



OPEN ACCESS

ORIGINAL ARTICLE

# Enteric fungal microbiota dysbiosis and ecological alterations in colorectal cancer

Olabisi Oluwabukola Coker,<sup>1</sup> Geicho Nakatsu,<sup>1</sup> Rudin Zhenwei Dai,<sup>1</sup> William Ka Kei Wu,<sup>1,2</sup> Sunny Hei Wong,<sup>1</sup> Siew Chien Ng,<sup>1</sup> Francis Ka Leung Chan,<sup>1</sup> Joseph Jao Yiu Sung,<sup>1</sup> Jun Yu<sup>1</sup>

► Additional material is published online only. To view please visit the journal online (<http://dx.doi.org/10.1136/gutjnl-2018-317178>).

<sup>1</sup>State Key Laboratory of Digestive Disease, Department of Medicine and Therapeutics, Li Ka Shing Institute of Health Sciences, CUHK Shenzhen Research Institute, The Chinese University of Hong Kong, Hong Kong

<sup>2</sup>Department of Anaesthesia and Intensive Care, The Chinese University of Hong Kong, Hong Kong

## Correspondence to

Dr Jun Yu, Department of Medicine and Therapeutics, Prince of Wales Hospital, the Chinese University of Hong Kong, Shatin, Hong Kong; [junyu@cuhk.edu.hk](mailto:junyu@cuhk.edu.hk)

Received 12 July 2018

Revised 30 October 2018

Accepted 30 October 2018

Published Online First

24 November 2018



© Author(s) (or their employer(s)) 2019. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

**To cite:** Coker OO, Nakatsu G, Dai RZ, et al. *Gut* 2019;**68**:654–662.

## ABSTRACT

**Objectives** Bacteriome and virome alterations are associated with colorectal cancer (CRC). Nevertheless, the gut fungal microbiota in CRC remains largely unexplored. We aimed to characterise enteric mycobiome in CRC.

**Design** Faecal shotgun metagenomic sequences of 184 patients with CRC, 197 patients with adenoma and 204 control subjects from Hong Kong were analysed (discovery cohort: 73 patients with CRC and 92 control subjects; validation cohort: 111 patients with CRC, 197 patients with adenoma and 112 controls from Hong Kong). CRC-associated fungal markers and ecological changes were also validated in additional independent cohorts of 90 patients with CRC, 42 patients with adenoma and 66 control subjects of published repository sequences from Germany and France. Assignment of taxonomies was performed by exact k-mer alignment against an integrated microbial reference genome database.

**Results** Principal component analysis revealed separate clusters for CRC and control ( $p < 0.0001$ ), with distinct mycobiomes in early-stage and late-stage CRC ( $p = 0.0048$ ). Basidiomycota:Ascomycota ratio was higher in CRC ( $p = 0.0042$ ), with increase in Malasseziomycetes ( $p < 0.0001$ ) and decrease in Saccharomycetes ( $p < 0.0001$ ) and Pneumocystidomycetes ( $p = 0.0017$ ). Abundances of 14 fungal biomarkers distinguished CRC from controls with an area under the receiver-operating characteristic curve (AUC) of 0.93 and validated AUCs of 0.82 and 0.74 in independent Chinese cohort V1 and European cohort V2, respectively. Further ecological analysis revealed higher numbers of co-occurring fungal intrakingdom and co-exclusive bacterial–fungal correlations in CRC ( $p < 0.0001$ ). Moreover, co-occurrence interactions between fungi and bacteria, mostly contributed by fungal Ascomycota and bacterial Proteobacteria in control, were reverted to co-exclusive interplay in CRC ( $p = 0.00045$ ).

**Conclusions** This study revealed CRC-associated mycobiome dysbiosis characterised by altered fungal composition and ecology, signifying that the gut mycobiome might play a role in CRC.

## INTRODUCTION

Gut microbiota is a complex ecosystem that plays an important role in the pathogenesis of colorectal cancer (CRC), a leading cause of cancer-related deaths worldwide.<sup>1 2</sup> Nevertheless, the role of

## Significance of this study

### What is already known on this subject?

- Donor gut microbiota from patients with colorectal cancer (CRC) has been shown to induce tumourigenesis in germ-deficient mice models.
- Bacteriome and virome alterations are associated with CRC.
- The involvement of gut fungal microbiota in CRC is largely unexplored.

### What are the new findings?

- We identified CRC-associated faecal fungal dysbiosis with increased Basidiomycota:Ascomycota ratio in patients with CRC compared with healthy subjects.
- Fungal class Malasseziomycetes was enriched in CRC while classes Saccharomycetes and Pneumocystidomycetes were depleted.
- Abundances of 14 fungal biomarkers distinguished CRC from controls with an area under the receiver-operating characteristic curve (AUC) of 0.93 and validated AUCs of 0.82 and 0.74 in two independent cohorts.
- Ecological analysis revealed a higher number of co-occurring fungal intrakingdom correlations and more co-exclusive correlations between fungi and bacteria in CRC compared with control indicating that synergistic intrafungal and antagonistic bacterial–fungal ecological associations may play role in colorectal carcinogenesis.

