



OPEN

## Alterations in co-abundant bacteriome in colorectal cancer and its persistence after surgery: a pilot study

Chin-Wen Png<sup>1,4</sup>, Yong-Kang Chua<sup>2,4</sup>, Jia-Hao Law<sup>2</sup>, Yongliang Zhang<sup>1</sup> & Ker-Kan Tan<sup>2,3</sup>✉

There is growing interest in the role of gut microbiome in colorectal cancer (CRC), ranging from screening to disease recurrence. Our study aims to identify microbial markers characteristic of CRC and to examine if changes in bacteriome persist after surgery. Forty-nine fecal samples from 25 non-cancer (NC) individuals and 12 CRC patients, before and 6-months after surgery, were collected for analysis by bacterial 16S rRNA gene sequencing. Bacterial richness and diversity were reduced, while pro-carcinogenic bacteria such as *Bacteroides fragilis* and *Odoribacter splanchnicus* were increased in CRC patients compared to NC group. These differences were no longer observed after surgery. Comparison between pre-op and post-op CRC showed increased abundance of probiotic bacteria after surgery. Concomitantly, bacteria associated with CRC progression were observed to have increased after surgery, implying persistent dysbiosis. In addition, functional pathway predictions based on the bacterial 16S rRNA gene data showed that various pathways were differentially enriched in CRC compared to NC. Microbiome signatures characteristic of CRC comprise altered bacterial composition. Elements of these dysbiotic signatures persists even after surgery, suggesting possible field-change in remnant non-diseased colon. Future studies should involve a larger sample size with microbiome data collected at multiple time points after surgery to examine if these dysbiotic patterns truly persist and also correlate with disease outcomes.

The human microbiome is a second genome of the human body where microorganisms establish a symbiotic relationship with the host in complementing metabolic deficits, protection from invading pathogens and maintaining immune homeostasis<sup>1</sup>. There is increasing evidence showing that gut microbiota influenced CRC development through the virulence factors of pathogenic bacterial strains, and recent reviews and papers have suggested a close correlation between certain microbial strains and CRC<sup>2–6</sup>. Bacteria such as *Bacteroides fragilis*, *Fusobacterium nucleatum*, and *Escherichia coli* have been identified to be enriched in patients with CRC<sup>7,8</sup>. Various molecular events as well as changes to the tumor microenvironment which are related to the enrichment of specific bacterial species were found to promote CRC tumorigenesis in various animal and *ex-vivo* models<sup>9–11</sup>. These specific microbial strains associated with CRC provided promising opportunities to develop diagnostic tools or treatment biomarkers for CRC.

However, the relationship between the bacteriome and CRC appear complex as there is unlikely a single pathogenic bacteria that is related to CRC, unlike the direct causal relationship that is established between *Helicobacter pylori* and gastric cancer<sup>12</sup>. Different microbial strains that are associated with CRC have been reported, some of which were not reproducible in other studies<sup>13</sup>. For example, Dejea et al. found that there was no consistent bacterial genus associated with tumors by high-throughput sequencing in 30 CRC and 6 adenoma human samples<sup>13</sup>. In another study, the widely studied *Fusobacterium* that is increasingly associated with CRC was only found to be elevated in 20–30% of CRC patients and was not a consistent finding<sup>14</sup>. On the other hand, other studies found differing microbes that triggered neoplasia and also proposed associated mechanisms. For example, *Bacteroides fragilis* was found to promote Th17 development which limits the availability of IL2 in the

<sup>1</sup>Department of Microbiology and Immunology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore. <sup>2</sup>Division of Colorectal Surgery, Department of Surgery, National University Hospital, Singapore, Singapore. <sup>3</sup>Department of Surgery, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore. <sup>4</sup>These authors contributed equally: Chin-Wen Png and Yong-Kang Chua. ✉email: surtkk@nus.edu.sg

	Cancer (% or SD) n = 12	Healthy (% or SD) n = 25
Age, mean	63.8 (± 9.3)	61.6 (± 8.9)
<b>Gender</b>		
Male	7 (58.3%)	14 (56.0%)
Female	5 (41.7%)	11 (44.0%)
Body mass index, mean	25.8 (± 4.0)	25.2 (± 4.4)
Adjuvant chemotherapy	7 (58.3%)	–

**Table 1.** Overview of participant clinical characteristics.

Subject	Age	Gender	Race	Tumor location	TNM stage	Surgery
M02-P	57	Male	Malay	Sigmoid	3	Anterior resection
M03-P	74	Female	Chinese	Descending	3	Left hemicolectomy
M04-P	56	Female	Malay	Sigmoid	1	Anterior resection
M06-P	60	Female	Malay	Sigmoid	3	Anterior resection
M07-P	67	Male	Chinese	Sigmoid	2	Anterior resection
M08-P	46	Male	Chinese	Rectum	3	Anterior resection
M09-P	59	Female	Indian	Rectum	2	Anterior resection
M11-P	65	Male	Chinese	Sigmoid	2	Anterior resection
M12-P	65	Male	Indian	Rectosigmoid	1	Anterior resection
M13-P	69	Male	Chinese	Sigmoid	3	Anterior resection
M15-P	82	Female	Chinese	Rectosigmoid	3	Anterior resection
M16-P	66	Male	Chinese	Sigmoid	1	Anterior resection

**Table 2.** Clinical data of cancer group patients.

local microenvironment, while *Escherichia coli* releases colibactin which is a genomic product of polyketide synthase island that is carcinogenic and promotes CRC development<sup>6,15</sup>.

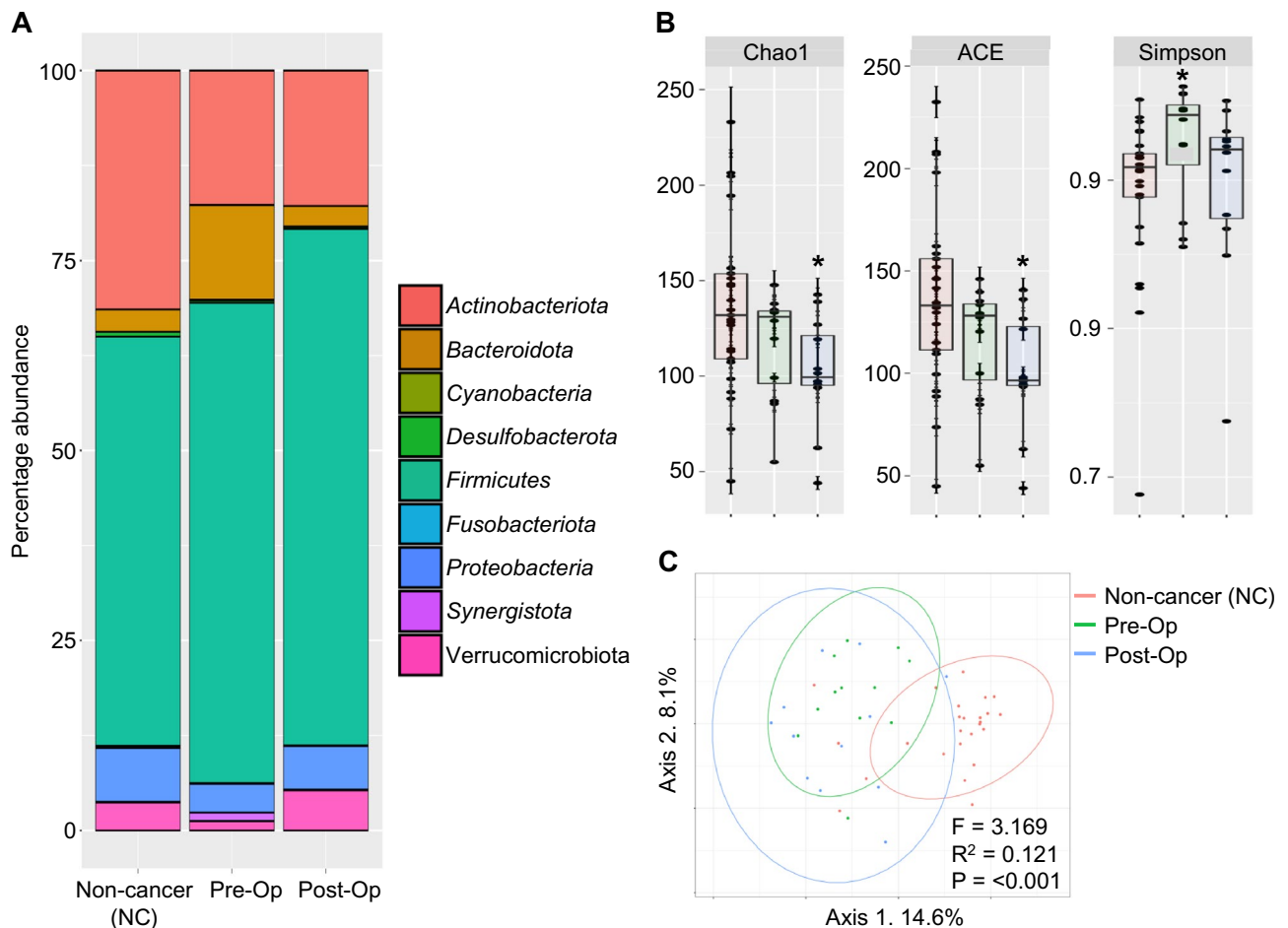
A common finding in all studies was the alterations or imbalance in gut microbial composition unique to CRC patients, which was otherwise known as microbial dysbiosis. As Katarzyna et al. highlighted, microbial “dysbiosis” is a loosely defined term that can refer to microbial compositional changes, imbalance, or alterations in specific microbial lineages. It can refer to taxonomic or functional dysbiosis<sup>16</sup>. Dysbiosis may not be adequate to the task of establishing microbiota causality in disease as such causal relationships need to be made on firm scientific basis and fulfill the “commensal Koch’s postulates”<sup>17,18</sup>. However, existing literature suggests that the initiation and development of CRC appear to be fueled by a collective microbial dysbiosis with interactions amongst numerous microbes, rather than propelled by specific pathogenic microbes.

