Research Paper

Whole-exome sequencing of rectal cancer identifies locally recurrent mutations in the Wnt pathway

Yi Yang[1,*], Xiaodong Gu[1,*], Zhenyang Li[1,*], Chuang Zheng[1], Zihao Wang[1], Minwei Zhou[1], Zongyou Chen[1], Mengzhen Li[2], Dongbing Li[2], Jianbin Xiang[1]

[1]Department of General Surgery, Huashan Hospital, Fudan University, Shanghai 200040, China
[2]MyGene Diagnostics Co., Ltd, Guangzhou 510000, China
*Equal contribution

Correspondence to: Dongbing Li, Jianbin Xiang; **email:** ldb06138233@163.com, https://orcid.org/0000-0002-5227-9643; xjbzhw@163.com, https://orcid.org/0000-0001-6150-9052

ABSTRACT

Locally recurrent rectal cancer (LRRC) leads to a poor prognosis and appears as a clinically predominant pattern of failure. In this research, whole-exome sequencing (WES) was performed on 21 samples from 8 patients to search for the molecular mechanisms of LRRC. The data was analyzed by bioinformatics. Gene Expression Profiling Interactive Analysis (GEPIA) and Human Protein Atlas (HPA) were performed to validate the candidate genes. Immunohistochemistry was used to detect the protein expression of LEF1 and CyclinD1 in LRRC, primary rectal cancer (PRC), and non-recurrent rectal cancer (NRRC) specimens. The results showed that LRRC, PRC, and NRRC had 668, 794, and 190 specific genes, respectively.
FGFR1 **and** ***MYC*** **have copy number variants (CNVs) in PRC and LRRC, respectively. LRRC specific genes were mainly enriched in positive regulation of transcription from RNA polymerase II promoter, plasma membrane, and ATP binding. The specific signaling pathways of LRRC were Wnt signaling pathway, gap junction, and glucagon signaling pathway, etc. The transcriptional and translational expression levels of genes including NFATC1, PRICKLE1, SOX17, and WNT6 related to Wnt signaling pathway were higher in rectal cancer (READ) tissues than normal rectal tissues. The** ***PRICKLE1*** **mutation (c.C875T) and** ***WNT6*** **mutation (c.G629A) were predicted as "D (deleterious)". Expression levels of LEF1 and cytokinin D1 proteins: LRRC > PRC > NRRC > normal rectal tissue. Gene variants in the Wnt signaling pathway may be critical for the development of LRRC. The present study may provide a basis for the prediction of LRRC and the development of new therapeutic drugs.**

INTRODUCTION

Locally recurrent rectal cancer (LRRC) is defined as the recurrence of rectal cancer (READ) only within the pelvis after radical resection [1]. LRRC is associated with symptoms of significant morbidity, such as rectal bleeding, bowel obstruction, chronic pain, fistulas, malodorous tumor discharge, tenesmus, and pelvic sepsis [2, 3]. LRRC generally occurs within 2-3 years after the initial surgery [4]. Non-recurrent rectal cancer (NRRC) is defined as READ that has

not recurred within 3 years after radical surgery. Historically, the introduction of total mesenteric excision (TME) has greatly reduced the local recurrence rate from 30% to 10% [3]. Combined with chemoradiotherapy, the local control (LC) rate for clinically localized READ ranges between 92% and 96% [5, 6]. Up to 50% of recurrent READ patients have synchronously diagnosed distant metastases [7, 8]. Meanwhile, nearly half of the patients with recurrence of READ are limited in the pelvis, which is defined as LRRC [8–10].

The causes and mechanism of LRRC have become a focus of current clinical medical research, highlighting the need for prevention of LRRC. Risk factors for LRRC occurrence include anastomotic leak, incomplete resection, intraoperative tumor perforation, high-grade pathology, and lack of adjuvant chemoradiotherapy [11]. LRRC is associated with short-term mortality, high reoperation rate, and additive healthcare costs [12]. Without treatment, the life expectancy of LRRC patients is limited and the quality of life is usually poor. Three major strategies for reducing the recurrence of READ are wider surgical coverage and neoadjuvant radiotherapy and neoadjuvant chemo-therapy [13, 14]. However, neoadjuvant chemo-radiation may significantly increase morbidity. Since LRRC portends a significantly worse oncologic outcome and quality of life, remarkable advances in surgical techniques and adjuvant chemoradiotherapy are improving the overall survival (OS) of LRRC. Although modern surgical techniques and multimodal therapies have proved to be effective, LRRC shows a significant treatment challenge. Some research showed a 5-year survival rate of 35% to 50% after surgery for LRRC [15, 16]. Chemotherapy alone can prolong OS to12 to 15 months, but it is not curable [8]. About 40% of LRRC patients have an opportunity for surgical treatment, but performing the operation can be a challenging option under such a situation [17]. Reoperation is very difficult and often needs to be combined with organ resection. The incidence of perioperative complications is specifically high, which is rare in domestic medical units. Therefore, how to prevent LRRC is the key to improve the survival and life quality of READ patients.

The long-term outcome of surgical treatment mainly depends on the clear margin resection [16]. We speculated that the molecular mechanism of LRRC may also be related to the occurrence and prognosis of LRRC. The signal pathway related to distant metastasis of READ has been studied in detail as a hot spot. However, the molecular mechanism of LRRC has not been reported yet. Whole-exome sequencing (WES) can reveal mutations within exon coding regions. This study was based on WES to explore mutations associated with LRRC. 5 trios of LRRC and 3 pairs of NRRC samples were sequenced by WES. The data was analyzed systematically by bioinformatics. Gene Expression Profiling Interactive Analysis (GEPIA) and Human Protein Atlas (HPA) were performed to validate the candidate genes. Immunohistochemistry was used to detect the protein expression of candidate genes in LRRC, PRC, and NRRC specimens. Our findings provided a basis for predicting LRRC and developing new therapeutic agents.

MATERIALS AND METHODS

Patients and specimens

The present study was reviewed and approved by the Ethics Committee of Huashan Hospital affiliated to Fudan University. We collected 21 samples, including 15 tissue samples from five LRRC patients (PRC, adjacent normal rectal tissue, and LRRC) and three NRRC patients (NRRC and adjacent normal rectal tissue) from May 2019 to November 2019 at Huashan Hospital, Fudan University. HE-stained sections from each sample were subjected to an independent pathology review to confirm that the tumor specimens were histologically consistent with LRRC and that the adjacent tissue specimens contained no tumor cells. All participants provided written informed consent for genetic analysis.

Whole-exome sequencing

For every individual, the genomic DNA of cells from one or two regions of the PRC, LRRC, and matched normal rectal tissue samples were sequenced. DNA was extracted from formalin-fixed and paraffin-embedded (FFPE) tissue blocks using MagPure FFPE DNA/RNA LQ Kit (Magen, Guangzhou, China). The DNA was then subjected to an additional quality and quantity evaluation step by utilizing NanoDrop ultra-micro spectrophotometer (Thermo Fisher Scientific, Wilmington, DE, USA) and Qubit 1X dsDNA HS Assay Kit (Thermo Fisher Scientific, Wilmington, DE, USA). Exome-coding DNA was captured with Agilent SSEL XTHS Human All Exon V6 (Agilent Technologies, Santa Clara, CA, USA), and the libraries were sequenced on an Illumina NovaSeq Platform (Illumina, San Diego, CA, USA), which produced 350-bp paired-end reads.

Sequencing data analysis

Raw sequencing data were filtered using the Trim Galore program to remove low quality reads at both ends of the sequencing [18]. The FastQC package (http://www.bioinformatics.babraham.ac.uk/projects/fastqc) was used to assess the quality score distribution of the sequencing reads. Read sequences were mapped to the human reference genome (GRCH37/hg19) using Burrows-Wheeler Aligner (BWA, v.0.7.15) with the default parameters, and duplicates were marked and discarded using Picard (http://broadinstitute.github.io/picard). After alignment by BWA, the reads were subjected to local realignment and recalibration using the Genome Analysis Toolkit (GATK). Variants and genotypes calling were also performed using Genome Analysis Toolkit (GATK, v3.7).

Point substitutions in all tumor samples were counted and the proportion of each point substitution (C> A, C> G, C> T, T> A, T> C, and T> G) was calculated for both PRC and LRRC samples. The results were visualized by GraphPad Prism 8.

ANNOVAR was utilized to annotate all called variants. The SNVs of the adjacent normal rectal tissue in PRC, LRRC, and NRRC samples were removed for further analysis. Mutations were screened before being subjected to pathway enrichment analysis. The minor allele frequency (MAF) of variants was evaluated in the 1000 Genome Project (http://www.ncbi.nlm.nih.gov/variation/tools/1000genomes/), ExAC (http://exac.broadinstitute.org/), variants of MAF<1% and annotated "." (No annotation information in the database) were retained. Subsequently, the pathogenicity of variants was predicted according to SIFT (http://sift.jcvi.org), Polyphen2 (http://genetics.bwh.harvard.edu/pph2) and FATHMM (http://fathmm.biocompute.org.uk/). Variants predicted by the three tools as "D (damaging)" were retained.

CNVs were identified using open source software called CNVkit (v0.9.7), a tool kit to infer and visualize the copy numbers from targeted DNA sequencing data. Genes of copy number over "3" were selected, the copy number of "2" meant no CNV, the copy number over "2" meant a copy gain in some paired chromosomes. The CNVs of the adjacent normal rectal tissue in PRC, LRRC, and NRRC samples were removed for further analysis.

Enrichment analyses of gene ontology (GO) and the Kyoto Encyclopedia of Genes and Genomes (KEGG)

The specific genes of PRC, LRRC, and NRRC samples were used to perform GO and KEGG enrichment analysis by DAVID 6.8 (https://david.ncifcrf.gov/). The results were considered as statistically significant if *P* value< 0.05. The top 15 enriched GO terms were visualized by BMKCloud (http://www.biocloud.net/), a free online platform for data analysis. The top 15 KEGG pathways were visualized by R ggplot2 package. GO functional analysis was divided into three parts: biological process (BP), cellular component (CC), and molecular function (MF).

Mutations of driver genes

To look for potential LRRC drivers, mutations in 155 TCGA rectal adenocarcinoma (https://www.cbioportal.org/) and 339 MSK READ patients were included in the analysis for comparison. Forty-eight genes with a high frequency of occurrence in both cases (TCGA Freq ≥ 10%; MSK Freq ≥ 5%) were selected. The mutation patterns of these genes in PRC and LRRC tissues were then compared.

mRNA and protein expression levels of hub genes

To validate the expression of the key DEGs, the Gene Expression Profiling Interactive Analysis (GEPIA) website (http://gepia2.cancer-pku.cn/#index) was applied to analyze the data of RNA sequencing expression based on thousands of samples from the GTEx project and TCGA [19]. The association between overall survival (OS) and the genes expressed in GC patients was determined using GEPIA. The lower and upper 50% of gene expression were set as the standard for analysis. Log-rank test results with P<0.05 were regarded as statistically significant. Besides, the GEPIA was employed to visualize the mRNA expression of hub genes in tumors and normal samples.

The Human Protein atlas (HPA) database (https://www.proteinatlas.org/) is a free online database that provides abundant transcriptome and proteome data on human normal or pathological tissues through RNA-sequence analysis and immunohistochemical analysis. In the present study, the protein expression and distribution of hub genes were investigated in GC tissues and compared with normal tissues in HPA [18].

