CANCER
COMMUNICATIONS
Open Access

Germline mutational profile of Chinese patients under 70 years old with colorectal cancer

Teng-Jia Jiang[1,†] | **Fang Wang[2,†]** | **Ying-Nan Wang[3,†]** | **Jia-Jia Hu[3]** | **Pei-Rong Ding[4]** | **Jun-Zhong Lin[4]** | **Zhi-Zhong Pan[4]** | **Gong Chen[4]** | **Jian-Yong Shao[2]** | **Rui-hua Xu[1]** | **Qi Zhao[3,*]** | **Feng Wang[1,*]**

[1] State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Department of Oncology, Sun Yat-sen University Cancer Center, Guangzhou, Guangdong 510060, P. R. China

[2] Department of Molecular Diagnosis, Sun Yat-sen University Cancer Center, Guangzhou, Guangdong 510060, P. R. China

[3] Department of Experimental Research, Sun Yat-sen University Cancer Center, Guangzhou, Guangdong 510060, P. R. China

[4] Department of Colorectal Surgery, Sun Yat-sen University Cancer Center, Guangzhou, Guangdong 510060, P. R. China

Correspondence
Dr. Feng Wang, State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Department of Oncology, Sun Yat-sen University Cancer Center, Guangzhou 510060, Guangdong, P. R. China.
Email: wangfeng@sysucc.org.cn
Dr. Qi Zhao, State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Department of Experimental Research, Sun Yat-sen University Cancer Center, Guangzhou 510060, Guangdong, P. R. China.
Email: zhaoqi@susycc.org.cn

*To whom correspondence should be addressed

†These authors contributed equally to this work.

Abstract
Background: Inherited susceptibility accounts for nearly one-third of colorectal cancer (CRC) predispositions and has an 80%-100% lifetime risk of this disease. However, there are few data about germline mutations of hereditary CRC-related genes in Chinese patients with CRC. This study aimed to assess the prevalence of gene mutations related to cancer susceptibility among Chinese patients with CRC, differences between Chinese and Western patients, and the phenotype-genotype correlation.
Methods: We retrospectively collected tumor samples from 526 patients with CRC under 70 years old who underwent hereditary CRC genetic testing. A series of bioinformatic analyses, as well as statistical comparisons, were performed.
Results: We found that 77 patients (14.6%) harbored functional variants of the 12 genes. The mutation frequencies of the top 5 mutated genes were 6.5% for MutL homolog 1 (*MLH1*), 5.1% for MutS homolog 2 (*MSH2*), 1.0% for *MSH6*, 0.8% for *PMS1* homolog 2 (*PMS2*), and 0.8% for APC regulator of the WNT signaling pathway (*APC*). Our data showed much higher rates of mutations of *MSH6* and *PMS2* genes among all mismatch repair (MMR) genes as compared with those in Western populations. Mutations in *MLH1*, *MSH2*, and *MSH6* were found to be mutually exclusive. Patients with *MLH1* or *MSH2* mutations had higher frequencies of personal history of cancer (*MLH1*: 20.6% vs. 8.7%; *MSH2*: 25.9% vs. 8.6%) and family history of cancer than those without these mutations (*MLH1*: 73.5% vs. 48.4%; *MSH2*: 70.4% vs. 48.9%), and the lesions were more prone to occur on the right

side of the colon than on the left side (*MLH1*: 73.5% vs. 29.3%; *MSH2*: 56.0% vs. 31.0%). The proportion of stage I/II disease was higher in patients with *MLH1* mutations than in those without *MLH1* mutations (70.6% vs. 50.7%), and the rate of polyps was higher in patients with *APC* mutations than in those with wild-type *APC* (75.0% vs. 17.4%).

Conclusion: These results provide a full-scale landscape of hereditary susceptibility over 12 related genes in CRC patients and suggest that a comprehensive multi-gene panel testing for hereditary CRC predisposition could be a helpful analysis in clinical practice.

KEYWORDS
colorectal cancer, genetic testing, germline mutation, hereditary CRC syndromes

1 | BACKGROUND

Colorectal cancer (CRC) is a common cancer worldwide. As of 2018, CRC was the second leading cause of cancer death affecting women (9.5%) and the third affecting men (10.9%) [1]. The etiology of CRC is both genetic and environmental. Accordingly, inherited susceptibility accounts for nearly one-third of CRC predispositions [2,3], playing a crucial role in CRC risk [4,5]. In developed countries, the average lifetime risk of CRC is approximately 2%–5% [6]. However, the lifetime risk of CRC increases up to 20% when the patient has a familial history of CRC and reaches 80%–100% in patients with hereditary CRC syndromes [4,7,8]. The hereditary risk factors of CRC are becoming a hot research topic globally.

Generally, hereditary CRC syndromes are divided into three subgroups: polyposis syndromes, nonpolyposis syndromes, and other syndromes related to CRC [9,10]. The main type of polyposis syndrome is familial adenomatous polyposis (FAP), which is an autosomal-dominant CRC syndrome that is characterized by the formation of hundreds to thousands of adenomatous colorectal polyps in early adolescence. Attenuated familial adenomatous polyposis (AFAP) is a less-severe form of the disease showing fewer polyps. Other polyposis syndromes have also been defined, such as mutY DNA glycosylase (*MUTYH*)-associated polyposis (MAP). Above CRC hereditary syndromes are caused by corresponding mutations [APC regulator of the WNT signaling pathway (*APC*) in FAP/AFAP and *MUTYH* in MAP]. Another main subgroup is Lynch syndrome (LS), which is also called hereditary non-polyposis colorectal cancer (HNPCC). It is a common form of inherited CRC and accounts for 2.0%–5.5% of the overall CRC burden [4,11,12]. LS is an autosomal-dominant inherited disease that increases the risk of many types of cancer, such as endometrial cancer, urinary tract

cancer, gastric cancer, and especially CRC [13,14]. LS is related to mismatch repair (MMR) genes [MutL homolog 1 (*MLH1*), MutS homolog 2 (*MSH2*), *MSH6*, and *PMS1* homolog 2 (*PMS2*)] and large deletions of 3′ end of epithelial cell adhesion molecule (*EPCAM*). Our gene testing panel also included *PMS1* and *MLH3*, but we found no solid evidence to prove that germline mutations of *PMS1* and *MLH3* could cause a predisposition to LS [15–17]. Therefore, we removed these two genes from our analysis to avoid potential disputes. The third subgroup is comprised of juvenile polyposis syndrome (JPS), Peutz-Jeghers syndrome (PJS), phosphatase and tensin homolog (*PTEN*)-hamartoma tumor syndrome (PHTS) [18], and oligodontia-colorectal cancer syndrome (ODCRCS) [19], and is related to specific genes [bone morphogenetic protein receptor type 1A (*BMPR1A*) and SMAD family member 4 (*SMAD4*) in JPS, serine/threonine kinase 11 (*STK11*) in PJS, *PTEN* in PHTS, and axin 2 (*AXIN2*) in ODCRCS]. Previous studies have evaluated the prevalence of hereditary CRC syndromes mainly in Western countries, which were confined to several populations with specific syndromes, such as LS or MAP, and sequenced using a gene testing panel with limited genes for hereditary CRC [20–22]. Because these studies either emphasized restrictive CRC syndromes or lacked sufficient number of samples, a larger cohort for the prevalence assessment of inherited cancer susceptibility among CRC patients is needed.

