Surgical Treatment for Colorectal Cancer Partially Restores Gut Microbiome and Metabolome Traits

Hirotsugu Shiroma

Tokyo institute of Technology School of Life Science and Technology <https://orcid.org/0000-0001-9575-0469>

Satoshi Shiba

Division of Cancer Genomics, National Cancer Center Research Institute, Chuo-ku, Tokyo, Japan

Pande Putu Erawijantari

Tokyo institute of Technology School of Life Science and Technology

Hiroyuki Takamaru

Endoscopy Division, National Cancer Center Hospital, Tokyo, Japan

Masayoshi Yamada

Endoscopy Division, National Cancer Center Hospital, Tokyo, Japan

Taku Sakamoto

Endoscopy Division, National Cancer Center Hospital, Tokyo, Japan

Yukihide Kanemitsu

Department of Colorectal Surgery, National Cancer Center Hospital, Tokyo, Japan

Sayaka Mizutani

Tokyo institute of Technology School of Life Science and Technology

Tomoyoshi Soga

Institute for Advanced Biosciences, Keio University, Yamagata, Japan

Yutaka Saito

Endoscopy Division, National Cancer Center Hospital, Tokyo, Japan

Tatsuhiro Shibata

National Cancer Institute Division of Cancer Prevention

Shinji Fukuda

Institute for Advanced Biosciences, Keio University, Yamagata, Japan

Shinichi Yachida

Division of Cancer Genomics, National Cancer Center Research Institute, Chuo-ku, Tokyo, Japan

Takuji Yamada (✉ takuji@bio.titech.ac.jp)

School of Life Science and Technology, Tokyo Institute of Technology, Tokyo, Japan

<https://orcid.org/0000-0002-9622-1849>

Keywords: Colorectal cancer, surgery, metagenome, human gut microbiome, metabolome

Posted Date: April 15th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-400629/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at mSystems on March 21st, 2022. See the published version at <https://doi.org/10.1128/msystems.00018-22>.

Abstract

Background: Postoperative colorectal cancer (CRC) patients are at increased risk of developing metachronous CRC. Despite accumulating evidence indicating that the gut microbiota and metabolites can promote CRC carcinogenesis, the influence of surgery for CRC on the gut microbiota and metabolites remains partially understood. We hypothesized that if surgery does not eliminate the bacteria and metabolites promoting CRC carcinogenesis, these bacteria and metabolites might be associated with the development of metachronous CRC. To test this hypothesis, we collected 170 fecal samples from 85 CRC patients in pre- and approximately one year postsurgery status, and performed shotgun metagenomics sequencing and capillary electrophoresis time-of-flight mass spectrometry-based metabolomics analyses and compared pre- and postsurgery status.

Results: CRC-associated bacteria such as *Parvimonas micra* and *Fusobacterium nucleatum* were significantly ($P < 0.005$) decreased after surgery. On the other hand, cholate, carcinogenesis-associated deoxycholate, its biotransformed genes (*bai* operon) from cholate to deoxycholate, and the contributing bacterium (*Clostridium scindens*) were significantly ($P < 0.005$) increased. Additionally, these alterations were only observed in postleft surgery. Cholate and glycocholate were significantly ($P < 0.005$) increased in postright surgery. We also developed a method for potential CRC risk assessment based on the gut microbiota and metabolomic compositions using a random forest machine learning algorithm and then applied it to postoperative patients. The estimated CRC risk based on the random forest algorithm was partially restored postsurgery.

Conclusions: Our results indicate that the high potential CRC risk in CRC postoperative patients is associated with metabolites derived from the gut microbiota, which targets interventions to reduce the CRC risk.

Background

Colorectal cancer (CRC) is the third most common cancer, with over 1.8 million new cases and approximately 881,000 deaths per year worldwide[1]. Its incidence has substantially increased over the past five years, accompanied by a gradual increase in CRC surgeries over the five past years in Japan[2]. Postoperative patients are reported to be at an increased risk of developing metachronous CRC, which is newly diagnosed CRC[3, 4]. The cumulative incidence of metachronous CRC is approximately 1.8% after 5 years in the Western world, rising to 3.4% after 10 years[5]. Therefore, quantitative evaluation of CRC development potential in postoperative patients is strongly required.

Accumulating evidence indicates that the carcinogenesis and progression of CRC are linked to the gut microbiota, in addition to genetic and other environmental factors. The gut microbiota and its metabolites not only reflect cancerous intestinal conditions but also directly affect CRC carcinogenesis. For instance, our previous study showed that the abundance of several species and their estimated growth rates were elevated along with CRC progression[6]. Several studies that compared CRC tissue with non-CRC mucosa

from the same patients showed that most of the amino acids (e.g., serine) were elevated in CRC tissue[7, 8]. Altogether, these results suggest that the presence of tumors, especially large malignant tumors in the advanced stage of CRC, and associated changes in cancerous intestinal conditions (e.g., stricture and inflammation) might lead to further enrichment of several bacteria and bacterial metabolites. In addition, several microbial activities might result in CRC carcinogenesis through host DNA damage. For example, genotoxic hydrogen sulfide producers, such as *Bilophila wadsworthia* and *Desulfovibrio* spp., have promoted CRC carcinogenesis through DNA damage[9]. Deoxycholate (DCA), which is a secondary bile acid and can be biotransformed by a specific group of gut microbes (e.g., *Clostridium scindens* and *Clostridium hiranonis*), was also reported to be associated with liver cancer through DNA damage[10] and hypothesized to promote CRC carcinogenesis[11]. Overall, revealing the relationship between the gut microbiota and/or metabolites and cancerous intestinal conditions is necessary for reducing the risk of microbe-derived CRC carcinogenesis.

In addition, the gut microbiota structure might be associated with postoperative outcomes[12], including CRC[3, 4]. Despite the high risk of metachronous CRC in postoperative patients, endoscopic or surgical resection of tumors is still considered the only curative option for CRC treatment[13]. Whereas there is an advanced understanding of the role of the gut microbiota in CRC carcinogenesis, the influence of surgery on the gut microbiota and metabolites remains partially understood. If treatment does not eliminate the bacteria and metabolites promoting CRC carcinogenesis, these bacteria and metabolites might be associated with the development of metachronous CRC. At this time, there have been only a few studies that have compared the gut environment associated with pre- and postsurgical status[14, 15].

Here, we collected 170 fecal samples from 85 CRC patients pre- and approximately one year post-surgery to investigate the influence of surgery on the gut environment. We characterized the fecal metagenomic and metabolomic alterations between pre- and postsurgical status. We also compared the bacteria observed to be affected by surgery with previously reported CRC-associated bacteria[6]. Additionally, we developed a method for potential CRC risk assessment based on the gut microbiota and its genes and metabolomic compositions using a random forest algorithm and then applied it to postoperative patients. To uncover the influence of surgery on the human gut environment, this study aimed to characterize the alteration in the CRC-associated gut microbiota and metabolites as the results of surgical treatment and to evaluate potential CRC risk in postoperative patients.

Results

Patients characterization

We collected fecal samples from 96 CRC patients in the pre- and postsurgical (N = 85) or Endoscopic submucosal dissection (ESD) treatment (N = 11) groups (days: 373 ± 182 , median $\pm$ s.d., **Figure S1a**, **Table 1**, **Table S1**). Classification into five groups (Stage 0, Stage I, Stage II, Stage III, and Stage IV) was carried out according to the patient's clinicopathologic findings. This classification was based on the Union for International Cancer Control (UICC) TNM Classification of Malignant Tumors. Eleven CRC

patients (Stage 0, N = 11) underwent ESD treatment. Eighty-five patients (Stage 0, N = 2; Stage I, N = 20; Stage II, N = 37; Stage III, N = 24; Stage IV, N = 2) underwent surgical treatment (right-sided operation, N = 20; left-sided operation, N = 65). Furthermore, 22 CRC patients underwent chemotherapy after surgical treatment (Stage III, N = 21, Stage IV, N = 1). Based on the postoperative colonoscopic status, these 96 patients who underwent surgical resection or ESD were classified into 3 groups according to the colonoscopic and histopathologic findings: (1) no abnormality groups (N = 81, neither marked colonoscopic findings nor a few polyps, up to two small polyps (< 5 mm)); (2) MP (N = 14, multiple polypoid adenomas with low-grade dysplasia, more than three adenomas, mostly with more than high-grade dysplasia); and (3) Stage 0 (N = 1, a patient who had undergone surgical treatment had already developed metachronous CRC during fecal sampling).

