HLA-DRB5	42
HLA-DQB2	42
HLA-DPA1	42
HLA-DPB1	42
IRF4	40
IRF1	39
OASL	38
OAS2	38
OAS3	37
GBP2	37
TRIM21	35

FIGURE 6: Validation of survival analysis and DEGs related to APOE and CD8+ T cells. A-B: survival analysis between APOE low and high expression groups in GSE39582 and GSE71187. C-D: analysis of DEGs related to APOE and CD8+ T cells in GSE39582 and GSE71187.

high cholesterol was a risk factor for colon cancer. In addition, a significant difference in pathologic_N was found between the low and high expression groups, and the proportion of pathologic_N3 in the high expression group was higher than that in the low expression group, which further indicated that high cholesterol was a risk factor for colon cancer.

CD8+ T cells play an extremely crucial role in antitumor immunity; however, tumor-infiltrating T cells often lose their effector functions. Ma et al. reported that cholesterol induced CD8+ T cell exhaustion by modulating endoplasmic reticulum stress pathways (IRE1/XBP1) in the tumor microenvironment [13]. What is more, CD8+ T cells were polarized into IL-9-secreting (Tc9) cells to exert antitumor responses, while cholesterol or its derivatives inhibited IL-9 expression by activating liver *X* receptors (LXRs) and leading

to LXR SUMOylation and reduced p65 binding to IL-9 promoter [31]. The characteristics of exhausted CD8+ T cells were attenuating antitumor responses and decreasing infiltrating density. Therefore, high cholesterol decreased the infiltration of CD8+ T cells. In this study, we also analyzed the key markers involved in the anticolon cancer response of CD8+ T cells through cholesterol metabolism regulation using bioinformatics analysis. The results showed that genes including the *HLA* gene family, *B2M*, and *IRF4* had a higher degree in the PPI network, and *STAT5A* was regulated by numerous miRNAs in the ceRNA network. The *HLA* gene family plays a major role in regulating immune responses in cancers. For instance, Tsai et al. revealed that *HLA-DQA1* and *HLA-DQB1* may participate in the development of oral cancer [32]. Bianchini et al. showed that the *HLA-E* gene strongly supports a potential tumor-evading immune
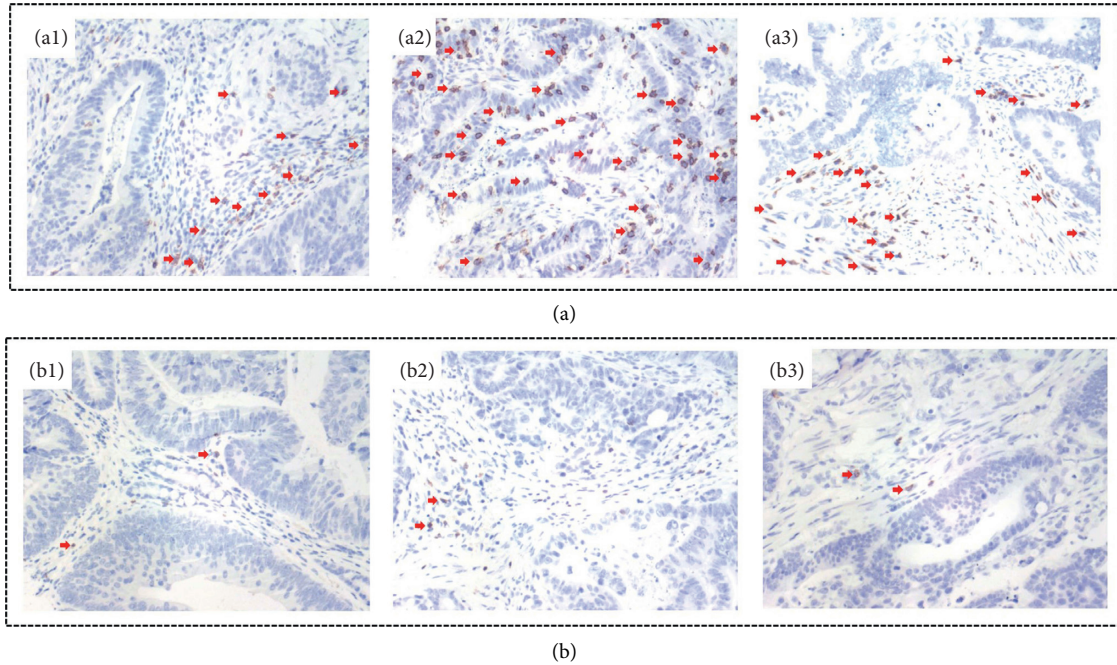
FIGURE 7: IHC staining of CD8. (a). IHC staining for CD8 in the high-APOE expression group. (b). IHC staining for CD8 in the low-APOE expression group. The magnification was X200. The red arrows represented positive CD8.

