



doi:10.7659/j.issn.1005-6947.2024.04.009  
http://dx.doi.org/10.7659/j.issn.1005-6947.2024.04.009  
China Journal of General Surgery, 2024, 33(4):592-602.

· 专题研究 ·

## MLH1 蛋白与近端散发性结肠癌瘤内菌群特征关系的生物信息学分析

蒋萍<sup>1</sup>, 朱安超<sup>1</sup>, 刘营营<sup>2</sup>, 李宗敏<sup>1</sup>, 刘玉艳<sup>1</sup>, 单景军<sup>1</sup>, 何英<sup>3</sup>

(黑龙江省哈尔滨市第一医院 1. 病理科 3. 消化科, 黑龙江 哈尔滨 150010; 2. 黑龙江省医院 病理科, 黑龙江 哈尔滨 150036)

### 摘要

**背景与目的:** mutL 同源物 1 (MLH1) 基因的突变会导致 DNA 错配修复 (MMR) 系统失活, 从而增加结直肠癌的发生风险。此外, 越来越多的研究表明, 肠道菌群的组成和功能异常与结直肠癌的发生和发展密切相关。然而, MLH1 蛋白表达与肠道菌群是否有关联目前尚不十分明确。因此, 本研究通过分析中国东北地区不同 MLH1 蛋白表型近端散发性结肠癌 (SCC) 患者的肿瘤组织中微生物组的差异, 探讨两者之间的潜在关系。

**方法:** 收集黑龙江省哈尔滨市第一医院和黑龙江省医院 2020—2021 年 407 例近端 SCC 肿瘤组织样本与临床资料, 用免疫组化法筛选出其中 MLH1 蛋白缺失病例 (缺失组) 与 MLH1 蛋白完整病例 (对照组)。采用 16S rRNA 基因测序技术, 对提取的肠道肿瘤组织内微生物 DNA, 进行生物信息学分析, 并分析临床病理特征以及特定类群的物种与菌群多样性的关系。

**结果:** 共筛选出缺失组病例 20 例, 对照组 18 例, 初步临床数据分析显示, 肿瘤越大, MLH1 蛋白缺失的风险越高 ( $P < 0.05$ )。不同 MLH1 状态下肿瘤组织内菌群的  $\alpha$  多样性除 Shannon 指数外 ( $P = 0.042$ ), 其他指数差异均无统计学意义 (均  $P > 0.05$ )。门类水平的微生物组  $\beta$  多样性分析显示两组之间的菌群组成无明显差异 ( $P = 0.076$ )。属类水平上  $\beta$  多样性分析结果显示, 两组间差异大于两组内差异 ( $P = 0.04$ ), 通过菌属丰度之间的比较, 发现粪球菌属的菌群丰度可能会促进 MLH1 蛋白缺失 (校正后的  $P < 0.01$ )。然而, 无论各临床病理参数或是不同丰度的关键物种之间, Shannon 指数均无明显差异 (均  $P > 0.05$ )。

**结论:** 近端 SCC 患者 MLH1 蛋白表型与肠道菌群的组成和多样性密切相关。此外, 鉴定出粪球菌属可能成为该人群 MLH1 蛋白表达缺失的关键物种, 并为今后该病的研究与防治提供新的切入点。

### 关键词

结肠肿瘤; DNA 错配修复; MutL 同源蛋白 1; 微生物群

中图分类号: R735.3

**基金项目:** 黑龙江省卫生健康委科技计划基金资助项目 (20210101040241)。

**收稿日期:** 2023-05-08; **修订日期:** 2023-08-17。

**作者简介:** 蒋萍, 黑龙江省哈尔滨市第一医院副主任医师, 主要从事消化系统肿瘤临床病理方面的研究 (朱安超为共同第一作者)。

**通信作者:** 何英, Email: drheying@163.com

## Bioinformatics analysis of the relationship between MLH1 protein and tumor microflora characteristics in proximal sporadic colon cancer

JIANG Ping<sup>1</sup>, ZHU Anchao<sup>1</sup>, LIU Yingying<sup>2</sup>, LI Zongmin<sup>1</sup>, LIU Yuyan<sup>1</sup>, SHAN Jingjun<sup>1</sup>, HE Ying<sup>3</sup>

(1. Department of Pathology 3. Department of Gastroenterology, Harbin First Hospital, Harbin 150010, China; 2. Department of Pathology, Heilongjiang Provincial Hospital, Harbin 150036, China)

### Abstract

**Background and Aims:** Mutations in the mutL homolog 1 (MLH1) gene can lead to inactivation of the DNA mismatch repair (MMR) system, increasing the risk of colorectal cancer. Additionally, growing evidence suggests that alterations in the composition and function of the intestinal microbiota are closely associated with the occurrence and development of colorectal cancer. However, the relationship between MLH1 protein expression and the intestinal microbiota remains unclear. Therefore, this study aimed to explore the potential relationship between them by analyzing the differences in the microbial composition of tumor tissues between patients with proximal sporadic colon cancer (SCC) and different MLH1 protein phenotypes in Northeast China.

**Methods:** Tumor tissue samples and clinical data were collected from 407 patients with proximal SCC treated in Harbin First Hospital and Heilongjiang Provincial Hospital between 2020 and 2021. Immunohistochemistry was used to screen for cases with MLH1 protein deficiency (deficiency group) and intact MLH1 protein (control group). Microbial DNA extracted from the intestinal tumor tissues was analyzed using 16S rRNA gene sequencing technology for bioinformatics analysis. The relationship between clinicopathologic features and the diversity of specific taxa and microbial diversity was analyzed.

**Results:** A total of 20 cases were screened in the deficiency group, and 18 cases were screened in the control group. Preliminary analysis of clinical data showed that the larger the tumor, the higher the risk of MLH1 protein deficiency ( $P<0.05$ ). The  $\alpha$ -diversity of the microbial community within tumor tissues under different MLH1 statuses showed no statistically significant differences except for the Shannon index ( $P=0.042$ ). Other diversity indices had no significant difference (all  $P>0.05$ ). The  $\beta$ -diversity analysis of microbial communities at the phylum level showed no significant differences between the two groups ( $P=0.076$ ). At the genus level,  $\beta$ -diversity analysis showed that the differences between the two groups were greater than those within the groups ( $P=0.04$ ). A comparison of the abundance of bacterial genera revealed that the abundance of the genus *Coprococcus* spp may promote MLH1 protein deficiency (adjusted  $P<0.01$ ). However, there were no significant differences in the Shannon diversity index between various clinicopathologic variables or key species with different abundances (all  $P>0.05$ ).

**Conclusion:** The MLH1 protein phenotype of proximal SCC patients is closely related to the composition and diversity of the intestinal microbiota. Moreover, *Coprococcus* spp was identified as a potential key species associated with MLH1 protein loss in this population, providing a new perspective for future research, prevention, and treatment of this disease.

