

# Patterns of germline and somatic mutations in 16 genes associated with mismatch repair function or containing tandem repeat sequences

Shih-Ching Chang<sup>1,2</sup> | Yuan-Tzu Lan<sup>1,2</sup>  | Pei-Ching Lin<sup>3,4</sup> | Shung-Haur Yang<sup>1,5</sup> | Chien-Hsing Lin<sup>6</sup> | Wen-Yi Liang<sup>7</sup> | Wei-Shone Chen<sup>1,2</sup> | Jeng-Kai Jiang<sup>1,2</sup>  | Jen-Kou Lin<sup>1,2</sup>

<sup>1</sup>Division of Colon & Rectal Surgery, Department of Surgery, Taipei Veterans General Hospital, Taipei, Taiwan

<sup>2</sup>Department of Surgery, Faculty of Medicine, School of Medicine, National Yang-Ming University, Taipei, Taiwan

<sup>3</sup>Department of Clinical Pathology, Yang-Ming Branch, Taipei City Hospital, Taipei, Taiwan

<sup>4</sup>Department of Health and Welfare, University of Taipei, Taipei, Taiwan

<sup>5</sup>National Yang-Ming University Hospital, Yilan, Taiwan

<sup>6</sup>Division of Genomic Medicine, National Health Research Institutes, Zhunan, Taiwan

<sup>7</sup>Department of Pathology, Taipei Veterans General Hospital, Taipei, Taiwan

## Correspondence

Yuan-Tzu Lan, Division of Colon & Rectal Surgery, Department of Surgery, Taipei Veterans General Hospital, No. 201, Sec. 2, Shipai Rd, Beitou District, Taipei 11217, Taiwan.

Email: ytlan@vghtpe.gov.tw

## Funding information

Taipei Veterans General Hospital (V105C-043, V106C-027, V107C-004). Ministry of Science and Technology, Taiwan (105-2314-B-075-010-MY2). Taipei City Hospital (10601-62-059, 10701-62-050). The funding body did not play role in the design of the study, and collection, analysis and interpretation of data and in writing the manuscript.

## Abstract

**Background:** We assumed that targeted next-generation sequencing (NGS) of mismatch repair-associated genes could improve the detection of driving mutations in colorectal cancers (CRC) with microsatellite instability (MSI) and microsatellite alterations at selected tetranucleotide repeats (EMAST) and clarify the somatic mutation patterns of CRC subtypes.

**Material and methods:** DNAs from tumors and white blood cells were obtained from 81 patients with EMAST(+)/MSI-high (MSI-H), 78 patients with EMAST(+)/microsatellite stable (MSS), and 72 patients with EMAST(-)/MSI-H. The germline and somatic mutations were analyzed with a 16-genes NGS panel.

**Results:** In total, 284 germline mutations were identified in 161 patients. The most common mutations were in *EPCAM* (24.8%), *MSH6* (24.2%), *MLH1* (21.7%), and *AXIN2* (21.7%). Germline mutations of *AXIN2*, *POLE*, *POLD1*, and *TGFBR2* also resulted in EMAST and MSI. EMAST(+)/MSI-H tumors had a significant higher mutation number ( $205.9 \pm 95.2$  mut/MB) than tumors that were only EMAST(+) or MSI-H ( $118.6 \pm 64.2$  and  $106.2 \pm 54.5$  mut/MB, respectively; both  $P < .001$ ). In patients with *AXIN2* germline mutations, the number of pathological somatic mutations in the tumors was significantly higher than those without *AXIN2* germline mutations ( $176.7 \pm 94.2$  mut/MB vs  $139.6 \pm 85.0$  mut/MB,  $P = .002$ ).

**Conclusion:** Next-generation sequencing could enhance the detection of familial CRC. The somatic mutation burden might result from not only the affected genes in germline mutations but also through the dysfunction of downstream effectors. The *AXIN2* gene might associate with hypermutation in tumors. Further in vitro experiments to confirm the causal relationship is deserved.

## KEYWORDS

colorectal cancer, EMAST, MMR, MSI, mutation

## 1 | INTRODUCTION

Since 2006, colorectal cancer (CRC) has been the most common cancer in Taiwan, with an incidence of 44.8/100 000.<sup>1</sup> In 2015, there were more than 15 000 new CRC patients diagnosed, with an annual incidence of 66.3/100 000.<sup>2</sup> With the increasing incidence of CRC, several reports, including ours, found that approximately 15%-20% of CRC patients have at least one first-degree relative with the disease, and this type of disease is often considered familial.<sup>3-5</sup>

The most common of hereditary CRC syndrome, Lynch syndrome, is associated with deficient mismatch repair (MMR) genes.<sup>6-8</sup> However, no germline mutations in MMR genes have been found in nearly half the patients with Lynch syndrome, as defined based on the Amsterdam criteria.<sup>9-12</sup> Additionally, there are several familial clusters of CRC in which no germline mutation in other genes could be identified.<sup>13-15</sup> Traditionally, the identification of Lynch syndrome has been based on the Bethesda criteria or Amsterdam criteria, followed by microsatellite analysis by genotyping or immunohistochemistry of MMR proteins; based on which MMR protein is not expressed, germline mutations are investigated by Sanger sequencing.<sup>16</sup> In addition to germline mutations of MMR genes, somatic mutations of MMR genes and hypermethylation promoter of *MLH1* have also been implicated as mechanisms underlying microsatellite instability (MSI).<sup>17</sup> Even with extensive molecular analysis, some reports showed that no definite cause including germline mutations, somatic mutations or *MLH1* promoter hypermethylation could be found in nearly one-third of patients with an abnormal MMR status or MSI and a Lynch syndrome-consistent family history.<sup>18,19</sup>

Another variant of MSI, elevated microsatellite alterations at selected tetranucleotide repeats (EMAST), possibly originating from *MSH3* dysfunction or inflammation, have been reported to have a unique phenotype in CRC.<sup>20-23</sup> *MSH3* dysfunction is a complex defect in cancer cells that not only generates EMAST but also may contribute to chromosomal instability and aneuploidy.<sup>20</sup> EMAST(+) cancers share features with MSI+ CRCs<sup>24</sup>; however, evidence has shown conflicting survival results of stage II/III CRC patients after adjuvant 5-FU-based chemotherapy regardless of EMAST status<sup>25</sup> compared to the survival results of patients with MSI-high (MSI-H) tumors.<sup>26</sup> Investigations of the relationships between EMAST and chemotherapy response using targeted gene mutation analysis to gain insight into the mechanisms related to inflammation-induced compartmentalization and inactivation of *MSH3* are ongoing.