TABLE 2: Immunohistochemical evaluation.

CRC	High-APOE	Low-APOE	Total	X^2	P value
CD8 ++/+++	18	8	26	—	—
CD8−/+	2	12	14	—	—
Total	20	20	40	10.989	0.001

Criterion: —: 0–10%; +: 11–30%; ++: 31–70%; +++: 70–100%.

surveillance strategy in colon cancer tissues [33]. Moreover, Benevolo et al. found that the *HLA-A*, *HLA-B*, and *HLA-C* expression in colon cancer is associated with prognosis [34]. Blum et al. suggested that, in patients with stage II colon cancer, the Map7/B2M expression ratio is a prognostic factor for survival [35]. *IRF4* is considered an oncogene in lymphoid malignancy and multiple myeloma [36], and Zhang et al. showed that overexpression of circPIP5K1A attenuated the expression of *IRF4* in the progression of colon cancer development [37]. As a transcription factor, the activation or phosphorylation of STAT is associated with many cancers. Slattery et al. showed that STAT5A was associated with colon cancer progression [38]. In addition, nucleotide analysis showed that the upregulated closely related genes were involved in EBV infection and cell adhesion molecule KEGG pathways. EBV is associated with several malignant tumors. Guan et al. reported a positive association between EBV infection and colon cancer [39]. It has been reported that a variety of cell adhesion molecules are involved in cell–cell and cell–matrix interactions in colon cancer, and some cell–cell and cell–matrix interactions determine colon cancer behavior [40]. Based on these results, we speculate that genes including the *HLA* gene family, *B2M*, *IRF4*, and *STAT5A* that may be involved in EBV infection and cell adhesion molecule pathways play important roles in the anticolon cancer response of CD8+ T cells by regulating cholesterol metabolism.

However, there are some limitations to this study. We need to conduct related experiments such as cell biology experiments and animal studies to validate the multiple candidate targets and signaling pathways obtained in this bioinformatics analysis.

5. Conclusion

In summary, genes including the *HLA* gene family, *B2M*, *IRF4*, and *STAT5A* might be the key genes involved in the anticolon cancer response of CD8+ T cells via regulation of cholesterol metabolism. These findings can guide clinical decision-making for colon cancer treatment as well as future research on colon cancer.

Abbreviation

KEGG: Kyoto Encyclopedia of Genes and Genomes
GO: Gene Ontology
GEO: Gene Expression Omnibus
PPI: Protein–protein interaction
BP: Biological process
ceRNA: Competing endogenous RNA
WGCNA: Weighted gene coexpression network analysis
DEGs: Differentially expressed genes
APOE: Apolipoprotein E
HLA: Histocompatibility leukocyte antigen
STAT5A: Signal transducer and activator of transcription 5A
IRF4: Interferon regulatory factor 4
B2M: Beta-2-microglobulin.

Data Availability

Raw data were obtained from TCGA repositories and https://xenabrowser.net/datapages/. Validation data were obtained from GEO repositories (GSE3958 and GSE71187) (https://www.ncbi.nlm.nih.gov/geo). Analyzed data are available within supplementary information or from the authors upon reasonable request.

Additional Points

1. The prognosis of colon cancer differed between the high and low cholesterol groups. 2. The probability of CD8+ T cell distribution differed significantly between the high and low cholesterol groups. 3. Genes including the *HLA* gene family, *B2M*, and *IRF4* might be the key genes involved in the anticolon cancer response of CD8+ T cells by regulating cholesterol metabolism.

Ethical Approval

The clinical protocols involving the patients and the informed consent form were approved by the Ethics Committee of Huzhou Central Hospital (no. 20191101-01).

Conflicts of Interest

The authors declare that no conflicts of interest.

Authors' Contributions

All the authors participated in the conception and design of the study. Dong Liang and Han Shuwen conceived and drafted the manuscript. Yang Xi and Jin Yin analyzed the data. Wang Yangyanqiu and Zhou Qing drafted the figure. Dong Liang, Chen Gong, and Han Shuwen wrote and revised the article. All the authors read and approved the paper.