Surgical treatment alters the gut microbiota and metabolites

We first verified whether the compositions of the gut microbiota and its genes and metabolites were changed between pre- and postsurgical treatment groups by Permutational multivariate analysis of variance (PERMANOVA) based on the Bray-Curtis distance. The microbial genus profile was significantly different between the pre- and postsurgical treatment groups (Fig. 1a, $P = 1.00 \times 10^{-4}$, **Table S2**). The Kyoto Encyclopedia of Genes and Genomes (KEGG) Orthology (KO) and metabolome profiles also showed similar patterns (Fig. 1b, c, $P = 1.00 \times 10^{-4}$, $P = 2.00 \times 10^{-4}$, respectively).

Next, we examined how much the gut microbiota and its genes and metabolites changed with surgical treatment by comparing Bray-Curtis dissimilarity between the same individuals. Because fecal samples from our healthy controls at different time points were not obtained, we downloaded publicly available metagenome[16] and metabolome[17] data from healthy control samples at different time points (**see Methods, Table S3**). For taxonomic profiles, Bray-Curtis dissimilarity between the same patients in pre- and postsurgical status was significantly higher than that between healthy controls (Fig. 1d, one-sided Wilcoxon rank-sum test, $P = 5.44 \times 10^{-15}$, **Table S2**). Genes and metabolome profile also showed a similar pattern when we compared the surgical treatment group with healthy controls (Fig. 1e, f, one-sided Wilcoxon rank-sum test, $P = 1.21 \times 10^{-8}$, $P = 1.45 \times 10^{-5}$, respectively).

Human genome content was decreased in the post-surgical status

A fecal occult blood test is the current standard for CRC screening tests[18]. In addition, a study on inflammatory bowel disease reported that the amount of human DNA in fecal samples was elevated[19]. This study hypothesized that the DNA source was either epithelial cells or bloodshed into feces. Our previous study also showed that Stage 0, I/II, and III/IV patients had higher human genome content than healthy controls, and the level was gradually elevated along with CRC progression[6].

We investigated the changes in the human genome content between pre- and postsurgical treatment samples. The human genome content was significantly decreased in the postsurgical treatment sample

(Figure S2c, one-sided Wilcoxon signed-rank test, $P = 2.47 \times 10^{-4}$, Table S2). We also observed a higher human genome content in right-sided colon cancer than in left-sided colorectal cancer (Figure S2d, one-sided Wilcoxon rank-sum test, $P = 0.0466$).

Microbial alterations between pre- and post-surgical treatment

We identified 114 species with relative abundances that were significantly ($P < 0.005$) different between pre- and postsurgical treatment, of which 90 species showed significantly decreased abundances while 24 species showed significantly increased abundances (Fig. 2a, one-sided Wilcoxon signed-rank test, Table S2).

We hypothesized that species that were elevated in each of the multistep CRC progression stages were decreased in postsurgical status. To assess this hypothesis, we compared healthy controls in our cohort[6] with patients with multistep CRC progression (MP, Stage 0, Stage I/II, and Stage III/IV) stages using a one-sided Wilcoxon rank-sum test. Ninety species were significantly ($P < 0.005$) elevated in at least one of the multistep CRC progression stages (one-sided Wilcoxon rank-sum test, Table S2). Among the 90 significantly elevated species, 14 species were significantly decreased in postsurgical status. In particular, universal CRC markers in a meta-analysis of CRC[20], *Parvimonas micra* [ref_mOTU_v2_1145], *Gemella morbillorum* [ref_mOTU_v2_4513], *Fusobacterium nucleatum* subsp. *animalis* [ref_mOTU_v2_0776], and *Peptostreptococcus stomatis* [ref_mOTU_v2_4614] were significantly elevated in Stage I/II and Stage III/IV and decreased in postsurgical status (Fig. 2a, b). Unknown *Dialister* [meta_mOTU_v2_5867], unknown *Peptostreptococcaceae* [meta_mOTU_v2_5742], and *Solobacterium moorei* [ref_mOTU_v2_0531] were also significantly elevated in Stage I/II and Stage III/IV and decreased in postsurgical status (Fig. 2a). These species were also reported as universal CRC markers in a meta-analysis of CRC[20].

Furthermore, we also examined 24 species that were significantly ($P < 0.005$) increased in postsurgical status (one-sided Wilcoxon signed-rank test, Table S2). *Ruminococcus gnavus* [ref_mOTU_v2_0280], the 3-oxo-DCA-producing species from DCA[21], *B. wadsworthia* [ref_mOTU_v2_1149], a genotoxic hydrogen sulfide-producing species using taurine from taurocholate[9], *C. scindens* [ref_mOTU_v2_0883], a genotoxic DCA-producing species[22], were also increased in postsurgical status (Fig. 2b).

Changes in the estimated bacterial growth rate between pre- and postsurgical treatment

We carried out an estimation of the growth rate in 24 species that showed significantly increased relative abundances in postsurgical status by using GRiD[23], then compared it in pre- and postsurgical status. Four species were elevated in postsurgical status (one-sided Wilcoxon rank-sum test, $P < 0.05$, Table S2). Two species belonged to *Clostridium*, namely, *Clostridium clostridioforme*, *C. scindens*, and two species (*R. gnavus* and *Dialister invisus*), which were elevated in postsurgical status (Fig. 2c).

Next, we investigated the association between bacterial relative abundance and estimated growth rate. The relative abundances of four species (*R. gnavus*, *C. scindens*, *Dialister invisus*, and *Bifidobacterium longum*) were positively correlated with their estimated growth rates in pre- and postsurgical treatment samples (both Spearman correlation coefficients, > 0.4 , **Table S2**). On the other hand, the relative abundances of two species (*Flavonifractor plautii* and *B. wadsworthia*) were negatively correlated with their estimated growth rates in pre- and postsurgical treatment samples (both Spearman correlation coefficients, < -0.4).

Changes in the fecal metabolome profile between pre- and postsurgical status

A total of 60 metabolites were significantly ($P < 0.005$) altered between pre- and postsurgical status (Fig. 3a, one-sided Wilcoxon signed-rank test, **Table S2**). First, we focused on 43 metabolites that were significantly decreased in postsurgical status. Among them, 13 metabolites were significantly ($P < 0.005$) elevated in at least one of the multistep CRC progression stages, as previously reported[6] (one-sided Wilcoxon rank-sum test). In particular, one of the amino acids, serine, was significantly elevated in Stages 0, I/II, and III/IV and decreased in postsurgical status (Fig. 3b). Furthermore, the amino acid-related metabolites Gly-Leu and urocanate were also significantly elevated in Stages I/II and III/IV and decreased in postsurgical status (Fig. 3b). These metabolites were reported as advanced-stage CRC markers[6]. The other amino acids, namely, methionine and cysteine, and the related metabolite, N,N-dimethylglycine were also significantly decreased in postsurgical status (Fig. 3a).

We also investigated 17 metabolites that were significantly ($P < 0.005$) increased in postsurgical status (one-sided Wilcoxon signed-rank test, **Table S2**). Interestingly, DCA, which is well known to be associated with CRC carcinogenesis[11], and cholate, which can be biotransformed to DCA, were significantly increased in postsurgical status (Fig. 3c).

Finally, we focused on the alteration of the other bile acids-related metabolites between pre- and postsurgical treatment. We examined changes in taurocholate and glycocholate between pre- and postsurgical treatment because these metabolites can be biotransformed to cholate by bacterial metabolism. The levels of taurocholate and glycocholate were not different between pre- and postsurgical status (Fig. 3c, Wilcoxon signed-rank test, $P = 0.568$, $P = 0.202$, respectively). Next, we calculated the amount of total bile acids as the sum of taurocholate, glycocholate, cholate, and DCA and then compared it between pre- and postsurgical status. The amount of total bile acids can be evaluated as the biosynthesis of bile acids or reabsorption capacity. The total bile acid content was significantly increased after surgery (Fig. 3d, one-sided Wilcoxon signed-rank test, $P = 1.86 \times 10^{-7}$, **Table S2**). Moreover, we calculated the ratio of DCA to cholate and then compared it between pre- and postsurgical status. This ratio may reflect the biotransformation rate to DCA from cholate. As a result, there was no significant difference between pre- and postsurgical status (Fig. 3e, one-sided Wilcoxon signed-rank test, $P = 0.396$, **Table S2**).

