Original Article

MiR-195 restrains lung adenocarcinoma by regulating CD4+ T cell activation via the CCDC88C/Wnt signaling pathway: a study based on the Cancer Genome Atlas (TCGA), Gene Expression Omnibus (GEO) and bioinformatic analysis

Cheng Yuan[1], Liyang Xiang[1], Rui Bai[1], Kuo Cao[1], Yanping Gao[1], Xueping Jiang[1], Nannan Zhang[1], Yan Gong[2,3], Conghua Xie[1,4,5]

[1]Department of Radiation and Medical Oncology, [2]Department of Biological Repositories, [3]Human Genetics Resource Preservation Center of Hubei Province, Human Genetics Resource Preservation Center of Wuhan University, [4]Hubei Key Laboratory of Tumour Biological Behaviors, [5]Hubei Cancer Clinical Study Center, Zhongnan Hospital of Wuhan University, Wuhan 430071, China

Contributions: (I) Conception and design: C Yuan, C Xie; (II) Administrative support: Y Gong, C Xie; (III) Provision of study materials or patients: C Yuan, L Xiang, R Bai; (IV) Collection and assembly of data: C Yuan, L Xiang, R Bai, K Cao, Y Gao, X Jiang, N Zhang; (V) Data analysis and interpretation: C Yuan, L Xiang, R Bai; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Dr. Conghua Xie. Department of Radiation and Medical Oncology, Zhongnan Hospital of Wuhan University, 169 Donghu Road, Wuhan 430071, China. Email: chxie_65@whu.edu.cn; Dr. Yan Gong. Department of Biological Repositories, Zhongnan Hospital of Wuhan University, 169 Donghu Road, Wuhan 430071, China. Email: yan.gong@whu.edu.cn.

Background: To systematically identity microRNA signatures, as well as miRNA-gene axes, for lung adenocarcinoma (LUAD) and to explore the potential biomarkers and mechanisms associated with the LUAD immune responses.

Methods: LUAD-related data were obtained from the Gene Expression Omnibus (GEO) and The Cancer Genome Atlas (TCGA), and these data were then used to identify the differentially expressed miRNAs that were downregulated in tumor tissues. Summary receiver operating characteristic curve analysis, survival analysis and meta-analysis were applied to evaluate the clinical significance and diagnostic value of the identified miRNAs. The presumed targets of the integrated-signature miRNAs were identified via 3 different target prediction algorithms: TargetScan, miRDB and DIANA-TarBase. Immunologic signature gene sets were enriched by gene set enrichment analysis (GSEA). Tumor-infiltrating lymphocytes were profiled by the Tumor IMmune Estimation Resource (TIMER). After pathway enrichment analysis using the Gene Ontology and Kyoto Encyclopedia of Genes and Genomes databases, pathway-gene networks were constructed using Cytoscape software.

Results: After integrated analysis of 4 GEO data sets (GSE48414, GSE51853, GSE63805 and GSE74190) and TCGA databases, miR-195 was identified as a potential clinical diagnostic marker. A total of 287 miR-195 target genes were screened, and 3 functional gene sets (GSE13485, GSE21379 and GSE29164) were enriched. GSE21379 was associated with the upregulation of CD4+ T cells in tumors, and the core genes were validated via the TIMER database. The CCDC88C expression level was significantly correlated with CD4+ T cell activation (partial.cor =0.437, P<0.001). Enrichment analysis revealed that CCDC88C was significantly enriched in the Wnt signaling pathway.

Conclusions: MiR-195, as a suppressor of lung adenocarcinoma, regulates CD4+ T cell activation via CCDC88C.

Keywords: miR-195; immunologic signatures; bioinformatic analysis; CCDC88C; lung adenocarcinoma

Submitted Jan 22, 2019. Accepted for publication May 10, 2019.
doi: 10.21037/atm.2019.05.54
View this article at: http://dx.doi.org/10.21037/atm.2019.05.54

Introduction

Non-small cell lung cancer (NSCLC), which accounts for 85% of all lung cancer cases, is one of the leading causes of cancer-related deaths worldwide (1,2). More specifically, lung adenocarcinoma (LUAD) is the most common subtype of NSCLC. Despite improvements in early disease detection and the development of chemotherapeutic and targeted treatments, the overall survival rate of LUAD patients remains poor (2). In recent years, immunotherapy has attracted increasing attention from oncologists. T cells are important mediators of tumor immunity, and in most types of solid tumors, T cell infiltration is a favorable prognostic marker (3,4). Immunotherapy to boost T cell functionality in tumors is rapidly becoming established as a standard treatment (5), and the immunotherapy focus has been on recruiting tumor infiltrating T cells (6). CD4+ T cells secrete a variety of cytokines that have direct effector functions and activate other immune cells (such as B cells, dendritic cells and even CD8 T Cells) (7,8). In lung cancer, tumor-infiltrating CD4+ T cells plays an essential role in the immune response (9). CD4+ T cells affect tumors by allowing CD8+ T cells entry to tumor sites (10) and infected mucosa (11); furthermore, they are also required for the inhibition of angiogenesis at tumor sites (12).

MicroRNAs (miRNAs) are small noncoding RNAs that regulate gene expression by degrading or inhibiting translation of their target transcripts, thereby affecting processes such as cell proliferation, differentiation and apoptosis (13). Changes in miRNA expression were reported as biomarkers for LUAD risk and prognosis (14), and miRNA-based biomarkers with prognostic or predictive potential for tumor responsiveness to immunocheckpoint inhibitors were recently described (15). Cell-specific miRNA expression patterns and the roles of miRNAs in the LUAD microenvironment have not been fully elucidated. Patients with similar clinical features often have different outcomes, suggesting an underlying relationship between LUAD development and genetic variations. The identification of new specific biomarkers that can be used to monitor tumor progression and treatment sensitivity, as well as to predict patient survival, will help overcome these challenges and improve outcomes in LUAD patients (16,17).

Gene expression profiling has become a new and effective method to identify prognostic markers and molecular targets for therapies (18). Dysregulated miRNAs in LUAD can be identified using miRNA expression profiling. The aim of our study was to use bioinformatic analysis of a large clinical dataset to systematically identity microRNA signatures, as well as miRNA-gene axes, related to LUAD and to explore potential biomarkers and mechanisms associated with LUAD immune responses.

Methods

Microarray profiles from the Gene Expression Omnibus (GEO) database

LUAD-related microarray profiles (up to November 2018) were obtained from the GEO database (http://www.ncbi.nlm.nih.gov/geo/). The search criterion of GEO Databases was shown in *Table S1*. Microarrays that met the following criteria were collected: (I) studies including at least 60 samples and (II) examination of miRNA expression in both cancerous tissue and adjacent noncancerous tissue from LUAD patients. Microarrays without useful data for analysis were excluded. Differentially expressed miRNAs (DEMs) between LUAD cancerous tissue and adjacent noncancerous tissue samples in each GEO dataset were ranked by the Robust Multi-Array Average and Linear Models for Microarray package and annotated by converting the different probe IDs to gene IDs.

Integrated analysis of miRNA expression datasets

The RobustRankAggreg (RRA) package was used to identify DEMs between LUAD cancerous tissue and adjacent noncancerous tissue samples. The adjusted P value and Log_2-fold change (FC) were specified as 0.05 and 1, respectively. One-sided test was applied to classify the downregulated DEMs. We selected the top 10 significantly downregulated DEMs for further studies.

miRNA-seq data from The Cancer Genome Atlas (TCGA) database

Publicly available miRNA-seq data on miRNA levels in LUAD cancer tissue and adjacent noncancerous tissue samples were directly downloaded from the TCGA data portal (http://cancergenome.nih.gov/). We obtained the miRNA profiles of 209 LUAD cancer tissue samples and 45 adjacent noncancerous tissue samples together with the clinical information (level 3) of the corresponding patients. DEMs between the LUAD samples with pathological stages of I–IV and adjacent noncancerous tissue samples were identified by calculating the FC ($|\log_2(FC)| > 2$ and adjusted

P value <0.05) with the R package edgeR.

Integrated analysis the GEO profiles and the TCGA miRNA-seq data

The top 10 DEMs identified as significantly downregulated in the GEO database were entered into the TCGA database for further verification. DEMs that showed consistent expression in GEO were selected for statistical analysis. Independent Student's *t*-tests were performed to calculate the differences in the miRNA levels between LUAD cancerous tissue and adjacent noncancerous tissue. P<0.05 was considered statistically significant.

Diagnosis and prognosis analysis

A receiver operating characteristic (ROC) curve built on a univariate classification model based on the DEM expression profiles across independent TCGA datasets were used to predict LUAD. Kaplan-Meier plots of the overall survival for a discriminatory median DEM expression profile based on TCGA sequencing data were used to assess prognostic accuracy. P values were calculated using the log-rank test.

MiRNAs meeting the above diagnostic and prognostic criteria were introduced into multiple linear regression models for further analysis. The relative miRNA levels were treated as an independent variable, and the diagnosis results were treated as a dependent variable. A linear regression equation was constructed to identify miRNAs with independent diagnostic value.

Pairwise meta-analysis and diagnostic meta-analysis

A comprehensive meta-analysis was performed using Stata 14.0 software (Stata Corporation, College Station, TX), combining the TCGA data and GEO datasets. The pooled data in the meta-analysis were assessed by the standard mean difference (SMD) with a 95% confidential interval (CI). Heterogeneity among the eligible microarrays was evaluated by chi-squared and I-squared tests. The effect model was then determined according to the heterogeneity. Specifically, a fixed effects model was conducted for the meta-analysis when the heterogeneity was low (I^2≤50% and P>0.1), and a random effects model was selected if apparent heterogeneity existed (I^2>50% or P≤0.1). A bivariate-mixed model was used to estimate the ROC curve, and the area under curve (AUC) was also estimated to optimize cut-off points.

Target prediction and functional analysis of miRNA

The presumed targets of the integrated-signature miRNAs were identified by 3 different target prediction algorithms: TargetScan, miRDB and DIANA-TarBase. Unique genes with target sites in 3' UTR sequences were included. To assess the possible functions, we searched the Gene Ontology (GO) database, the Kyoto Encyclopedia of Genes and Genomes (KEGG) database and the Database for Annotation, Visualization and Integrated Discovery (DAVID). A P value less than 0.01 was defined as the cutoff criterion for KEGG pathways enriched in the target gene set.

Gene set enrichment analysis (GSEA)

The enrichment analyses for immunologic signature gene sets were conducted with GSEA v3.0 for the target genes. The enriched pathways were arranged in the order of their normalized enrichment scores (NESs).

Immunocyte infiltration in the tumor microenvironment

The core enriched genes have been packaged into the web-accessible resource TIMER (Tumor IMmune Estimation Resource; https://cistrome.shinyapps.io/timer/), to enable further exploration of the impacts of the core enriched genes on immunocyte infiltration in tumor microenvironments.

Results

Collection of microarray datasets from GEO

The flow chart for the study selection for this integrated analysis is shown in *Figure 1*. We searched the GEO database, and the GEO microarrays can be regarded as a training dataset to screen for DEMs in LUAD. Finally, 4 GEO datasets (accession numbers GSE48414, GSE51853, GSE63805 and GSE74190) were included in the present study, and the characteristics of the studies based on the GEO dataset are presented *Table 1*.

