Journal of Gastrointestinal Cancer
https://doi.org/10.1007/s12029-023-00925-4

RESEARCH

Gut Microbiome Composition in Lynch Syndrome With and Without History of Colorectal Neoplasia and Non-Lynch Controls

S. B. Rifkin[1,2] · M. A. Sze[3] · K. Tuck[1] · E. Koeppe[1,2] · E. M. Stoffel[1,2] · P. D. Schloss[3]

Accepted: 24 February 2023

Abstract

Background While Lynch syndrome (LS) is a highly penetrant colorectal cancer (CRC) syndrome, there is considerable variation in penetrance; few studies have investigated the association between microbiome and CRC risk in LS. We analyzed the microbiome composition among individuals with LS with and without personal history of colorectal neoplasia (CRN) and non-LS controls.

Methods We sequenced the V4 region of the 16S rRNA gene from the stool of 46 individuals with LS and 53 individuals without LS. We characterized within community and in between community microbiome variation, compared taxon abundance, and built machine learning models to investigate the differences in microbiome.

Results There was no difference within or between community variations among LS groups, but there was a statistically significant difference in both within and between community variation comparing LS to non-LS. *Streptococcus* and *Actinomyces* were differentially enriched in LS-CRC compared to LS-without CRN. There were numerous differences in taxa abundance comparing LS to non-LS; notably, *Veillonella* was enriched and *Faecalibacterium* and *Romboutsia* were depleted in LS. Finally, machine learning models classifying LS from non-LS controls and LS-CRC from LS-without CRN performed moderately well.

Conclusions Differences in microbiome composition between LS and non-LS may suggest a microbiome pattern unique to LS formed by underlying differences in epithelial biology and immunology. We found specific taxa differences among LS groups, which may be due to underlying anatomy. Larger prospective studies following for CRN diagnosis and microbiome composition changes are needed to determine if microbiome composition contributes to CRN development in patients with LS.

Keywords Lynch syndrome (LS) · Colorectal cancer (CRC) · Colorectal neoplasia (CRN) · Microbiome · DNA mismatch repair (MMR)

Introduction

Lynch syndrome (LS) is an autosomal dominant, highly penetrant colorectal cancer (CRC) syndrome caused by inherited defects in DNA mismatch repair (MMR) genes (*MLH1*, *MSH2*, *MSH6*, *PMS2*, and *EPCAM*). There is considerable variation in phenotypes of colorectal neoplasia (CRN) among carriers of LS-associated germline variants [1]. This heterogeneity has been attributed to patient factors, such as age, sex, specific MMR genetic mutation, and/or health behaviors including diet, physical activity, and smoking; however, the relative impact of these on the severity of CRN is still unclear.

Accumulating evidence suggests that features of the microbiome contribute to sporadic colon cancer. Specific bacteria with virulent toxins have been associated with CRC, including *Bacteroides fragilis* (ETBF), *Escherichia coli* (loci/colibactin/pks + Ec), *Fusobacterium nucleatum* (FadA adhesin/invasin), and other bacteria (e.g., Porphyromonas), but the majority of studies have described global alterations in microbiome [2–11]. Recurrent patterns of dysbiosis, or imbalanced gut microbiome, found in CRC cases include lower α diversity,

✉ S. B. Rifkin
srifkin@med.umich.edu

1 Department of Internal Medicine, Rogel Cancer Center, University of Michigan, Ann Arbor, MI 48109, USA

2 Division of Gastroenterology and Hepatology, University of Michigan School of Medicine, Ann Arbor, MI 48109, USA

3 Department of Immunology and Microbiology, University of Michigan, Ann Arbor, MI 48109, USA

Published online: 13 June 2023

differences in β diversity, and recurring differences in functional taxon abundances, namely decreased levels of beneficial commensal bacteria and increased pathogenic bacteria [12]. Furthermore, increasing dysbiosis may track with the adenoma-adenocarcinoma spectrum and excision of colonic lesions has been shown to normalize the microbiome [13, 14].

Few studies have investigated the dynamics of the microbiome composition in LS. These studies have not shown a consistent difference in the microbiome composition comparing MMR pathogenic variant carriers with CRC and/or adenoma to carriers without prior CRN. One larger study followed a cohort of 100 patients with LS for 1 year and found that the microbiome in baseline stool sample differed by baseline diagnosis of adenoma, but also noted that history of colon resection was more strongly associated with observed differences in microbiome composition [15]. Meanwhile, several smaller studies have noted a difference between individuals with and without LS [15–18].

Our aim was to determine whether there were differences in the microbiome composition among individuals with LS with and without personal history of CRN and non-LS controls.

Methods

Study Design

Subjects We enrolled 46 individuals with genetically confirmed LS from the University of Michigan cancer genetics clinic. We enrolled 53 individuals without LS and without a history of CRN (non-LS controls) recruited from the local Ann Arbor community. Individuals with diagnoses of HIV/AIDS, hepatitis B or C, and/or inflammatory bowel disease were excluded. All study participants provided informed consent. The study protocol was approved by the University of Michigan IRB (HUM00113700).

Data Collection At enrollment, participants filled out questionnaires eliciting information on demographics, medical comorbidities, current medications, and family history. Subjects provided a stool sample for sequencing using a Zymo stool collection kit and DNA/RNA Shield Collection tube.

16S rRNA Gene Sequencing and Analysis

DNA was extracted using a 96-well Soil DNA isolation kit and an epMotion 5075 automated pipetting system. We amplified the V4 region of the 16S rRNA gene using custom barcoded primers and sequenced it using Illumina MiSeq sequencer [19]. The 16S rRNA gene sequences were curated using the *mothur* software package (version 1.47.0) [20]. We merged paired-end reads into contigs, aligned them to the SILVA 16S rRNA sequence database, and removed low-quality sequences and chimeras [21]. We classified sequences by training a naive Bayesian classifier with a 16S rRNA gene training set (Ribosomal Database Project (RDP)) [22]. We clustered sequences into operational taxonomic units (OTUs) using a 97% similarity cutoff with the OptiClust clustering algorithm [23].

Statistical Analysis

All statistical analyses were performed in R (version 4.0.2). Baseline characteristics were characterized and compared according to study groups: LS subjects with a history of CRC (LS-CRC), a history of adenoma (LS-adenoma), no history of either CRC or adenoma (LS-without CRN), and non-Lynch controls (non-LS controls). These baseline variables were assessed using medians and interquartile ranges for continuous variables, and as frequencies and percentages for categorical variables.

