Int J Clin Exp Pathol 2026;19(6):278-291
www.ijcep.com /ISSN:1936-2625/IJCEP0172583

Original Article

Pathological features and prognosis based on microsatellite stability status in colorectal cancer patients: a retrospective analysis

Xinxin Liu[1], Yuming Yang[1*], Honggang Zheng[1,2*], Yanping Zhang[1]

[1]Department of Pathology, Tianjin Nankai Hospital, Tianjin, The People's Republic of China; [2]Tianjin King Med Diagnostics Laboratory Co., Ltd., Tianjin, The People's Republic of China. [*]Equal contributors.

Received March 11, 2026; Accepted May 20, 2026; Epub June 15, 2026; Published June 30, 2026

Abstract: Background: Colorectal cancer (CRC) is one of the most common and life-threatening intestinal tumors throughout the world. Its incidence and mortality have been going up in recent decades, and more accurate and personalized treatment for CRC are also being developed continuously. Early screening, molecular typing, and risk grouping have become critical for improving patient survival by using sensitive and special biological markers. Methods: We analyzed the expression of four mismatch repair (MMR) genes, namely MutL homolog 1 (MLH1), MutS homolog 2 (MSH2), MutS homolog 6 (MSH6), and PMS1 homolog 2 (PMS2), in CRC and normal tissues using the Tumor, Normal and Metastatic (TNM plot) database. We used the Kaplan-Meier analysis for prognosis, and related pathways were annotated by the Kyoto Encyclopedia of Genes and Genomes. We conducted a retrospective study with 115 CRC patients, who were divided into microsatellite stable (MSS) and microsatellite instability (MSI) groups by immunohistochemistry (IHC). The diagnostic value of MMR markers was assessed through receiver operating characteristic (ROC) curve analysis. Results: Online data showed that MMR genes were more highly expressed in CRC than in the normal group ($P<0.05$). Low expression of these genes was linked to shorter recurrence free survival (RFS, $P<0.05$), and the related signaling pathways also exert an impact on cell cycle regulation. In clinical samples, loss of MLH1 expression was found in female patients, those aged ≥60 years and lymph node-negative cases ($P<0.05$). MSI-positive tumors were mostly located in the right-sided colon ($P<0.05$). The combined model of MSH6 and PMS2 had the highest area under the curve (AUC) of 0.981 (95% confidence interval (CI): 0.936-0.997, $P<0.001$). Conclusion: Combined detection of MSH6 and PMS2 serves as a reliable biomarker for determining the microsatellite status in CRC. This study provides a basis for the precise diagnosis and targeted therapy of CRC.

Keywords: Deoxyribonucleic acid mismatch repair system, colorectal cancer pathological features, retrospective analysis

Introduction

Colorectal cancer (CRC) is the third most diagnosed cancer in the world [1]. Recent studies from the United States show that CRC is increasingly seen in younger people, and more patients are diagnosed at advanced stages [2]. CRC still represents the second leading cause of cancer-related deaths worldwide and in the United States [3]. Thus, it remains a major global health problem. Consequently, molecular classification has become increasingly important for the targeted therapy of CRC. Patient grouping based on deficient mismatch repair (dMMR) or high microsatellite instability (MSI-H) status has important clinical value [4].

The deoxyribonucleic acid (DNA) mismatch repair (MMR) system consists of two key protein complexes, including MutS (MSH2, MSH4, MSH5, and MSH6), and the MutL (MLH1, MLH3, PMS1, and PMS2); these complexes can find base pair mismatches and bring downstream repair proteins to the position of DNA damage [5, 6]. The dimerization of MLH1 and PMS2 can change the structure of the MutL complex, and this change is needed to activate its endonuclease activity [7, 8]. Complete MMR function keeps the accuracy of DNA replication, and damage to this system will cause genomic instability and more gene mutations [9, 10].

Tumors with dMMR or MSI-H usually have high tumor mutation burden, many tumor-infiltrating

https://doi.org/10.62347/XJVC9794

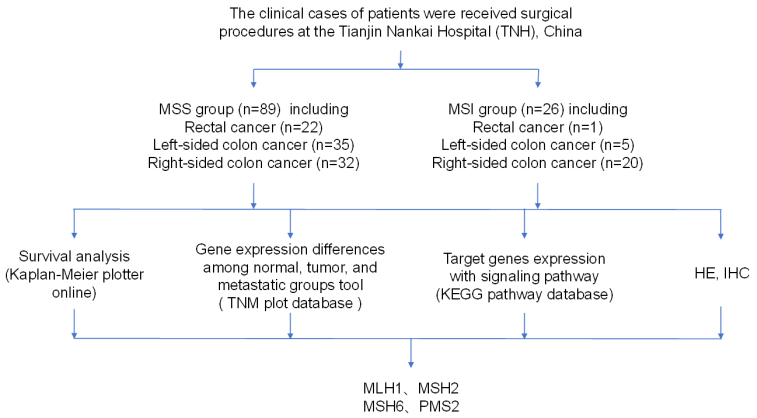

Figure 1. Flow chart illustrating the experimental design of this study. Notes: CRC, Colorectal cancer; MSS, microsatellite stable; MSI, microsatellite instability; MLH1, MutL homolog 1; MSH2, MutS homolog 2; MSH6, MutS homolog 6; PMS2, PMS1 homolog 2; H&E, Hematoxylin-eosin; IHC, immunohistochemistry; TNM plot, Tumor, Normal, and Metastatic tissues plot; KEGG, Kyoto Encyclopedia of Genes and Genomes.

lymphocytes, and related gene changes, so they can respond well to immune checkpoint inhibitor treatment. On the contrary, most patients with pMMR or MSS have low tumor mutation burden and little T-cell infiltration, and they are not sensitive to or may even be resistant to immune checkpoint inhibitor therapy [11, 12].

Testing MMR and MSI status is of great importance to diagnosis, treatment choice and prognosis evaluation of CRC and other tumors [13], biomarkers related to MMR have been used in many kinds of cancers, such as gastric cancer, esophageal cancer [14], endometrial carcinoma [15], salivary gland tumors [16], sarcomas in Lynch syndrome (LS) patients [17], CRC [18], and castration-resistant prostate cancer [19]. Based on these previous studies, our study aimed to clarify the clinical value of microsatellite stability and optimize MMR-based testing methods for more accurate CRC diagnosis.

Materials and methods

Clinical samples

We carried out this retrospective single-center study at Tianjin Nankai Hospital, China, wirh 115 patients who received surgical treatment for CRC from July 2023 to May 2025. There were 23 cases of rectal cancer, 40 cases of left-sided colon cancer and 52 cases of right-sided colon cancer in this study (**Figure 1**). The study used archived formalin-fixed paraffin-embedded (FFPE) samples and collected clinicopathological information. This study was approved by the Ethics Committee of Tianjin Nankai Hospital (approval number: NKYY_YXKT_IRB_2025_132_01).

In this study, patients' ages ranged from 32 to 91 years, with a mean age of 68.71± 12.00 years. The cohort included 62 males and 53 females. All 115 cases were histologically classified according to the World Health Organization (WHO) [20].

Inclusion criteria: (1) Patients admitted for the first time who underwent surgical resection and paraffin-embedded pathological examination; (2) Patients with complete clinical and pathological records. Exclusion criteria: (1) Patients with metastatic CRC or metastatic tumors from other organs; (2) Patients with uncertain pathological diagnoses or extremely low tumor cellularity in specimens; (3) Patients who had received preoperative radiotherapy, chemotherapy, immunotherapy, or other anti-tumor treatments.