### Key words

Colonic Neoplasms; DNA Mismatch Repair; MutL Protein Homolog 1; Microbiota

**CLC number:** R735.3

在发达国家，结直肠癌（colorectal cancer, CRC）的发病率和病死率呈逐年上升趋势，同时也是中国的第四大恶性肿瘤，其防治已成为迫切需要解决的问题<sup>[1-2]</sup>。从遗传学的观点，可将CRC分为家族遗传性CRC和散发性结肠癌（sporadic colon cancer, SCC）。研究<sup>[3-4]</sup>发现，微卫星不稳定性（microsatellite instability, MSI）大约占SCC发病的15%~20%，而mutL同源物1（mutL homolog 1, MLH1）基因启动子异常甲基化是SCC患者重要的诊断依据。此外，为了更好地筛选SCC的人群和评估预后情况，也可通过确定BRAF突变状态和MSI状态来实现<sup>[5]</sup>。对不同部位（左半或右半）CRC进行分子生物学研究，发现BRAF突变以及MSI在近端CRC的发生率较高<sup>[6-7]</sup>。近年来，大量的研究发现该肿瘤相关的生物学特征，如微生物组、甲基化、共识分子亚型等，是影响预后的重要角色<sup>[8]</sup>。其中，微生物组和MSI作为肠癌免疫治疗中重要的标志物<sup>[9-10]</sup>，越来越受到人们的关注。另外，这两个标志物的研究还被认为可以辅助判读与该肿瘤有关的治疗效果。与菌群失调相关的炎症和致癌物的形成是微生物引起癌变作用的两个主要原因<sup>[11]</sup>。而肠道菌群产生的代谢物（包括胆汁酸和短链脂肪酸），考虑是参与微生物与宿主之间串扰的重要成分<sup>[12]</sup>。但到目前为止，SCC肿瘤组织中的MLH1蛋白表达与肠道菌群是否具有关联性尚不十分明确。本研究针对中国东北地区人群，通过16S rRNA基因测序分析近端SCC患者在不同MLH1状态下的肿瘤组织内微生物组的变化，揭示近端SCC患者肠道菌群的差异与MLH1蛋白缺失的潜在关系。

## 1 材料与方法

### 1.1 患者人群与样本收集

研究对象为2020—2021年来自中国东北地区在黑龙江省哈尔滨市第一医院和黑龙江省医院手术切除的CRC患者。收集患者肿瘤石蜡包埋样本和临床数据。根据遗传性结直肠癌临床诊治和家系管理中国专家共识<sup>[13]</sup>的诊断流程，以及美国癌症联合委员会（American Joint Committee on Cancer, AJCC）临床分期<sup>[14]</sup>进行样本的筛选。患者纳入和排除标准为：（1）患者在手术标本切除前没有接受化疗、放射治疗或抗生素治疗；（2）基于AJCC第8版

对TNM分类信息进行检索，并根据本版的TNM确定CRC的诊断；（3）已经或经常使用非类固醇抗炎药、他汀类药物或益生菌的患者被排除在外；（4）排除的其他标准还包括慢性肠道疾病、其他感染迹象、食物过敏和饮食限制。本研究已获得黑龙江省哈尔滨市第一医院伦理委员会批准（批准号：Q2021-064）。

### 1.2 免疫组化染色

采用MLH1（迈新公司，中国福州）、PMS2（迈新公司，中国福州）、MSH2（迈新公司，中国福州）、MSH6（迈新公司，中国福州）和BRAF（基因科技，中国上海）的单抗经4 mm切片后进行免疫组化检测。每张切片均配加阳性对照，操作过程参照试剂说明书完成，再使用OLYMPUS BX51显微镜对染色的组织切片进行观察判读。染色表达部位为核阳性，所有抗体都为阳性判读为错配修复蛋白表达完整，四种抗体中至少有1种抗体为阴性则判读为错配修复（mismatch repair, MMR）蛋白表达缺失。所有病理切片的检阅由2名副高级职称的病理诊断医师完成。

### 1.3 DNA提取及细菌16S rRNA测序

送检的肿瘤组织用福尔马林固定，再经石蜡包埋。每个样本采用8 μm连续切片10张制成蜡膜并放置于无菌微管中，然后在-20℃下保存，直到用于16S rRNA基因测序。所有的磁珠、试管和非酶试剂在使用前均经过紫外线消毒。用FastDNA<sup>®</sup> SPIN Kit（MP Biomedicals，美国）Soil提取试剂盒，从组织中提取细菌基因组DNA。然后使用通用引物（515F 5'-GTG CCA GCM GCC GCG G-3'和806R 5'-GGA CTA CHV GGG TWT CTA AT-3'）经PCR扩增16S rRNA基因V4区域，同时加入接头和样品特定的条形码序列。扩增条件为：95℃ 3 min，然后95℃ 30 s变性，55℃ 30 s退火，72℃ 30 s延伸，共30个循环。聚合酶链式反应产物在2%琼脂糖凝胶上电泳。用AxyPrep DNA凝胶回收试剂盒（Axygen Biosciences，美国），再次纯化扩增产物并进行定量。磁珠回收PCR产物得到最终的文库。利用Miseq PE300/NovaSeq PE250平台（美吉生物医药科技有限公司，中国上海）进行测序。参照笔者之前的研究<sup>[15]</sup>，将本次入组的原始数据上传至新的Figshare数据库（序列号：10.6084/m9.figshare.21153547）。



## 1.4 生物信息学分析

使用USEARCH 11以97%的阈值将序列聚类为可操作分类单元 (operational taxonomic units, OTUs), 去除嵌合体后将具有代表性的OTUs通过比对Silva 138 (<http://www.arb-silva.de>) 物种分类数据库得出, 分类置信度为0.7, 以供进一步分析。应用核糖体数据库项目 (Ribosomal Database Project, RDP) 分类器 v.2.13 对以下数据库中的OTUs代表序列进行分类: Greengene v135 和 RDP v11.5。为了揭示样本中各菌群在进化过程中的亲缘关系, 采用FastTree (version 2.1.3) 软件, 选择最大简约法MP挑选高丰度OTUs, 并在线比对NCBI数据库, Bootstrap值设置为500次重复。本次研究中, 将相对丰度的百分比值均<1%的菌群物种划定为Others, 进而筛选出优势菌群。主坐标分析 (principal co-ordinates analysis, PCOA), 是一种通过采用不同的距离算法对肠癌组织中的菌群进行降低维度的一种分析方法, 可以了解两组菌群的差异性与相关性。选用的距离算法是Bray-Curtis, 运用Mann-Whitney *U* test秩和检验对两组菌群进行不同水平的物种丰度比较, 并经FDR校正后测评物种的显著性差异, 校正后的*P*值 (即*q*值) <0.01。根据阳性病例的相对丰度, 以中位数为分界点, 将菌群总数水平分为高、低两个亚组, 未检出特定菌群的病例为阴性。采用Fisher检验和单因素方差检验分析各临床参数、特定类

群丰度与物种多样性的关系。统计和作图运用的软件是R语言 (version 3.3.1)。

## 1.5 统计学处理

临床参数的采集运用SPSS 21.0对基本数据进行分析。对符合正态分布的计量资料采用均数±标准差 ( $\bar{x} \pm s$ ) 表示, 组间比较采用*t*检验; 计数资料用例数 (百分比) [*n* (%)]表示, 组间比较采用Fisher检验。*P*<0.05为差异具有统计学意义。

## 2 结果

### 2.1 近端SCC患者肿瘤组织中免疫组化染色的表达

采用免疫组化SP法检测显示, MLH1、PMS2、MSH2及MSH6四种蛋白抗体主要表达于细胞核, 呈棕色 (图1)。407例患者中MMR蛋白表达缺失的患者为31例, 占总比的7.6%。而MLH1、PMS2、MSH2、MSH6蛋白缺失的患者占总MMR蛋白表达缺失的比例分别为64.5% (20/31)、22.6% (7/31)、9.7% (3/31)、3.2% (1/31)。在MMR修复蛋白中MLH1与PMS2偶联, 而MSH2与MSH6偶联。即MLH1 (或MSH2) 蛋白缺失同时PSM2 (或MSH6) 蛋白也会联合缺失, 而PSM2 (或MSH6) 蛋白缺失仅为单蛋白缺失。再对MLH1蛋白缺失的所有组织样本进行BRAF染色, 均显示为棕色细胞质, 呈阳性表达。