Moreover, limited data, focusing on only MMR genes, have been published about germline hereditary susceptibility in Chinese CRC patients [23]. In our present research, we collected tumor samples from CRC patients diagnosed at the Sun Yat-sen University Cancer Center (SYSUCC; Guangzhou, Guangdong, China), who were recommended to undergo a hereditary CRC 14-gene screening. These 14 genes are associated with the syndromes and

phenotypes of both nonpolyposis syndrome (LS) and some of the polyposis syndromes mentioned above. The current study mainly aimed to assess the prevalence of gene mutations related to cancer susceptibility among Chinese patients with CRC, to find the differences between Chinese and Western patients, and to explore the phenotype-genotype correlation. This comprehensive germline mutation assessment of Chinese patients with CRC may provide evidence support for clinical practice, facilitate primary prevention, and increase the health benefits of patients and their families.

2 | MATERIALS AND METHODS

2.1 | Patient selection

All patients had undergone hereditary CRC genetic testing at SYSUCC between October 2014 and August 2016 were eligible for enrollment. They had been referred for genetic counseling when themselves or their family members had some high-risk features for CRC [e.g., young age at diagnosis, personal/family history of cancer or polyps, tumor microsatellite instability (MSI), and MMR deficiency]. Several studies recommended universal LS screening to CRC patients up to 70 years of age [24,25], which is cost-effective. Therefore, we selected patients under 70 years old in the analysis. This study was approved by the Ethics Committee of SYSUCC. All patients signed informed consent during genetic consulting to allow the use of their data in clinical research.

2.2 | Genetic testing

Ethylenediaminetetraacetic acid (EDTA)-anticoagulated peripheral blood samples were collected after diagnosis and before treatment. Genetic testing, which included 14 genes (namely, *MLH1*, *MLH3*, *MSH2*, *MSH6*, *PMS1*, *PMS2*, *APC*, *AXIN2*, *STK11*, *EPCAM*, *PTEN*, *SMAD4*, *MUTYH*, and *BMPR1A*) was carried out at the Molecular Diagnostics Department of SYSUCC using the Illumina HiSeq 2000 (Illumina, San Diego, CA, USA; target sequence capture and next-generation sequencing). Variants were classified per the Ambry five-tier variant classification protocol [pathogenic mutation; variant, likely pathogenic; variant of unknown significance (VUS); variant, likely benign; and benign], which is based on guidelines published by the American College of Medical Genetics and Genomics, the Association for Molecular Pathology, and the International Agency for Research on Cancer [26,27]. VUS was classified if evidence was insufficient to support either a pathogenic or a benign interpretation. Samples from

individuals who carried pathogenic or likely pathogenic alterations were further used for germline analysis.

2.3 | Grouping

The young and old groups were categorized by the cut-off age of 50 years, which was recommended as the optimal age to initiate colonoscopy screening [28], and then divided into subgroups bases on genetic testing results. For the mutation group, the patients were included if at least one of the tested genes were mutated. Patients without any mutations were categorized into the non-mutation group.

2.4 | Data collection

Clinical data were retrieved from the medical record archive of SYSUCC. We collected the following information about the patients: age, gender, smoking and drinking history (excluding occasional smoking and drinking in social engagements), family history, pathological grade, TNM stage (the 7th edition of the American Joint Committee on Cancer TNM staging system), overall survival [OS, defined as the duration from the completion of first treatment (e.g., chemotherapy, chemoradiotherapy, and surgery) till the last follow-up by October 18, 2018], tumor location (ascertained by an intraoperative probe), MMR protein expression [determined using immunohistochemistry (IHC)], genetic testing results, and so on.

To determine differences in the distributions of MMR genes between different ethnicities, we also collected datasets from previous studies.

2.5 | Statistical analysis

Statistical analyses were carried out using the R programming language (version R 3.5.1, https://www.R-project.org). The Chi-square (χ^2) test and Fisher's exact test were used to compare differences in categorical parameters between different groups. All P values were two-sided, and P values of < 0.05 were considered statistically significant. A gene mutation profile was presented using the Complex-Heatmap [29] R package, and the mutation fraction was plotted using the 'ggplot2' R package.

Gene mutation exclusive analysis was performed using the "maftools" package [30]. The "*somaticInteractions*()" function of the package can detect mutually exclusive or co-occurring sets of genes, which performs pairwise Fisher's exact test to detect significant pairs of genes to calculate P values. The "*oncostrip*()" function can depict the distribution of the mutually exclusive genes in the cohort.

3 | RESULTS

3.1 | Demographic and clinicopathological characteristics

A total of 526 patients with pathologically confirmed CRC who underwent genetic testing were included. Of the 526 patients, 313 (59.5%) were men, and 261 (49.6%) were under 50 years old (Table 1). The median age at first CRC diagnosis was 50 years, ranging from 15 to 70 years. One hundred twenty-three (23.4%) patients had a smoking history, 42 (8.0%) had a drinking history, and 190 (36.1%) had a family history of cancer. Thirty-one (5.9%) patients had personal history of LS-related cancers. Highly/moderately differentiated adenocarcinoma accounted for 338 (64.3%) patients. The proportions of patients at stage I/II ($n = 269$, 51.1%) and at III/IV ($n = 248$, 47.1%) were almost equivalent. The majority ($n = 352$, 66.9%) of patients had the primary tumor on the left side of the colon (the splenic flexure, descending colon, sigmoid, and rectum), and 167 (31.7%) on the right side (the cecum, ascending colon, hepatic flexure, and transverse colon). As of October 18, 2018, 488 (92.8%) patients were alive. Only 140 (26.6%) patients had deficient MMR status.

We divided the cohort into two groups by the age 50 to explore the clinical differences between old and young patients [28]. The sample sizes of patients in the two groups were nearly equal (261 vs. 265). As shown in Table 1, the two groups had significant differences in most clinical characteristics. The old group had significantly higher frequencies of smoking ($P = 0.020$) and drinking history ($P = 0.009$), personal history of LS-related cancer ($P < 0.001$), left-side primary cancer ($P < 0.001$), and well-differentiated cancer ($P = 0.014$) than the young group. No significant differences were found in family history, polyps, TNM stage, OS, and MMR status.

Then, we assessed the differences between mutation carriers and non-carriers (Table 2). The median age of mutation carriers was 50 years (range, 15 to 70 years), and that of non-carriers was 45 years (range, 19 to 66 years). We found that patients with mutations had significantly higher frequencies of personal history of cancer ($P < 0.001$), family history of cancer ($P < 0.001$), polyps ($P = 0.006$), low TNM stage ($P = 0.035$), and right-side primary cancer ($P < 0.001$) than patients with no detected gene mutations.

For 417 patients with complete MMR IHC staining results, we calculated the sensitivity of IHC by pairing lost proteins and related genes. Six patterns of MMR IHC were identified: *MLH1* loss alone or with *PMS2* loss (pattern 1, $n = 71$), *MSH2* loss alone or with *MSH6* loss (pattern 2, $n = 43$), *PMS2* loss alone (pattern 3, $n = 14$), *MSH6* loss alone (pattern 4, $n = 3$), loss of all four MMR proteins (pattern 5, $n = 0$), and unpaired loss (i.e., *MSH2*/*PMS2* or three MMR genes; pattern 6, $n = 9$). For patterns 1 to 4, the sensitivity and specificity of MSI testing were as follows: 75.0% and 87.9% for pattern 1, 86.9% and 94.2% for pattern 2, 75.0% and 97.3% for pattern 3, 33.3% and 99.5% for pattern 4.