Hematoxylin-eosin (H&E) and immunohistochemistry (IHC) technique

All 115 patients with difficult-to-evaluate CRC were retrospectively surveyed with H&E examination. Tissue sections 4-6 μm thick from FFPE were prepared for IHC, and these sections were processed with the streptavidin-peroxidase method. Antibodies against MLH1, MSH2, MSH6, and PMS2 (MXB Biotechnologies, China) were applied according to the manufacturers' protocols, and the suitable positive and negative controls were included in each experiment (Phosphate-Buffered Saline (PBS), Ethylenediaminetetraacetic acid (EDTA), Hematoxylin, 3,3'-Diaminobenzidine (DAB), Ventana Medical Systems, Inc., the United States of America).

Computational expression data and signaling pathway

RNA-sequencing data was used to characterize the expression of target genes in normal and cancer tissues across various tumor types.

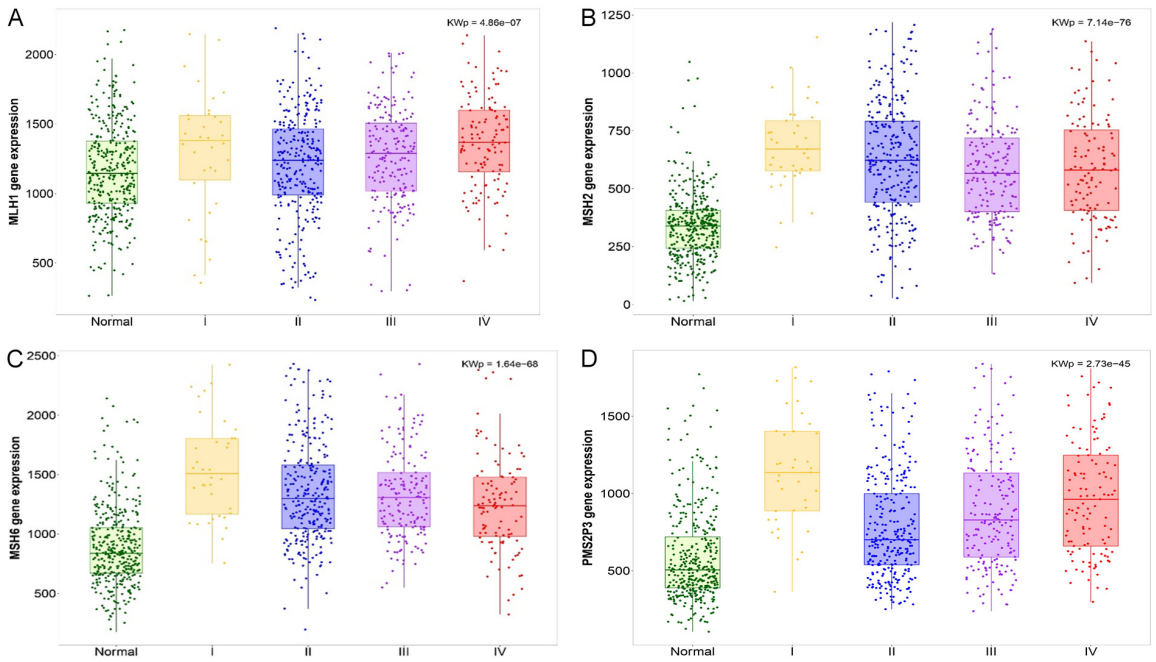

Figure 2. Differential gene expression analysis across pan-tumor, normal, and metastatic (TNM plot) tissues. A-D. Using gene array data, the combined expression of MLH1, MSH2, MSH6, and PMS2 was higher in colon tumor tissues across all the clinical stages relative to normal tissues (Kruskal-Wallis (KW) test, all P<0.001).

Pan-cancer box plots were used the Tumor, Normal, and Metastatic tissues (TNM) online tool, which were matched with clinical data of colon cancer samples retrieved from the Kaplan-Meier plotter database. Meanwhile, these data were used to investigate the relationships among the expression of MLH1, MSH2, MSH6 and PMS2. Additionally, we analyzed overall survival (OS, 1061 CRC samples) and recurrence-free survival (RFS, 1336 CRC samples) of patients. Finally, the Kyoto Encyclopedia of Genes and Genomes (KEGG) database was utilized to determine the correlations between these four genes and various signaling pathways.

Statistical analysis

Data were presented as ranges, means ± standard deviations, medians, absolute counts (n), and relative counts (%). All these data analyses were performed using Microsoft Excel. The Kruskal-Wallis (KW) test for TNM plot: differential gene expression analysis across pan-tumor, normal, and metastatic tissues. The chi-square test was applied to analyze the association between categorical variables. Correlation analysis was performed using the Pearson correlation test, and these analyses were performed using SPSS 22.0 software. Survival analyses for OS and RFS were executed using log-rank test (Kaplan-Meier plotter from a public database). The receiver operating characteristic (ROC) curve, generated using MedCalc software (version 23.2.1, MedCalc Software Ltd., Ostend, Belgium), was applied to determine the optimal cut-off value of variables, with the area under the curve (AUC) indicating predictive performance. Statistical significance was established at P<0.05.

Results

Differential expression of four MMR genes in tumor and normal tissues

TNM plot pan-cancer database was analyzed and compared to the expression patterns of four key MMR genes- *MLH1*, *MSH2*, *MSH6*, and *PMS2*- between CRC and adjacent normal colorectal tissues. The results demonstrated that the expression levels of all four MMR genes were notably elevated in CRC tissues compared with normal controls, and the differences were statistically significant (P<0.05; **Figure 2A-D**).

Survival analysis of four MMR genes in CRC

To evaluate the prognostic value of four MMR genes in CRC, Kaplan-Meier survival analysis

was performed. Based on median RNA expression levels of each target gene, CRC patients were categorized into two subgroups: the high- and low-expression group. The findings indicated that CRC patients with higher MLH1 expression had significantly higher RFS (P=0.0024; **Figure 3A1**, **3A2**), while OS (P=0.24) had no significant difference. Similarly, patients with higher MSH2 expression had significantly higher RFS (P<0.001; **Figure 3B1**, **3B2**), while OS (P=0.17) had no significance. Additionally, patients with higher MSH6 expression had significantly higher RFS (P<0.001; **Figure 3C1**, **3C2**), while OS (P=0.12) had no significant difference. Finally, patients with higher PMS2 expression had significantly higher RFS (P=0.0017; **Figure 3D1**, **3D2**), while OS (P=0.12) had no significant difference.

Pathway enrichment analysis of four MMR genes and proteins

KEGG pathway analysis revealed that four MMR genes were significantly enriched in the mismatch repair pathway. This pathway is responsible for detecting DNA replication errors, recruiting repair proteins, maintaining genomic stability, and regulating cell cycle progression and apoptosis (**Figure 4A**). Functional, MMR proteins suppress CRC development by preserving genome integrity; defects such as mutations or epigenetic silencing promote MSI and drive carcinogenesis, particularly in Lynch syndrome (LS) associated CRC (**Figure 4B**).