Due to the heterogeneity in the sample types in the GSE microarrays, the common DEMs were examined separately (*Figure 2*). The downregulated DEMs in each GSE are presented in http://fp.amegroups.cn/cms/atm.2019.05.54-1.pdf. *Tables S2-S4*. There were inconsistencies in the DEMs obtained from each GSE microarray. Therefore, the RRA

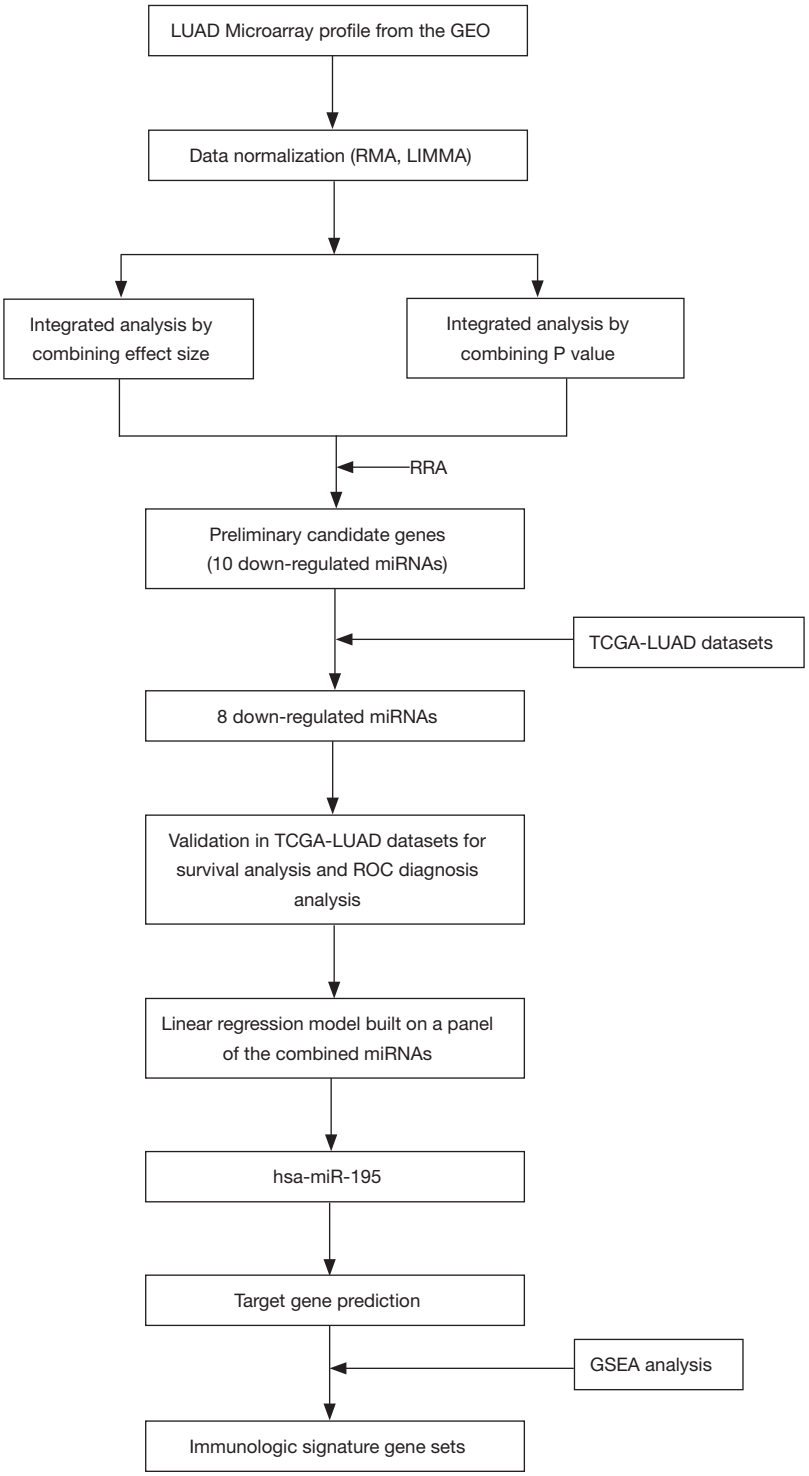

Figure 1 The flowchart of the integrated analysis and functional validation.

Table 1 Datasets used in finally quantitative synthesis and integrated analysis

No.	Author	Publication year	Country	Assay type	Tumor site	No. of samples (Pairs)	Sample size (normal/tumor)	Platform	Source accession	PMID
1	Bjaanæs	2014	Norway	Agilent-031181 Unrestricted_Human_ miRNA	LUAD	20	174 (154/20)	GPL16770	GSE48414	24599520
2	Tai	2015	Japan	Agilent-015508 Human miRNA Microarray	NSCLC	5	131 (126/5)	GPL7341	GSE51853	26483346
3	Robles	2016	USA	NanoString nCounter Human miRNA assay	LUAD	31	62 (31/31)	GPL18410	GSE63805	26134223
4	Lu	2015 (last update date)	China	Agilent-019118 Human miRNA Microarray	SCLC, LUAD, LUSC	44	126 (82/44)	GPL19622	GSE74190	–

LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma; SCLC, small cell lung carcinoma; NSCLC, non-small cell lung cancer; Pairs, tumor tissues and paired adjacent noncancerous tissues from the same patient.

package was used to perform an integrated analysis of the 4 GSE microarrays to identify co-downregulated DEMs. There were 27 significantly downregulated miRNAs. The hierarchical clustering of the top 10 miRNAs is shown in *Figure 3*.

Integrated-signature miRNAs showed clinical prognostic significance in LUAD patients

We further validated the top 10 downregulated DEMs in TCGA-LUAD samples (209 LUAD cancerous tissue samples and 45 adjacent noncancerous tissue samples, http://fp.amegroups.cn/cms/atm.2019.05.54-2.pdf). Only miR-195, miR-451, miR-144, miR-218, miR-133b, miR-145, miR-143 and miR-497 were significantly downregulated in LUAD tumors (*Figure 4*). The diagnostic efficiency and prognostic value of each of these miRNAs were estimated via ROC curve analysis and Kaplan-Meier survival analysis, respectively. Ultimately, we selected 3 miRNAs (miR-143, miR-195 and miR-218) with high diagnostic efficiency (AUC >0.8, P<0.05) and prognostic value (logrank P<0.05) (*Figure 5*). We next optimized the accuracy by using a linear regression model built on a panel of the combined miRNAs. By constructing the linear regression equation LUAD risk score = –0.0.02267miR-143-0.1115miR-195-0.04098miR-218+2.3699, miR-195 was identified as the most significant independent variable (P=0.0006, *Table 2*).

Combining the TCGA data and GEO datasets, the results of pairwise meta-analyses indicated that miR-195 was overexpressed in adjacent noncancerous tissue samples (SMD =–1.69, 95% CI: –1.92 to –1.46, *Figure 6*). Furthermore, the results of a diagnostic meta-analysis suggested that miR-195 offers high diagnostic efficiency (AUC =0.9180, *Figure 7A*; the pooled sensitivity =0.97, 95% CI: 0.95–0.99, *Figure 7B*; the pooled specificity =0.65, 95% CI: 0.57–0.72, *Figure 7C*).

Target gene prediction coupled with pathway analysis

To explore the biological mechanisms of miR-195 in LUAD, we performed target gene prediction coupled with pathway analysis. A total of 287 target genes (http://fp.amegroups.cn/cms/atm.2019.05.54-3.pdf) were identified via TargetScan, miRDB and DIANA-TarBase, and these genes were then subjected to GO and KEGG analyses. The results of the GO term analysis included the biological process (BP), cellular component (CC) and molecular function (MF) groups. The target genes were mainly enriched in protein binding, beta-catenin binding, ubiquitin protein ligase activity and activin binding in the MF group; nucleoplasm, cytoplasm, cytosol and nucleus in the CC group; and protein phosphorylation and Wnt signaling pathway in the BP group (*Table S5* and *Figure 8*). The results of the GO and KEGG analysis indicated that the most significantly enriched terms were “protein binding” and “cell cycle”. The top 50 genes with significant differences in their expression levels are shown along with their functions in *Figure 8B*. All of the target genes were analyzed using the KEGG pathway website and the clusterProfiler package of

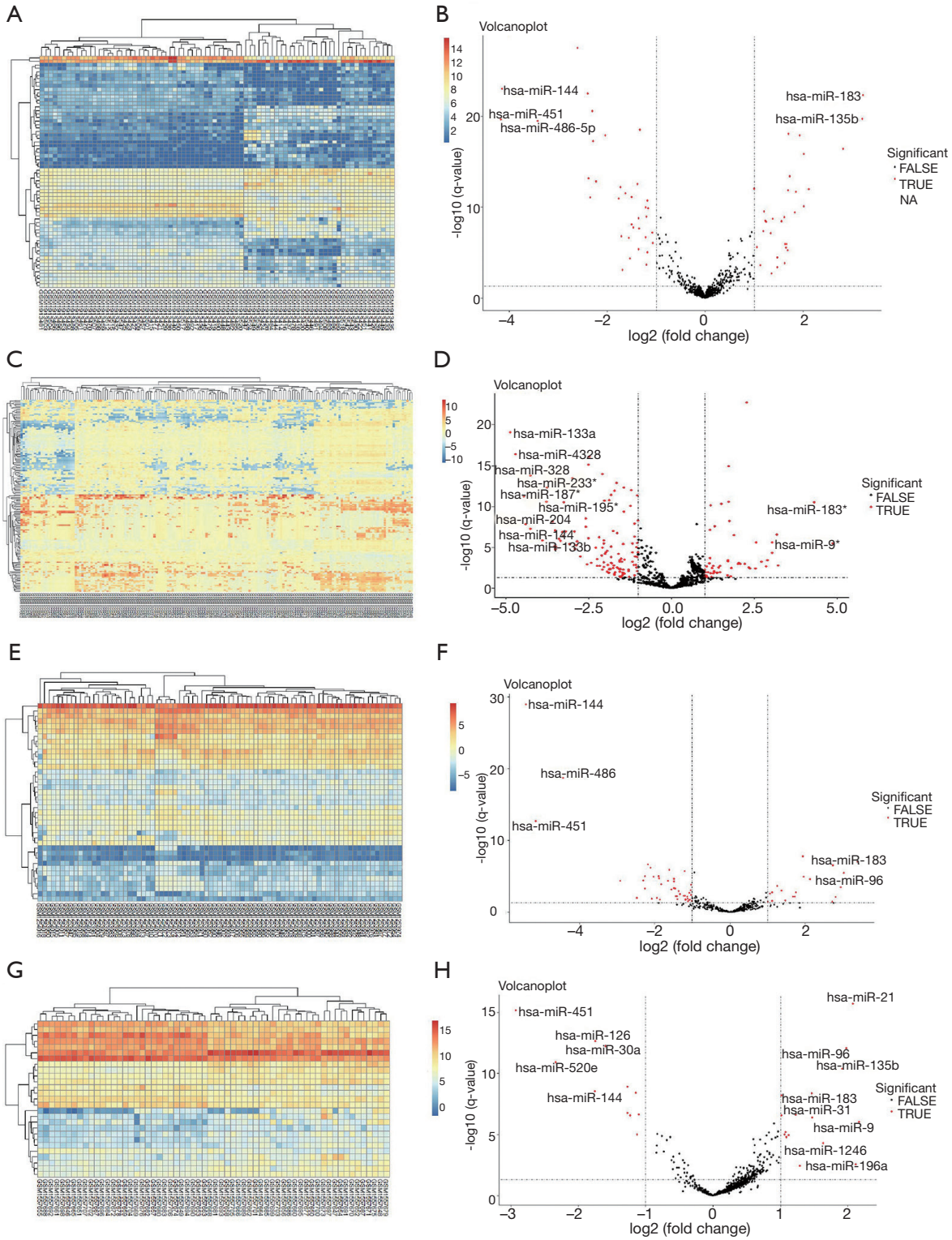

Figure 2 Clustering of the genes in LUAD cancerous tissue samples *vs.* adjacent noncancerous tissue samples across each independent dataset. Each column represents a sample and each row represents the expression level of a miRNA. The color scale represents the raw Z score, ranging from blue (low expression) to red (high expression). The red dots in the volcano plot represent the miRNAs that are significantly different. (A and B) GSE48414; (C and D) GSE51853; (E and F) GSE63805; (G and H) GSE74190.

Figure 3 Clustering integrated analysis of the downregulated DEMs in the expression datasets by RobustRankAggreg package.