Immunohistochemistry

The Formalin-Fixed and Paraffin-Embedded (FFPE) samples were used in this study. 3-mm tumor sections were incubated with commercial rabbit polyclonal antibodies against LEF1 (ab137872, abcam) and cyclin D1 (ab16663, abcam) at 1/100 dilution overnight at 4° C. Then, the sections were conjugated with HRP-Sheep Anti-Rabbit IgG-HRP-Sheep Anti-Mouse IgG antibody (BOSTER BA1056, 1:500 dilution;) at room temperature for 2 h, then covered by 3, 3-diaminobenzidine (DAB) (Vector Laboratories, Burlingame, CA). Elivision plus kit for immunohistochemistry (IHC) (KIT-990, MXB) was used in this study.

Statistical analyses

Fisher's exact test was used to assess differences in the count data. Top 15 significant GO terms of BP, MF, and CC were listed according to the p-value. Top 15 significant KEGG pathways were listed according to the p-value. P<0.05 was considered statistically significant.

RESULTS

WES data

The clinical characteristics of the eight patients are listed in Supplementary Table 1. Age range is 45-71

years. The gender included 8 males. The tumor size included 4 patients (2cm), 3 patients (3cm), and 1 patient (4cm), and 1 patient (5cm). The pathologic 4 stage I, 3 stage II, and 1 stage III. The sequencing quality of the WES was analyzed and the raw data, Q30 (proportion of mapped reads) and mean depth of each sample are shown. As shown in Supplementary Table 2, all samples had Q30 ratio >90%, good sequencing quality, >90% of mapped reads, and an average sequencing depth > 200 x, which was sufficient to identify mutations.

Mutational signatures of LRRC and PRC tumors

All point substitutions were divided into six groups (C>A, C>G, C>T, T>A, T>C and T>G) according to the direction of the mutation. In total, 97862 and 47292 somatic substitutions were identified in all PRC and LRRC tissues by comparing them with the matched adjacent normal samples. The proportion of each group was different, but the proportion of all six mutant groups did not differ significantly between LRRC and matched PRC tissues (Figure 1).

The common variations in LRRC and PRC tumors

After removing the control SNVs from PRC and LRRC samples, the focus was on common variations. In PRC, four genes, including *MYCN, SCRIB, SNAPC4, and MED15*, were identified with eight mutations (Table 1). Notably, the SCRIB mutation (c.T233G) was predicted as "D (deleterious)" by SIFT and Polyphen2 software. In the LRRC, two genes, including *SCRIB* and *RUSC2* (Table 2), were found to have three mutations. Interestingly, LRRC had the same *SCRIB* mutation as

Figure 1. Divergent mutational features in LRRC and PRC samples of LRRC patients. Six mutational subtypes in LRRC and PRC tissues. LRRC tissues are presented in orange, and the PRC tissues are presented in cyan.

PRC. In addition, two frame shift mutations were found in the *RUSC2* gene (c.831_833del, c.3465_3467del).

The biological process and pathways related to PRC, LRRC, and NRRC

In the PRC samples, there were 110,198 variants after removing the control mutations, 2388 variants after screening, and 1479 genes after further deduplication. In the LRRC samples, there were 100061 variants after removing the control variants, 2076 variants after screening, and 1360 genes after further deduplication. In the NRRC samples, there were 23067 variants after removing the control variants, 573 variants after screening, and 481 genes after further deduplication. The PRC, LRRC, and NRRC had 794, 668, and 190 specific genes, respectively (Figure 2). There were 472 likely pathogenic genes shared by the PRC and LRRC samples. The genes including 668 genes and 472 genes were important for the development of LRRC.

In PRC samples, GO functional analysis showed that the specific genes were mainly enriched in the biological process of positive regulation of negative regulation of apoptotic process and apoptotic cell clearance (Figure 3A and Supplementary Table 3). The genes were mainly enriched in signaling pathways such as proteoglycans in cancer, ECM-receptor interaction, ErbB signaling pathway, and protein digestion and absorption (Figure 3B and Supplementary Table 4). In LRRC samples, specific genes were mainly enriched in the biological process including angiogenesis and regulation of transcription from RNA polymerase II promoter (Figure 4A and Supplementary Table 5). The genes were mainly enriched in the signaling pathways such as glucagon signaling pathway, gap junction, axon guidance, and metabolic pathways (Figure 4B and Supplementary Table 6). In NRRC samples, specific genes were mainly enriched in the biological process including regulation of cell proliferation and positive regulation of extrinsic apoptotic signaling pathway in the absence of ligand (Figure 5A and Supplementary Table 7). The genes were mainly enriched in the signaling pathways such as peroxisome, metabolic pathways, and butanoate metabolism (Figure 5B and Supplementary Table 8).

Wnt signaling pathway is one of the specific pathways in LRRC

Venn analysis of the enriched KEGG pathway was performed on five sets of samples. As shown in Figure 6, there are 7 specific pathways in LRRC, including Wnt signaling pathway, Gap junction, Glucagon signaling pathway, Axon guidance, Thyroid hormone synthesis, Morphine addiction and Glycine, serine and threonine metabolism.
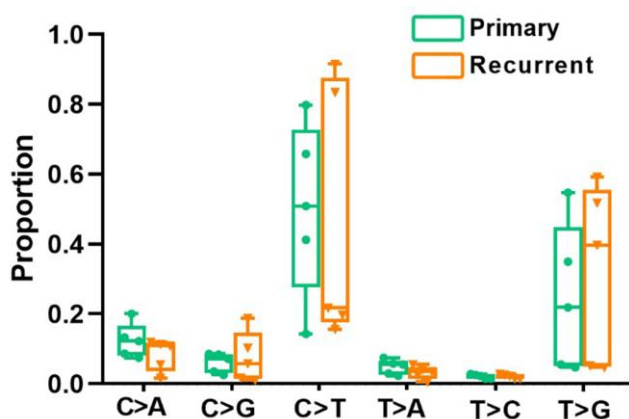
Table 1. Common variations in PRC samples.

Gene	Mutation type	Variants	1000g2015_eas	ExAC eas	SIFT	Polyphen2
MYCN	Frame shift insertion	c.87dupC	-	-	-	-
SCRIB	Nonsynonymous	c.233T>G	-	-	D	D
SNAPC4	Nonsynonymous	c.3158T>A	-	-	D	B
		c.1035G>C	-	-	-	-
		c.1197G>C				
MED15	Synonymous	c.1170G>C	-	-	-	-
		c.1248G>C				
		c.1368G>C				

Table 2. Common variations in LRRC samples.

Gene	Mutation type	Variants	1000g2015_eas	ExAC eas	SIFT	Polyphen2
SCRIB	Nonsynonymous	c.233T>G	-	-	D	D
RUSC2	Non-frameshift deletion	c.831_833del c.3465_3467del	-	-	-	-
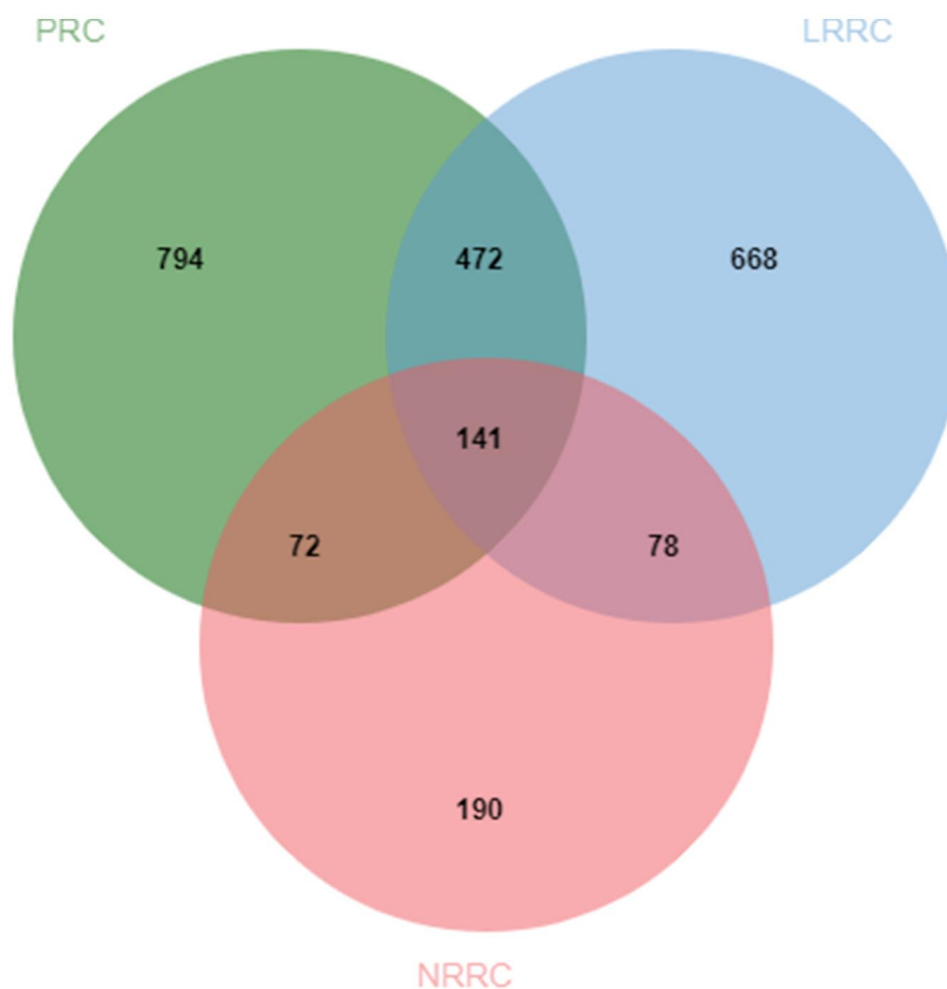
Figure 2. Venn diagram of likely pathogenic genes associated with PRC, LRRC, and NRRC. PRC, primary rectal cancer; NRRC, non-recurrent rectal cancer; LRRC, local recurrent rectal cancer.