Differences in microsatellite stability and clinicopathological features in CRC

IHC examination results demonstrated that MMR-positive protein expression manifested as yellowish-brown granular staining localized in the cytoplasm and cell membrane of tumor cells (**Figure 5**). Subgroup analyses were performed according to clinical pathological parameters, including gender, age, lymph node status, pTNM stage, tumor grade, and size. The results indicated that there were no significant differences in MMR protein expression among these subgroups (all P>0.05). Notably, a significant difference in tumor location was observed between the MSI and MSS (P<0.05; **Table 1**).

Pathological features of MSI CRC and correlation analysis

Stratification analysis revealed that loss-MLH1 was closely associated with gender, age, and lymph node status in MSI tumors, with all differences reaching statistical significance (all P< 0.05; **Table 2**). Notably, significant correlations were found among most pair-wise combinations of MMR proteins, except for the combination of MLH1 and MSH2, which had no significant difference (P>0.05; **Table 3**).

ROC curve analysis of four MMR proteins and their combinations

The diagnostic efficiency of single MMR indicators was assessed via ROC analysis. The AUC of MLH1 was 0.808 (95% CI: 0.724-0.875); MSH2 was 0.538 (95% CI: 0.443-0.632); MSH6 was 0.558 (95% CI: 0.462-0.650); and PMS2 was 0.923 (95% CI: 0.858-0.965). Each of these molecules possessed certain predicative potency for CRC. In the combined detection models of dual proteins, the combination of MSH6 and PMS2 presented the best diagnostic performance, with an AUC of 0.981 (95% CI: 0.936-0.997) and an optimal cutoff value of 0.9615. The combination of MSH2 and PMS2 showed an AUC of 0.962 (95% CI: 0.908-0.989, cutoff: 0.9231), while the MLH1 and PMS2 yielded an AUC of 0.942 (95% CI: 0.883-0.977, cutoff: 0.8846). Moreover, the paired (MLH1 and MSH6) and (MLH1 and MSH2) exhibited moderate diagnostic capacity, with respective AUC values of 0.865 and 0.846. By contrast, the MSH2 and MSH6 combination produced the poorest predictive effect, with an AUC of only 0.558 and a cutoff value of 0.1154 (**Figure 6**; **Table 4**).

Discussion

Colorectal cancer (CRC) remains a common malignant tumor found around the world, and epidemiological predictions show that the yearly number of new CRC cases will rise to 3.2 million by the year 2040 [21]. CRC is projected to rank third in cancer incidence and second in cancer-associated mortality worldwide [22]. In the present study consisting of 115 CRC participants, the MSI subtype occupied 22.6% of all cases (26/115). We analyzed stage III-IV patients, and found that 92.3% of MSI tumors (24/26) were higher than 91.0% of MSS (81/89), these results are in line with the common situation that most colorectal cancer patients are diagnosed at a late stage [2], and they also suggest that MSI-type CRC may show stronger malignant development features. The-

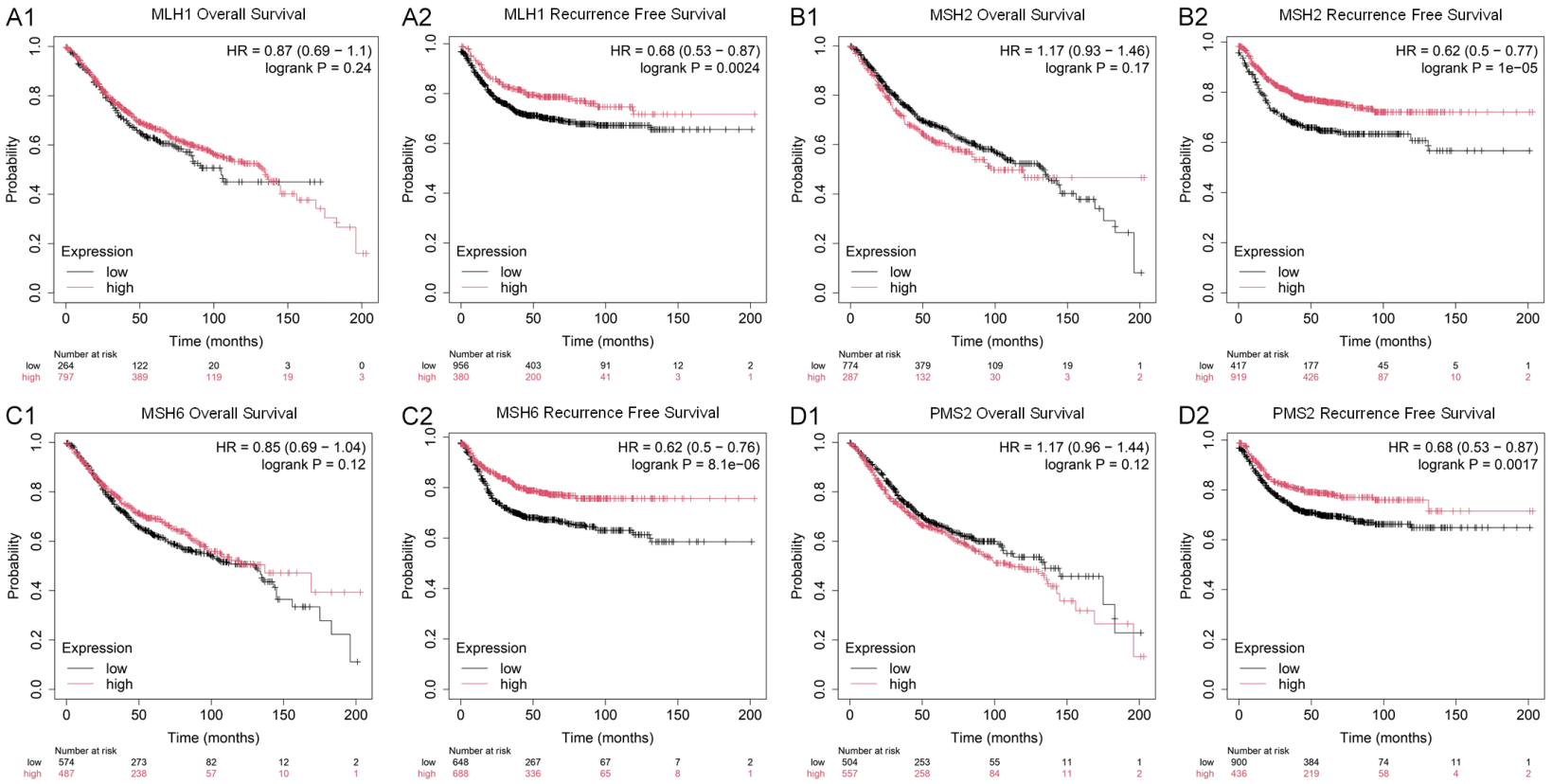

Figure 3. Survival analysis of *MLH1, MSH2, MSH6, and PMS2* expression in colon cancer. Survival analysis of *MLH1* in colon cancer indicated that patients with higher *MLH1* expression had favorable overall survival (OS) and significantly improved recurrence free survival (RFS), though the difference in OS did not reach statistical (OS: P=0.24, A1; RFS: P=0.0024, A2). Survival analysis of *MSH2* in colon cancer indicated that higher *MSH2* expression was associated with significantly better RFS (P<0.001, B2), whereas it tended to correlate with poorer OS, with no statistical significance (P=0.17, B1). Survival analysis of *MSH6* in colon cancer indicated that patients with higher *MSH6* expression had numerically better OS and significantly enhanced RFS, though the OS difference was not statistically significant (OS: P=0.12, C1; RFS: P<0.001, C2). Survival analysis of *PMS2* in colon cancer indicated that patients with higher *PMS2* expression were linked to significantly improved RFS (P=0.0017, D2), while it showed a trend toward worse OS without statistical significance (P=0.12, D1).