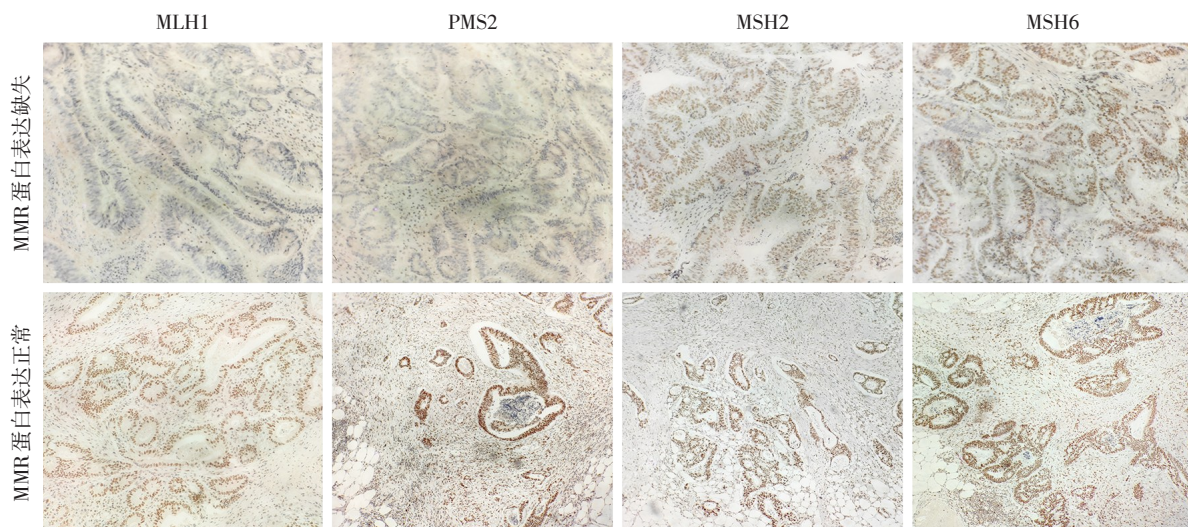


图1 肿瘤组织中不同MMR蛋白的免疫组化染色 (×200)

Figure 1 Immunohistochemical staining of different MMR proteins in tumor tissues (×200)

## 2.2 临床数据分析结果

选择上述20例MLH1蛋白缺失的近端SCC肿瘤组织（缺失组）与相同分期MLH1蛋白完整的肿瘤组织18例（对照组），并收集到的临床信息和相应的临床病理结果。临床数据分析结果显示，MLH1蛋白缺失与肿瘤大小有关（ $P=0.03$ ），即肿瘤最大直径 $\geq 4$  cm，MLH1蛋白缺失的风险较高，而其他各项水平上均无统计学意义（ $P>0.05$ ）（表1）。

表1 近端SCC患者的临床特征

Table 1 Clinical features of patients with proximal SCC

特征	缺失组(n=20)	对照组(n=18)	P
年龄(岁, $\bar{x} \pm s$ )	62.8 $\pm$ 13.25	63.5 $\pm$ 7.73	0.85
性别[n(%)]			
男	11(55.0)	11(61.1)	0.75
女	9(45.0)	7(38.9)	
肿瘤大小[cm, n(%)]			
<4	2(10.0)	8(44.4)	0.03
$\geq 4$	18(90.0)	10(55.6)	
分化程度[n(%)]			
高/中	17(85.0)	12(66.7)	0.26
低	3(15.0)	6(33.3)	
T分期[n(%)]			
T1/T2	3(15.0)	4(22.2)	0.69
T3/T4	17(85.0)	14(77.8)	
N分期[n(%)]			
N0	17(85.0)	11(61.1)	0.14
N1/N2	3(15.0)	7(38.9)	
M分期[n(%)]			
M0	20(100.0)	16(88.9)	0.22
M1	0(0.0)	2(11.1)	

## 2.3 近端SCC肿瘤组织内菌群高通量测序的数据结果

根据PCR检测结果，本项目合格的样本将进入后续建库上机测序。按最小样本OTUs序列数据抽平后，数据统计显示一共获得2 054 807条有效序列，碱基数为526 085 287 bp，序列平均长度为256。经OUTs物种分类信息统计，分类群包括49个菌群门类，153个菌群纲类，371个菌群目类，646个菌群科类，1 394个菌群属类，6 355个OTUs。

## 2.4 近端SCC肿瘤组织内菌群的稀释曲线与多样性分析

所有样品的稀释曲线（rarefaction curve）与Shannon-Wiener曲线分析在此测序深度上，曲线都已经趋向平坦（图2A-B），表明测序的深度足够

深，可以充分分析肠道肿瘤组织内菌群的组成。此外，与对照组相比，两组之间菌群关系的Venn图显示，49个菌群门中有10个是缺失组肿瘤组织唯一的菌群（图2C），1 394个菌群属中有411个是缺失组肿瘤组织唯一的菌群（图2D）。

运用Wilcoxon rank-sum检验对两组测序的有效序列做Sobs、Chao1、Simpson和Shannon指数运算。如表2所示，反映群落丰富度的Sobs和Chao1指数在缺失组中略高于对照组，反映群落多样性的Simpson指数在缺失组中略低于对照组，但差异均无统计学意义（均 $P>0.05$ ）；而反映群落多样性的Shannon指数，缺失组明显高于对照组（ $P<0.05$ ）。以上结果表明，两组在肿瘤组织内微生物的群落丰富度类似，而群落的多样性表现不同。

## 2.5 不同近端SCC肿瘤组织内菌群在门类水平上的菌群组成

在菌群门的分类水平上（图3A），利用Bray-Curtis距离矩阵的算法进行PCOA分析（图3B），散点图所示PC1解释率为67.85%，PC2解释率为16.79%，两组内差异大于两组间差异，无明显的统计学意义（ $P=0.076$ ）。对照组中菌群的置信椭圆几乎覆盖了缺失组的置信椭圆，表明两组间菌群相似性很高。本实验说明，在缺失组患者与对照组患者的肠道菌群门类水平上组间与组内无差异。在缺失组与对照组中包含的优势菌门，为变形菌门（Proteobacteria）、放线菌门（Actinobacteria）、厚壁菌门（Firmicutes）以及拟杆菌门（Bacteroidetes）共四个菌群，分别占菌群总数的96.92%和99.55%，其中两组中相对丰度最高的菌群是Proteobacteria。

## 2.6 不同近端SCC肿瘤组织内菌群在属类水平上的菌群组成

在菌群属的分类水平上，利用Bray-Curtis距离矩阵的算法进行PCOA分析（图4A），散点图所示PC1解释率为63.18%，PC2解释率为11.69%，两组间差异大于两组内差异，具有明显的统计学意义（ $P=0.04$ ）。对照组菌群的置信椭圆与缺失组的置信椭圆有一定的偏差，说明两组间菌群具有差异。在缺失组和对照组中优势菌属总共有11个（图4B），占比分别为93.24%和97.76%。其中两组中相对丰度最高的菌属是红球菌属（Rhodococcus）。通过系统发生进化树分析发现，Rhodococcus菌属和链霉菌属（Streptomyces）来源于Actinobacteriota菌门，芽孢杆菌属（Bacillus）来源于Firmicutes菌门，