We quantified the variation in microbial communities using metrics for α-diversity (inverse Simpson index) to quantify the variation within groups and β-diversity-based distance metric (Bray–Curtis) to quantify the variation between groups (distance). Inverse Simpson index and Bray–Curtis distance values were calculated with rarefaction using the *mothur* software package. We compared differences in the inverse Simpson index according to study groups (LS-CRC, LS-adenoma, LS-without CRN, and non-LS controls), history of surgery (including total colectomy, left hemicolectomy, right colectomy, and no surgery), and MMR pathogenetic variant (*MLH1*, *MSH2*, *MSH6*, *PSM2*, *EPCAM*) using pairwise Kruskal–Wallis test. We also modeled α-diversity (within community diversity) as the independent variable using a multivariable linear regression model with study group (LS-CRC, LS-adenoma, LS-without CRN, and non-LS controls) as the dependent variable while also controlling for age, BMI, and sex. We then compared differences in the Bray–Curtis distance by study group (LS-CRC, LS-adenoma, LS-without CRN, and non-LS controls), history of colon surgery, and MMR pathogenetic variant using pairwise PERMANOVA. To visualize the Bray–Curtis distance values, non-metric multidimensional scaling (NMDS) was calculated, and results were plotted using the vegan R package. We repeated these analyses comparing LS groups to non-LS and compared each individual LS study groups to non-LS to identify which group was driving a difference.

We then identified taxa that were differentially abundant between study groups (LS-CRC, LS-adenoma, LS-without CRN, and non-LS controls), history of colon surgery, and

among MMR pathogenic variants in patients with LS using beta-binomial regression with the Corncob package in R. This analysis models associations between covariates of interest and both the mean taxon abundance and variation in taxon abundance allowing for overdispersion, which is characteristic of microbiome data [24]. The Corncob package uses the Benjamini-Hochberg (B-H) method to adjust *p*-values in the multiple correlation analyses and controls for the expected false discovery rate at 0.05. We repeated these analyses comparing all LS groups to non-LS controls. Associations between groups and taxa with adjusted *p*-values less than 0.05 were considered statistically significant.

We then trained machine learning classification models comparing (1) LS with a history of CRC to LS without a history of CRC or adenoma, (2) LS with a history of adenoma to LS without a history of CRC or adenoma, and (3) all individuals with LS to non-LS controls. We first combined information from the individual taxa and evaluated the ability to classify LS patient using logistic regression models from the mikropml package in R [25]. We then created a model that incorporated taxa as well as key demographic risk factors including age, sex, and BMI to determine if these demographic risk factors improved the model. We computed the area under the receiver operator characteristic curve (AUROC) to measure the ability of each model (taxa alone and taxa plus demographic factors) to correctly classify subjects to their true groups. We compared the cross-validated AUROC to the testing AUROC to evaluate the generalizability of each model and to check whether the models were overfit.

Results

Baseline Characteristics of Study Participants

The 99 study participants were grouped as follows: 17 LS-adenoma, 10 LS-CRC, 19 LS-without CRN, and 53 non-LS controls. Subjects were 83% female and had a median age of 40.5 years and a median BMI of 26.6 kg/m^2 (see Table 1). LS-CRC and LS-without CRN had a lower BMI compared to LS-adenoma and non-LS controls. Sex and tobacco use were not different across these three groups. Patients with LS were older than non-LS controls. Patients with LS-CRC were more likely to have a *MLH1* and *MSH2* and less likely to have a *MSH6* gene defect.

Associations Between History of CRN and Gut Microbiome

We calculated within community diversity or alpha diversity using inverse Simpson index and found the median value was 3.62. Among individuals with LS, there was no difference when we compared LS-CRC to LS-without CRN or LS-adenoma to LS-without CRN (Fig. 1A). Because we noticed a difference between LS and non-LS, we repeated these analyses comparing each LS group (LS-CRC, LS-adenoma, LS-without CRN) to non-LS to determine which group was driving the difference. The inverse Simpson index was higher in LS-CRC relative to non-LS controls (*p*-value = 0.008) and LS-without CRN compared to non-LS

Table 1 Baseline characteristics of LS with and without CRN and non-LS controls

Characteristic	Non-LS controls, $N = 53$[a]	LS-no CRN, $N = 19$[a]	LS-adenoma, $N = 17$[a]	LS-CRC, $N = 10$[a]	*p*-value[b]
Age	36 (31, 40)	52 (38, 60)	56 (51, 66)	64 (55, 69)	<0.001
BMI	28 (24, 32)	25 (23, 28)	31 (24, 32)	24 (21, 27)	0.03
Sex					0.3
Female	47 (89%)	15 (79%)	12 (71%)	9 (90%)	
Male	6 (11%)	4 (21%)	4 (24%)	1 (10%)	
Surgery					<0.001
No surgery		19 (100%)	17 (100%)	0 (0%)	
L hemicolectomy		0 (0%)	0 (0%)	2 (20%)	
R hemicolectomy		0 (0%)	0 (0%)	2 (20%)	
Colectomy		0 (0%)	0 (0%)	6 (60%)	
Gene					0.01
EPCAM		1 (5.3%)	1 (5.9%)	1 (10%)	
MLH1		3 (16%)	4 (24%)	3 (30%)	
MSH2		2 (11%)	7 (41%)	3 (30%)	
MSH6		11 (58%)	5 (29%)	0 (0%)	
PMS2		2 (11%)	0 (0%)	3 (30%)	

[a]Median (IQR); *n* (%)

[b]Kruskal-Wallis rank sum test; Fisher's exact test

controls (p-value = 0.007), but there was no difference between LS-adenoma and non-LS controls. There was a significant difference in inverse Simpson index according to history of colon surgery (p-value = 0.02).

PERMANOVA analyses were performed to compare the overall community structure of the gut microbiota of all samples based on the OTU relative abundance (Fig. 2). When we performed PERMANOVA analysis among patients with LS, we found no differences in beta diversity or community distance according to history of adenoma (p-value = 0.64) or CRC (p-value = 0.11) compared to without CRN. We did find a difference in community structure according to a history of colon surgery (p-value < 0.001); however, there were only 2 people with a history of left hemicolectomy and 2 people with history of right hemicolectomy. These small numbers could have made the community structures appear more different because of sampling error due to smaller numbers not representing intraindividual heterogeneity of these populations.

Variation in community structure was associated with LS status compared to non-LS controls (Fig. 2), even when we corrected for age (p-value < 0.001). Again, we wanted to determine which LS group was responsible for the observed difference between the community structures in LS and non-LS. Therefore, we repeated the PERMANOVA analyses comparing each LS group (LS-CRC, LS-adenoma, LS-without CRN) to non-LS. We found a significant difference in community structure when we compared LS-CRC to non-LS controls (p-value < 0.001) and LS-adenoma to non-LS controls (p-value = 0.003), but no difference between LS-without CRN to non-LS controls (p-value = 0.11). The difference in community structure between LS and non-LS was driven by LS-CRC and LS-adenoma and not LS-without CRC. In converse, we found no differences in community structure when we compared by pathogenic variant groups among LS (Fig. 2).