The variations of genes related to Wnt signaling pathway in LRRC

In the present study, 24 genes were enriched in Wnt signaling pathway of LRRC samples (Table 3). And 12 genes were enriched only in Wnt signaling pathway of LRRC samples (Supplementary Table 6). The variations of the 12 genes were listed in Table 4. 15 variations of the 12 genes were predicted as "D (deleterious)" by SIFT, Polyphen2, and FATHMM software, which were critical to the development of LRRC.
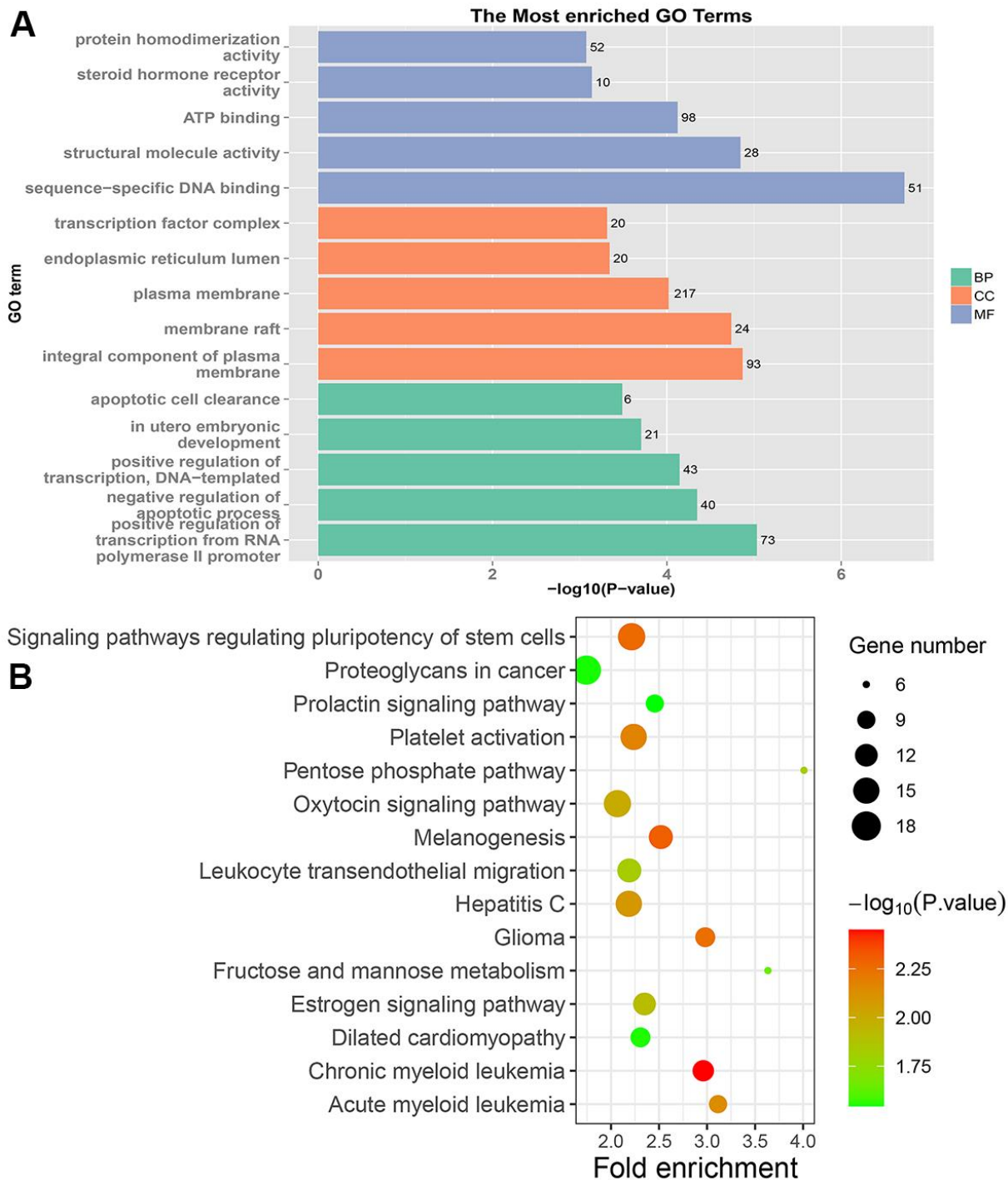

Figure 3. GO and KEGG enrichment analysis of PRC samples. (**A**) The top 15 enriched GO terms. (**B**) The top 15 significant KEGG pathways. Fold enrichment represents the degree of enrichment. The size of the bubble indicates the number of genes. The depth of bubble color indicates the level of significance.

Analysis of the genes related to Wnt signaling pathway via GEPIA and HPA databases

As shown in Figure 7, the expression levels of *NFATC1, PRICKLE1, SOX17,* and *WNT6* in CRC tissues were significantly lower than in normal tissues.

The protein expression levels of the genes in CRC were explored using the HPA database (Figure 8). The protein levels of *NFATC1* were not expressed in CRC tissues and normal tissues. The protein level of PRICKLE1 was not expressed in CRC tissues, whereas the low protein expression level of PRICKLE1 was
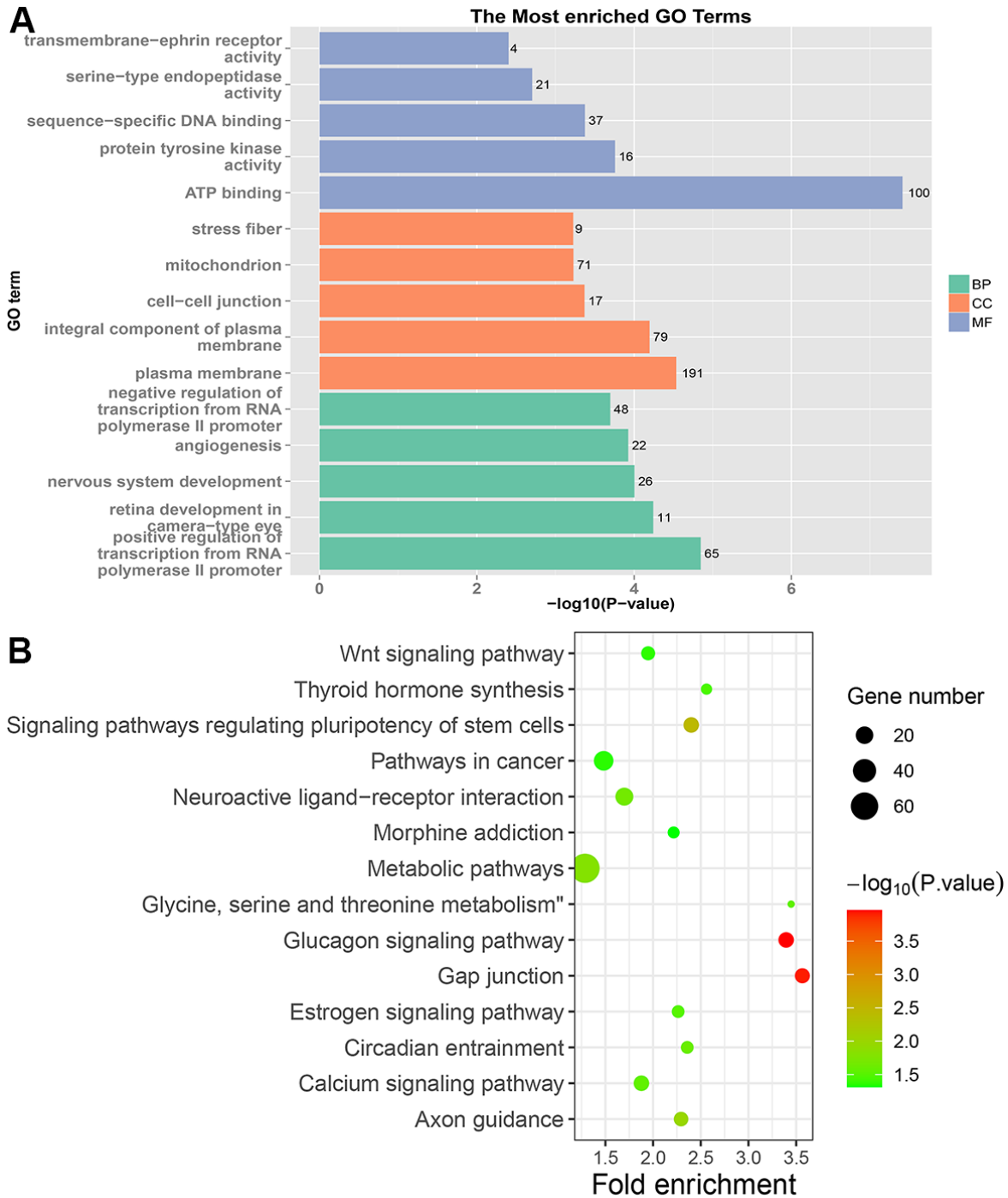

Figure 4. GO and KEGG enrichment analysis of LRRC samples. (**A**) The top 15 enriched GO terms. (**B**) The top 15 significant KEGG pathways. Fold enrichment represents the degree of enrichment. The size of the bubble indicates the number of genes. The depth of bubble color indicates the level of significance.

observed in normal tissues. The protein levels of SOX17 and WNT6 were not expressed in CRC tissues, whereas the low protein expression levels of SOX17 and WNT6 were observed in normal tissues. In summary, the present results indicated that the transcriptional and translational expression levels of the

hub genes were overexpressed in patients with CRC. *LEF1* and *cyclin D1* are key genes in the Wnt signaling pathway. IHC was performed to test LEF1 and cyclin D1 protein expression in adjacent normal rectal tissue, PRC and LRRC FFPE samples. As shown in Figures 9, 10, the expression of LEF1 and cyclin D1
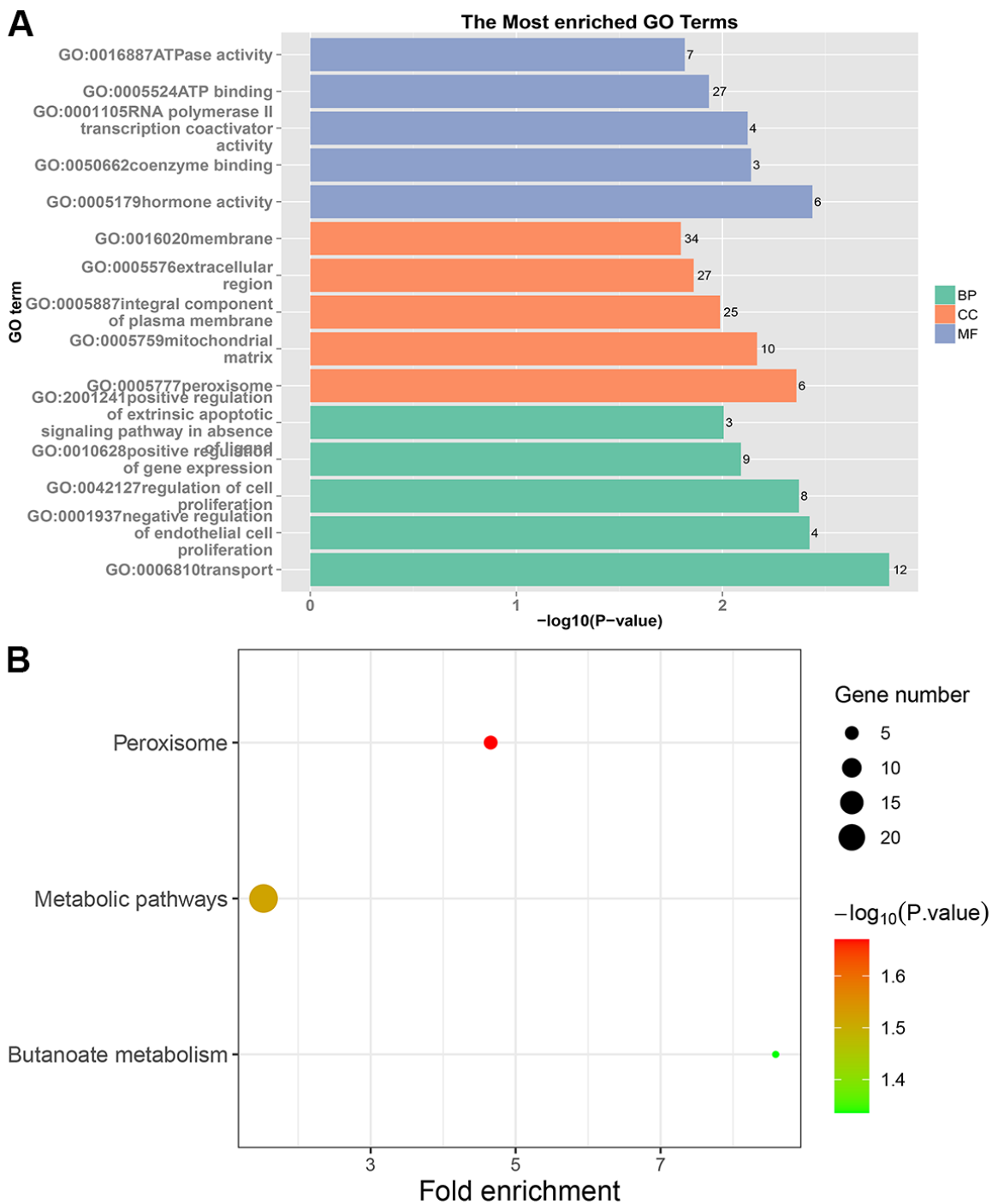

Figure 5. GO and KEGG enrichment analysis of NRRC samples. (**A**) The top 15 enriched GO terms. (**B**) The top 15 significant KEGG pathways. Fold enrichment represents the degree of enrichment. The size of the bubble indicates the number of genes. The depth of bubble color indicates the level of significance.

proteins was higher in PRC compared to adjacent normal rectal tissue; the expression of LEF1 and cyclin D1 proteins was higher in LRRC compared to PRC; and the expression of LEF1 and cyclin D1 proteins was higher in PRC compared to NRRC. The above results suggest that the Wnt signaling pathway may play an important role in the development of LRRC.