The technique of next-generation sequencing (NGS) has improved quickly, and a number of new disease genes, including both rare and common variants, have been identified through this technique.<sup>27-29</sup> The clinical use of NGS is becoming feasible because the associated expense has

dramatically decreased. Several extensive analyses of Lynch-like syndrome by NGS showed the association of MMR deficiency with mutations or epimutations of several other genes, including *POLD1*, *POLE*,<sup>30,31</sup> and *AXINI2*.<sup>32</sup> In addition, a high tumor mutation burden has been reported in CRC with dMMR or MSI.<sup>33</sup> To further investigate the germline and somatic mutation of MMR-associated genes, we enrolled patients with EMAST and patients with MSI, and analyzed the tumor mutation burden and germline mutations with a 16 MMR-associated gene panel using NGS.

## 2 | MATERIALS AND METHODS

### 2.1 | Ethics statement

The study was approved by the Institutional Review Board of Taipei Veterans General Hospital (number 2017-06-004BC) and the data were analyzed anonymously.

### 2.2 | Patient sample collection

We obtained information on 81 patients with EMAST(+) and MSI-H CRC, 78 patients with EMAST(+) and microsatellite stable (MSS) CRC, and 72 patients with EMAST(-) but MSI-H CRC from a database of 1505 CRC patients. Another 60 patients without EMAST(+) or MSI-H CRC were randomly selected from the database. In total, DNA samples from 291 paired tumor and white blood cells samples (WBCs) were obtained from the Biobank of Taipei Veterans General Hospital.

This prospectively collected clinical database consisted of 1505 patients with CRCs who received surgery at the Taipei Veterans General Hospital between 2000 and 2010.<sup>34-36</sup> This database prospectively collected clinical information including age, sex, personal and family medical histories, tumor locations, and CEA levels. The pathological data consisted of TNM stage, differentiation, mucinous histology, and the presence of inflammation in the stroma and lymphovascular invasion. The exclusion criteria were as follows: death within 30 days of surgery, preoperative radiochemotherapy, or emergency operations. The molecular data, including microsatellite status, EMAST status, and 12 genes, including 139 hot spots, were analyzed by MassArray.<sup>34,36</sup> The protocols for analyzing MSI, EMAST, and somatic mutations were the same as those used in previous studies.<sup>24,34,36</sup> The MSI markers consisted of D5S345, D2S123, BAT25, BAT26, and D17S250. The five tetranucleotide markers included D20S82, D20S85, D8S321, D9S242, and MYCL1 to determine the EMAST status. Both MSI-H and EMAST(+) were defined as at least two positive markers out of the aforementioned five markers. The twelve genes analyzed by MassArray were

*APC, FBXW7, TP53, KRAS, NRAS, HRAS, BRAF, PI3KCA, PTEN, AKT1, TGFBR2, and SMAD4.*

## 2.3 | Next-generation sequencing

We used the high-throughput genome sequencer Illumina HiSeq2500 system, to comprehensively explore the DNA sequences of all exons of 16 well-known DNA repair-related genes, namely, *AXIN1, AXIN2, BAX, CTNNA1, EPCAM, EXO1, MLH1, MSH2, MSH3, MSH6, PCNA, PMS1, PMS2, POLD1, POLE, and TGFBR2*, which were selected based on previous studies.<sup>37,38</sup> A total of 100 ng of DNA from each individual was used to construct a sample library using a Roche KAPA Library Preparation Kit (Roche). Each DNA sample was fragmented and used to prepare the DNA library by performing end-repairing, a-overhang addition, adaptor ligation, and size selection (150–350 bp). The target DNA of the DNA repair-related genes was enriched using the probe-based methods. The probes were synthesized by Integrated DNA Technologies (USA) according to our previously designed probe sequences, and the capture procedure was performed following the IDT guidelines. After the probe-based enrichment, a library of each pool was amplified with 14 cycles. The amplified libraries were quantified using a qPCR system and pooled into a new 1.5-mL tube as a 10 nmol/L pooled DNA library. The final pool was used for sequencing (illumina HiSeq2500 sequencer, 2 × 100 bp). The raw output of each individual sample was 250 Mb, and the average depth of the target regions was >1000×. The sequence of each read was trimmed based on the quality score (Q30), and any read less than 45 bp in length was discarded before the following analysis. Reads were aligned to the human hg19 reference genome using BWA-MEM (<http://bio-bwa.sourceforge.net/>), and the GATK Unified Genotyper (GATKLite version 2.3-9) was used for calling variants. After variant calling, we used Variant Effect Predictor ([http://grch37.ensembl.org/Homo\\_sapiens/Tools/VEP](http://grch37.ensembl.org/Homo_sapiens/Tools/VEP)) to annotate the identified variants for the subsequent statistical analysis. The mean read depth was 2404.2 ± 1457.7 and 4119.4 ± 2991.1 in the germline and somatic mutation analyses, respectively.

## 2.4 | Statistical analysis

A chi-square test and two-tailed Fisher's exact test were used to compare the genotype frequencies of the clinicopathological features. Numerical values were compared using Student's *t* test and one-way ANOVA. The data are expressed as the mean ± standard deviation. Statistical significance was defined as *P* < .05. Statistical analyses were performed using SPSS for Windows (version 16.0).

## 3 | RESULTS

### 3.1 | Clinicopathological features of patients with/without germline mutations

As shown in Table 1, patients with germline mutations had a significantly higher frequency of a proximal tumor location and stage I and II disease. The other pathological features were similar between patients with/without germline mutations.

### 3.2 | Patterns and frequency of germline mutations

In 231 patients with either EMAS(+)- or MSI-H-CRC, 284 pathological germline mutations (3 deletion, 18 insertion frameshift and 263 missense variants) were identified in 161 patients. Seventy patients had no germline mutation. Of the 60 patients without EMAS(+)- or MSI-H-CRC, we did not find any germline mutations using this 16-gene panel. Therefore, the panel for 16 DNA repair-related genes panel was estimated to detect pathologic germline mutations in at least 10.6% of the patients (161 of 1505 patients) in this database. The other 70 patients with either EMAS(+)- or MSI-H tumors with clinical suspected Lynch syndrome, but no germline mutations identified through genetic testing were proposed as having Lynch-like syndrome.<sup>39</sup> In addition to germline mutations of MMR genes, Lynch-like syndrome could result from somatic mutations of MMR genes

**TABLE 1** Comparison of clinicopathological features of 231 patients carrying tumors either EMAS(+)- or MSI-H with/without germline mutation

Clinicopathological features	Without GM (n = 70) (%)	With GM (n = 161) (%)	<i>P</i> value
Age	70.8 ± 14.1	67.9 ± 13.1	.456
Male Gender	37 (52.9)	96 (59.6)	.441
Location			
Proximal colon	21 (30.0)	78 (48.4)	.003 <sup>a</sup>
Distal colon	30 (42.9)	35 (21.7)	
Rectum	19 (27.1)	48 (29.8)	
Poor differentiation	9 (12.9)	23 (14.3)	.773
Lymphovascular invasion	11 (15.7)	36 (22.4)	.329
Mucinous Histology	11 (15.7)	24 (14.9)	.875
TNM stage			
I	17 (15.7)	16 (9.9)	.011 <sup>a</sup>
II	23 (32.9)	81 (50.3)	
III	19 (27.1)	46 (28.6)	
IV	11 (15.7)	18 (11.2)	

Abbreviation: GM, germline mutation.