3.2 | Mutational profiles of patients with CRC under 70 years

Out of the 526 patients, 77 (14.6%) had at least one mutation in one of the 12 genes listed in Figure 1A, covering seven diseases related to hereditary CRC. In total, 77 mutations were identified in 77 patients consisting exclusively of pathogenic mutations and likely pathogenic variants. Among the 526 patients, *MLH1* was the most frequently mutated ($n = 34$, 6.5%), followed by *MSH2* ($n = 27$, 5.1%), *MSH6* ($n = 5$, 1.0%), and *PMS2* ($n = 4$, 0.8%). These four genes are commonly associated with LS. Mutations in *EPCAM*, *PTEN*, *BMPR1A*, and *SMAD4* were not detected in our cohort. Apart from LS-related genes, *APC* mutations were detected in 4 (0.8%) patients, and all mutations were associated with non-Lynch diseases accounted for 1.4% of all patients. Interestingly, no individuals with two or more mutations or likely pathogenic variants were found in our cohort. In addition, frameshift and nonsense mutations were the most common alterations (both with a frequency of 37.3%), followed by missense mutations (13.0%), splicing mutations (6.5%), and deletion mutations (5.2%). Among all 12 genes, only the *MLH1* gene covered five mutational types.

A total of 262 VUSs were detected in 191 (36.3%) of the 526 patients (Supplementary Figure S1). The genes with high VUS frequencies were *MUTYH* ($n = 50$), *MSH6* ($n = 41$), and *MLH1* ($n = 32$). Among all 262 VUSs, 115 (43.9%) occurred in one of the MMR genes associated with LS. Sixteen patients with deficient MMR carried VUS in MMR genes that corresponded to their abnormal MMR protein levels, suggesting that these mutations may display potential pathogenicity (Supplementary Table S1).

The top five mutated genes in the old group were *MLH1*, *MSH2*, *MSH6*, *PMS2*, and *AXIN2*, while in the young group, the order was *MLH1*, *MSH2*, *PMS2*, *MSH6*, and *APC*; notably, the frequency of *APC* mutations was much higher in the young group than in the old group (1.5% vs. 0%, Figure 1B and 1C). Nonsense mutation was the most common mutation in the young group (40.4%), whereas frameshift mutation was the most common mutation in the old group (43.3%). Nevertheless, there was no significant difference in the distribution of mutation types between the two groups ($P = 0.247$, Figure 2A). The young

TABLE 1 Clinical characteristics of the 526 CRC Chinese patients who underwent hereditary genetic testing

Characteristic	Whole cohort [cases (%)]	Young group [cases (%)]	Old group [cases (%)]	P-value
Total	526	261	265	
Age				
< 50 years	261 (49.6)	NA	NA	
≥50 years	265 (50.4)	NA	NA	
Gender				0.033
Male	313 (59.5)	143 (54.8)	170 (64.2)	
Female	213 (40.5)	118 (45.2)	95 (35.8)	
Smoking history				0.020
Yes	123 (23.4)	49 (18.8)	74 (27.9)	
No	325 (61.8)	170 (65.1)	155 (58.5)	
Unknown	78 (14.8)	42 (16.1)	36 (13.6)	
Drinking history				0.009
Yes	42 (8.0)	12 (4.6)	30 (11.3)	
No	403 (76.6)	205 (78.5)	198 (74.7)	
Unknown	81 (15.4)	44 (16.9)	37 (14.0)	
Family history of cancer				0.103
CRC	78 (14.8)	41 (15.7)	37 (14.0)	
non-CRC cancer	82 (15.6)	37 (14.2)	45 (17.0)	
CRC and non-CRC cancer	30 (5.7)	9 (3.4)	21 (7.9)	
No/unknown history	336 (63.9)	174 (66.7)	162 (61.1)	
Personal history of cancer				<0.001
Multiple CRC	23 (4.4)	7 (2.7)	16 (6.0)	
Endometrial cancer	6 (1.1)	0 (0)	6 (2.3)	
Ovarian cancer	2 (0.4)	2 (0.8)	0 (0)	
Other cancer	19 (3.6)	2 (0.8)	17 (6.4)	
No/unknown history	476 (90.5)	250 (95.8)	226 (85.3)	
Colorectal polyps				0.140
Present	94 (17.9)	40 (15.3)	54 (20.4)	
Absent/unknown	432 (82.1)	221 (84.7)	211 (79.6)	
Tumor differentiation				0.014
High/moderate	338 (64.3)	154 (59.0)	184 (69.4)	
Poor	188 (35.7)	107 (41.0)	81 (30.6)	
TNM stage				0.053
I/II	269(51.1)	122 (46.7)	147 (55.5)	
III/IV	248(47.1)	134 (51.3)	114 (43.0)	
Unknown	9 (1.7)	5 (1.9)	4 (1.5)	
Primary tumor location				<0.001
Left side of the colon	352 (66.9)	161 (61.7)	191 (72.1)	
Right side of the colon	167 (31.8)	100 (38.3)	67 (25.3)	
Both sides of the colon	7 (1.3)	0 (0)	7 (2.6)	
Survival status				0.720
Survival	488 (92.8)	242 (92.7)	246 (92.8)	
Dead	33 (6.3)	15 (5.7)	18 (6.8)	
Missing	5 (1.0)	4 (1.5)	1 (0.4)	

(Continues)

TABLE 1 (Continued)

Characteristic	Whole cohort [cases (%)]	Young group [cases (%)]	Old group [cases (%)]	*P*-value
MMR status				0.148
Proficient	281 (53.4)	129 (49.4)	152 (57.4)	
Deficient	140 (26.6)	75 (28.7)	65 (24.5)	
Unknown	105 (20.0)	57 (21.8)	48 (18.1)	

Abbreviations: CRC, colorectal cancer; NA, not applicable; MMR, mismatch repair.

group had a higher gene mutation rate than the old group (18.0% vs. 11.3%, $P = 0.036$, Figure 2B).

3.3 | Comparison of mutation frequency between the present Chinese cohort and Western cohorts

To compare the mutation prevalence between Chinese and Western cohorts, we collected mutation data of 3411 patients from six published studies [31–36] (Supplementary Table S2), which were all diagnosed with inherited or family-related colorectal cancers. With the mutation statuses of four genes (*MLH1*, *MSH2*, *MSH6*, and *PMS2*) available in all six datasets, we performed a comparison of MMR gene mutation distributions (two-tailed Fisher's exact test) between our study subjects and those from published studies. As shown in Figure 3, the mutation fractions of the four genes were largely different between our study subjects and those reported by Bonadona et al. [31], Møller et al. [33], Moreira et al. [34], and Sjursen et al. [35] (all $P < 0.01$). In contrast to the majority of Western cohorts, patients with *MSH6* and *PMS2* mutations, which play less important roles in the hereditary risks of LS, accounted for a higher proportion (38.5%) in our Chinese dataset.