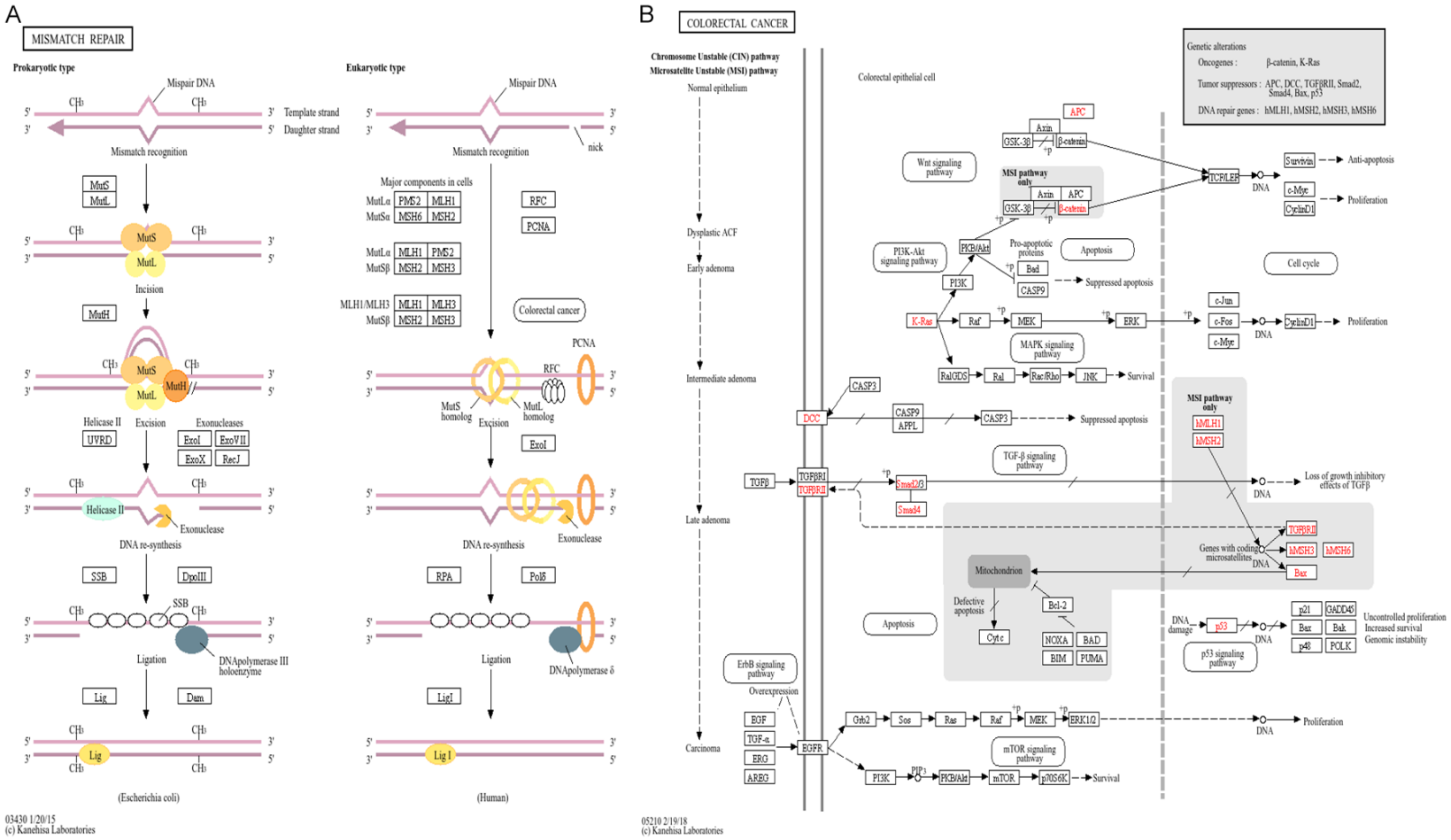

Figure 4. Target genes of *MLH1, MSH2, MSH6, and PMS2* in the Mismatch repair pathway (A) and colorectal cancer signaling pathway (B).