Microbial gene alterations between pre- and postsurgical status

Because cholate and DCA were significantly increased in postsurgical status, we focused on the genes associated with the biosynthesis of these metabolites.

First, we investigated alterations in bile salt hydrolase (K01442), which can biotransform taurocholate/glycocholate to cholate. The relative abundance of bile salt hydrolase was not significantly different between pre- and postsurgical status (Fig. 4a, Wilcoxon signed-rank test, $P = 0.277$, **Table S2**).

Next, we examined alterations in the *bai* operon, which can biotransform cholate to DCA. A previous study reported that the *bai* operon is composed of eight genes, of which six genes (*bai B*, *CD*, *E*, *F*, *A*, and *H*) are required for biotransformation of DCA to cholate[24]. We applied a previously described similar approach[20], a metagenome annotation framework based on hidden Markov models (HMMs)[25] and homology search to identify and quantify the *bai* operon from the metagenome-assembled genomes (MAGs)-based gene catalog (**See Supplementary information**). Using the MAGs-based gene catalog has advantages for accurate annotation of genes based on gene arrangement and quantifying gene abundance based on operons. First, we succeeded in the identification of 11 DCA-producing species that contain six *bai* genes in 22547 MAGs derived from 716 metagenomic samples [cutoffs for quality of genomes: completeness > 50 & contamination < 10] (**Table S4**). Among them, seven MAGs consisted of six *bai* genes in the same contig (Fig. 4c). Next, the quantification of each *bai* gene was carried out by using the sum of each gene derived from the seven MAGs. Finally, we compared the abundance of the *bai* operon between pre- and postsurgical status. As a result, the *bai* operon was significantly increased in postsurgical status (Fig. 4d, one-sided Wilcoxon signed-rank test, $P = 4.88 \times 10^{-3}$, **Table S2**). Additionally, a majority of the contributors were *C. scindens* (Fig. 4e).

Influence of differences in the type of operation (right-sided versus left-sided) on bile acid metabolism

Right-sided operations not only resected the right colon but also part of the terminal ileum (Fig. 5a, b), which might result in the inhibition of bile acids reabsorption[26]. Thus, we also investigated the influence of the difference in operations (right-sided vs. left-sided) on bile acid metabolism (right-sided, $N = 20$; left-sided, $N = 65$).

First, we examined the differences in the alterations in the secondary bile acid (DCA), primary bile acids (cholate, glycocholate, and taurocholate), the ratio of DCA to cholate, and total bile acids between pre- and postsurgical treatment for different operation types. DCA was increased in both patients who underwent right-sided and patients who underwent left-sided operations (Fig. 5d, one-sided Wilcoxon signed-rank test, $P = 0.0301$, $P = 4.00 \times 10^{-5}$, respectively, **Table S2**). Cholate was also significantly increased in patients after both right-sided and left-sided operations (Fig. 5d, one-sided Wilcoxon signed-rank test, $P = 1.68 \times 10^{-4}$, $P = 5.31 \times 10^{-4}$, respectively). Glycocholate was significantly increased in patients

post-right-sided operation (Fig. 5d, one-sided Wilcoxon signed-rank test, $P = 1.54 \times 10^{-3}$). Taurocholate was increased in patients post-right-sided operation, whereas it tended to be decreased in patients post-left-sided operation (Fig. 5d, one-sided Wilcoxon signed-rank test, $P = 0.0327$, $P = 0.0366$, respectively). The concentration of total bile acids was significantly increased in patients after both right-sided and left-sided operations (Fig. 5e, one-sided Wilcoxon signed-rank test, $P = 3.62 \times 10^{-7}$, $P = 1.30 \times 10^{-4}$, respectively).

Next, we investigated the differences in the alterations in bile acid-related genes and *C. scindens*, the major contributor of the *bai* operon between pre- and postsurgical treatment for different operation types. Bile salt hydrolase was not significantly different both patients who underwent right-sided and patients who underwent left-sided operations (Fig. 5g, Wilcoxon signed-rank test, $P = 0.956$, $P = 0.179$, respectively, **Table S2**). The relative abundance of the *bai* operon and *C. scindens* were significantly increased in patients post-left-sided operation (Fig. 5h, i, one-sided Wilcoxon signed-rank test, $P = 1.90 \times 10^{-3}$, $P = 4.85 \times 10^{-3}$). The estimated growth rate of *C. scindens* was higher in patients post-left-sided operation than in patients pre-left-sided operation (Fig. 5j, one-sided Wilcoxon rank-sum test, $P = 0.036$).

Estimation of potential CRC risk based on the gut microbiota and metabolome profiles

In our previous study, we identified metagenomic and metabolomic signatures of CRC progression and examined their potential as CRC diagnostic biomarkers[6]. These signatures can be used in a statistical framework to calculate the CRC probability of a new patient having CRC[20, 27, 28]. We refer to CRC probability as ‘normalized probability’ hereafter.

We built non-stage (all CRC stages together) and stage-specific random forest-based binary classifiers to distinguish cases in each of the stages (CRC, N = 201, MP, N = 61; Stage 0, N = 57; Stage I/II, N = 85; Stage III/IV, N = 59) from healthy controls (H, N = 245) (**see Methods, Table S5**). Notably, a number of species, including *P. micra* [ref_mOTU_v2_1145], *P. stomatis* [ref_mOTU_v2_4614], *F. nucleatum* subsp. *animalis* [ref_mOTU_v2_0776], and *G. morbillorum* [ref_mOTU_v2_4513], were identified as top metagenomic and metabolomic signatures in CRC, Stage I/II, and Stage III/IV classifiers (Fig. 6a).

While surgical treatment involves resection of CRC lesions, postoperative patients are reported to be at increased risk of developing metachronous CRC[3, 4]. We hypothesized that the gut microbiota and metabolites would not be improved to the same extent as those of healthy controls after tumor resection. Therefore, we applied estimation of potential CRC risk methods in 83 pairs of pre- and postsurgical metagenomic and metabolomic samples, then examined the alterations in normalized probability between pre- and postsurgical treatment. In a non-stage-specific CRC classifier, the resulting normalized probabilities were significantly decreased in postsurgical treatment samples compared to pretreatment samples (Fig. 6b **panels Classifier and Applied “CRC”**, one-sided Wilcoxon signed-rank test, $P = 5.43 \times 10^{-9}$, **Table S2**). In stage-specific classifiers, for Stage I/II and Stage III/IV cases, the normalized probability was also significantly decreased in postsurgical treatment samples compared to presurgical treatment

samples (Fig. 5b **panels Applied “Stage I/II” and “Stage III/IV”**, one-sided Wilcoxon signed-rank test, $P=7.29\times10^{-5}$, $P=2.98\times10^{-8}$, respectively). However, for the MP and Stage 0 classifiers, the score of postsurgical treatment samples was not significantly altered from that of presurgical samples (Fig. 6c **panels Classifier “MP” Classifier “Stage 0”**, Wilcoxon signed-rank test, $P=0.245$, $P=0.325$, respectively). In particular, in the Stage 0 classifier, postsurgical treatment samples had a significantly higher median normalized probability than healthy control samples (one-sided Wilcoxon rank-sum test, $P=1.90\times10^{-13}$). These findings suggest that the normalized probability derived from the Stage 0 classifier could reflect the potential CRC risk in postoperative patients.

Application of the method to the estimation of CRC risk in gastrectomy patients

Gastrectomy patients have been reported to be at an increased risk of metachronous CRC[29]. Thus, we applied the Stage 0 classifier to estimate the potential CRC risk based on the gut microbiome and metabolome profiles from our gastrectomy cohort[30]. The normalized probability was significantly higher in the gastrectomy patients than in the healthy controls (**Figure S3 panel Stage 0**, one-sided Wilcoxon rank-sum test, $P=9.17\times10^{-10}$, **Table S2**). This finding indicates that normalized probabilities derived from the Stage 0 classifier could reflect the potential CRC risk in gastrectomy patients.

