Check for updates

Identification of lymph node metastasis-related genes and patterns of immune infiltration in colon adenocarcinoma

Haoxiang Zhang[1,2†], Guibin Zhao[3†], Guangwei Zhu[1,2,4]
and Jianxin Ye[1,2]*

[1]Department of Gastrointestinal Surgery 2 Section, Institute of Abdominal Surgery, Key Laboratory of Accurate Diagnosis and Treatment of Cancer, The First Affiliated Hospital, Fujian Medical University, Fuzhou, China, [2]Department of Gastrointestinal Surgery 2 Section, National Regional Medical Center, Fujian Medical University, Fuzhou, China, [3]Department of Gastrointestinal Surgery, Mindong Hospital Affiliated to Fujian Medical University, Fuan, China, [4]Key Laboratory of Ministry of Education for Gastrointestinal Cancer, Fujian Medical University, Fuan, China

Backgrounds: Colon adenocarcinoma(COAD) is one of the most common tumors of the digestive tract. Lymph node metastasis (LNM) is a well-established prognostic factor for COAD. The mechanism of COAD lymph node metastasis in immunology remains unknown. The identification of LNM-related biomarkers of COAD could help in its treatment. Thus, the current study was aimed to identify key genes and construct a prognostic signature.

Methods: Gene expression and clinical data were obtained from The Cancer Genome Atlas (TCGA) database. Differentially expressed genes were calculated by using R software. GO functional and KEGG pathway enrichment analysis were processed. The CIBERSORT algorithm was used to assess immune cell infiltration. STRING database was used to screen key genes and constructed a protein-protein interaction network (PPI network). The LASSO-Cox regression analysis was performed based on the components of the PPI network. The correlation analysis between LNM-related signature and immune infiltrating cells was then investigated. TISIDB was used to explore the correlation between the abundance of immunomodulators and the expression of the inquired gene.

Results: In total, 394 differentially expressed genes were identified. After constructing and analyzing the PPI network, 180 genes were entered into the LASSO-Cox regression model, constructing a gene signature. Five genes(PMCH, LRP2, NAT1, NKAIN4, and CD1B) were identified as LNM-related genes of clinical value. Correlation analysis revealed that LRP2 and T follicular helper cells (R=0.34, P=0.0019) and NKAIN4 and T follicular helper cells (R=0.23, P=0.041) had significant correlations. Immunologic analysis revealed that LRP2 and NKAIN4 are potential coregulators of immune checkpoints in COAD.

Conclusion: In general, this study revealed the key genes related to lymph node metastasis and prognostic signature. Several potential mechanisms and therapeutic and prognostic targets of lymph node metastasis were also demonstrated in COAD.

KEYWORDS

immune cell infiltration, lymph node metastasis, TCGA, bioinformatic analysis, colon adenocarcinoma

Introduction

Colorectal cancer (CRC) is the third most commonly diagnosed neoplasm, with the second-highest cancer-associated mortality rate globally (1). Colorectal cancer includes cancers of both the colon and the rectum and is the most common form of gastrointestinal cancer. About 90% of colorectal cancers are adenocarcinomas originating from the colorectal mucosal epithelium in the world (2). The treatment of colorectal cancer includes surgery, chemotherapy, radiotherapy, and other biological, immunological treatments. Although therapies for colorectal cancer have improved in recent years, the survival rate of colorectal cancer patients is still not satisfactory (3). The development of a novel therapeutic target is required to increase the survival rates of colorectal cancer patients. Therefore, we conducted a genetic and immunologic study on colon adenocarcinoma(COAD).

Lymph node metastasis (LNM) is an independent risk factor affecting the prognosis of patients with COAD (4–6). The extent of lymph node metastasis affects the postoperative staging and subsequent treatment measures of COAD patients. The TNM system of the American Joint Committee on Cancer/International Union Against Cancer Classification (AJCC/UICC) recommends a minimal assessment of 12 lymph nodes for accurate staging. Currently, some studies have investigated the mechanism of COAD lymph node metastasis. The combination of p53 deficiency and azoxymethane (AOM) promotes tumor development, including the growth of invasive cancers and lymph node metastasis (7). Smad4 and VEGF-C are involved in lymphangiogenesis and lymphatic metastasis (8). FHL2 overexpression of myofibroblasts in sporadic colon cancer was significantly linked to lymphatic metastasis (9). The tumor microenvironment of COAD is associated with tumor lymph node metastasis. Interferon regulatory factor 1 (IRF1) was associated with tumor stage, lymph node metastasis, and distant metastasis. Furthermore, there are significant associations between IRF1 and CD8+ T cells, T cell, dendritic cells, T-helper 1 cells, and T cell exhaustion (10). It appears that the mechanism of lymph node metastasis in COAD is complicated and is related to immune-related genes and tumor microenvironment.

We conducted a study on lymph node metastasis-related genes (LNM-related genes) in COAD and revealed the relationship between LNM-related genes and immune infiltrating cells in COAD. In our study, datasets for colon adenocarcinoma samples were downloaded from The Cancer Genome Atlas (TCGA) public database. We screened out the differentially expressed genes between N+ (COAD patients with lymph node metastasis) and N0 (COAD patients without lymph node metastasis) samples and constructed a protein-protein interaction network(PPI) to screen out LNM-related genes to construct a survival prognosis model. The co-expression analysis of genes with prognostic significance and immune infiltrating cells of COAD was conducted to illustrate the potential mechanisms of lymph node metastasis in COAD.

Materials and methods

Identification of differentially expressed genes

COAD (TCGA-COAD) RNA-seq data were downloaded from the TCGA public data platform(The Cancer Genome Atlas) (https://portal. gdc.cancer.gov/), including 473 cases of COAD tumor tissues and 41 cases of normal tissues. Normal tissue data were eliminated, and selected high-quality samples in duplicate samples(1. Frozen tissue is preferred to formalin-fixed paraffin-embedded tissue; 2. Tissue with larger plate value). Finally, 447 tumor tissue with clinical information were selected for the identification of differentially expressed genes. Demographic and clinical data of 447 patients with COAD are summarized(including age, gender, stage, and TNM classification) (Table 1). Tissue samples were divided into N+ group (n=181) with positive lymph node metastasis and N0 group without lymph node metastasis (n=266). Limma R package (version: 3.40.2) of R software was used for differential expression analysis between the two groups. The adjusted P-value was analyzed to correct for false-positive results in TCGA. FDR< 0.05 and |log2 fold change|> 1 were defined as the thresholds for differential expression analysis. A graphical representation of the complete workflow is provided in Figure 1.

To further confirm the underlying function of potential targets, the data were analyzed by functional enrichment. Gene Ontology (GO) is a widely-used tool for annotating genes with functions, especially molecular function (MF), biological pathways (BP), and cellular components (CC). Kyoto Encyclopedia of Genes and Genomes (KEGG) Enrichment Analysis is a practical resource for analyzing gene functions and associated high-level genome functional information. To better understand the carcinogenesis of mRNA, the ClusterProfiler package (version: 3.18.0) in R was employed to analyze the GO function of potential targets and enrich the KEGG pathway. Ethical approval from an institutional review board was not required because the survey data are publicly available. The study was conducted following the Declaration of Helsinki (as revised in 2013).

Construction of a protein-protein interaction network

STRING is a public database that contains interactions between known and predicted proteins (https://string-db.org/). The database covers protein interactions from 2031 species containing a total of 9,643,763 proteins and 1,380,838,440 interactions. In addition to the experimental data, the results of text mining from PubMed abstracts and the synthesis of other database data are also included in the database. It also uses bioinformatics methods to predict the results. The PPI network was constructed by the STRING database based on differentially expressed genes between the N+ group and the N0 group. The PPI network was then visualized using Cytoscape software version 3.8.0 (www.cytoscape.org/).