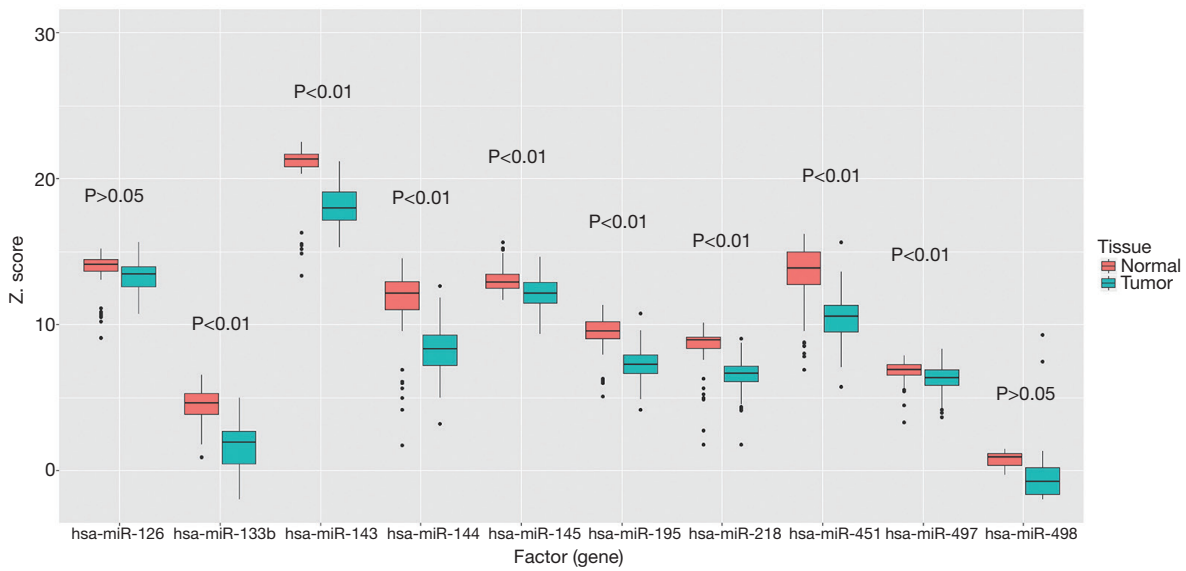

Figure 4 Plots of the expression levels of the downregulated miRNAs in tumor and normal tissue samples (TCGA dataset). The expression values of the miRNAs are log2-transformed.

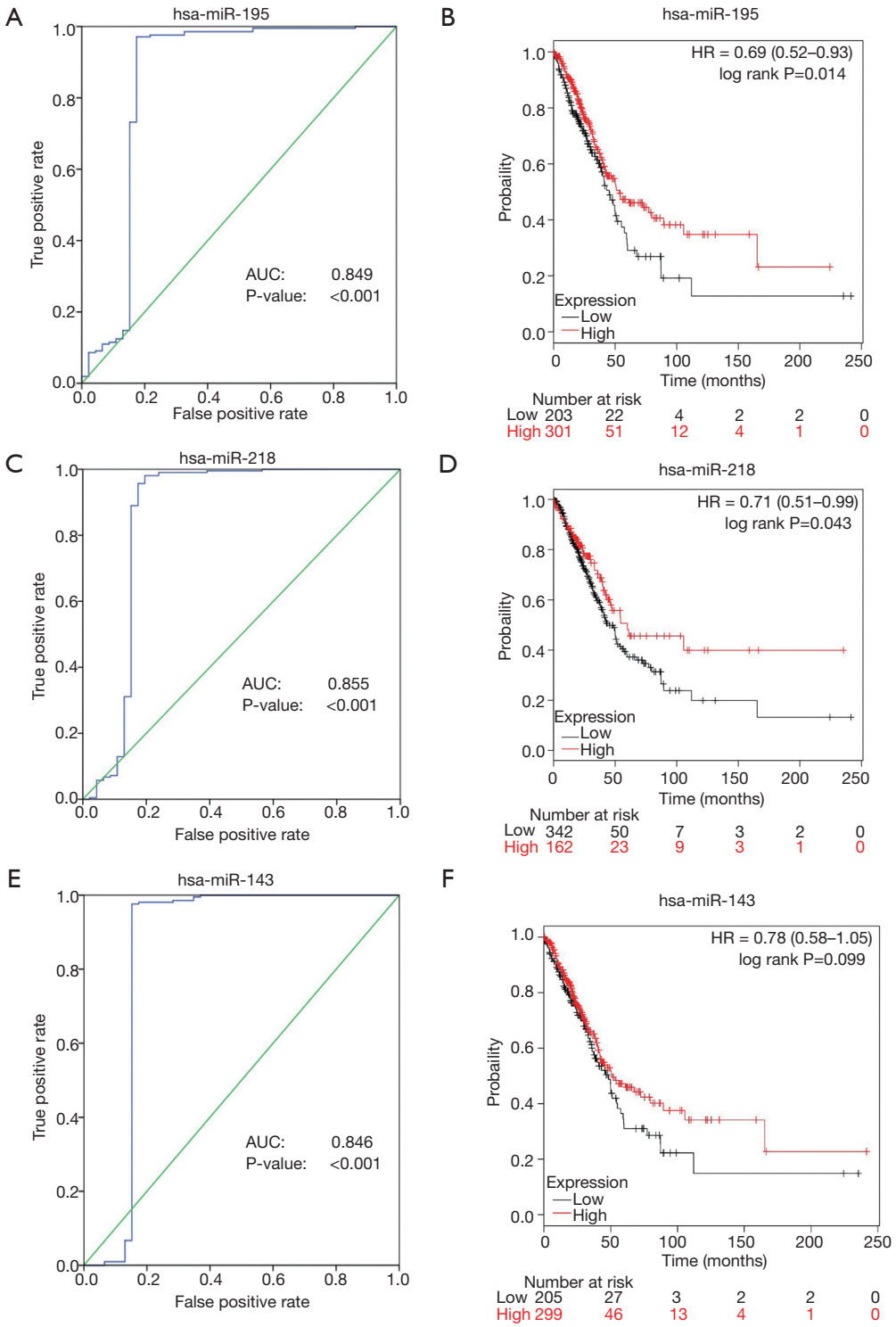

Figure 5 Diagnostic analysis and survival analysis. The ROC curve was built on a univariate classification model based on miRNA expression levels across independent TCGA datasets to predict LUAD. Kaplan-Meier plots of overall survival for a discriminatory median DEM expression profile, based on TCGA sequencing data, to assess prognostic accuracy. The P values were calculated using the log-rank test.

Table 2 LUAD risk score was built using a linear regression model

Independent variables	Coefficient	Std. Error	t	P
miR-143	−0.02267	0.02111	−1.074	0.2840
miR-195	−0.1115	0.03192	−3.491	0.0006
miR-218	−0.04098	0.02263	−1.811	0.0714

LUAD, lung adenocarcinoma. Linear regression equation: Y =−0.02267miR-143 −0.1115miR-195 −0.04098miR-218 +2.3699.

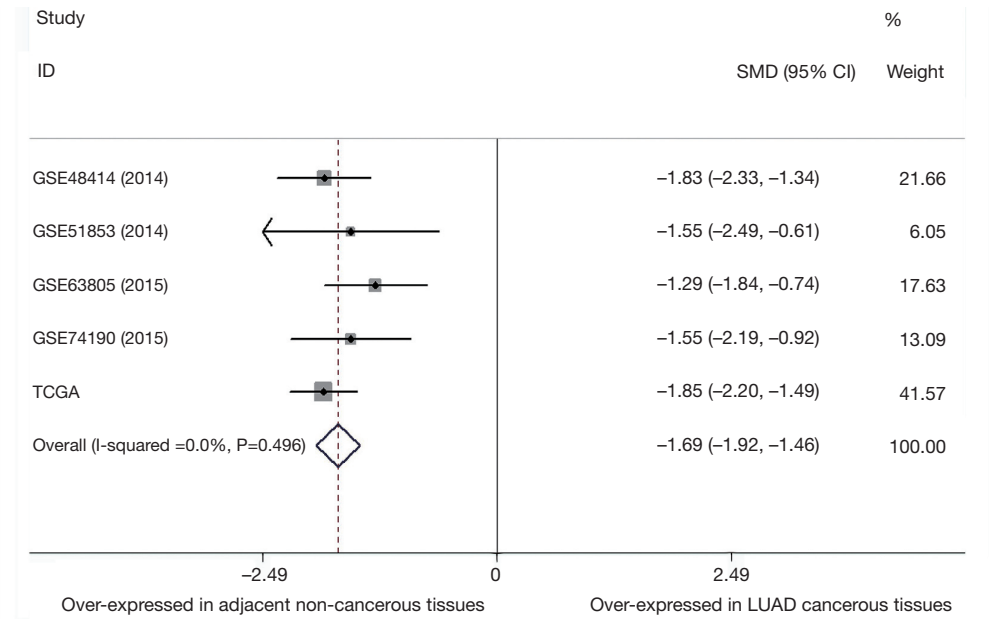

Figure 6 Forest plots summarizing miR-195 downregulation in the 5 datasets in the integrated analysis. Each row represents a study with a standardized mean difference (SMD) between LUAD and normal tissue, and the 95% confidence interval (CI) is shown. The size of the gray box is proportional to the relative effects of each dataset. The dotted vertical line at 0 represents the null hypothesis. The diamonds represent the overall, combined mean difference for miR-195. Thus, negative values indicate the downregulation of miR-195 in LUAD.

the R software, and only genes with P values less than 0.01 were included. The largest number genes were enriched in the PI3K-Akt signaling pathway (*Figure 9*). The specific links between each gene and its function are shown in *Figure 9C*.

GSEA in immunologic signature gene sets

To characterize the potential mechanisms of immunologic function associated with the miR-195 target genes, GSEA was used to obtain the biological processes enriched in immunologic signature gene sets. Then, 3 functional gene sets were enriched (GSE13485, GSE21379 and GSE29164, *Figure 10*), and they were all upregulated in the tumor tissue samples. The core genes of the 3 immunologic signature gene sets are shown in *Tables S6-S8*.

GSE21379 was associated with upregulation of CD4+ T cells in tumors, and the core enrichment genes were validated via the TIMER database. The correlations between the expression levels of 14 genes (*OSBPL3*, *IVNS1ABP*, *USP42*, *VEGFA*, *BAG4*, *GGA3*, *BTRC*, *CCDC88C*, *NOTCH2*, *MAFK*, *CAMSAP1*, *PRKAR2A*, *MOB4*, *DDX3Y* and *FRYL*) and CD4+ T cell infiltration were examined. The expression levels of CCDC88C were significantly correlated with CD4+ T cell activation (partial. cor =0.437, P<0.001, *Figure S1*).

Discussion

The present study, based on GEO and TCGA analysis,

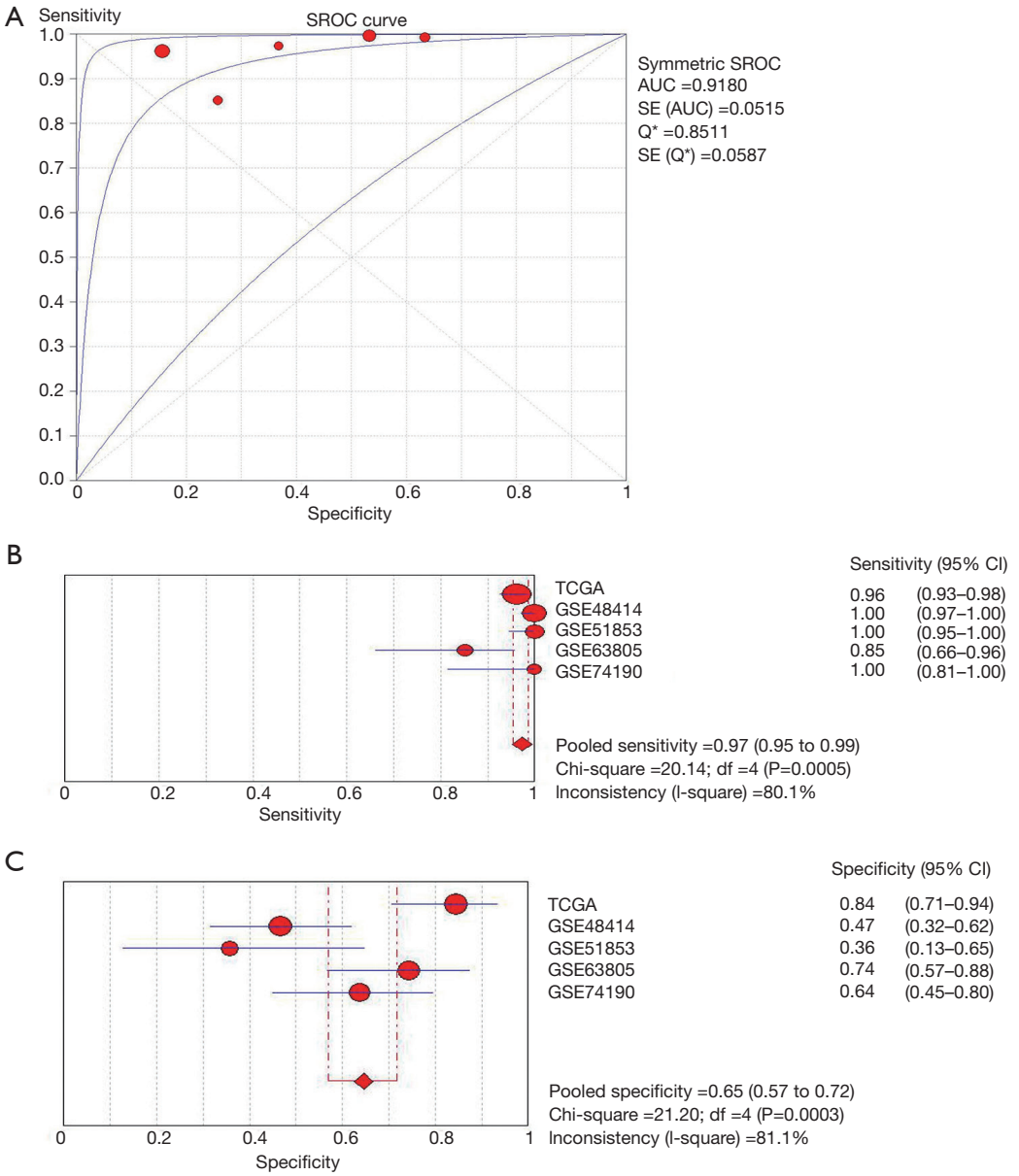

Figure 7 Diagnostic meta-analysis. (A) SROC curves for miR-195 in LUAD diagnosis. (B) The sensitivity of miR-195 (pooled sensitivity =0.97, 95% CI: 0.95–0.99). (C) The specificity of miR-195 (pooled specificity =0.65, 95% CI: 0.57–0.72).

revealed that miR-195 was overexpressed in adjacent noncancerous tissue samples, and that it had high diagnostic efficiency. These results are consistent with those of a previous study (19) that showed that miR-195 suppressed tumor cell growth, migration and invasion and was associated with better survival outcomes in LUAD patients.