Rikenellaceae\_RC9\_gut\_group 菌属来源于 Bacteroidota 菌门, unclassified\_norank\_Bacteria 菌属来源于 unclassified\_k\_norank\_d\_Bacteria 菌门, 其余菌属都来源于 Proteobacteria 菌门 (图4C)。

接下来, 通过 Wilcoxon 符号秩检验, 具有差异的菌属总共有 127 个, 在 11 个优势菌属中最终筛

选出 1 个具有明显差异的优势菌属 ( $q < 0.01$ ), 即粪球菌属 (Coprococcus) (图5)。与对照组相比, 缺失组中 Coprococcus 菌属的丰度明显升高, 并且发现该菌属仅存在于缺失组。从上述结果中考虑这类菌属在 SCC 患者肿瘤组织内的存在, 可能是影响 MLH1 蛋白缺失的关键 (或核心) 物种。

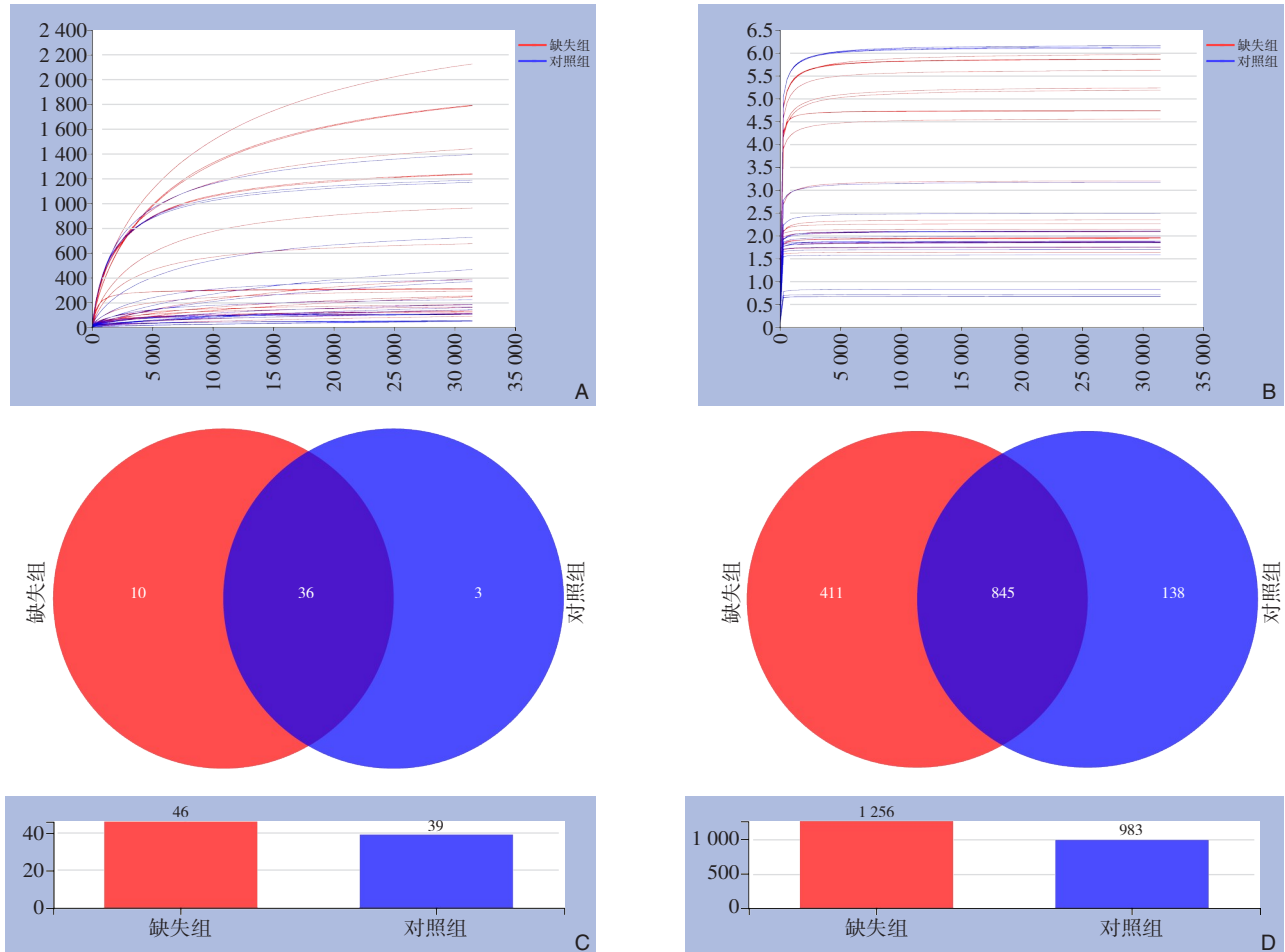


图2 缺失组和对照组肿瘤组织内微生物群落丰富度和多样性 A: Sobs 序列数估计的丰富度之间的稀疏曲线; B: Shannon 序列数估计的丰富度之间的稀疏曲线; C: 门类水平上, 代表了两组肿瘤组织间共享和独特的分类群的 Venn 图; D: 属类水平上两组的菌群关系的 Venn 图

Figure 2 Microbial community richness and diversity in tumor tissues of the deficiency and control groups A: Rarefaction curves of richness estimated by Sobs sequence numbers; B: Rarefaction curves of richness estimated by Shannon sequence numbers; C: Venn diagram representing shared and unique taxonomy between the two groups at the taxonomic level; D: Venn diagram showing the microbial community relationship between the two groups at the genus level

表2 两组间  $\alpha$  多样性分析

Table 2 Analysis of  $\alpha$ -diversity between two groups

指数	缺失组	对照组	P
Sobs	312.15±219.68	204.89±177.18	0.081
Chao	338.87±223.48	228.71±179.10	0.099
Simpson	0.16±0.11	0.30±0.05	0.052
Shannon	3.12±1.26	2.23±1.47	0.042



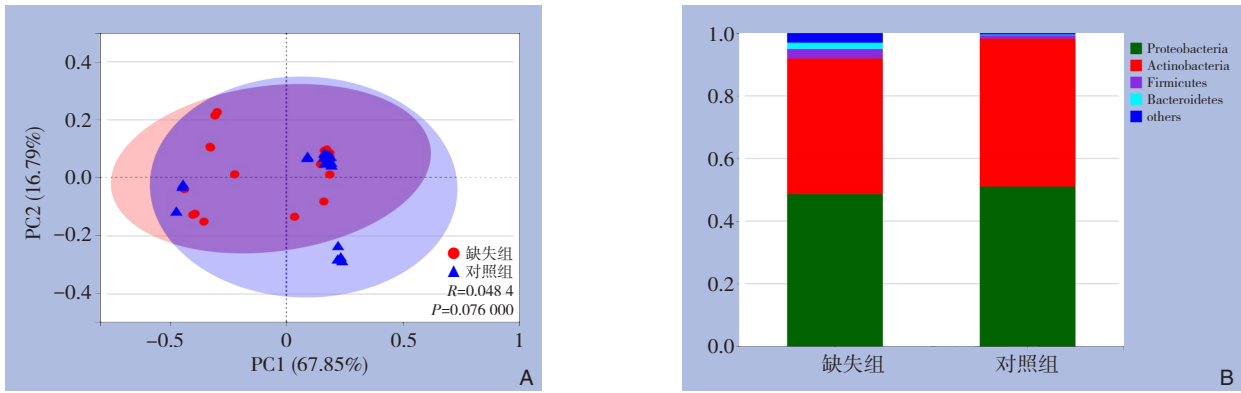


图 3 菌群门类的物种组成与样本间的分析 A: 两组样本间的PCOA分析; B: 两组菌群的柱状图

Figure 3 Analysis of species composition of microbial phyla between samples A: PCOA analysis between the two groups of samples; B: Bar chart of microbial communities between the two groups