Analysis of driver genes in LRRC

Based on the database, 48 possible driver genes were selected. As shown in Figure 11, Mutations in the genes, including *MUC17, TTN, SYNE1, MUC16, FAT4, FLG, CSMD1, FAT3, RYR1, COL6A3, NEB, OBSCN, ZFHX4, TCF7L2, ERN3, NOTCH3, KMT2C,* and *PTRRT,* were detected in both LRRC and PRC samples from five LRRC patients. Mutations in PEG3, NRAS, and BRAF did not occur in the PRC samples of the five patients. In LRRC samples, PEG3 mutations were present in 4 samples, NRAS in 1 sample, and BRAF in 1 sample. The other 27 mutations were present in some of the LRRC or PRC samples.

The CNVs in PRC and LRRC

After removing the control CNVs from PRC and LRRC samples, the focus was on specific CNVs. As shown in Table 5, *FGFR1* and *MYC* were found to be genes specific for CNVs in PRC and LRRC, respectively. Besides, *FGFR3* CNVS were occurred in both PRC and LRRC. The copy numbers of *FGFR3* in PRC and LRRC tumors were 3.542 and 3.548, respectively. Due to the sample size, after removing control CNVs, no CNVs was found in NRRC. There were no significant CNVs related to Wnt pathway.

DISCUSSION

WES has a high sensitivity to common, rare, and low-frequency mutations. It could find the most disease-related mutations in the exon region, and only needs to sequence about 1% of the genome [20]. WES had been widely used in CRC to detect the mutational landscape of CRC populations and provided novel insights into the treatment and prognosis of CRC in the clinic [21].
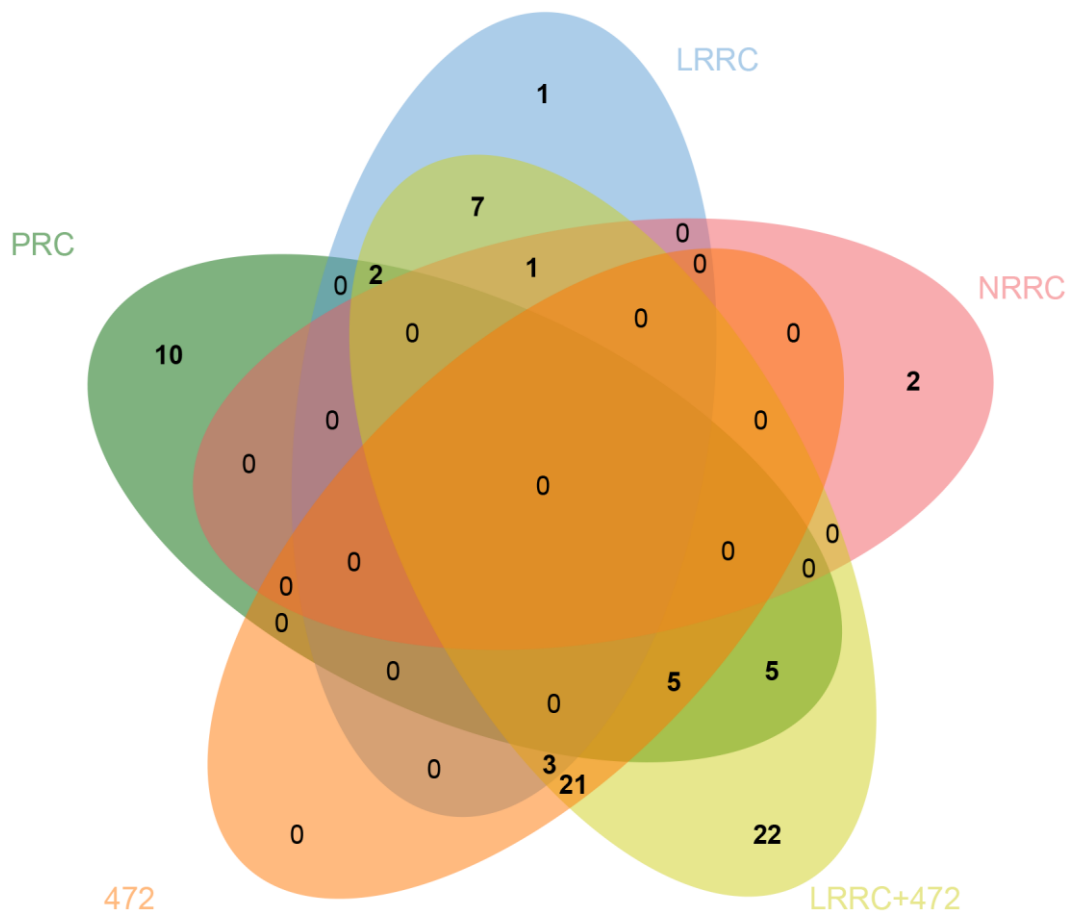

Figure 6. Venn diagram of specific pathways in PRC, LRRC, NRRC, 472, and LRRC+472. 472, 472 common genes; LRRC+472, LRRC+ 472 common genes.

Table 3. The specific pathways in LRRC samples.

Term	Count	%	P-Value
hsa04310: Wnt signaling pathway	24	0.013840352	0.000623
hsa04540: Gap junction	23	0.01326367	0.000001
hsa04922: Glucagon signaling pathway	22	0.012686989	0.000032
hsa04360: Axon guidance	22	0.012686989	0.001176
hsa04918: Thyroid hormone synthesis	16	0.009226901	0.000375
hsa05032: Morphine addiction	15	0.00865022	0.013851
hsa00260: Glycine, serine and threonine metabolism	9	0.005190132	0.011322

Table 4. Variations of genes related to Wnt signaling pathway in LRRC Samples.