<sup>a</sup>chi-square test with statistically significant difference, *P* < .05.

or methylation of the *MLH1* promoter.<sup>17</sup> We analyzed the somatic mutations of five MMR genes (*MLH1*, *MSH2*, *EPCAM*, *MSH6*, *PMS2*) in the 70 patients with Lynch-like syndrome and found 68 somatic mutations in 34 patients (Supplemental Excel). Our study did not obtain information on *MLH1* promoter methylation, but we found 18 *BRAF*<sup>V600E</sup> mutations (9 had somatic mutations of MMR genes) in 70 patients with Lynch-like syndrome. Because *BRAF* mutations are correlated with *MLH1* methylation,<sup>40,41</sup> we could assume that sporadic MSI tumors could develop through *MLH1* promoter methylation and somatic MMR mutations to induce MSI carcinogenesis.

In 161 patients with CRC with germline mutations, the most commonly mutated genes were *EPCAM* (40, 24.8%), *MSH6* (39, 24.2%), followed by *MLH1* (35, 21.7%), *AXIN2* (35, 21.7%), *POLD1* (28, 17.4%), and *MSH2* (24, 14.9%). Fifteen patients had two mutations in the same gene (5 *AXIN2*, 1 *MLH1*, 6 *MSH2*, 1 *MSH6*, and 2 *POLD1*). No mutation was noted in *PCNA* (Table 2; Figure 1).

### 3.3 | Germline mutations associated with EMAST and MSI genotypes

Studies demonstrated that in patients with Lynch syndrome with dysfunctional MMR genes, 90% had MSI.<sup>8,17</sup> As shown in Table 2, we found that the majority of individual germline gene mutations have different impacts on the EMAST or MSI genotypes. From our results, we could conclude that the majority of EMAST and MSI result from not only MMR dysfunction but also germline mutations of *AXIN2*, *POLE*, *POLD1*, and *TGFBR2*. These findings need to be confirmed in an in vitro study.

In 161 CRC patients with germline mutations, the mutation number ranged from 1 to 4; 81 patients had a single germline mutation, 47 patients had two germline mutations, 23 patients had three germline mutations, and 10 patients had four germline mutations (Table 3). According to the EMAST and MSI tumor classifications, patients with germline mutations had similar frequencies of these three molecular subtypes. EMAST(+) with MSI-H:36.0%, EMAST(+) with MSS: 32.3%, and EMAST(−) with MSI-H: 31.7%,  $P = .856$ . Additionally, the number of germline mutations was not associated with the EMAST or MSI classification (Table 3).

### 3.4 | Patterns and frequencies of somatic mutations

There were 3083 somatic mutations, including 434 5' UTR variants, 540 splice region variants and 290 synonymous variants in the coding region in patients with either EMAST(+) or MSI-H tumor. Among these somatic mutations, 1819 pathological mutations (654 frameshift, 1004 missense, 1 start-loss, and 160 nonsense variants) were found, including 1374 mutations in 161 patients with germline mutations and 445 mutations in 70 patients without germline mutations. After the elimination of the 284 germline mutations, there were 1090 somatic mutations in 161 patients with germline mutations. All total numbers of somatic mutations and numbers of pathological somatic mutations per tumor were similar between patients with and without germline mutations ( $270.7 \pm 125.1$  mut/MB vs  $257.8 \pm 111.4$  mut/MB,  $P = .454$ ;  $146.1 \pm 89.0$  mut/MB vs  $143.2 \pm 83.6$  mut/MB,  $P = .820$ ,

**TABLE 2** Distributions of 16 gene mutations in 161 patients with germline mutation and their associations with EMAST and MSI subtypes

Gene	Mutation no.	Case No. (%) n = 161	EMAST(+) (%) n = 110	MSI-H (%) n = 109
<i>AXIN1</i>	2	2 (1.2)	—	—
<i>AXIN2</i>	40	35 (21.7)	24 (68.6)	26 (74.3)
<i>BAX</i>	1	1 (0.6)	—	—
<i>CTNNB1</i>	1	1 (0.6)	—	—
<i>EPCAM</i>	40	40 (24.8)	24 (60.0)	33 (82.5)
<i>EXO1</i>	3	3 (1.3)	—	—
<i>MLH1</i>	36	35 (21.7)	26 (74.3)	19 (54.3)
<i>MSH2</i>	30	24 (14.9)	19 (79.2)	14 (58.3)
<i>MSH3</i>	9	9 (5.6)	5 (55.6)	5 (66.7)
<i>MSH6</i>	40	39 (24.2)	23 (59.0)	27 (69.1)
<i>PCNA</i>	0	0 (0.0)	—	—
<i>PMS1</i>	5	5 (3.1)	—	—
<i>PMS2</i>	17	17 (10.6)	13 (76.5)	13 (76.5)
<i>POLD1</i>	30	28 (17.4)	19 (67.9)	19 (67.9)
<i>POLE</i>	15	15 (9.3)	7 (46.7)	10 (66.7)
<i>TGFBR2</i>	15	15 (9.3)	10 (66.7)	12 (80.0)





**FIGURE 1** Distributions of 16 MMR-related gene mutations in 161 cases with germline mutation

Table 3). Additionally, tumor with different numbers of germline mutations had similar numbers of somatic mutations ( $P = .902$ , Table 4).

The downstream numbers of somatic mutations were associated with the EMAST-MSI subtypes. Among all patients with or without germline mutations, EMAST(+) and MSI-H tumors had a mean of  $205.9 \pm 95.2$  somatic mut/MB, which was significantly higher than the mean number of somatic mutations in tumors that were only EMAST(+) or only MSI-H ( $118.6 \pm 64.2$ ,  $106.2 \pm 54.5$  mut/MB, respectively; both  $P < .001$ , Table 5). The somatic mutation patterns in EMAST and MSI subtype tumors were similar between patients with and without germline mutations (Table 6). The genes with higher numbers of somatic mutations were *AXIN2*, *BAX*, *MSH3*, *MSH6*, *PMS1*, *POLD1*, *POLE*, and *TGFBR2*. In particular, the *MSH3* mutations were almost located in EMAST(+) and MSI-H tumors. As shown in the Table 5, the somatic mutation frequencies of individual genes, except *PMS1*, were higher in EMAST(+) and MSI-H tumors than in other subtypes of tumors.