It is also inconclusive if the state of microbial dysbiosis associated with CRC continues to persist after curative surgery. Few studies have investigated the microbial alterations after curative CRC surgery. Kong et al. reported a reduced ratio of *Bacteroidetes/Firmicutes* after surgery, which could contribute to intestinal inflammation<sup>19</sup>. Tumor-associated microbes including *Enterococcus* and *Fusobacterium* were also found to be reduced. However, beneficial obligate anaerobes such as *Prevotella* were also reduced postoperatively. It is postulated that if the microbial dysbiosis is associated with the cancer, then resection of the tumour should not only remove the lesion but also reverse the dysbiosis that accompanied CRC. On the other hand, persistent dysbiosis after surgery may suggest possible field change in the “non-diseased” colon that predisposed the patient to CRC in the first place.

In this study, we thus seek to examine alterations in co-abundant bacterial strains in CRC patients before (pre-op) and after (post-op) surgery and to also compare them to non-cancer (NC) individuals. Using 16S rRNA gene sequencing, we first performed the classical community analysis and statistical tests to compare the gut microbial community structure and composition between CRC patients before and after surgery and NC individuals. We then applied functional network analysis to further examine the differences among them.

## Results

**Study population.** In this study, we have included a total of 37 individuals, of which 12 were CRC patients and 25 were NC controls. Using T statistics (non-centrality parameter) power calculation, the number of samples included in this study provide sufficient power to detect effect size of 2, at 95% confidence with 80% power<sup>26</sup>. The participant and clinical data are summarized in Tables 1 and 2, respectively. Bacterial 16S rRNA gene sequencing of 49 fecal samples from CRC patients and NC controls were performed to investigate the compositional changes of the microbiome in CRC patients pre- and post-operatively (pre- and post-op). A total of 2,559,498 filtered sequences were obtained with an average of 52,234 sequences per sample (range 3531–98,096). A rarefaction curve for richness is plotted to analyse the sequencing depth of the samples (Supplementary Fig. 1).

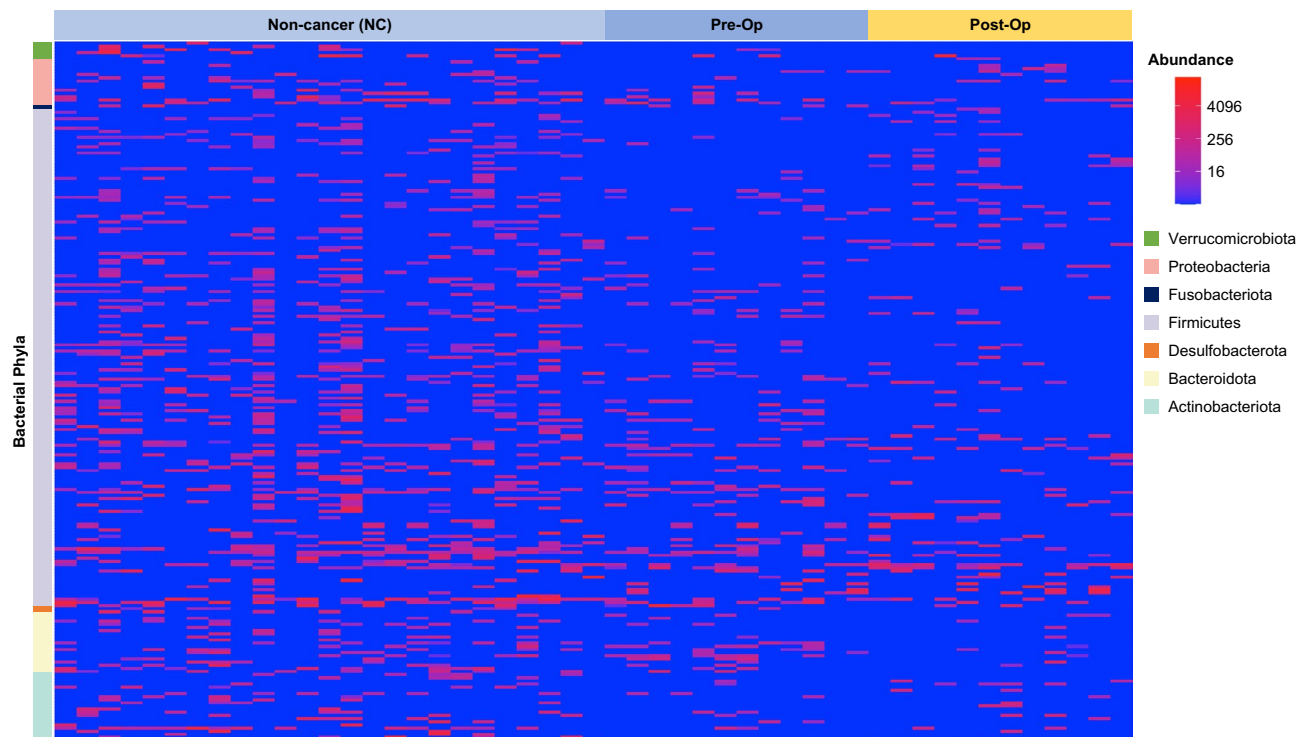


**Figure 1.** Fecal bacterial compositions are different between NC and CRC patients. (A) Bar chart shows the top bacterial phyla present in all samples. (B) Comparisons of bacterial alpha diversity in CRC to NC control group. \* denotes  $p < 0.05$  (Wilcoxon rank sum test). (C) Principal coordinates analysis (PCoA) of bacterial beta diversity derived from Bray–Curtis distances among specimens is shown. F statistics,  $R^2$  & p values of the comparison are presented (PERMANOVA).

**Bacterial diversity is altered in CRC patients compared to NC individuals.** *Firmicutes*, *Actinobacteria*, *Verrucomicrobia* and *Bacteroidetes* are amongst the most abundant bacteria phyla found in all samples (Fig. 1A), Global microbial alpha diversity was assessed using Chao1, ACE and Simpson across 3 groups. Chao1 and ACE indices were significantly decreased in post-op patients ( $p < 0.05$ ). In addition, an increase in Simpson index was observed for pre-op patients ( $p < 0.05$ ) when compared to other groups in the analysis (Wilcoxon rank sum test, Fig. 1B). Using principal coordinate analysis (PCoA) of the Bray–Curtis dissimilarity distances, beta diversity analysis of the groups showed that there is a clear segregation between CRC patients and NC (Fig. 1C). Furthermore, microbial composition of post-op CRC patients presented a more scattered distribution, indicating a different microbial composition as compared to pre-op CRC patients and NC. In all, the data suggest reduced fecal bacterial species richness and diversity in the CRC patients compared to the NC.

**Taxonomic differences in pre- and post-operative CRC patients compared to NC controls.** An overview of differentially abundant bacteria in CRC patients and NC fecal specimens was shown in Fig. 2. Based on the heatmap representation of bacterial abundance, it is clear that alterations in the abundance of amplicon sequence variants (ASVs) belonging to *Actinobacteria*, *Firmicutes*, *Proteobacteria*, *Verrucomicrobia* and *Bacteroidetes* phyla across the different groups were present. Of note, several specific ASVs belonging to *Actinobacteria* and *Firmicutes* were increased, whereas some *Bacteroidetes* ASVs were reduced in post-op CRC patients compared to pre-op patients.

Further detailed analysis revealed the level of changes in abundance of specific bacterial ASVs in pre- and post-op CRC patients compared to NC (Fig. 3, Supplementary Data 1). As shown in Fig. 3A, pre-op CRC patients have a higher abundance of several *Bacteroidetes* ASVs. *Bacteroides caccae*, *Bacteroides vulgatus* and *Bacteroides fragilis* were found to be increased in pre-op CRC patients ( $\alpha \leq 0.0001$ , Fig. 3A, pre-op column). In addition, other bacteria species including, ASV 769 *Odoribacter splanchnicus* ASV 694 *Barnesiella intestinihominis*, ASV 194 *Anaerostipes hadrus*, ASV 766 *Blautia* spp. and ASV 583 *Eubacterium hallii* group were increased by more than 2 Log<sub>2</sub> fold in pre-op CRC compared to NC (Fig. 3B)<sup>27–32</sup>. However, these differences were no longer found



**Figure 2.** Heatmap showing differentially abundant bacterial phyla across NC and CRC patients. Top differentially abundant bacterial phyla in CRC compared to NC control are shown ( $\alpha < 0.01$ , Wald test).