3.4 | Associations between molecular features and clinicopathological characteristics

Correlation analyses were performed to explore the associations between mutated genes and clinicopathological features of the 526 patients, in which some clinicopathological features (personal history of cancer, family history of cancer, and TNM stage) were reclassified into two categories (Figure 4). Among all patients, those with *MLH1* or *MSH2* mutations had higher frequencies of personal history of cancer (*MLH1*: 20.6% vs. 8.7%; *MSH2*: 25.9% vs. 8.6%), family history of cancer (*MLH1*: 73.5% vs. 48.4%; *MSH2*: 70.4% vs. 48.9%), and right side primary cancer (*MLH1*: 73.5% vs. 29.3%; *MSH2*: 56.0% vs. 31.0%) than

patients without these gene mutations. The proportion of stage I/II disease was higher in patients with *MLH1* mutations than in those without *MLH1* mutations (70.6% vs. 50.7%), and the rate of polyps was higher in patients with *APC* mutations than in those with wild-type *APC* (75.0% vs. 17.4%). Detailed reclassified data are provided in Supplementary Table S3. In addition, we explored the prognostic values of the mutations, but none of them had a significant ability to predict OS (Supplementary Figure S2). By the way, we only kept *MLH1* and *MSH2* in further analyses, excluding other genes, because patients with corresponding mutations were too low to support the survival analysis.

3.5 | Gene exclusive analysis

Many disease-causing genes in cancer showed strong exclusiveness or co-occurrences in their mutation patterns. We analyzed the distribution of these mutations in the 77 patients who had at least one mutation in one of the 12 genes. A set of genes, including *MLH1, MSH2*, and *MSH6*, were found to be mutually exclusive in CRC ($P < 0.01$, Figure 5A). No patients harbored coincident mutations in this set of genes. This set of genes as a group was also confirmed to be mutually exclusive in mutation profiles ($P < 0.01$, Figure 5B).

4 | DISCUSSION

Studies on hereditary CRC in China are limited. The present study provides a comprehensive description of the hereditary genetic risks of CRC among Chinese patients. In the current study, the proportion of hereditary CRC in CRC patients under 70 years old was 14.6% (77 of 526). Patients with *MLH1* or *MSH2* mutations had higher frequencies of personal/family history of cancer and right side primary cancer than patients without these gene mutations. Different from Western patients, Chinese patients had a higher rate of *MSH6* and *PMS2* mutations (38.5%). Mutations in *MLH1, MSH2*, and *MSH6* were found to be mutually exclusive.

TABLE 2 Demographic and clinicopathological characteristics between mutation carriers and non-carriers

| Characteristic	Mutation non-carriers [cases (%)]	Mutation carriers [cases (%)]	P-value
Total	449	77	
Age group			0.036
< 50 years	214 (47.7)	47 (61.0)	
≥50 years	235 (52.3)	30 (39.0)	
Gender			0.617
Male	265 (59.0)	48 (62.3)	
Female	184 (41.0)	29 (37.7)	
Smoking history			0.542
Yes	104 (23.2)	19 (24.7)	
No	282 (62.8)	43 (55.8)	
Unknown	63 (14.0)	15 (19.5)	
Drinking history			0.641
Yes	35 (7.8)	7 (9.1)	
No	347 (77.3)	56 (72.7)	
Unknown	67 (14.9)	14 (18.2)	
Personal history of cancer			<0.001
Multiple CRC	15 (3.3)	8 (10.4)	
Endometrial cancer	1 (0.2)	5 (6.5)	
Ovarian cancer	1 (0.2)	1 (1.3)	
Other cancer	17 (3.8)	2 (2.6)	
No/unknown history	415 (92.4)	61 (79.2)	
Family history of cancer			<0.001
CRC	51 (11.4)	27 (35.1)	
Non-CRC cancer	77 (17.1)	5 (6.5)	
CRC and non-CRC cancer	24 (5.3)	6 (7.8)	
No/unknown history	297 (66.1)	39 (50.6)	
Colorectal polyps			0.006
Present	71 (15.8)	23 (29.9)	
Absent/unknown	378 (84.2)	54 (70.1)	
Tumor differentiation			0.122
High/moderate	295 (65.7)	43 (55.8)	
Poor	154 (34.3)	34 (44.2)	
TNM stage			0.035
I/II	220 (49.0)	49 (63.6)	
III/IV	220 (49.0)	28 (36.4)	
Unknown	9 (2.0)	0 (0)	
Primary tumor location			<0.001
Left side of the colon	321 (71.5)	31 (40.3)	
Right side of the colon	123 (27.4)	44 (57.1)	
Both sides of the colon	5 (1.1)	2 (2.6)	
Survival status			0.070
Survival	413 (92.0)	75 (97.4)	
Dead	32 (7.1)	1 (1.3)	
Missing	4 (0.9)	1 (1.3)	

Abbreviations: CRC, colorectal cancer.
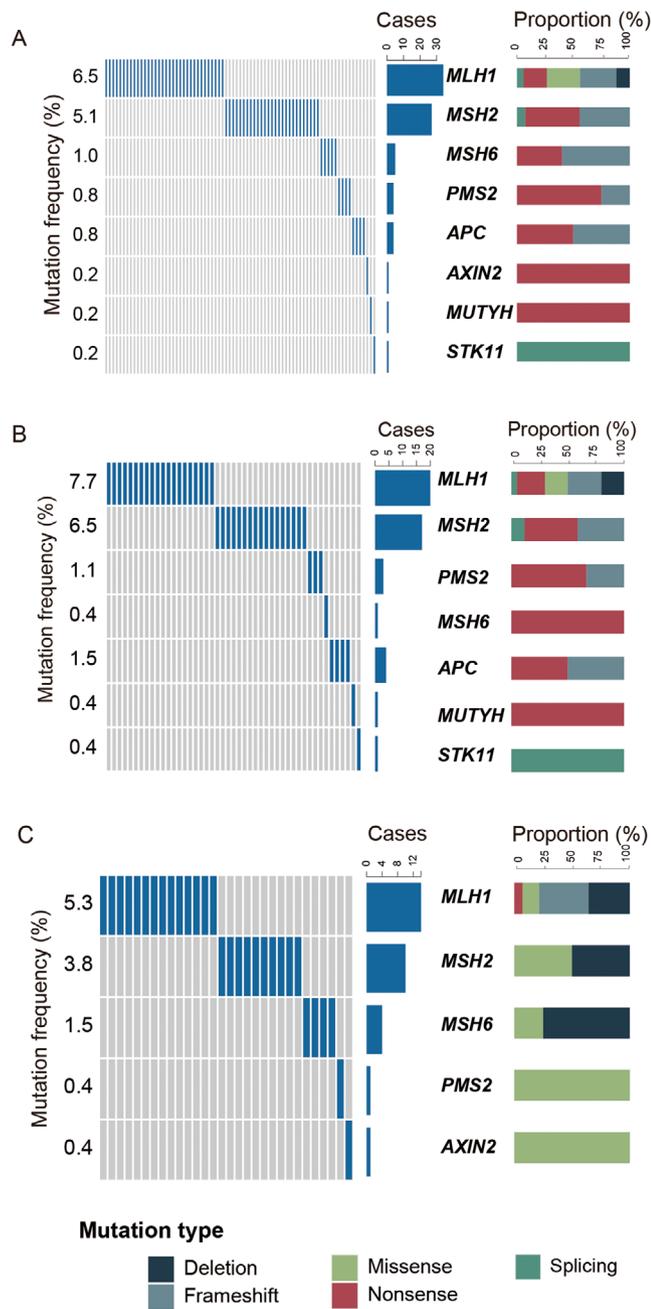
FIGURE 1 Repertoire of germline genetic alterations of colorectal cancer in the present cohort and mutation profiles in different age groups. A. Eight hereditary genes were mutated in 77 of the 526 patients. Stacked bar charts (right) indicate the mutation types for each gene. B. Young group (age < 50 years, *n* = 261): Seven hereditary genes were mutated in 47 patients. C. Old group (age ≥ 50 years, *n* = 265): Five hereditary genes were mutated in 30 patients. Stacked bar charts (right) indicate the mutation types for each gene. Abbreviations: *MLH1*, MutL homolog 1; *MSH2*, MutS homolog 2; *MSH6*, MutS homolog 6; *PMS2*, PMS1 homolog 2; *APC*, APC regulator of WNT signaling pathway; *AXIN2*, axin 2; *MUTYH*, mutY DNA glycosylase; *STK11*, serine/threonine kinase 11
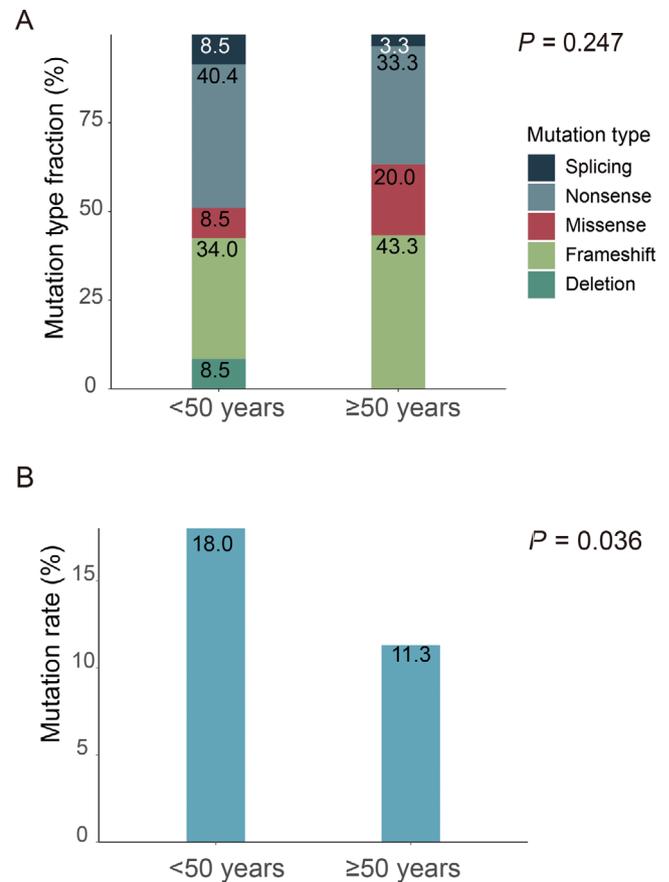
FIGURE 2 Comparison of mutation types. Bar plots show mutation type fraction (A) and mutation rate (B) in the young and old groups.