Discussion

The gut microbiota and its metabolites have been associated with CRC progression[31] and CRC carcinogenesis[9]. However, how surgical treatment alters the gut microbiota and metabolites is not completely understood. Here, we showed the alterations in the gut microbiota and metabolites due to surgical treatment for CRC. Our study indicated that 14 species (e.g., *F. nucleatum*) and 13 metabolites (e.g., serine), which were elevated at least in one of the multistep CRC progression stages, were significantly decreased in postsurgical status. Previous studies demonstrated that *F. nucleatum*, a protumorigenic bacterium[32], could potentiate CRC cells[31]. In addition, serine is an energy resource for the production of pyruvate in CRC tissue, which plays a role in CRC cell growth[33]. The species and metabolites that were decreased in postsurgical status, which were elevated in advanced-stage CRC, may reflect the improvement in cancerous intestinal conditions in postoperative patients. The decrease in other universal CRC markers derived from bacteria[20] (e.g., *P. micra*, *G. morbillorum*, and *P. stomatis*) and advanced-stage CRC markers derived from metabolites[6] (e.g., urocanate, Gly-Leu, and N,N-dimethylglycine) in postsurgical status support the idea that surgery might improve cancerous intestinal conditions.

In contrast, we also observed increased CRC carcinogenesis-associated metabolite, its producing genes, and its contributor in the post-surgical status. Our study showed that DCA, its producing genes (*bai* operon), and its contributors (*C. scindens*) significantly increased in the post-surgical status. Moreover, the estimated growth rate of *C. scindens* was also elevated in the post-surgical status. Several studies showed that DCA was associated with liver cancer[10] and hypothesized to promote CRC

carcinogenesis[11]. Altogether, the DCA enrichment in the postoperative patients might be associated with the increasing risk of developing metachronous CRC, suggesting that the surgical treatment incompletely recovers cancerous intestinal conditions.

In addition, the alteration in bile acid metabolism between pre- and postoperative status may be affected by the difference in the type of operation. In particular, right-sided operations not only resect the right colon but also part of the terminal ileum, which might result in the inhibition of bile acids reabsorption[26]. The specific influence of the right-sided operation is consistent with the increase in total bile acids, glycocholate, taurocholate, cholate, and DCA in patients post-right-sided operation. These findings were consistent with the high prevalence of diarrhea in patients who underwent the right-sided operation[34]. However, the concentration of total bile acids in fecal samples was also significantly increased in patients post-left-sided operation. There was no significant difference in total serum cholesterol, which reflects bile acid biosynthesis in the liver, between pre- and post-left-sided operative status (Fig. 5k, Wilcoxon signed-rank test, $P = 0.406$, **Table S2**). The mechanism for the enrichment of total bile acids in left-sided surgical treatment is yet to be understood; however, we hypothesized that the decrease in bile acid reabsorption capacity might be caused by the surgical operation in patients post-left-sided surgery. In addition, we observed an increase in cholate and a decrease in taurocholate in patients after the left-sided operation. Further work is needed to investigate the expression level of bile salt hydrolase to identify contributors, as there was no difference in the amount of the associated gene between pre-left-sided and post-left-sided operative status. The increase in *C. scindens* with the enrichment of cholate might result in an increase in DCA. The elevation of the estimated growth rate of *C. scindens* in patients post-left-sided operation and post-right-sided operation may support this idea. Trends for the decrease in the relative abundance of the *bai* operon and *C. scindens* in patients post-right-sided operation might be explained by the decrease in DCA-producing species in the right colon owing to resection[35]. Collectively, these results suggested that modification of the cholate concentration in the colorectum could be a potential target to reduce the DCA concentration.

A previous study demonstrated that the CRC probability based on the gut microbiota from a CRC classifier in CRC cases was decreased after surgical treatment[15]. Thus, we also developed a method to assess the potential CRC risk based on the gut microbiome and metabolome compositions using a random forest algorithm. Our results showed that in CRC, Stage I/II, and Stage III/IV cases, normalized probabilities derived from CRC, Stage I/II, and Stage III/IV classifiers were significantly decreased in postsurgical status. These suggest that normalized probabilities derived from CRC, Stage I/II, and Stage III/IV classifiers might reflect cancerous intestinal conditions caused by the presence of tumors. Trends of elevated normalized probabilities derived from these classifiers in multistep CRC progression samples and contributors (*P. micra*, *P. stomatis*, *G. morbillorum*, and *F. nucleatum*) of these classifiers along with CRC progression support the idea that these classifiers might reflect cancerous intestinal conditions (**Figure S4a, d, e**, Fig. 5a). Further validation of these classifiers is important to develop a tool to diagnose the presence of CRC. Since MP and Stage 0 (intramucosal carcinoma) metagenomic signatures are derived from gut environments in the absence of tumors, they are more likely to be associated with the initiation of CRC carcinogenesis. Thus, we hypothesized that MP and Stage 0 classifiers could be used to

estimate potential early CRC risk based on the gut microbiota and metabolites. In fact, the normalized probabilities derived from MP and Stage 0 classifiers were not decreased in postsurgical status. Besides, postsurgical treatment cases also had significantly higher normalized probability derived from the Stage 0 classifier than the healthy controls. Taken together, these results support the idea that the normalized probability derived from the Stage 0 classifier can be used as a measurement to estimate the potential risk of CRC, particularly for postsurgical treatment cases. It is difficult, however, to further characterize the clinical relevance of the risk to reflect metachronous CRC.

We acknowledge that our study also presents a limitation. Our approach revealed the confounding effect of chemotherapy in surgery on the gut microbiome and on metabolome alterations (**Figure S5, Table S2**), however, could not perfectly eliminate the possible confounding effect of antibiotics. Patients who underwent surgery were treated used antibiotics within 2 days before and after surgical treatment to prevent infection by pathogenic bacteria. However, a previous study indicated that the majority of gut microbiota compositions recovered to preantibiotic baseline levels within 1.5 months after treatment[36]. The fecal samples in our study were collected approximately one year after treatment. Thus, antibiotic usage might not have affected a majority of the observed differences.

In conclusion, our study led to an understanding of associations between the gut microbiota and its metabolites with cancerous intestinal conditions. We found incomplete recovery of gut environmental conditions by surgical treatment. The developed method could be used to estimate potential CRC risk. Our findings showed that species and metabolites that were elevated in advanced-stage CRC were significantly decreased after surgery. However, DCA, which is known to be associated with CRC carcinogenesis, and its biosynthetic genes and producing species were significantly increased in postsurgical status. To uncover the detailed molecular mechanism underlying the effect of DCA in metachronous CRC development, further follow-up studies are required. Our methods for potential CRC risk estimation might be used for risk assessment of metachronous CRC in postoperative patients.

Methods

Study participants and fecal sample collection

This study was conducted with participants undergoing total colonoscopy in the National Cancer Center Hospital, Tokyo, Japan. We excluded patients with hereditary or suspected hereditary diseases (e.g., lynch syndrome and familial adenomatous polyposis) from this study. Fecal samples were collected from 248 healthy controls, 67 MP, and 257 CRCs participants (Stage 0, N=73; Stage I/II, N=110; and Stage III/IV, N=74) as mentioned in our previous study[6]. Our previous study reported a difference in gut microbiota profiles between pre- and post-surgical treatment in 28 CRC patients[6]. Additionally, fecal samples from 57 CRC patients were newly collected in this study to investigate the differences in gut microbiome and metabolome between pre- and post-surgical treatment in detail (**Table S1**). In total, we collected fecal samples from these 85 CRC patients in approximately one year after they underwent surgical treatment (patients in Stage 0/I/II/III/IV). Among them, 20 CRC patients underwent the right-sided operation, while

65 CRC patients underwent the left-sided operation. Furthermore, 22 CRC patients underwent chemotherapy after surgical treatment (patients in Stage III/IV). Capecitabine, oxaliplatin, and FOLFOX were used for chemotherapy. None of the patients underwent radiation therapy. In addition, we collected 11 fecal samples from CRC patients who underwent ESD before and about one year after the treatment.