Acknowledgments

Supplementary Materials

Supplementary Table 1. Related DEGs of APOE. Module genes obtained from WGCNA were intersected with CD8+ T cell-related genes, and a total of 320 closely related DEGs of APOE was obtained. Supplementary Table 2. Basic information of 40 CRC patients. Raw data were obtained from TCGA repositories and, for our purposes, from https://xenabrowser.net/datapages/. Analyzed data are available within supplementary information or from the authors upon reasonable request. (*Supplementary Materials*)

References

[1] A. Yaghoubi, M. Khazaei, A. Avan, S. M. Hasanian, and S. Soleimanpour, "The bacterial instrument as a promising therapy for colon cancer," *International Journal of Colorectal Disease*, vol. 35, no. 4, pp. 595–606, 2020.

[2] M. Arnold, M. S. Sierra, M. Laversanne, I. Soerjomataram, A. Jemal, and F. Bray, "Global patterns and trends in colorectal cancer incidence and mortality," *Gut*, vol. 66, no. 4, pp. 683–691, 2017.

[3] T. Watanabe, K. Muro, Y. Ajioka et al., "Japanese Society for Cancer of the Colon and Rectum (JSCCR) guidelines 2016 for the treatment of colorectal cancer," *International Journal of Clinical Oncology*, vol. 23, no. 1, pp. 1–34, 2018.

[4] S. Choi and A. J. Snider, "Diet, lipids and colon cancer," *Cellular Nutrient Utilization and Cancer*, vol. 347, pp. 105–144, 2019.

[5] L. Bailie, M. B. Loughrey, and H. G. Coleman, "Lifestyle risk factors for serrated colorectal polyps: a systematic review and meta-analysis," *Gastroenterology*, vol. 152, no. 1, pp. 92–104, 2017.

[6] S. J. O'Keefe, J. V. Li, L. Lahti et al., "Fat, fibre and cancer risk in African Americans and rural Africans," *Nature Communications*, vol. 6, p. 6342, 2015.

[7] L. C. Vinikoor, R. C. Millikan, J. A. Satia et al., "trans-Fatty acid consumption and its association with distal colorectal cancer in the North Carolina Colon Cancer Study II," *Cancer Causes & Control*, vol. 21, no. 1, pp. 171–180, 2010.

[8] A. Giuliani, F. Ferrara, M. Scimò, F. Angelico, L. Olivieri, and L. Basso, "Adipose tissue fatty acid composition and colon cancer: a case-control study," *European Journal of Nutrition*, vol. 53, no. 4, pp. 1029–1037, 2014.

[9] H. Rodriguez-Broadbent, "Mendelian randomisation implicates hyperlipidaemia as a risk factor for colorectal cancer," *International Journal of Cancer*, vol. 140, no. 12, pp. 2701–2708, 2017.

[10] A. Ribas and J. D. Wolchok, "Cancer immunotherapy using checkpoint blockade," *Science*, vol. 359, 2013.

[11] S. Gabrilovich and A. Dmitry, "Mechanisms and functional significance of tumour-induced dendritic-cell defects," *Science*, vol. 4, no. 12, pp. 941–952, 2009.

[12] W. Yang, Y. Bai, Y. Xiong et al., "Potentiating the antitumour response of CD8+ T cells by modulating cholesterol metabolism," *Nature*, vol. 531, no. 7596, pp. 651–655, 2016.

[13] X. Ma, E. Bi, Y. Lu et al., "Cholesterol induces CD8+ T cell exhaustion in the tumor microenvironment," *Cell Metabolism*, vol. 30, no. 1, pp. 143–156, 2019, e5.

[14] B. Wang, X. Rong, E. N. D. Palladino et al., "Phospholipid remodeling and cholesterol availability regulate intestinal stemness and tumorigenesis," *Cell Stem Cell*, vol. 22, no. 2, pp. 206–220, 2018, e4.

[15] L. In Kyu, "RORα regulates cholesterol metabolism of CD8+T cells for anticancer immunity," *Cancers*, vol. 12, no. 7, p. 1733, 2020.

[16] G. Mikhail, "Kolonin. Bad cholesterol uptake by CD36 in T-cells cripples anti-tumor immune response," *Immunometabolism*, vol. 3, no. 4, Article ID e210028, 2021.

[17] Y. Wei, "Potentiating the antitumour response of CD8+ T cells by modulating cholesterol metabolism," *Nature*, vol. 531, no. 7596, pp. 651–655, 2016.

[18] K. Tomczak, P. Czerwińska, and M. Wiznerowicz, "The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge," *Contemporary Oncology*, vol. 19, no. 1A, pp. A68–77, 2015.

[19] J. Harrow, A. Frankish, J. M. Gonzalez et al., "GENCODE: the reference human genome annotation for the ENCODE Project," *Genome Research*, vol. 22, no. 9, pp. 1760–1774, 2012.