| Gene	Mutation type	AAChange	SIFT	FATHMM	Frequency
CHD8	nonsynonymous	NM_001170629: exon30: c.G5564T: p.R1855L	D	D	0.012788
CHD8	nonsynonymous	NM_001170629: exon17: c.G3542A: p.G1181E	D	D	0.051282
FZD5	nonsynonymous	NM_003468: exon2: c.C728T: p.S243L	D	D	0.019231
FZD5	nonsynonymous	NM_003468: exon2: c.C155T: p.P52L	D	D	0.042857
NFATC1	nonsynonymous	NM_001278673: exon4: c.C236T: p.T79M	D	D	0.015842
PLCB3	nonsynonymous	NM_001184883: exon15: c.G1736A: p.R579H	D	D	0.018927
PRICKLE1	nonsynonymous	NM_001144881: exon7: c.C875T: p.P292L	D	D	0.051471
PRICKLE2	nonsynonymous	NM_198859: exon7: c.A908C: p.Q303P	D	D	0.023973
PRICKLE2	nonsynonymous	NM_198859: exon8: c.G2081A: p.R694H	D	D	0.010539
PRKACA	nonsynonymous	NM_001304349: exon6: c.G727A: p.D243N	D	D	0.036364
SMAD4	nonsynonymous	NM_005359: exon11: c.C1373T: p.A458V	D	D	0.035088
SOX17	nonsynonymous	NM_022454: exon2: c.G1075A: p.D359N	D	D	0.034483
TCF7L1	nonsynonymous	NM_031283: exon8: c.A920C: p.H307P	D	D	0.05291
VANGL2	nonsynonymous	NM_020335: exon2: c.C23T: p.S8L	D	D	0.023729
WNT6	nonsynonymous	NM_006522: exon3: c.G629A: p.G210D	D	D	0.055046
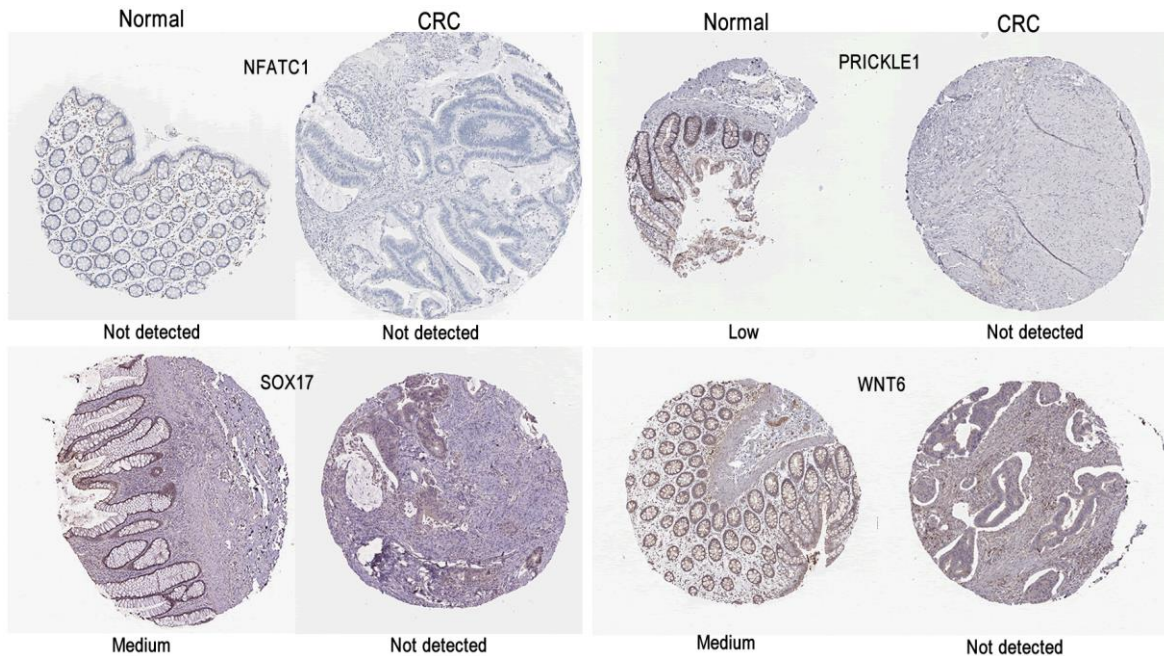
Figure 7. Validation of the mRNA expression levels of *NFATC1, PRICKLE1, SOX17,* and *WNT6* in READ tissues and normal rectal tissues using GEPIA. The red box represents READ samples (92), and the gray box represents normal samples (318). READ, rectal adenocarcinoma. Significance markers: ***, p<0.001.
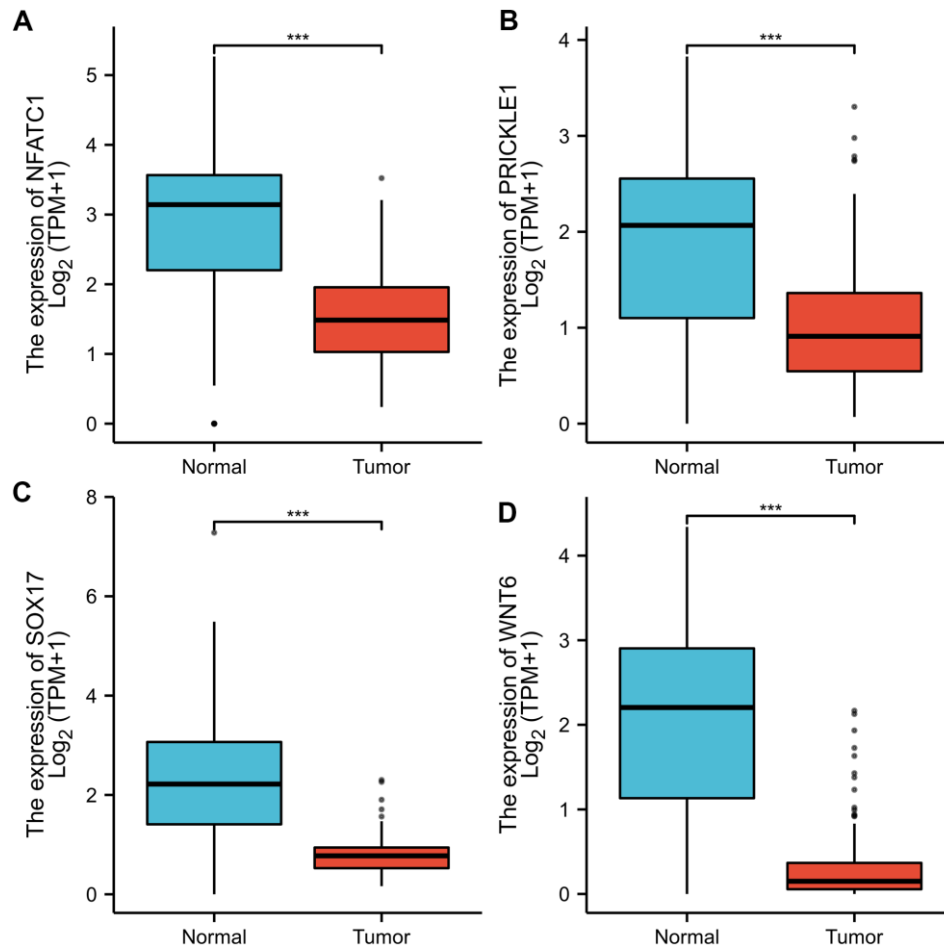
Figure 8. Representative immunohistochemistry images of (**A**) NFATC1; (**B**) PRICKLE1; (**C**) SOX7; (**D**) WNT6 in CRC and colorectal tissues derived from the HPA database. HPA, Human Protein Atlas.
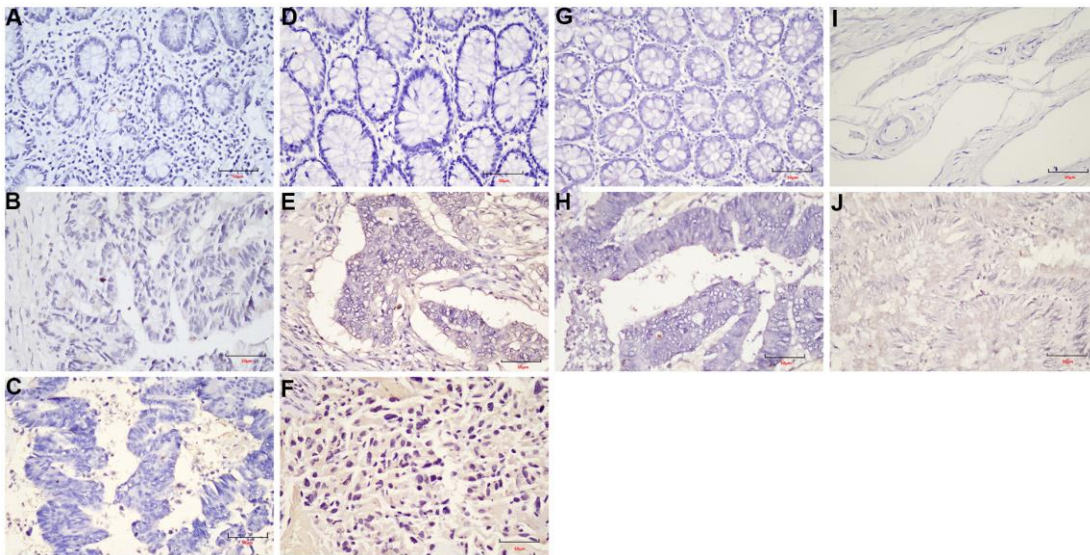

Figure 9. The expression of *LEF1* in FFPE samples. (**A**) DST-LRRC; (**B**) DST-PRC; (**C**) DST-rectal tissue; (**D**) LZQ-LRRC; (**E**) LZQ-PRC; (**F**) LZQ-rectal tissue; (**G**) LHC-NRRC; (**H**) LHC-rectal tissue; (**I**) SCC-NRRC; (**J**) SCC-rectal tissue. Scale bar=50μm.
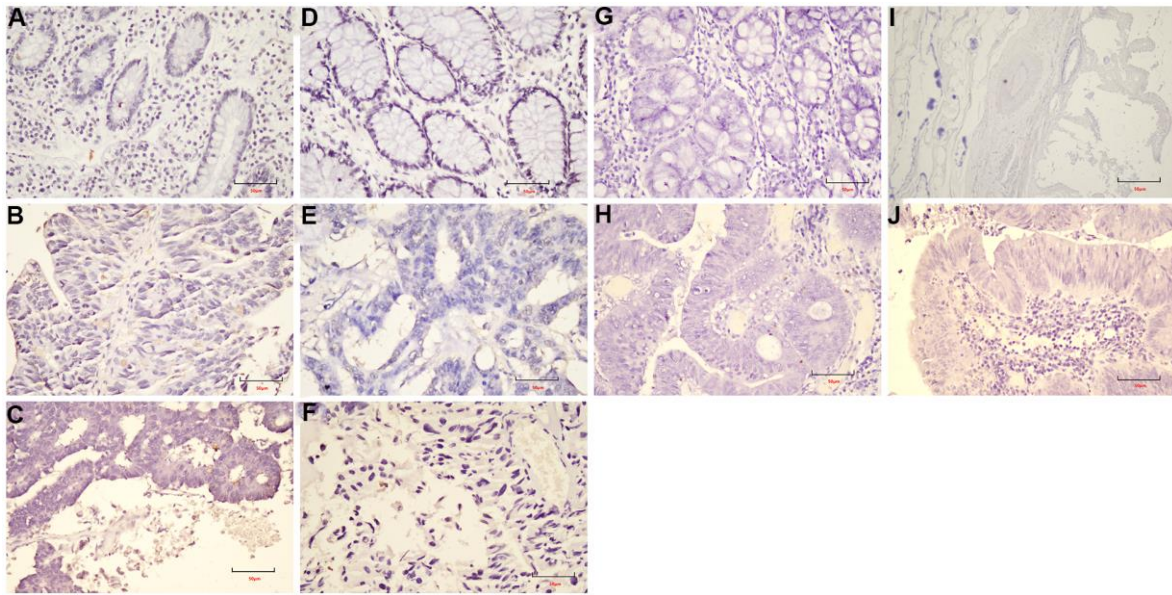
Figure 10. The expression of *cyclin D1* in FFPE samples. (**A**) DST-LRRC; (**B**) DST-PRC; (**C**) DST-rectal tissue; (**D**) LZQ-LRRC; (**E**) LZQ-PRC; (**F**) LZQ-rectal tissue; (**G**) LHC-NRRC; (**H**) LHC-rectal tissue; (**I**) SCC-NRRC; (**J**) SCC-rectal tissue. Scale bar=50μm.
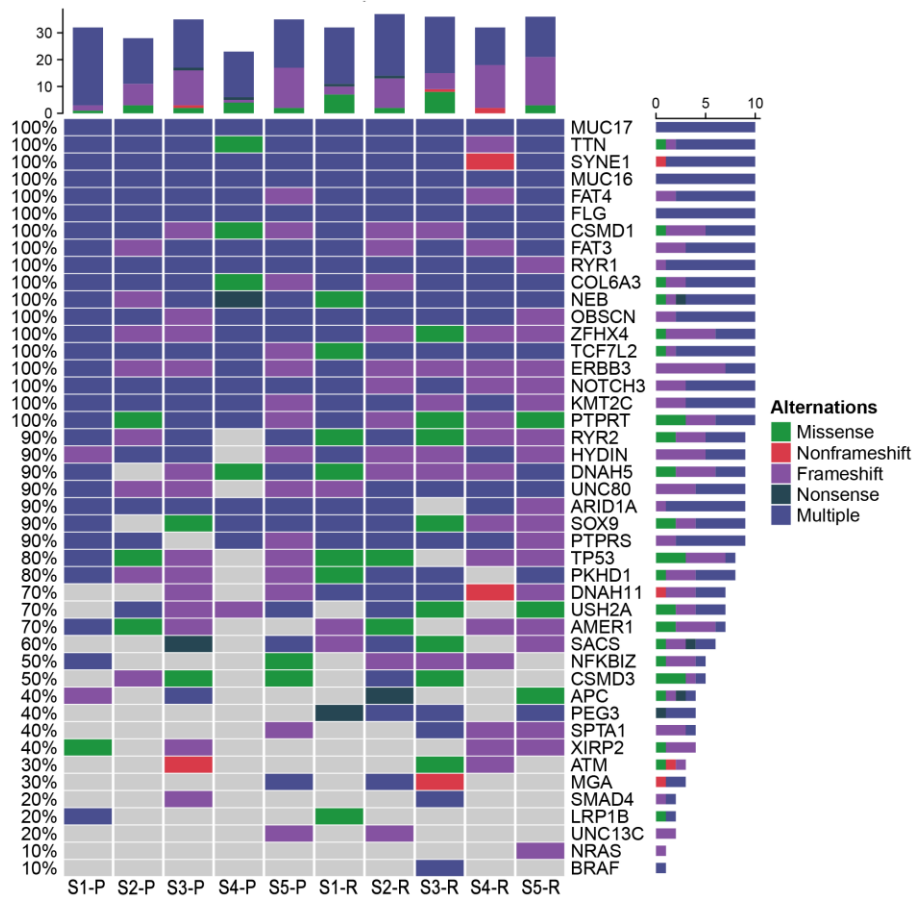

Figure 11. Potential driver genes related to LRRC. Forty-eight genes included (1) genes reported as significantly mutated genes in the previous study in CRC; (2) genes with a higher mutation rate in the cohort (TCGA Freq≥10%; MSK Freq≥5%).

Table 5. The CNVs of PRC and LRRC samples.