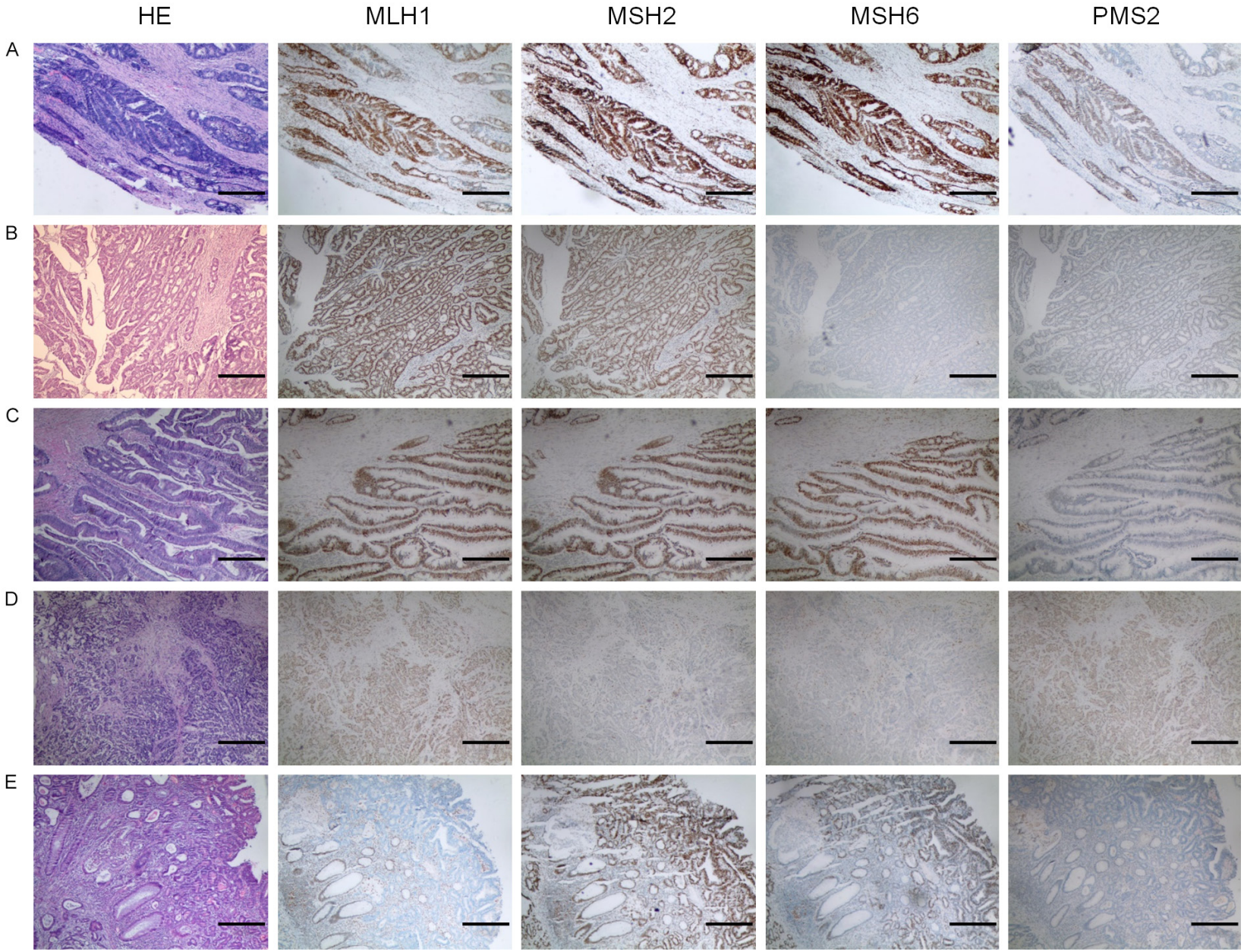
HE
MLH1
MSH2
MSH6
PMS2
A
B
C
D
E

Figure 5. Expression of MMR proteins in tumor tissues of CRC patients. A. Positive expression of MLH1, MSH2, MSH6, and PMS2 in CRC tissues. B. Positive expression of MLH1, MSH2, and PMS2, with negative expression of MSH6 in CRC tissues. C. Positive expression of MLH1, MSH2, and MSH6, with negative expression of PMS2 in CRC tissues. D. Positive expression of MLH1, and PMS2, with negative expression of MSH2 and MSH6 in CRC tissues. E. Positive expression of MSH2 and MSH6, with negative expression of MLH1 and PMS2 in CRC tissues (All scale bars = 40 μm, HE staining, ×400 magnification; IHC staining, ×400 magnification).

Table 1. Association between microsatellite stability status differences and clinicopathological features in CRC

Items	MSS (n=89)	MSI (n=26)	χ^2	*p*-value
Gender 0			3.228	0.072
Male	52	10		
Female	37	16		
Age (years)			0.338	0.561
<60	16	6		
≥60	73	20		
Lymph node			0.424	0.515
Positive	31	7		
Negative	58	19		
pTNM stage			0.043	0.836
I-II	8	2		
III-IV	81	24		
Tumor differentiation grade			0.904	0.342
Poorly differentiated	3	2		
Moderately to well differentiated	86	24		
Tumor location			14.188	<0.001
Rectum	22	1		
Left-sided colon	35	5		
Right-sided colon	32	20		
Tumor size (cm)			1.270	0.260
<3.5	10	1		
≥3.5	79	25		

Notes: CRC, Colorectal cancer; MSS, microsatellite stable; MSI, microsatellite instability.

refore, we need to carry out more deep research on the internal molecular control mechanisms.

Genome instability, especially chromosomal instability (CIN), is a key characteristic of colorectal cancer. DNA damage and impaired DNA repair signaling are recognized as major endogenous factors driving tumor progression [23]. Genes involved in DNA repair and genome stability are key contributors to hereditary tumor syndromes [24], and modern oncology research increasingly focuses on DNA repair defects as actionable therapeutic targets [25]. KEGG analysis showed that the four MMR genes were mainly concentrated in the mismatch repair pathway. This pathway is used to recognize DNA replication errors, recruit related proteins, repair DNA, regulate the cell cycle and modulate cell apoptosis. Normal MMR proteins can inhibit the development of CRC. In contrast, gene changes or silent states may cause MSI and enhance tumorigenesis, especially in LS-related CRC [26]. These results can prove that MMR defects lead to gene instability by affecting many biological processes, at the same time, damaged MMR function may also weaken the homologous recombination repair process. Meanwhile, MMR functional defects may impair the homologous recombination repair process. This provides a promising molecular target for developing combined therapeutic strategies targeting the DNA repair system.

DNA repair pathways usually function in a mutually exclusive fashion, a feature that exerts considerable impact on the design of genetic

Table 2. Association between MSI related gene expression and clinicopathological features in CRC

Items	MLH1		χ^2	*p*-value	MSH2		χ^2	*p*-value	MSH6		χ^2	*p*-value	PMS2		χ^2	*p*-value
	(+)	(-)			(+)	(-)			(+)	(-)			(+)	(-)		
Gender			6.289	0.009			0.122	0.727			0.038	0.846			0.038	0.846
Male	7	3			9	1			9	1			9	1		
Female	3	13			15	1			14	2			14	2		
Age (years)			6.635	0.010			0.885	0.347			3.363	0.057			3.363	0.057
<60	5	1			5	1			4	2			2	4		
≥60	5	15			19	1			19	1			1	19		
Lymph node			4.398	0.036			0.798	0.372			1.249	0.264			1.249	0.264
Positive	5	2			7	0			7	0			0	7		
Negative	5	14			17	2			16	3			3	16		
pTNM stage			1.354	0.245			0.181	0.671			0.283	0.595			0.283	0.595
I-II	0	2			2	0			2	0			0	2		
III-IV	10	14			22	2			21	3			3	21		
Tumor differentiation grade			0.038	0.846			3.140	0.076			1.578	0.209			1.578	0.209
Poorly differentiated	1	2			2	1			2	1			1	2		
Moderately to well differentiated	9	14			22	1			21	2			2	21		
Tumor location			1.706	0.426			0.650	0.723			1.017	0.601			1.017	0.601
Rectum	1	0			1	0			1	0			0	1		
Left-sided colon	2	3			5	0			5	0			0	5		
Right-sided colon	7	13			18	2			17	3			3	17		
Tumor size (cm)			0.650	0.420			0.087	0.768			0.136	0.713			0.136	0.713
<3.5	0	1			1	0			1	0			0	1		
≥3.5	10	15			23	2			22	3			3	22		

Notes: MLH1, MutL homolog 1; MSH2, MutS homolog 2; MSH6, MutS homolog 6; PMS2, PMS1 homolog 2.

Table 3. Correlation analysis of MLH1, MSH2, MSH6, and PMS2 in MSI CRC

Correlation pair	r	SE	95% CI	*p*-value	Statistical significance
MLH1 vs MSH2	-0.365	0.113	[-0.640, -0.200]	0.067	-
MLH1 vs MSH6	-0.457*	0.129	[-0.732, -0.228]	0.019	*
MLH1 vs PMS2	0.457*	0.129	[0.228, 0.732]	0.019	*
MSH2 vs MSH6	0.799**	0.158	[0.527, 1.000]	0.000	**
MSH2 vs PMS2	-0.799**	0.158	[-1.000, -0.527]	0.000	**
MSH6 vs PMS2	-1.000**	0.000	[-1.000, -1.000]	0.000	**

Notes: CI, 95% confidence interval; SE, standard error; vs, versus (Pearson, *P<0.05; **P<0.01).