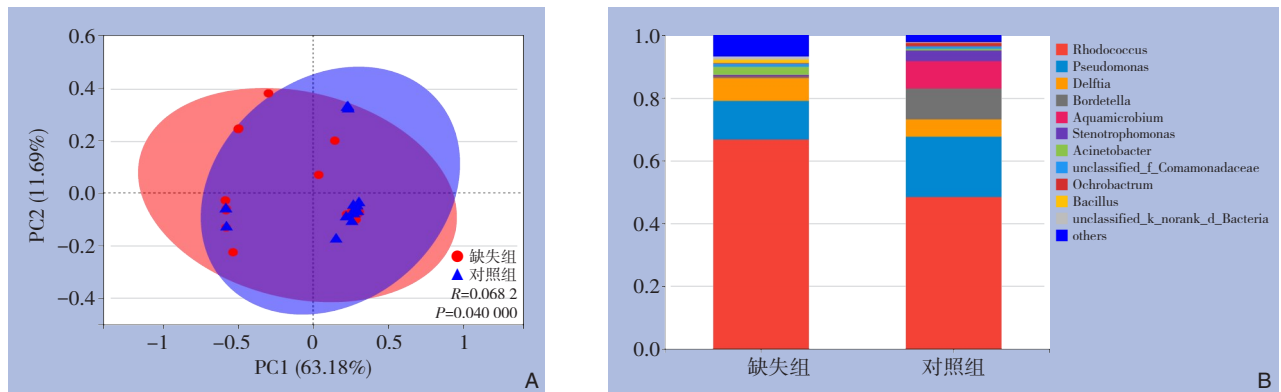


图 4 菌群属类的物种组成与样本间的分析 A: 两组样本间的PCOA分析; B: 两组菌群的柱状图; C: 系统发生进化树分析

Figure 4 Analysis of species composition of microbial genera between samples A: PCOA analysis between the two groups of samples; B: Bar chart of microbial communities between the two groups; C: Phylogenetic tree analysis

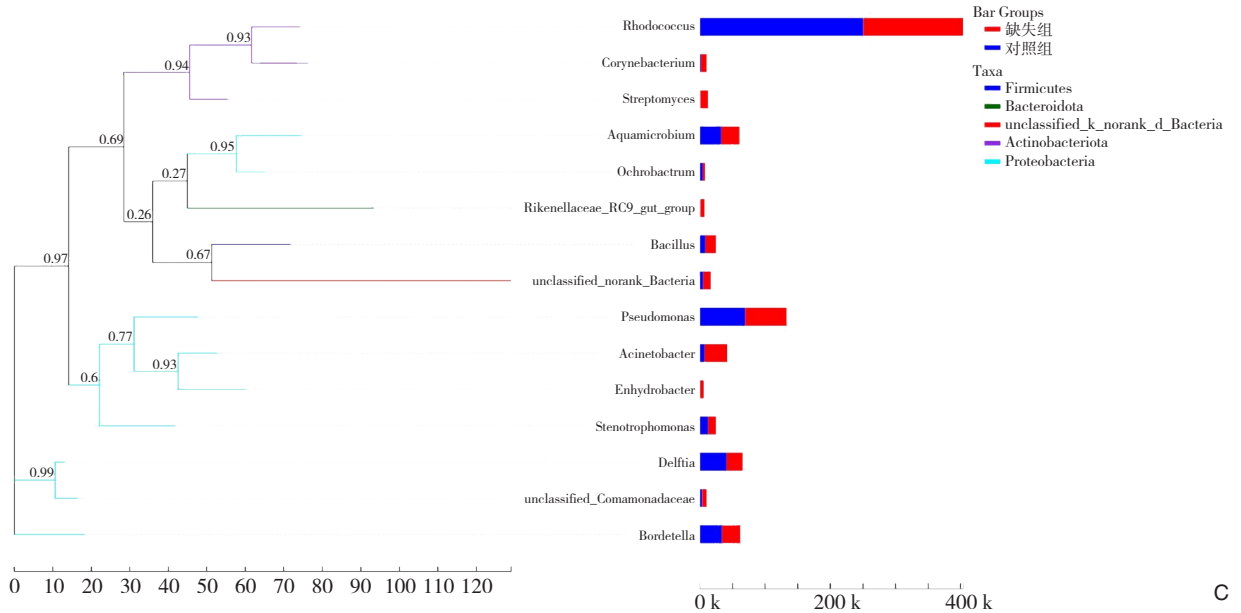


图 4 菌群属类的物种组成与样本间的分析 A: 两组样本间的PCOA分析; B: 两组菌群的柱状图; C: 系统发生进化树分析

Figure 4 Analysis of species composition of microbial genera between samples A: PCOA analysis between the two groups of samples; B: Bar chart of microbial communities between the two groups; C: Phylogenetic tree analysis

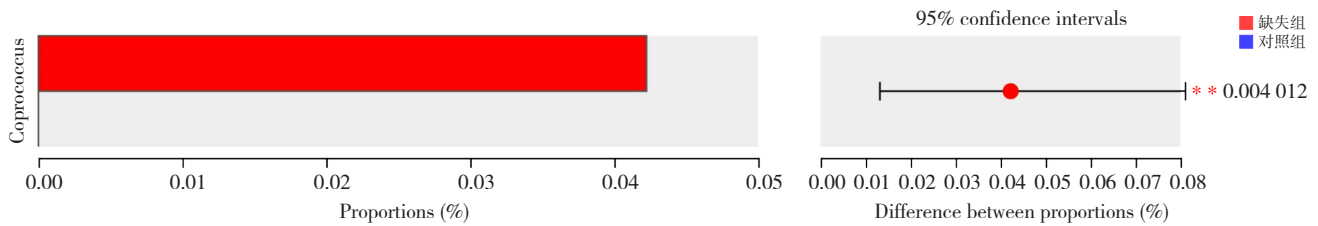


图5 在菌属丰度的分析中,具有明显差异的物种比较

Figure 5 Comparison of species with significant differences in genus abundance analysis

### 2.7 肿瘤临床参数、关键菌群与菌群多样性的关系

在确定了缺失和对照之间的核心微生物组后,进行分层分析,以调查微生物α多样性(Shannon指数)是否与两组的临床病理参数(肿瘤直径、侵袭深度、淋巴转移、远处转移)有关。在临床病理参数与组间Shannon指数的分析中,均显示无统计学意义(均 $P>0.05$ )。随后,分析Shannon指数与不同丰度的关键物种之间的相关性。在缺失组内不同Coprococcus菌属的富集程度与Shannon指数的差异无统计学意义( $P>0.05$ ),而在对照组内Coprococcus菌属也均为阴性(表3)。

表3 两组间α多样性与肿瘤特征、关键菌属的关系

Table 3 The relationship between α-diversity and tumor characteristics, as well as key bacterial genera, between the two groups