If we do not consider the rare mutations in *CTNNB1* (1 patient) and *EXO1* (3 patients), tumors carrying germline mutations of *AXIN2* had the highest somatic mutation number

in this 16-gene panel. The mutation frequencies of the other genes are shown in Table S1. In patients carrying *AXIN2* germline mutations, the total number of somatic mutations and the number of pathological somatic mutations in tumors were  $315.5 \pm 122.4$  and  $176.7 \pm 94.2$  mut/MB, respectively. Both numbers were significantly higher than those in patients without *AXIN2* germline mutations ( $258.2 \pm 119.0$  mut/MB,  $P = .010$ ;  $139.6 \pm 85.0$  mut/MB,  $P = .020$ ), indicating that dysfunction of *AXIN2* could induce the hypermutating status in tumors.

### 3.5 | Identification of functional solitary germline mutations

Of the 81 single-gene mutations in patients with germline mutations, 40 patients had rare single-gene mutation, including two *AXIN1* mutations, 12 *AXIN2* mutations, one *BAX* mutation, one *CTNNB1* mutation, two *MSH3* mutations, two *PMS1* mutations, four *POLD1* mutations, nine *POLE* mutations, and seven *TGFBR2* mutations. In addition to five major MMR genes (*MLH1*, *MSH2*, *EPCAM*, *MSH6*, and *PMS2*) and three additional genes with rare mutations (*AXIN1*, *BAX*,

**TABLE 3** The number of germline mutation was not associated with EMAST and MSI subtypes

No. of GM	Case no.	EMAST(+) MSI-H (%)	EMAST(+) MSS (%)	EMAST(-) MSI-H (%)	P value
0	70	23 (32.9)	24 (34.2)	23 (32.9)	.876
1	81	30 (37.0)	28 (34.6)	23 (28.4)	
2	47	15 (31.9)	15 (31.9)	17 (36.2)	
3	23	7 (30.4)	8 (34.8)	8 (34.8)	
4	10	6 (60.0)	1 (10.0)	3 (30.0)	
≥1	161	58 (36.0)	52 (32.3)	51 (31.7)	.856

Abbreviation: GM, Germline mutation.

**TABLE 4** Numbers of all and pathological somatic mutation per colorectal tumor in 161 patients with different germline mutation numbers

No. of GM	Case no.	No. of All SM (per tumor)	P value	No. of PSM (per tumor)	P value
0	70	257.8 ± 111.4	0.902	143.2 ± 83.6	.982
1	81	269.4 ± 114.7		150.5 ± 83.5	
2	47	264.5 ± 146.4		138.6 ± 99.6	
3	23	281.3 ± 81.3		145.0 ± 71.0	
4	10	287.1 ± 186.8		148.0 ± 124.1	
≥1	161	270.7 ± 125.1	0.454	146.1 ± 89.0	.820

Abbreviations: GM, germline mutation; SM, somatic mutation; PSM, pathological somatic mutation.

**TABLE 5** The numbers of downstream somatic mutation were associated with EMAST-MSI subtypes

	Total	EMAST(+) MSI-H	EMAST(+) MSS	EMAST(-) MSI-H
With GM				
Case no.	161	58	52	51
No. of SM (mut/MB)	146.1 ± 89.1	212.0 ± 100.5	122.3 ± 64.0	106.9 ± 54.5
Without GM				
Case no.	70	23	24	23
No. of SM (mut/MB)	143.2 ± 83.6	195.9 ± 107.6	121.9 ± 53.4	112.4 ± 54.5
All patients				
Case no.	231	81	76	74
No. of SM (mut/MB)	—	205.9 ± 95.2 <sup>a,b</sup>	118.6 ± 64.2 <sup>a</sup>	106.2 ± 54.5 <sup>b</sup>

Abbreviations: GM, germline mutation; SM, somatic mutation.

<sup>a</sup>Statistically significant difference ( $P < .001$ ).

<sup>b</sup>Statistically significant difference ( $P < .001$ ).

and *CTNNB1*), another seven genes were included to identify the associated gene mutations, which are shown in Table S2 and Figure S1. Eleven (31.4%) *AXIN2* mutations, eight (53.3%) *POLE* mutations and six (40%) *TGFBR2* mutations could result in the MSI-H or EMAST(+) phenotypes without any accompany germline mutations of 16 MMR-related genes. The molecular and clinicopathological features of solitary *AXIN2*, solitary *TGFBR2* and solitary *POLE* were demonstrated in Figure S2. The majority of *EXO1*, *MSH3*, *PMS1*, and *POLD1* mutations might need other accompanying MMR-related germline gene mutations to result in the necessary molecular changes, including MSI-H, EMAST(+) or somatic mutations. Therefore, mutations in *AXIN2*, *POLE*, and *TGFBR2* might be driving mutations of familial colon cancer syndrome or there may be other genes involved we did not detect.

### 3.6 | Rare germline mutations

Four patients had rare germline mutations in *AXINI* (2 patients), *BAX* (1 patient), and *CTNNB1* (1 patient). The clinicopathological features are shown in Table S3. These tumors did not have mucinous histology or lymphovascular invasion.

## 4 | DISCUSSION

This study provided several major contributions to the knowledge of familial CRC. First, targeted sequencing by NGS could increase the detection of familial CRC including Lynch syndrome. Second, the somatic mutation burden occurs due to the dysfunction of downstream effectors but not the affected gene with germline mutation. Third, the dysfunction of *AXIN2* might affect both the canonical APC pathway and the hypermutation mechanism. Fourth, we demonstrated the similarities and differences between the clinicopathological features of patients with CRC with and without germline mutations. Fifth, we described the characteristics of patients with rare germline mutations.