The mutation landscape in our cohort represented a particular pattern in which all patients harbored sole mutation. The mutation rate of non-MMR genes was 1.3% (7 out of 526), which displays the advantage of comprehensive multi-gene panel over the confined panel in detecting CRC hereditary gene mutations. More potential patients and their families harboring germline mutations would be found and consequently seek professional counseling by genetic experts to reduce future cancer risks. Patients in the young group had a significantly higher frequency of gene mutations than those in the old group (18.0% vs. 11.3%, $P < 0.05$), which is consistent with clinical observations. These data indicate the necessity to perform genetic testing in young patients to identify hereditary risks.

The molecular genetic background of many inherited CRC syndromes has been clarified [5,7,8]. LS is a well-described hereditary cancer syndrome caused by germline mutations in DNA MMR genes (*MLH1*, *MSH2*, *MSH6*, and *PMS2*) or by a deletion in the *EPCAM* gene, which regulates methylation of the *MSH2* promoter [14,37,38]. We found that 13.3% of patients in the current study harbored LS-associated gene mutations.
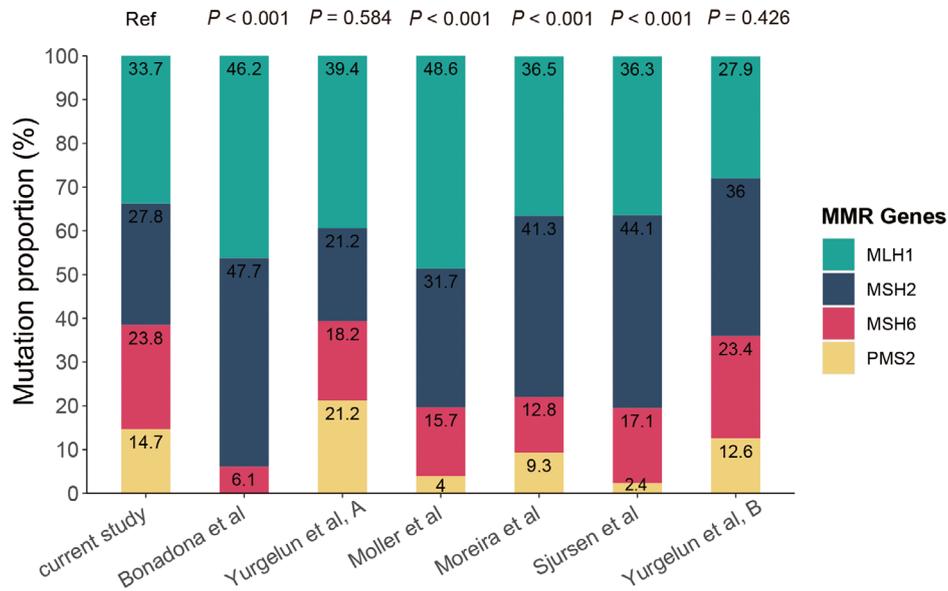
FIGURE 3 MMR gene mutation distributions compared with prior studies. Fisher's exact test was used to compare mutation frequencies of mismatch repair genes from this study with those from previous studies
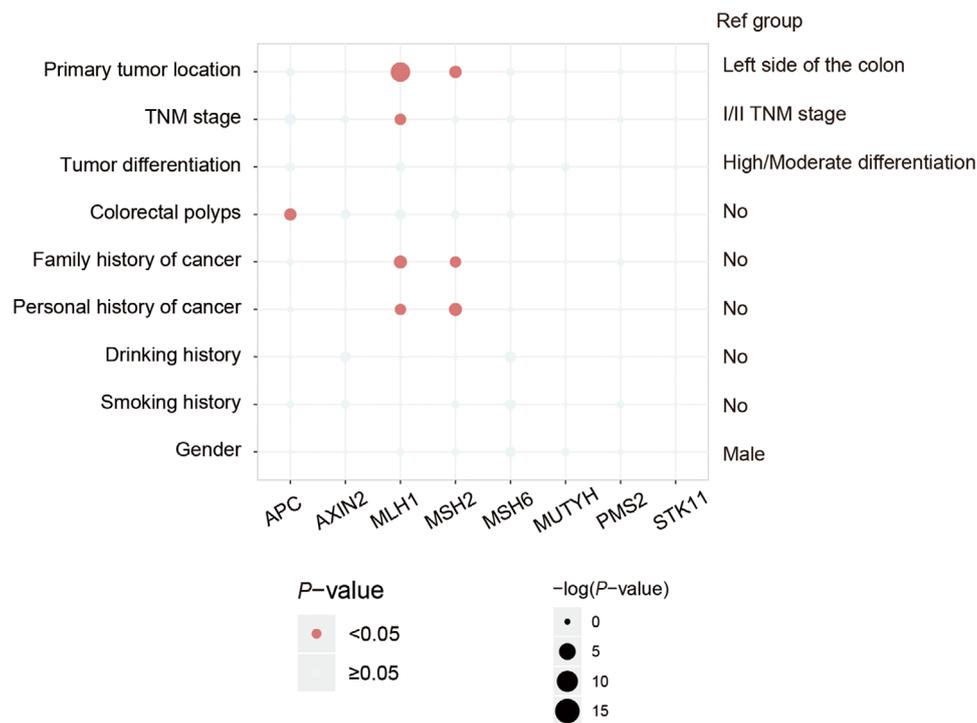