特征	α多样性(Shannon指数)			
	缺失组(n=20)	P	对照组(n=18)	P
肿瘤大小(cm)				
<4	3.09	0.97	2.39	0.70
≥4	3.12		2.10	
分化程度				
高/中	3.21	0.44	2.07	0.33
低	2.58		3.01	
T分期				
T1/T2	3.09	0.97	2.59	0.66
T3	3.12		2.16	
N分期				
N0	3.34	0.06	2.31	0.78
N1/N2	1.87		2.10	
Coprococcus spp.				
阴性/低丰度	2.91	0.20	未检出	—
高丰度	3.75		未检出	

注:为了评估不同微生物区系的丰度水平,使用微生物区系的平均丰度作为分类的分界点

Note: Using the average abundance of microbial communities as the cutoff point for classification to evaluate the abundance levels of different microbial communities

### 3 讨论

本研究显示,在MMR蛋白表达缺失的病例中MLH1蛋白缺失的表达最高。假设缺失组病例的肿瘤组织内可能存在不同于对照组病例的特定类群,并且这些菌群在每组病例中物种的组成和多样性之间也存在潜在的关联。为此,采用16S rRNA基因测序技术对肿瘤组织内微生物组的组成和多样性进行分析。发现在近端SCC患者不同MLH1蛋白表达的α多样性比较中,除Shannon指数外,其他指数均未显示总体物种的含量或多样性的重大变化。然而,在两组间同类水平上菌群的配对分析中,显示作为优势菌群的Coprococcus菌属相对丰度存在显著差异。此外,每组内这些关键物种相对丰度的差异也与Shannon多样性指数具有密切的联系。

通过人类癌症基因组的大规模测序发现,MSI的发生主要是由于DNA MMR受损所致<sup>[16-17]</sup>。在一些情况下,部分缺陷的MMR蛋白系统是生命相容,实际上可能是对生存有益<sup>[18]</sup>。目前,对于手术切除的样本在国内外均采用免疫组化的方法进行MMR的检测,作为评估SCC患者预后生存状况和临床免疫治疗的重要标准。在本研究的队列中,MSI SCC的发生率为7.6%,与西方报道的发生率相比,更接近亚洲其他地区报道的5.9%~7.8%的发生率<sup>[3,19-20]</sup>。这些SCC患者MMR基因突变模式存在较大差异的原因可能与不同种族、不同地域有关<sup>[21]</sup>。同时发现MLH1蛋白缺失的病例是SCC MMR缺失蛋白患者的主要部分,结果与其他的人群研究相一致<sup>[22-24]</sup>。同样,在临床肿瘤分期中,与近端SCC有关的MMR蛋白表达缺失发现主要发生在I期和II期,其次是III期,并且淋巴结转移的发生率也较低<sup>[19,25]</sup>。

研究<sup>[26]</sup>发现,肠癌患者的肠道微生物群失衡(失调),显示“坏”微生物的数量增加,而“好”



微生物的数量减少。当菌群失调时,不仅会影响宿主体内黏膜的完整性,也会调节炎症性信号通路和免疫细胞的应答,说明肠道菌群可能与肿瘤的致病机制之间存在一定的因果关系<sup>[27]</sup>。

在菌群门类水平上,本研究发现缺失组和对照组在近端 SCC 肿瘤组织中微生物的组成与之前报道<sup>[28-29]</sup>相似,均发现优势菌群为 Proteobacteria、Firmicutes、Actinobacteria 和 Bacteroidetes 总群落占比 >90%。研究<sup>[30]</sup>发现, Firmicutes/Bacteroidetes 比值的增加会引起低度炎症,导致 IL-6 基因表达和 DNA 的损伤增加,并通过增加 MLH1 甲基化状态导致该基因表达减少。但是,本研究中门类水平的微生物组在  $\beta$  多样性分析中,显示不同组之间的菌群组成无显著差异,表示两组间菌群结构相似。Jin 等<sup>[28]</sup>对 329 例肠癌患者的肿瘤组织分析,发现四个优势菌属为 Acinetobacter 属、贪铜菌属、鞘脂菌属和鞘脂单胞菌属。但本研究中,对近端 SCC 组织样本的属类菌群的组成分析,显示优势菌属共有 11 个,其中丰度最高的菌属是 Rhodococcus。随后,笔者将两组病例的菌群属进行物种差异性比较,发现 Coprococcus 菌属与缺失病例成正比,并且与 IMS 病例无关。在菌群属类水平下, Ai 等<sup>[31]</sup>发现 Lachnocostridium 和 Coprococcus 在 CRC 患者与健康人群之间具有显著差异。因此,证实上述 Coprococcus 菌属作为关键物种与不同 MLH1 蛋白表达的近端 SCC 患者具有密切的关联。更重要的是,考虑在近端 SCC 患者肠道瘤内关键菌属 Coprococcus 的富集,对 MLH1 蛋白缺失可能会存在潜在的风险,但还需要进一步地验证。此外,研究<sup>[32-33]</sup>报道,有一些产丁酸的细菌,如 Coprococcus 菌属(属于假单胞菌门)能够刺激 MSH2 或 MLH1 缺乏的结肠上皮细胞异常过度增殖,并通过提高丁酸水平增强  $\beta$ -连环蛋白的活性。基于上述的研究结果,我们考虑肠道肿瘤组织内微生物组的差异性可能与不同地域的外部环境和个人饮食习惯等因素的影响有关。

根据微生物多样性的分析结果,展示出在缺失组病例的群落多样性方面, Shannon 指数显著高于对照组病例,而 Simpson 指数则相反并且无显著性差异。在群落丰富度方面,缺失组病例的 Sobs 指数和 Chao1 指数都仅略高于对照组病例,并且均无显著性差异。既往报道<sup>[34]</sup>,发现 Chao1 和 Simpson 指数在 MMR 蛋白表达缺失组中呈显著上

调,但 Shannon 指数仅在 MMR 蛋白表达缺失组中具有增加趋势而无显著差异。可以看出,肠癌患者体内菌群的多样性可能与肠道肿瘤微环境具有潜在的关联。随后,为了更深入地揭示特定肠道微环境中不同参数的物种状况,笔者再针对 Shannon 指数与各组中病理临床数据及特定微生物群的关系进行研究。但结果表明,近端 SCC 患者的各个临床病理特征和关键菌属 DNA 可能不会对瘤内群落多样性产生影响。

目前,肠癌免疫治疗中出现的生物标志物主要分为四种类型:肿瘤突变、预先存在的免疫反应、程序性死亡受体 1 (programmed cell death 1, PD-1) 配体 (programmed death ligand 1, PD-L1) 表达和微生物群<sup>[35]</sup>。与此同时,肠癌患者的肿瘤免疫疗法已经显示出持续的临床反应,并正在引发癌症治疗的范式转变。其中, PD-1 及其配体 PD-L1 常作为免疫检查点抑制剂治疗中重要的生物靶点而备受关注。既往已有大量研究证实 MMR 蛋白表达缺失也对预测 PD-1/PD-L1 抑制剂的治疗反应有一定价值<sup>[36-37]</sup>,本研究主要针对 MLH1 蛋白缺失的近端 SCC 临床患者癌组织的样本进行研究。