### How might it impact on clinical practice in the foreseeable future?

- Our analysis highlights the role of fungi in colorectal tumourigenesis and the potential use of fungi faecal markers in diagnosing CRC.

microbial components other than gut bacteria, such as fungi, is largely unexplored in CRC, partly due to their relatively lower abundance and lack of well-characterised reference genomes.<sup>3 4</sup> As a result, the understanding of their contributions towards colorectal carcinogenesis is limited. Increased capacity of high-throughput sequencing methods has provided access to non-bacterial components of the gut microbiota. In this regard, we recently

reported on virome signatures associated with CRC based on shotgun metagenomic analyses.<sup>5</sup> Although the fungal composition of the mammalian microbiota was estimated to comprise less than 1% of all commensal microbial species by genomic equivalence,<sup>4</sup> emerging studies have demonstrated its importance in the gut.<sup>6–11</sup> It assumes a stable role in the development and maintenance of host immune system, but can be altered in both intestinal and extraintestinal diseases.<sup>6–11</sup> Thus, investigation into enteric mycobiome in patients with CRC as compared with healthy subjects is worthwhile and would provide greater understanding of yet unknown mechanisms by which the gut microbiota contributes to colon tumourigenesis beyond the bacterial kingdom.

Distinct gut fungal communities and diversities were recently described as being associated with IBDs and liver cirrhosis in humans.<sup>8–9–12</sup> Moreover, the potential of commensal fungi to functionally replace intestinal bacteria in maintaining host immune responses in the gut by reversing disease susceptibility induced by bacterial depletion has been demonstrated in a mouse model.<sup>13</sup> Consistently, a recent study identified fungal compositional differences in tissue biopsies of 27 subjects with colorectal adenoma and adjacent tissues,<sup>6</sup> implicating altered host-associated fungal populations in the development of CRC. Understanding mycobial diversity and abundance in CRC compared with healthy subjects is therefore important to elucidate the potential contribution of fungal species in colorectal tumourigenesis. Importantly, faecal fungal dysbiosis might open up new opportunities for using fungal species as non-invasive diagnostic biomarkers for CRC.

Here, we characterised CRC-associated mycobiome by analysing faecal metagenomic sequences of 73 patients with CRC and 92 healthy subjects with validation in two independent cohorts of a total of 201 patients with CRC, 239 patients with adenoma and 178 controls. For the first time, we show that faecal mycobiome profile is predictive of CRC, independent of bacterial and other clinical parameters. Additionally, ecological co-occurrence network analysis demonstrates CRC-specific fungal intrakingdom and fungi–bacteria interactions.

## MATERIALS AND METHODS

### Patient cohorts

We analysed faecal shotgun metagenomic sequences from 585 individuals comprising faecal shotgun metagenomic sequences

of 184 patients with CRC, 197 patients with adenoma and 204 control subjects from Hong Kong (discovery cohort: 73 patients with CRC and 92 control subjects; independent Chinese validation cohort V1: 111 patients with CRC, 197 patients with adenoma and 112 controls from Hong Kong) who had undergone standard colonoscopy examinations at the Prince of Wales Hospital at the Chinese University of Hong Kong.<sup>14</sup> Presentations of digestive symptoms to the outpatient gastroenterology clinics and asymptomatic individuals ages over 50 years were the recruitment criteria. All CRC subjects had intact colonic lesions as at the time of stool collection. Stool samples were collected and stored at  $-20^{\circ}\text{C}$  within 4 hours and  $-80^{\circ}\text{C}$  within 24 hours for long-term storage. Qiagen QIAmp DNA Stool Mini Kit (Qiagen) was used for DNA extraction according to the manufacturers' instructions. The diagnostic potential of our fungal markers was further demonstrated in faecal metagenomics dataset comprising 90 CRC, 42 adenoma and 66 healthy subjects from Germany and France designated as European cohort V2.<sup>15</sup> The demographic and clinical details of all samples are shown in [table 1](#).

### Metagenomics sequencing

Whole-genome shotgun sequencing of Chinese discovery and validation samples were carried out on Illumina HiSeq 2000 (Illumina, San Diego, California, USA), and sequences from Germany and France were sequenced on HiSeq 2000/2500 platform (Illumina, San Diego, California, USA),<sup>15</sup> which are available on public repository. Briefly, bases with Phred quality of less than 3, reads with average of quality less than 15 for every 4 bases and reads with lengths less than 36 bases were discarded using Trimmomatic software (V.0.36).<sup>16</sup> Any adapter sequences in paired-end reads were clipped by checking for simple and palindromic matches of 10 and 30 bases with a library of universal Illumina Nextera and Truseq adapter sequences if an adapter-read pair had a mismatch count of 2 or less. Mammalian genomes (hg38, felCat8, canFam3, mm10, rn6, susScr3, galGal4 and bosTau8; UCSC Genome Browser)<sup>17</sup> and 6877 bacterial plasmids (NCBI RefSeq database accessed on 12 May 2016), 2116 complete plastomes (NCBI RefSeq database accessed on 12 May 2016) and 6093 UNiVec sequences (NCBI RefSeq database accessed on 12 May 2016), which are potential habitat-associated/laboratory-associated or extrachromosomal sequence contaminants, were removed after alignment using Bowtie2 V.2.2.9 with 'very-sensitive' default settings.