Differentially Abundant Taxa by History of Adenoma, Colon Cancer, Colon Surgery, and LS Status

Veillonella was depleted in LS-adenoma compared to LS-without CRN and is the only taxon difference between these groups, whereas *Veillonella*, *Streptococcus*, and *Actinomyces* were differentially enriched in LS-CRC compared to LS-without CRN (Figs. 3 and 4). Notably, we identified the same differentially abundant genera associated with LS with history of colon surgery compared to LS without surgery, *Streptococcus* and *Actinomyces* (Figs. 5 and 6). We found several taxa that differed between LS and non-LS controls (Fig. 4). Compared to non-LS controls, *Veillonella*, 2 members of *Streptococcus*, *Lactococcus*, *Lachnospiraceae*, and *Clostridium* XIVa were enriched and 2 members of *Faecalibacterium*, *Ruminococcus*, and *Bacteroides* were depleted in total LS. When we looked separately at each group compared to non-LS controls, *Veillonella* and *Lachnospiraceae* were enriched in LS-without CRN and *Romboutsia* and *Faecalibacterium* were depleted (Fig. 4). *Parabacteroides*, *Veillonell*a, and *Erysipelotrichaceae* were depleted and unclassified

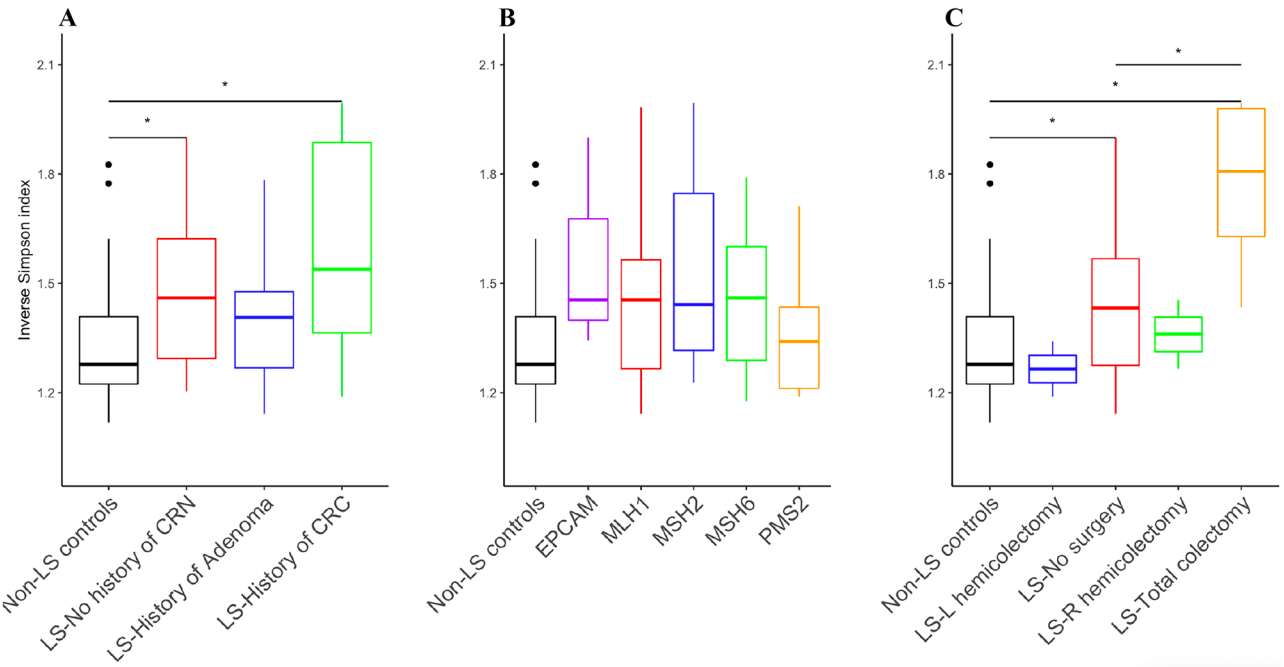

Fig. 1 Alpha diversity by **A** group status, **B** MMR pathogenic variant, and **C** surgery history

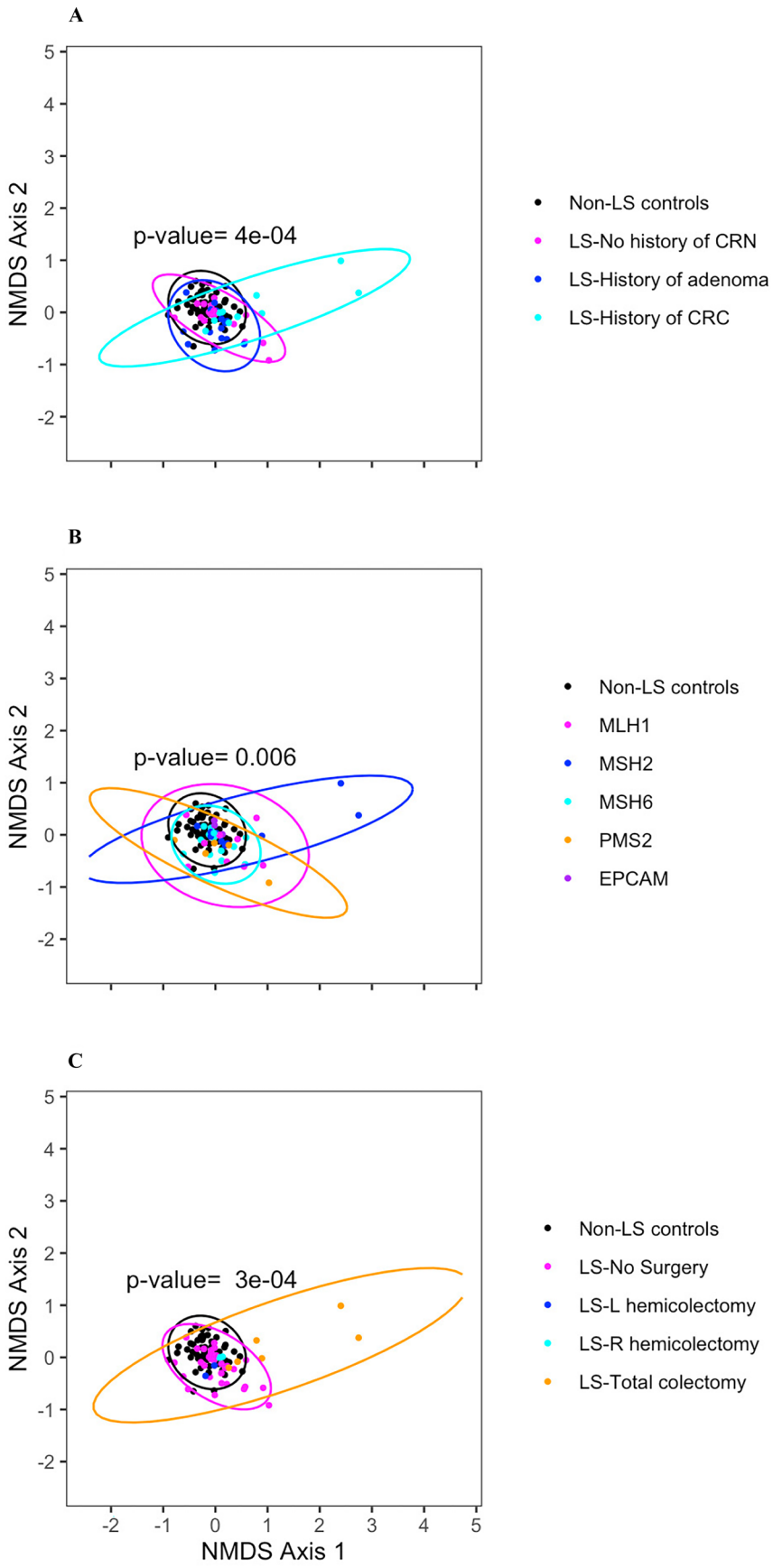

Fig. 2 Beta diversity (between community) assessed with PERMANOVA by **A** group status, **B** MMR pathogenic variant, and **C** surgery history