Survival analysis based on the components of the PPI network

All components of the PPI network were incorporated into the Cox regression model, and to assure multifaceted models were not overfitting, the Lasso regression was used. All patients were consequently classified into a high- and low-risk group by utilizing the median risk score calculated by LASSO as the cutoff value. Kaplan-Meier survival curves were compared using the log-rank test. Receiver operating characteristic (ROC) curves were

TABLE 1 Characteristics of the included COAD patients obtained from the TCGA database.

Basic information	N+ group (n=181)	N0 group(n=266)
Age	66 (median)	70 (median)
Gender		
Female	91	145
Male	90	121
AJCC Stage		
I & II	0	251
III & IV	177	8
Unknow	4	7
T classification		
Tis	0	1
T1 & T2	9	77
T3 & T4	172	188
N classification		
N0	0	266
N1	102	0
N2	79	0
M classification		
M0	102	228
M1	53	8
MX	26	30

COAD; colon adenocarcinoma; TCGA, the The Cancer Genome Atlas; AJCC, American Joint Committee on Cancer.

constructed to estimate the area under the ROC curve. A genomic nomogram was developed based on the LASSO-Cox regression model analysis results to predict the 1, 3, 5-year overall survival(OS). The nomogram provided a graphical representation of the factors, which can calculate the risk of death for an individual patient by the points associated with each risk factor through the "rms" R package. The calibration curve was used to evaluate the consistency/calibration, that is, the difference between the predicted and true values. Genes with prognostic significance in the PPI network are defined as lymph node metastasis-related genes (LNM-related genes) in COAD.
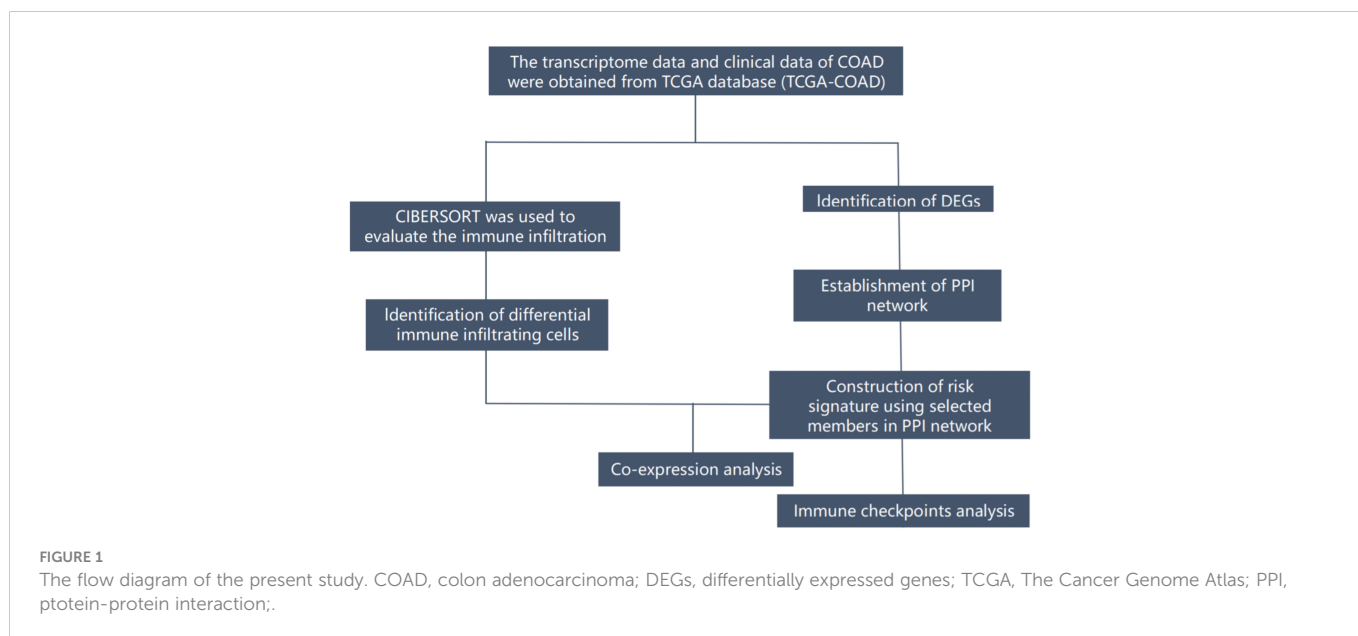
Immune cell infiltration

We analyzed immune cell infiltration in each tumor sample using the CIBERSORT method ($p < 0.05$). CIBERSORT is a deconvolution algorithm that can predict the abundance of 22 immune cell types based on the gene expression data. WilcoxTest was used to screen the differential immune cell infiltration between the N0 and N+ groups.). Then, Pearson's correlation coefficient was used to assess the correlations between the differential immune infiltrating cells and LNM-related genes.

Immunologic analysis

The investigated immunoinhibitors were collected according to Garraway et al.'s study (11–15). Based on five studies, we selected immune checkpoints with the top ten mutation rates. Pearson's correlation analysis was used to evaluate the relationship between LNM-related genes and immune checkpoints. Each Spearman correlation between inquired gene and a distinct immunoinhibitor in an individual cancer type was integrated into the indicated heatmap.

Statistical analysis

All statistical analysis were performed using R 4.11 software. PPI network was constructed by "STRING." Univariate and multivariate analysis were performed with log-rank test and Cox regression, using

FIGURE 1
The flow diagram of the present study. COAD, colon adenocarcinoma; DEGs, differentially expressed genes; TCGA, The Cancer Genome Atlas; PPI, ptotein-protein interaction;.