**Table 1** Demographic and clinical details of samples

Cohort	Factor	Control	CRC	Adenoma	P values	Sample collection	Sequencing platform
Chinese cohort 1 (discovery)	Sample size	92	73	NA	NA	No antibiotics and no invasive medical intervention for 3 months; no vegetarian diet; no history of cancer or inflammatory disease of intestine	Sequencing platform: Illumina HiSeq 2000/2500; sequencing target depth: 5 GB; read length: 100 bp
	Age	58.51 (7.55)	65.90 (10.61)	NA	<0.0001		
	Gender	Male: 51Female: 41	Male: 47Female: 26	NA	0.316		
	BMI	23.87 (3.31)	24.07 (3.18)	NA	0.697		
Chinese cohort 2 validation (V1)	Sample size	112	111	197	NA	No previous colon or rectal surgery, CRC, inflammatory or infectious injuries of the intestine; no need for emergency colonoscopy	
	Age	62.39 (6.51)	73.3 (10.50)	66.11 (5.94)	<0.0001		
	Gender	Male: 44 Female: 68	Male: 59 Female: 52	Male: 119 Female: 78	0.002		
	BMI	23.91 (3.34)	24.29 (3.17)	24.51 (3.48)	0.322		
German and French validation (V2)	Sample size	66	90	42	NA	No previous colon or rectal surgery, CRC, inflammatory or infectious injuries of the intestine; no need for emergency colonoscopy	
	Age	58.75 (12.96)	68.44 (12.22)	62.95 (9.11)	0.012		
	Gender	Male: 32Female: 32	Male: 53 Female: 35	Male: 30 Female: 12	0.087		
	BMI	24.72 (3.19)	25.89 (4.29)	25.90 (4.10)	0.056		

BMI, Body Mass Index; CRC, colorectal cancer; NA, not applicable.

### Sequence taxonomic annotation

Taxonomy was assigned to metagenomic reads using k-mer-based algorithms implemented in Kraken taxonomic annotation pipeline.<sup>18</sup> A custom database comprising 7505 bacterial and 1083 fungal reference genomes from NCBI (<http://www.ncbi.nlm.nih.gov>), FungiDB (<http://fungidb.org>), Ensemble (<http://fungi.ensembl.org>) and Broad Institute (<http://broadinstitute.org>) (accessed on 22 October 2017) was built using Jellyfish program by counting distinct 31-mer in the reference libraries, with each k-mer in a read mapped to the lowest common ancestor of all reference genomes with exact k-mer matches. Each query was thereafter classified to a taxon with the highest total hits of k-mer matched by pruning the general taxonomic trees affiliated with mapped genomes. The final metagenomic read counts were normalised by cumulative sum scaling method using *metagenomeSeq* R/Bioconductor package.<sup>19</sup>

### Statistical analysis

MaAsLin (multivariate association with linear models), a multivariate statistical framework,<sup>20</sup> was used to identify associations between clinical metadata and microbial abundance while controlling for confounders, namely age, gender and Body Mass Index (BMI). Co-occurrence and co-exclusion relationships within fungal and between fungal and bacterial kingdoms were estimated using SparCC algorithm,<sup>21</sup> known for its robustness to gut microbiome data compositionality. P values less than 0.05 after multiple comparisons correction using false discovery rate (FDR) method were considered significant. The inclusion criterion for plotting features was having a minimum absolute magnitude of interkingdom or intrakingdom correlation of 0.6 and FDR of less than 0.05 in either one or both of control and CRC. For visualisation, correlations with absolute coefficients below 0.3 were masked to show all signals that are exclusively significant. Taxonomies were sorted by the levels of phylum and each phylum was clustered separately by classes. Areas under receiver-operating characteristic curves (AUCs) were determined with standard linear regression using one-standard error rule method, and adjusted for age, gender and BMI. Two-tailed Mann-Whitney U test was used to determine statistically significant differences between case and control.

## RESULTS

### Enteric fungal dysbiosis is associated with CRC

The discovery cohort comprised 165 subjects, namely 73 patients with CRC and 92 healthy control subjects (table 1). About 10.34% of reads were filtered out by Trimmomatic trimming and contaminant removal steps leaving an average of 27 640 420 high-quality paired reads per sample for microbial taxonomic classification (online supplementary table 1). To validate the quality of our reference database, we analysed alpha and beta diversities of bacterial fraction of the microbiota. In agreement with our previous study,<sup>14</sup> alpha diversity indices were significantly reduced in CRC compared with control while beta diversity analysis showed separate clusters for CRC and control ( $p < 0.005$ , online supplementary figure 1). Rarefaction plots reached plateau for fungal species in both CRC and control samples (online supplementary figure 2). We then assessed and compared the enteric fungal composition in our cohort. Both unsupervised principal component analysis and supervised redundancy analysis (RDA) revealed that the gut mycobiome discriminated CRC and control into two significantly distinct groups ( $p = 0.0006$  and  $p = 0.001$ , respectively; figure 1A,B), indicating that CRC exhibits unique fungal profiles

from healthy subjects. Interestingly, RDA also revealed that the fungal components formed different clusters for control, early-stage (TNM stage I and II) CRC and late-stage (TNM stage III and IV) CRC, depicting a CRC-stage-specific disruption in mycobiome homeostasis ( $p = 0.001$ ; figure 1C). The fungal alpha diversities as measured by species richness and Shannon diversity, however, were not significantly different between CRC and control (figure 1D), unlike the observation with bacteria. Taken together, these results indicate alterations in gut fungal composition in association with colorectal carcinogenesis.