Our series showed that the incidence of familial CRC with identified germline mutations was as high as 10%. Of these 161 patients, 100 patients (62.1%) fulfilled the Bethesda criteria, which was a lower percent than was found in another Asian study.<sup>42</sup> The major reason for this lower percent was inadequate or inaccurate of personal and family histories. This limitation has been reported previously.<sup>43,44</sup> Considering the five commonly mutated MMR genes associated with Lynch syndrome, namely, *MLH1*, *MSH2*, *EPCAM*, *MSH6*, and *PMS2*, the incidence of detected Lynch syndrome was higher up to 7.97% (120 in 1505 patients had at least one

**TABLE 6** The somatic mutation patterns of EFAST and MSI subtypes between cases with and without germline mutations

No. of SM (%)	With GM (n = 161)			Without GM (n = 70)		
	EMAST(+) MSI-H (n = 58)	EMAST(+) MSS (n = 52)	EMAST(−) MSI-H (n = 51)	EMAST(+) MSI-H (n = 23)	EMAST(+) MSS (n = 24)	EMAST(−) MSI-H (n = 23)
<i>AXIN1</i>	21 (36.2)	2 (3.8)	1 (2.0)	2 (8.7)	1 (4.2)	0 (0.0)
<i>AXIN2</i>	26 (44.8)	6 (11.5)	5 (9.8)	9 (39.1)	6 (25.0)	5 (21.7)
<i>BAX</i>	48 (82.8)	19 (36.5)	16 (31.4)	18 (78.3)	12 (50.0)	9 (39.1)
<i>CTNNB1</i>	2 (3.4)	4 (7.7)	3 (5.9)	3 (13.0)	0 (0.0)	0 (0.0)
<i>EPCAM</i>	7 (12.1)	6 (11.5)	2 (3.9)	9 (39.1)	2 (8.3)	1 (4.3)
<i>EXO1</i>	17 (29.3)	2 (3.8)	2 (3.9)	4 (17.4)	2 (8.3)	3 (13.0)
<i>MLH1</i>	16 (27.6)	5 (9.6)	5 (9.8)	4 (17.4)	4 (16.7)	2 (8.7)
<i>MSH2</i>	15 (25.9)	8 (15.4)	3 (5.9)	2 (8.7)	1 (4.2)	6 (26.1)
<i>MSH3</i>	55 (94.8)	2 (3.8)	4 (7.8)	19 (82.6)	2 (8.3)	3 (13.0)
<i>MSH6</i>	35 (60.3)	7 (13.5)	14 (27.5)	19 (82.6)	6 (25.0)	4 (17.4)
<i>PCNA</i>	1 (1.7)	1 (1.9)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
<i>PMS1</i>	91 (156.9)	102 (196.2)	100 (196.1)	51 (221.7)	42 (175.0)	43 (187.0)
<i>PMS2</i>	16 (27.6)	5 (9.6)	3 (5.9)	4 (17.4)	2 (8.3)	2 (8.7)
<i>POLD1</i>	52 (89.7)	19 (36.5)	10 (19.6)	10 (43.5)	5 (20.8)	6 (26.1)
<i>POLE</i>	64 (110.3)	44 (84.6)	21 (41.2)	14 (60.9)	22 (91.7)	10 (43.5)
<i>TFGBR2</i>	91 (156.9)	54 (103.8)	58 (113.7)	36 (156.5)	25 (104.2)	23 (100.0)

Abbreviations: GM, germline mutation; SM, somatic mutation.

mutation in these five MMR genes). This result was higher than in previous studies (3%-6%), including ours.<sup>4,17,42,45-49</sup> The traditional method of Lynch syndrome screening was the application of the revised Bethesda criteria, detection of MSI or IHC of MMR proteins and then sequencing. Recently, with the advent of NGS, germline testing has progressively moved from mutational analysis of single genes toward a multigene panel analysis that could enhance the detection rate of Lynch syndrome. The other causes of the higher incidence of Lynch syndrome in our series were that we included patients with EFAST phenotypes and target sequences of MMR-associated genes, including a high frequency of *EPCAM* mutations.

Seventy patients were not found to have germline mutations but had MSI or EFAST. We found that 23.1% had *BRAF* mutations and 50.0% (34/68) had somatic MMR mutations. Patients with CRC with somatic MMR mutations also had MSI.<sup>18</sup> As *BRAF* mutations are correlated with *MLH1* methylation,<sup>40,41</sup> we could assume that sporadic MSI tumors could occur through *MLH1* methylation and somatic MMR mutations, resulting in MSI carcinogenesis.

Eight patients had more than two germline mutations and were considered to have constitutional mismatch repair deficiency syndrome (biallelic germline mutations). Apart from an earlier age of onset of CRC ( $32.1 \pm 13.1$ ), the other characteristics of these patients were identical to those of patients with Lynch syndrome. A previous study showed that patients with biallelic germline mutations had often had several

tumors, starting in childhood.<sup>50</sup> We could not confirm this in our patients because we did not have the relevant information in our database.

EFAST result in tetranucleotide instability, and the detection of EFAST and MSI-H was defined as the presence of instability in at least two of the five markers. A previous study demonstrated the association between EFAST and the loss of *MSH3* nuclear expression in CRC.<sup>20,51</sup> Further, an in vitro study provided the evidence that *MSH3* knockdown could increase dinucleotide or tetranucleotide instability.<sup>23</sup> However, our germline analysis demonstrated that only 4.5% (5/110) of EFAST patients had *MSH3* germline mutations, all of which were missense variants. We did not have data on *MSH3* nuclear expression. We could not conclude that these missense variants are associated with the loss of *MSH3* nuclear expression and the induced EFAST phenomenon. Because of the low frequency of *MSH3* germline mutations, we believe that the role of *MSH3* is limited in the EFAST phenomenon, which conflicts with the results of other reports.<sup>20,23,51</sup>

Furthermore, we found that patients with EFAST and MSI had higher somatic mutation rates with/without germline variants. Our result is reasonable because if mononucleotide, dinucleotide, and tetranucleotide frameshift mutations are not corrected, the underlying dysfunction of MMR proteins or other DNA repair-associated proteins is more severe than in situations in which only tetranucleotide frameshifts or mono- or dinucleotide frameshifts are not corrected. In the

patients without germline mutations, the number of *POLE* somatic mutations (60.9%) in EMAST(+)-MSI-H tumors was lower than in EMAST(+)-MSS tumors (91.7%), but higher than in EMAST(-)-MSI-H tumors (43.5%). This result implies that EMAST, but not MSI, trigger *POLE* somatic mutations. For patients with germline mutations, *PMS1* somatic mutations were as high as 156.9% in EMAST(+)-MSI-H tumors, but lower than those in EMAST(+)-MSS (196.2%) and EMAST(-)-MSI-H tumors (196.1%). This result means that the *PMS1* somatic mutation burden does not occur through the EMAST or MSI mechanism but rather through the dysfunction of genes with germline mutations.<sup>52,53</sup> Until now, there was no report linking EMAST and the tumor mutation burden. Our results deserve further study.