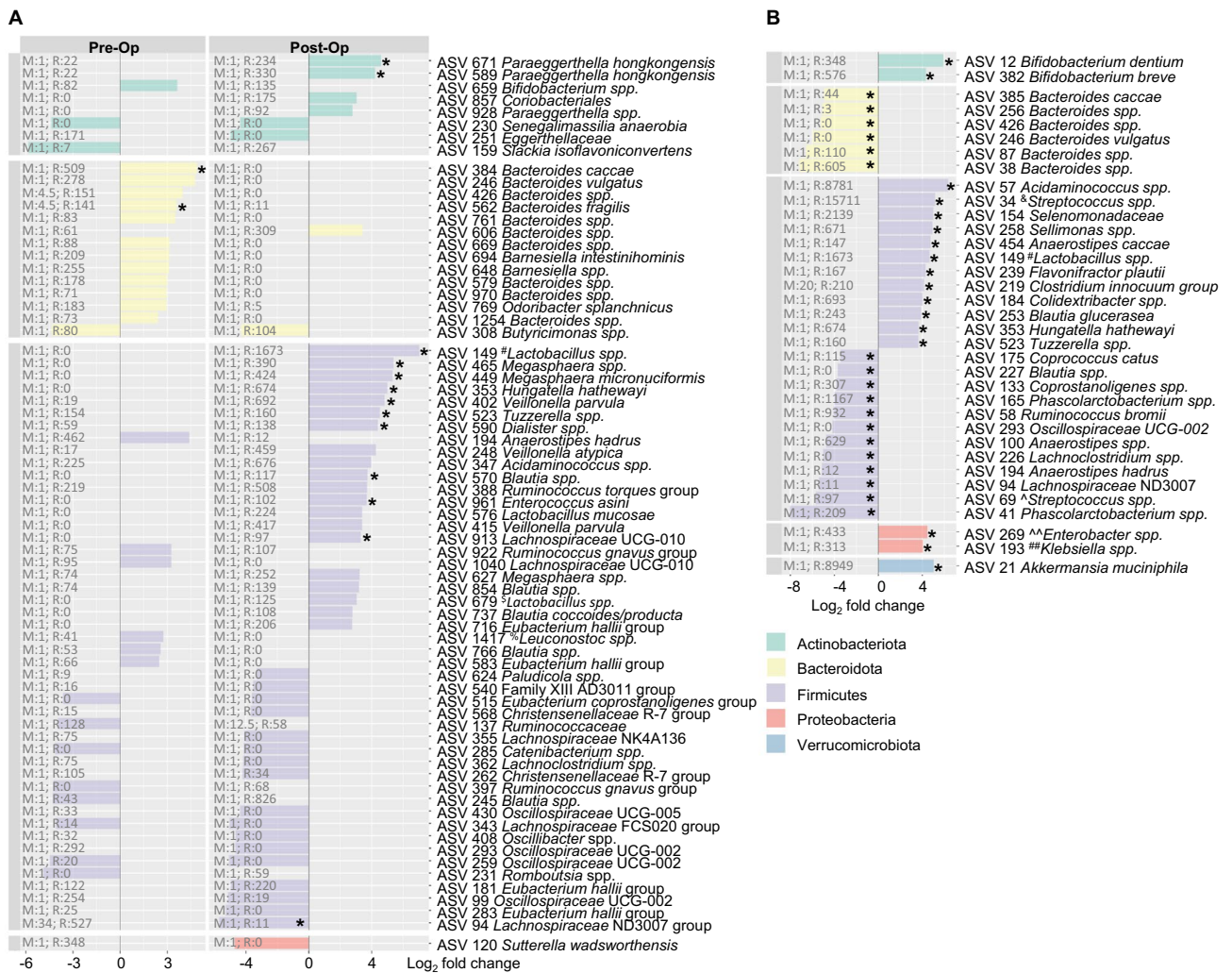
in post-op CRC compared to NC. Of note, the abundance of *Bacteroides* spp., *B. caccae* and *B. vulgatus* were reduced in post-op CRC compared to pre-op based on paired analysis (Fig. 3B). In contrast, fecal samples from pre-op patients have reduced ASV 308 *Butyricimonas* spp., ASV 515 *Eubacterium coprostanoligenes* group, ASV 285 *Catenibacterium* spp. and ASV 343 *Lachnospiraceae* FCS020 group, and these bacterial groups remained reduced even after post-op. Additionally, paired analysis revealed the influence of surgery on specific bacterial ASVs in CRC patients. These included the increase in ASV 269 *Enterobacter* spp., ASV 193 *Klebsiella* spp. and ASV 21 *Akkermansia muciniphila*.

**Microbial functional features associated with CRC development.** *Overview of pathway analysis between groups.* Gene content and functional pathways were inferred from the bacterial 16S rRNA gene sequencing data using Picrust2. The top differentially abundant KEGG (KOs) profiles from NC, pre- and post-op groups were compared using ANOVA Tukey–Kramer post-hoc comparison with Benjamini–Hochberg correction<sup>33</sup>. PCoA plot (Fig. 4A, Supplementary Data 2) showed that the KOs profiles from NC are significantly different from both pre- and post-op CRC (Benjamini–Hochberg FDR < 0.05).

There were 66 and 300 differentially abundant KOs identified in pre- and post-op CRC compared to NC control group respectively (BH adjusted  $P \leq 0.0015$ , Supplementary Data 3 and 4). These KOs were mapped to KEGG pathways including metabolic pathways (23% in pre-op, 34% in post-op); biosynthesis of secondary metabolites (10% in both pre-op and post-op) and microbial metabolism in diverse environments (3% in pre-op, 8% in post-op). The differentially abundant KOs were also assigned to BRITE hierarchies. The most represented BRITE hierarchies included; enzymes (51% in pre-op, 34% in post-op), two-component system (6% in pre-op, 3% in post-op), prokaryotic defense system (5% in pre-op, 1% in post-op) and transcription machinery (5% in pre-op, 1% in post-op).

Interestingly, KOs including K00392 sir; sulfite reductase (ferredoxin) was found to be significantly reduced in pre- and post-op ( $-5.3$  and  $-6.3$   $\text{Log}_2$  fold change respectively, adj.  $P \leq 0.0001$ ) compared to NC (Fig. 4B,C, Supplementary Fig. 2 and Supplementary Data 2), which mapped to the MODULE “M00176 Assimilatory sulfate reduction, sulfate = > H<sub>2</sub>S”. The importance of the altered sulfur metabolism was also identified in subsequent detailed pathway enrichment analysis using clusterProfiler R package<sup>34</sup>. Additionally, ABC transporters remain over-represented in post-op (when compared with NC or pre-op), thus suggesting surgery did not change the bacteria group carrying out this function in patients’ gut. Key KOs in ABC transporter include the trehalose/maltose transport system, cellobiose transport system, urea transport system and fluoroquinolone transport system (Supplementary Data 3 and 4).

*Pre-op CRC vs NC.* As mentioned above, pathway enrichment analysis revealed that “Sulfur metabolism” was significantly lower in pre-op CRC compared to NC controls (Fig. 5A and Supplementary Fig. 2). K00392 is highly associated with pathways including “Sulfur metabolism” and “Microbial metabolism in diverse environ-