### Taxonomic alterations of enteric fungi in CRC

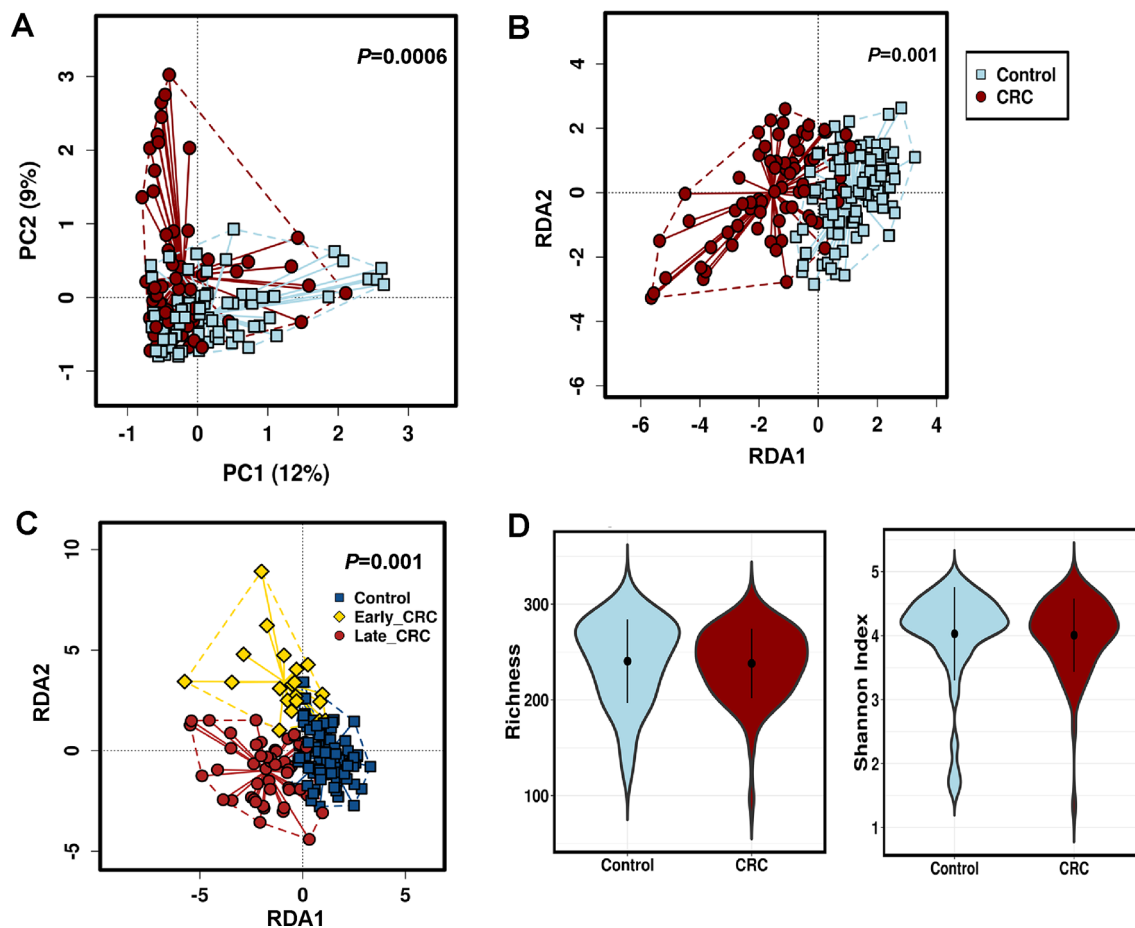
Consistent with previous studies and as a validation for our analysis, we observed enrichment and depletion of bacteria phyla Fusobacteria and Firmicutes, respectively, in CRC compared with control (online supplementary figure 1). Among the fungal taxa, the phylum Ascomycota dominated the mycobiota, while Basidiomycota was observed as the second most abundant phylum (figure 2A) in both case and control groups. Moreover, the Basidiomycota:Ascomycota ratio, which reflects fungal dysbiosis,<sup>12</sup> was skewed with higher values in CRC than control (figure 2B,  $p = 0.0042$ ). Additional differences were observed at lower taxonomic levels, including significant depletion of Saccharomycetes and Pneumocystidomycetes classes and enrichment of Malasseziomycetes class in CRC compared with control (figure 2C, online supplementary table 2, supplementary data). At the family level, Lipomycetaceae was significantly reduced in CRC while Pisolithaceae, Marasmiaceae, Malasseziaceae, Erysiphaceae, Pseudorotiaceae and Chaetomiacae were increased (figure 2D, online supplementary table 3). Furthermore, at the genus level, six fungal features were enriched in CRC, including *Malassezia*, *Moniliophthora*, *Rhodotorula*, *Acremonium*, *Thielaviopsis* and *Pisolithus* (figure 2E, online supplementary table 4).

### Fungal species as CRC diagnostic markers

To identify gut fungal taxa with potential value for CRC diagnosis, we investigated fungal alterations at the species level. After correction for confounders, namely age, gender and BMI, the abundances of 38 fungal species were observed to be significantly different in CRC when compared with control. Twenty-nine of these were CRC-enriched while nine others were CRC-depleted (figure 3A, online supplementary table 5). The diagnostic potential of species with greater than twofold abundance changes, including *Aspergillus flavus*, *Kwoniella mangrovensis*, *Pseudogymnoascus* sp. VKM F-4518, *Debaryomyces fabryi*, *A. sydowii*, *Moniliophthora perniciosus*, *K. heavenensis*, *A. ochraceoroseus*, *Talaromyces islandicus*, *Malassezia globosa*, *Pseudogymnoascus* sp. VKM F-4520, *A. rambellii*, *Pneumocystis murina* and *Nosemia apis* were investigated. We observed that these markers classified CRC and controls with an AUC of 0.93. They also distinguished control from early-stage CRC with an AUC of 0.91 (figure 3B).