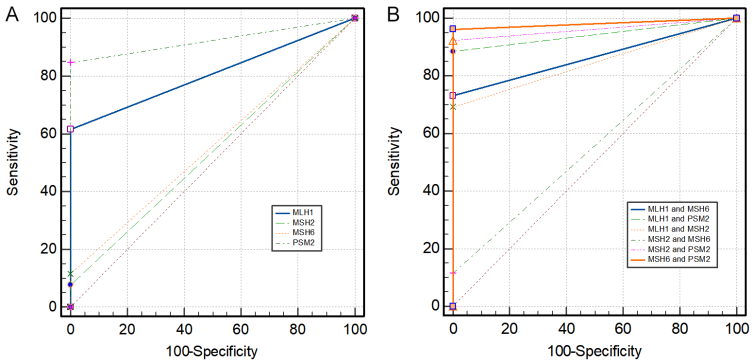

Figure 6. Receiver operating characteristics (ROC) curves for MLH1, MSH2, MSH6, PMS2, and their combination in CRC. A. ROC curves for MLH1, MSH2, MSH6, PMS2. B. Predictive performance of different combined models of MLH1, MSH2, MSH6, PMS2 in CRC. Notes: Among patients with colorectal cancer (as shown in **Table 4**), the MSH6 and PMS2 combination showed the highest predictive accuracy, followed by MSH2 and PMS2, and MLH1 and PMS2 (all P<0.001). The MSH2 and MSH6 model demonstrated no significant predictive value (P=0.394).

screening approaches and the stratification of therapeutic regimens [27]. For instance, hereditary breast, ovarian, pancreatic, and prostate cancers are linked to homologous recombination deficiency (HRD) and have long been categorized as "HRD-associated cancers" [28]. On the other hand, hereditary CRC and endometrial cancers are distinguished by MSI, which arises from defects in the MMR system, and they display a unique mutational profile [28, 29]. Our clinical data further explain these different features. Among the 115 CRC patients enrolled in this research, MSI tumors were located in the right-sided colon, accounting for 76.9% (20/26), which was higher than 19.2% (5/26) in the left-sided colon. On the other hand, MSS tumors tended to occur in the left colon, with a proportion of 39.3% (35/89). This result is consistent with the well-established observation that "MSI-positive CRCs are mostly originated from the right-sided colon" [30, 31], and it further strengthens the clinical rationale that "tumor location can act as a preliminary screening indicator for MSI status" [32, 33]. In clinical practice, this finding provides support for optimizing the CRC screening workflow, specifically advocating for the strategy of "initial location-based assessment followed by subtype-specific stratification" [34].

The DNA mismatch repair (MMR) system is an evolutionarily conserved cellular mechanism responsible for rectifying replication-related errors, thereby preserving genomic integrity [35]. Mismatch formation can be triggered by multiple factors, such as incorrect nucleotide incorporation by DNA polymerase, strand recombination events, and DNA damage induced by chemical or physical agents [35]. MMR genes are designated based on their bacterial counterparts, which include MutS homologs (MSH2, MSH3, MSH4, MSH5, MSH6) and MutL homologs (MLH1, MLH3, PMS1, PMS2) [36]. Our IHC results were consistent with the known molecular interactions among MMR proteins. The simultaneous loss of MLH1 and PMS2 was the most common pattern, accounting for 61.5% (16/26) of cases, followed by the isolated PMS2 loss at 26.9% (7/26). In contrast, the combined loss of MSH2 and MSH6 (7.7%, 2/26) and isolated MSH6 loss (3.9%, 1/26) were relatively rare. This expression profile aligns with the well-established biological requirement that MLH1 must form a dimer with PMS2 to fulfill its functional role [37]. Mutations in the MLH1 gene impair the stability of PMS2, resulting in the concurrent downregulation of both proteins [38, 39]. These observations further confirm the functional

Table 4. Receiver operating characteristics (ROC) curves for predictive performance of different combined models of MLH1, MSH2, MSH6, PMS2 in CRC

Variables	AUC	SE	95% CI	Cut-off Value	*p*-value
MLH1 and MSH6	0.865	0.0533	0.789-0.922	0.7308	<0.001
MLH1 and PMS2	0.942	0.0370	0.883-0.977	0.8846	<0.001
MLH1 and MSH2	0.846	0.0561	0.767-0.907	0.6923	<0.001
MSH2 and MSH6	0.558	0.0677	0.462-0.650	0.1154	0.394
MSH2 and PMS2	0.962	0.0307	0.908-0.989	0.9231	<0.001
MSH6 and PMS2	0.981	0.0220	0.936-0.997	0.9615	<0.001

Notes: AUC, area under the curve. Compared with individual markers (**Figure 6A**), paired combinations (**Figure 6B**) improved diagnostic accuracy, with the MSH6/PMS2 combination performing best.

cooperation among MMR proteins, and this cooperation is of great significance for clinical diagnostic practices.

We confirmed that MMR proteins form heterodimeric complexes, which can identify mismatched bases, respond to small insertions and deletions in genes, and assist in repairing such gene structural abnormalities. Deficiency in MMR function can be caused by multiple pathways: acquired somatic mutations in MMR genes often lead to gene inactivation, among which hypermethylation of the MLH1 promoter-resulting in transcriptional silencing of the MLH1 gene-is the most common mechanism [40, 41]. Germline mutations in MMR genes also contribute substantially to MMR deficiency (dMMR). Pathogenic germline variants in MLH1, MSH2, MSH6, PMS2, or 3'-end deletions of the EPCAM gene, which induce hypermethylation of the MSH2 promoter, are well-recognized key etiological factors in hereditary CRC and LS [41]. We analyzed the clinical characteristics and further clarified the associations with MMR protein expression. Specifically, MLH1 protein loss was more common in patients aged ≥60 years (57.7%, 15/26) than in those younger than 60 years (3.9%, 1/26; P=0.01). Additionally, MLH1 loss occurred more frequently in cases without lymph node metastasis (53.8%, 14/26) than in those with lymph node metastasis (7.7%, 2/26; P=0.036). These trends indicate that MLH1 loss is related to age-related risks, which may be attributed to the accumulation of epigenetic changes such as MLH1 methylation, and it is also linked to a lower metastatic potential. MLH1 loss may serve as a potential marker for prognostic stratification, though further validation with survival data is required. Notably, while MLH1 hypermethylation is generally associated with sporadic CRC [42], it has been documented in rare cases of LS. Specifically, germline MLH1 mutations in LS most commonly result in the concurrent loss of MLH1 and PMS2, followed by the loss of MSH2 and MSH6 driven by germline MSH2 mutations [43].

In clinical practice, the combined detection models (including MSH6 and PMS2, MSH2 and PMS2, MLH1 and PMS2) exhibit high predictive efficacy, which provides important biological markers for judging tumor biological behaviors and guiding clinical decisions. Among these models, the MSH6 and PMS2 combined model has the highest predictive accuracy, making it the optimal combined detection strategy in clinical practice to improve the accuracy of predictive evaluation. The superior predictive effect of the MSH6 and PMS2 combined model may be associated with the mutual cooperation between these two proteins in the MMR pathway: MSH6 can accurately identify DNA mismatches, while PMS2 assist in smoothly completing the DNA repair process; the combined detection of both proteins can fully reflect the functional state of the MMR system. The high AUC values of the MSH2 and PMS2, MLH1 and PMS2 combined models also confirm that PMS2 serves as a core component of the MutLα complex, which maintains MMR function; combined detection with other MMR proteins can greatly improve the predictive accuracy. In addition, the optimal cut-off values of each combined model calculated in this study (ranging from 0.1154 to 0.9615) can serve as a reliable reference for clinical judgment of positive and negative results. This finding validates the clinical significance of the notion that "combined detection of multiple MMR proteins improves the accuracy of MSI status determination" [44], providing quantitative evidence to

optimize MMR testing protocols for accurate CRC. This is especially important in clinical work, where we need to distinguish between sporadic and hereditary CRC; this step is crucial for formulating personalized treatment.