[20] R. W. Mahley, "Central nervous system lipoproteins," *Arteriosclerosis, Thrombosis, and Vascular Biology*, vol. 36, no. 7, pp. 1305–1315, 2016.

[21] G. K. Smyth, "LIMMA: linear models for microarray data in bioinformatics and computational biology solutions using r and bioconductor," *Statistics for Biology and Health*, vol. 43, 2005.

[22] A. Kassambara, "ggpubr:"ggplot2" based publication ready plots," *R package version 0.1*, vol. 6, 2017.

[23] A. M. Newman, C. L. Liu, M. R. Green et al., "Robust enumeration of cell subsets from tissue expression profiles," *Nature Methods*, vol. 12, no. 5, pp. 453–457, 2015.

[24] 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.

[25] M. Ashburner, C. A. Ball, J. A. Blake et al., "Gene ontology: tool for the unification of biology," *Nature Genetics*, vol. 25, no. 1, pp. 25–29, 2000.

[26] M. Kanehisa and S. Goto, "KEGG: kyoto encyclopedia of genes and genomes," *Nucleic Acids Research*, vol. 28, no. 1, pp. 27–30, 2000.

[27] D. Szklarczyk, "The string database in 2017: quality-controlled protein–protein association networks, made broadly accessible," *Nucleic acids research*, vol. 15, Article ID gkw937, 2016.

[28] P. Shannon, "Cytoscape: a software environment for integrated models of biomolecular interaction networks," *Genome Research*, vol. 13, no. 11, pp. 2498–2504, 2003.

[29] K. A. Steinmetz and J. D. Potter, "Egg consumption and cancer of the colon and rectum," *European Journal of Cancer Prevention*, vol. 3, no. 3, pp. 237–246, 1994.

[30] R. J. Jacobs, P. W. Voorneveld, L. L. Kodach, and J. C. Hardwick, "Cholesterol metabolism and colorectal cancers," *Current Opinion in Pharmacology*, vol. 12, no. 6, pp. 690–695, 2012.

[31] X. Ma, E. Bi, C. Huang et al., "Cholesterol negatively regulates IL-9-producing CD8+ T cell differentiation and antitumor activity," *Journal of Experimental Medicine*, vol. 215, no. 6, pp. 1555–1569, 2018.

[32] S.-C. Tsai, M.-C. Sheen, and B.-H. Chen, "Association between HLA-DQA1, HLA-DQB1 and oral cancer," *The Kaohsiung Journal of Medical Sciences*, vol. 27, no. 10, pp. 441–445, 2011.

[33] M. Bianchini, E. Levy, C. Zucchini et al., "Comparative study of gene expression by cDNA microarray in human colorectal cancer tissues and normal mucosa," *International Journal of Oncology*, vol. 29, no. 1, pp. 83–94, 2006.

[34] M. Benevolo, M. Mottolese, G. Piperno et al., "HLA-A, -B, -C expression in colon carcinoma mimics that of the normal colonic mucosa and is prognostically relevant," *The American Journal of Surgical Pathology*, vol. 31, no. 1, pp. 76–84, 2007.

[35] C. Blum, A. Graham, M. Yousefzadeh et al., "The expression ratio of Map7/B2M is prognostic for survival in patients with stage II colon cancer," *International Journal of Oncology*, vol. 33, no. 3, pp. 579–584, 2008.

[36] G. Gualco, L. M. Weiss, and C. E. Bacchi, "Mum1/Irf4," *Applied Immunohistochemistry & Molecular Morphology*, vol. 18, no. 4, pp. 301–310, 2010.

[37] Q. Zhang, C. Zhang, J.-X. Ma, H. Ren, Y. Sun, and J.-Z. Xu, "Circular RNA PIP5K1A promotes colon cancer development through inhibiting miR-1273a," *World Journal of Gastroenterology*, vol. 25, no. 35, pp. 5300–5309, 2019.

[38] M. L. Slattery, A. Lundgreen, S. A. Kadlubar, K. L. Bondurant, and R. K. Wolff, "JAK/STAT/SOCS-signaling pathway and colon and rectal cancer," *Molecular Carcinogenesis*, vol. 52, no. 2, pp. 155–166, 2013.

[39] X. Guan, Y. Yi, Y. Huang et al., "Revealing potential molecular targets bridging colitis and colorectal cancer based on multidimensional integration strategy," *Oncotarget*, vol. 6, no. 35, pp. 37600–37612, 2015.

[40] M. V. Agrez, "Cell adhesion molecules and colon cancer," *ANZ Journal of Surgery*, vol. 66, no. 12, pp. 791–798, 1996.