FIGURE 4 Associations between molecular features and clinicopathological characteristics. The analysis for associations between gene mutations and clinicopathological characteristics of CRC patients performed using the Wald chi-square (χ^2) test. All clinical features were reclassified into two categories. In the dot plot, a red dot indicates a significant association between two variables. The bigger the dot is, the more significant the association is

The MMR gene mutation frequency between Chinese and Western patients was not in complete accordance. Our data presented that Chinese patients showed a much higher proportion of mutations of *MSH6* and *PMS2* genes than Western patients. On one hand, this discrepancy may be attributed to racial differences. On the other hand, this phenomenon indicates the importance of the database for annotation of gene mutations that contribute to
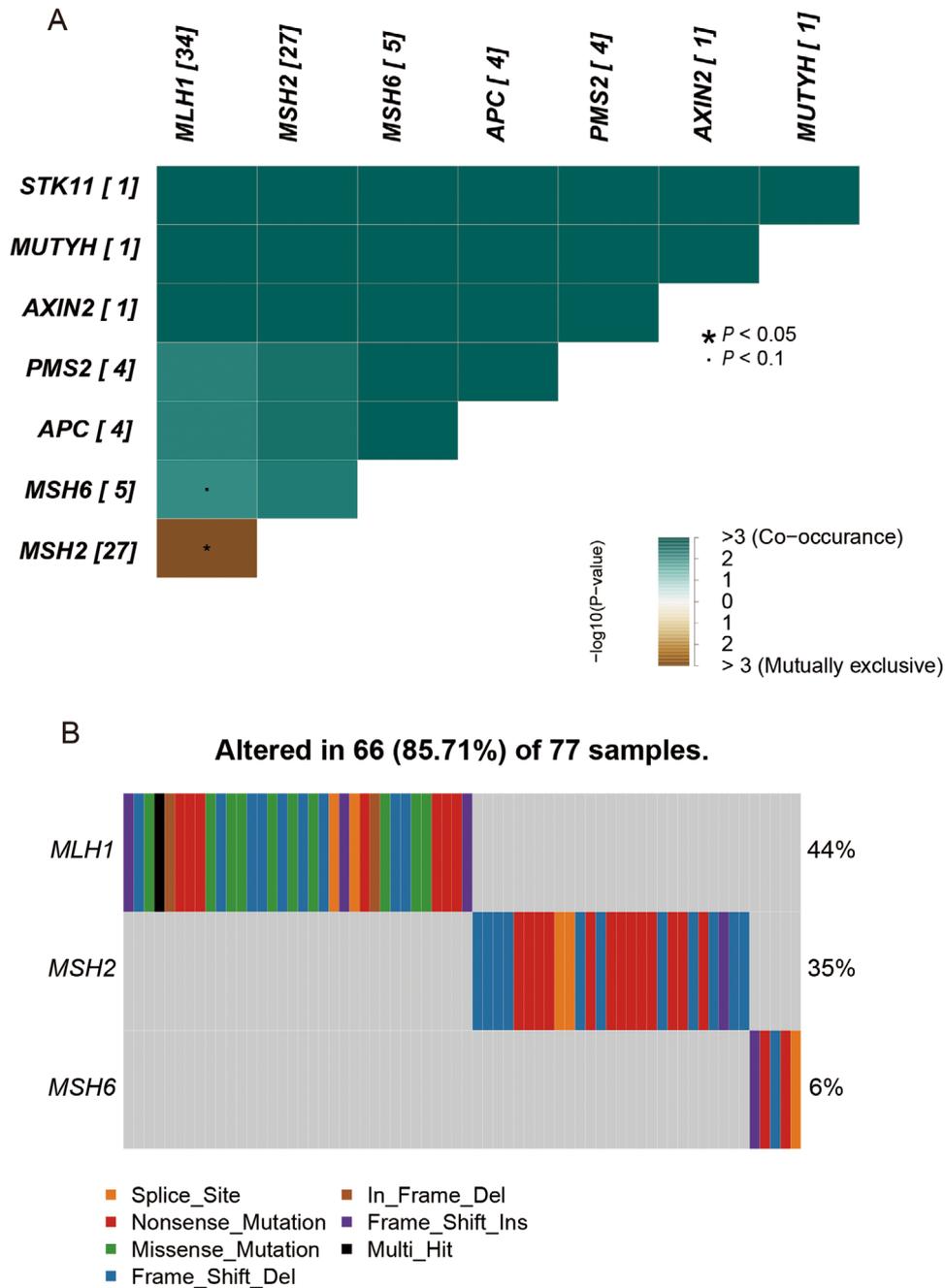
FIGURE 5 Mutual exclusive analysis of the selected genes. A. Mutual exclusive analysis of hereditary CRC genes. The numbers in square brackets following genes indicate the number of patients with corresponding mutant gene. B. Mutation exclusiveness distribution of *MLH1*, *MSH2*, and *MSH6*

corresponding CRC syndrome. With the development and enrichment of a gene mutation database, we may detect *MSH6* and *PMS2* mutations which were not detectable when western studies were published (before 2017). More panel testing data are required to improve the detection rate of pathogenic and likely pathogenic mutations.

Of note, several patients in our cohort did not show related clinical phenotypes for a specific syndrome, but the corresponding mutations were detected. For instance,

2 of 5 FAP/MAP patients lacked the phenotype of diffuse colorectal polyposis, which implies that the presence of polyps is not a perfect indicator for hereditary susceptibility [20]. MMR gene mutations accounted for a large proportion (90.9%) of all gene mutations. Some mutation carriers did not meet strict criteria for assessing hereditary CRC risks (e.g., Amsterdam criteria or Bethesda guidelines), suggesting that such history and phenotypic characteristics may have limited utility in identifying hereditary

mutation carriers. In this circumstance, panel-based testing may be particularly more useful than syndrome-based testing as they could detect more patients with potential cancer risks. A recent study also suggested that a threshold should be discussed at which systematic gene panel testing would be more efficient irrespective of phenotype [32]. Considering cost-effectiveness, the loose standard of gene panel testing should be encouraged with decreasing sequencing costs for CRC patients.

In our cohort, APC mutations were highly associated with the polyp history, and MLH1 and MSH2 mutations were highly associated with a personal/family history of cancer. These findings were consistent with clinical observations and previous studies [39,40]. Patients with MLH1 mutations were prone to be within the early TNM stage. A previous study observed better clinical outcome in patients with MLH1- or MSH2-negative CRC at stage II or III compared with MLH1- or MSH2-positive patients [41], suggesting that MLH1/MSH2 mutations may provide useful prognostic information for the management of stage II and III CRC patients.

The alterations of mutually exclusive genes that affect the same pathway tend not to co-occur in the same patients. MLH1 is the most important susceptibility gene for LS. The role of MLH1 in MMR has often been described as that of a "molecular matchmaker", which involves coupling mismatch recognition with downstream steps of MMR. hMutSa, the predominant mismatch-binding factor in humans, consists of two proteins, MSH2 and MSH6. MSH2 is the second most frequently detected MMR gene. MSH6 and MSH2 work through forming heterodimers, and MSH6 appears to be the subunit responsible for mismatch recognition in hMutSa complex [42]. MLH1 and MSH2 belong to the same pathway, the "mismatch pathway". No patient had mutations of these three genes simultaneously, suggesting that pathogenic mutation of only one of them could be enough to promote CRC tumorigenesis.