the survival package of R. There were two outcome variables (status and survival time). The expression of the components of the PPI network and immune cell infiltration were assumed to be independent variables in Cox regression used for two survival analysis. LASSO-Cox regression model analysis was performed using R software's package "glmnet" (version 4.11). Receiver operating characteristic (ROC) curves of 1, 3, 5 years were plotted, and the area under the curve (AUC) values were calculated using the R package "survivalROC." Nomograms and calibration plots were done with "rms" package of R software. Differences in gene expression between the two groups were evaluated using the Mann-Whitney U-test.
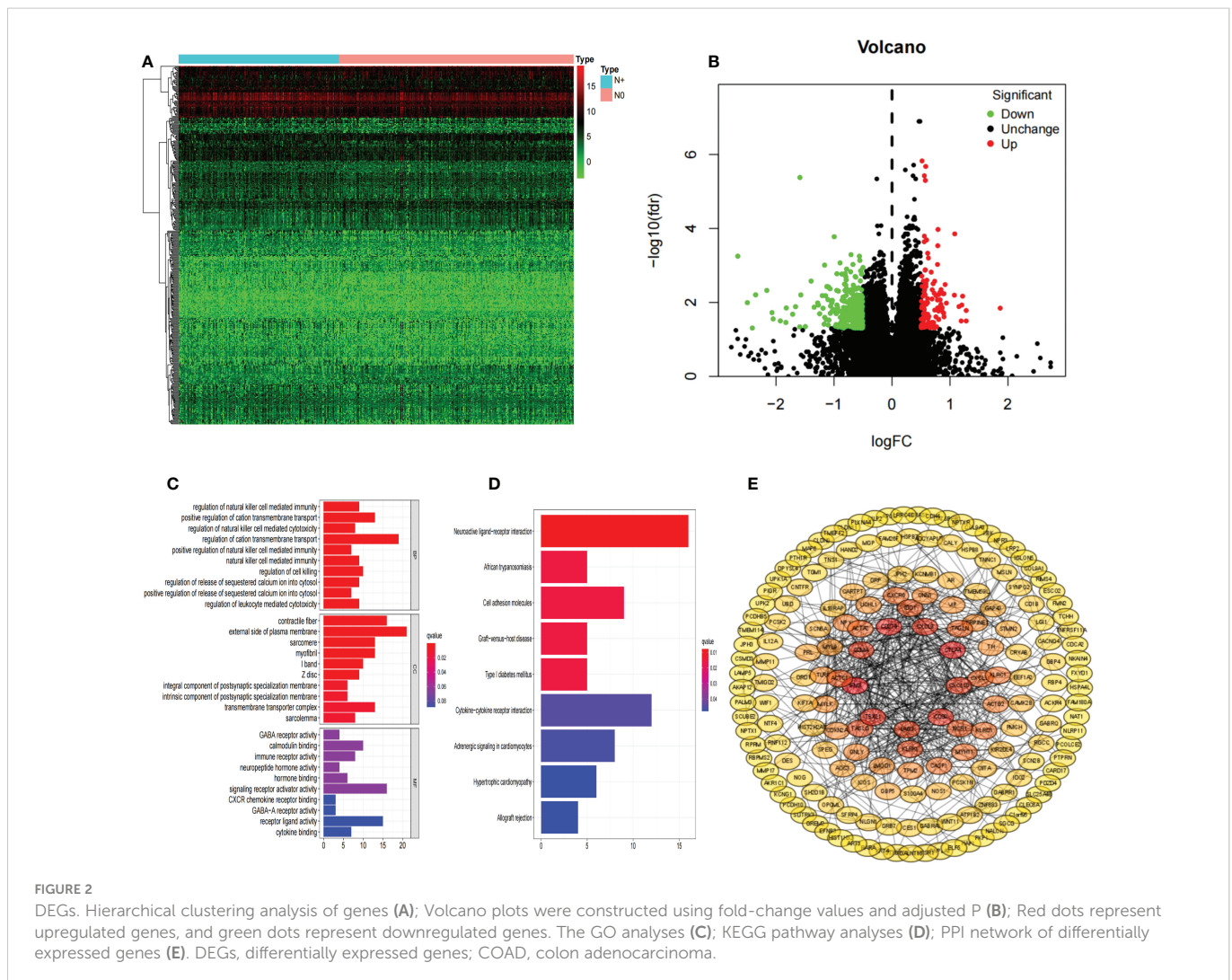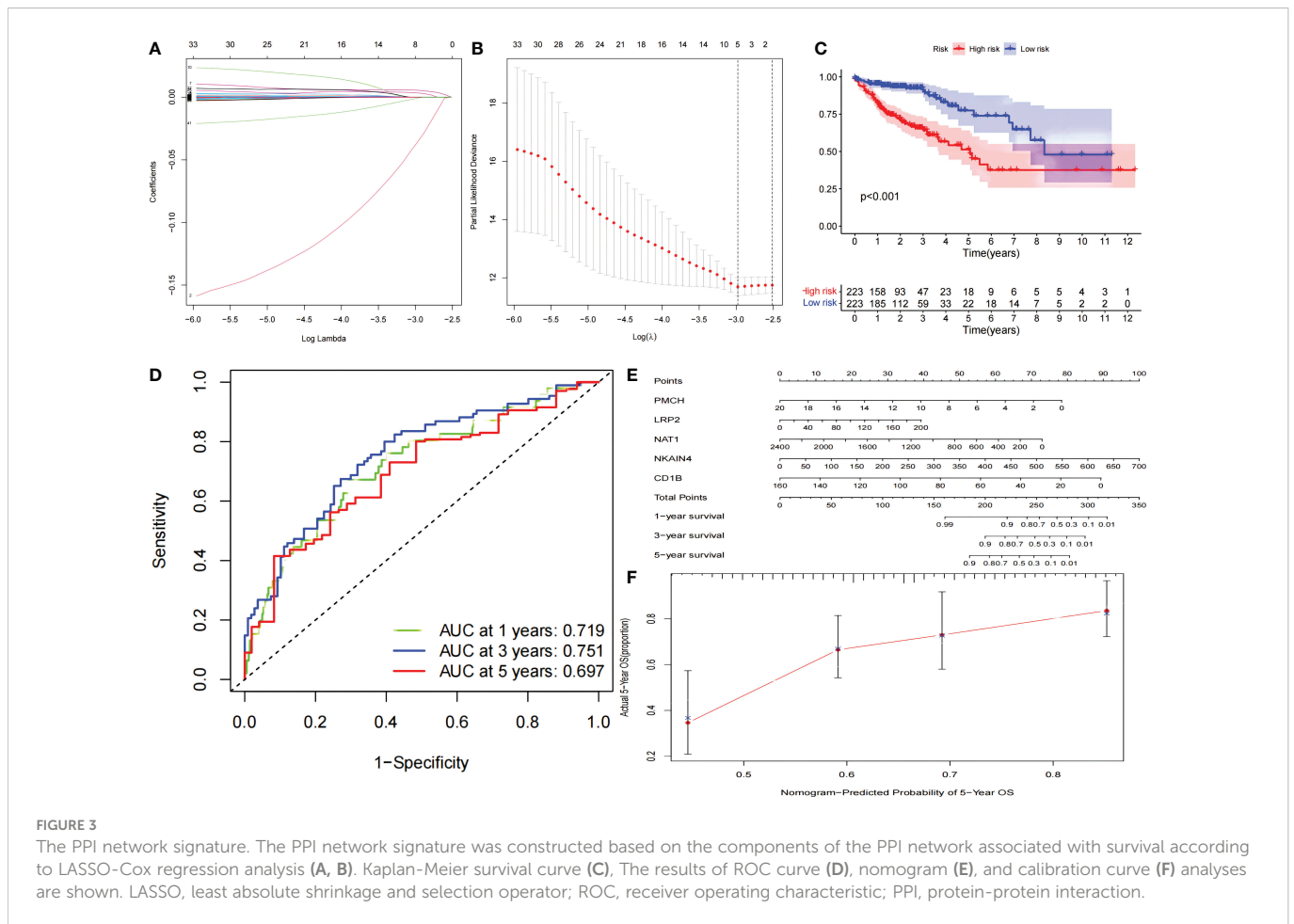
Results

DEGs

We identified 394 differentially expressed genes classified as upregulated (115) or downregulated (279) in the COAD N+ group versus N0 group tissues (Figures 2A, B). We next analyzed DEGs by GO and enrichment in KEGG pathways (Figures 2C, D).

PPI network construction and survival analysis

A PPI network was constructed composed of 180 DEGs based on the STRING database (Figure 2E). Univariate Cox proportional hazards regression analysis, LASSO logistic regression analysis, multivariate Cox regression analysis model were used to screen out the characteristic genes associated with survival (Figures 3A, B). The P-value, HR, and 95% CI of each variable are shown in Table 2. The performance of the risk score was evaluated by dividing the COAD samples in the TCGA into two subgroups, high-risk and low-risk, using the median risk score (1.14) calculated by LASSO as a cutoff. Risk stratification by risk score showed that patients in the high-risk subgroup had a shorter OS than those in the low-risk subgroup (Figure 3C). ROC curve showed that OS of COAD patients was perfectly predicted by the risk score [area under the curve (AUC) for 1-year survival =0.719; AUC for 3-year survival =0.751; and AUC for 5-year survival =0.697] (Figure 3D). A nomogram predicted the probability of 1-, 3-, and 5-year OS (Figure 3E). The calibration curve indicated that prediction using the nomogram was accurate (Figure 3F).

FIGURE 2

DEGs. Hierarchical clustering analysis of genes (A); Volcano plots were constructed using fold-change values and adjusted P (B); Red dots represent upregulated genes, and green dots represent downregulated genes. The GO analyses (C); KEGG pathway analyses (D); PPI network of differentially expressed genes (E). DEGs, differentially expressed genes; COAD, colon adenocarcinoma.

FIGURE 3

The PPI network signature. The PPI network signature was constructed based on the components of the PPI network associated with survival according to LASSO-Cox regression analysis **(A, B)**. Kaplan-Meier survival curve **(C)**, The results of ROC curve **(D)**, nomogram **(E)**, and calibration curve **(F)** analyses are shown. LASSO, least absolute shrinkage and selection operator; ROC, receiver operating characteristic; PPI, protein-protein interaction.