Nevertheless, most previous basic studies focused on one miRNA-195 target gene, i.e., *CHEK1* (19), *IRS1* (20), or *MMP14* (21). Our KEGG pathway analysis found that the largest number genes were enriched in the PI3K-Akt signaling pathway, including the following genes: *CCNE1*, *FGF2*, *PIK3R1*, *AKT3*, *RPS6KB1*, *PHLPP2*, *ITGA2*, *YWHAQ*, *YWHAH*, *CCND1*, *PRKAA1*, *MYB*, *RAF1*, *INSR*, *VEGFA*, *LAMC1* and *CHUK*. Because miRNAs are mainly negative regulators of their target genes, these upregulated genes (*CCNE1*, *RPS6KB1*, *ITGA2*, *YWHAQ*, *PRKAA1*, *INSR*, *VEGFA*, *LAMC1* and *CHUK*) should be given attention in future studies of LUAD. The PI3K-

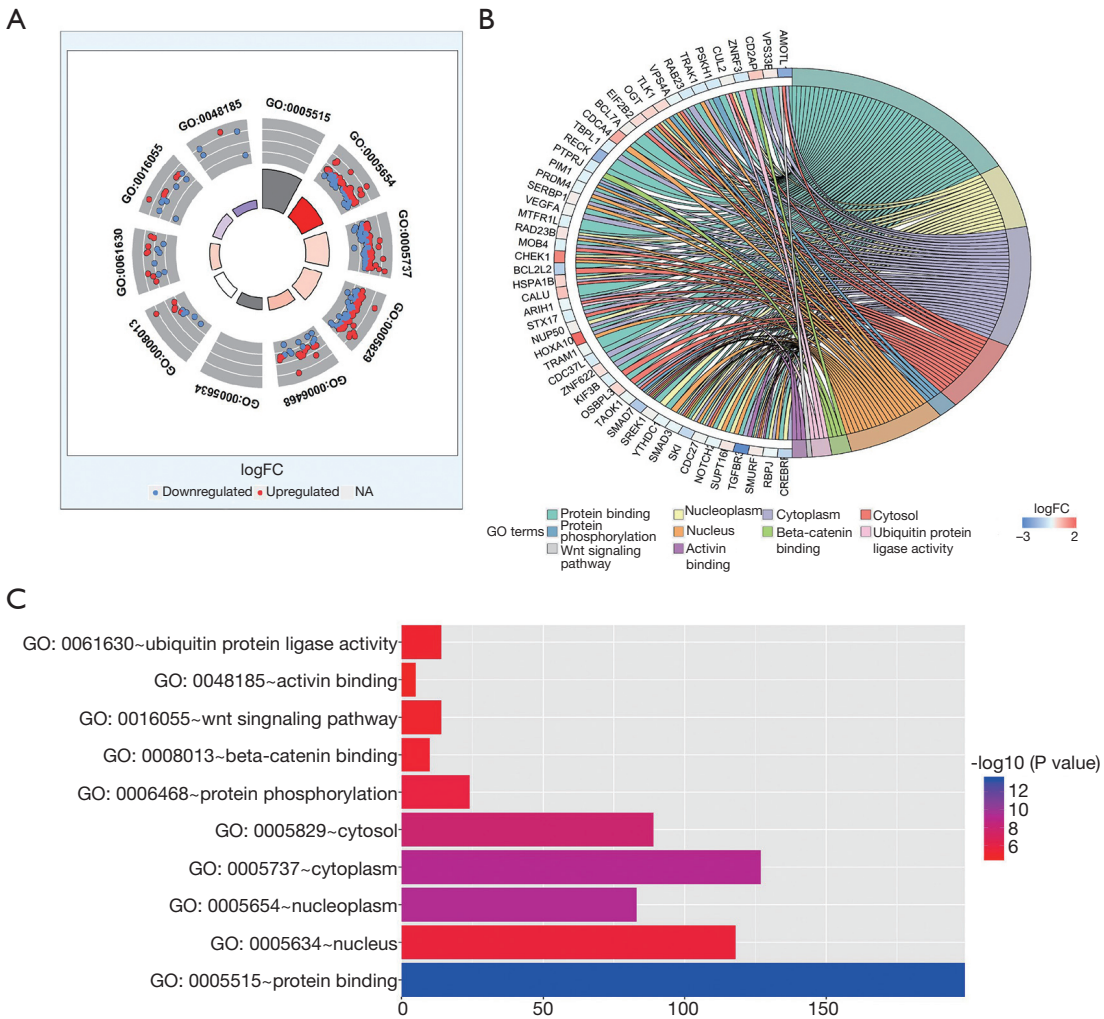

Figure 8 Gene Ontology terms of 287 overlapping miR-195 target genes. (A) Each point represents a gene, and the colors represent the expression level (red indicates upregulated expression and blue indicates downregulated expression). (B) The top 50 genes identified via functional enrichment with their Log FC values. (C) Significantly enriched GO terms of the miR-195 target genes based on their functions.

Akt signaling pathway is essential for maintaining cell growth, survival, death and metabolism, and it is commonly activated during cancer initiation and progression (22). In addition, the PI3K-Akt signaling pathway can regulate the proliferation, migration, invasion, apoptosis and angiogenesis of lung cancer cells (23), and activation of the PI3K-Akt signaling pathway may be a therapeutic molecular target for lung cancer (24). In addition to the relationship between *VEGFA* and the PI3K-Akt signaling pathway in lung cancer (25), the biological behaviors of lung cancer involving the PI3K-Akt signaling pathway remain to be investigated.

The GSEA-based identification of an immunologic signature gene set was an important objective of this study. GSE13485 was mainly related to a vaccine response, and GSE29164 was based on data collected during immunotherapy for melanoma. Therefore, we focused on the relationship between the genes and immune processes contained in GSE21379. As shown in *Table S7*, 15 core genes (*OSBPL3*, *IVNS1ABP*, *USP42*, *VEGFA*, *BAG4*, *GGA3*, *BTRC*, *CCDC88C*, *NOTCH2*, *MAFK*, *CAMSAP1*, *PRKAR2A*, *MOB4*, *DDX3Y* and *FRYL*) were involved in the upregulation of CD4+ T cells in tumor tissue. Previous studies (26,27) showed that CD4+ T cells induced cytotoxic programming of CD8+ T cells, which then suppress tumor growth via IFN-γ secretion or direct killing of the

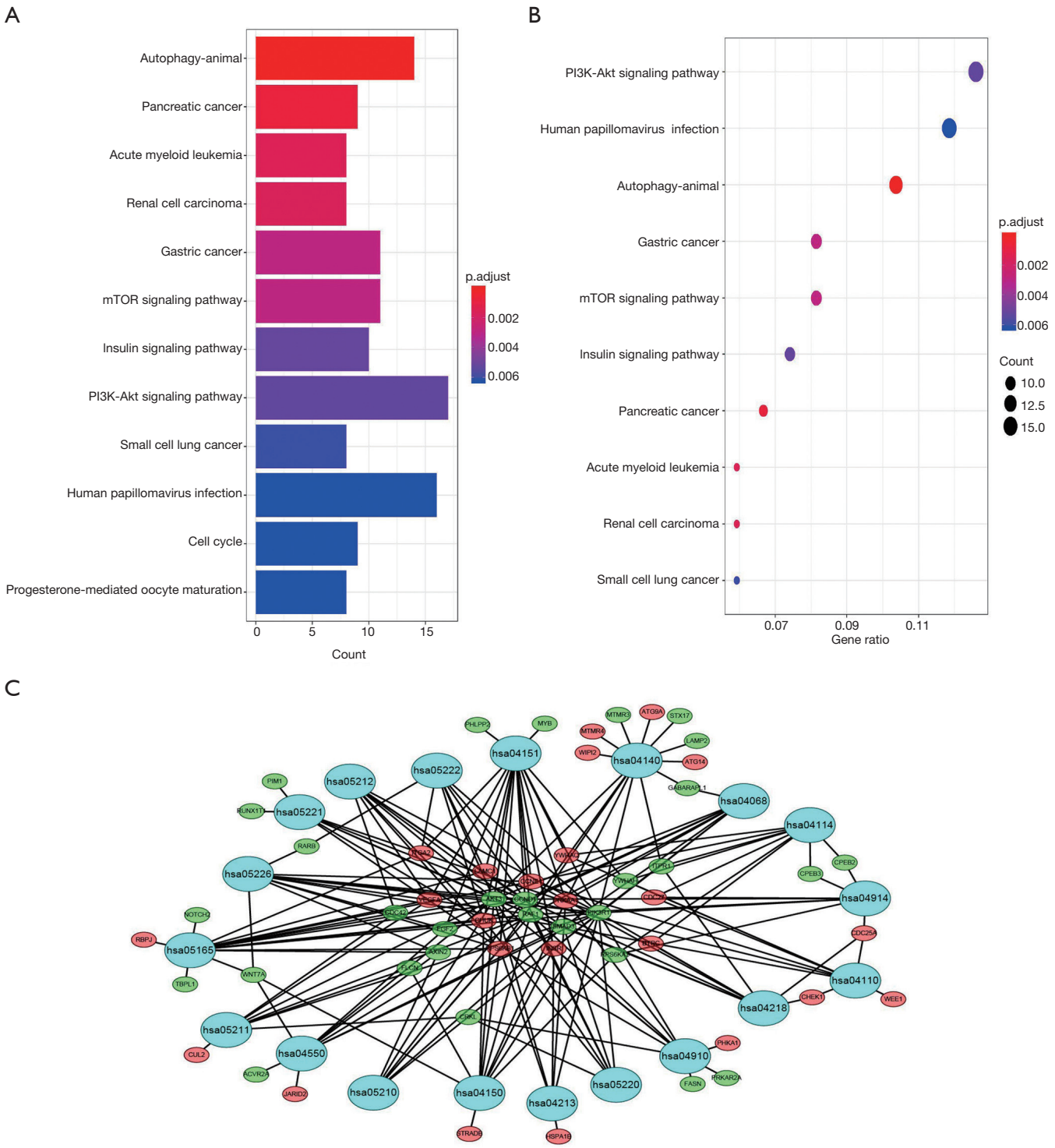

Figure 9 Significant signaling pathway analysis of miR-195 target genes performed with the KEGG pathway website and R software packages. (A and B) Representative dot plots of the pathway enrichment analysis of the miR-195 target genes. Gene ratio = count/set size. (C) The relationship between the genes and KEGG pathways. The red, green and blue circles denote upregulated genes, downregulated genes and the KEGG pathway ID, respectively.