### Validation of faecal fungal diagnostic markers of CRC in independent cohorts

The performance of the fungal classifiers identified in the discovery cohort was investigated in independent Chinese cohort V1, comprising 112 control, 197 adenoma and 111 CRC subjects. Interestingly, our fungal markers distinctly stratified CRC subjects from the control with an AUC of 0.82 and control from early-stage CRC, with an AUC of 0.81 (figure 4A). Furthermore, our fungal markers were validated in an ethnically different European cohort V2 (German and French; control,  $n = 66$ ; CRC,  $n = 90$ ; early-stage CRC,  $n = 32$ ), from which AUCs of 0.74 and 0.72 were obtained



**Figure 1** Alteration of enteric fungi in colorectal cancer (CRC). (A) Unsupervised principal component analysis of Bray-Curtis distance showing the stratification of CRC from control samples by fungi compositional profile. Control,  $n=92$ ; CRC,  $n=73$ . PC1 and PC2 represent the top two principal coordinates that reflect most of the diversity, given as a percentage. Groups were compared using Mann-Whitney U test. (B) Supervised multivariate redundancy analysis of fungi compositional profile stratifying patients with CRC ( $n=73$ ) from control ( $n=92$ ). RDA1 and RDA2 represent the top two coordinates that reflect most of the diversity, given as a percentage. Groups were compared using Mann-Whitney U test. (C) Stage-specific enteric fungal compositional multivariate redundancy analysis in control,  $n=92$ ; Early\_CRC,  $n=20$ ; Late\_CRC,  $n=54$ . RDA1 and RDA2 represent the top two coordinates that reflect most of the diversity, given as a percentage. Groups were compared using Kruskal-Wallis test. (D) Richness and Shannon Index diversity indices between CRC ( $n=73$ ) and control ( $n=92$ ).

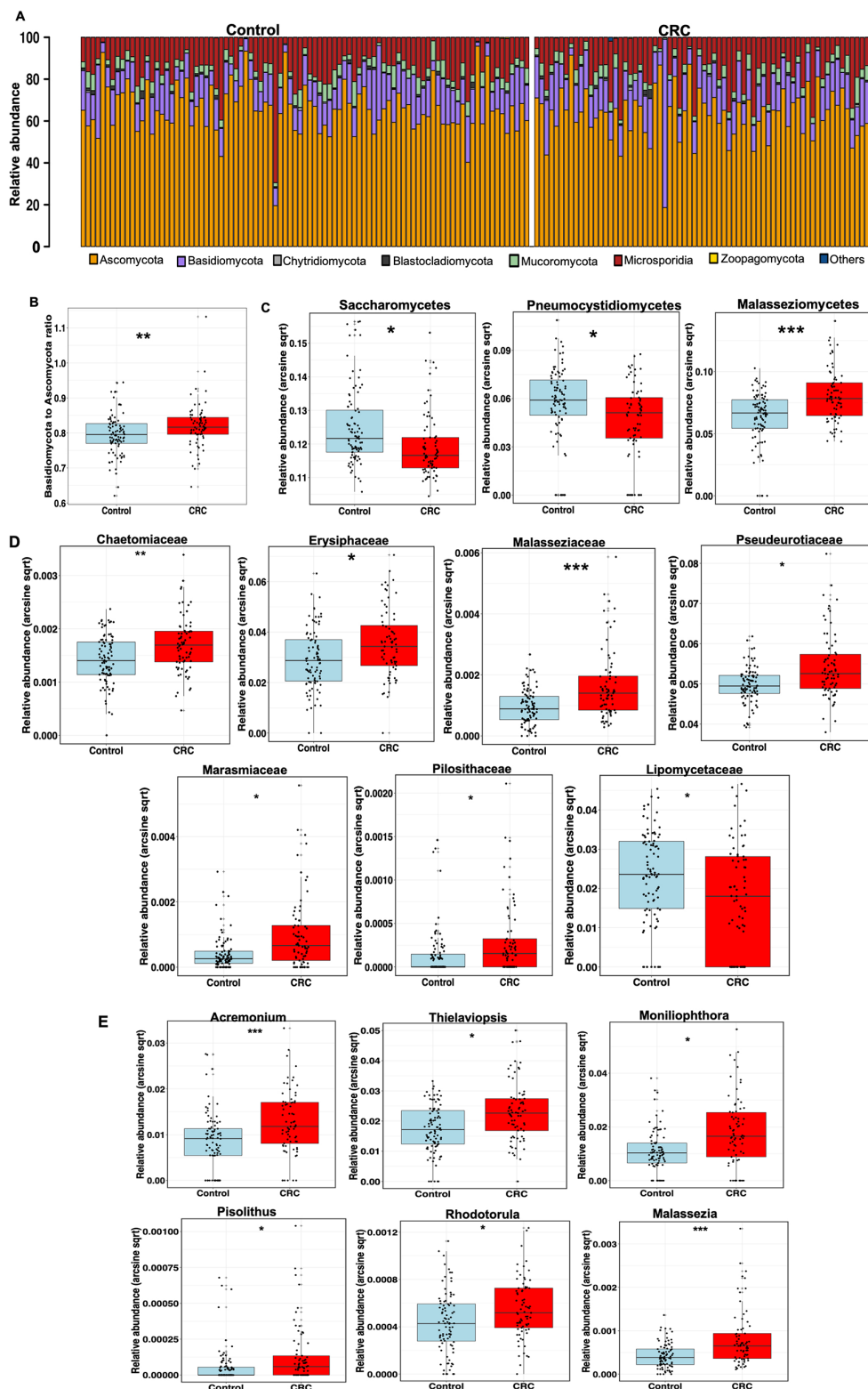
for CRC versus control and early-stage CRC versus control, respectively (figure 4B). However, the fungal markers could only classify adenoma subjects from control with AUCs of 0.63 and 0.6, respectively, in cohorts V1 and V2 (figure 4A,B).