Sample	Chromosome	Start	End	Gene	Depth	Copy_number
P1-primary	Chr4	1803536	1803776	FGFR3	677.633	3.12
P2-primary	Chr4	1803536	1803776	FGFR3	493.65	3.04
P3-primary	Chr4	1803536	1803776	FGFR3	699.067	3.27
P3-primary	Chr11	69588723	69588963	FGF4	655.746	3.24
P3-primary	Chr8	38282001	38282241	FGFR1	886.567	3.79
P4-primary	Chr4	1803284	1803510	FGFR3	414.624	4.15
P5-primary	Chr4	1803536	1803776	FGFR3	986.671	4.13
P5-primary	Chr11	69588723	69588963	FGF4	580.725	3.04
P1-recurrent	Chr4	1803284	1803510	FGFR3	273.934	3.71
P2-recurrent	Chr4	1803536	1803776	FGFR3	543.558	3.37
P3-recurrent	Chr4	1803284	1803510	FGFR3	200.765	3.38
P4-recurrent	Chr4	1803536	1803776	FGFR3	772.854	3.92
P4-recurrent	Chr11	69588723	69588963	FGF4	494.054	3.08
P5-recurrent	Chr4	1803536	1803776	FGFR3	472.738	3.36
P5-recurrent	Chr8	128750764	128751028	MYC	2600.53	12.58

LRRC refers to the recurrence of READ after radical resection. In addition to negative distal and circumferential margins, radical resection also requires no lateral or distant lymph node metastasis. Attention should also be paid to tumor-free operations, including irrigation before anastomosis. However, the evolution of LRRC at the genome level remains unknown. In the present study, WES was performed on PRC, LRRC, and NRRC samples to detect the mutational characteristics of LRRC and discover the specific genes and related pathways. The results of the study showed that LRRC tissues exhibit different mutation profiles.

Many studies have explored the relationship between mutations in key driver genes and CRC metastasis [22]. The specific genes of PRC were found, including *MYCN, SCRIB, SNAPC4,* and *MED15*. *MYCN* is a member of the *MYC* family. The amplification of *MYCN* is related to many tumors, most notably neuroblastoma [23]. A frame shift mutation (c.87dupC) was found in the *MYCN* gene in 5 PRC samples, which was a clinically unknown mutation. *SCRIB* is a membrane protein that is involved in the maintenance of the apical basal cell polarity of epithelial tissues. It plays a tumor-suppressive role in the progression of skin and liver cancer [24]. Notably, a pathogenic mutation of *SCRIB* (c.T233G) was found in both PRC and LRRC samples. *SCRIB* affects tumor development by negatively modulating the Wnt/β-catenin signaling pathway [25]. It is suggested that SCRIB may affect the occurrence and development of READ through the Wnt/β-catenin signaling pathway. *SNAPC4* gene encodes the largest subunit of the small nuclear RNA-activating protein (SNAP) complex, and its role in cancer is unknown. A nonsynonymous mutation was found in

the *SNAPC4* gene (c.T3158A) was damaged (SIFT). MED15 is part of the multiprotein mediator complex, which plays a cancer-promoting role in urothelial bladder cancer (BCa) and renal cell carcinoma (RCa).

Two specific genes of LRRC samples were found, including *SCRIB* and *RUSC2*. Non-frameshift mutations of the *RUSC2* (c.831_833del, c.3465_3467del) were found in LRRC samples. *RUSC2* interacts with the SHD domain of GIT2 and reduces GIT2 degradation, which regulates lung cancer progression through EGFR signaling [26]. Whether RUSC2 affects the progression of LRRC through EGFR signaling needs further research.

FGFR1 and *MYC* were the specific genes of CNVs for PRC and LRRC, respectively. Genetic aberrations in FGFRs have been reported in a variety of cancers, including gastric, lung, and breast cancers [27]. FGFR1 amplification was previously shown to be associated with resistance to endocrine therapy, shorter time to distant metastasis, and shorter overall survival in HR[+] breast cancer [28]. MYC amplification plays an important role in the progression of CRC [29].

The pathways involved in PRC and LRRC are different. The specific pathways enriched in the PRC were the pentose phosphate pathway, fructose and mannose metabolism, proteoglycans in cancer, ECM-receptor interaction, and protein digestion and absorption. They were involved in the occurrence, development, and metastasis of CRC [30–33]. The pathogenic genes of NRRC were mainly enriched in peroxisome and butanoate metabolism. In CRC tissues, the peroxisome

proliferator-activated receptor (PPAR) signaling pathway was down-regulated [34]. Several studies have shown that PPARγ activation promotes cell cycle arrest, apoptosis, and differentiation in many human tumors, and selective synthetic ligands have been shown to act as potential antitumor drugs [35]. It is suggested that the peroxisome limits the LRRC through the PPAR signaling pathway.

In this study, the specific signaling pathways of LRRC were Wnt signaling pathway, gap junction, glucagon signaling pathway, axon guidance, thyroid hormone synthesis, morphine addiction, glycine, serine and threonine metabolism. The present study focused on the Wnt signaling pathway. The *PRICKLE1* mutation (c.C875T) and *WNT6* mutation (c.G629A) were predicted as "D (deleterious)" by SIFT, Polyphen2 and FATHMM software, which were critical to the occurrence of LRRC (Table 4). The expression of LEF1 and cyclin D1 proteins of Wnt signaling pathway was higher in PRC compared to NRRC. Carbonic anhydrase IV (CA4) inhibits the Wnt signaling pathway by targeting the WTAP-WT1-TBL1 axis and is a novel tumor suppressor in CRC [36]. Agrin (AGRN) may act as an oncogenic indicator of READ through activation of the WNT pathway, which could help in the development of optimal treatments for READ [37]. Activation of Wnt signaling as a mechanism of chemoresistance in recurrent small cell lung cancer (SCLC) [38]. PRC1, a novel Wnt target, functions in a positive feedback loop that reinforces Wnt signaling to promote early Hepatocellular carcinoma (HCC) recurrence [39]. AKIP1 is a novel regulator of Wnt/β-catenin signaling and early relapse of HCC [40]. These findings suggest that the development of small molecule drugs that target the Wnt pathway may be important for LRRC.

Neoplastic transformation is frequently associated with a loss of gap junction intercellular communication and a reduction in the expression of connexins in various tumor types [41]. Furthermore, gap junctions may have distinct functional roles in cell growth and cell invasion, as a gap junction inhibitor decreases the invasion of prostate cancer cells [42]. Studies have found that glucagon signaling pathway may be involved in the progression of CRC [43]. Synthesis and growth of tumor proteins can be stimulated by glucagon *in situ* [44]. Several lines of evidence indicate that axon guidance genes are involved not only in neural development but also in cancer development. ROBO1 and ROBO2, crucial regulators of axon guidance, are considered potential tumor suppressor genes [45]. Multiple studies have shown a significant connection between hypothyroidism and pancreatic, gastric, and breast cancer [46]. Deregulation of Thyroid Hormones

(THs) system in Colorectal Cancer (CRC) suggests that these hormones may play roles in CRC pathogenesis [47]. Morphine promotes tumorigenesis and cetuximab resistance via EGFR signaling activation in human colorectal cancer [48]. Glycolysis and glycine, serine and threonine were activated in CRC, and these alterations may promote cell proliferation [49].

Besides, this study found two KEGG pathways shared by PRC and LRRC, including signaling pathways regulating pluripotency of stem cells and estrogen signaling pathway. Cancer stem cells (CSCs) are involved in the occurrence and recurrence of CRC and have been identified [50]. Estrogen signaling pathway may be involved in the occurrence of CRC, but the specific mechanism is not yet clear [51]. These two pathways may be involved in the progression from PRC to LRRC.

There are certain limitations in the present study. The number of patients used for this study was too small. Although this study can't yield a definitive conclusion, Wnt signaling pathway may play a critical role in LRRC. Although we used some tractable methods to validate the possible biological significance of Wnt signaling pathway mutations in LRRC, the specific functions and molecular mechanisms of the Wnt signaling pathway in LRRC need to be further explored.

CONCLUSIONS

We described the mutation characteristics of LRRC, PRC, and NRRC in the present study. Our results showed that LRRC, PRC, and NRRC had different genomic characteristics and involved different pathways. The signaling pathway of Wnt signaling pathway, Gap junction, Glucagon signaling pathway, Axon guidance, Thyroid hormone synthesis, Morphine addiction, and Serine and threonine metabolism may be related to the occurrence of LRRC. The genes including *NFATC1, PRICKLE1, SOX17,* and *WNT6,* related to Wnt signaling pathway, may play a critical role in LRRC. Further translational and clinical research is imperative to investigate new therapeutic strategies for LRRC.

AUTHOR CONTRIBUTIONS

Writing – review and editing, Yi Yang, Xiaodong Gu, Zhenyang Li, Chuang Zheng, Zihao Wang, Minwei Zhou, Zongyou Chen, Mengzhen Li, Dongbing Li and Jianbin Xiang.

CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest.

FUNDING

REFERENCES

1. Beyond TME Collaborative. Consensus statement on the multidisciplinary management of patients with recurrent and primary rectal cancer beyond total mesorectal excision planes. Br J Surg. 2013; 100:1009–14.
https://doi.org/10.1002/bjs.9192 PMID:23754654

2. Harji DP, Sagar PM. Advancing the surgical treatment of locally recurrent rectal cancer. Br J Surg. 2012; 99:1169–71.
https://doi.org/10.1002/bjs.8879
PMID:22829061

3. Mirnezami AH, Sagar PM, Kavanagh D, Witherspoon P, Lee P, Winter D. Clinical algorithms for the surgical management of locally recurrent rectal cancer. Dis Colon Rectum. 2010; 53:1248–57.
https://doi.org/10.1007/DCR.0b013e3181e10b0e
PMID:20706067

4. Bhangu A, Ali SM, Cunningham D, Brown G, Tekkis P. Comparison of long-term survival outcome of operative vs nonoperative management of recurrent rectal cancer. Colorectal Dis. 2013; 15:156–63.
https://doi.org/10.1111/j.1463-1318.2012.03123.x
PMID:23190113

5. Bouchard P, Efron J. Management of recurrent rectal cancer. Ann Surg Oncol. 2010; 17:1343–56.
https://doi.org/10.1245/s10434-009-0861-2
PMID:20041351

6. van Gijn W, Marijnen CA, Nagtegaal ID, Kranenbarg EM, Putter H, Wiggers T, Rutten HJ, Påhlman L, Glimelius B, van de Velde CJ, and Dutch Colorectal Cancer Group. Preoperative radiotherapy combined with total mesorectal excision for resectable rectal cancer: 12-year follow-up of the multicentre, randomised controlled TME trial. Lancet Oncol. 2011; 12:575–82.
https://doi.org/10.1016/S1470-2045(11)70097-3
PMID:21596621

7. Bakx R, Visser O, Josso J, Meijer S, Slors JF, van Lanschot JJ. Management of recurrent rectal cancer: a population based study in greater Amsterdam. World J Gastroenterol. 2008; 14:6018–23.
https://doi.org/10.3748/wjg.14.6018 PMID:18932280

8. Palmer G, Martling A, Cedermark B, Holm T. A population-based study on the management and outcome in patients with locally recurrent rectal cancer. Ann Surg Oncol. 2007; 14:447–54.
https://doi.org/10.1245/s10434-006-9256-9
PMID:17139457