Conclusion

Our study confirmed that MSI testing and MMR protein expression are closely associated with the clinical study of CRC. These detection biomarkers can effectively distinguish tumor characteristics and assist clinicians in assessing disease progression and patient prognosis. By investigating the molecular interactions between PMS2 with MSH6 and MSH2, we can enrich basic research data and promote the popularization and standardization of combined MMR protein detection in clinical practice. Meanwhile, our research has obvious deficiencies: all patients and samples were collected from a single hospital, leading to a limited sample scope. Further multi-center studies with a larger number of cases are needed to prove whether our results can be suitable for a wider clinical setting.

Acknowledgements

This work was supported by the Department of Pathology, Tianjin Nankai Hospital, Tianjin, China and Tianjin King Med Diagnostics Laboratory Co., Ltd., The People's Republic of China.

Disclosure of conflict of interest

None.

Address correspondence to: Xinxin Liu, Department of Pathology, Tianjin Nankai Hospital, No. 6, Changjiangdao, Nankai District, Tianjin 300100, The People's Republic of China. Tel: +86-22-2743-5308; E-mail: 645845259@qq.com

References

[1] Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I and Jemal A. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2024; 74: 229-263.

[2] Siegel RL, Wagle NS, Cercek A, Smith RA and Jemal A. Colorectal cancer statistics, 2023. CA Cancer J Clin 2023; 73: 233-254.

[3] Menon G and Cagir B. Colon Cancer. In: StatPearls. Treasure Island (FL): StatPearls Publishing Copyright © 2025, StatPearls Publishing LLC.; 2025.

[4] Fan WX, Su F, Zhang Y, Zhang XL, Du YY, Gao YJ, Li WL, Hu WQ and Zhao J. Oncological characteristics, treatments and prognostic outcomes in MMR-deficient colorectal cancer. Biomark Res 2024; 12: 89.

[5] Strand M, Prolla TA, Liskay RM and Petes TD. Destabilization of tracts of simple repetitive DNA in yeast by mutations affecting DNA mismatch repair. Nature 1993; 365: 274-276.

[6] Michaels ML, Cruz C, Grollman AP and Miller JH. Evidence that MutY and MutM combine to prevent mutations by an oxidatively damaged form of guanine in DNA. Proc Natl Acad Sci U S A 1992; 89: 7022-7025.

[7] Kadyrov FA, Dzantiev L, Constantin N and Modrich P. Endonucleolytic function of MutLalpha in human mismatch repair. Cell 2006; 126: 297-308.

[8] Kadyrov FA, Holmes SF, Arana ME, Lukianova OA, O'Donnell M, Kunkel TA and Modrich P. Saccharomyces cerevisiae MutLalpha is a mismatch repair endonuclease. J Biol Chem 2007; 282: 37181-37190.

[9] Kramer B, Kramer W and Fritz HJ. Different base/base mismatches are corrected with different efficiencies by the methyl-directed DNA mismatch-repair system of E. coli. Cell 1984; 38: 879-887.

[10] Nik-Zainal S, Alexandrov LB, Wedge DC, Van Loo P, Greenman CD, Raine K, Jones D, Hinton J, Marshall J, Stebbings LA, Menzies A, Martin S, Leung K, Chen L, Leroy C, Ramakrishna M, Rance R, Lau KW, Mudie LJ, Varela I, McBride DJ, Bignell GR, Cooke SL, Shlien A, Gamble J, Whitmore I, Maddison M, Tarpey PS, Davies HR, Papaemmanuil E, Stephens PJ, McLaren S, Butler AP, Teague JW, Jönsson G, Garber JE, Silver D, Miron P, Fatima A, Boyault S, Langerød A, Tutt A, Martens JW, Aparicio SA, Borg Å, Salomon AV, Thomas G, Børresen-Dale AL, Richardson AL, Neuberger MS, Futreal PA, Campbell PJ and Stratton MR; Breast Cancer Working Group of the International Cancer Genome Consortium. Mutational processes molding the genomes of 21 breast cancers. Cell 2012; 149: 979-993.

[11] Le DT, Uram JN, Wang H, Bartlett BR, Kemberling H, Eyring AD, Skora AD, Luber BS, Azad NS, Laheru D, Biedrzycki B, Donehower RC, Zaheer A, Fisher GA, Crocenzi TS, Lee JJ, Duffy SM, Goldberg RM, de la Chapelle A, Koshiji M, Bhaijee F, Huebner T, Hruban RH, Wood LD, Cuka N, Pardoll DM, Papadopoulos N, Kinzler KW, Zhou S, Cornish TC, Taube JM, Anders RA, Eshleman JR, Vogelstein B and Diaz LA Jr. PD-1