### Perturbed intrakingdom and interkingdom ecological networks in CRC

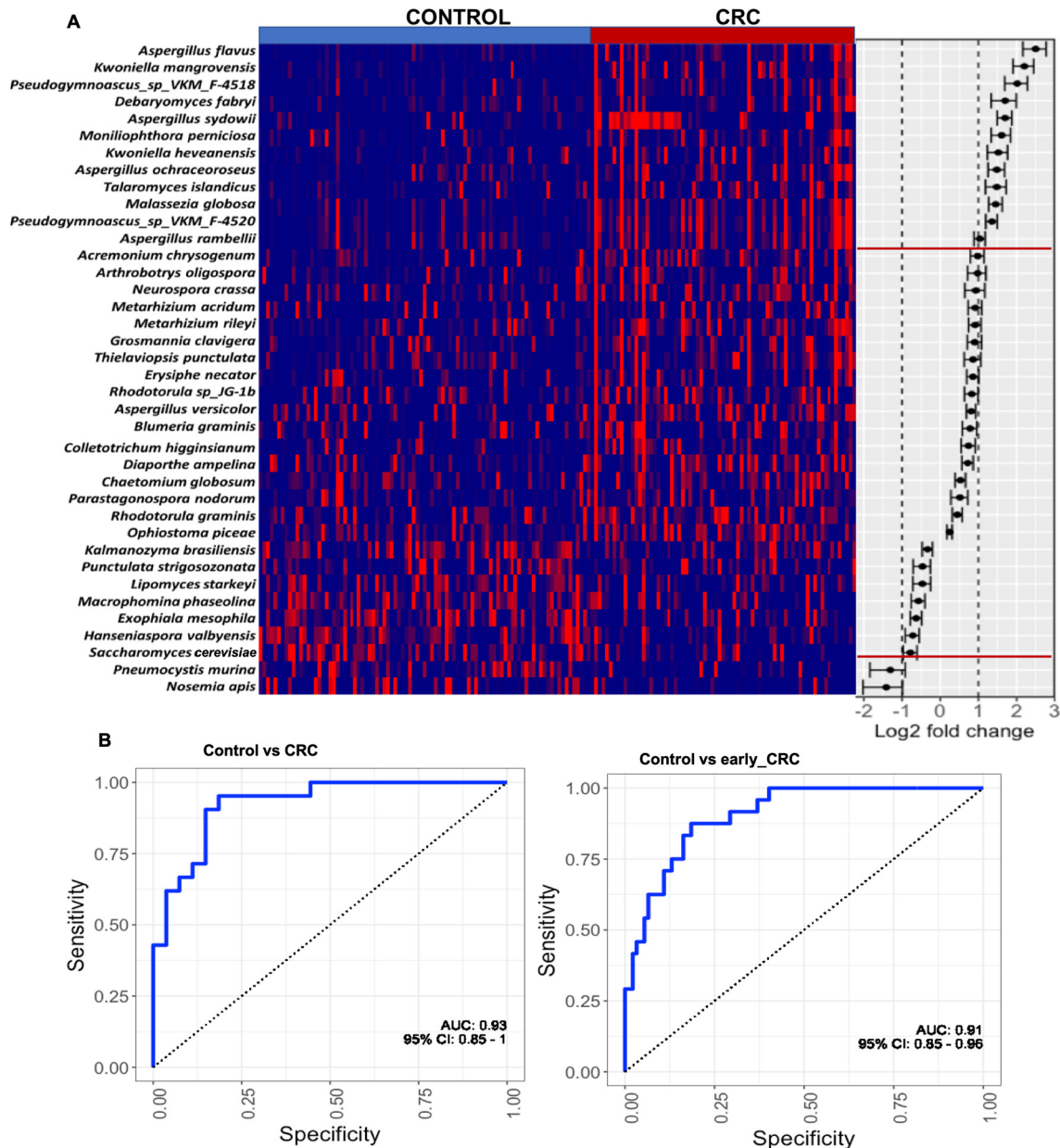
Due to the complex and multifactorial nature of CRC, we investigated the potential alterations of polymicrobial ecological interactions in colorectal tumourigenesis by estimating fungal intrakingdom and bacterial–fungal interkingdom correlations. We observed a considerable difference within the fungal kingdom network, with significantly more positive and stronger interactions in CRC ( $p<0.0001$ , figure 5A, online supplementary tables 6–8). At the phylum level, only Ascomycota and Mucoromycota were significantly correlated in control. New correlations among the Ascomycota, Basidiomycota and Mucoromycota phyla of the fungal kingdom were observed in CRC. Furthermore, markedly different interkingdom ecological networks were observed between CRC and control. Correlations between fungi and bacteria were both fewer and mostly positive in control than in CRC, in which more and stronger co-exclusive relationships were observed ( $p<0.00001$ ,

figure 5A, online supplementary tables 6–10). In control, classes of Ascomycota phylum, including Leotiomycetes, Sordariomycetes, Eurotiomycetes and an unclassified class of Mucoromycota, were found to correlate synergistically with Alphaproteobacteria, Gammaproteobacteria, Deltaproteobacteria and Betaproteobacteria of the bacterial Proteobacteria phylum. Conversely, additional fungal classes of Basidiomycota and Mucoromycota phyla were observed to participate in new interkingdom interactions in CRC, such as the Pucciniomycetes, Exobasidiomycetes and Malasseziomycetes, which formed significant co-exclusive network with major bacterial classes including Alphaproteobacteria, Gammaproteobacteria, Deltaproteobacteria and Cytophagia (figure 5A). Moreover, interkingdom co-occurrence interactions involving fungal classes Leotiomycetes, Sordariomycetes and Eurotiomycetes in control were reversed to become co-exclusive interactions in CRC (figure 5A), indicating that gut microbiota homeostasis may involve a balanced and synergistic interkingdom relationship in healthy subjects which becomes disrupted in CRC. The enhanced intrafungal co-occurrence and bacteria–fungi co-exclusive correlation networks in CRC compared with control were further validated in cohort V1 ( $p=0.004414$  and  $p<0.00001$ , respectively; online supplementary tables 11–15) and V2 ( $p=0.00006$  and  $p<0.00001$ , respectively;