As reported, IHC testing has approximately only 62%–78% sensitivity for identifying tumors with germline mutations [43]. In the present study, IHC sensitivities were high enough for identifying MLH1, MSH2, and PMS1 mutations, but extremely low for MSH6 mutations. The low sensitivity for MSH6 mutation identification may be explained by the majority of MSH6 mutations being VUSs, which may have possible pathogenicity [44].

The major strength of the present study includes the use of a large, consecutive cohort of Chinese patients along with detailed clinicopathological data. These data were used to perform an in-depth and comprehensive examination of their associations with the corresponding gene mutation profile. The use of a clinical laboratory with extensive experience in gene panel testing and professional interpretation of germline cancer susceptibility gene variants allowed for comprehensive and rapid analysis.

We also recognized the limitations in the present study. The major limitation was the clinic-based cohort with some high-risk features recruited from a large academic center, which may limit the generalization to population-based patients. Point mutations of PTEN, BMPR1A, and SMAD4 as well as truncated mutations in EPCAM were not found in our study. One possible reason is the sample size limitation of our cohort. The other reason may be the low mutation rates of these genes in the Chinese population. In addition, although our gene panel was comprised of the majority of high-penetrance cancer susceptibility genes (i.e., MMR genes, APC, MUTYH, STK11, PTEN, BMPR1A, and SMAD4) adopted by the majority of commercially available multigene panels, other emerging cancer susceptibility genes [such as DNA polymerase delta 1, catalytic subunit/ DNA polymerase epsilon, catalytic subunit (POLD1/POLE) [45–47], BRF1 RNA polymerase III transcription initiation factor subunit (BRF1) [48], and FANCD2 and FANCI associated nuclease 1 (FAN1) [49]] also show a promising future in the area of CRC susceptibility.

In conclusion, the present study used a multigene panel testing to display the comprehensive molecular landscape of hereditary CRC susceptibility in the Chinese population and the genotype-phenotype association within hereditary syndromes. Our results also suggest the necessity of multigene panel testing in CRC patients with or without high risks for identifying patients, as well as their relatives, with CRC susceptibility gene mutations for further supervision.

DATA AVAILABILITY STATEMENT

The authenticity of this article has been validated by uploading the key raw data onto the Research Data Deposit public platform (www.researchdata.org.cn), with the approval RDD number as RDDA2020001501.

AUTHOR CONTRIBUTIONS

Conception and design: QZ, Feng W

Acquisition of data: QZ, TJJ, Fang W, YNW

Analysis and interpretation of data (including statistical analysis, biostatistics, computational analysis): QZ, TJJ, Fang W, PRD, ZZP, GC, JYS, JZL, Feng W, RHX.

Writing, review, and/or revision of the manuscript: Feng W, QZ, TJJ, QNW, Fang W.

Study supervision: Feng W, QZ.

All authors read and approved the manuscript for publication.

CONFLICT OF INTEREST

none declared.

REFERENCES

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68(6):394-424.

2. Martin-Morales L, Feldman M, Vershinin Z, Garre P, Caldes T, Levy D. SETD6 dominant negative mutation in familial colorectal cancer type X. Hum Mol Genet. 2017;26(22):4481-93.

3. Short E, Thomas LE, Hurley J, Jose S, Sampson JR. Inherited predisposition to colorectal cancer: towards a more complete picture. J Med Genet. 2015;52(12):791-6.

4. Rustgi AK. The genetics of hereditary colon cancer. Genes Dev. 2007;21(20):2525-38.

5. Jasperson KW, Tuohy TM, Neklason DW, Burt RW. Hereditary and familial colon cancer. Gastroenterology. 2010;138(6):2044-58.

6. Rawla P, Sunkara T, Barsouk AJPdG. Epidemiology of colorectal cancer: Incidence, mortality, survival, and risk factors. Prz Gastroenterol 2019;14(2):89.

7. Lynch HT, de la Chapelle A. Hereditary colorectal cancer. N Engl J Med. 2003;348(10):919-32.

8. Peters U, Bien S, Zubair N. Genetic architecture of colorectal cancer. Gut. 2015;64(10):1623-36.

9. Boland CR. Taking the starch out of hereditary colorectal cancer. Lancet Oncol. 2012;13(12):1179-80.

10. Stoffel EM, Boland CR. Genetics and Genetic Testing in Hereditary Colorectal Cancer. Gastroenterology. 2015;149(5):1191-203 e2.

11. Aaltonen LA. Hereditary intestinal cancer. Semin Cancer Biol. 2000;10(4):289-98.

12. Hampel H, Stephens JA, Pukkala E, Sankila R, Aaltonen LA, Mecklin JP, et al. Cancer risk in hereditary nonpolyposis colorectal cancer syndrome: later age of onset. Gastroenterology. 2005;129(2):415-21.

13. Lynch HT, Lynch PM, Lanspa SJ, Snyder CL, Lynch JF, Boland CR. Review of the Lynch syndrome: history, molecular genetics, screening, differential diagnosis, and medicolegal ramifications. Clin Genet. 2009;76(1):1-18.

14. Houlleberghs H, Goverde A, Lusseveld J, Dekker M, Bruno MJ, Menko FH, et al. Suspected Lynch syndrome associated MSH6 variants: A functional assay to determine their pathogenicity. PLoS Genet. 2017;13(5):e1006765.

15. Liccardo R, De Rosa M, Izzo P, Duraturo FJGr, practice. Novel implications in molecular diagnosis of Lynch syndrome. Gastroenterol Res Pract. 2017;2017:2595098.

16. Ou Jianghua, Rasmussen Merete, Westers Helga, Andersen Sofie D., Jager Paul O., Kooi Krista A., Niessen Renée C., Eggen Bart J. L., Nielsen Finn C., Kleibeuker Jan H., Sijmons Rolf H., Rasmussen Lene J., Hofstra Robert M. W.. Biochemical characterization ofMLH3missense mutations does not reveal an apparent role of MLH3 in Lynch syndrome. *Genes, Chromosomes and Cancer.* 2009;48: 4:340–350. http://doi.org/10.1002/gcc.20644.

17. Silva Felipe Cavalcanti Carneiro da, Valentin Mev Dominguez, Ferreira Fábio de Oliveira, Carraro Dirce Maria, Rossi Benedito Mauro. Mismatch repair genes in Lynch syndrome: a review. *Sao Paulo Medical Journal.* 2009;127: 1:46–51. http://doi.org/10.1590/s1516-31802009000100010.

18. Kanth P, Grimmett J, Champine M, Burt R, Samadder NJ. Hereditary Colorectal Polyposis and Cancer Syndromes: A Primer on Diagnosis and Management. Am J Gastroenterol. 2017;112(10):1509-25.

19. Lammi L, Arte S, Somer M, Jarvinen H, Lahermo P, Thesleff I, et al. Mutations in AXIN2 cause familial tooth agenesis and predispose to colorectal cancer. Am J Hum Genet. 2004;74(5):1043-50.

20. Grover S, Kastrinos F, Steyerberg EW, Cook EF, Dewanwala A, Burbidge LA, et al. Prevalence and Phenotypes of APC and MUTYH Mutations in Patients With Multiple Colorectal Adenomas. JAMA. 2012;308(5):485-92.