Our results, similar to those of other studies,<sup>54,55</sup> demonstrated that the tumors with EMAST or MSI were located predominantly in the proximal colon and had mucinous histology or poor differentiation but were diagnosed at an earlier stage, most commonly in stage II. These specific phenotypes were considered to be the result of MSI that cause a change in downstream genes rather than of the mechanism that induced the MSI. A previous study confirmed that hereditary MSI patients had better outcomes due to earlier stage disease at diagnosis compared to sporadic MSI patients.<sup>56</sup>

For the solitary *AXIN2*, *TGFBR2*, and *POLE* germline mutations, our previous study collected a database of CRC-associated gene mutations. We deciphered the molecular and clinicopathological features of patients with solitary *AXIN2*, *TGFBR2*, and *POLE* germline mutations (Figure S2). In the solitary *AXIN2* germline mutation patients, their mutation loci were all located in the GSK3b,  $\beta$ -catenin and PP2Ac-binding domains. Furthermore, the tumors had several *APC*, *TP53*, and *KRAS* mutations. Additionally, a previous study suggested that wild-type *AXIN2* inhibits the Wnt signaling pathway; however, mutated *AXIN2* dominantly activates Wnt signaling in cancer cells.<sup>57</sup> These results implied that *AXIN2* might be an initiator of the canonical pathway through dysfunction of the  $\beta$ -catenin destruction complex. More interestingly, germline mutations of *AXIN2* had a high tumor mutation burden in our results. As in a previous study, this partly explained the fact that *AXIN2* germline mutation are responsible for hereditary cancer without known molecular causes.<sup>58</sup> Although *AXIN2* explained some causes of MSI and the EMAST phenomenon, the somatic *AXIN2* mutation rate was not very high. Only 24.7% of tumors had somatic *AXIN2* mutations. Because the *AXIN2* gene containing short coding oligonucleotide repeats, somatic mutations in the *AXIN2* gene have been reported in MSI CRC,<sup>59</sup> but the functional consequences of these *AXIN2* changes remain unclear. Therefore, functional studies of the effects of *AXIN2* mutations on the canonical and MMR pathway deserve further study.

Based on a high mutation frequency in coding mononucleotide tracts,<sup>60</sup> *TGFBR2* was considered to be the MSI

CRC target gene.<sup>61</sup> Indeed, our results showed that all 231 patients with EMAST or MSI had somatic mutations in the *TGFBR2* gene (Table 5). A previous study presented the evidence that familial CRC originated partly from germline *TGFBR2* mutations.<sup>62</sup> However, another study did not find germline *TGFBR2* mutation associated with patients with early-onset CRC or HNPCC.<sup>63</sup> We still found 15 patients with *TGFBR2* germline mutations and six patients with solitary *TGFBR2* germline mutations. Three of these six patients with *TGFBR2* germline mutations had the same second-hit mutation in *TGFBR2* (c.382\_383delAA) in their tumors and these patients had similar clinicopathological features, including proximal tumor locations, early TNM stages and the EMAST(+)-MSI-H subtype (Figure S2). In the other three patients, the second-hit mutation was in other loci of *TGFBR2*, and the clinicopathological features were different except for the presence of MSI-H. These six germline mutation loci are all located in the intracellular and protein kinase genes. A recent study showed that cell lines harboring *TGFBR2* mutants failed to respond to exogenous TGF- $\beta$  stimulation, and re-expression of wild-type *TGFBR2* restored canonical TGF- $\beta$  signaling and proliferative inhibition, confirming the mutational loss of TGF- $\beta$  tumor suppressive activity.<sup>64</sup> Therefore, germline *TGFBR2* mutations could be one of the causes of familial CRC without MMR dysfunction.

Patients with/without *POLE* germline mutations had similar number of total and pathological somatic mutations in their tumors. Eight patients had solitary germline *POLE* mutations. Six of those eight patients had a second-hit *POLE* mutation in their tumors. These eight patients did not have unique clinicopathological features, but had mutations in the EGFR signaling pathway, including mutations in *KRAS*, *BRAF*, and *PIK3CA* (6/8) (Figure S2). A recent study showed that hypermutant cancers were enriched for defects in mismatch repair pathway genes *POLE* and *POLD1*.<sup>65</sup> Different *POLE* germline mutations could result in different mutation burdens in tumors and that not all tumors with *POLE* mutations were associated with hypermutation; drive mutations of *POLE* were uncovered outside the exonuclease domain, indicating that other domains may be responsible for proof reading.<sup>65</sup> In our study, only two of eight solitary germline mutations were located in the exonuclease domain. This might explain the similar mutation burdens in tumors in patients with/without *POLE* germline mutations.

There were several advantages to this study. First, we extracted MSI or EMAST patients from a large database with detailed clinicopathological features and molecular signatures. Second, the study design involved somatic and germline mutation analysis. Third, we analyzed 16 MMR-associated genes via NGS with high read depth, looking for both germline and somatic mutations. However, the lack of immunohistochemistry to support the loss or dysfunction of proteins induced by the mutations is the major limitation of this study.



Finally, our results provided evidence that NGS could enhance the detection rate of Lynch syndrome, which was underestimated in the past. The somatic mutation burden might occur through the dysfunction of downstream effectors rather than directly through the affected genes. We also identified several genes with germline mutations that may explain the familiar CRC. The *AXIN2* gene has dual functions in the canonical pathway and is associated with hypermutation in tumors. Our findings deserve to do in vitro experiments to confirm this causal relationship.

### CONFLICT OF INTERESTS

All the authors declare no competing interests, financial or non-financial, in relation to the work described.

### AUTHOR CONTRIBUTIONS

Yuan-Tzu Lan involved in study design and writing – original draft. Shih-Ching Chang involved in study design and writing – review and editing. Pei-Ching Lin and Chien-Hsing Lin involved in molecular analysis including MSI, EMAS, and NGS. Wen-Yi Liang analyzed the pathological slides. Jeng-Kai Jiang, Wei-Shone Chen, Shung-Haur Yang, and Jen-Kou Lin collected clinical data and samples.

### ORCID

Yuan-Tzu Lan  <https://orcid.org/0000-0001-7350-6529>

Jeng-Kai Jiang  <https://orcid.org/0000-0002-8214-4934>

### DATA AVAILABILITY STATEMENT

The datasets generated and/or analyzed during this study are available for noncommercial use from the corresponding author on reasonable request.