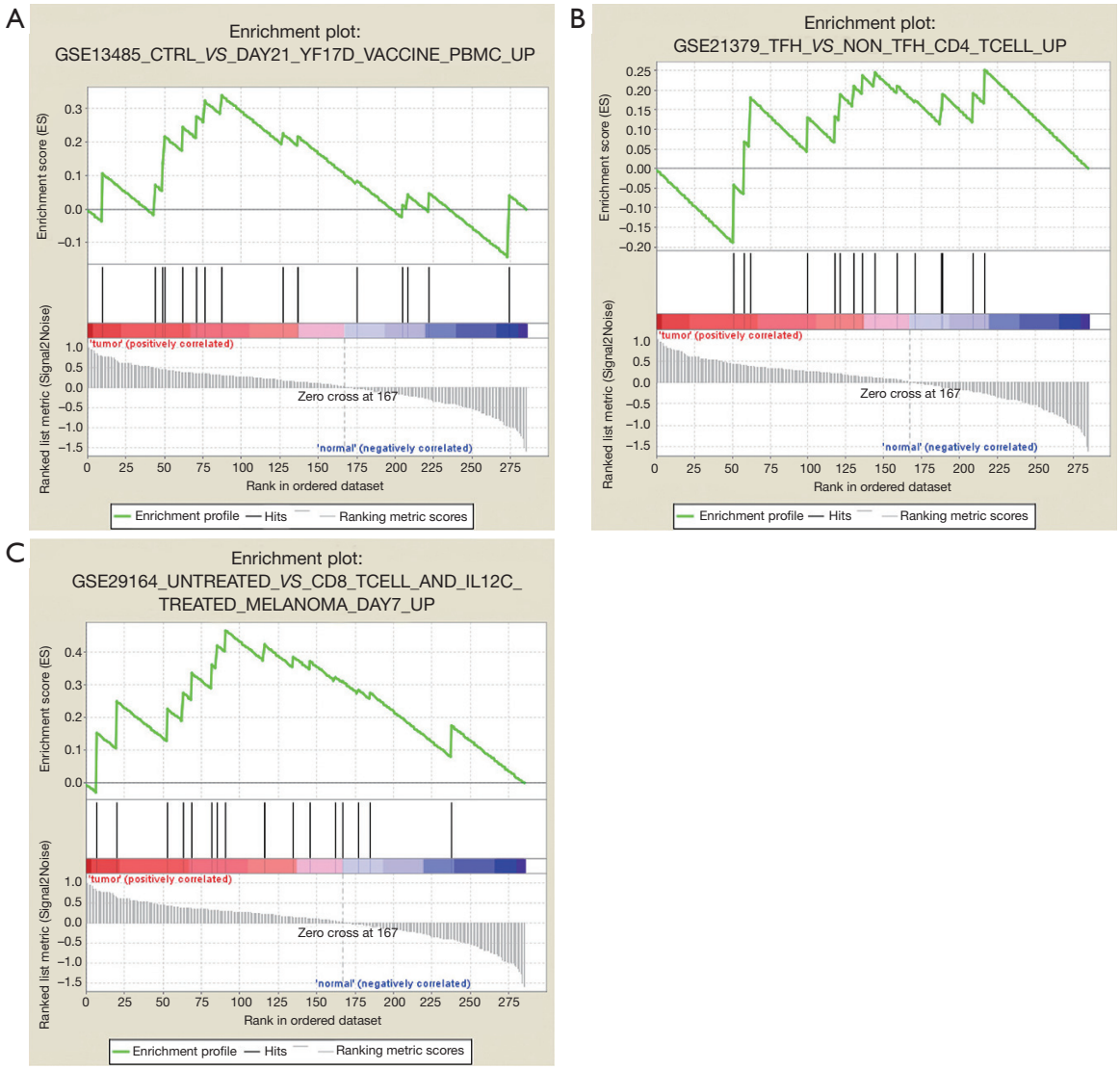

Figure 10 Gene set enrichment analysis (GSEA) of TCGA dataset immunologic signature gene sets.

tumor cells (28,29). However, the effects of CD4+ T cell infiltration on the biological behaviors of tumors are not consistent. The presence of CD4+ T cells in the tumor microenvironment was linked to poor outcomes in prostate cancer patients (30) as well as in patients with renal cell carcinoma (31). CD4+ T cells recruited in mammary cancer enhanced metastasis (32). Among the immunologic signature gene sets in GSE13485, NOTCH2 (33-35) and VEGFA (36) had significant effects on CD4+ T cells. However, the immune regulatory mechanisms of the other genes in tumors and lung cancer are not fully elucidated. Therefore, our study provides a clue for studying the genetic regulation of CD4+ T cells and lung cancer immunity.

Our results also demonstrated that the CCDC88C expression level was significantly correlated with CD4+ T cell activation. Enomoto *et al.* (37) found that CCDC88C (coiled-coil domain containing 88C) encodes a member of the hook-related proteins involved in the regulation of the Wnt signaling pathway. These results are consistent with our enrichment analysis. Furthermore, the Wnt signaling pathway controls inflammatory responses induced by multiple factors, such as pathogenic bacteria via Toll-like receptors (38,39), and it might be involved in the impaired T-cell homeostasis present in a variety of immune system diseases, such as rheumatoid arthritis and

systemic lupus erythematosus (40). Inhibition of the Wnt signaling pathway enhanced CD4+ T cell infiltration into the central nervous system by increasing the expression of vascular cell adhesion molecule-1 and the transcytosis protein Caveolin-1, as well as by promoting endothelial transcytosis (41). A previous study (42) showed that both Wnt3a and β-catenin were overexpressed by tumor-infiltrating and nontumor-infiltrating CD4+ or CD8+ T cells. Wnt3a blockade inhibited the differentiation of naive T cells but could not rescue the dysfunction of differentiated T cells in the tumor environment. The canonical Wnt signaling pathway blocks T cell differentiation and plays an important role in phenotypic maintenance of naive T cells and stem cell-like memory T cells in human peripheral blood (43); however, its effects on tumor-infiltrating lymphocytes in non-small cell lung cancer are still unclear. Based on the results of our bioinformatic analysis and previous literature reports, we conclude that CCDC88C might regulate CD4+ T cell activation via the Wnt signaling pathway.

However, this conclusion should be treated with caution. The GO enrichment analysis showed that CCDC88 was enriched in the Wnt signaling pathway, but this pathway was not significantly enriched in the KEGG results. In general, the biological process results from the GO analysis have many similar functions to those identified via the KEGG pathway analysis. Since the two types of enrichment analysis are based on different databases, there may be some inconsistencies in the results. However, this inconsistency could represent a cross-complement that provides verification of the two methods.

The tumor microenvironment, with its individual immune cells, may play key roles in tumor progression. Cancer development is driven by the accumulation of random mutations that lead to increased dysregulation of several key pathways. Therefore, it is very important to use bioinformatics approaches to identify key genes that shape tumor immune microenvironments.

Acknowledgments

Funding: This work was supported by grants from the Chinese National Natural Science Foundation (grant No. 81572967, 81372498, and 81800429), Hubei Natural Science Foundation (grant No. 2013CFA006), and Zhongnan Hospital of Wuhan University Science, Technology and Innovation Seed Fund (grant No. znpy2016050, znpy2017001, and znpy2017049), National key clinical speciality construction program of China [No. (2013)544], Wuhan City Huanghe Talents Plan and the Fundamental Research Funds for the Central Universities (grant No. 2042018kf0065).

Footnote

Conflicts of Interest: The authors have no conflicts of interest to declare.

References

1. Blandin Knight S, Crosbie PA, Balata H, et al. Progress and prospects of early detection in lung cancer. Open Biol 2017;7:170070.
2. Siegel RL, Miller KD, Jemal A. Cancer Statistics, 2017. CA Cancer J Clin 2017;67:7-30.
3. Fridman WH, Pagès F, Sautès-Fridman C, et al. The immune contexture in human tumours: impact on clinical outcome. Nat Rev Cancer 2012;12:298-306.
4. Becht E, Giraldo NA, Dieu-Nosjean MC, et al. Cancer immune contexture and immunotherapy. Curr Opin Immunol 2016;39:7-13.
5. Oja AE, Piet B, van der Zwan D, et al. Functional Heterogeneity of CD4+ Tumor-Infiltrating Lymphocytes With a Resident Memory Phenotype in NSCLC. Front Immunol 2018;9:2654.
6. Bruno TC, Ebner PJ, Moore BL, et al. Antigen-Presenting Intratumoral B Cells Affect CD4+ TIL Phenotypes in Non-Small Cell Lung Cancer Patients. Cancer Immunol Res 2017;5:898-907.
7. Kamphorst AO, Ahmed R. CD4 T-cell immunotherapy for chronic viral infections and cancer. Immunotherapy 2013;5:975-87.
8. Swain SL, McKinstry KK, Strutt TM. Expanding roles for CD4$^+$ T cells in immunity to viruses. Nat Rev Immunol 2012;12:136-48.
9. Hiraoka K, Miyamoto M, Cho Y, et al. Concurrent infiltration by CD8+ T cells and CD4+ T cells is a favourable prognostic factor in non-small-cell lung carcinoma. Br J Cancer 2006;94:275-80.
10. Bos R, Sherman LA. CD4+ T-cell help in the tumor milieu is required for recruitment and cytolytic function of CD8+ T lymphocytes. Cancer Res 2010;70:8368-77.
11. Nakanishi Y, Lu B, Gerard C, et al. CD8(+) T lymphocyte mobilization to virus-infected tissue requires CD4(+) T-cell help. Nature 2009;462:510-3.
12. Rakhra K, Bachireddy P, Zabuawala T, et al.

CD4(+) T cells contribute to the remodeling of the microenvironment required for sustained tumor regression upon oncogene inactivation. Cancer Cell 2010;18:485-98.
13. Pritchard CC, Kroh E, Wood B, et al. Blood cell origin of circulating microRNAs: a cautionary note for cancer biomarker studies. Cancer Prev Res (Phila) 2012;5:492-7.
14. Świtlik WZ, Szemraj J. Circulating miRNAs as non-invasive biomarkers for non-small cell lung cancer diagnosis, prognosis and prediction of treatment response. Postepy Hig Med Dosw (Online) 2017;71:649-62.
15. Halvorsen AR, Sandhu V, Sprauten M, et al. Circulating microRNAs associated with prolonged overall survival in lung cancer patients treated with nivolumab. Acta Oncol 2018;57:1225-31.
16. Vargas AJ, Harris CC. Biomarker development in the precision medicine era: lung cancer as a case study. Nat Rev Cancer 2016;16:525-37.
17. Neal JW, Gainor JF, Shaw AT. Developing biomarker-specific end points in lung cancer clinical trials. Nat Rev Clin Oncol 2015;12:135-46.
18. Khan J, Wei JS, Ringnér M, et al. Classification and diagnostic prediction of cancers using gene expression profiling and artificial neural networks. Nat Med 2001;7:673-9.
19. Liu B, Qu J, Xu F, et al. MiR-195 suppresses non-small cell lung cancer by targeting CHEK1. Oncotarget 2015;6:9445-56.
20. Wang Y, Zhang X, Zou C, et al. miR-195 inhibits tumor growth and angiogenesis through modulating IRS1 in breast cancer. Biomed Pharmacother 2016;80:95-101.
21. Li M, Ren CX, Zhang JM, et al. The Effects of miR-195-5p/MMP14 on Proliferation and Invasion of Cervical Carcinoma Cells Through TNF Signaling Pathway Based on Bioinformatics Analysis of Microarray Profiling. Cell Physiol Biochem 2018;50:1398-413.
22. Koundouros N, Poulogiannis G. Phosphoinositide 3-Kinase/Akt Signaling and Redox Metabolism in Cancer. Front Oncol 2018;8:160.
23. Zhou Y, Li S, Li J, et al. Effect of microRNA-135a on Cell Proliferation, Migration, Invasion, Apoptosis and Tumor Angiogenesis Through the IGF-1/PI3K/Akt Signaling Pathway in Non-Small Cell Lung Cancer. Cell Physiol Biochem 2017;42:1431-46.
24. Matsuda S, Nakagawa Y, Kitagishi Y, et al. Reactive Oxygen Species, Superoxide Dimutases, and PTEN-p53-AKT-MDM2 Signaling Loop Network in Mesenchymal Stem/Stromal Cells Regulation. Cells 2018;7:E36.
25. Chen CH, Lai JM, Chou TY, et al. VEGFA upregulates FLJ10540 and modulates migration and invasion of lung cancer via PI3K/AKT pathway. PLoS One 2009;4:e5052.
26. Ahrends T, Bąbała N, Xiao Y, et al. CD27 Agonism Plus PD-1 Blockade Recapitulates CD4+ T-cell Help in Therapeutic Anticancer Vaccination. Cancer Res 2016;76:2921-31.
27. Ahrends T, Spanjaard A, Pilzecker B, et al. CD4+ T Cell Help Confers a Cytotoxic T Cell Effector Program Including Coinhibitory Receptor Downregulation and Increased Tissue Invasiveness. Immunity 2017;47:848-61.e5.
28. Quezada SA, Simpson TR, Peggs KS, et al. Tumor-reactive CD4(+) T cells develop cytotoxic activity and eradicate large established melanoma after transfer into lymphopenic hosts. J Exp Med 2010;207:637-50.
29. Friedman KM, Prieto PA, Devillier LE, et al. Tumor-specific CD4+ melanoma tumor-infiltrating lymphocytes. J Immunother 2012;35:400-8.
30. McArdle PA, Canna K, McMillan DC, et al. The relationship between T-lymphocyte subset infiltration and survival in patients with prostate cancer. Br J Cancer 2004;91:541-3.
31. Bromwich EJ, McArdle PA, Canna K, et al. The relationship between T-lymphocyte infiltration, stage, tumour grade and survival in patients undergoing curative surgery for renal cell cancer. Br J Cancer 2003;89:1906-8.
32. Tan W, Zhang W, Strasner A, et al. Tumour-infiltrating regulatory T cells stimulate mammary cancer metastasis through RANKL-RANK signalling. Nature 2011;470:548-53.
33. Briseño CG, Satpathy AT, Davidson JT, et al. Notch2-dependent DC2s mediate splenic germinal center responses. Proc Natl Acad Sci U S A 2018;115:10726-31.
34. Oh SJ, Ahn S, Jin YH, et al. Notch 1 and Notch 2 synergistically regulate the differentiation and function of invariant NKT cells. J Leukoc Biol 2015;98:781-9.
35. Auderset F, Schuster S, Coutaz M, et al. Redundant Notch1 and Notch2 signaling is necessary for IFNγ secretion by T helper 1 cells during infection with Leishmania major. PLoS Pathog 2012;8:e1002560.
36. Nakano T, Inoue Y, Shimojo N, et al. Lower levels of hsa-mir-15a, which decreases VEGFA, in the CD4+ T cells of pediatric patients with asthma. J Allergy Clin Immunol 2013;132:1224-7.e12.
37. Enomoto A, Ping J, Takahashi M. Girdin, a novel actin-binding protein, and its family of proteins possess versatile functions in the Akt and Wnt signaling pathways. Ann N Y Acad Sci 2006;1086:169-84.
38. Trinath J, Holla S, Mahadik K, et al. The WNT signaling