**Figure 2** Taxonomic distribution of enteric fungi in colorectal cancer (CRC). (A) Relative abundance of dominant enteric fungi phyla in control,  $n=92$  and CRC,  $n=73$ . The dominant phyla are Ascomycota, Basidiomycota and Mucoromycota in both groups. (B) Basidiomycota:Ascomycota relative abundance ratio, an index of fungal dysbiosis, is higher in CRC,  $n=73$  than control,  $n=92$  ( $*p<0.001$ , Mann-Whitney U test). (C) Differentially abundant fungal classes between CRC ( $n=73$ ) and control ( $n=92$ ). Differences in abundance were detected using multivariate statistical model and corrected for confounders including age, Body Mass Index (BMI) and gender ( $*Q<0.20$ ,  $**Q<0.10$ ,  $***Q<0.05$ , arcsin-sqrt transformation of relative abundance are derivatives of MaAsLin). (D) Differentially abundant fungal families between CRC,  $n=73$  and control,  $n=92$ . Differences in abundance were detected using multivariate statistical model and corrected for confounders including age, BMI and gender ( $*Q<0.20$ ,  $**Q<0.10$ ,  $***Q<0.05$ , arcsin-sqrt transformation of relative abundance are derivatives of MaAsLin). (E) Differentially abundant fungal genera between CRC,  $n=73$  and control,  $n=92$ . Differences in abundance were detected using multivariate statistical model and corrected for confounders including age, BMI and gender ( $*Q<0.20$ ,  $**Q<0.10$ ,  $***Q<0.05$ , arcsin-sqrt transformation of relative abundance are derivatives of MaAsLin).



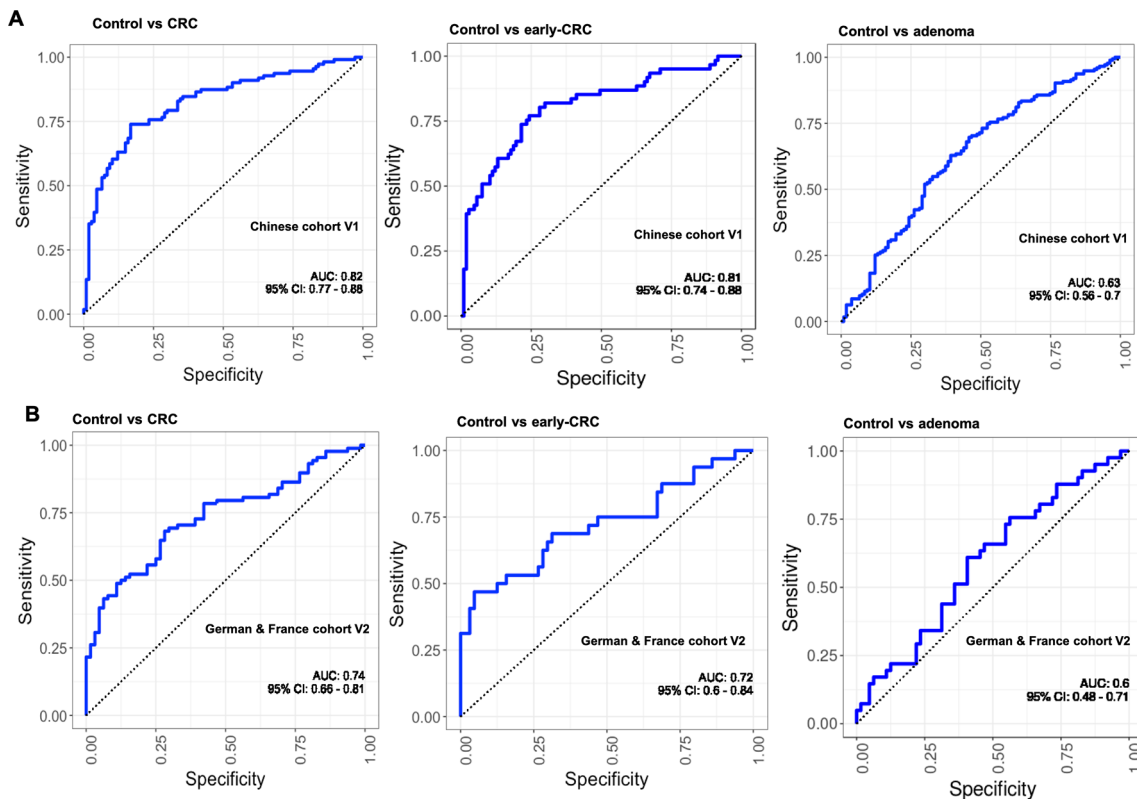
**Figure 3** Fungal species as CRC diagnostic markers. (A) Heatmap showing the relative fold change of fungal species in colorectal cancer (CRC), n=73 against mean normalised relative abundance of control group, n=92 in discovery cohort. Species with absolute log2 fold change greater than 1 were selected as CRC fungal metagenomics markers. (B) Diagnostic performance of fungal markers for classification of CRC from control by logistic regression, corrected for confounders including age, Body Mass Index (BMI) and gender in discovery cohort. The markers achieved an area under receiver-operating characteristic curve (AUC) of 0.93 for the classification of control, n=92 from CRC, n=73, and AUC of 0.91 for the classification of control, n=92 from early CRC, n=20. .

online supplementary tables 16–20). Taken together, these results suggest a complex relationship between fungi and bacteria in the gut, in which their antagonistic associations may contribute to colorectal tumorigenesis.