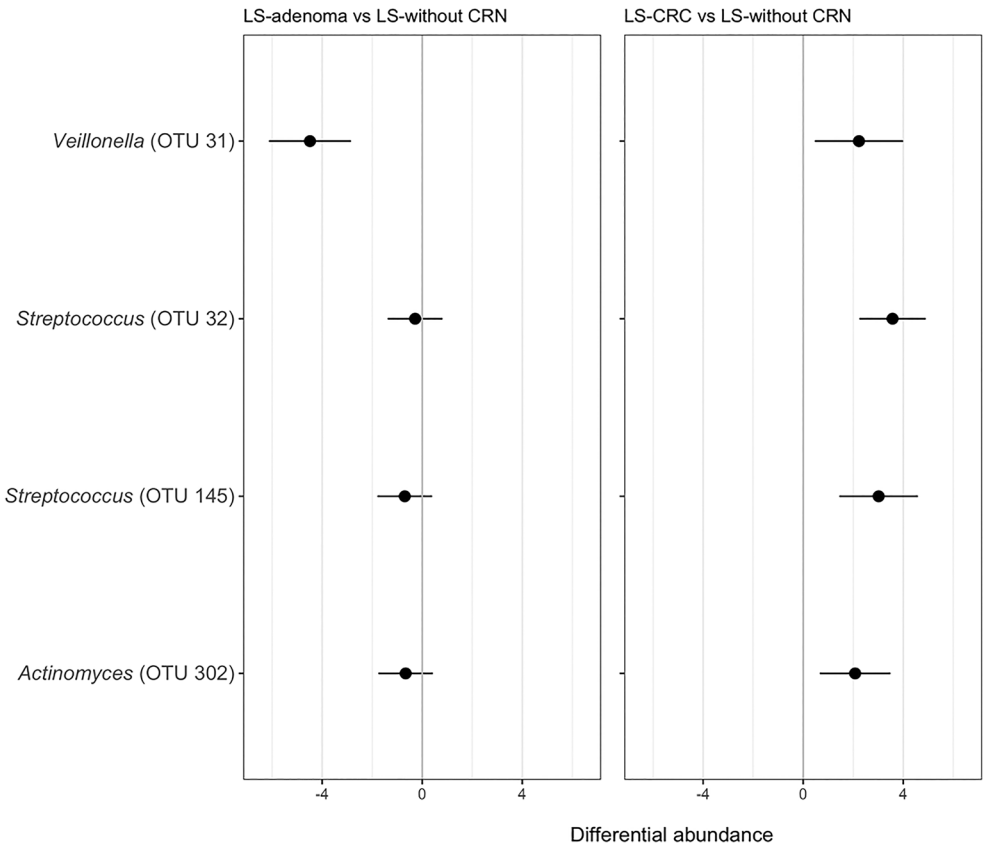

Fig. 3 Differential abundance of OTUs by study group among LS assessed with beta-binomial regression

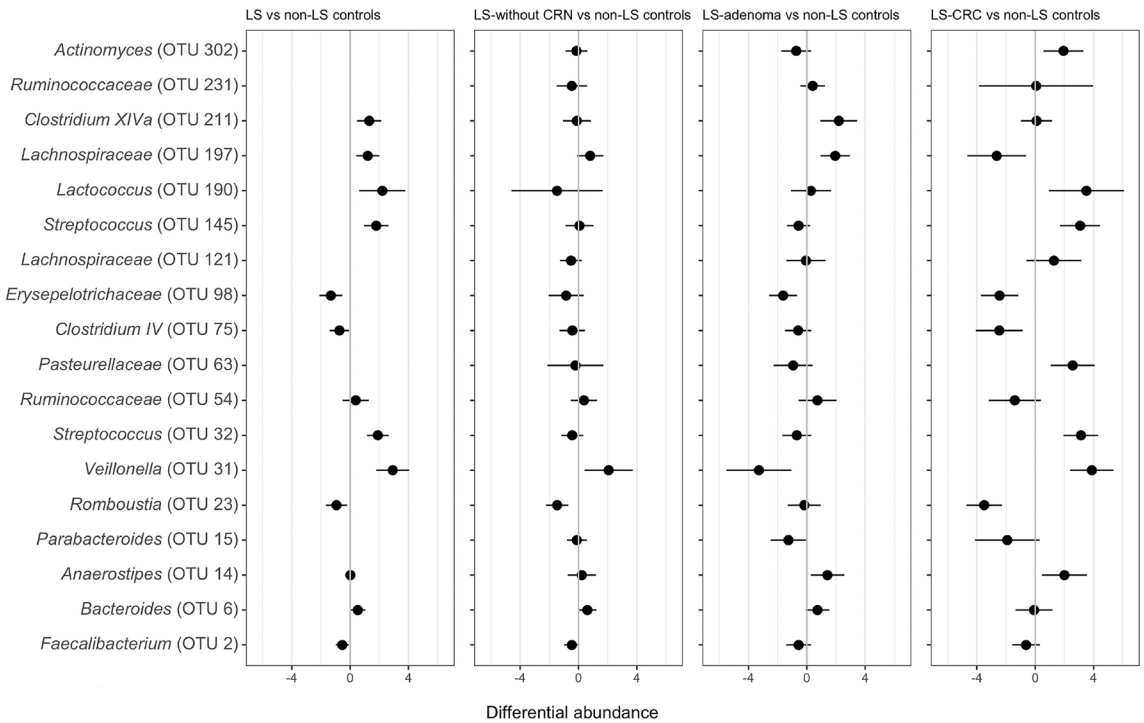

Fig. 4 Differential abundance of OTUs by study group assessed with beta-binomial regression