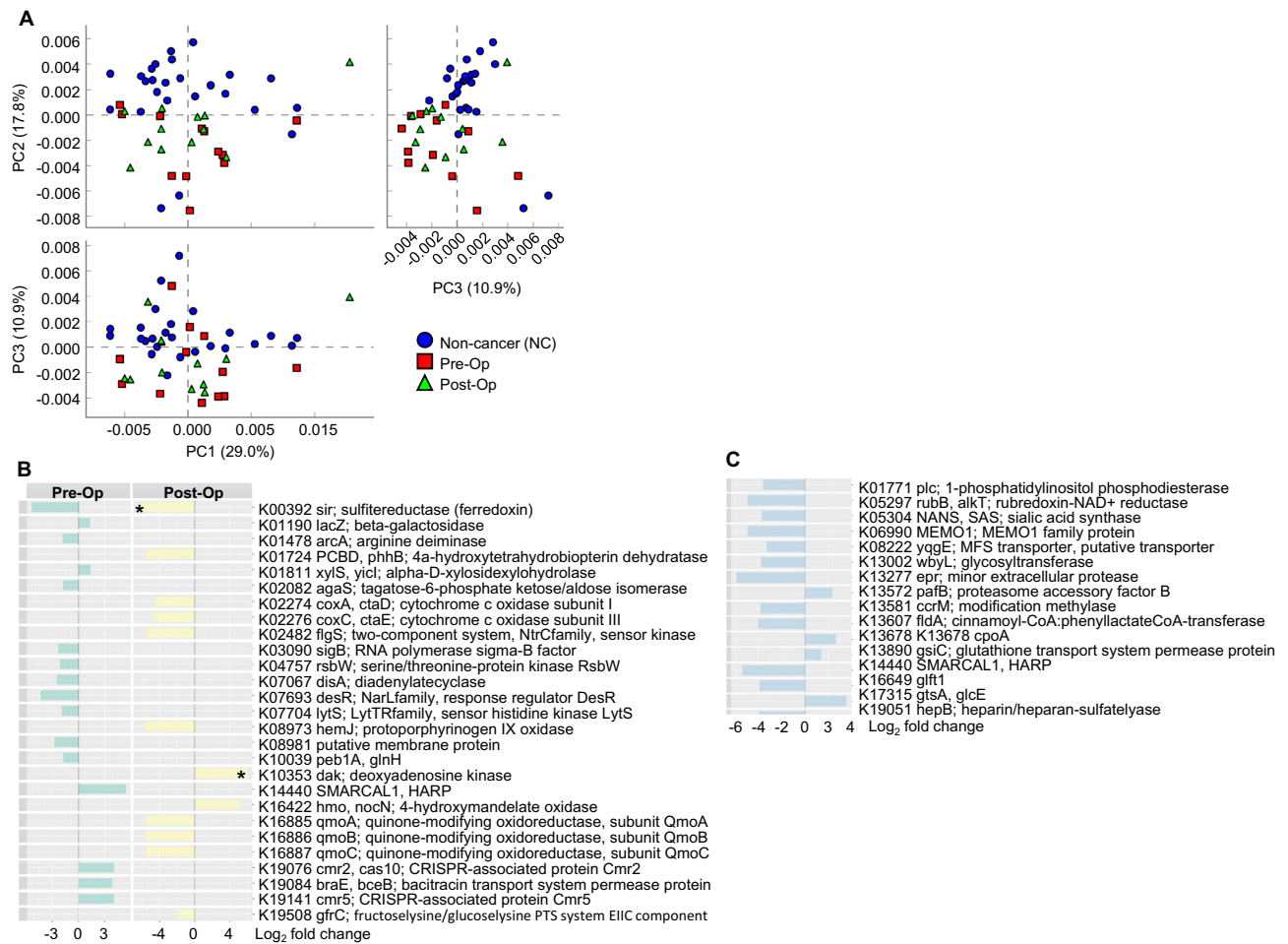


**Figure 3.** Abundance of specific bacterial ASVs are different in CRC patients compared to NC control and changes after surgery. (A) Bar chart shows bacterial ASVs that are significantly different in CRC patient pre-op vs post-op compared to NC controls ( $\alpha \leq 0.0001$ , \*denotes Benjamini–Hochberg adj  $< 0.0001$ ). “M;R” denotes median and range of Deseq2 normalised counts in pre-op & post-op groups. (B) Bar chart shows bacterial ASVs that are significantly different in CRC post-op vs pre-op (Paired test,  $\alpha \leq 0.0001$ , \*denotes Benjamini–Hochberg adj  $< 0.0001$ ). “M;R” denotes median and range of Deseq2 normalised counts in post-op group. Plots showing detail media & range in NC, pre-op & post-op are summarised in Supplementary Data 1 and 2. Abbreviated ASV annotation as below; ASV149 #*Lactobacillus acidophilus/casei/crispatus/gallinarum*; ASV679 \$*Lactobacillus crustorum/farciminis/formosensis/heilongjiangensis/mindensis/musae/nantensis*; ASV1417%*Leuconostoc carnosum/citreum/garlicum/holzappelii/lactis*; ASV\_34 &*Streptococcus alactolyticus/equinus/galloyticus/macedonicus/pasteuri/pasteurianus*; ASV\_69 ^*Streptococcus lutetiensis/salivarius/thermophilus*; ASV\_269 ^^*Enterobacter aerogenes/asburiae/bugandensis/cancerogenus/cloacae/hormaechei/ludwigii/mori/roggenkampii/soli/tabaci*; ASV\_193 ##*Klebsiella pneumoniae/quasivariicola/variicola*.

ments”, indicating a buildup of sulfate in the colonic environment due to the inability to metabolize sulfate to sulfur amino acids<sup>35,36</sup>.

**Post-op CRC vs NC.** When comparing post-op to NC, enrichment analysis showed that Lipoarabinomannan (LAM) biosynthesis was suppressed in post-op compared to NC control group. Lipoarabinomannan is a virulence factor commonly associated with *Mycobacteria tuberculosis* infection. It is also expressed by many other *Actinomyces* bacteria including *Gordonia*, *Rhodococcus*, *Tsakumurella*, *Amycolatopsis* and *Corynebacterium*. While the sequencing data did not have sufficient resolution to identify each of these *Actinomyces* bacteria (*M. tuberculosis* was not detected), we found that the suppression of (LAM) biosynthesis was due to reduced abundance of specific KOs (Supplementary Fig. 2) and ASVs belonging to the Actinomycetaceae family, in particular ASV581 and ASV1102.

Furthermore, KOs related to “Oxidative phosphorylation”, “Thermogenesis” and “Valine, leucine & isoleucine degradation” were suppressed in post-op CRC patients ( $\alpha = 0.000001$ ) (Fig. 5B and Supplementary Fig. 2). KO

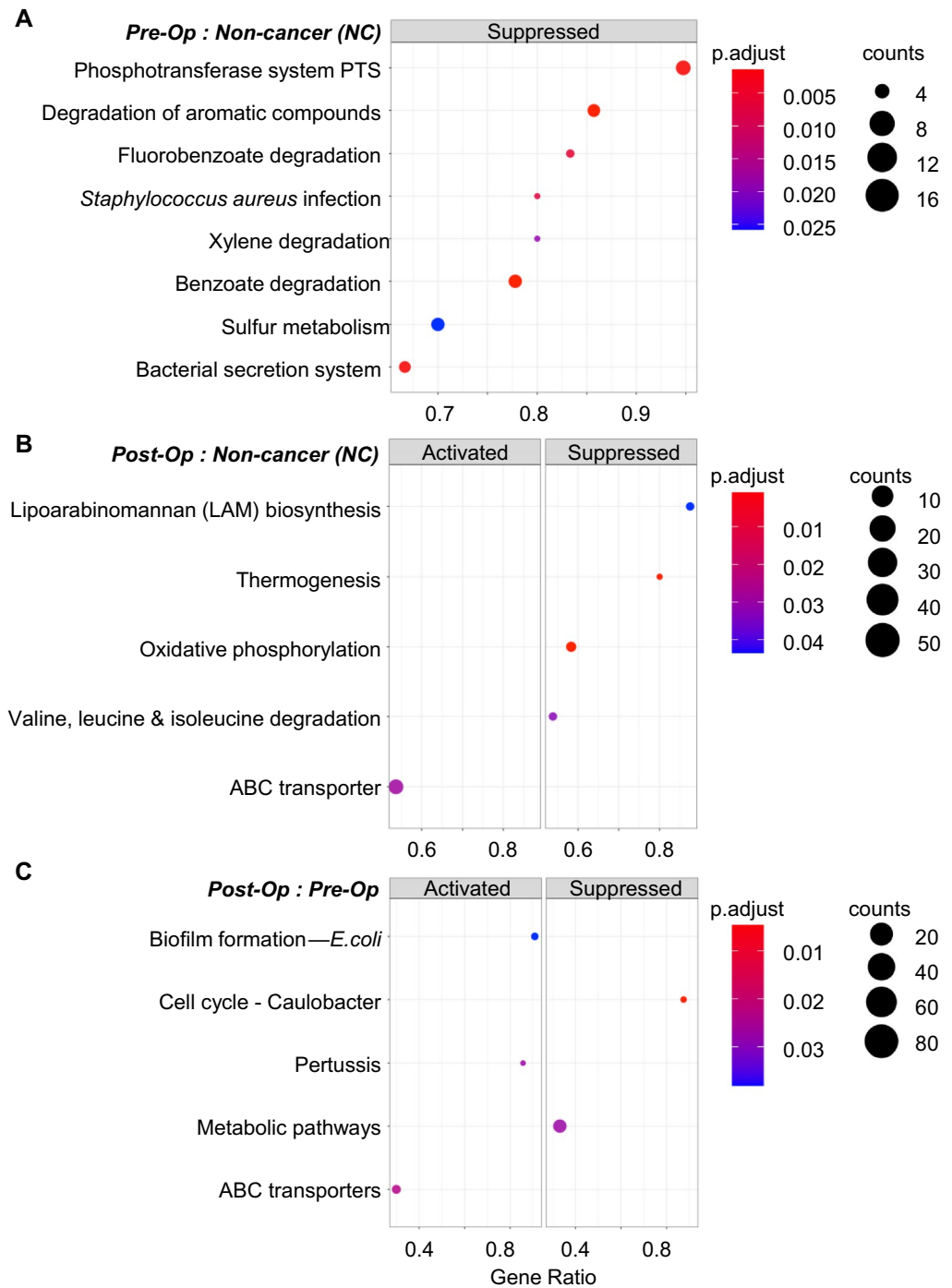


**Figure 4.** Changes in functional profile of bacterial composition in NC and CRC patients. **(A)** Principal coordinate analysis (PCoA) of PICRUSt2-projected functional profiles (level 3, KEGG orthology (KOs)) between NC, CRC patients pre- and post-op (ANOVA Tukey–Kramer post hoc comparison, Benjamini–Hochberg correction,  $p_{adj} < 0.05$ ). **(B)** Bar chart shows top differentially represented KOs in CRC patients pre- and post-op compared to NC control ( $\alpha < 0.000001$ , all comparisons Benjamini–Hochberg  $adj < 0.001$ , \* $p_{adj} < 0.0001$ ). **(C)** Bar chart shows top differentially represented KOs in CRC post-op compared to pre-op (Paired test,  $p < 0.001$ ).