Immune cell infiltration

We used CIBERSORT to calculate the proportions of 22 immune cell subsets in each COAD sample downloaded from TCGA. Barplots and heatmaps were used to visualize the data (Figures 4A, B). The proportions of T follicular helper cells (P=0.012) and monocytes (P=0.002) were different in the N+ group and N0 group tissues (Figure 4C). We evaluated the correlations in the immune cell populations (Figure 4D). Finally, to further explore the potential role of the PPI network and immune cell signature, we evaluated the correlations between our risk signature and immune cell populations. LRP2 and T follicular helper cells (R=0.34, P=0.0019), and NKAIN4 and T follicular helper cells (R=0.23, P=0.041) had significant correlations (Figures 5A–C).

Relationship between LNM-related genes and immune checkpoints in COAD

Intriguingly, the genomic investigation revealed that LRP2 and NKAIN4 are involved in the alteration of immune checkpoints in COAD. The general landscape of LRP2, NKAIN4, and immune checkpoint alteration in COAD was compactly visualized, including fusion, amplification, deep deletion, truncating, and missense mutations (Figure 6A). The detailed relationship between LRP2, NKAIN4, and each representative immune checkpoint was

individually presented, as indicated in Figure 6B. Of note, the LRP alteration showed a statistically significant co-occurrence rather than mutual exclusivity with extensive immune checkpoints, such as SYNE1, OBSCN, TTN, MUC17, MUC16, PIK3CA, and FLG. The NKAIN4 alteration showed a statistically significant co-occurrence rather than mutual exclusivity with extensive immune checkpoints, such as OBSCN, MUC16, and MUC17. These findings strongly indicate that LRP2 and NKAIN4 are potential coregulators of immune checkpoints in COAD.

Discussion

Colon adenocarcinoma is one common malignancy of the digestive tract worldwide. Radiation therapy, chemotherapy and immunotherapy have been included in the treatment in addition to surgery (13, 16). Immunotherapy has become the latest research hotspot in recent years, and new immune therapeutic targets have also been identified (17, 18). Lymph node metastasis is one of the vital elements of the TNM stage. The present study identified key genes and constructed a prognostic signature to elucidate the mechanism of lymph node metastasis in COAD and shed new light on the development of immunological therapy for COAD patients.

In the present study, we identified 394 differentially expressed genes in COAD N+ group versus N0 group tissues and constructed a PPI network. Based on the components in the PPI network, the

TABLE 2 LASSO-Cox analysis.

ID	HR	95% CI	p value	coef
PMCH	0.858	0.756-0.975	0.019	-0.152169986
LRP2	1.007	1.001-1.013	0.011	0.007619784
NAT1	0.998	0.997-0.999	0.029	-0.001180128
NKAIN4	1.005	1.002-1.008	<0.001	0.00554486
CD1B	0.978	0.958-0.999	0.043	-0.021669547
HR, hazard ratio; CI, Confidence Interval.				

prognostic risk model for COAD patients was constructed. Ultimately, five genes(PMCH, CD1B, NAT1, NKAIN4, and LRP2, P<0.05) were incorporated into the final prognostic signatures. GO analysis revealed that DEGs were mainly enriched in immune-related pathways. Therefore, the correlation analysis between LNM-related signature and immune infiltrating cells was then investigated. We found that Tfh cells in COAD were associated with NKAIN4 and LRP2 expression.

T follicular helper (Tfh) cells are the main B cell-helper in mediating B cell differentiation in the germinal center (19–21). Tfh cells were associated with poor prognosis in Kidney Renal Clear Cell Carcinoma (22). Higher infiltrating fractions of memory Tfh cells were present in the high-risk prognostic group in hepatocellular carcinoma (23). The exhaustion of Tfh cells induced by intra-tumoral PDL1 resulted in impaired B cell function, facilitating the progression of advanced HCC (24). In addition, Tfh cells abundance improves the prognosis of patients with breast cancer or colorectal cancer (25, 26). Tfh cells are likely to be associated with the immune tolerance of tumors. We infer that the proportion of Tfh cells failed to serve as independent prognostic markers because COAD tissue with lymph node metastasis may be highly immune-tolerant, thereby inhibiting the immune effect of B cells. This mechanism contributes to tumor invasion and metastasis.

PMCH(promelanin-concentrating hormone) may act as a neurotransmitter or neuromodulator in neuronal functions that regulate goal-directed behavior, such as food intake and general arousal (27). In renal clear cell carcinoma, the low expression of PMCH is significantly related to higher tumor stages, more chances of metastasis, and poor prognosis (28). This study found that the high expression of the PMCH gene is a good prognostic factor for the survival of COAD patients.

Findings have shown that leukemia and several solid tumors contain a rare subpopulation of cells with specific surface CD markers that has the potential to self-renew, sustain cancer growth, and predict prognosis (29, 30). Low CD1B expression correlated with poorer biochemical recurrence-free survival in prostate cancer. CD1B confers prostate cancer progression and may help improve clinical prognostic stratification (31). The TME-related signature consisting of CD1B(HR 0.027, 95%CI 0.002-0.284, p=0.003) is effective to predict survival and to provide individualized treatment for COAD patients (32). The high expression of CD1B exhibited a significant correlation with longer OS in colorectal cancer (33). Guo et al. found that CD1B, which is involved in resting and activated dendritic cells, may be potential novel biomarkers for immunotherapy (34). The current

research results on CD1B are consistent with the results of our research, indicating that the high expression of CD1B is a good prognostic factor for COAD. Moreover, CD1B is related to resting and activated dendritic cells and may be used as an immunotherapy target.

The NAT1 gene, located on chromosome 8p22, encodes one of two human enzymes known to metabolize arylamine- and hydrazine-type drugs (35). In several cancers, the expression of NAT1 is associated with cell proliferation in vitro and with survival in vivo (36). A previous study has reported that low NAT1 expression resulted in a distinctly poor response to chemotherapy in breast cancer patients (37). NAT1 showed significant positive correlations with OS and could be considered to be protective genes in COAD (38). In COAD, NAT1 has been found to be enriched in the caffeine metabolism pathway (39). NAT1 may function as a tumor suppressor gene, and that a high level of methylation could reduce the mRNA expression of this gene in COAD tissues (40).

The high expression of NKAIN4 and LRP2 is related to the poor prognosis of COAD patients and has a co-expression relationship with the immune infiltrating cells related to lymph node metastasis.