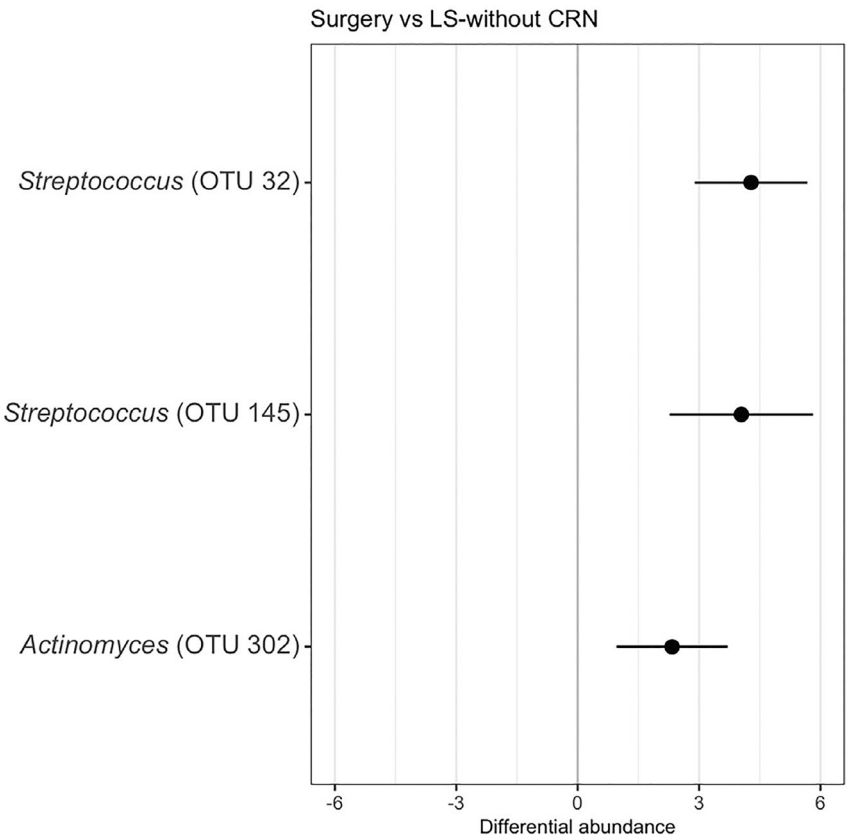

Fig. 5 Differential abundance among LS by history of colon surgery assessed with beta-binomial regression: any surgery versus no surgery

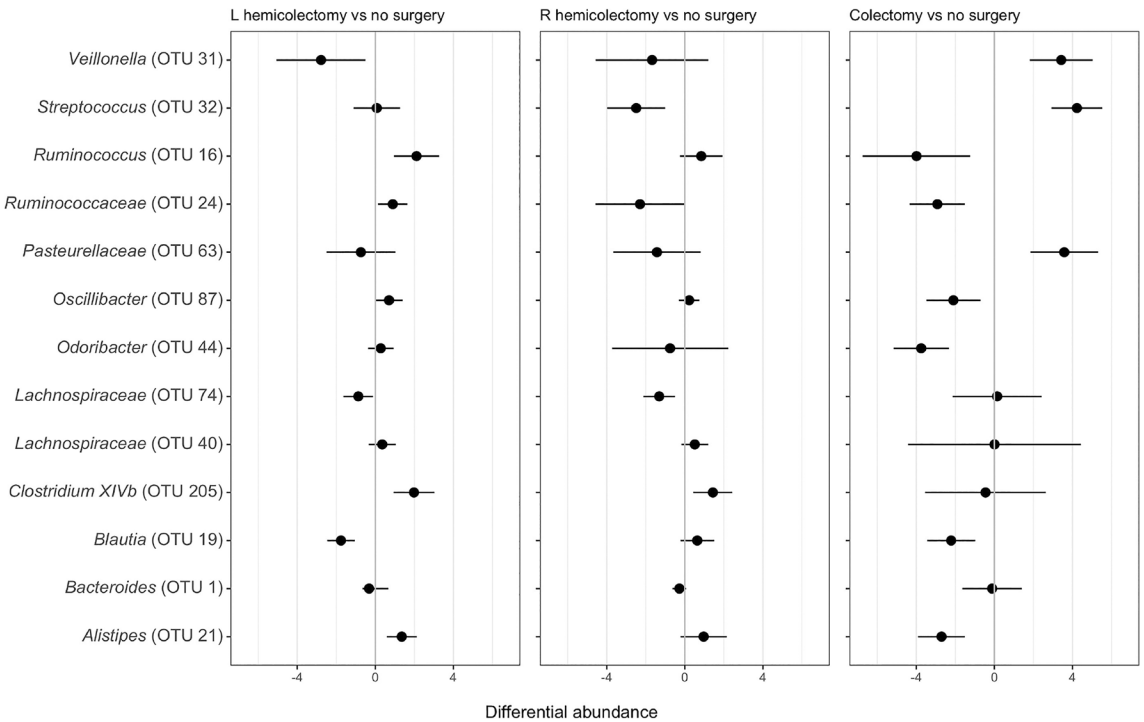

Fig. 6 Differential abundance among LS by history of colon surgery assessed with beta-binomial regression

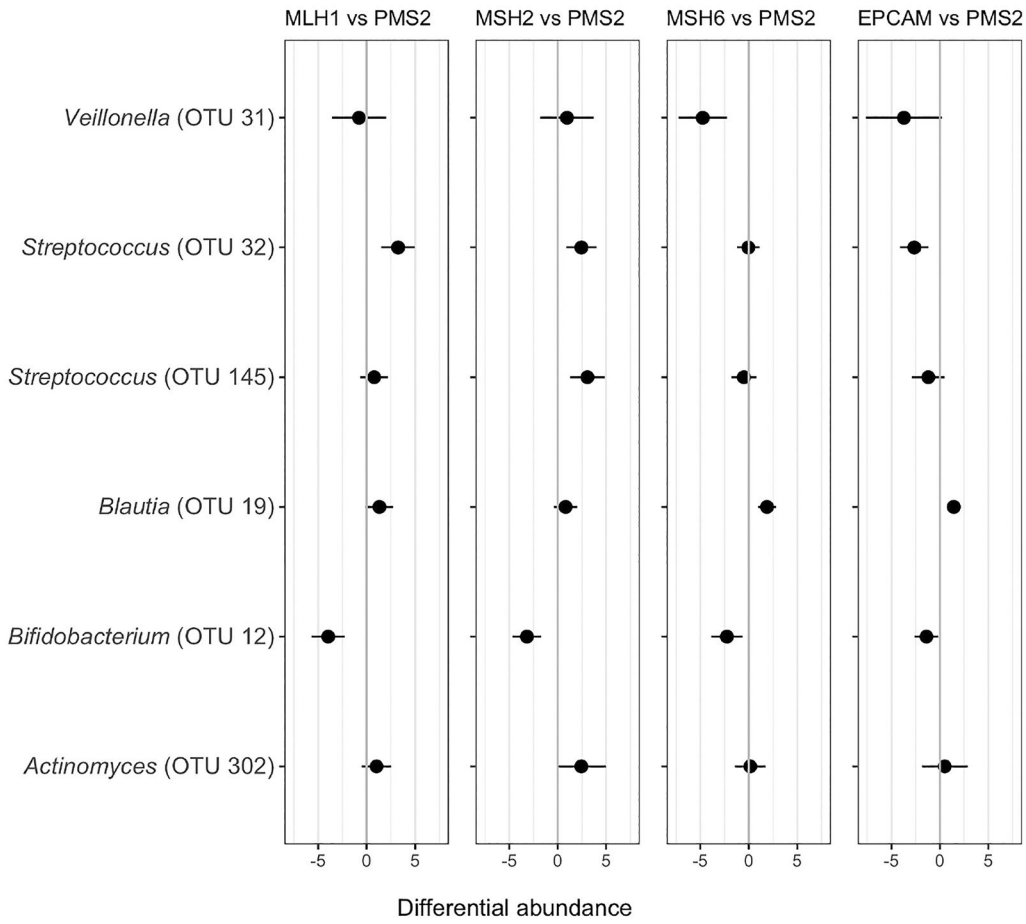

Fig. 7 Differential abundance of OTUs by LS pathogenic variant assessed with beta-binomial regression

Lachnospiraceae and *Clostridium* XIVa were enriched in LS-adenoma (Fig. 4). *Veillonella*, *Pasteurellaceae*, *Streptococcus*, *Lactococcus*, and *Actinomyces* were enriched and *Parabacteroides*, *Romboutsia*, *Clostridium* IV, *Erysipelotrichaceae*, and *Lachnospiraceae* were depleted in LS-without CRN. We found taxon abundance differences by pathogenic variant groups. *MLH1*, *MSH2*, and *MSH6* were depleted in *Bifidobacterium*, while *MSH2* was enriched in *Actinomyces* and both *MLH1* and *MS2* were enriched in *Streptococcus* compared to *PSM2* (Fig. 7).