Fecal samples were collected immediately at first defecation after bowel cleansing at the hospital[37]. We then placed the fecal samples directly on dry ice and stored them at -80°C for metagenomic and metabolomic analyses. Colonoscopy was performed after bowel cleansing.

Whole-genome shotgun sequencing and quantification of fecal metabolites

Whole-genome shotgun sequencing was carried out on the extracted genomic DNA from fecal samples (**Supplementary Methods**). Capillary electrophoresis time-of-flight mass spectrometry (CE-TOF MS) was performed to quantify the fecal metabolites as previously described[38–40] (**Supplementary Methods**). We quantified 397 metabolites in 694 samples (**Table S1**).

Taxonomic and functional profiling

We obtained high-quality reads from raw reads (**Supplementary Methods, Figure S1b, Table S6**). Reads that were mapped to the human genome were subsequently removed (**Figure S2a, b**).

High-quality reads were used to obtain taxonomic and functional profiles. Taxonomic profiles were generated by mOTUs2 profiler version 2.0.0[41] (**Supplementary Methods**). Functional profiles were generated by our *in-house* pipeline. In brief, we generated MAGs and its driven gene catalog to quantify gene and KO[42] profiles (**Supplementary Methods**).

Estimation of microbial replication rates

False positives in the comparison of bacterial relative abundance could occur because alterations in the abundance of one bacterium could lead to changes in the relative abundances of other bacteria[43]. Therefore, we estimated the replication rates using the GRiD[23] version 1.2 with parameter single. The basic algorithm is based on the calculation of the ratio between coverage at the peak (*ori*) and trough (*ter*) on the reference genome as the replication rate because the bacterial genome is bidirectionally replicated from the origin of replication (*ori*) to the terminus region (*ter*). GRiD can estimate the growth rate using draft genome through reordering multiple contigs using these coverage depths at low sequencing coverage (coverage > 0.2).

The estimation of growth rate using GRiD required the reference genome and coverage of each species. Estimation of replication rates was successfully computed for 19 out of 24 species that had significantly

($P < 0.005$) increased their relative abundance after surgical treatment (one-sided Wilcoxon signed-rank test). These mOTU-annotated species corresponded to their MAGs based on NCBI taxonomy and utilized as reference genomes (**Table S2**).

For the statistical test, samples with low coverage (coverage < 0.2) were omitted because its replication rate could not be estimated. Among 19 species, two species were omitted (*Lactobacillus casei* and *Pyramidobacter piscolens*) owing to the low detections in total samples (N=120, N=62 in 716 samples, respectively).

Building the classifier

To estimate the potential CRC risk of each metagenomic and metabolomic sample, we built a classifier based on the random forest algorithm[44] utilizing species, KOs, and metabolome profiles. Because our previous study showed that the gut microbiome and metabolome profiles varied according to the stage of CRC progression[6], we built five classifiers (healthy controls vs. MP, Stage 0, Stage I/II, Stage III/IV, and CRC) using the RandomForestClassifier function from the sklearn.ensemble module in the Python package 'Scikit-learn' version 0.19.1. We did not use metagenomic and metabolomic samples from pre- and postsurgical treatment to build the classifiers.

We estimated the CRC probability using metagenome and metabolome profiles from each classifier. We first estimated the CRC probability in each sample, which was utilized for building each classifier. Leave-one-out cross-validation was carried out by the LeaveOneOut function from sklearn.model_selection in the Python package 'Scikit-learn'. Then, the CRC probability of each sample was estimated by the predict_proba function for each classifier. We then estimated the value in each sample that was not utilized for building each classifier by the predict_proba function for each classifier. We also estimated the CRC probability for pre- and postsurgical treatment samples and gastrectomy samples.

The CRC probability might be affected by the performance of the classifier and the ratio of the number of samples between healthy controls and each category in each classifier. Thus, we computed the difference between the CRC probability and minimum CRC probability. The value was later divided by the difference between the maximum and minimum probability in each classifier to obtain the normalized probability.

The decision boundary in each classifier was determined by the median of the normalized probability in healthy controls for each classifier. We defined the samples as having a high potential CRC risk if their normalized probability was higher than the decision boundary. Classifier contributors of each classifier were determined based on the RandomForestClassifier feature importance function sklearn.ensemble module from the Python package 'Scikit-learn'.

For evaluation of the classifier performance, each classifier was validated by 10 randomized 10-fold stratified cross-validations, which were carried out by the StratifiedKFold function from sklearn.model_selection in the Python package 'Scikit-learn'. The accuracy of each classifier was

examined using the average AUC in each test. We optimized the filter of average relative abundance or concentration, `n_estimators`, and `max_depth` hyperparameter to maximize the AUC in each classifier for species, KOs, and metabolome profiles (**Table S6**). For the combination model, we also optimized the `n_estimators`, `max_depth`, and number of contributors derived from each optimized classifier (**Table S6**).

Visualization of the taxonomic tree

A taxonomic tree based on the mOTUs2 maker genes was downloaded from <https://motu-tool.org/data/mOTUs.treefile>. We filtered out low-abundance mOTUs (average of mOTUs in all metagenomic samples $< 10^{-6}$) and nonsignificantly different mOTUs based on the statistical test between different groups (e.g., $P > 0.005$) from the taxonomic tree by the `drop.tip` function in the R package 'ape'. The taxonomic tree in the species was later visualized by iTOL version 5.5[45].

Statistical analyses

Low-abundance microbial features (species and KOs) were discarded [cutoffs of average relative abundance in 716 samples: species < 0.00001 , KO < 0.0000001]. The metabolites that were detected in less than 5% of 694 samples were also discarded.

To investigate the overall effect of ESD and surgical treatments in the gut microbiome and metabolome profiles, Bray-Curtis dissimilarity analysis within the same individual was carried out by the `vegdist` function parameter with the method `bray` in the R package 'vegan'. PERMANOVA was also performed to investigate the difference between pre- and postsurgical treatment groups by the `adonis` function parameter with 9999 permutations and the `bray` method in the R package 'vegan'.

Nonmetric multidimensional scaling (NMDS) was carried out to visualize the effects of surgical treatment by the `metaMDS` function parameters with 20 `trymax` from the R package 'vegan'.

To identify the microbial feature alterations due to surgical treatment, a one-sided Wilcoxon signed-rank test was performed by the `wilcoxsign_test` function parameter with the exact distribution from the R package 'coin' (**Table S2**). To identify CRC-associated microbial features, we compared healthy controls with each of the multistep CRC progression stage (MP, Stage 0, Stage I/II, Stage III/IV) groups using a one-sided Wilcoxon rank-sum test by the `Wilcox_test` function from the R package 'coin' (**Table S2**).

$P < 0.005$ was considered statistically significant in all statistical tests. Furthermore, all P values were corrected by the Benjamini-Hochberg method, which is an estimation of the false discovery rate (FDR), to obtain the FDR-corrected P value (q value) (**Table S2**).

Declarations

Acknowledgements

We are thankful to all the participants and their families who participated in this study. We are thankful to Mr. Z. Nakagawa for inspiring discussions and I. Take, M. Sezawa, M. Iwahara, and M. Komori for expert technical assistance.

Ethics approval and consent to participate

Informed consent was obtained from the institutional review boards of each participating institute (National Cancer Center, 2013-244; Tokyo Institute of Technology, 2014018, Keio University, Shonan Fujisawa Campus, 78).

Competing of interests

S.F. and T.Y. are founders of Metabologenomics. The company is focused on the design and control of the gut environment for human health. The company had no control over the interpretation, writing or publication of this work. The terms of these arrangements are being managed by Keio University and Tokyo Institute of Technology according to their conflict of interest policies.

Consent for publication

Not applicable.

Availability of data and material

Nucleotide sequences are available in the DDBJ Sequence Read Archive (DRA) as DRA006684, DRA008156, and DRA011152. The cohort from Erawijantari *et al.* is available in the DRA as DRA007281 and DRA008243. The cohort from Voigt *et al.* is available in the European Nucleotide Archive (ENA) database as ERP009422. The 22547 MAGs derived from 716 metagenomic samples are available at

http://matsu.bio.titech.ac.jp/CRC_MAGs/CRC_MAGs.tar.gz.