NKAIN4 is named for its interaction with the β1 subunit of Na (+)/K(+) transporting ATPase (sodium-potassium pump). In recent years, the role of the sodium-potassium pump in tumors has attracted the attention of related scholars (41). As a transmembrane protein complex, a sodium-potassium pump can remove 3 Na+ from the cell by decomposing 1 ATP and moving 2 K+ into the cell. This function is necessary for cells to maintain osmotic pressure and establish an electrochemical gradient (42). In recent years, studies have found that sodium-potassium pumps also act as signal transduction molecules, playing an important role in cell growth, differentiation, survival, metastasis, invasion, and intercellular and cell-to-interstitial connections forming a scaffold for protein interactions (43). The role of the sodium-potassium pump in tumors has attracted widespread attention, and many studies have found that its activity is enhanced in a variety of tumors (44). Some scholars have tried to use the sodium-potassium pump subunit as a tumor biomarker. It is proposed that the subunit is highly expressed in lung cancer, kidney cancer, colorectal cancer, and medulloblastoma, and the expression of β subunit is increased in cervical cancer and colorectal cancer (45–47). We speculate that in COAD, NKAIN4 further promotes the activity and function of the sodium-potassium pump by activating the expression of the β1 subunit, thereby promoting tumor proliferation and progression. In our research, we also found that the expression of NKAIN is correlated with Tfh cells, although the correlation between
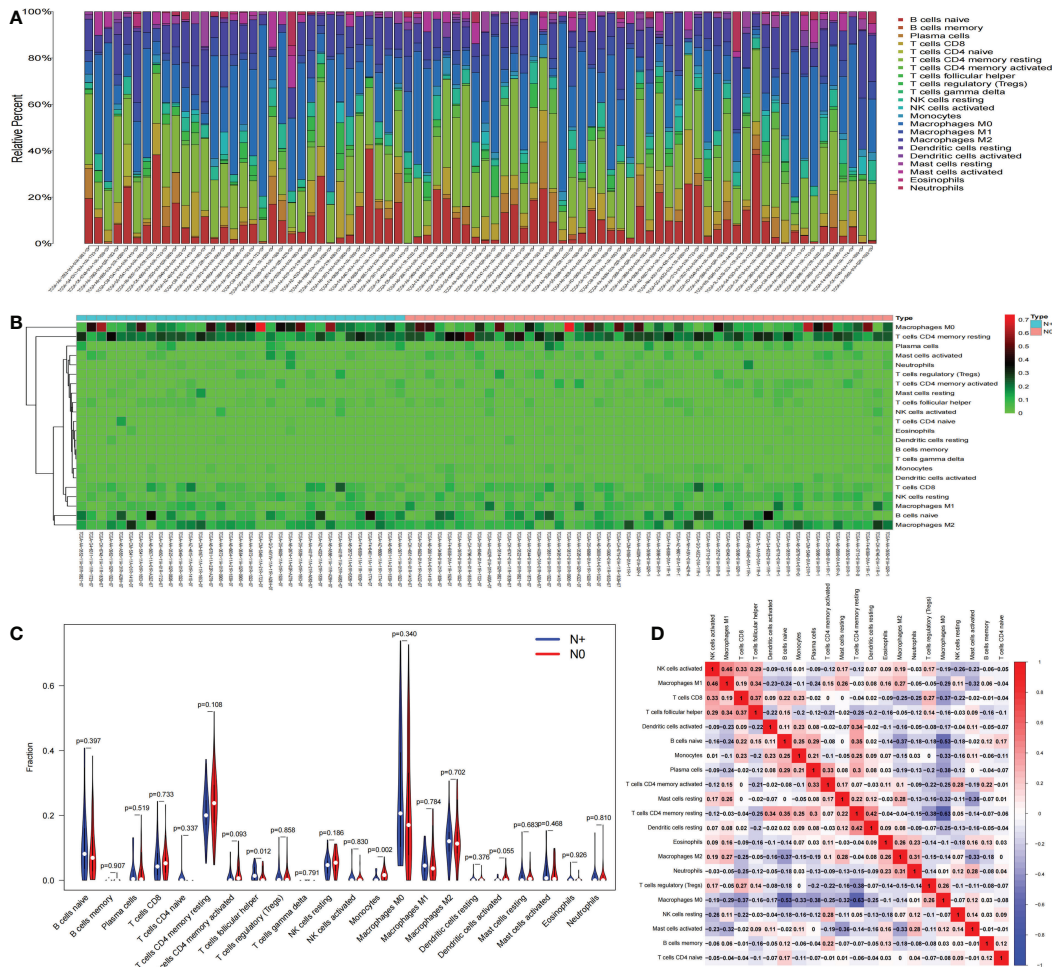
FIGURE 4
Immune cell infiltration. The composition of immune cells was assessed by the CIBERSORT algorithm in COAD **(A, B)**. The proportions of T follicular helper cells and monocytes were different in N+ and N0 group tissues **(C)**. The results of the correlation analysis of immune infiltrating cells are shown as a heatmap **(D)**. CIBERSORT, cell-type identification by estimating relative subsets of RNA transcripts; COAD, colon adenocarcinoma.

them is not so strong(R=0.23, P<0.05), suggesting that there may be some regulation between NKAIN and Tfh cells in COAD.

LRP2(LDL receptor-related protein 2), a member of the LDL receptor family, is a crucial regulator in the sonic hedgehog pathway, an important developmental process (48). Megalin, encoded by the gene LRP2, is an abundantly present cell surface protein in the kidney that binds DBP to mediate the internalization of 25(OH)D into the cytosol (49). Calcitriol may prevent cancer progression by multiple levels, including reducing cell proliferation, increasing cell differentiation, and apoptosis (50–52). Calcitriol suppresses
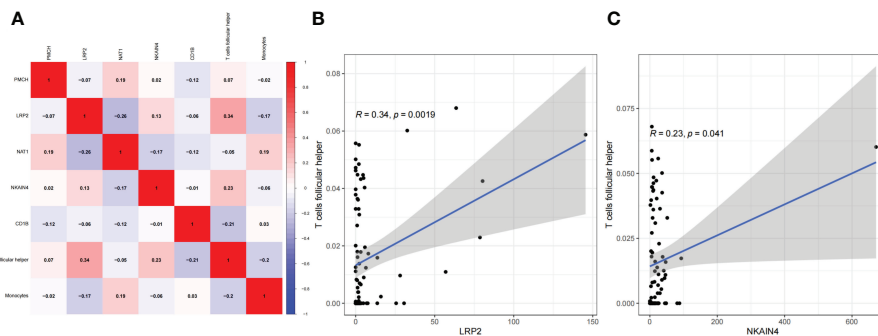

FIGURE 5
Coexpression analysis. The results of the correlation analysis of LNM-related genes expression and immune infiltrating cell are shown as a heatmap **(A)**. The correlations between LRP2 and T follicular helper cells, and NKAIN4 and T follicular helper cells **(B, C)**; LNM,lymph node metastasis.