K08973 hemJ; protoporphyrinogen IX oxidase was reduced by 4 Log<sub>2</sub> fold. This KO is highly associated with “Heme biosynthesis, plants and bacteria, glutamate”. Studies have demonstrated that glutamine stimulates the expression of HO-1, a protein that is involved in maintaining intestinal cell proliferation and repair<sup>37</sup>. The presence of pathogenic bacteria such as *Enterobacter spp.*, *Klebsiella spp.*, and *Hungatella hathewayi* (Fig. 3B) as well as a compromised immune function suggests that the colonic environment remains in a state of dysbiosis and is prone to damage after surgery.

Interestingly, thermogenesis is suppressed in post-op CRC patients. Studies have reported that amino acids valine, leucine and isoleucine participate in lipogenesis, lipolysis and immune function of the gut<sup>38</sup>. Lipogenesis and lipolysis can be circumvented through supplementation of microbial butyrate. However, butyrate producing bacteria, *Lachnospiraceae*, was observed to have diminished (Fig. 3B) thus subjecting the gut environment to a compromised adaptive thermogenesis<sup>39</sup>. Similarly, reduction in K02274 coxA, ctaD; cytochrome c oxidase subunit I and K02276 coxC, ctaE; cytochrome c oxidase subunit III, which are associated with “Oxidative phosphorylation”, were also observed. Studies have demonstrated that mitochondrial mutations affect downstream processes of the electron transport chain system, which indirectly contribute to an increase in production of reactive oxygen species leading to a compromised repair system<sup>40,41</sup>.

**Pre- vs post-op CRC.** To evaluate the influence of surgery on microbial dynamics, further analysis were carried out by comparing the changes in predicted function profile between pre- and post-op using paired sample test. A total of 25 differentially abundant KOs were identified (BH adjusted  $P \leq 0.0015$ , Supplementary Data 3). These KOs were mapped to KEGG pathways including metabolic pathways (28%); ABC transporters (8%); Carbon metabolism (8%). KOs K19076 cmr2, cas10; CRISPR-associated protein Cmr2, K19141 cmr5; CRISPR-associated protein Cmr5 and K14440 SMARCAL1, HARP were more abundant (3 Log<sub>2</sub> fold) in pre-op patients com-



**Figure 5.** KEGG gene enrichment analysis. (A,B) Dot plots showing KEGG pathways that were enriched in CRC pre- and post-op compared to NC controls respectively. (C) Dot plot shows enriched KEGG pathways in CRC patients pre- and post-op. p. adjust and counts denote Benjamini–Hochberg adjusted p values and gene counts respectively.

pared to NC. These KOs are highly associated with BRITE functions including “Prokaryotic defense system” and “Chromosome and associated proteins” ( $\alpha < 0.000001$ ) (Fig. 4B, pre-op column, Supplementary Data 2 and 5).

Interestingly, genes associated with PATHWAY function “ABC transporters”, K13890 *gsiC*; glutathione transport system permease protein and K17315 *gtsA*; glucose/mannose transport system substrate-binding protein, were found to be over-represented in post-op, suggesting that surgery leads to increase in bacteria that may be responsible for this function in CRC patients (Supplementary Data 3, 4 and 5). Enrichment analysis revealed genes related to “cell cycle—*Caulobacter*” and “metabolic pathways” were suppressed in post-op CRC patients when compared to their pre-op state ( $\alpha = 0.001$ ) (Fig. 5C and Supplementary Fig. 2) as demonstrated with a 4 Log<sub>2</sub> fold reduced expression of K13581 *ccrM*; modification methylase, K01771 *plc*; 1-phosphatidylinositol

phosphodiesterase, K05304 NANS, SAS; sialic acid synthase and K16649 glf1 (Fig. 4C). As the above-mentioned pathways are related to cellular process and metabolism, their suppression suggests that the cell cycle of the gut microbiome is faulty.

## Discussion

In this pilot study, we compared the gut microbiota of CRC patients (pre- and post-op) with NC individuals. As a pilot study, the number of patients included is small. Therefore, we have taken measures to ensure that only the most reliable data is presented. With a small sample size, only sufficiently large enough effects can be considered genuine. Based on power calculation, results with effect size of 2 would provide 95% confidence with 80% power (see “Results”). Therefore, we have only reported differences that were greater than effect size 2 and at a strict alpha cut-off at  $\alpha \leq 0.0001$ . With these statistical criteria, differences in bacterial composition were identified between CRC patients and NC individuals. The variations in gut microbiota of pre- and post-op patients may suggest possible field change in the “non-diseased” colon that predisposed them to CRC in the first place, due to the findings of persistent microbial dysbiosis after surgery. This study provides insights to alterations in the gut microbiota of post-op CRC patients.

Firstly, we compared the global gut microbiota composition between CRC patients (pre- and post-op) and non-cancer (NC) controls. We found reduction in fecal bacterial species richness and diversity in the CRC patients compared to the NC control group (Fig. 1) which is in line with data published by Cong et al.<sup>42</sup>. However, some studies showed no significant changes in diversity<sup>43–45</sup>. The differences in outcomes may be attributed to different sample sizes or analytical methods. Interestingly, we noticed significant changes in certain microbiota when comparing paired microbial communities between CRC patients pre- and post-op. This may suggest that surgery for CRC not only removes the tumor but is also capable of altering associated gut microbiota communities.

Following which, we analysed the differences in microbiota composition between pre-op, post-op and NC groups. We observed that pre-op CRC group has a significantly higher abundance of several *Bacteroidetes* ASVs and *Odoribacter splanchnicus* compared to NC<sup>19,41</sup>. There is existing literature which suggests microbial dysbiosis having a role in the development of CRC<sup>46–49</sup>. For instance, *Bacteroidetes* and *Odoribacter splanchnicus* are suggested to promote colorectal carcinogenesis. Others have reported association of specific *Bacteroidetes* spp. including *B. caccae* and *B. vulgatus* are known to be involved in destabilizing the colonic wall of the gut, potentially resulting in the progression of CRC<sup>46,47</sup>. At the same time, *Odoribacter* levels is shown to correlate with somatic mutations and cell proliferation resulting in poorer CRC prognosis<sup>50,51</sup>. Interestingly, differences in *Bacteroidetes* ASVs and *Odoribacter splanchnicus* were no longer observed after surgery when compared to NC individuals (Fig. 3A).

At the same time, there was an increased abundance of probiotic bacteria such as *Bifidobacterium* spp. (*Bifidobacterium dentium* and *Bifidobacterium breve*), *Lactobacillus* spp, *Akkermansia muciniphila*, *Anaerostipes caccae* and *Colidextribacter* spp. after surgery (paired analysis between post- and pre-op, Fig. 3B). These probiotics have been demonstrated to fortify the colonic mucus layer, competitively exclude pathogenic bacteria and regulate the colonic environment through regulating of immune cells for anti-tumour activity<sup>52–57</sup>. The increased abundance of beneficial gut genera after surgery such as short-chain fatty acid (SCFA) producing microbiota (*Roseburia* spp. and *Blautia* spp.) and gut barrier enhancer (*Lactobacillus* spp.) suggest that surgery may partially “revert” the gut back to its “NC” pre-cancerous state to a certain degree<sup>30,58</sup>. In addition, presence of pathogenic bacteria such as *Enterobacter* spp., *Klebsiella* spp. and *Hungatella hathewayi* could indicate that the colonic environment is still in a state of dysbiosis as these bacteria are known to be oncogenic<sup>59–61</sup>. Of note, increased *Klebsiella* spp. was also previously reported in post-op compared to pre-op in association with Pertussis (KEGG pathway) enrichment<sup>42</sup>. Microbiota associated with CRC progression such as *Clostridium innocuum*, *Eubacterium brachy*, *Eggerthella* spp. were also observed to be significantly abundant in post-operative CRC patients<sup>40,62,63</sup>. Further, we observed that curative surgery resulted in the reversion of some bacterial strains (i.e., *Bacteroidetes* spp.) to a “non-cancer” level but not others (i.e., *Eubacterium coprostanoligenes* group). On the other hand, certain bacterial groups which were found in equal abundance in NC patients were increased after surgery (i.e., *A. muciniphila*). This illustrates a possible field-change in microbial composition in the colon of CRC patients, where there is persistence of oncogenic microbial species even after surgery.