Classification of CRC and Adenoma History Using OTUs in a Machine Learning Model

We compared the cross-validation AUROC to the testing AUROC for each model (Fig. 8). The testing models were performed as well as the training models. Therefore, we were able to conclude that the models were not overfit (Table 2). The model comparing LS-CRC to LS-without CRN built with taxa and demographic factors (AUROC = 0.833, IQR: 0.5, 1.0) performed better than taxa alone (AUROC = 0.667, IQR: 0.5, 1.0). However, the confidence intervals were large and overlapped due to the small number of individuals included in these models. The model designed to classify LS-adenoma from LS-without CRN performed poorly (AUROC = 0.50, IQR: 0.33, 0.67). The model performance improved slightly when demographic factors were added (AUROC = 0.56, IQR: 0.44, 0.88). The models built to classify LS and non-LS controls performed well (AUROC = 0.69, IQR: 0.66, 0.72) and there was no change in AUROC when demographics were added to the model (AUROC = 0.68, IQR: 0.66, 0.70). The models that included both taxa and demographic factors demonstrated improved performance classifying LS according to history of CRN compared to taxa only models. Models built to distinguish LS from non-LS controls showed no improvement in model performance after adding demographics factors.

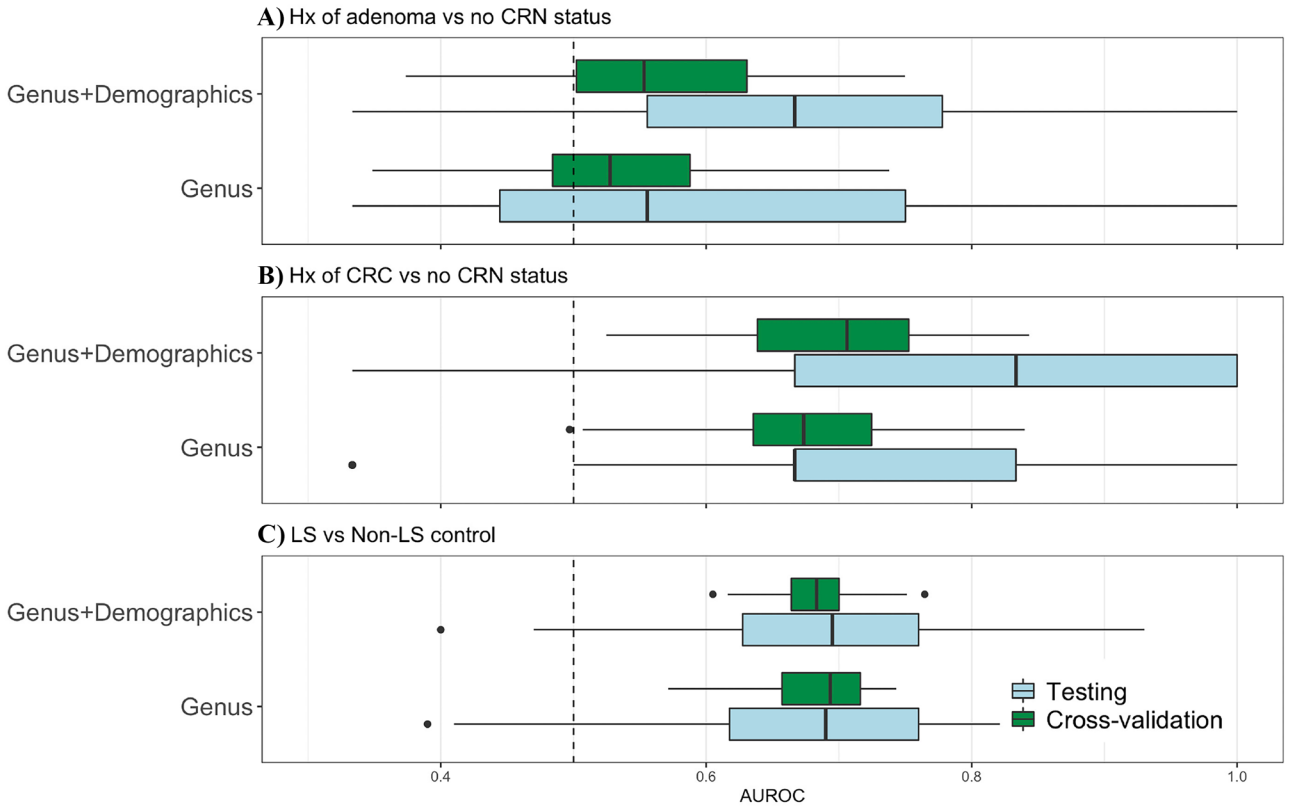

Fig. 8 Performance of machine learning models as measured by area under the ROC curve (AUROC) of **A** Hx of adenoma vs no Hx of CRN, **B** Hx of CRC vs no Hx of CRN, and **C** LS vs. non-LS controls

Table 2 The training and testing AUC for models classifying LS-CRC from LS-without CRC, LS-adenoma from LS-without CRC, and LS from non-LS controls

Machine learning models	Training[a]	Testing[a]
LS-CRC vs LS-without CRN		
Genus	0.68 (0.64, 0.71)	0.67 (0.5, 1.0)
Genus + demographic	0.69 (0.64, 0.74)	0.83 (0.5, 1.0)
LS-adenoma vs LS-without CRN		
Genus	0.53 (0.48, 0.59)	0.50 (0.33, 0.67)
Genus + demographic	0.55 (0.50, 0.63)	0.56 (0.44, 0.78)
LS vs non-LS controls		
Genus	0.69 (0.68, 0.70)	0.69 (0.66, 0.72)
Genus + demographic	0.67 (0.67,0.70)	0.68 (0.66, 0.70)

[a]Median (IQR)

Conclusion

Our study found significant differences in the microbiome between LS with or without CRN and non-Lynch controls and few between LS with CRC from LS-without CRN. We found no global differences (alpha or beta diversity) between LS with CRN (either CRC or adenoma) and without CRN. There were taxon-specific differences comparing individuals with LS and history of CRC and LS without any CRN, *Streptococcus* and *Actinomyces*. Notably, two *Streptococcus* species, *Streptococcus bovis* and *Streptococcus gallolyticus*, have been frequently linked with colon cancer [26]. Additionally, *Actinomyces* has been associated with CRN in a systematic review [27]. However, we also found these same changes when we evaluated by history of surgery. Post-surgical anatomy may be an overriding factor confounding the association between history of colon cancer and microbiome composition as all individuals with a history of CRC underwent colorectal surgery.