pathway contributes to dectin-1-dependent inhibition of Toll-like receptor-induced inflammatory signature. Mol Cell Biol 2014;34:4301-14.
39. Silva-García O, Valdez-Alarcón JJ, Baizabal-Aguirre VM. The Wnt/β-catenin signaling pathway controls the inflammatory response in infections caused by pathogenic bacteria. Mediators Inflamm 2014;2014:310183.
40. Ye H, Zhang J, Wang J, et al. CD4 T-cell transcriptome analysis reveals aberrant regulation of STAT3 and Wnt signaling pathways in rheumatoid arthritis: evidence from a case-control study. Arthritis Res Ther 2015;17:76.
41. Lengfeld JE, Lutz SE, Smith JR, et al. Endothelial Wnt/β-catenin signaling reduces immune cell infiltration in multiple sclerosis. Proc Natl Acad Sci U S A 2017;114:E1168-77.
42. Schinzari V, Timperi E, Pecora G, et al. Wnt3a/β-Catenin Signaling Conditions Differentiation of Partially Exhausted T-effector Cells in Human Cancers. Cancer Immunol Res 2018;6:941-52.
43. Tang YY, Sheng SY, Lu CG, et al. Effects of Glycogen Synthase Kinase-3β Inhibitor TWS119 on Proliferation and Cytokine Production of TILs From Human Lung Cancer. J Immunother 2018;41:319-28.

Cite this article as: Yuan C, Xiang L, Bai R, Cao K, Gao Y, Jiang X, Zhang N, Gong Y, Xie C. MiR-195 restrains lung adenocarcinoma by regulating CD4+ T cell activation via the CCDC88C/Wnt signaling pathway: a study based on the Cancer Genome Atlas (TCGA), Gene Expression Omnibus (GEO) and bioinformatic analysis. Ann Transl Med 2019;7(12):263. doi: 10.21037/atm.2019.05.54

Supplementary

Table S1 Search criterion of GEO Databases

Search No.	Search criterion of GEO Databases （up to December 28, 2018）	Items found
#1	(((((Lung Neoplasms[MeSH Terms]) OR Carcinoma, Non-Small-Cell Lung[MeSH Terms]) OR Adenocarcinoma of Lung[MeSH Terms]) OR NSCLC) OR lung Adenocarcinoma) OR LUAD	35,294
#2	((("micrornas"[MeSH Terms] OR mirna[All Fields]) OR ("micrornas"[MeSH Terms] OR microrna[All Fields]) OR ("micrornas"[MeSH Terms] OR mirnas[All Fields]) OR ("micrornas"[MeSH Terms] OR micrornas[All Fields]))	61,229
#3	(profile[All Fields] OR profiles[All Fields] OR profiling[All Fields] OR pattern[All Fields])	116,098
#4	"Homo sapiens"[porgn]	1,577,257
#5	#1 AND #2 AND #3 AND #4	804

Table S2 The down-regulated DEMs in GSE51853

Gene	adj.P.Val	P.Value	t	B	logFC
hsa-miR-144	5.28E-27	1.12E-29	−17.59234	56.946	−5.418055
hsa-miR-486	4.58E-17	1.95E-19	−11.81435	33.799	−4.432312
hsa-miR-451	3.46E-11	2.21E-13	−8.7426	20.045	−5.15291
hsa-miR-126	1.69E-05	2.16E-07	−5.65619	6.504	−2.175387
hsa-miR-30a-5p	3.91E-05	6.65E-07	−5.38563	5.41	−2.105199
hsa-miR-101	6.84E-05	1.31E-06	−5.21951	4.751	−1.602091
hsa-miR-140	9.57E-05	2.04E-06	−5.11034	4.324	−1.203233
hsa-miR-185	1.30E-04	3.18E-06	−4.9987	3.892	−0.932981
hsa-miR-133b	3.43E-04	1.06E-05	−4.69054	2.726	−2.249404
hsa-miR-218	3.43E-04	1.10E-05	−4.68321	2.699	−2.024709
hsa-miR-565	3.43E-04	1.02E-05	−4.70235	2.77	−1.887203
hsa-miR-133a	5.43E-04	1.96E-05	−4.53051	2.137	−1.87426
hsa-miR-557	6.05E-04	2.32E-05	−4.48689	1.979	−1.484434
hsa-miR-30a-3p	8.74E-04	3.72E-05	−4.36026	1.525	−2.301978
hsa-miR-139	9.36E-04	4.27E-05	−4.32292	1.392	−2.906636
hsa-miR-145	9.36E-04	4.38E-05	−4.31593	1.368	−1.641525
hsa-miR-143	1.05E-03	5.13E-05	−4.27304	1.216	−1.482877
hsa-miR-523	2.76E-03	1.41E-04	−3.9927	0.253	−1.825452
hsa-miR-16	2.85E-03	1.52E-04	−3.97165	0.183	−1.031358
*hsa-miR-373**	7.39E-03	4.40E-04	−3.66185	−0.826	−1.480641
hsa-miR-338	8.44E-03	5.21E-04	−3.61174	−0.984	−2.256666
hsa-miR-497	8.68E-03	5.54E-04	−3.59313	−1.042	−1.45217
hsa-miR-659	9.81E-03	6.47E-04	−3.54613	−1.188	−1.554404
hsa-miR-498	1.28E-02	8.72E-04	−3.45468	−1.468	−1.043826
hsa-miR-195	1.60E-02	1.16E-03	−3.36634	−1.733	−1.354666
hsa-miR-516-3p	2.43E-02	1.97E-03	−3.19723	−2.226	−1.069751
hsa-miR-378	3.64E-02	3.09E-03	−3.04791	−2.644	−1.319273
hsa-miR-638	4.37E-02	3.81E-03	−2.9774	−2.836	−1.030528

Table S3 The down-regulated DEMs in GSE63805

Gene	adj.P.Val	P.Value	t	B	logFC
hsa-miR-451	2.43E-13	6.62E-16	–10.4946	25.8539	–2.9249
hsa-miR-126	5.66E-11	2.31E-13	–9.07384	20.1336	–1.746
hsa-miR-30a	1.04E-10	5.66E-13	–8.85938	19.2579	–1.60025
hsa-miR-520e	1.53E-09	1.25E-11	–8.12161	16.2338	–2.3309
hsa-miR-30d	1.21E-07	1.31E-09	–7.0094	11.6836	–1.26736
hsa-miR-144	2.36E-07	2.89E-09	–6.81965	10.9147	–1.75169
hsa-miR-30b	2.87E-07	4.01E-09	–6.74053	10.5952	–1.14825
hsa-miR-145	9.59E-06	1.70E-07	–5.82196	6.9507	–1.27152
hsa-miR-218	1.18E-05	2.70E-07	–5.70542	6.4995	–1.22811
hsa-miR-223	1.18E-05	2.38E-07	–5.73764	6.624	–1.10043
hsa-miR-497	2.76E-04	9.01E-06	–4.79815	3.1126	–0.82562
hsa-miR-135a	2.98E-04	1.06E-05	–4.75562	2.9603	–1.1259
hsa-miR-195	3.48E-04	1.37E-05	–4.68458	2.7075	–0.75577
hsa-miR-16	3.85E-04	1.68E-05	–4.63018	2.5152	–0.6351
hsa-miR-30c	5.95E-04	3.00E-05	–4.47107	1.9595	–0.68193
hsa-miR-143	1.50E-03	8.38E-05	–4.18252	0.9793	–0.83223
hsa-miR-99a	4.57E-03	2.68E-04	–3.84362	–0.1218	–0.72796
ebv-miR-BART9	4.95E-03	3.03E-04	–3.80685	–0.2377	–0.84051
hsa-miR-140-5p	6.15E-03	3.94E-04	–3.72758	–0.4852	–0.5014
hsa-miR-302e	9.09E-03	6.44E-04	–3.57649	–0.9474	–0.58935
hsa-miR-302f	1.62E-02	1.28E-03	–3.35862	–1.5905	–0.61706
hsa-let-7b	1.65E-02	1.33E-03	–3.34684	–1.6245	–0.44182
hsa-miR-23a	2.16E-02	1.88E-03	–3.23388	–1.9458	–0.57929
hsa-miR-26a	2.40E-02	2.17E-03	–3.18665	–2.0777	–0.42085
hsa-let-7d	3.42E-02	3.31E-03	–3.04369	–2.4683	–0.36884
hsa-miR-125b	4.74E-02	5.24E-03	–2.88447	–2.8871	–0.53226
hsa-miR-335	4.93E-02	5.67E-03	–2.85642	–2.9591	–0.50533