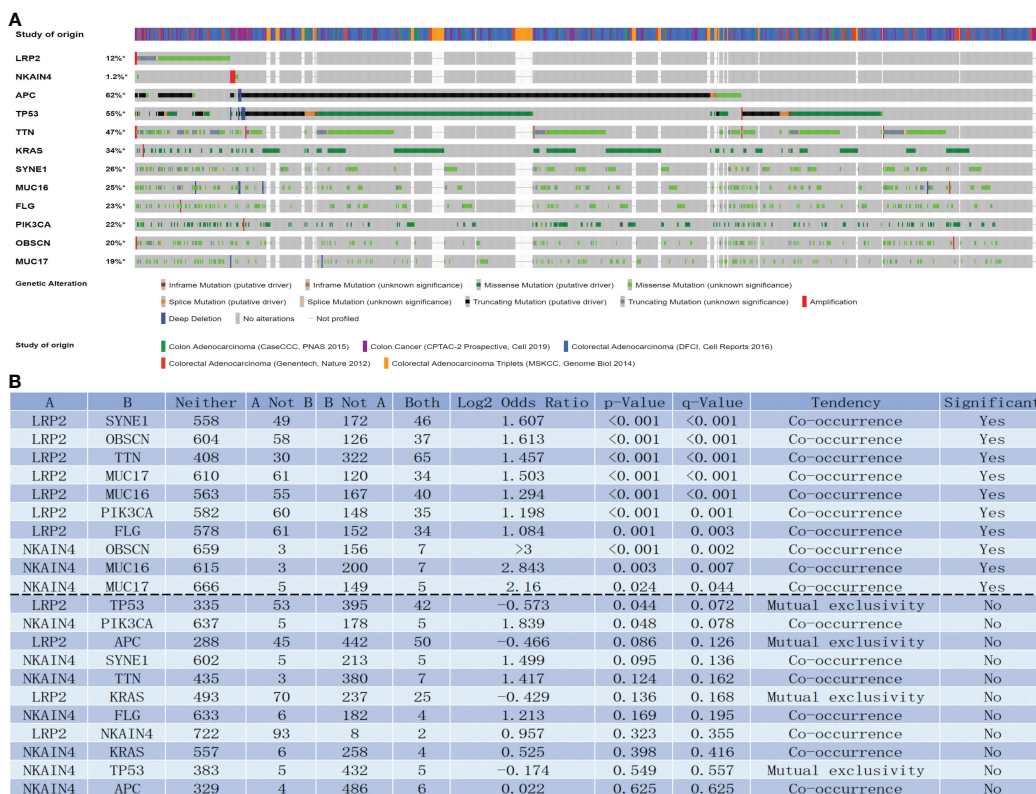| A	B	Neither	A Not B	B Not A	Both	Log2 Odds Ratio	p-Value	q-Value	Tendency	Significant
LRP2	SYNE1	558	49	172	46	1.607	<0.001	<0.001	Co-occurrence	Yes
LRP2	OBSCN	604	58	126	37	1.613	<0.001	<0.001	Co-occurrence	Yes
LRP2	TTN	408	30	322	65	1.457	<0.001	<0.001	Co-occurrence	Yes
LRP2	MUC17	610	61	120	34	1.503	<0.001	<0.001	Co-occurrence	Yes
LRP2	MUC16	563	55	167	40	1.294	<0.001	<0.001	Co-occurrence	Yes
LRP2	PIK3CA	582	60	148	35	1.198	<0.001	0.001	Co-occurrence	Yes
LRP2	FLG	578	61	152	34	1.084	0.001	0.003	Co-occurrence	Yes
NKAIN4	OBSCN	659	3	156	7	>3	<0.001	0.002	Co-occurrence	Yes
NKAIN4	MUC16	615	3	200	7	2.843	0.003	0.007	Co-occurrence	Yes
NKAIN4	MUC17	666	5	149	5	2.16	0.024	0.044	Co-occurrence	Yes
LRP2	TP53	335	53	395	42	-0.573	0.044	0.072	Mutual exclusivity	No
NKAIN4	PIK3CA	637	5	178	5	1.839	0.048	0.078	Co-occurrence	No
LRP2	APC	288	45	442	50	-0.466	0.086	0.126	Mutual exclusivity	No
NKAIN4	SYNE1	602	5	213	5	1.499	0.095	0.136	Co-occurrence	No
NKAIN4	TTN	435	3	380	7	1.417	0.124	0.162	Co-occurrence	No
LRP2	KRAS	493	70	237	25	-0.429	0.136	0.168	Mutual exclusivity	No
NKAIN4	FLG	633	6	182	4	1.213	0.169	0.195	Co-occurrence	No
LRP2	NKAIN4	722	93	8	2	0.957	0.323	0.355	Co-occurrence	No
NKAIN4	KRAS	557	6	258	4	0.525	0.398	0.416	Co-occurrence	No
NKAIN4	TP53	383	5	432	5	-0.174	0.549	0.557	Mutual exclusivity	No
NKAIN4	APC	329	4	486	6	0.022	0.625	0.625	Co-occurrence	No

FIGURE 6
LNM-related genes with immune checkpoints in COAD. LNM-related genes and immune checkpoint alteration in COAD. Compact visualization of cases with multiple genetic alterations of LNM-related genes and immune checkpoints (origined from five studies) were individually shown by cBioPortal as indicated, including fusion, amplification, deep deletion, truncating mutation, and missense mutation **(A)**. Mutual-exclusivity analysis between LNM-related genes and multiple-immune checkpoints in COAD. The altered relationship between LNM-related genes and each immune checkpoint, such as co-occurrence and mutual exclusivity, was presented as indicated **(B)**. The detailed log2 odds ratio, P-value, Q-value, tendency, and significance were individually presented in each panel. Q-value < 0.05 was considered to be statistically significant (indicated as yes, and others as no); LNM, lymph node metastasis; COAD, colon adenocarcinoma.

glycolysis and cell growth in human colorectal cancer cells, suggesting an inhibitory role of the biologically active form of vitamin D in colorectal cancer progression (53). Therefore, the reduction of vitamin D may promote lymph node metastasis in COAD. The SNP of LRP2 can cause a decrease in vitamin D (54), but we did not find a significant difference in mutation frequency in our sample (Figure S1). Thus, we infer that the relevant mechanism between LRP2 and COAD lymph node metastasis may not be related to reducing vitamin D but to other effects of LRP2, such as the reuptake of many ligands and signal transduction. Among other tumors, there are also some studies on LRP2. Higher expression of LRP2 was reported as a favorable prognostic factor in renal cell carcinoma (55). Germ-line polymorphisms in the LRP2 gene may be associated with an increased risk for recurrence in prostate cancer (56). Somatic mutations in LRP2 have been identified in gastric cancer (57). It can be seen that the research on the invasion and prognosis of LRP in tumor tissues is still inconclusive. This study found that LRP2 is related to the worse survival rate of COAD patients and is related to Tfh cells.

In COAD tissues with lymph node metastasis, there are more infiltrating Tfh cells. Moreover, the LRP2 gene and NKAIN gene have a significant correlation with Tfh cells. We preliminarily speculate that COAD tissues with lymph node metastasis may have a higher level of immune tolerance, which may cause increased Tfh cells infiltration. Further studies are warranted in connections between genes and Tfh cells. We also found that the NKAIN4 and LRP2 alteration showed a statistically significant co-occurrence rather than mutual exclusivity with extensive immune checkpoints. These findings strongly indicate that LRP2 and NKAIN4 are potential coregulators of immune checkpoints in COAD.

Our study shows that LRP2 and NKAIN4 genes can regulate the lymph node metastasis of COAD, and may regulate the infiltration of Tfh cells to mediate immune tolerance. LRP2, NKAIN4 and Tfh cells can show the migration ability of COAD cells, so the prognosis of patients with COAD can be predicted clinically by detecting the above three molecules. In addition, targeted drugs against LRP2 and NKAIN4 may also improve the immune tolerance of cancer cells, thereby reducing the proliferation, invasion and migration capabilities of cancer cells, and improving the prognosis of patients.

This study mainly has the following shortcomings: 1. The data used for analysis comes from public databases, there is no uniform criterion for data standardization, and the lack of clinical data in some

patients may cause bias. 2. No further experimental verification. 3. There are few studies on LRP2, NKAIN4, and Tfh cells in COAD, and the mechanism of action cannot be fully elucidated. The present study has the following strengths: 1.To the best of our knowledge, this is the first study on the relationship between LNM-related genes and immune infiltrating cells in COAD, which is innovative. 2. The clinical prediction model constructed has good predictive performance. 3. Our findings also provide ideas for future targeted therapy and immunotherapy.

This study constructed a prognostic model of genes related to lymph node metastasis in COAD and found that PMCH, CD1B, NAT1, NKAIN4, and LRP2 have prognostic, predictive values. Then, we performed a correlation analysis between genes with prognostic significance and differential immune infiltrating cells in COAD tissues. It is found that LRP2 and NKAIN4 are related to the infiltration of Tfh cells in COAD. Further studies have found that LRP2 and NKAIN4 are potential coregulators of immune checkpoints. This study helps to clarify the mechanism of lymph node metastasis in COAD, and in-depth research is needed to demonstrate and clarify the detailed mechanism.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding author.

Author contributions

HXZ, GBZ: Conception, Design of the work, Analysis and Writing. GWZ: Supervision, Reviewing and Editing. JXY: Writing, Reviewing and Editing. All authors contributed to the article and approved the submitted version.

Funding

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2022.907464/full#supplementary-material

SUPPLEMENTARY FIGURE 1
The mutation of LRP2 mRNA in different groups of samples, in which the horizontal axis represents the samples of different groups, the vertical axis represents the distribution of the mRNA mutation, where different colors represent different mutation types, and the right side represents significance, P value is the test method.