Funding

This work was supported in part by grants from the National Cancer Center Research and Development Fund (2020-A-4), the Japan Agency for Medical Research and Development (AMED) (JP18ek0109187 to S.F., S.Y., and T.Y.; JP19cm0106464 to S.Y. and T.Y.; JP20gm1010009 to S.F.; JP20ck0106546 to Y.S., S.Y., and T.Y.; JP20 cm0106477 to S.S., S.Y., and T.Y.; and 20jk0210009 to T.S. and S.Y.); JST-ERATO (JPMJER1902 to S.F.); JST-AIP AccelerationResearch (JPMJCR19U3 to S.Y. and T.Y.); JSPS (Japan Society for the Promotion of Science) KAKENHI (142558 and 221S0002 to T.Y.; 18H04805 to S.F.; 20H03662 to S.Y.); the Food Science Institute Foundation (to S.F.); the Program for the Advancement of

Research in Core Projects under Keio University's Longevity Initiative (to S.F.); Integrated Frontier Research for Medical Science Division, Institute for Open and Transdisciplinary Research Initiatives, Osaka University (to S.Y.); Joint Research Project of the Institute of Medical Science, the University of Tokyo (T.S. and S.Y.); the Takeda Science Foundation (to S.Y. and S.F.); the Suzuken Memorial Foundation (to S.Y.); Yasuda Memorial Medical Foundation (to S.Y.) and Yakult Bio-Science Foundation (to S.Y.).

Author's contributions

H.S., S.Y., T. Shibata. and T.Y. contributed to the study concept and design. S.Y., S.S., H.T., M.Y., T. Sakamoto., Y.K. and Y.S. collected clinical samples and information. T. Soga. and S.F. performed metabolome analysis. H.S. performed bioinformatics analyses on metagenome and metabolome data. H.S., P.P.E., S.Y., T.Y., S.S., S.F. and S.M. wrote the manuscript. S.Y. and T.Y. supervised the study. All authors read and approved the final manuscript.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* 2018;68:394–424.
2. Kakeji Y, Takahashi A, Udagawa H, Unno M, Endo I, Kunisaki C, et al. Surgical outcomes in gastroenterological surgery in Japan: Report of National Clinical database 2011–2016. *Annals of Gastroenterological Surgery.* Wiley-Blackwell; 2018;2:37.
3. Yabuuchi Y, Imai K, Hotta K, Ito S, Kishida Y, Yamaguchi T, et al. Higher incidence of metachronous advanced neoplasia in patients with synchronous advanced neoplasia and left-sided colorectal resection for colorectal cancer. *Gastrointest Endosc.* 2018;88:348–59.e1.
4. Mulder SA, Kranse R, Damhuis RA, Ouwendijk RJT, Kuipers EJ, van Leerdam ME. The incidence and risk factors of metachronous colorectal cancer: an indication for follow-up. *Dis Colon Rectum.* 2012;55:522–531.
5. Bouvier AM, Latournerie M, Jooste V, Lepage C, Cottet V, Faivre J. The lifelong risk of metachronous colorectal cancer justifies long-term colonoscopic follow-up. *Eur J Cancer.* 2008;44:522–527.
6. Yachida S, Mizutani S, Shiroma H, Shiba S, Nakajima T, Sakamoto T, et al. Metagenomic and metabolomic analyses reveal distinct stage-specific phenotypes of the gut microbiota in colorectal cancer. *Nat Med.* 2019;25:968–976.
7. Hirayama A, Kami K, Sugimoto M, Sugawara M, Toki N, Onozuka H, et al. Quantitative metabolome profiling of colon and stomach cancer microenvironment by capillary electrophoresis time-of-flight mass spectrometry. *Cancer Res.* 2009;69:4918–4925.
8. Mal M, Koh PK, Cheah PY, Chan ECY. Metabotyping of human colorectal cancer using two-dimensional gas chromatography mass spectrometry. *Anal Bioanal Chem.* 2012;403:483–493.

9. Ridlon JM, Wolf PG, Gaskins HR. Taurocholic acid metabolism by gut microbes and colon cancer. *Gut Microbes*. 2016;7:201–15.
10. Yoshimoto S, Loo TM, Atarashi K, Kanda H, Sato S, Oyadomari S, et al. Obesity-induced gut microbial metabolite promotes liver cancer through senescence secretome. *Nature*. 2013;499:97–101.
11. Ajouz H, Mukherji D, Shamseddine A. Secondary bile acids: an underrecognized cause of colon cancer. *World J Surg Oncol*. 2014;12:164.
12. Guyton K, Alverdy JC. The gut microbiota and gastrointestinal surgery. *Nat Rev Gastroenterol Hepatol*. 2017;14:43–54.
13. Al Bandar MH, Kim NK. Current status and future perspectives on treatment of liver metastasis in colorectal cancer. *Oncol Rep*. 2017;37:2553–2564.
14. Ohigashi S, Sudo K, Kobayashi D, Takahashi T, Nomoto K, Onodera H. Significant changes in the intestinal environment after surgery in patients with colorectal cancer. *J Gastrointest Surg*. 2013;17:1657–1664.
15. Sze MA, Baxter NT, Ruffin IV MT, Rogers MAM, Schloss PD. Normalization of the microbiota in patients after treatment for colonic lesions. *Microbiome*. 2017;5:150.
16. Voigt AY, Costea PI, Kultima JR, Li SS, Zeller G, Sunagawa S, et al. Temporal and technical variability of human gut metagenomes. *Genome Biol*. 2015;16:73.
17. Nagata N, Tohya M, Fukuda S, Suda W, Nishijima S, Takeuchi F, et al. Effects of bowel preparation on the human gut microbiome and metabolome. *Scientific Reports*. 2019;9:4042.
18. Levin B, Lieberman DA, McFarland B, Andrews KS, Brooks D, Bond J, et al. Screening and surveillance for the early detection of colorectal cancer and adenomatous polyps, 2008: a joint guideline from the American Cancer Society, the US Multi-Society Task Force on Colorectal Cancer, and the American College of Radiology. *Gastroenterology*. 2008;134:1570–1595.
19. Lewis JD, Chen EZ, Baldassano RN, Otley AR, Griffiths AM, Lee D, et al. Inflammation, Antibiotics, and Diet as Environmental Stressors of the Gut Microbiome in Pediatric Crohn's Disease [Internet]. *Cell Host & Microbe*. 2017; 18: 489–500.
20. Wirbel J, Pyl PT, Kartal E, Zych K, Kashani A, Milanese A, et al. Meta-analysis of fecal metagenomes reveals global microbial signatures that are specific for colorectal cancer. *Nat Med*. 2019;25:679–689.
21. Devlin AS, Fischbach MA. A biosynthetic pathway for a prominent class of microbiota-derived bile acids. *Nat Chem Biol*. 2015;11:685–690.
22. Ridlon JM, Harris SC, Bhowmik S, Kang D-J, Hylemon PB. Consequences of bile salt biotransformations by intestinal bacteria. *Gut Microbes*. 2016;7:22–39.
23. Emiola A, Oh J. High throughput in situ metagenomic measurement of bacterial replication at ultra-low sequencing coverage. *Nat Commun*. 2018;9:4956.
24. Funabashi M, Grove TL, Wang M, Varma Y, McFadden ME, Brown LC, et al. A metabolic pathway for bile acid dehydroxylation by the gut microbiome. *Nature*. 2020;582:566–570.