While there were no global differences among LS comparing CRN to no history, we found both global and multiple taxon level differences between LS and non-LS controls. LS carriers had a higher alpha diversity compared to non-LS controls. Similarly, there were significant differences in community structure. *Veillonella* has been associated with inflammatory conditions including Crohn's disease and hepatic encephalopathy and was enriched in LS. LS carriers can also have dysplastic crypts that may provoke chronic low-level inflammation despite being histologically undetectable. Lower abundances of bacteria

regularly associated with being anti-inflammatory including *Faecalibacterium* and *Romboutsia* were also depleted in LS [28]. These findings could be due to the significant differences in age between these groups. *Veillonella* was enriched in both LS-without CRN and history of CRC compared to non-LS controls.

Because we observed more differences in microbiome composition comparing LS to non-LS controls than comparing LS groups to each other, it was unexpected that the machine learning models classifying (1) LS from non-LS controls and (2) LS-CRC from LS-without CRN performed equally well. The differences in microbiome composition between LS and non-LS controls might suggest that underlying microbiome composition of LS carriers is different potentially due to underlying genetic pathogenetic variants leading to changes in the colorectal epithelium or mucosal immunity. We observed fewer differences in microbiome composition among LS groups according to history of CRN, but the LS-CRC microbiome was different enough from LS-without CRN to allow a machine model to distinguish between these two groups. The observed differences between LS-CRC and LS-without CRN and the ability of our machine learning model to decipher the two groups might also be due to history of colon resection, which was associated with similar differences in microbiome composition. Conversely, we did not observe either global or taxon level differences between LS-adenoma and LS-without CRN and furthermore the model we built to classify these two groups did not perform well.

Other studies have similarly found that the microbiome compositions of subgroups of LS according to CRN history were like each other but different compared to the microbiome composition of non-LS controls. Mori et al. did not observe a difference between LS with CRC and LS with endometrial cancer (GC) but noted a statistically significant difference in the community structure between non-LS controls compared to LS. The authors hypothesized there might be an underlying fecal microbiota pattern associated with LS [16]. Similarly, Lu et al. found no difference in microbiome composition among LS carriers with and without cancer, but observed that LS carriers were enriched in *B. fragilis* and *Parabacteroides distasonis*, and Pseudomonadaceae family compared to non-LS controls [17]. Ferrarese et al. found stool microbiome from LS carriers was enriched in Bacteroidetes and Proteobacteria and depleted in in the Firmicutes and *Ruminococcaceae* compared to non-LS controls [18].

Fewer studies have noted differences in microbiome composition among patients with LS with adenomas and CRC compared to LS with no CRN. Gonzalez et al. found non-significant difference in several genera when they compared LS with cancer to LS without cancer, but sample size was very limited ($n = 8$) [29]. In the largest study to date, Yan et al. found that *Clostridiaceae* was depleted and *Desulfovibrio* was enriched in LS with baseline adenomas compared to no adenomas at baseline. They also observed that history of surgery was the most significant contributor to differences in microbiome composition [15].

Only one other study explored whether different LS pathogenic variants were characterized by different gut microbial populations. Yan et al. found that *MLH1* and *MSH2* mutation carriers were depleted in *Clostridiales* and MLH1 were enriched in *Blautia* and *Oscillospira* [15]. We found that *MLH1*, *MSH2*, and *MSH6* were depleted in *Bifidobacterium* while *MLH1* and *MS2* were enriched in *Actinomyces* and *MSH2* was enriched in *Streptococcus* compared to *PSM2.* As noted above, we found similar differences in *Streptococcus* and *Actinomyces* when we evaluated by history of surgery. Post-surgical anatomy may also be confounding the association between LS pathogenic variant and microbiome composition as individuals with *MLH1* and *MSH2* were more likely to be diagnosed with cancer and require curative colorectal surgery compared to *PSM2*.

While correlation studies in humans have not yet provided evidence that microbiome contribute to cancer development in LS, LS animal models provide evidence that the microbiome may drive carcinogenesis in LS. A study conducted in a *MSH2* Lynch mouse model ($APC^{Min}/MSH2^{-/-}$) demonstrated that bacterial-derived butyrate might drive hyperproliferation in MMR-deficient cecal epithelial cells marked by deregulated beta-catenin activity [30]. By treating these mice with antibiotics or a diet low in carbohydrates, the study authors reduced both butyrate levels and the total number of polyps by 75%. This finding may explain the lack of any protective effect observed in LS carriers randomized to resistant starch despite increasing their levels of butyrate concentration in the CAPP2 randomized control trial in contrast to the protective effect observed in sporadic CRC [31]. Another *MSH2* knockout mouse model exposed to conventional microbiome exhibited increased epithelial turnover rates, increased rate of spontaneous mutations in MSH2-deficient crypts, and increased microsatellite instability compared to specific pathogen free (SPF mice) [32]. The authors hypothesized that bacterial presence in MMR-deficient crypts drives epithelial turnover and subsequent mutations in DNA.

One strength of this study is that it is the largest study of LS to include non-LS controls for comparisons. Prior studies comparing LS patient to non-LS controls were based on small cohorts with less than 10 subjects in each groups making their findings subject to internal and external validity issues. Our similar findings of notable differences between LS and controls with a larger sample size strengthen and confirm this repeated finding.

There were several limitations to our study. The controls were significantly younger compared to patients with

LS, which might confound the differences in microbiome composition we observed between these two groups. Age is an important contributing factor in microbiome variation. Furthermore, the study populations were female predominant, and therefore, our results might not be generalizable. Additionally, we did not collect stool prospectively and our investigation of differences in microbiome composition within LS carriers was based on prior diagnoses of colorectal adenomas and cancer. The stool microbiome is more significant and informative if collected prior to adenoma detection as our hypothesis is that microbiome composition may lead to the risk of polyp and cancer formation. As this was a cross-sectional study, stool collected might not accurately reflect the composition at the time adenoma or cancer was forming. Another limitation was that the sample size was still too small to do multigroup analyses and adjust for covariates. Larger studies with prospective stool collection are needed to untangle how microbiome protects against or contributes to carcinogenesis in LS.

In conclusion, we found significant global and specific differences in the microbiome composition when we compared individuals with LS to non-LS controls. This finding confirms observations from prior studies done with significantly smaller sample sizes. Further studies investigating how differences in the epithelial biology and immunology between LS carriers and non-LS controls might shape differences in the microbiome composition are warranted. We found no global differences and few specific taxa differences when we compared LS-CRC to LS-without CRN. While animal models suggest microbiome may contribute to colorectal carcinogenesis, no clinical studies in people have documented an association between CRC and microbiome in LS. A longitudinal collection of stool microbiome in LS to detect the contribution of stool microbiome composition to carcinogenesis in LS would be invaluable to investigating the role of microbiome composition in CRC development.

Acknowledgements We thank Katherine McBride who helped process the samples.