References

1. Siegel RL, Miller KD, Goding Sauer A, Fedewa SA, Butterly LF, Anderson JC, et al. Colorectal cancer statistics, 2020. *CA Cancer J Clin* (2020) 70(3):145–64. doi: 10.3322/caac.21601

2. Rawla P, Sunkara T, Barsouk A. Epidemiology of colorectal cancer: Incidence, mortality, survival, and risk factors. *Prz Gastroenterol* (2019) 14(2):89–103. doi: 10.5114/pg.2018.81072

3. Kekelidze M, D'Errico L, Pansini M, Tyndall A, Hohmann J. Colorectal cancer: current imaging methods and future perspectives for the diagnosis, staging and therapeutic response evaluation. *World J Gastroenterol* (2013) 19(46):8502–14. doi: 10.3748/wjg.v19.i46.8502

4. Huh JW, Kim YJ, Kim HR. Distribution of lymph node metastases is an independent predictor of survival for sigmoid colon and rectal cancer. *Ann Surg* (2012) 255(1):70–8. doi: 10.1097/SLA.0b013e31823785f6

5. Elferink MAG, Siesling S, Visser O, Rutten HJ, van Krieken J, Tollenaar R, et al. Large Variation between hospitals and pathology laboratories in lymph node evaluation in colon cancer and its impact on survival, a nationwide population-based study in the Netherlands. *Ann Oncol* (2011) 22(1):110–7. doi: 10.1093/annonc/mdq312

6. Pei JP, Zhang CD, Liang Y, Zhang C, Wu KZ, Zhao ZM, et al. Novel nomograms individually predicting overall survival of non-metastatic colon cancer patients. *Front Oncol* (2020) 10:733. doi: 10.3389/fonc.2020.00733

7. Neufert C, Heichler C, Brabletz T, Scheibe K, Boonsanay V, Greten FR, et al. Inducible mouse models of colon cancer for the analysis of sporadic and inflammation-driven tumor progression and lymph node metastasis. *Nat Protoc* (2021) 16(1):61–85. doi: 10.1038/s41596-020-00412-1

8. Li X, Liu B, Xiao J, Yuan Y, Ma J, Zhang Y. Roles of VEGF-c and Smad4 in the lymphangiogenesis, lymphatic metastasis, and prognosis in colon cancer. *J Gastrointest Surg* (2011) 15(11):2001–10. doi: 10.1007/s11605-011-1627-2

9. Gullotti L, Czerwitzki J, Kirfel J, Propping P, Rahner N, Steinke V, et al. FHL2 expression in peritumoural fibroblasts correlates with lymphatic metastasis in sporadic but not in HNPCC-associated colon cancer. *Lab Invest* (2011) 91(12):1695–705. doi: 10.1038/labinvest.2011.109

10. Shao YJ, Ni JJ, Wei SY, Weng XP, Shen MD, Jia YX, et al. IRF1-mediated immune cell infiltration is associated with metastasis in colon adenocarcinoma. *Med (Baltimore)* (2020) 99(37):e22170. doi: 10.1097/MD.0000000000022170

11. Giannakis M, Mu XJ, Shukla SA, Qian ZR, Cohen O, Nishihara R, et al. Genomic correlates of immune-cell infiltrates in colorectal carcinoma. *Cell Rep* (2016) 15(4):857–65. doi: 10.1016/j.celrep.2016.03.075

12. Seshagiri S, Stawiski EW, Durinck S, Modrusan Z, Storm EE, Conboy CB, et al. Recurrent r-spondin fusions in colon cancer. *Nature* (2012) 488(7413):660–4. doi: 10.1038/nature11282

13. Brannon AR, Vakiani E, Sylvester BE, Scott SN, McDermott G, Shah RH, et al. Comparative sequencing analysis reveals high genomic concordance between matched primary and metastatic colorectal cancer lesions. *Genome Biol* (2014) 15(8):454. doi: 10.1186/s13059-014-0454-7

14. Guda K, Veigl ML, Varadan V, Nosrati A, Ravi L, Lutterbaugh J, et al. Novel recurrently mutated genes in African American colon cancers. *Proc Natl Acad Sci U S A* (2015) 112(4):1149–54. doi: 10.1073/pnas.1417064112

15. Vasaikar S, Huang C, Wang X, Petyuk VA, Savage SR, Wen B, et al. Proteogenomic analysis of human colon cancer reveals new therapeutic opportunities. *Cell* (2019) 177 (4):1035–49 e19. doi: 10.1016/j.cell.2019.03.030

16. Le DT, Kim TW, Van Cutsem E, Geva R, Jager D, Hara H, et al. Phase II open-label study of pembrolizumab in treatment-refractory, microsatellite instability-High/ Mismatch repair-deficient metastatic colorectal cancer: KEYNOTE-164. *J Clin Oncol* (2020) 38(1):11–9. doi: 10.1200/JCO.19.02107

17. Salem M, Wallace C, Velegraki M, Li A, Ansa-Addo E, Metelli A, et al. GARP dampens cancer immunity by sustaining function and accumulation of regulatory T cells in the colon. *Cancer Res* (2019) 79(6):1178–90. doi: 10.1158/0008-5472.CAN-18-2623

18. Li H, Bullock K, Gurjao C, Braun D, Shukla SA, Bosse D, et al. Metabolomic adaptations and correlates of survival to immune checkpoint blockade. *Nat Commun* (2019) 10(1):4346. doi: 10.1038/s41467-019-12361-9

19. Linterman MA, Beaton L, Yu D, Ramiscal RR, Srivastava M, Hogan JJ, et al. IL-21 acts directly on b cells to regulate bcl-6 expression and germinal center responses. *J Exp Med* (2010) 207(2):353–63. doi: 10.1084/jem.20091738

20. Fazilleau N, Mark L, McHeyzer-Williams LJ, McHeyzer-Williams MG. Follicular helper T cells: lineage and location. *Immunity* (2009) 30(3):324–35. doi: 10.1016/j.immuni.2009.03.003

21. Crotty S. Follicular helper CD4 T cells (TFH). *Annu Rev Immunol* (2011) 29:621–63. doi: 10.1146/annurev-immunol-031210-101400

22. Deng L, Wang P, Qu Z, Liu N. The construction and analysis of ceRNA network and immune infiltration in kidney renal clear cell carcinoma. *Front Genet* (2021) 12:667610. doi: 10.3389/fgene.2021.667610

23. Tang B, Zhu J, Zhao Z, Lu C, Liu S, Fang S, et al. Diagnosis and prognosis models for hepatocellular carcinoma patient's management based on tumor mutation burden. *J Adv Res* (2021) 33:153–65. doi: 10.1016/j.jare.2021.01.018

24. Zhou ZQ, Tong DN, Guan J, Tan HW, Zhao LD, Zhu Y, et al. Follicular helper T cell exhaustion induced by PD-L1 expression in hepatocellular carcinoma results in impaired cytokine expression and b cell help, and is associated with advanced tumor stages. *Am J Trans Res* (2016) 8(7):2926.

25. Bindea G, Mlecnik B, Tosolini M, Kirilovsky A, Waldner M, Obenauf AC, et al. Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. *Immunity* (2013) 39(4):782–95. doi: 10.1016/j.immuni.2013.10.003

26. Gu-Trantien C, Loi S, Garaud S, Equeter C, Libin M, de Wind A, et al. CD4(+) follicular helper T cell infiltration predicts breast cancer survival. *J Clin Invest* (2013) 123 (7):2873–92. doi: 10.1172/JCI67428

27. Yang T, Kasagi S, Takahashi A, Mizusawa K. Effects of background color and feeding status on the expression of genes associated with body color regulation in the goldfish carassius auratus. *Gen Comp Endocrinol* (2021) 312:113860. doi: 10.1016/j.ygcen.2021.113860

28. Wang Y, Yang J, Zhang Q, Xia J, Wang Z. Extent and characteristics of immune infiltration in clear cell renal cell carcinoma and the prognostic value. *Transl Androl Urol* (2019) 8(6):609–18. doi: 10.21037/tau.2019.10.19

29. Brabletz T, Jung A, Spaderna S, Hlubek F, Kirchner T. Opinion: migrating cancer stem cells - an integrated concept of malignant tumour progression. *Nat Rev Cancer* (2005) 5(9):744–9. doi: 10.1038/nrc1694