Table S4 The down-regulated DEMs in GSE74190

Gene	adj.P.Val	P.Value	t	B	logFC
hsa-miR-126	2.51E-25	3.05E-28	−16.97988	53.87438	−2.606823
hsa-miR-144	4.12E-21	1.00E-23	−14.29145	43.57032	−4.144946
*hsa-miR-126**	8.40E-21	3.07E-23	−14.01579	42.46067	−2.39406
hsa-miR-218	4.23E-19	2.58E-21	−12.94572	38.06224	−2.302513
hsa-miR-451	2.56E-18	2.18E-20	−12.44139	35.94098	−4.1685
hsa-miR-486-5p	3.19E-18	3.11E-20	−12.3584	35.58912	−3.417527
hsa-miR-140-3p	2.83E-17	3.10E-19	−11.8239	33.30445	−1.32449
hsa-miR-30a	9.21E-17	1.28E-18	−11.49705	31.89247	−2.039716
hsa-miR-139-5p	3.51E-16	5.55E-18	−11.16276	30.43748	−2.283491
*hsa-miR-144**	3.30E-12	6.82E-14	−9.05712	21.08639	−2.372553
*hsa-miR-30a**	6.81E-12	1.49E-13	−8.8839	20.30921	−2.226259
hsa-miR-223	1.35E-11	3.11E-13	−8.72153	19.58067	−1.375561
hsa-miR-133b	2.77E-11	6.74E-13	−8.55086	18.81514	−1.724604
hsa-miR-498	9.59E-11	2.92E-12	−8.22664	17.36267	−1.626849
hsa-miR-145	2.62E-10	8.29E-12	−7.99557	16.33003	−1.498766
hsa-miR-338-3p	2.66E-10	8.75E-12	−7.9835	16.27615	−2.341491
*hsa-miR-187**	3.38E-10	1.15E-11	−7.92251	16.00411	−1.756869
hsa-miR-497	5.83E-10	2.06E-11	−7.79354	15.42967	−1.173224
hsa-miR-195	2.93E-09	1.11E-10	−7.4183	13.76591	−1.213904
*hsa-miR-30c-2**	3.81E-09	1.48E-10	−7.35278	13.47683	−1.157946
hsa-miR-101	2.82E-08	1.20E-09	−6.88137	11.41267	−1.192292
hsa-miR-130a	3.10E-08	1.40E-09	−6.84717	11.26416	−0.910084
hsa-miR-143	3.56E-08	1.65E-09	−6.80954	11.10094	−1.357178
*hsa-miR-145**	1.50E-07	7.86E-09	−6.45194	9.56274	−1.493834
hsa-miR-638	3.53E-07	1.89E-08	−6.24868	8.69977	−1.459486
hsa-miR-134	4.20E-07	2.30E-08	−6.2029	8.50672	−1.344233
hsa-miR-1225-5p	1.11E-06	6.47E-08	−5.96045	7.4926	−1.074196
*hsa-miR-373**	1.75E-06	1.07E-07	−5.84219	7.00352	−0.792895
hsa-miR-140-5p	1.79E-06	1.11E-07	−5.83253	6.96371	−0.959252
hsa-miR-1	2.68E-06	1.73E-07	−5.72638	6.52833	−1.710784
hsa-miR-557	3.12E-06	2.05E-07	−5.68569	6.36231	−1.208886
hcmv-miR-UL70-3p	3.65E-06	2.44E-07	−5.64371	6.19157	−1.558056
*hsa-miR-150**	4.03E-06	2.75E-07	−5.6152	6.07592	−0.934027
hsa-miR-99a	9.35E-06	6.83E-07	−5.39385	5.18677	−0.869441
hsa-miR-1224-5p	9.36E-06	6.96E-07	−5.38952	5.16951	−0.759061
hsa-miR-30c	1.05E-05	7.94E-07	−5.3571	5.04074	−0.595558
hsa-miR-30b	1.28E-05	9.81E-07	−5.3049	4.83411	−1.063555
hsa-miR-328	1.48E-05	1.17E-06	−5.26169	4.66376	−0.915212
hsa-miR-572	4.99E-05	4.20E-06	−4.94081	3.42087	−1.352622
*hsa-miR-143**	9.37E-05	8.10E-06	−4.77178	2.78276	−0.606271
hsv1-miR-LAT	1.18E-04	1.03E-05	−4.70915	2.54942	−1.172914
hsa-miR-100	2.31E-04	2.13E-05	−4.51769	1.84711	−0.871147
hsa-miR-30d	2.40E-04	2.28E-05	−4.49993	1.78283	−0.719427
kshv-miR-K12-3	2.60E-04	2.50E-05	−4.47566	1.69523	−0.925876
hsa-miR-27a	3.83E-04	3.92E-05	−4.35454	1.2623	−0.492615
hsa-miR-940	4.01E-04	4.15E-05	−4.33896	1.20712	−0.690085
hsa-miR-34b	4.52E-04	4.79E-05	−4.29959	1.06828	−1.530313
hsa-miR-874	8.69E-04	9.43E-05	−4.11264	0.4199	−0.649814
hsa-miR-142-5p	1.32E-03	1.48E-04	−3.98467	−0.01312	−0.983926
hsa-miR-551b	1.94E-03	2.20E-04	−3.87121	−0.38936	−1.196408
hsa-miR-197	2.83E-03	3.35E-04	−3.74868	−0.7873	−0.552384
*hsa-let-7d**	2.99E-03	3.57E-04	−3.72965	−0.84832	−0.743634
hsa-miR-150	3.41E-03	4.12E-04	−3.68732	−0.98323	−0.744552
hsa-miR-335	3.94E-03	4.89E-04	−3.63569	−1.1463	−0.855696
hsa-miR-630	4.56E-03	5.72E-04	−3.58841	−1.29421	−0.92997
hsa-miR-34c-5p	6.41E-03	8.35E-04	−3.47238	−1.65119	−1.682751
*hsa-miR-125b-2**	7.58E-03	1.03E-03	−3.40859	−1.84376	−0.71629
hsa-miR-636	7.75E-03	1.08E-03	−3.39336	−1.88935	−0.677055
*hsa-miR-1226**	9.92E-03	1.40E-03	−3.31004	−2.13604	−0.685631
hsa-miR-188-5p	1.18E-02	1.72E-03	−3.24472	−2.32617	−0.539625
hsa-miR-125b	1.29E-02	1.94E-03	−3.20596	−2.43762	−0.506143
hsa-miR-877	1.37E-02	2.08E-03	−3.18318	−2.50263	−0.562507
hsa-miR-203	1.42E-02	2.21E-03	−3.16338	−2.55886	−0.790212
hsa-miR-601	1.51E-02	2.37E-03	−3.13944	−2.62647	−0.679305
*hsa-miR-149**	2.16E-02	3.64E-03	−2.99634	−3.02219	−0.533993
hsa-miR-936	2.22E-02	3.81E-03	−2.98072	−3.0645	−0.368881
hsa-miR-139-3p	2.36E-02	4.10E-03	−2.95532	−3.1329	−0.721107
hsa-miR-671-5p	2.50E-02	4.47E-03	−2.92589	−3.21161	−0.475478
hsa-miR-342-3p	2.94E-02	5.52E-03	−2.85223	−3.40574	−0.266087
hsa-miR-564	2.96E-02	5.58E-03	−2.8487	−3.41496	−0.530209
hsa-miR-125a-3p	3.26E-02	6.30E-03	−2.80559	−3.52659	−0.725482
hsa-miR-517c	3.56E-02	6.94E-03	−2.77146	−3.61399	−0.213062
hsa-miR-939	3.72E-02	7.30E-03	−2.75326	−3.66025	−0.512994
hsa-miR-29c	4.58E-02	9.20E-03	−2.66924	−3.87046	−0.455756

*, refers to the minor product of the 2 miRNAs derived from the same miRNA hairpin. Usually, its concentration is lower than the major one in the cell.

Table S5 The results of GO term analysis included the BP, CC and MF group

Category	ID	Term	Genes	adj_pval
MF	GO:0005515	Protein binding	*AMOTL1, VPS33B, CD2AP, ZNRF3, CUL2, PSKH1, TRAK1, RAB23, VPS4A, TLK1, OGT, EIF2B2, BCL7A, CDCA4, TBPL1, RECK, PTPRJ, PIM1, PRDM4, SERBP1, VEGFA, MTFR1L, RAD23B, MOB4, CHEK1, BCL2L2, HSPA1B, CALU, ARIH1, STX17, NUP50, HOXA10, TRAM1, CDC37L1, ZNF622, KIF3B, OSBPL3, TAOK1, SMAD7, SREK1, YTHDC1, SMAD3, SKI, CDC27, NOTCH2, SUPT16H, TGFBR3, SMURF1, RBPJ, CREBRF, CCNT2, CLSPN, ZBTB33, CPEB3, ZMAT3, CCNT1, SHOC2, PDIA6, BAG4, CCNE1, TMEM100, MYB, SIK1, USP15, FGF2, TLE4, RBBP6, CARD10, PDIK1L, RAB11FIP2, CCND1, ALOX12, PPP6C, ZBTB10, ATG14, FKBP1A, NUFIP2, MTMR3, FASN, AGO1, CSDE1, PRKAA1, PCMT1, PEX13, AGO4, PEX12, KIF21A, MLLT6, MTMR4, SSRP1, PPP1R11, AFF4, STRADB, RAPH1, ITPR1, RPS6KA3, PHF19, YWHAH, RASSF2, YWHAQ, ZBTB5, KATNAL1, WNT7A, USP42, RNMT, STK38, BTRC, CBX4, CASK, CBX2, RPS6KB1, FLCN, VCL, PRKAR2A, SPRED1, INSR, ZNF449, RET, SPTLC1, ATG9A, GABPA, WEE1, RNF138, NAPG, LITAF, CACUL1, PPIL1, SNRPB2, UBE2V1, NR2C2, AMER1, HOXA3, DDX3X, ELK4, AXIN2, ENTPD1, ARHGDIA, GABARAPL1, YIPF6, ITGA2, MAFK, WIPI2, CYB561A3, C2ORF42, LAMP2, CDC42SE2, SALL1, CPNE1, ABL2, TBC1D20, GPATCH8, PLEKHA1, KIF23, CDK17, SEC24A, FOXK1, E2F7, CBFA2T3, PHC3, CDC42, PLRG1, LRRFIP2, CHUK, AKT3, ARL2, TBL1XR1, SYNRG, RUNX1T1, STXBP3, PURA, RFWD2, ACVR2A, TARBP2, ARHGAP32, CRKL, BTG2, CPSF7, KPNA1, PHLPP2, PPM1A, CEP55, N4BP1, ETNK1, PAFAH1B1, PAFAH1B2, TNRC6B, PIK3R1, PTPRD, UBE4A, HMBOX1, SIRT4, RAF1, UBE2Q1, TAB3, CDC25A, LSM11, IRF4, GGA3, RNF41*	5.23E-11
CC	GO:0005654	Nucleoplasm	*CREBRF, KIF23, CCNT2, ACOX1, CLSPN, ZBTB33, RNMT, ZBTB34, EZH1, BTRC, E2F7, CCNT1, CBX4, SHOC2, CBX2, RPS6KB1, PNISR, CBFA2T3, PHC3, CUL2, CCNE1, PLRG1, SPRED1, RARB, OGT, AKT3, CHUK, TBL1XR1, GABPA, RUNX1T1, TLE4, UBN2, RBBP6, WEE1, RFWD2, PDIK1L, CCND1, RAB11FIP2, ASH1L, CPSF7, KPNA1, RAD23B, LITAF, ZBTB10, SNRPB2, PPM1A, CHEK1, HSPA1B, IVNS1ABP, NR2C2, MTMR3, HOXA3, ELK4, NUP50, AGO1, ETNK1, PRKAA1, AXIN2, SSRP1, JARID2, SMAD7, SREK1, SMAD3, SKI, MAFK, WIPI2, CDC27, CDC25A, TAB3, NOTCH2, RPS6KA3, PHF19, RPRD2, DMTF1, SALL1, SUPT16H, SH3BGRL2, LSM11, SMURF1, IRF4, RBPJ, IPPK, PLEKHA1*	6.63E-07
CC	GO:0005737	Cytoplasm	*STK38, BTRC, CASK, RPS6KB1, PNISR, AMOTL1, FLCN, VPS33B, CD2AP, ZNRF2, PRKAR2A, TRAK1, RAB23, VPS4A, RARB, SPRED1, EIF2B2, TBPL1, RET, PIM1, UBR3, WEE1, TRIM35, PRDM4, SERBP1, VEGFA, RAD23B, MOB4, LITAF, UBE2V1, LRIG2, HSPA1B, ARIH1, DDX3X, ELK4, STX17, NUP50, DDX3Y, HOXA10, AXIN2, CDC37L1, ARHGDIA, ZNF622, TAOK1, SMAD7, SMAD3, SKI, WIPI2, CDC27, HSPE1-MOB4, CDC42SE2, DMTF1, SALL1, PRICKLE2, CPNE1, TGFBR3, SMURF1, RBPJ, PLEKHA1, CREBRF, CLSPN, ZBTB33, CPEB2, CPEB3, SHOC2, CDC42, BZW1, ACTR2, CAMSAP1, SIK1, USP15, FGF2, CHUK, AKT3, SYNRG, RAP2C, CCDC88C, FBXL20, RUNX1T1, SRPRB, PURA, CARD10, RFWD2, ACVR2A, TARBP2, CCND1, ASH1L, KPNA1, ALOX12, MYO5A, PHLPP2, USP3, ATG14, FKBP1A, IVNS1ABP, NUFIP2, MTMR3, CYP26B1, CSDE1, AGO1, FASN, ETNK1, PCMT1, PRKAA1, AGO4, PAFAH1B2, KIF21A, PIK3R1, SSRP1, UBE4A, NUCKS1, HMBOX1, PPP1R11, RAF1, STRADB, RAPH1, CDC25A, TAB3, RPS6KA3, YWHAH, RASSF2, HSPA4L, SH3BGRL2, YWHAQ, IRF4, KATNAL1, PSAT1*	9.09E-07
CC	GO:0005829	Cytosol	*KIF23, SEC24A, BTRC, CAPZA2, CASK, RPS6KB1, AMOTL1, ZNRF2, VCL, ARHGAP12, CDC42, CUL2, ACTR2, BAG4, CCNE1, PRKAR2A, VPS4A, SPRED1, OGT, EIF2B2, CHUK, ARL2, RAP2C, STXBP3, RBBP6, RFWD2, TARBP2, CCND1, ARHGAP32, CRKL, BTG2, KPNA1, ALOX12, MYO5A, PPP6C, MOB4, PHLPP2, ATG14, PHKA1, SNX16, PPM1A, UBE2V1, FKBP1A, BCL2L2, CHEK1, HSPA1B, TPM2, ARIH1, MTMR3, AMER1, MTHFR, STX17, FASN, AGO1, ETNK1, PCMT1, PRKAA1, PAFAH1B1, AGO4, PAFAH1B2, TNRC6B, AXIN2, MTMR4, PIK3R1, ARHGDIA, GABARAPL1, KIF3B, OSBPL3, TAOK1, SMAD7, RAF1, SMAD3, STRADB, WIPI2, CDC27, UBE2Q1, CDC25A, TAB3, RPS6KA3, YWHAH, HSPA4L, YWHAQ, CPNE1, SMURF1, IRF4, IPPK, PSAT1, ABL2, RNF41*	2.21E-05
BP	GO:0006468	Protein phosphorylation	*CCNT2, CDK17, RET, STK38, TAOK1, CCNT1, PHKA1, PIM1, RAF1, CASK, RPS6KB1, STRADB, PDIK1L, CCNE1, CCND1, RPS6KA3, PSKH1, RASSF2, TLK1, PRKAA1, SIK1, CHUK, PIK3R1, AKT3*	0.002328
CC	GO:0005634	Nucleus	*RNMT, STK38, BTRC, CBX4, CBX2, RPS6KB1, LUZP1, FLCN, TRAK1, VPS4A, TLK1, OGT, RARB, CDCA4, ZNF449, GABPA, PIM1, PISD, WEE1, TRIM35, PRDM4, SERBP1, RNF138, RAD23B, UBE2V1, DSCR3, CHEK1, NR2C2, AMER1, HOXA3, DDX3X, ELK4, DDX3Y, HOXA10, AXIN2, ARHGDIA, PLAG1, SMAD7, YTHDC1, SMAD3, ITGA2, SKI, MAFK, CDC27, NOTCH2, C2ORF42, ZFHX4, DMTF1, SALL1, SUPT16H, CPNE1, SMURF1, RBPJ, KIF23, CCNT2, ACOX1, ZBTB33, CPEB2, FOXK1, CPEB3, E2F7, CCNT1, SHOC2, CBFA2T3, PHC3, BAG4, CCNE1, PLRG1, MKX, SIK1, MYB, FGF2, USP15, ARL2, TBL1XR1, TLE4, UBN2, RBBP6, PURA, TARBP2, CCND1, ZNF691, ASH1L, CPSF7, KPNA1, IFT74, ZNF275, PHLPP2, USP3, ATG14, PPM1A, NUFIP2, ZNF704, AGO1, PRKAA1, AGO4, MLLT6, PIK3R1, USP31, ZNF267, SSRP1, UBE4A, NUCKS1, JARID2, HMBOX1, PPP1R11, RAF1, STRADB, UBE2Q1, CDC25A, RPS6KA3, RASSF2, HSPA4L, ZBTB5, LSM11, IRF4, KATNAL1, IPPK*	0.004008
MF	GO:0008013	Beta-catenin binding	*PTPRJ, TBL1XR1, AMER1, SMAD7, BTRC, SALL1, SMAD3, AXIN2, CD2AP, VCL*	0.008021
MF	GO:0061630	Ubiquitin protein ligase activity	*CACUL1, BTRC, UBR3, UBE2V1, RBBP6, ZNRF3, ZNRF2, RFWD2, CDC42, ARIH1, CUL2, RNF138, SMURF1, RNF41*	0.012324
BP	GO:0016055	Wnt signaling pathway	*ZBTB33, CCDC88C, BTRC, PPM1A, TLE4, AMOTL1, CCNE1, AMER1, DDX3X, RNF138, LRRFIP2, PRKAA1, AXIN2, WNT7A*	0.018071
MF	GO:0048185	Activin binding	*ACVR2A, SMAD7, TGFBR3, FKBP1A, SMURF1*	0.038549