9. Pawlik TM, Skibber JM, Rodriguez-Bigas MA. Pelvic exenteration for advanced pelvic malignancies. Ann Surg Oncol. 2006; 13:612–23.
https://doi.org/10.1245/ASO.2006.03.082
PMID:16538402

10. Heriot AG, Tekkis PP, Darzi A, Mackay J. Surgery for local recurrence of rectal cancer. Colorectal Dis. 2006; 8:733–47.
https://doi.org/10.1111/j.1463-1318.2006.01018.x
PMID:17032318

11. Cai Y, Li Z, Gu X, Fang Y, Xiang J, Chen Z. Prognostic factors associated with locally recurrent rectal cancer following primary surgery (Review). Oncol Lett. 2014; 7:10–16.
https://doi.org/10.3892/ol.2013.1640 PMID:24348812

12. Ashraf SQ, Burns EM, Jani A, Altman S, Young JD, Cunningham C, Faiz O, Mortensen NJ. The economic impact of anastomotic leakage after anterior resections in English NHS hospitals: are we adequately remunerating them? Colorectal Dis. 2013; 15:e190–98.
https://doi.org/10.1111/codi.12125 PMID:23331871

13. Fujita S, Mizusawa J, Kanemitsu Y, Ito M, Kinugasa Y, Komori K, Ohue M, Ota M, Akazai Y, Shiozawa M, Yamaguchi T, Bandou H, Katsumata K, et al, and Colorectal Cancer Study Group of Japan Clinical Oncology Group. Mesorectal Excision With or Without Lateral Lymph Node Dissection for Clinical Stage II/III Lower Rectal Cancer (JCOG0212): A Multicenter, Randomized Controlled, Noninferiority Trial. Ann Surg. 2017; 266:201–07.
https://doi.org/10.1097/SLA.0000000000002212
PMID:28288057

14. Sebag-Montefiore D, Stephens RJ, Steele R, Monson J, Grieve R, Khanna S, Quirke P, Couture J, de Metz C, Myint AS, Bessell E, Griffiths G, Thompson LC, Parmar M. Preoperative radiotherapy versus selective postoperative chemoradiotherapy in

patients with rectal cancer (MRC CR07 and NCIC-CTG C016): a multicentre, randomised trial. Lancet. 2009; 373:811–20.
https://doi.org/10.1016/S0140-6736(09)60484-0
PMID:19269519

15. Harris CA, Solomon MJ, Heriot AG, Sagar PM, Tekkis PP, Dixon L, Pascoe R, Dobbs BR, Frampton CM, Harji DP, Kontovounisios C, Austin KK, Koh CE, et al. The Outcomes and Patterns of Treatment Failure After Surgery for Locally Recurrent Rectal Cancer. Ann Surg. 2016; 264:323–29.
https://doi.org/10.1097/SLA.0000000000001524
PMID:26692078

16. Westberg K, Palmer G, Hjern F, Johansson H, Holm T, Martling A. Management and prognosis of locally recurrent rectal cancer - A national population-based study. Eur J Surg Oncol. 2018; 44:100–07.
https://doi.org/10.1016/j.ejso.2017.11.013
PMID:29224985

17. Tanis PJ, Doeksen A, van Lanschot JJ. Intentionally curative treatment of locally recurrent rectal cancer: a systematic review. Can J Surg. 2013; 56:135–44.
https://doi.org/10.1503/cjs.025911 PMID:23517634

18. Lin C, Zhang Y, Zhang K, Zheng Y, Lu L, Chang H, Yang H, Yang Y, Wan Y, Wang S, Yuan M, Yan Z, Zhang R, et al. Fever Promotes T Lymphocyte Trafficking via a Thermal Sensory Pathway Involving Heat Shock Protein 90 and α4 Integrins. Immunity. 2019; 50:137–51.e6.
https://doi.org/10.1016/j.immuni.2018.11.013
PMID:30650373

19. Wei J, Wang Y, Shi K, Wang Y. Identification of Core Prognosis-Related Candidate Genes in Cervical Cancer via Integrated Bioinformatical Analysis. Biomed Res Int. 2020; 2020:8959210.
https://doi.org/10.1155/2020/8959210
PMID:32258155

20. Veltman JA, Brunner HG. De novo mutations in human genetic disease. Nat Rev Genet. 2012; 13:565–75.
https://doi.org/10.1038/nrg3241
PMID:22805709

21. Liu Z, Yang C, Li X, Luo W, Roy B, Xiong T, Zhang X, Yang H, Wang J, Ye Z, Chen Y, Song J, Ma S, et al. The landscape of somatic mutation in sporadic Chinese colorectal cancer. Oncotarget. 2018; 9:27412–22.
https://doi.org/10.18632/oncotarget.25287
PMID:29937994

22. Huang D, Sun W, Zhou Y, Li P, Chen F, Chen H, Xia D, Xu E, Lai M, Wu Y, Zhang H. Mutations of key driver genes in colorectal cancer progression and metastasis. Cancer Metastasis Rev. 2018; 37:173–87.
https://doi.org/10.1007/s10555-017-9726-5
PMID:29322354

23. Zhao Z, Shelton SD, Oviedo A, Baker AL, Bryant CP, Omidvarnia S, Du L. The PLAGL2/MYCN/miR-506-3p interplay regulates neuroblastoma cell fate and associates with neuroblastoma progression. J Exp Clin Cancer Res. 2020; 39:41.
https://doi.org/10.1186/s13046-020-1531-2
PMID:32087738

24. Kapil S, Sharma BK, Patil M, Elattar S, Yuan J, Hou SX, Kolhe R, Satyanarayana A. The cell polarity protein Scrib functions as a tumor suppressor in liver cancer. Oncotarget. 2017; 8:26515–31.
https://doi.org/10.18632/oncotarget.15713
PMID:28460446

25. Daulat AM, Wagner MS, Walton A, Baudelet E, Audebert S, Camoin L, Borg JP. The Tumor Suppressor SCRIB is a Negative Modulator of the Wnt/β-Catenin Signaling Pathway. Proteomics. 2019; 19:e1800487.
https://doi.org/10.1002/pmic.201800487
PMID:31513346

26. Duan B, Cui J, Sun S, Zheng J, Zhang Y, Ye B, Chen Y, Deng W, Du J, Zhu Y, Chen Y, Gu L. EGF-stimulated activation of Rab35 regulates RUSC2-GIT2 complex formation to stabilize GIT2 during directional lung cancer cell migration. Cancer Lett. 2016; 379:70–83.
https://doi.org/10.1016/j.canlet.2016.05.027
PMID:27238570

27. Katoh M. Fibroblast growth factor receptors as treatment targets in clinical oncology. Nat Rev Clin Oncol. 2019; 16:105–22.
https://doi.org/10.1038/s41571-018-0115-y
PMID:30367139

28. Turner N, Pearson A, Sharpe R, Lambros M, Geyer F, Lopez-Garcia MA, Natrajan R, Marchio C, Iorns E, Mackay A, Gillett C, Grigoriadis A, Tutt A, et al. FGFR1 amplification drives endocrine therapy resistance and is a therapeutic target in breast cancer. Cancer Res. 2010; 70:2085–94.
https://doi.org/10.1158/0008-5472.CAN-09-3746
PMID:20179196

29. Ben-David E, Bester AC, Shifman S, Kerem B. Transcriptional dynamics in colorectal carcinogenesis: new insights into the role of c-Myc and miR17 in benign to cancer transformation. Cancer Res. 2014; 74:5532–40.
https://doi.org/10.1158/0008-5472.CAN-14-0932
PMID:25125661

30. Shibuya N, Inoue K, Tanaka G, Akimoto K, Kubota K. Augmented pentose phosphate pathway plays critical roles in colorectal carcinomas. Oncology. 2015; 88:309–19.
https://doi.org/10.1159/000369905
PMID:25591719

31. Brown DG, Rao S, Weir TL, O'Malia J, Bazan M, Brown RJ, Ryan EP. Metabolomics and metabolic pathway networks from human colorectal cancers, adjacent mucosa, and stool. Cancer Metab. 2016; 4:11.
https://doi.org/10.1186/s40170-016-0151-y
PMID:27275383

32. Zhou F, Tang D, Xu Y, He H, Wu Y, Lin L, Dong J, Tan W, Dai Y. Identification of microRNAs and their Endonucleolytic Cleavaged target mRNAs in colorectal cancer. BMC Cancer. 2020; 20:242.
https://doi.org/10.1186/s12885-020-06717-4
PMID:32293320

33. Wang J, Yu S, Chen G, Kang M, Jin X, Huang Y, Lin L, Wu D, Wang L, Chen J. A novel prognostic signature of immune-related genes for patients with colorectal cancer. J Cell Mol Med. 2020; 24:8491–504.
https://doi.org/10.1111/jcmm.15443 PMID:32564470

34. Zhang X, Yao J, Shi H, Gao B, Zhang L. LncRNA TINCR/microRNA-107/CD36 regulates cell proliferation and apoptosis in colorectal cancer via PPAR signaling pathway based on bioinformatics analysis. Biol Chem. 2019; 400:663–75.
https://doi.org/10.1515/hsz-2018-0236
PMID:30521471

35. Blanquicett C, Roman J, Hart CM. Thiazolidinediones as anti-cancer agents. Cancer Ther. 2008; 6:25–34.
PMID:19079765

36. Zhang J, Tsoi H, Li X, Wang H, Gao J, Wang K, Go MY, Ng SC, Chan FK, Sung JJ, Yu J. Carbonic anhydrase IV inhibits colon cancer development by inhibiting the Wnt signalling pathway through targeting the WTAP-WT1-TBL1 axis. Gut. 2016; 65:1482–93.
https://doi.org/10.1136/gutjnl-2014-308614
PMID:26071132

37. Wang ZQ, Sun XL, Wang YL, Miao YL. Agrin promotes the proliferation, invasion and migration of rectal cancer cells via the WNT signaling pathway to contribute to rectal cancer progression. J Recept Signal Transduct Res. 2021; 41:363–70.
https://doi.org/10.1080/10799893.2020.1811325
PMID:32862766

38. Wagner AH, Devarakonda S, Skidmore ZL, Krysiak K, Ramu A, Trani L, Kunisaki J, Masood A, Waqar SN, Spies NC, Morgensztern D, Waligorski J, Ponce J, et al. Recurrent WNT pathway alterations are frequent in relapsed small cell lung cancer. Nat Commun. 2018; 9:3787.
https://doi.org/10.1038/s41467-018-06162-9
PMID:30224629

39. Chen J, Rajasekaran M, Xia H, Zhang X, Kong SN, Sekar K, Seshachalam VP, Deivasigamani A, Goh BK, Ooi LL, Hong W, Hui KM. The microtubule-associated protein PRC1 promotes early recurrence of hepatocellular carcinoma in association with the Wnt/β-catenin signalling pathway. Gut. 2016; 65:1522–34.
https://doi.org/10.1136/gutjnl-2015-310625
PMID:26941395

40. Cui Y, Wu X, Lin C, Zhang X, Ye L, Ren L, Chen M, Yang M, Li Y, Li M, Li J, Guan J, Song L. AKIP1 promotes early recurrence of hepatocellular carcinoma through activating the Wnt/β-catenin/CBP signaling pathway. Oncogene. 2019; 38:5516–29.
https://doi.org/10.1038/s41388-019-0807-5
PMID:30936461

41. Mesnil M. Connexins and cancer. Biol Cell. 2002; 94:493–500.
https://doi.org/10.1016/s0248-4900(02)00025-4
PMID:12566222