### REFERENCES

- Healthy statistics: Cancer Registry Annual Report in Taiwan area. The Executive Yuan, 2009.
- Healthy statistics: Cancer Registry Annual Report in Taiwan Area. The Executive Yuan, 2017.
- Fuchs CS, Giovannucci EL, Colditz GA, Hunter DJ, Speizer FE, Willett WC. A prospective study of family history and the risk of colorectal cancer. *N Engl J Med*. 1994;331:1669-1674.
- Chang SC, Lin PC, Yang SH, Wang HS, Liang WY, Lin JK. Taiwan hospital-based detection of Lynch syndrome distinguishes 2 types of microsatellite instabilities in colorectal cancers. *Surgery*. 2010;147:720-728.
- Bass AJ, Meyerhardt JA, Chan JA, Giovannucci EL, Fuchs CS. Family history and survival after colorectal cancer diagnosis. *Cancer*. 2008;112:1222-1229.
- Lynch HT, Lanspa S, Smyrk T, Boman B, Watson P, Lynch J. Hereditary nonpolyposis colorectal cancer (Lynch syndromes I & II). Genetics, pathology, natural history, and cancer control, Part I. *Cancer Genet Cytogenet*. 1991;53:143-160.
- Burt RW, DiSario JA, Cannon-Albright L. Genetics of colon cancer: impact of inheritance on colon cancer risk. *Annu Rev Med*. 1995;46:371-379.
- Boland PM, Yurgelun MB, Boland CR. Recent progress in Lynch syndrome and other familial colorectal cancer syndromes. *CA Cancer J Clin*. 2018;68:217-231.
- Shiovitz S, Copeland WK, Passarelli MN, et al. Characterisation of familial colorectal cancer Type X, Lynch syndrome, and non-familial colorectal cancer. *Br J Cancer*. 2014;111:598-602.
- Valle L. Genetic predisposition to colorectal cancer: where we stand and future perspectives. *World J Gastroenterol*. 2014;20:9828-9849.
- Lindor NM, Rabe K, Petersen GM, et al. Lower cancer incidence in Amsterdam-I criteria families without mismatch repair deficiency: familial colorectal cancer type X. *JAMA*. 2005;293:1979-1985.
- Yamaguchi T, Furukawa Y, Nakamura Y, et al. Comparison of clinical features between suspected familial colorectal cancer type X and Lynch syndrome in Japanese patients with colorectal cancer: a cross-sectional study conducted by the Japanese Society for Cancer of the Colon and Rectum. *Jpn J Clin Oncol*. 2015;45:153-159.
- Win AK, Jenkins MA, Dowty JG, et al. Prevalence and penetrance of major genes and polygenes for colorectal cancer. *Cancer Epidemiol Biomarkers Prev*. 2017;26:404-412.
- Chubb D, Broderick P, Dobbins SE, et al. Rare disruptive mutations and their contribution to the heritable risk of colorectal cancer. *Nat Commun*. 2016;7:11883.
- Dai J, Shen W, Wen W, et al. Estimation of heritability for nine common cancers using data from genome-wide association studies in Chinese population. *Int J Cancer*. 2017;140:329-336.
- Furtado LV, Samowitz WS. Colorectal cancer molecular profiling: from IHC to NGS in search of optimal algorithm. *Virchows Arch*. 2017;471:235-242.
- Hampel H, Frankel WL, Martin E, et al. Screening for the Lynch syndrome (hereditary nonpolyposis colorectal cancer). *N Engl J Med*. 2005;352:1851-1860.
- Geurts-Giele WRR, Leenen CHM, Dubbink HJ, et al. Somatic aberrations of mismatch repair genes as a cause of microsatellite-unstable cancers. *J Pathol*. 2014;234:548-559.
- Carethers JM, Stoffel EM. Lynch syndrome and Lynch syndrome mimics: The growing complex landscape of hereditary colon cancer. *World J Gastroenterol*. 2015;21:9253-9261.
- Carethers JM, Koi M, Tseng-Rogenski SS. EMAS is a form of microsatellite instability that is initiated by inflammation and modulates colorectal cancer progression. *Genes (Basel)*. 2015;6:185-205.
- Tseng-Rogenski SS, Chung H, Wilk MB, Zhang S, Iwaizumi M, Carethers JM. Oxidative stress induces nuclear-to-cytosol shift of hMSH3, a potential mechanism for EMAS in colorectal cancer cells. *PLoS ONE*. 2012;7:e50616.
- Campregher C, Schmid G, Ferk F, et al. MSH3-deficiency initiates EMAS without oncogenic transformation of human colon epithelial cells. *PLoS ONE*. 2012;7:e50541.
- Haugen AC, Goel A, Yamada K, et al. Genetic instability caused by loss of MutS homologue 3 in human colorectal cancer. *Cancer Res*. 2008;68:8465-8472.
- Lee HS, Park KU, Kim DW, et al. Elevated microsatellite alterations at selected tetranucleotide repeats (EMAS) and microsatellite instability in patients with colorectal cancer and its clinical features. *Curr Mol Med*. 2016;16:829-839.
- Hamaya Y, Guarinos C, Tseng-Rogenski SS, et al. Efficacy of adjuvant 5-fluorouracil therapy for patients with EMAS-positive stage II/III colorectal cancer. *PLoS ONE*. 2015;10:e0127591.