Lastly, we examined the functional features inferred from bacterial 16S rRNA gene. Previous studies have shown that metabolic pathways are much more consistent across individuals compared to the differences in bacterial abundance<sup>63</sup>. A healthy gut microbiota may contain specific microbial combinations and essential metabolic pathways that together maintain a two-way beneficial relationship with the host<sup>64</sup>. In our study, we observed differential enrichment of various pathways in patients when compared between the groups. For instance, the inability to metabolise sulfate in the colonic environment was found to be significantly suppressed in pre-op CRC patients compared to NC. In addition, correlation between increased pathogenic bacteria and infectious disease associated pathways (e.g., Pertussis) may be indicative of adverse post-surgery response. Furthermore, correlation of some suppressed pathways to the presence of pathogenic bacteria are indications of how the gut microbiota in CRC patients responded to surgery, thereby influencing its functional stability<sup>40</sup>.

Besides utilising microbial strains which are found in greater abundance in CRC, equal emphasis should be given to other significant strains that are reduced after surgery. The microbial dysbiotic signatures associated with CRC will therefore comprise a specific pattern of alteration in bacteriome. This contrasts with the existing trend of simply identifying specific microbial strains as a complementary tool for the screening and diagnosis of CRC as proposed by various studies. Moreover, these specific patterns can also be used to screen for disease recurrence in CRC patients after surgery, where the return of specific microbial markers can indicate recurrent disease. A follow-up study on the current cohort can seek to examine possible changes in bacteriome in association with disease remission and recurrence.



More importantly, it will be interesting to examine if a similar pattern of microbial dysbiosis is present in patients with adenomatous polyps, which are precursors of CRC, and if these signatures are lost after polypectomy is performed. Studies have reported specific microbes associated with the presence of colonic polyps, which supports the presence of a unique microbial profile even in patients with pre-malignant polyps<sup>65,66</sup>. If found, these microbial dysbiotic signatures may perhaps have the potential to be used as a predictor for patients with polyps and who thus require more urgent colonoscopy. This can help streamline CRC screening and has the potential to guide the allocation and prioritization of colonoscopy, which is costly and not without risks. It may even help guide the indication for surveillance colonoscopy by identifying patients with a greater chance of having recurrent or persistent polyps.

Our study has a number of limitations. Other than the aforementioned small sample size, all recruited CRC patients were diagnosed with left-sided colon cancer. As right and left-sided colon cancers are known to have different molecular biological characteristics and possibly differentially expressed microbial species, the results of our study may not be applicable to right-sided colon cancers<sup>63</sup>. On the other hand, samples from NC group were collected only at a single timepoint and follow up sampling of NC to match the CRC group with exact surgery (e.g. colectomy) was not performed. Although matched NC sampling may provide better comparison with CRC patients, clinical indication for surgery in benign conditions in the NC group is rare (e.g. elective anterior resection for diverticular disease complicated by stricture or fistula). Therefore, the NC controls included are next to ideal for the purpose of this study. They have normal colonoscopy and do not have a history of inflammatory bowel or autoimmune disease (see “Methods”, “Study participants and fecal sample collection”), which could confound the results in this study. Dietary information should also be taken into consideration particularly for CRC patients as changes and differences in diet may influence gut microbiome composition and the risk of CRC<sup>67</sup>. Therefore, inclusion of differences as well as changes in dietary patterns may provide better assessment and identification of confounders when identifying microbial changes before and after surgery in CRC patients.

Specimens used for microbial analysis in this pilot study was based on stool samples, which is postulated to reflect the cancer microbiota due to the downstream shedding of cancer cells. However, it might not fully represent the altered microbial profile of the cancer tissue and we are also unable to study changes in the tumor microenvironment. The bacterial 16S rRNA gene sequencing method, which is a widely employed method of examination of gut microbiome, has technical limitations as 16S rRNA gene amplicon sequencing provides less coverage (absent of other microbes). Unlike shotgun methodology, which can sequence all the genomic material present in the samples, 16S rRNA gene sequencing may not provide a comprehensive assessment of the microbial population with fine resolution at lower taxonomic level. However, both 16S and shotgun methodologies have demonstrated close consensus for bacterial phyla detection<sup>68,69</sup>. We have also analysed our sequencing data using the Dada2 software package to improve the reliability of results. This method was shown to detect fine scale sequence variations with better accuracy and sensitivity compared to other OTU-based output analysis<sup>22,70,71</sup>. The limited resolution provided by 16S rRNA gene sequencing potentially impacted downstream analysis including functional profiling, which can only be “predicated or inferred” using curated pipelines such as Picrust2 employed in this pilot study<sup>23</sup>. Validation with experimental data is important to confirm the changes in functional pathways and their possible influence in the various patient groups. This applies to shotgun sequencing data as well even though all genetic material that encodes for microbial products can potentially be determined by shotgun sequencing. Shotgun results are suitable for downstream bioinformatic analysis to determine the range of functions that can be carried by the organisms. However, the extent of functional activity is not guaranteed regardless of how well the sequencing coverage is, and experimental validation to confirm each of the functional pathways is still required. Therefore, the functional profiles presented in our manuscript should be carefully interpreted since it is an “inferred” data from a small sample population. To note, this study has carefully presented only data that is most reliable based on statistical testing as mentioned above (e.g. effect size > 2 and  $\alpha \leq 0.0001$ ).

Although there are limitations in samples size and methodology, our pilot study provides important preliminary insights to the changes of bacterial compositions in CRC with respect to surgical treatment in local Singapore cohort. Our study clearly showed the presence of bacterial composition alteration in NS vs CRC and after surgical intervention. Further validation studies should be carried out in a larger sample size and include both left- and right sided CRC, with a longer duration of follow-up to examine possible predictive microbial signature for disease remission or recurrence. In addition, shotgun sequencing can be employed to assess changes in other microbes including virus and eukaryotes. Further analysis to determine presence of network or co-abundance relationships between different microbial species will help to provide an in-depth evaluation of possible microbial interaction that could influence functional outcomes. The changes in functional outcomes can be measured by metabolomics or lipidomics to assess key microbial and host metabolites such as short chain fatty acid, bile acids, and indole derivatives.

## Conclusion

Microbiome signatures characteristic of CRC likely involves a set of associated microbial alterations which comprise both significant increases and reductions in specific bacterial strains. Elements of this unique microbial dysbiosis of CRC persists even after surgery, suggesting possible field-change in the remnant non-diseased colon. Future studies should involve a larger sample size with shotgun microbiome sequencing data collected at multiple time points after surgery to examine if these dysbiotic patterns truly persist and also correlate with disease outcomes.

## Methods

**Study participants and fecal sample collection.** Ethics approval to conduct this study was obtained by the National Healthcare Group domain specific review board (NHG-DSRB), reference number: 2017/01257. All methods were carried out in accordance with relevant guidelines and regulations as outlined by National Healthcare Group domain specific review board, Singapore.

All patients aged 21 and above who were scheduled to undergo elective diagnostic or screening colonoscopy between May 2018 and November 2018 were invited to participate in this study. Informed consent was obtained from all participants and/or their legal guardian(s) prior to their participation in the study. Participants who were found to have CRC at colonoscopy were assigned to the “Cancer” group, while those who did not have cancer and had normal colonoscopy results were assigned to the “Non-cancer” group. Exclusion criteria were patients who were unable to provide written consent for participation in the study, had pre-existing family history of familial adenomatous polyposis (FAP) or hereditary non-polyposis colorectal cancer (HNPCC), diagnosed inflammatory bowel disease, autoimmune diseases and/or consumption of a prolonged course of antibiotics (> 3 days) during the three months prior to the study or after the surgery.

To compare changes in the gut microbiome profiles of CRC patients before and after surgery, fecal samples at 2 sampling time points were collected from the cancer group. The 2 time points were 1. Prior to colonoscopy (pre-op), before initiation of any cancer treatment, and 2. Six months after surgery (post-op). On the other hand, only one fecal sample prior to colonoscopy was collected from each individual in the Non-cancer group (NC). As there have been reports of immediate microbial alterations after mechanical bowel preparation<sup>20</sup> which generally recover after 14 days<sup>21</sup>, the stool samples collected prior to colonoscopy for both groups were taken before consumption of polyethylene glycol. A total of 49 fecal samples were collected directly into sterile fecal collection tubes containing RNAlater® (Invitrogen, Lithuania) and 4 sterile glass beads measuring 5 mm in diameter (Merck, Germany). Samples collected by the patients were mailed to the laboratory within one week ( $\pm$  three days). Pre-op samples were collected 24 h ( $\pm$  12 h) pre-op, and frozen at  $-80^{\circ}\text{C}$  within 2 h ( $\pm$  1.5 h) hours, whereas post-op samples were collected at 6 months ( $\pm$  5 days) post-op by the patients for mailing to the laboratory. All mailed in samples were stored in  $-80^{\circ}\text{C}$  within 24 h ( $\pm$  12 h) upon receipt in the laboratory. Due to the longitudinal nature of this study, all samples were stored at  $-80^{\circ}\text{C}$  for an average of 6 months ( $\pm$  2 months) before bulk DNA extraction. All samples are not subjected to any freeze–thaw process before DNA extraction. All DNA extracted were immediately quantified and stored at  $-20^{\circ}\text{C}$  for 7 days before delivery to external vendors for sequencing. The DNA were packaged in dry ice for delivery to the sequencing vendors (reception by vendors within 24 h).