21. Woods MO, Younghusband HB, Parfrey PS, Gallinger S, McLaughlin J, Dicks E, et al. The genetic basis of colorectal cancer in a population-based incident cohort with a high rate of familial disease. Gut. 2010;59(10):1369-77.

22. Bertario L, Russo A, Sala P, Varesco L, Giarola M, Mondini P, et al. Multiple approach to the exploration of genotype-phenotype correlations in familial adenomatous polyposis. J Clin Oncol. 2003;21(9):1698-707.

23. Dong Lin, Jin Xianglan, Wang Wenmiao, Ye Qiurong, Li Weihua, Shi Susheng, Guo Lei, Ying Jianming, Zou Shuangmei. Distinct clinical phenotype and genetic testing strategy for Lynch syndrome in China based on a large colorectal cancer cohort. *International Journal of Cancer.* 2020;146: 11:3077–3086. http://doi.org/10.1002/ijc.32914.

24. Vasen HF, Blanco I, Aktan-Collan K, Gopie JP, Alonso A, Aretz S, et al. Revised guidelines for the clinical management of Lynch syndrome (HNPCC): recommendations by a group of European experts. Gut. 2013;62(6):812-23.

25. Ladabaum U, Wang G, Terdiman J, Blanco A, Kuppermann M, Boland CR, et al. Strategies to identify the Lynch syndrome among patients with colorectal cancer: a cost-effectiveness analysis. Ann Intern Med. 2011;155(2):69-79.

26. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med. 2015;17(5):405-24.

27. Plon SE, Eccles DM, Easton D, Foulkes WD, Genuardi M, Greenblatt MS, et al. Sequence variant classification and reporting: recommendations for improving the interpretation of cancer susceptibility genetic test results. Hum Mutat. 2008;29(11):1282-91.

28. Chen C, Stock C, Hoffmeister M, Brenner H. Optimal age for screening colonoscopy: a modeling study. Gastrointest Endosc. 2019;89(5):1017-25.e12.

29. Gu Z, Eils R, Schlesner MJB. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics 2016;32(18):2847-9.

30. Mayakonda A, Lin D-C, Assenov Y, Plass C, Koeffler HPJGr. Maftools: efficient and comprehensive analysis of somatic variants in cancer. Genome Res. 2018;28(11):1747-56.

31. Bonadona V, Bonaïti B, Olschwang S, Grandjouan S, Huiart L, Longy M, et al. Cancer risks associated with germline mutations in MLH1, MSH2, and MSH6 genes in Lynch syndrome. JAMA. 2011;305(22):2304-10.

32. Yurgelun MB, Kulke MH, Fuchs CS, Allen BA, Uno H, Hornick JL, et al. Cancer susceptibility gene mutations in individuals with colorectal cancer. J Clin Oncol. 2017;35(10):1086.

33. Møller P, Seppälä T, Bernstein I, Holinski-Feder E, Sala P, Evans DG, et al. Cancer incidence and survival in Lynch syndrome patients receiving colonoscopic and gynaecological surveillance: first report from the prospective Lynch syndrome database. Gut. 2017;66(3):464-72.

34. Moreira L, Balaguer F, Lindor N, De la Chapelle A, Hampel H, Aaltonen LA, et al. Identification of Lynch syndrome among patients with colorectal cancer. JAMA. 2012;308(15): 1555-65.

35. Sjursen W, McPhillips M, Scott RJ, Talseth-Palmer BAJMg, medicine g. Lynch syndrome mutation spectrum in New South Wales, Australia, including 55 novel mutations. Mol Genet Genomic Med. 2016;4(2):223-31.

36. Yurgelun MB, Allen B, Kaldate RR, Bowles KR, Judkins T, Kaushik P, et al. Identification of a variety of mutations in cancer predisposition genes in patients with suspected Lynch syndrome. Gastroenterology 2015;149(3):604-13. e20.

37. Valle L. Recent Discoveries in the Genetics of Familial Colorectal Cancer and Polyposis. Clin Gastroenterol Hepatol. 2017;15(6):809-19.

38. Mathers JC, Movahedi M, Macrae F, Mecklin JP, Moeslein G, Olschwang S, et al. Long-term effect of resistant starch on cancer risk in carriers of hereditary colorectal cancer: an analysis from the CAPP2 randomised controlled trial. Lancet Oncol. 2012;13(12):1242-9.

39. Dempsey KM, Broaddus R, You YN, Noblin SJ, Mork M, Fellman B, et al. Is it all Lynch syndrome?: An assessment of family history in individuals with mismatch repair-deficient tumors. Genet Med. 2015;17(6):476-84.

40. Michils G, Tejpar S, Thoelen R, van Cutsem E, Vermeesch JR, Fryns JP, et al. Large deletions of the APC gene in 15% of mutation-negative patients with classical polyposis (FAP): a Belgian study. Hum Mutat. 2005;25(2):125-34.

41. Lanza G, Gafà R, Santini A, Maestri I, Guerzoni L, Cavazzini L. Immunohistochemical Test for MLH1 and MSH2 Expression Predicts Clinical Outcome in Stage II and III Colorectal Cancer Patients. J Clin Oncol. 2006;24(15):2359-67.

42. Peltomäki P. Lynch Syndrome Genes. Familial Cancer. 2005;4(3):227-32.

43. Buza N, Ziai J, Hui P. Mismatch repair deficiency testing in clinical practice. Expert Rev Mol Diagn. 2016;16(5):591-604.

44. Okkels Henrik, Lindorff-Larsen Karen, Thorlasius-Ussing Ole, Vyberg Mogens, Lindebjerg Jan, Sunde Lone, Bernstein Inge, Klarskov Louise, Holck Susanne, Krarup Henrik Bygum. MSH6 Mutations are Frequent in Hereditary Nonpolyposis Colorectal Cancer Families With Normal pMSH6 Expression as Detected by Immunohistochemistry. *Applied Immunohistochemistry & Molecular Morphology.* 2012;20: 5:470–477. http://doi.org/10.1097/pai.0b013e318249739b.

45. Bellido F, Pineda M, Aiza G, Valdés-Mas R, Navarro M, Puente DA, et al. POLE and POLD1 mutations in 529 kindred with familial colorectal cancer and/or polyposis: review of reported cases and recommendations for genetic testing and surveillance. Genet Med. 2016;18(4):325.

46. Palles C, Cazier J-B, Howarth KM, Domingo E, Jones AM, Broderick P, et al. Germline mutations affecting the proofreading domains of POLE and POLD1 predispose to colorectal adenomas and carcinomas. Nat Genet. 2013;45(2):136.

47. Buchanan DD, Stewart JR, Clendenning M, Rosty C, Mahmood K, Pope BJ, et al. Risk of colorectal cancer for carriers of a germline mutation in POLE or POLD1. Genet Med. 2018;20(8):890.

48. Bellido F, Sowada N, Mur P, Lazaro C, Pons T, Valdes-Mas R, et al. Association Between Germline Mutations in BRF1, a Subunit of the RNA Polymerase III Transcription Complex, and Hereditary Colorectal Cancer. Gastroenterology. 2018;154(1):181-94 e20.

49. Segui N, Mina LB, Lazaro C, Sanz-Pamplona R, Pons T, Navarro M, et al. Germline Mutations in FAN1 Cause Hereditary Colorectal Cancer by Impairing DNA Repair. Gastroenterology. 2015;149(3):563-6.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.