25. Eddy SR. Accelerated Profile HMM Searches. PLoS Comput Biol. 2011;7:e1002195.
26. Kurien M, Evans KE, Leeds JS, Hopper AD, Harris A, Sanders DS. Bile acid malabsorption: an under-investigated differential diagnosis in patients presenting with diarrhea predominant irritable bowel syndrome type symptoms. Scand J Gastroenterol. 2011;46:818–822.
27. Zeller G, Tap J, Voigt AY, Sunagawa S, Kultima JR, Costea PI, et al. Potential of fecal microbiota for early-stage detection of colorectal cancer. Mol Syst Biol. 2014;10:766.
28. Thomas AM, Manghi P, Asnicar F, Pasolli E, Armanini F, Zolfo M, et al. Metagenomic analysis of colorectal cancer datasets identifies cross-cohort microbial diagnostic signatures and a link with choline degradation. Nat. Med. 2019;25:667–678.
29. Eom BW, Lee HJ, Yoo MW, Cho JJ, Kim WH, Yang H-K, et al. Synchronous and metachronous cancers in patients with gastric cancer. J Surg Oncol. 2008;98:106–110.
30. Erawijantari PP, Mizutani S, Shiroma H, Shiba S, Nakajima T, Sakamoto T, et al. Influence of gastrectomy for gastric cancer treatment on faecal microbiome and metabolome profiles. Gut. 2020;69:1404–1415;
31. Kostic AD, Gevers D, Pedamallu CS, Michaud M, Duke F, Earl AM, et al. Genomic analysis identifies association of *Fusobacterium* with colorectal carcinoma. Genome Research. 2012;22:292–298.
32. Bullman S, Pedamallu CS, Sicinska E, Clancy TE, Zhang X, Cai D, et al. Analysis of *Fusobacterium* persistence and antibiotic response in colorectal cancer. Science. 2017;358:1443–1448.
33. Ohshima K, Nojima S, Tahara S, Kurashige M, Kawasaki K, Hori Y, et al. Serine racemase enhances growth of colorectal cancer by producing pyruvate from serine. Nat. Metabolism. 2020;2:81–96.
34. Bertelsen CA, Larsen HM, Neuenschwander AU, Laurberg S, Kristensen B, Emmertsen KJ. Long-term Functional Outcome After Right-Sided Complete Mesocolic Excision Compared With Conventional Colon Cancer Surgery: A Population-Based Questionnaire Study. Dis Colon Rectum. 2018;61:1063–1072.
35. Thomas LA, Veysey MJ, Bathgate T, King A, French G, Smeeton NC, et al. Mechanism for the transit-induced increase in colonic deoxycholic acid formation in cholesterol cholelithiasis. Gastroenterology. 2000;119:806–815.
36. Palleja A, Mikkelsen KH, Forslund SK, Kashani A, Allin KH, Nielsen T, et al. Recovery of gut microbiota of healthy adults following antibiotic exposure. Nat Microbiol. 2018;3:1255–1265.
37. Nishimoto Y, Mizutani S, Nakajima T, Hosoda F, Watanabe H, Saito Y, et al. High stability of faecal microbiome composition in guanidine thiocyanate solution at room temperature and robustness during colonoscopy. Gut. 2016;65:1574–1575.
38. Soga T, Baran R, Suematsu M, Ueno Y, Ikeda S, Sakurakawa T, et al. Differential metabolomics reveals ophthalmic acid as an oxidative stress biomarker indicating hepatic glutathione consumption. J Biol Chem. 2006;281:16768–16776.
39. Soga T, Igarashi K, Ito C, Mizobuchi K, Zimmermann HP, Tomita M. Metabolomic profiling of anionic metabolites by capillary electrophoresis mass spectrometry. Anal Chem. 2009;81:6165–6174.

40. Ishii C, Nakanishi Y, Murakami S, Nozu R, Ueno M, Hioki K, et al. A Metabologenomic Approach Reveals Changes in the Intestinal Environment of Mice Fed on American Diet. *Int J Mol Sci*. 2018;19:4079.
41. Milanese A, Mende DR, Paoli L, Salazar G, Ruscheweyh H-J, Cuenca M, et al. Microbial abundance, activity and population genomic profiling with mOTUs2. *Nat Commun*. 2019;10:1014.
42. Kanehisa M, Goto S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research*. 2000;28:27–30.
43. Knight R, Vrbanac A, Taylor BC, Aksenov A, Callewaert C, Debelius J, et al. Best practices for analysing microbiomes. *Nat Rev Microbiol*. 2018;16:410–22.
44. Breiman L. Random Forests. *Machine Learning*. 2001;45:5–32.
45. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res*. 2019;47:W256–W259.

Tables

Due to technical limitations, table 1 is only available as a download in the Supplemental Files section.

Figures

Figure 1

Distinct fecal microbiome, gene, and metabolome compositions between pre- and postsurgical treatment CRC patients. a-c, NMDS analysis based on the Bray-Curtis distance was carried out to visualize the influence of surgical treatment on the overall community structure of the genus-level taxonomic profile (a), KO-level functional profile (b), and metabolome profile (c) in pre- and postsurgical treatment samples. d-f, Violin and box plots show the Bray-Curtis dissimilarity of the genus-level taxonomic profile (d), KO-level functional profile (e), and metabolome profile (f) of fecal samples at two different time points within the same participants. The boxes in box plots represent 25th-75th percentiles, black lines indicate the median, and whiskers extend to the lowest and highest values within 1.5 times the interquartile range. The Voigt et al. datasets obtained from healthy individuals at two different time points were used as the control for taxonomic and functional profile analysis (d, e). The Nagata et al. datasets were used as control data from a healthy individual cohort for metabolomic analysis (f). PERMANOVA shows a difference between presurgical and postsurgical treatment in community structure in each profile (a, b, c). Statistical analysis was performed by a one-sided Wilcoxon rank-sum test (d, e, f). Significant differences are denoted as follows: +++: elevation ($P < 0.005$); ++: elevation ($P < 0.01$); +: elevation ($P < 0.05$); ---: depletion ($P < 0.005$); --: depletion ($P < 0.01$); -: depletion ($P < 0.05$).

Figure 2

Different enrichment patterns of the fecal microbiota between pre- and postsurgical treatment. a, Significantly different species between different groups were detected by two methods. First, the pre-surgical treatment (n=85) and post-surgical treatment (n=85) patients were compared by a one-sided Wilcoxon signed-rank test ($P < 0.005$). Second, the healthy controls (H: n=251) were compared with MP (MP: n=67), Stage 0 (0: n=73), Stage I/II (I/II: n=143), and Stage III/IV (III/IV: n=86) independently by a one-sided Wilcoxon rank-sum test ($P < 0.005$). Phylogenetic relationships among 197 significantly different species were visualized in a phylogenetic tree complemented with 3 layers of the heatmap. The edges in the phylogenetic tree represent five phyla (Firmicutes: green; Fusobacteria: pink; Actinobacteria: blue; Proteobacteria: orange; Bacteroidetes: purple). The first layer of the heatmap represents the relative abundance in each species (Legend). The second layer represents the P value, which was obtained by comparing healthy controls and each stage for each species (red: elevated; blue: depleted). The third layer also represents the P value, which was obtained by comparing pre- and postsurgical treatment groups for each species (red: increase; blue: decrease). b, c, Each box plot shows the relative abundances (b) or estimated growth rate (c) of species (M: multistep CRC progression; S: between pre- and postsurgical treatment). Each line in boxplot S shows alteration patterns between pre- and postsurgical treatment within the same patients (increase: red; decrease: blue; neither-increase nor decrease: green). The sizes of points in boxplot S reflect the distribution of the population in each category. The boxes in box plots M and S (c) represent the 25th-75th percentiles, black lines indicate the median, and whiskers extend to the lowest and highest values within 1.5 times the interquartile range. The number of samples is represented at the bottom of each category. A one-sided Wilcoxon rank-sum statistical test was performed to characterize elevation or depletion patterns in each stage compared to healthy controls (panel M, b, c), while a one-sided Wilcoxon signed-rank test was performed to characterize the increasing or decreasing trend in the pre- and postsurgical groups (panel S, b, c). Significant differences characteristics are denoted as follows: +++: elevation or increase ($P < 0.005$); ++: elevation or increase ($P < 0.01$); +: elevation

or increase ($P < 0.05$); ---: depletion or decrease ($P < 0.005$); --: depletion or decrease ($P < 0.01$); -: depletion or decrease ($P < 0.05$).