由于近年来研究<sup>[38-39]</sup>报道,一些肠道微生物组可以通过塑造宿主免疫来诱导 CD8<sup>+</sup>T 细胞反应,进而调节 PD-1/PD-L1 抑制剂的抗肿瘤疗效。本研究结果显示, Coprococcus 菌属是此次研究的关键物种。此外,该菌群代谢产物丁酸盐可通过 ID2 依赖的 IL-12 信号通路调节抗肿瘤 CD8<sup>+</sup>T 细胞应答,提示肠道微生物代谢物可以促进抗癌免疫,从而充分提高治疗效果<sup>[40-41]</sup>。笔者认为上述文献,虽然揭示了肠道微生物代谢物的操控可能是有效治疗癌症的一部分,但如何确定丁酸盐对小鼠和患者的补充剂量仍有待于进一步的研究。

尽管,本研究设计对假设进行了严格的测试,但同样也具有一定的局限性。由于不能从这些病例样本中确定关键物种的真实性能,未来需要再对 SCC 患者的大样本队列进行验证。同时,针对本次观察性研究中瘤内菌群特征与 MLH1 蛋白缺失之间的因果关系也并未涉及。此外,本研究的人群主要来自中国东北地区,这可能会限制我们的研究结果对其他种族和地区人群的推广。

综上所述,本次研究主要针对近端 SCC 患者不同 MLH1 蛋白表达的肿瘤组织内菌群的组成结构与物种的多样性进行分析。发现 Coprococcus 菌属

可能是导致该类肿瘤内 MLH1 蛋白缺失的关键物种,并且各项临床病理参数以及不同丰富度的该物种与菌群多样性无关。此研究通过探究肠道菌群与宿主肿瘤微环境之间的关系,为肿瘤预防和精准治疗提供了一种新的策略。

利益冲突:所有作者均声明不存在利益冲突。

作者贡献声明:蒋萍、朱安超负责收集与分析数据、撰写文章;何英负责总体把控、批改文章;刘营营、李宗敏、刘玉艳和单景军负责提供病例及数据。

## 参考文献

- [1] Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries[J]. *CA Cancer J Clin*, 2021, 71(3):209–249. doi: 10.3322/caac.21660.
- [2] Chen W, Zheng R, Baade PD, et al. Cancer statistics in China, 2015[J]. *CA Cancer J Clin*, 2016, 66(2):115–132. doi: 10.3322/caac.21338.
- [3] Nakayama Y, Iijima T, Inokuchi T, et al. Clinicopathological features of sporadic MSI colorectal cancer and Lynch syndrome: a single-center retrospective cohort study[J]. *Int J Clin Oncol*, 2021, 26(10):1881–1889. doi: 10.1007/s10147-021-01968-y.
- [4] Nouri Nojaded J, Hashemzadeh S, Samadi Kafil H, et al. Evaluation of microsatellite instability in tumor and tumor marginal samples of sporadic colorectal cancer using mononucleotide markers[J]. *EXCLI J*, 2018, 17: 945–951. doi: 10.17179/excli2018-1455.
- [5] Stachler MD, Rinehart E, Lindeman N, et al. Novel molecular insights from routine genotyping of colorectal carcinomas[J]. *Hum Pathol*, 2015, 46(4):507–513. doi: 10.1016/j.humpath.2015.01.005.
- [6] Missiaglia E, Jacobs B, D'Ario G, et al. Distal and proximal colon cancers differ in terms of molecular, pathological, and clinical features[J]. *Ann Oncol*, 2014, 25(10):1995–2001. doi: 10.1093/annonc/mdu275.
- [7] Lee MS, Menter DG, Kopetz S. Right versus left colon cancer biology: integrating the consensus molecular subtypes[J]. *J Natl Compr Canc Netw*, 2017, 15(3): 411–419. doi: 10.6004/jnccn.2017.0038.
- [8] Molina-Cerrillo J, San Román M, Pozas J, et al. BRAF mutated colorectal cancer: new treatment approaches[J]. *Cancers (Basel)*, 2020, 12(6):1571. doi: 10.3390/cancers12061571.
- [9] Qu R, Zhang Y, Ma Y, et al. Role of the gut microbiota and its metabolites in tumorigenesis or development of colorectal cancer[J]. *Adv Sci (Weinh)*, 2023, 10(23):e2205563. doi: 10.1002/advs.202205563.
- [10] Amodio V, Lamba S, Chilà R, et al. Genetic and pharmacological modulation of DNA mismatch repair heterogeneous tumors promotes immune surveillance[J]. *Cancer Cell*, 2023, 41(1):196–209. doi: 10.1016/j.ccell.2022.12.003.
- [11] Cai J, Sun L, Gonzalez FJ. Gut microbiota-derived bile acids in intestinal immunity, inflammation, and tumorigenesis[J]. *Cell Host Microbe*, 2022, 30(3):289–300. doi: 10.1016/j.chom.2022.02.004.
- [12] Sepich-Poore GD, Zitvogel L, Straussman R, et al. The microbiome and human cancer[J]. *Science*, 2021, 371(6536): eabc4552. doi: 10.1126/science.abc4552.
- [13] 中国抗癌协会大肠癌专业委员会遗传学组. 遗传性结直肠癌临床诊治和家系管理中国专家共识[J]. *实用肿瘤杂志*, 2018, 33(1): 3–16. doi: 10.13267/j.cnki.syzlzz.2018.01.002  
Genetics Group of the Committee of Colorectal Cancer, China Anticancer Association. The Chinese expert consensus on clinical diagnosis, treatment and pedigree management of hereditary colorectal cancer[J]. *Journal of Practical Oncology*, 2018, 33(1):3–16. doi: 10.13267/j.cnki.syzlzz.2018.01.002.
- [14] Weiser MR. AJCC 8th edition: colorectal cancer[J]. *Ann Surg Oncol*, 2018, 25(6):1454–1455. doi: 10.1245/s10434-018-6462-1.
- [15] Zhu A, Liu Y, Li Z, et al. Diagnosis and functional prediction of microbial markers in tumor tissues of sporadic colorectal cancer patients associated with the MLH1 protein phenotype[J]. *Front Oncol*, 2022, 12:1116780. doi: 10.3389/fonc.2022.1116780.
- [16] Cao R, Yang F, Ma SC, et al. Development and interpretation of a pathomics-based model for the prediction of microsatellite instability in Colorectal Cancer[J]. *Theranostics*, 2020, 10(24): 11080–11091. doi: 10.7150/thno.49864.
- [17] Le DT, Durham JN, Smith KN, et al. Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade[J]. *Science*, 2017, 357(6349):409–413. doi: 10.1126/science.aan6733.
- [18] Voskarides K. An evolutionary explanation for antibiotics' association with increased colon cancer risk[J]. *Evol Med Public Health*, 2022, 10(1):214–220. doi: 10.1093/emph/eoac018.
- [19] Vilar E, Gruber SB. Microsatellite instability in colorectal cancer—the stable evidence[J]. *Nat Rev Clin Oncol*, 2010, 7(3): 153–162. doi: 10.1038/nrclinonc.2009.237.
- [20] 谢仲鹏, 李海荣, 吴余, 等. 海南地区结直肠癌中错配修复蛋白及多药耐药蛋白的表达及意义[J]. *临床与实验病理学杂志*, 2020, 36(10):1205–1208. doi: 10.13315/j.cnki.cjcep.2020.10.016.  
Xie ZP, Li HR, Wu Y, et al. Expression and significance of mismatch repair protein and multidrug resistance protein in colorectal cancer in Hainan area[J]. *Journal of Clinical and Experimental Pathology*, 2020, 36(10):1205–1208. doi: 10.13315/j.cnki.cjcep.2020.10.016.
- [21] Yan WY, Hu J, Xie L, et al. Prediction of biological behavior and prognosis of colorectal cancer patients by tumor MSI/MMR in the Chinese population[J]. *Onco Targets Ther*, 2016, 9:7415–7424. doi: 10.2147/OTT.S117089.
- [22] Ghanipour L, Jirström K, Sundström M, et al. Associations of defect mismatch repair genes with prognosis and heredity in sporadic colorectal cancer[J]. *Eur J Surg Oncol*, 2017, 43(2):311–