**Bacterial 16S rRNA gene sequencing.** Fecal samples were freshly thawed and vortexed to produce a homogenate and washed once using 1X phosphate buffer saline (Vivantis Technologies, Malaysia). Genomic DNA was extracted using DNeasy PowerSoil Pro Kit (Qiagen, Germany) according to manufacturers’ protocol. The purity and DNA yield for each fecal sample was measured using a nanodrop.

In this study, the V3–V4 region of 16S rRNA was sequenced to analyse composition of the gut microbiota of fecal samples. A PCR targeting the V3–V4 region using forward (5′-CCTACGGGNGGCWGCAG) and reverse primers (5′-GACTACHVGGGTATCTAATCC) with overhang adaptors, were performed as recommended by the 16S metagenomics library kit by Illumina. The quality and quantity of the amplicons were measured using Agilent 4200 TapeStation, picogreen and nanodrop. All samples passed the quality control measurement and proceeded for a second round of PCR step for library preparation. Library qualities were measured using Agilent 4200 TapeStation, picogreen and qPCR. Libraries that passed the quality control measurement were pooled as recommended by Illumina and were sequenced using the MiSeq platform using 2  $\times$  300 PE format.

**Microbiome profiling.** Raw fastq files were quality filtered, merged, demultiplexed and denoised using DADA2 (version 1.14.1) microbiome pipeline with default parameters<sup>22</sup>. Output from DADA2, known as amplicon sequence variants (ASVs), were cleaned up by removing chimeras. ASVs is a clustering method for better precision and improve resolution of sequence identity. ASVs were then assigned to their respective taxonomy using Silva (version 138) classifier.

Statistical analysis was carried out in R (version 3.6.3) and analysis of microbiome diversities was carried out using the Phyloseq package (version 1.30.0). Alpha diversity was calculated using Chao1, ACE and Simpson. Diversity differences between CRC (at each sampling time point) and NC groups were analysed using Wilcoxon rank sum test. Inter-group similarities were visualized using principal coordinate analysis (PCoA) of the Bray–Curtis dissimilarity distances. The DeSeq2 analysis package (version 1.26.0—statistical pipeline based on range of distribution) was used to identify microbiota that showed significant differential abundances across NC, pre-op and post-op. Read counts/sample and counts across samples were normalised by sample-specific factors determined by median ratio of ASV or KO counts (median of ratio method). The data shown in the figures are the average of the normalised count values, corrected for confounders including gender, race and age of the patients over all samples. For comparisons between pre-Op and post-Op, paired analysis was employed to reflect sampling from the same patient. Predicted functional information from the gut microbiome data was generated using PICRUSt2 (version 2.3.0) pipeline with default parameters<sup>23–25</sup>.

## Data availability

The datasets generated and/or analysed during the current study are available in the NCBI Sequence read archive (SRA) as part of accession BioProject number PRJNA662014.

Received: 1 December 2021; Accepted: 17 May 2022

Published online: 14 June 2022

## References

1. Neish, A. S. Microbes in gastrointestinal health and disease. *Gastroenterology* **136**(1), 65–80 (2009).
2. Marchesi, J. R. *et al.* Towards the human colorectal cancer microbiome. *PLoS ONE* **6**(5), e20447 (2011).
3. Marchesi, J. R. *et al.* The gut microbiota and host health: A new clinical frontier. *Gut* **65**(2), 330–339 (2016).
4. Scanlan, P. D. *et al.* Culture-independent analysis of the gut microbiota in colorectal cancer and polyposis. *Environ. Microbiol.* **10**(3), 789–798 (2008).
5. Ou, J. *et al.* Diet, microbiota, and microbial metabolites in colon cancer risk in rural Africans and African Americans. *Am. J. Clin. Nutr.* **98**(1), 111–120 (2013).
6. Arthur, J. C. *et al.* Intestinal inflammation targets cancer-inducing activity of the microbiota. *Science* **338**(6103), 120–123 (2012).
7. Dejea, C. M. *et al.* Patients with familial adenomatous polyposis harbor colonic biofilms containing tumorigenic bacteria. *Science* **359**(6375), 592–597 (2018).
8. Kostic, A. D. *et al.* Genomic analysis identifies association of *Fusobacterium* with colorectal carcinoma. *Genome Res.* **22**(2), 292–298 (2012).
9. Yang, Y. *et al.* *Fusobacterium nucleatum* increases proliferation of colorectal cancer cells and tumor development in mice by activating toll-like receptor 4 signaling to nuclear factor- $\kappa$ B, and up-regulating expression of MicroRNA-21. *Gastroenterology* **152**(4), 851–66.e24 (2017).
10. Kostic, A. D. *et al.* *Fusobacterium nucleatum* potentiates intestinal tumorigenesis and modulates the tumor-immune microenvironment. *Cell Host Microbe*. **14**(2), 207–215 (2013).
11. Dharmani, P., Strauss, J., Ambrose, C., Allen-Vercoe, E. & Chadee, K. *Fusobacterium nucleatum* infection of colonic cells stimulates MUC2 mucin and tumor necrosis factor alpha. *Infect Immun.* **79**(7), 2597–2607 (2011).
12. Ishaq, S. & Nunn, L. *Helicobacter pylori* and gastric cancer: A state of the art review. *Gastroenterol. Hepatol. Bed Bench* **8**(Suppl 1), S6–S14 (2015).
13. Dejea, C. M. *et al.* Microbiota organization is a distinct feature of proximal colorectal cancers. *Proc. Natl. Acad. Sci. USA* **111**(51), 18321–18326 (2014).
14. Flemer, B., Herlihy, M., O’Riordain, M., Shanahan, F. & O’Toole, P. W. Tumour-associated and non-tumour-associated microbiota: Addendum. *Gut Microbes*. **9**(4), 369–373 (2018).
15. Geis, A. L. *et al.* Regulatory T-cell response to enterotoxigenic *Bacteroides fragilis* colonization triggers IL17-dependent colon carcinogenesis. *Cancer Discov.* **5**(10), 1098–1109 (2015).
16. Hooks, K. B. & O’Malley, M. A. Dysbiosis and its discontents. *MBio* **8**(5), e01492 (2017).
17. Bäckhed, F. *et al.* Defining a healthy human gut microbiome: Current concepts, future directions, and clinical applications. *Cell Host Microbe*. **12**(5), 611–622 (2012).
18. Brüssow, H. Problems with the concept of gut microbiota dysbiosis. *Microb. Biotechnol.* **13**(2), 423–434 (2020).
19. Kong, C. *et al.* Alterations in intestinal microbiota of colorectal cancer patients receiving radical surgery combined with adjuvant CapeOx therapy. *Sci. China Life Sci.* **62**(9), 1178–1193 (2019).
20. Watanabe, M. *et al.* Randomized clinical trial of the influence of mechanical bowel preparation on faecal microflora in patients undergoing colonic cancer resection. *Br. J. Surg.* **97**(12), 1791–1797 (2010).
21. Nagata, N. *et al.* Effects of bowel preparation on the human gut microbiome and metabolome. *Sci. Rep.* **9**(1), 4042 (2019).
22. Callahan, B. J. *et al.* DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**(7), 581–583 (2016).
23. Douglas, G. M. *et al.* PICRUSt2: An improved and extensible approach for metagenome inference. *bioRxiv* <https://doi.org/10.1101/672295> (2019).
24. Ye, Y. & Doak, T. G. A parsimony approach to biological pathway reconstruction/inference for genomes and metagenomes. *PLoS Comput. Biol.* **5**(8), e1000465 (2009).
25. Douglas, G. M. *et al.* PICRUSt2 for prediction of metagenome functions. *Nat. Biotechnol.* **38**(6), 685–688 (2020).
26. Kohn, M. A. & Senyak, J. *Sample Size Calculators*. <https://www.sample-size.net/> (2021)
27. Yazici, C. *et al.* Race-dependent association of sulfidogenic bacteria with colorectal cancer. *Gut* **66**(11), 1983–1994 (2017).
28. Cheng, W. Y., Wu, C.-Y. & Yu, J. The role of gut microbiota in cancer treatment: Friend or foe?. *Gut* **69**(10), 1867 (2020).
29. Ai, D. *et al.* Identifying Gut microbiota associated with colorectal cancer using a zero-inflated lognormal model. *Front. Microbiol.* **10**, 826 (2019).
30. Wu, Y. *et al.* Identification of microbial markers across populations in early detection of colorectal cancer. *Nat. Commun.* **12**(1), 3063 (2021).
31. Liu, X. *et al.* *Blautia* a new functional genus with potential probiotic properties?. *Gut Microbes*. **13**(1), 1–21 (2021).
32. Engels, C., Ruscheweyh, H.-J., Beerenwinkel, N., Lacroix, C. & Schwab, C. The common gut microbe *Eubacterium hallii* also contributes to intestinal propionate formation. *Front. Microbiol.* **7**, 713 (2016).
33. Kanehisa, M. & Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**(1), 27–30 (2000).
34. Yu, G., Wang, L. G., Han, Y. & He, Q. Y. clusterProfiler: An R package for comparing biological themes among gene clusters. *OMICS* **16**(5), 284–287 (2012).
35. Dordević, D., Jančíková, S., Vitězová, M. & Kushkevych, I. Hydrogen sulfide toxicity in the gut environment: Meta-analysis of sulfate-reducing and lactic acid bacteria in inflammatory processes. *J. Adv. Res.* **27**, 55–69 (2021).
36. Kushkevych, I. *et al.* Recent advances in metabolic pathways of sulfate reduction in intestinal bacteria. *Cells* **9**(3), 698 (2020).
37. Oates, P.-S. & West, A.-R. Heme in intestinal epithelial cell turnover, differentiation, detoxification, inflammation, carcinogenesis, absorption and motility. *World J. Gastroenterol.* **12**(27), 4281–4295 (2006).
38. Zhang, S., Zeng, X., Ren, M., Mao, X. & Qiao, S. Novel metabolic and physiological functions of branched chain amino acids: A review. *J. Anim. Sci. Biotechnol.* **8**, 10 (2017).
39. Li, B. *et al.* Microbiota depletion impairs thermogenesis of brown adipose tissue and browning of white adipose tissue. *Cell Rep.* **26**(10), 2720–37.e5 (2019).
40. Ashton, T. M., McKenna, W. G., Kunz-Schughart, L. A. & Higgins, G. S. Oxidative phosphorylation as an emerging target in cancer therapy. *Clin. Cancer Res.* **24**(11), 2482 (2018).
41. Namslauer, I. & Brzezinski, P. A mitochondrial DNA mutation linked to colon cancer results in proton leaks in cytochrome C oxidase. *Proc. Natl. Acad. Sci. USA* **106**(9), 3402 (2009).
42. Cong, J. *et al.* A pilot study: Changes of gut microbiota in post-surgery colorectal cancer patients. *Front. Microbiol.* **9**, 2777 (2018).
43. Deng, X. *et al.* Comparison of microbiota in patients treated by surgery or chemotherapy by 16S rRNA sequencing reveals potential biomarkers for colorectal cancer therapy. *Front. Microbiol.* **9**, 1607 (2018).
44. Carr, P. R. *et al.* Lifestyle factors and risk of sporadic colorectal cancer by microsatellite instability status: A systematic review and meta-analyses. *Ann. Oncol.* **29**(4), 825–834 (2018).
45. Aries, V., Crowther, J. S., Drasar, B. S., Hill, M. J. & Williams, R. E. Bacteria and the aetiology of cancer of the large bowel. *Gut* **10**(5), 334–335 (1969).
46. Lucas, C., Barnich, N. & Nguyen, H. T. T. Microbiota, inflammation and colorectal cancer. *Int. J. Mol. Sci.* **18**(6), 1310 (2017).
47. Sánchez-Alcoholado, L. *et al.* Gut microbiota-mediated inflammation and gut permeability in patients with obesity and colorectal cancer. *Int. J. Mol. Sci.* **21**(18), 6782 (2020).
48. Gao, Z., Guo, B., Gao, R., Zhu, Q. & Qin, H. Microbiota dysbiosis is associated with colorectal cancer. *Front. Microbiol.* **6**, 20 (2015).