Figure 3

Different enrichment patterns of fecal metabolite between pre- and post-surgical treatment. a, Comparison of metabolite concentrations between the pre-surgical treatment ($n=83$) and post-surgical treatment ($n=83$) groups. The x-axis represents the average $\log_2$ transformed fold change between pre- and postsurgical treatment, while the y-axis represents the $-\log_{10}$ transformed P values obtained from the one-sided Wilcoxon signed-rank test. The horizontal dashed line shows a $-\log_{10}$ transformed P value of 0.005. The sizes of the circles represent the average concentration of each metabolite in pre- and postsurgical treatment samples (Legend). Elevation or depletion in at least one of the multistep CRC progression stages is colored red and blue, respectively. b-e, Each box plot shows the concentration of each metabolite (b, c), total bile acids (d), and the ratio of DCA to cholate (e) (M: multistep CRC progression; S: between pre- and postsurgical treatment). Each line in boxplot S shows alteration patterns between pre- and postsurgical treatment within the same patients (increase: red; decrease: blue; neither increase nor decrease; green). The sizes of points in boxplot S reflect the distribution of the population in each category. The boxes in box plots M and S represent the 25th-75th percentiles, black lines indicate the median, and whiskers extend to the lowest and highest values within 1.5 times the interquartile range. The number of samples is represented at the bottom of each category (b-e). A one-sided Wilcoxon rank-sum statistical test was performed to characterize elevation or depletion patterns in each stage compared to healthy controls (panel M, b-e). A one-sided Wilcoxon signed-rank statistical test was performed to characterize the increasing or decreasing trend in the pre- and postsurgical groups (panel S, b-d), while a

one-sided Wilcoxon rank-sum test was performed for e. Significant differences are denoted as follows: +++: elevation or increase ($P < 0.005$); ++: elevation or increase ($P < 0.01$); +: elevation or increase ($P < 0.05$); ---: depletion or decrease ($P < 0.005$); --: depletion or decrease ($P < 0.01$); -: depletion or decrease ($P < 0.05$).

Figure 4

Distinct enrichment patterns of bile acid-related genes between pre- and postsurgical treatment. a, d, Each box plot shows the relative abundances or $-\log_{10}$ transformed relative abundances of genes or operons, respectively (M: multistep CRC progression; S: between pre- and postsurgical treatment). Box plot M is ordered by multistep CRC progression stage (H: red; MP: yellow; Stage 0: green; Stage I/II: blue; Stage III/IV: purple). Box plot S is ordered by pre- and postsurgical treatment (pre: orange; post: white). The number of samples is represented at the bottom of each category. Each line in boxplot S shows alteration patterns between pre- and postsurgical treatment within the same patients (increase: red; decrease: blue). The sizes of points in boxplot S reflect the distribution of the population in each category. The boxes in box plots M and S represent the 25th-75th percentiles, black lines indicate the median, and whiskers extend to the lowest and highest values within 1.5 times the interquartile range. b, e, Each pie chart shows the percentage of contributors to genes or operons. The percentages of contributors are indicated in parentheses. c, Gene arrangement of the bai operon in seven metagenome-assembled genomes (MAGs). Samples are shown, and their category is indicated in parentheses. The color reflects the gene name (Legends). Taxonomic assignments for MAGs were carried out by GTDB-Tk (Methods). A one-sided Wilcoxon rank-sum statistical test was performed to characterize elevation or depletion patterns by comparing healthy controls and each stage (panel M, a, d). A one-sided Wilcoxon signed-rank statistical test was performed to characterize increasing or decreasing trends by comparing pre- and postsurgical treatments (panel S, a, d). Significant differences characteristics are denoted as follows: +++: elevation or increase ($P < 0.005$); ++: elevation or increase ($P < 0.01$); +: elevation or increase ($P < 0.05$); ---: depletion or decrease ($P < 0.005$); --: depletion or decrease ($P < 0.01$); -: depletion or decrease ($P < 0.05$).

Figure 5

Influence of right- or left-sided operations on bile acids metabolism. a, Overview of bile acid metabolism in the gastrointestinal tract. Metabolites were placed corresponding to their production sites. The black arrow represents biotransformation by bacterial metabolism. The red arrow represents the flow of bile acid reabsorption. The red plus sign represents the increase in metabolites or operons in the postsurgical treatment samples compared to the presurgical status (one-sided Wilcoxon signed-rank test, $P < 0.005$). b, The right-sided operation was defined by resection of not only the right colon but also part of the terminal ileum. c, The left-sided operation was defined by resection of part of the left colon. The black point represents the position of CRC (b, c). d-k, Each box plot shows the concentration of each metabolite (d, e); the ratio of DCA to cholate (f); the relative abundance of genes (g), operon (h) and species (i), the estimated growth rate of species (j), and total serum cholesterol (k). The sizes of the points in the boxplots reflect the distribution of the population in each category (d-k). The color of boxes in the box plot represents the right (green)- or left (yellow)-sided operation. Each line in the boxplot shows alteration patterns between pre- and postsurgical treatment within the same patients (increase: red; decrease: blue; neither-increase nor decrease: green). The boxes in box plots represent 25th-75th percentiles, black lines indicate the median, and whiskers extend to the lowest and highest values within 1.5 times the interquartile range. The number of samples is represented at the bottom of each category. A one-sided

Wilcoxon rank-sum statistical test was performed to characterize the elevation or depletion pattern between pre-right-sided operation and pre-left-sided operation samples or post-right-sided operation and post-left-sided operation samples (d-k). One-sided Wilcoxon signed-rank statistical test was performed to characterize the increase or decrease trend between pre- and post-right-sided operation samples or pre- and post-left-sided operation samples (d, e, g, h, k), while a one-sided Wilcoxon rank-sum test was performed in f and j. Significant difference characteristics are denoted as follows: +++: elevation or increase ($P < 0.005$); ++: elevation or increase ($P < 0.01$); +: elevation or increase ($P < 0.05$); ---: depletion or decrease ($P < 0.005$); --: depletion or decrease ($P < 0.01$); -: depletion or decrease ($P < 0.05$).

Figure 6

Application of classifier into samples obtained from pre- and post-surgical treatment. a, The heatmap shows the ranking of the top ten contributors in each classifier (healthy controls, n=245) vs. CRC (n=201), MP (n=61), Stage 0 (n=57), Stage I/II (n=85), and Stage III/IV (n=59), independently). The y-axis is ordered

by the ranking of contributors from the Stage III/IV classifier. Color represents the category of contributors (species: red; KO: blue; metabolite: black). Color and the number in the heatmap represent the ranking of contributors in each classifier (Legend). b, c, The classifiers were applied to the pre- and postsurgical treatment CRC samples, and the obtained normalized probability, which was based on the CRC probability, is shown on the y-axis (Methods). Each box plot represents the normalized probability from each classifier. Dashed lines show the median of normalized probability (red: healthy controls; black: each stage that was used in building the classifier). Normalized probability in each patient (CRC, Stage I/II, Stage III/IV, respectively) was obtained from each classifier (CRC, Stage I/II, Stage III/IV, respectively) (b). Normalized probability in surgical CRC patients was obtained from each classifier (MP, Stage 0, Stage I/II, and Stage III/IV) (c). Each box plot is ordered by pre- and postsurgical treatment. Each line in box plots shows alteration patterns between pre- and postsurgical treatment within the same patients (increase: red, decrease: blue). The sizes of the points in the boxplots reflect the distribution of the population in each category. The number of samples used in applying and building a classifier is represented at the bottom of each category. A one-sided Wilcoxon signed-rank statistical test was performed to characterize the increasing or decreasing trend between pre- and postsurgical treatment (b, c). Significant difference characteristics are denoted as follows: +++: increase ($P < 0.005$); ++: increase ($P < 0.01$); +: increase ($P < 0.05$); ---: decrease ($P < 0.005$); --: decrease ($P < 0.01$); -: decrease ($P < 0.05$).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementFigure1.pdf](#)
- [SupplementFigure2.pdf](#)
- [SupplementFigure3.pdf](#)
- [SupplementFigure4.pdf](#)
- [SupplementFigure5.pdf](#)
- [SupplementaryTable1.xlsx](#)
- [SupplementaryTable2.xlsx](#)
- [SupplementaryTable3.xlsx](#)
- [Table1.xlsx](#)
- [SupplementaryTable4.xlsx](#)
- [SupplementaryTable5.xlsx](#)
- [SupplementaryTable6.xlsx](#)
- [supplementaryinformation.docx](#)