321. doi: 10.1016/j.ejso.2016.10.013.
- [23] Yearsley M, Hampel H, Lehman A, et al. Histologic features distinguish microsatellite-high from microsatellite-low and microsatellite-stable colorectal carcinomas, but do not differentiate germline mutations from methylation of the MLH1 promoter[J]. *Hum Pathol*, 2006, 37(7): 831–838. doi: 10.1016/j.humpath.2006.02.009.
- [24] 李云凤, 田琳. 结直肠癌中错配修复蛋白的表达及其与临床病理特征的相关性分析[J]. *哈尔滨医科大学学报*, 2022, 56(3):285–288. doi: 10.20010/j.issn.1000–1905.2022.03.0285.
- Li YF, Tian L. Analysis of the expression of mismatch repair protein in colorectal cancer and its correlation with clinicopathological characteristics[J]. *Journal of Harbin Medical University*, 2022, 56(3): 285–288. doi: 10.20010/j.issn.1000–1905.2022.03.0285.
- [25] Battaglin F, Naseem M, Lenz HJ, et al. Microsatellite instability in colorectal cancer: overview of its clinical significance and novel perspectives[J]. *Clin Adv Hematol Oncol*, 2018, 16(11):735–745. doi: 10.1634/theoncologist.2017–0289.
- [26] Abdulla MH, Agarwal D, Singh JK, et al. Association of the microbiome with colorectal cancer development (Review)[J]. *Int J Oncol*, 2021, 58(5):17. doi: 10.3892/ijo.2021.5197.
- [27] 刘珂良, 王晓慧, 李思柔, 等. 肠道微生物群在肝胆胰恶性肿瘤中的研究进展[J]. *中国普通外科杂志*, 2023, 32(2):279–286. doi: 10.7659/j.issn.1005–6947.2023.02.014.
- Liu KL, Wang XH, Li SR, et al. Research progress of gut microbiome in hepatobiliary and pancreatic malignancies[J]. *China Journal of General Surgery*, 2023, 32(2): 279–286. doi: 10.7659/j.issn.1005–6947.2023.02.014.
- [28] Jin M, Shang F, Wu J, et al. Tumor-associated microbiota in proximal and distal colorectal cancer and their relationships with clinical outcomes[J]. *Front Microbiol*, 2021, 12: 727937. doi: 10.3389/fmicb.2021.727937.
- [29] 陆晓娟. 肠道菌群和遗传性非息肉病性结直肠癌相关性研究[D]. 北京: 中国人民解放军医学院, 2020. doi:10.27637/d.cnki.gjyjc.2019.000101.
- Lu XJ. Correlation between intestinal flora and hereditary nonpolyposis colorectal cancer[D]. Beijing: Chinese People's Liberation Army Medical College, 2020. doi: 10.27637/d.cnki.gjyjc.2019.000101.
- [30] Remely M, Ferk F, Sterneder S, et al. EGCG prevents high fat diet-induced changes in gut microbiota, decreases of DNA strand breaks, and changes in expression and DNA methylation of Dnmt1 and MLH1 in C57BL/6J male mice[J]. *Oxid Med Cell Longev*, 2017, 2017:3079148. doi: 10.1155/2017/3079148.
- [31] Ai DM, Pan HF, Li XX, et al. Identifying gut microbiota associated with colorectal cancer using a zero-inflated lognormal model[J]. *Front Microbiol*, 2019, 10:826. doi: 10.3389/fmicb.2019.00826.
- [32] van den Abbeele P, Belzer C, Goossens M, et al. Butyrate-producing *Clostridium* cluster XIVa species specifically colonize mucins in an in vitro gut model[J]. *ISME J*, 2013, 7(5):949–961. doi: 10.1038/ismej.2012.158.
- [33] Belcheva A, Irrazabal T, Robertson SJ, et al. Gut microbial metabolism drives transformation of MSH2-deficient colon epithelial cells[J]. *Cell*, 2014, 158(2): 288–299. doi: 10.1016/j.cell.2014.04.051.
- [34] 吴晶晶. dMMR/pMMR结直肠癌肠道菌群及与肿瘤微环境特征的相关性研究[D]. 武汉: 华中科技大学, 2023. doi:10.27157/d.cnki.ghzku.2021.000241.
- Wu JJ. Study on intestinal flora of dMMR/pMMR colorectal cancer and its correlation with tumor microenvironment characteristics[D]. Wuhan: Huazhong University of Science and Technology, 2023. doi:10.27157/d.cnki.ghzku.2021.000241.
- [35] Fan A, Wang B, Wang X, et al. Immunotherapy in colorectal cancer: current achievements and future perspective[J]. *Int J Biol Sci*, 2021, 17(14):3837–3849. doi: 10.7150/ijbs.64077.
- [36] Bai Z, Zhou Y, Ye Z, et al. Tumor-infiltrating lymphocytes in colorectal cancer: the fundamental indication and application on immunotherapy[J]. *Front Immunol*, 2021, 12:808964. doi: 10.3389/fimmu.2021.808964.
- [37] Jiang D, Song Q, Wei X, et al. PMS2 expression with combination of PD-L1 and TILs for predicting survival of esophageal squamous cell carcinoma[J]. *Front Oncol*, 2022, 12: 897527. doi: 10.3389/fonc.2022.897527.
- [38] Tanoue T, Morita S, Plichta DR, et al. A defined commensal consortium elicits CD8 T cells and anti-cancer immunity[J]. *Nature*, 2019, 565(7741): 600–605. doi: 10.1038/s41586–019–0878–z.
- [39] Gao YH, Bi DX, Xie RT, et al. *Fusobacterium nucleatum* enhances the efficacy of PD-L1 blockade in colorectal cancer[J]. *Signal Transduct Target Ther*, 2021, 6(1):398. doi: 10.1038/s41392–021–00795–x.
- [40] He Y, Fu LH, Li YP, et al. Gut microbial metabolites facilitate anticancer therapy efficacy by modulating cytotoxic CD8+ T cell immunity[J]. *Cell Metab*, 2021, 33(5):988–1000.e7. doi: 10.1016/j.cmet.2021.03.002.
- [41] Low L, Suleiman K, Shamdas M, et al. Gut dysbiosis in ocular mucous membrane pemphigoid[J]. *Front Cell Infect Microbiol*, 2022, 12:780354. doi: 10.3389/fcimb.2022.780354.

( 本文编辑 宋涛 )

本文引用格式: 蒋萍, 朱安超, 刘营营, 等. MLH1 蛋白与近端散发性结肠癌瘤内菌群特征关系的生物信息学分析[J]. *中国普通外科杂志*, 2024, 33(4):592–602. doi: 10.7659/j.issn.1005–6947.2024.04.009

Cite this article as: Jiang P, Zhu AC, Liu YY, et al. Bioinformatics analysis of the relationship between MLH1 protein and tumor microflora characteristics in proximal sporadic colon cancer[J]. *Chin J Gen Surg*, 2024, 33(4): 592–602. doi: 10.7659/j.issn.1005–6947.2024.04.009