26. Ribic CM, Sargent DJ, Moore MJ, et al. Tumor microsatellite-instability status as a predictor of benefit from fluorouracil-based adjuvant chemotherapy for colon cancer. *N Engl J Med*. 2003;349:247-257.
27. 1000 Genomes Project Consortium, Abecasis GR, Altshuler D, Auton A, et al. A map of human genome variation from population-scale sequencing. *Nature*. 2010;467:1061-1073.
28. Tennessen JA, Bigham AW, O'Connor TD, et al. Evolution and functional impact of rare coding variation from deep sequencing of human exomes. *Science*. 2012;337:64-69.
29. C Yuen RK, Merico D, Bookman M, et al. Whole genome sequencing resource identifies 18 new candidate genes for autism spectrum disorder. *Nat Neurosci*. 2017;20:602-611.
30. Haraldsdottir S, Hampel H, Tomsic J, et al. Colon and endometrial cancers with mismatch repair deficiency can arise from somatic, rather than germline, mutations. *Gastroenterology*. 2014;147:1308.e1-1316.e1.
31. Palles C, Cazier J-B, Howarth KM, et al. Germline mutations affecting the proofreading domains of POLE and POLD1 predispose to colorectal adenomas and carcinomas. *Nat Genet*. 2013;45:136-144.
32. Hansen MF, Johansen J, Sylvander AE, et al. Use of multigene-panel identifies pathogenic variants in several CRC-predisposing genes in patients previously tested for Lynch Syndrome. *Clin Genet*. 2017;92:405-414.
33. Timmermann B, Kerick M, Roehr C, et al. Somatic mutation profiles of MSI and MSS colorectal cancer identified by whole exome next generation sequencing and bioinformatics analysis. *PLoS ONE*. 2010;5:e15661.
34. Chang SC, Lin PC, Lin JK, et al. Mutation spectra of common cancer-associated genes in different phenotypes of colorectal carcinoma without distant metastasis. *Ann Surg Oncol*. 2016;23:849-855.
35. Lin JK, Lin PC, Lin CH, et al. Clinical relevance of alterations in quantity and quality of plasma DNA in colorectal cancer patients: based on the mutation spectra detected in primary tumors. *Ann Surg Oncol*. 2014;21(Suppl 4):S680-S686.
36. Lin JK, Chang SC, Yang YC, Li AF. Loss of heterozygosity and DNA aneuploidy in colorectal adenocarcinoma. *Ann Surg Oncol*. 2003;10:1086-1094.
37. Hsieh P, Yamane K. DNA mismatch repair: molecular mechanism, cancer, and ageing. *Mech Ageing Dev*. 2008;129:391-407.
38. Castiglia D, Bernardini S, Alvino E, et al. Concomitant activation of Wnt pathway and loss of mismatch repair function in human melanoma. *Genes Chromosomes Cancer*. 2008;47:614-624.
39. Rodriguez-Soler M, Perez-Carbonell L, Guarinos C, et al. Risk of cancer in cases of suspected lynch syndrome without germline mutation. *Gastroenterology*. 2013;144(5):926.e1-932.e1; quiz e13-e14.
40. McGivern A, Wynter CV, Whitehall VL, et al. Promoter hypermethylation frequency and BRAF mutations distinguish hereditary non-polyposis colon cancer from sporadic MSI-H colon cancer. *Fam Cancer*. 2004;3:101-107.
41. Parsons MT, Buchanan DD, Thompson B, Young JP, Spurdle AB. Correlation of tumour BRAF mutations and MLH1 methylation with germline mismatch repair (MMR) gene mutation status: a literature review assessing utility of tumour features for MMR variant classification. *J Med Genet*. 2012;49:151-157.
42. Lee J, Xiao YY, Sun YY, et al. Prevalence and characteristics of hereditary non-polyposis colorectal cancer (HNPCC) syndrome in immigrant Asian colorectal cancer patients. *BMC Cancer*. 2017;17:843.
43. Aitken J, Bain C, Ward M, Siskind V, MacLennan R. How accurate is self-reported family history of colorectal cancer? *Am J Epidemiol*. 1995;141:863-871.
44. Mitchell RJ, Brewster D, Campbell H, et al. Accuracy of reporting of family history of colorectal cancer. *Gut*. 2004;53:291-295.
45. Lorans M, Dow E, Macrae FA, Winship IM, Buchanan DD. Update on hereditary colorectal cancer: improving the clinical utility of multigene panel testing. *Clin Colorectal Cancer*. 2018;17:e293-e305.
46. Stoffel EM, Koeppe E, Everett J, et al. Germline genetic features of young individuals with colorectal cancer. *Gastroenterology*. 2018;154:897-905.e1.
47. Sunga AY, Ricker C, Espenschied CR, et al. Spectrum of mismatch repair gene mutations and clinical presentation of Hispanic individuals with Lynch syndrome. *Cancer Genet*. 2017;212:1-7.
48. Pearlman R, Frankel WL, Swanson B, et al. Prevalence and spectrum of germline cancer susceptibility gene mutations among patients with Early-Onset colorectal cancer. *JAMA Oncol*. 2017;3:464-471.
49. Vasen HFA, Velthuizen ME, Kleibeuker JH, et al. Hereditary cancer registries improve the care of patients with a genetic predisposition to cancer: contributions from the Dutch Lynch syndrome registry. *Fam Cancer*. 2016;15:429-435.
50. Bakry D, Aronson M, Durno C, et al. Genetic and clinical determinants of constitutional mismatch repair deficiency syndrome: report from the constitutional mismatch repair deficiency consortium. *Eur J Cancer*. 2014;50:987-996.
51. Lee SY, Chung H, Devaraj B, et al. Microsatellite alterations at selected tetranucleotide repeats are associated with morphologies of colorectal neoplasias. *Gastroenterology*. 2010;139:1519-1525.
52. Smith CE, Mendillo ML, Bowen N, et al. Dominant mutations in *S. cerevisiae* PMS1 identify the Mlh1-Pms1 endonuclease active site and an exonuclease 1-independent mismatch repair pathway. *PLoS Genet*. 2013;9:e1003869.
53. Heck JA, Argueso JL, Gemici Z, et al. Negative epistasis between natural variants of the *Saccharomyces cerevisiae* MLH1 and PMS1 genes results in a defect in mismatch repair. *Proc Natl Acad Sci USA*. 2006;103:3256-3261.
54. Raut CP, Pawlik TM, Rodriguez-Bigas MA. Clinicopathologic features in colorectal cancer patients with microsatellite instability. *Mutat Res*. 2004;568:275-282.
55. Ward R, Meagher A, Tomlinson I, et al. Microsatellite instability and the clinicopathological features of sporadic colorectal cancer. *Gut*. 2001;48:821-829.
56. Benatti P, Gafa R, Barana D, et al. Microsatellite instability and colorectal cancer prognosis. *Clin Cancer Res*. 2005;11:8332-8340.
57. Koch A, Hrychuk A, Hartmann W, et al. Mutations of the Wnt antagonist AXIN2 (Conductin) result in TCF-dependent transcription in medulloblastomas. *Int J Cancer*. 2007;121:284-291.
58. Lammi L, Arte S, Somer M, et al. Mutations in AXIN2 cause familial tooth agenesis and predispose to colorectal cancer. *Am J Hum Genet*. 2004;74:1043-1050.
59. Liu W, Dong X, Mai M, et al. Mutations in AXIN2 cause colorectal cancer with defective mismatch repair by activating beta-catenin/TCF signalling. *Nat Genet*. 2000;26:146-147.
60. Boland CR, Thibodeau SN, Hamilton SR, et al. A National Cancer Institute Workshop on Microsatellite Instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. *Cancer Res*. 1998;58:5248-5257.

61. Markowitz S, Wang J, Myeroff L, et al. Inactivation of the type II TGF-beta receptor in colon cancer cells with microsatellite instability. *Science*. 1995;268:1336-1338.
62. Lu SL, Kawabata M, Imamura T, et al. HNPCC associated with germline mutation in the TGF-beta type II receptor gene. *Nat Genet*. 1998;19:17-18.
63. Verma L, Porter TR, Richards FM, et al. Germline mutation analysis of the transforming growth factor beta receptor type II (TGFBR2) and E-cadherin (CDH1) genes in early onset and familial colorectal cancer. *J Med Genet*. 2001;38:E7.
64. Cammareri P, Rose AM, Vincent DF, et al. Inactivation of TGFbeta receptors in stem cells drives cutaneous squamous cell carcinoma. *Nat Commun*. 2016;7:12493.
65. Campbell BB, Light N, Fabrizio D, et al. Comprehensive analysis of hypermutation in human cancer. *Cell*. 2017;171:1042-1056.e10.

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

**How to cite this article:** Chang S-C, Lan Y-T, Lin P-C, et al. Patterns of germline and somatic mutations in 16 genes associated with mismatch repair function or containing tandem repeat sequences. *Cancer Med*. 2020;9:476–486. <https://doi.org/10.1002/cam4.2702>