42. Tate AW, Lung T, Radhakrishnan A, Lim SD, Lin X, Edlund M. Changes in gap junctional connexin isoforms during prostate cancer progression. Prostate. 2006; 66:19–31.
https://doi.org/10.1002/pros.20317 PMID:16114058

43. Wu Z, Liu Z, Ge W, Shou J, You L, Pan H, Han W. Analysis of potential genes and pathways associated with the colorectal normal mucosa-adenoma-carcinoma sequence. Cancer Med. 2018; 7:2555–66.
https://doi.org/10.1002/cam4.1484 PMID:29659199

44. Fiocca R, Rindi G, Capella C, Grimelius L, Polak JM, Schwartz TW, Yanaihara N, Solcia E. Glucagon, glicentin, proglucagon, PYY, PP and proPP-icosapeptide immunoreactivities of rectal carcinoid tumors and related non-tumor cells. Regul Pept. 1987; 17:9–29.
https://doi.org/10.1016/0167-0115(87)90029-2
PMID:2882565

45. Je EM, Gwak M, Oh H, Choi MR, Choi YJ, Lee SH, Yoo NJ. Frameshift mutations of axon guidance genes ROBO1 and ROBO2 in gastric and colorectal cancers with microsatellite instability. Pathology. 2013; 45:645–50.
https://doi.org/10.1097/PAT.0000000000000007
PMID:24247621

46. Brown AR, Simmen RC, Simmen FA. The role of thyroid hormone signaling in the prevention of digestive system cancers. Int J Mol Sci. 2013; 14:16240–57.
https://doi.org/10.3390/ijms140816240
PMID:23924944

47. Zohreh B, Masoumeh V, Fakhraddin N, Omrani GH. Apigenin-mediated Alterations in Viability and Senescence of SW480 Colorectal Cancer Cells Persist in The Presence of L-thyroxine. Anticancer Agents Med Chem. 2019; 19:1535–42.
https://doi.org/10.2174/1871520619666190704102708 PMID:31272364

48. Lu H, Zhang H, Weng ML, Zhang J, Jiang N, Cata JP, Ma D, Chen WK, Miao CH. Morphine promotes tumorigenesis and cetuximab resistance via EGFR signaling activation in human colorectal cancer. J Cell Physiol. 2021; 236:4445–54.
https://doi.org/10.1002/jcp.30161 PMID:33184860

49. Gu J, Xiao Y, Shu D, Liang X, Hu X, Xie Y, Lin D, Li H. Metabolomics Analysis in Serum from Patients with Colorectal Polyp and Colorectal Cancer by ^{1}H-NMR Spectrometry. Dis Markers. 2019; 2019:3491852.
https://doi.org/10.1155/2019/3491852
PMID:31089393

50. Szaryńska M, Olejniczak A, Kobiela J, Spychalski P, Kmieć Z. Therapeutic strategies against cancer stem cells in human colorectal cancer. Oncol Lett. 2017; 14:7653–68.
https://doi.org/10.3892/ol.2017.7261
PMID:29250169

51. Lv J, Li L. Hub Genes and Key Pathway Identification in Colorectal Cancer Based on Bioinformatic Analysis. Biomed Res Int. 2019; 2019:1545680.
https://doi.org/10.1155/2019/1545680
PMID:31781593

SUPPLEMENTARY MATERIALS

Supplementary Tables

Supplementary Table 1. Characteristics of the patients in this study.

Number	Age	Gender	Tumor size, cm	Tumor stage	Pathological description
LRRC (first operation)					
1-LJ	45	male	3	T2N0M0 I	adenocarcinoma
2-LZC	62	male	2	T2N0M0 I	adenocarcinoma
3-SYJ	69	male	2	T4aN0M0 IIB	adenocarcinoma
4-WLJ	57	male	4	T4aN2M0 IIIC	adenocarcinoma
5-WGR	71	male	2	T1N0M0 I	adenocarcinoma
NRRC					
1-SL	49	male	3	T4aN0M0 IIB	adenocarcinoma
2-SCC	61	male	2	T1N0M0 I	adenocarcinoma
3-LHC	70	male	5	T4aN0M0 IIB	adenocarcinoma

Supplementary Table 2. Quality of whole-exome sequencing.

Patient ID	Raw data (Mb)	Q30	[Total] fraction of mapped reads (%)	[Target] average depth
P1-N	17918.2	92.18%	95.58	243.04
P1-P	38812.2	93.14%	92.92	527.11
P1-R	12741.7	93.38%	97.00	187.55
P2-N	17989.6	92.32%	99.23	251.08
P2-P	20155.4	92.78%	98.32	250.19
P2-R	22565.7	92.70%	99.30	319.95
P3-N	18089.1	92.31%	99.10	241.69
P3-P	23193.9	92.85%	98.99	334.37
P3-R	26373.1	92.99%	96.58	370.91
P4-N	17481.1	92.58%	99.20	243.31
P4-P	15413.3	94.65%	96.44	224.44
P4-R	16435.2	95.02%	99.38	232.93
P5-N	16383.6	92.07%	98.67	214.39
P5-P	23762.6	92.84%	98.65	337.59
P5-R	18689.8	93.13%	99.34	247.94
P6-N	26541.1	93.68%	93.56	332.02
P6-NR	29528.5	94.17%	95.81	398.69
P7-N	15592.5	96.32%	93.19	217.23
P7-NR	15828.1	96.18%	93.51	224.44
P8-N	15890.8	95.41%	91.16	181.46
P8-NR	16316.7	96.16%	93.82	196.02

N, Normal rectal cancer; P, Primary rectal cancer; R, Local-recurrent rectal cancer; NR, Non-recurrent rectal cancer.

Supplementary Table 3. Top 15 significant GO terms of BP, MF and CC in PRC samples.

Category	Term	Count	P-Value
BP	positive regulation of transcription from RNA polymerase II promoter	73	9.28E-06
BP	negative regulation of apoptotic process	40	4.49E-05
BP	positive regulation of transcription, DNA-templated	43	7.15E-05
BP	in utero embryonic development	21	1.97E-04
BP	apoptotic cell clearance	6	3.24E-04
CC	integral component of plasma membrane	93	1.35E-05
CC	membrane raft	24	1.82E-05
CC	plasma membrane	217	9.55E-05
CC	endoplasmic reticulum lumen	20	4.54E-04
CC	transcription factor complex	20	4.85E-04
MF	sequence-specific DNA binding	51	1.88E-07
MF	structural molecule activity	28	1.43E-05
MF	ATP binding	98	7.53E-05
MF	steroid hormone receptor activity	10	7.25E-04
MF	protein homodimerization activity	52	8.40E-04

Supplementary Table 4. The KEGG pathways in PRC samples.

Term	Count	%	P-Value
Chronic myeloid leukemia	11	0.009228652	0.003544157
Melanogenesis	13	0.010906589	0.005012796
Signaling pathways regulating pluripotency of stem cells	16	0.013423494	0.005374073
Glioma	10	0.008389684	0.005678579
Platelet activation	15	0.012584526	0.006738809
Acute myeloid leukemia	9	0.007550716	0.00739768
Hepatitis C	15	0.012584526	0.008209968
Oxytocin signaling pathway	16	0.013423494	0.010028696
Estrogen signaling pathway	12	0.010067621	0.012404787
Leukocyte transendothelial migration	13	0.010906589	0.014795069
Pentose phosphate pathway	6	0.00503381	0.015005993
Fructose and mannose metabolism	6	0.00503381	0.022446568
Dilated cardiomyopathy	10	0.008389684	0.027945534
Proteoglycans in cancer	18	0.015101431	0.028185022
Prolactin signaling pathway	9	0.007550716	0.028536839
ECM-receptor interaction	10	0.008389684	0.034080997
ErbB signaling pathway	10	0.008389684	0.034080997
Prostate cancer	10	0.008389684	0.036317726
Protein digestion and absorption	10	0.008389684	0.036317726
HTLV-I infection	21	0.017618336	0.037392779
Rap1 signaling pathway	18	0.015101431	0.042070079
Hypertrophic cardiomyopathy (HCM)	9	0.007550716	0.046279642

Supplementary Table 5. Top 15 significant GO terms of BP, MF and CC in LRRC samples.

Category	Term	Count	P-Value
BP	positive regulation of transcription from RNA polymerase II promoter	65	1.42E-05
BP	retina development in camera-type eye	11	5.71E-05
BP	nervous system development	26	9.90E-05
BP	angiogenesis	22	1.19E-04
BP	negative regulation of transcription from RNA polymerase II promoter	48	2.01E-04
CC	plasma membrane	191	2.91E-05
CC	integral component of plasma membrane	79	6.36E-05
CC	cell-cell junction	17	4.27E-04
CC	mitochondrion	71	5.91E-04
CC	stress fiber	9	5.94E-04
MF	ATP binding	100	3.86E-08
MF	protein tyrosine kinase activity	16	1.75E-04
MF	sequence-specific DNA binding	37	4.22E-04
MF	serine-type endopeptidase activity	21	0.001976808
MF	transmembrane-ephrin receptor activity	4	0.003941594

Supplementary Table 6. Significant pathways in LRRC samples.

Term	Count	%	P-Value
Glucagon signaling pathway	15	0.0149	0.0001
Gap junction	14	0.0139	0.0001
Signaling pathways regulating pluripotency of stem cells	15	0.0149	0.0036
Axon guidance	13	0.0129	0.0106
Metabolic pathways	70	0.0695	0.0163
Neuroactive ligand-receptor interaction	21	0.0209	0.0218
Circadian entrainment	10	0.0099	0.0249
Calcium signaling pathway	15	0.0149	0.0280
Glycine, serine and threonine metabolism	6	0.0060	0.0281
Estrogen signaling pathway	10	0.0099	0.0315
Thyroid hormone synthesis	8	0.0079	0.0351
Wnt signaling pathway	12	0.0119	0.0430
Pathways in cancer	26	0.0258	0.0439
Morphine addiction	9	0.0089	0.0488

Supplementary Table 7. Top 15 significant GO terms of BP, MF, and CC in NRRC samples.

Category	Term	Count	P-Value
BP	GO:0006810~transport	12	0.001550891
BP	GO:0001937~negative regulation of endothelial cell proliferation	4	0.003777034
BP	GO:0042127~regulation of cell proliferation	8	0.004251828
BP	GO:0010628~positive regulation of gene expression	9	0.008125803
BP	GO:2001241~positive regulation of extrinsic apoptotic signaling pathway in absence of ligand	3	0.009862223
CC	GO:0005777~peroxisome	6	0.004372107
CC	GO:0005759~mitochondrial matrix	10	0.006789772
CC	GO:0005887~integral component of plasma membrane	25	0.010268471
CC	GO:0005576~extracellular region	27	0.013785056
CC	GO:0016020~membrane	34	0.015904909
MF	GO:0005179~hormone activity	6	0.003657841
MF	GO:0050662~coenzyme binding	3	0.007258267
MF	GO:0001105~RNA polymerase II transcription coactivator activity	4	0.007540464
MF	GO:0005524~ATP binding	27	0.011626528
MF	GO:0016887~ATPase activity	7	0.015231878

Supplementary Table 8. Significant pathways in NRRC samples.

Term	Count	%	P-Value
Peroxisome	5	0.017031133	0.021363124
Metabolic pathways	24	0.081749438	0.030427677
Butanoate metabolism	3	0.01021868	0.046182301