49. Feng, Q. *et al.* Gut microbiome development along the colorectal adenoma–carcinoma sequence. *Nat. Commun.* **6**(1), 6528 (2015).
50. Hiippala, K. *et al.* Novel *Odoribacter splanchnicus* strain and its outer membrane vesicles exert immunoregulatory effects in vitro. *Front. Microbiol.* **11**, 2906 (2020).
51. Oh, B. S. *et al.* Cell-free supernatant of *odoribacter splanchnicus* isolated from human feces exhibits anti-colorectal cancer activity. *Front. Microbiol.* **12**, 736343 (2021).
52. Engevik, M. A. *et al.* *Bifidobacterium dentium* fortifies the intestinal mucus layer via autophagy and calcium signaling pathways. *MBio* **10**(3), e01087-19 (2019).
53. Wei, H. *et al.* Antitumor mechanisms of bifidobacteria. *Oncol. Lett.* **16**(1), 3–8 (2018).
54. Zhong, L., Zhang, X. & Covasa, M. Emerging roles of lactic acid bacteria in protection against colorectal cancer. *World J. Gastroenterol.* **20**(24), 7878–7886 (2014).
55. Dutta, D. & Lim, S. H. Bidirectional interaction between intestinal microbiome and cancer: Opportunities for therapeutic interventions. *Biomark. Res.* **8**(1), 31 (2020).
56. Tarashi, S. *et al.* Gut bacteria and their metabolites: Which one is the defendant for colorectal cancer?. *Microorganisms* **7**(11), 561 (2019).
57. Mager, L. F. *et al.* Microbiome-derived inosine modulates response to checkpoint inhibitor immunotherapy. *Science* **369**(6510), 1481 (2020).
58. Yurdakul, D., Yazgan-Karataş, A. & Şahin, F. *Enterobacter* strains might promote colon cancer. *Curr. Microbiol.* **71**(3), 403–411 (2015).
59. Lai, H. C. *et al.* Increased incidence of gastrointestinal cancers among patients with pyogenic liver abscess: A population-based cohort study. *Gastroenterology* **146**(1), 129–37.e1 (2014).
60. Xia, X. *et al.* Bacteria pathogens drive host colonic epithelial cell promoter hypermethylation of tumor suppressor genes in colorectal cancer. *Microbiome* **8**(1), 108 (2020).
61. Chia, J. H. *et al.* *Clostridium innocuum* is a significant vancomycin-resistant pathogen for extraintestinal clostridial infection. *Clin. Microbiol. Infect.* **23**(8), 560–566 (2017).
62. Lazar, V. *et al.* Aspects of gut microbiota and immune system interactions in infectious diseases, immunopathology, and cancer. *Front. Immunol.* **9**, 1830 (2018).
63. Huttenhower, C. *et al.* Structure, function and diversity of the healthy human microbiome. *Nature* **486**(7402), 207–214 (2012).
64. Martiny, J. B., Jones, S. E., Lennon, J. T. & Martiny, A. C. Microbiomes in light of traits: A phylogenetic perspective. *Science* **350**(6261), aac9323 (2015).
65. Sheng, Q. *et al.* Characteristics of fecal gut microbiota in patients with colorectal cancer at different stages and different sites. *Oncol. Lett.* **18**(5), 4834–4844 (2019).
66. Sobhani, I. *et al.* Microbial dysbiosis and colon carcinogenesis: Could colon cancer be considered a bacteria-related disease?. *Ther. Adv. Gastroenterol.* **6**(3), 215–229 (2013).
67. Loke, Y. L., Chew, M. T., Ngeow, Y. F., Lim, W. W. D. & Peh, S. C. Colon carcinogenesis: The interplay between diet and gut microbiota. *Front. Cell. Infect. Microbiol.* **10**, 733 (2020).
68. Brumfield, K. D., Huq, A., Colwell, R. R., Olds, J. L. & Leddy, M. B. Microbial resolution of whole genome shotgun and 16S amplicon metagenomic sequencing using publicly available NEON data. *PLoS ONE* **15**(2), e0228899 (2020).
69. Escobar-Zepeda, A. *et al.* Analysis of sequencing strategies and tools for taxonomic annotation: Defining standards for progressive metagenomics. *Sci. Rep.* **8**(1), 12034 (2018).
70. Barnes, C. J. *et al.* Comparing DADA2 and OTU clustering approaches in studying the bacterial communities of atopic dermatitis. *J. Med. Microbiol.* **69**(11), 1293–1302 (2020).
71. Prodan, A. *et al.* Comparing bioinformatic pipelines for microbial 16S rRNA amplicon sequencing. *PLoS ONE* **15**(1), e0227434 (2020).

## Acknowledgements

This work was supported by the National University Health System Seed Fund [NUHSRO/2020/142/RO5+6/Seed-Sep/02]; and the The Lee Foundation.

## Author contributions

Conception and design: C.W.P., Y.K.C., J.H.L., K.K.T. Administrative support: C.W.P., Y.K.C., Y.Z., K.K.T. Provision of study materials or patients: Y.K.C., K.K.T. Collection and assembly of data: C.W.P., Y.K.C. Data analysis and interpretation: C.W.P., Y.K.C., J.H.L. Manuscript writing: All authors. Final approval of manuscript: All authors.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-14203-z>.

**Correspondence** and requests for materials should be addressed to K.-K.T.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022