blockade in tumors with mismatch-repair deficiency. N Engl J Med 2015; 372: 2509-2520.
[12] Overman MJ, McDermott R, Leach JL, Lonardi S, Lenz HJ, Morse MA, Desai J, Hill A, Axelson M, Moss RA, Goldberg MV, Cao ZA, Ledeine JM, Maglinte GA, Kopetz S and André T. Nivolumab in patients with metastatic DNA mismatch repair-deficient or microsatellite instability-high colorectal cancer (CheckMate 142): an open-label, multicentre, phase 2 study. Lancet Oncol 2017; 18: 1182-1191.
[13] Nojadeh JN, Behrouz Sharif S and Sakhinia E. Microsatellite instability in colorectal cancer. Excli J 2018; 17: 159-168.
[14] Shannon AB, Mehta R, Mok SR, Lauwers GY, Baldonado JJAR, Fontaine J, Pimiento JM and Sinnamon AJ. Clinical and pathologic response to neoadjuvant immunotherapy in dna mismatch repair protein-deficient gastroesophageal cancers. Ann Surg Oncol 2024; 31: 8616-8626.
[15] Valabrega G, Powell MA, Hietanen S, Miller EM, Novak Z, Holloway R, Denschlag D, Myers T, Thijs AM, Pennington KP, Gilbert L, Fleming E, Zub O, Landrum LM, Ataseven B, Gogoi R, Podzielinski I, Cloven N, Monk BJ, Sharma S, Herzog TJ, Stuckey A, Pothuri B, Secord AA, Chase D, Vincent V, Meyers O, Garside J, Mirza MR and Black D. Patient-reported outcomes in the subpopulation of patients with mismatch repair-deficient/microsatellite instability-high primary advanced or recurrent endometrial cancer treated with dostarlimab plus chemotherapy compared with chemotherapy alone in the ENGOT-EN6-NSGO/GOG3031/RUBY trial. Int J Gynecol Cancer 2025; 35: 101852.
[16] Alves GM, Severino-Lazo RJ, Amaral-Silva GK, Silveira FM and Carvalho MD. DNA mismatch repair system expression in salivary gland tumors: a systematic review. Med Oral Patol Oral Cir Bucal 2024; 29: e655-e664.
[17] Denu RA, Quintana-Perez CD, Wangsiricharoen S, Ingram DR, Wani KM, Lazar AJ, Ratan R, Roland CL and You YN. DNA mismatch repair deficiency as a biomarker in sarcoma. Surg Oncol Insight 2024; 1: 100091.
[18] Williams CJM, Peddle AM, Kasi PM, Seligmann JF, Roxburgh CS, Middleton GW and Tejpar S. Neoadjuvant immunotherapy for dMMR and pMMR colorectal cancers: therapeutic strategies and putative biomarkers of response. Nat Rev Clin Oncol 2024; 21: 839-851.
[19] Klümper N, Grünwald V, Hartmann A, Hölzel M and Eckstein M. The role of microsatellite instability/DNA mismatch repair deficiency and tumor mutational burden as biomarkers in predicting response to immunotherapy in castration-resistant prostate cancer. Eur Urol 2024; 86: 388-390.
[20] Nagtegaal ID, Odze RD, Klimstra D, Paradis V, Rugge M, Schirmacher P, Washington KM, Carneiro F and Cree IA. The 2019 WHO classification of tumours of the digestive system. Histopathology 2020; 76: 182-188.
[21] Siegel RL, Giaquinto AN and Jemal A. Cancer statistics, 2024. CA Cancer J Clin 2024; 74: 12-49.
[22] Xi Y and Xu P. Global colorectal cancer burden in 2020 and projections to 2040. Transl Oncol 2021; 14: 101174.
[23] Tubbs A and Nussenzweig A. Endogenous DNA damage as a source of genomic instability in cancer. Cell 2017; 168: 644-656.
[24] Curtin NJ. DNA repair dysregulation from cancer driver to therapeutic target. Nat Rev Cancer 2012; 12: 801-817.
[25] Russo A, Incorvaia L, Capoluongo E, Tagliaferri P, Galvano A, Del Re M, Malapelle U, Chiari R, Conte P, Danesi R, Fassan M, Ferrara R, Genuardi M, Ghiorzo P, Gori S, Guadagni F, Marchetti A, Marchetti P, Midiri M, Normanno N, Passiglia F, Pinto C, Silvestris N, Tallini G, Vatrano S, Vincenzi B, Cinieri S and Beretta G. The challenge of the molecular tumor board empowerment in clinical oncology practice: a position paper on behalf of the AIOM- SIAPEC/IAP-SIBioC-SIC-SIF-SIGU-SIRM Italian scientific societies. Crit Rev Oncol Hematol 2022; 169: 103567.
[26] Mas-Ponte D, McCullough M and Supek F. Spectrum of DNA mismatch repair failures viewed through the lens of cancer genomics and implications for therapy. Clin Sci (Lond) 2022; 136: 383-404.
[27] Dietlein F, Thelen L and Reinhardt HC. Cancer-specific defects in DNA repair pathways as targets for personalized therapeutic approaches. Trends Genet 2014; 30: 326-339.
[28] Groelly FJ, Fawkes M, Dagg RA, Blackford AN and Tarsounas M. Targeting DNA damage response pathways in cancer. Nat Rev Cancer 2023; 23: 78-94.
[29] Adar T, Rodgers LH, Shannon KM, Yoshida M, Ma T, Mattia A, Lauwers GY, Iafrate AJ, Hartford NM, Oliva E and Chung DC. Universal screening of both endometrial and colon cancers increases the detection of Lynch syndrome. Cancer 2018; 124: 3145-3153.
[30] Piredda ML, Ammendola S, Sciammarella C, Turri G, Bagante F, Fassan M, Mafficini A, Mombello A, Cataldi S, Paolino G, Mattiolo P, Florena AM, Genna M, Fior F, Cheng L, Lawlor RT, Scarpa A, Pedrazzani C and Luchini C. Colorectal cancer with microsatellite instability: right-sided location and signet ring cell histology are associated with nodal metastases, and extranodal extension influences disease-free survival. Pathol Res Pract 2021; 224: 153519.

[31] Chen Y, Chen Z, Huang J, Hu J, He X, Lan P and He X. Clinicopathological and molecular characteristics of early-onset vs late-onset colorectal cancer according to tumor location. Int J Clin Oncol 2022; 27: 749-755.
[32] Arakawa K, Hata K, Kawai K, Tanaka T, Nishikawa T, Sasaki K, Shuno Y, Kaneko M, Hiyoshi M, Emoto S, Murono K and Nozawa H. Predictors for high microsatellite instability in patients with colorectal cancer fulfilling the revised bethesda guidelines. Anticancer Res 2018; 38: 4871-4876.
[33] Sinicrope FA, Rego RL, Foster N, Sargent DJ, Windschitl HE, Burgart LJ, Witzig TE and Thibodeau SN. Microsatellite instability accounts for tumor site-related differences in clinicopathologic variables and prognosis in human colon cancers. Am J Gastroenterol 2006; 101: 2818-2825.
[34] Huang J, Xu L, Cai Y, Tan X, Lin H, Chen Z, Wang C, Deng W and Fu X. Certain pMMR colorectal cancer patients should undergo additional MSI-PCR testing to reduce the risk of misdiagnosing MSI-H and Lynch syndrome. BMC Cancer 2025; 25: 1103.
[35] Gupta D and Heinen CD. The mismatch repair-dependent DNA damage response: mechanisms and implications. DNA Repair (Amst) 2019; 78: 60-69.
[36] Marti TM, Kunz C and Fleck O. DNA mismatch repair and mutation avoidance pathways. J Cell Physiol 2002; 191: 28-41.
[37] Lin X, Bian C, Liang D, Li L, Tang Q and Li Q. An in-frame deletion mutation in MLH1 drives Lynch syndrome-associated colorectal cancer. BMC Cancer 2025; 25: 1165.
[38] Wu X, Platt JL and Cascalho M. Dimerization of MLH1 and PMS2 limits nuclear localization of MutLalpha. Mol Cell Biol 2003; 23: 3320-3328.
[39] Kosinski J, Hinrichsen I, Bujnicki JM, Friedhoff P and Plotz G. Identification of Lynch syndrome mutations in the MLH1-PMS2 interface that disturb dimerization and mismatch repair. Hum Mutat 2010; 31: 975-982.
[40] Umar A, Boland CR, Terdiman JP, Syngal S, de la Chapelle A, Rüschoff J, Fishel R, Lindor NM, Burgart LJ, Hamelin R, Hamilton SR, Hiatt RA, Jass J, Lindblom A, Lynch HT, Peltomaki P, Ramsey SD, Rodriguez-Bigas MA, Vasen HF, Hawk ET, Barrett JC, Freedman AN and Srivastava S. Revised bethesda guidelines for hereditary nonpolyposis colorectal cancer (Lynch syndrome) and microsatellite instability. J Natl Cancer Inst 2004; 96: 261-268.
[41] Boland CR and Goel A. Microsatellite instability in colorectal cancer. Gastroenterology 2010; 138: 2073-2087.e2073.
[42] Carnevali IW, Cini G, Libera L, Sahnane N, Facchi S, Viel A, Sessa F and Tibiletti MG. *MLH1* promoter methylation could be the second hit in lynch syndrome carcinogenesis. Genes (Basel) 2023; 14: 2060.
[43] Alpert L, Pai RK, Srivastava A, McKinnon W, Wilcox R, Yantiss RK, Arcega R, Wang HL, Robert ME, Liu X, Pai RK, Zhao L, Westerhoff M, Hampel H, Kupfer S, Setia N, Xiao SY, Hart J and Frankel WL. Colorectal carcinomas with isolated loss of PMS2 staining by immunohistochemistry. Arch Pathol Lab Med 2018; 142: 523-528.
[44] Fu G, Li J, Zhang Q, Lv C, Zhang Z, Wang X, Wu R and Chen L. Detecting microsatellite instability in cancer via multiplexed orthogonal gap-enhanced Raman tags. Chem Sci 2025; 16: 10881-10894.