Author Contribution MAS and PDS were involved in the conception and design of the study. KT, EMS, and SBR were involved in sample collection. SBR processed the samples and analyzed the data. All authors interpreted the data. SBR and PDS wrote the manuscript. All authors reviewed and revised the manuscript. All authors read and approved the final manuscript.

Availability of Data and Materials Datasets are available on reasonable request.

Declarations

Competing Interests The authors declare no competing interests.

References

1. Møller P, 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. https://doi.org/10.1136/gutjnl-2015-309675.
2. Tjalsma H, Boleij A, Marchesi JR, Dutilh BE. A bacterial driver-passenger model for colorectal cancer: beyond the usual suspects. Nat Rev Microbiol. 2012;10(8):575–82. https://doi.org/10.1038/nrmicro2819.
3. Wu S, et al. A human colonic commensal promotes colon tumorigenesis via activation of T helper type 17 T cell responses. Nat Med. 2009;15(9):1016–22. https://doi.org/10.1038/nm.2015.
4. Dejea CM, et al. Patients with familial adenomatous polyposis harbor colonic biofilms containing tumorigenic bacteria. Science. 2018;359(6375):592–7. https://doi.org/10.1126/science.aah3648.
5. Arthur JC, et al. Intestinal inflammation targets cancer-inducing activity of the microbiota. Science. 2012;338(6103):120–3. https://doi.org/10.1126/science.1224820.
6. De Almeida CV, de Camargo MR, Russo E, Amedei A. Role of diet and gut microbiota on colorectal cancer immunomodulation. World J Gastroenterol. 2019;25(2):151–62. https://doi.org/10.3748/wjg.v25.i2.151.
7. Zackular JP, Rogers MAM, Ruffin MT, Schloss PD. The human gut microbiome as a screening tool for colorectal cancer. Cancer Prev Res Phila Pa. 2014;7(11):1112–21. https://doi.org/10.1158/1940-6207.CAPR-14-0129.
8. Drewes JL, et al. High-resolution bacterial 16S rRNA gene profile meta-analysis and biofilm status reveal common colorectal cancer consortia. NPJ Biofilms Microbiomes. 2017;3:34. https://doi.org/10.1038/s41522-017-0040-3.
9. Flemer B, et al. Tumour-associated and non-tumour-associated microbiota in colorectal cancer. Gut. 2017;66(4):633–43. https://doi.org/10.1136/gutjnl-2015-309595.
10. Brim H, et al. Microbiome analysis of stool samples from African Americans with colon polyps. PloS One. 2013;8(12):e81352. https://doi.org/10.1371/journal.pone.0081352.
11. Lu Y, et al. Mucosal adherent bacterial dysbiosis in patients with colorectal adenomas. Sci Rep. 2016;6(26337). https://doi.org/10.1038/srep26337.
12. Levy M, Kolodziejczyk AA, Thaiss CA, Elinav E. Dysbiosis and the immune system. Nat Rev Immunol. 2017;17(4):219–32. https://doi.org/10.1038/nri.2017.7.
13. Zhang Y, et al. Changes in gut microbiota and plasma inflammatory factors across the stages of colorectal tumorigenesis: a case-control study. BMC Microbiol. 2018;18(1):92. https://doi.org/10.1186/s12866-018-1232-6.
14. Sze MA, Baxter NT, Ruffin MT, Rogers MAM, Schloss PD. Normalization of the microbiota in patients after treatment for colonic lesions. Microbiome. 2017;5(1):150. https://doi.org/10.1186/s40168-017-0366-3.
15. Yan Y, et al. Structure of the mucosal and stool microbiome in Lynch syndrome. Cell Host Microbe. 2020;27(4):585–600. https://doi.org/10.1016/j.chom.2020.03.005.
16. Mori G, et al. Gut microbiota analysis in postoperative Lynch syndrome patients. Front Microbiol. 2019;10:1746. https://doi.org/10.3389/fmicb.2019.01746.
17. Lu X, Kang Q, Jin P, Sheng J. The interactions between gut microbiota and Lynch syndrome. Clin Gastroenterol Hepatol. 2017;15(1):e48. https://doi.org/10.1016/j.cgh.2016.09.116.
18. Ferrarese R, et al. Oral and fecal microbiota in Lynch syndrome. J Clin Med. 2020;9(9):E2735. https://doi.org/10.3390/jcm9092735.
19. Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the

MiSeq Illumina sequencing platform. Appl Environ Microbiol. 2013;79(17):5112–20. https://doi.org/10.1128/AEM.01043-13.
20. Schloss PD, et al. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl Environ Microbiol. 2009;75(23):7537–41. https://doi.org/10.1128/AEM.01541-09.
21. Pruesse E, et al. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. Nucleic Acids Res. 2007;35(21):7188–96. https://doi.org/10.1093/nar/gkm864.
22. Wang Q, Garrity GM, Tiedje JM, Cole JR. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Appl Environ Microbiol. 2007;73(16):5261–7. https://doi.org/10.1128/AEM.00062-07.
23. Westcott SL, Schloss PD. OptiClust, an improved method for assigning amplicon-based sequence data to operational taxonomic units. mSphere 2017;2(2). https://doi.org/10.1128/mSphereDirect.00073-17.
24. Martin BD, Witten D, Willis AD. Modeling microbial abundances and dysbiosis with beta-binomial regression. Ann Appl Stat. 2020;14(1):94–115. https://doi.org/10.1214/19-aoas1283.
25. Topçuoğlu BD, Lesniak NA, Ruffin MT, Wiens J, Schloss PD. A framework for effective application of machine learning to microbiome-based classification problems. mBio. 11(3):e00434–20. https://doi.org/10.1128/mBio.00434-20.
26. Andres-Franch M, et al. Streptococcus gallolyticus infection in colorectal cancer and association with biological and clinical factors. PloS One. 2017;12(3):e0174305. https://doi.org/10.1371/journal.pone.0174305.
27. Yu L, et al. A systematic review of microbial markers for risk prediction of colorectal neoplasia. Br J Cancer. 2022;1–11. https://doi.org/10.1038/s41416-022-01740-7.
28. Mangifesta M, et al. Mucosal microbiota of intestinal polyps reveals putative biomarkers of colorectal cancer. Sci Rep. 2018;8(1):13974. https://doi.org/10.1038/s41598-018-32413-2.
29. Gonzalez A, Kapila N, Melendez-Rosado J, Liang H, Castro-Pavia F. An evaluation of the fecal microbiome in Lynch syndrome. J Gastrointest Cancer. 2021;52(1):365–8. https://doi.org/10.1007/s12029-021-00588-z.
30. Belcheva A, et al. Gut microbial metabolism drives transformation of MSH2-deficient colon epithelial cells. Cell. 2014;158(2):288–99. https://doi.org/10.1016/j.cell.2014.04.051.
31. Mathers JC, 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. https://doi.org/10.1016/S1470-2045(12)70475-8.
32. Pieters W, et al. Pro-mutagenic effects of the gut microbiota in a Lynch syndrome mouse model. Gut Microbes. 2022;14(1):2035660. https://doi.org/10.1080/19490976.2022.2035660.