30. Burkert J, Wright NA, Alison MR. Stem cells and cancer: an intimate relationship. *J Pathol* (2006) 209(3):287–97. doi: 10.1002/path.2016

31. Lee CH, Chen LC, Yu CC, Lin WH, Lin VC, Huang CY, et al. Prognostic value of CD1B in localised prostate cancer. *Int J Environ Res Public Health* (2019) 16(23):4723. doi: 10.3390/ijerph16234723

32. Chen W, Huang J, Xiong J, Fu P, Chen C, Liu Y, et al. Identification of a tumor microenvironment-related gene signature indicative of disease prognosis and treatment response in colon cancer. *Oxid Med Cell Longev* (2021) 2021:6290261. doi: 10.1155/2021/6290261

33. Wang Y, Li W, Jin X, Jiang X, Guo S, Xu F, et al. Identification of prognostic immune-related gene signature associated with tumor microenvironment of colorectal cancer. *BMC Cancer* (2021) 21(1):905. doi: 10.1186/s12885-021-08629-3

34. Guo J, Jin H, Xi Y, Guo J, Jin Y, Jiang D. The miR-582/CD1B axis is involved in regulation of dendritic cells and is associated with clinical outcomes in advanced lung adenocarcinoma. *BioMed Res Int* (2020) 2020:4360930. doi: 10.1155/2020/4360930

35. Butcher NJ, Minchin RF. Arylamine n-acetyltransferase 1: A novel drug target in cancer development. *Pharmacol Rev* (2012) 64(1):147–65. doi: 10.1124/pr.110.004275

36. Tiang JM, Butcher NJ, Cullinane C, Humbert PO, Minchin RF. RNAi-mediated knock-down of arylamine n-acetyltransferase-1 expression induces e-cadherin up-

regulation and cell-cell contact growth inhibition. *PloS One* (2011) 6(2):e17031. doi: 10.1371/journal.pone.0017031

37. Minchin RF, Butcher NJ. Trimodal distribution of arylamine n-acetyltransferase 1 mRNA in breast cancer tumors: association with overall survival and drug resistance. *BMC Genomics* (2018) 19(1):513. doi: 10.1186/s12864-018-4894-4

38. Zhang C, Zhao Z, Liu H, Yao S, Zhao D. Weighted gene Co-expression network analysis identified a novel thirteen-gene signature associated with progression, prognosis, and immune microenvironment of colon adenocarcinoma patients. *Front Genet* (2021) 12:657658. doi: 10.3389/fgene.2021.657658

39. Liu F, Ji F, Ji Y, Jiang Y, Sun X, Lu Y, et al. In-depth analysis of the critical genes and pathways in colorectal cancer. *Int J Mol Med* (2015) 36(4):923–30. doi: 10.3892/ijmm.2015.2298

40. Shi C, Xie LY, Tang YP, Long L, Li JL, Hu BL, et al. Hypermethylation of n-acetyltransferase 1 is a prognostic biomarker in colon adenocarcinoma. *Front Genet* (2019) 10:1097. doi: 10.3389/fgene.2019.01097

41. Mijatovic T, Dufrasne F, Kiss R. Na+/K+-ATPase and cancer. *Pharm Pat Anal* (2012) 1(1):91–106. doi: 10.4155/ppa.12.3

42. Mijatovic T, Van Quaquebeke E, Delest B, Debeir O, Darro F, Kiss R. Cardiotonic steroids on the road to anti-cancer therapy. *Biochim Biophys Acta* (2007) 1776(1):32–57. doi: 10.1016/j.bbcan.2007.06.002

43. Alevizopoulos K, Calogeropoulou T, Lang F, Stournaras C. Na+/K+ ATPase inhibitors in cancer. *Curr Drug Targets* (2014) 15(10):988–1000. doi: 10.2174/1389450115666140908125025

44. Wang HY, O'Doherty GA. Modulators of Na/K-ATPase: A patent review. *Expert Opin Ther Pat* (2012) 22(6):587–605. doi: 10.1517/13543776.2012.690033

45. Durlacher CT, Chow K, Chen XW, He ZX, Zhang X, Yang T, et al. Targeting Na (+)/K(+) -translocating adenosine triphosphatase in cancer treatment. *Clin Exp Pharmacol Physiol* (2015) 42(5):427–43. doi: 10.1111/1440-1681.12385

46. Baker Bechmann M, Rotoli D, Morales M, Maeso Mdel C, Garcia Mdel P, Avila J, et al. Na,K-ATPase isozymes in colorectal cancer and liver metastases. *Front Physiol* (2016) 7:9. doi: 10.3389/fphys.2016.00009

47. Seligson DB, Rajasekaran SA, Yu H, Liu X, Eeva M, Tze S, et al. Na,K-adenosine triphosphatase alpha1-subunit predicts survival of renal clear cell carcinoma. *J Urol* (2008) 179(1):338–45. doi: 10.1016/j.juro.2007.08.094

48. Christ A, Herzog K, Willnow TE. LRP2, an auxiliary receptor that controls sonic hedgehog signaling in development and disease. *Dev Dyn* (2016) 245(5):569–79. doi: 10.1002/dvdy.24394

49. Richards Z, Batai K, Farhat R, Shah E, Makowski A, Gann PH, et al. Prostatic compensation of the vitamin d axis in African American men. *JCI Insight* (2017) 2(2): e91054. doi: 10.1172/jci.insight.91054

50. DeLuca HF. Overview of general physiologic features and functions of vitamin d. *Am J Clin Nutr* (2004) 80(6 Suppl):1689S–96S. doi: 10.1093/ajcn/80.6.1689S

51. Iseki K, Tatsuta M, Uehara H, Iishi H, Ishiguro S. Inhibition of angiogenesis as a mechanism for inhibition by 1alpha-hydroxyvitamin D3 and 1,25-dihydroxyvitamin D3 of colon carcinogenesis induced by azoxymethane in wistar rats. *Int J Cancer* (1999) 81 (5):730–3. doi: 10.1002/(SICI)1097-0215(19990531)81:5<730::AID-IJC11>3.0.CO;2-Q

52. Mantell DJ, Owens PE, Bundred NJ, Mawer EB, Canfield AE. 1 alpha,25-dihydroxyvitamin D(3) inhibits angiogenesis *in vitro* and in vivo. *Circ Res* (2000) 87 (3):214–20. doi: 10.1161/01.RES.87.3.214

53. Huang CY, Weng YT, Li PC, Hsieh NT, Li CI, Liu HS, et al. Calcitriol suppresses warburg effect and cell growth in human colorectal cancer cells. *Life (Basel)* (2021) 11 (9):63. doi: 10.3390/life11090963

54. Shao B, Jiang S, Muyiduli X, Wang S, Mo M, Li M, et al. Vitamin d pathway gene polymorphisms influenced vitamin d level among pregnant women. *Clin Nutr* (2018) 37 (6 Pt A):2230–7. doi: 10.1016/j.clnu.2017.10.024

55. Gonias SL, Karimi-Mostowfi N, Murray SS, Mantuano E, Gilder AS. Expression of LDL receptor-related proteins (LRPs) in common solid malignancies correlates with patient survival. *PloS One* (2017) 12(10):e0186649. doi: 10.1371/journal.pone.0186649

56. Holt SK, Karyadi DM, Kwon EM, Stanford JL, Nelson PS, Ostrander EA. Association of megalin genetic polymorphisms with prostate cancer risk and prognosis. *Clin Cancer Res* (2008) 14(12):3823–31. doi: 10.1158/1078-0432.CCR-07-4566

57. Ge S, Li B, Li Y, Li Z, Lin S. Genomic alterations in advanced gastric cancer endoscopic biopsy samples using targeted next-generation sequencing. *Am J Cancer Res* (2017) 7(7):1540.