BP, biological process; CC, cellular component; MF, molecular function.

Table S6 immunologic signatures gene sets in GSE13485

Name	PROBE	Rank in gene list	Rank Metric Score	RUNNING ES	Core enrichment
row_0	CALU	10	0.757121444	0.103793696	Yes
row_1	RAD23B	44	0.474761099	0.0702462	Yes
row_2	VPS33B	49	0.431200683	0.13561504	Yes
row_3	ATG14	50	0.428349882	0.21521427	Yes
row_4	USP42	62	0.363170534	0.24211097	Yes
row_5	NAPG	71	0.336881727	0.2751926	Yes
row_6	ATG9A	76	0.323062837	0.32046643	Yes
row_7	ZBTB10	87	0.291141659	0.33766824	Yes
row_8	USP3	127	0.157543182	0.22303265	No
row_9	PAFAH1B2	137	0.130324796	0.21404028	No
row_10	PPM1A	175	-0.030661186	0.0832066	No
row_11	MTMR3	205	-0.182105184	0.010035696	No
row_12	SIK1	208	-0.207770154	0.04126505	No
row_13	ZBTB34	222	-0.288843781	0.04696972	No
row_14	SMAD7	274	-0.97839427	0.040590495	No

Table S7 Immunologic signatures gene sets in GSE21379

Name	PROBE	Rank in gene list	Rank Metric Score	RUNNING ES	Core enrichment
row_0	OSBPL3	51	0.428294897	–0.042887516	Yes
row_1	IVNS1ABP	58	0.391723037	0.067869164	Yes
row_2	USP42	62	0.363170534	0.18000917	Yes
row_3	VEGFA	100	0.254309952	0.12975562	Yes
row_4	BAG4	118	0.193770424	0.13276401	Yes
row_5	GGA3	121	0.189701915	0.18974267	Yes
row_6	BTRC	130	0.147670254	0.21032134	Yes
row_7	CCDC88C	136	0.130609587	0.23618208	Yes
row_8	NOTCH2	144	0.102264941	0.24504647	Yes
row_9	MAFK	159	0.049663894	0.21023504	Yes
row_10	CAMSAP1	171	-0.011271492	0.17346863	Yes
row_11	PRKAR2A	188	-0.107315883	0.15083629	Yes
row_12	MOB4	189	-0.114678584	0.18974242	Yes
row_13	DDX3Y	209	-0.214066073	0.19225629	Yes
row_14	FRYL	217	-0.249059677	0.25092262	Yes

Table S8 Immunologic signatures gene sets in GSE29164

Name	PROBE	Rank in gene list	Rank Metric Score	RUNNING ES	Core enrichment
row_0	ENTPD7	7	0.784201086	0.15258355	Yes
row_1	TLK1	20	0.619249582	0.24910028	Yes
row_2	OGT	53	0.418568671	0.2258615	Yes
row_3	NUP50	63	0.362397194	0.27502146	Yes
row_4	ATP13A3	69	0.342263281	0.33441314	Yes
row_5	UBN2	82	0.308478147	0.36018828	Yes
row_6	NUCKS1	85	0.298018843	0.42061958	Yes
row_7	JARID2	91	0.27646625	0.46503368	Yes
row_8	TAB3	116	0.201044887	0.4219091	No
row_9	FOXK1	135	0.132048801	0.38530102	No
row_10	RPS6KA3	146	0.097817853	0.3705305	No
row_11	TNRC6B	162	0.039751362	0.3240236	No
row_12	ZMAT3	167	-0.002394611	0.3097539	No
row_13	PRDM4	177	-0.038236551	0.28512442	No
row_14	RAF1	185	-0.066899218	0.27442694	No
row_15	VCL	238	-0.405213803	0.17407407	No

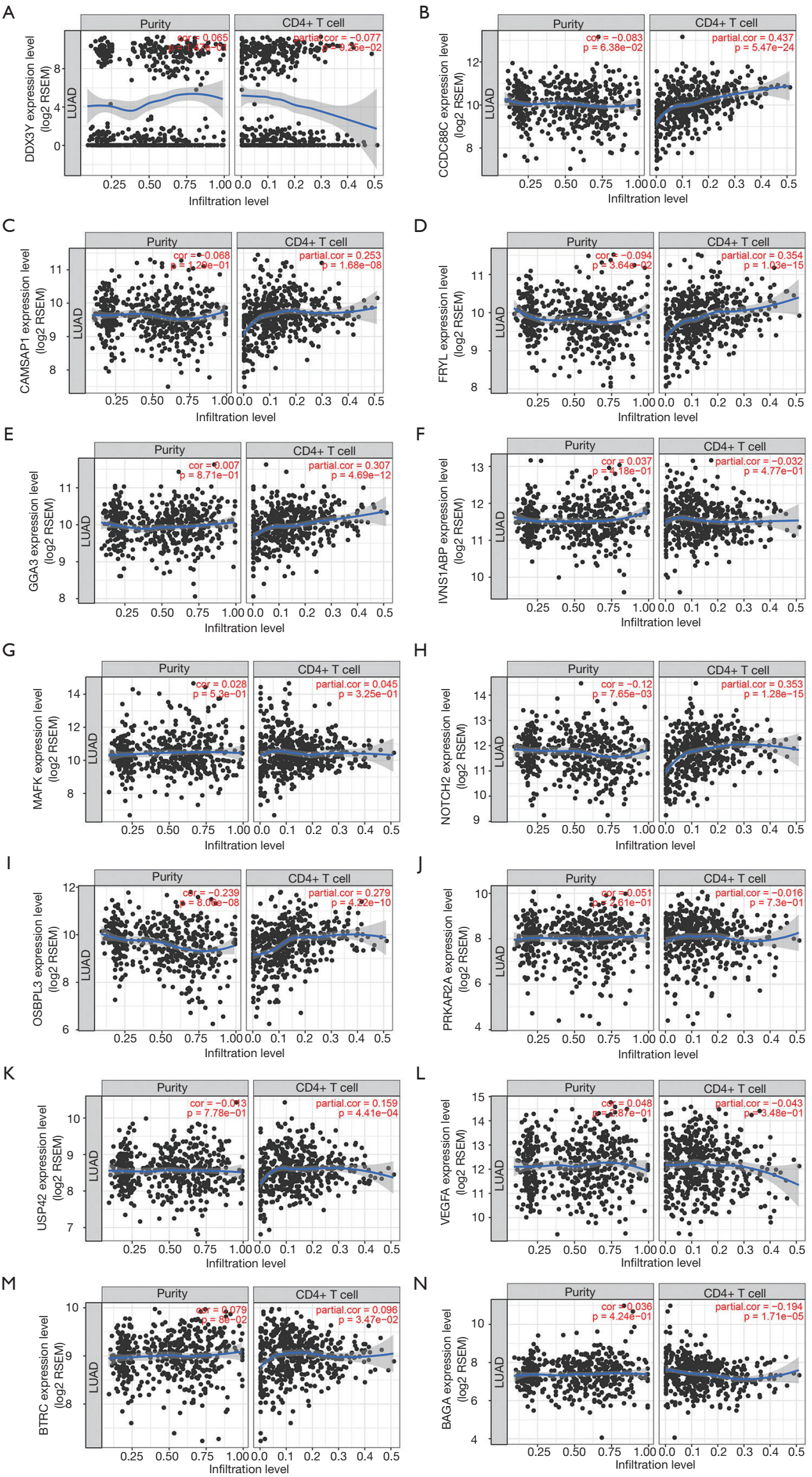

Figure S1 The correlation between CD4+ T cell infiltration and genes as profiled by TIMER. A total of 14 genes (*OSBPL3*, *IVNS1ABP*, *USP42*, *VEGFA*, *BAG4*, *GGA3*, *BTRC*, *CCDC88C*, *NOTCH2*, *MAFK*, *CAMSAP1*, *PRKAR2A*, *MOB4*, *DDX3Y* and *FRYL*) were used to explore the correlation between gene expression changes and CD4 + T cell infiltration. The expression levels of CCDC88C were significantly correlated with CD4+ T cell activation.