### Ecological interactions among differentially abundant bacterial and fungal families

To understand the potential interplay among differentially abundant bacteria and fungi in CRC and control, we performed additional bacterial differential abundance analysis (online supplementary table 21) and estimated the ecological network at the family taxonomy level. In control, co-occurrence interactions were observed between fungal Chaetomiaceae and bacterial

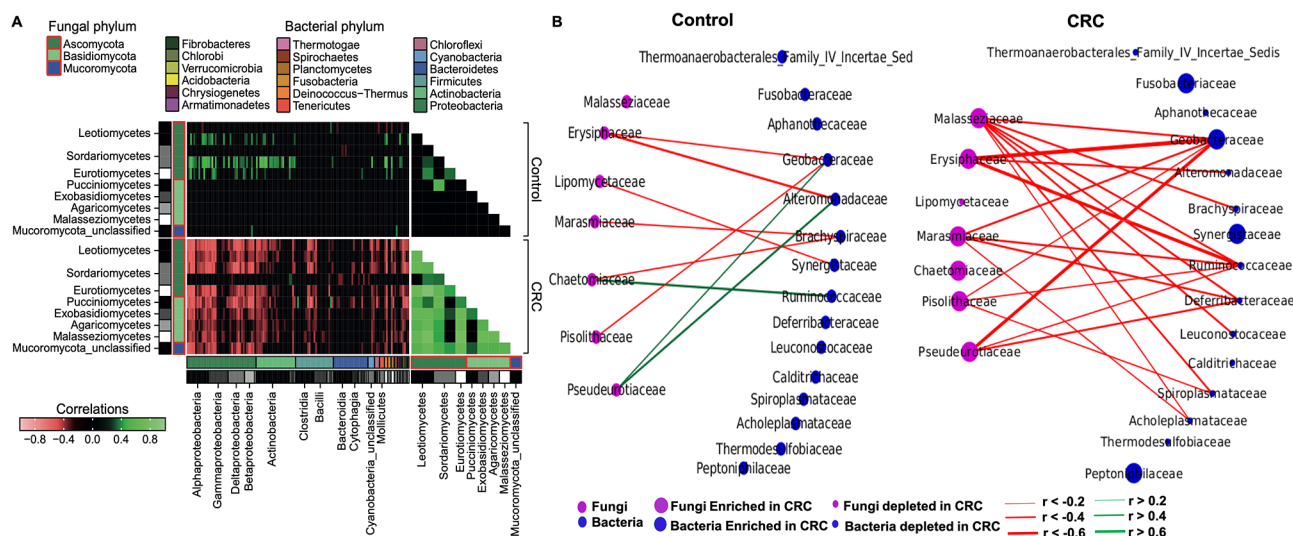
Ruminococcaceae, and between fungal Pseudeurotiaceae and bacterial Alteromonadaceae and Geobacteraceae (figure 5B), which reflect the contribution of synergistic interkingdom interactions to colon microbiota homeostasis. However, in CRC, with the enrichment of bacterial Geobacteraceae, Synergistaceae, Peptoniphilaceae and Fusobacteriaceae, the positive correlations between CRC-enriched fungal Chaetomiaceae and CRC-depleted bacterial Ruminococcaceae, and between CRC-enriched fungal Pseudeurotiaceae and CRC-depleted bacterial Alteromonadaceae no longer existed while that between fungal Pseudeurotiaceae and bacterial Geobacteriaceae was reversed to negative, despite both families being CRC-enriched. Moreover, only Geobacteraceae was involved in significant co-exclusive relationship with



**Figure 4** Validation of faecal fungal diagnostic markers of colorectal cancer (CRC) in independent cohorts. (A) The markers achieved an area under receiver-operating characteristic curve (AUC) of 0.82 for the classification of control, n=112 from CRC, n=111; AUC of 0.81 for the classification of control, n=111 from early\_CRC, n=61; and AUC of 0.63 for the classification of control, n=111 from adenoma, n=197 in Chinese cohort V1. (B) The markers achieved an AUC of 0.74 for the classification of control, n=66 from CRC, n=90; AUC of 0.71 for the classification of control, n=66 from early\_CRC, n=32; and AUC of 0.6 for the classification of control, n=66 from adenoma, n=42 in German & France cohort V2.

fungi in CRC among the CRC-enriched bacterial families, while other observed associations involved CRC-depleted bacterial families (figure 5B). Notably, our network analysis revealed that, in CRC, depleted bacterial families formed significant co-exclusive interactions with enriched fungal families, while enriched

bacterial families such as Fusobacteriaceae, Synergistaceae and Peptoniphilaceae did not correlate significantly with enriched or depleted fungal families. This suggests that dysbiosis in bacterial kingdom may enhance colonisation by certain fungi families in CRC as previously described.<sup>22</sup> On the other hand, within the



**Figure 5** Perturbed intrakingdom and interkingdom ecological networks in colorectal cancer (CRC). Correlations were estimated by SPaCC algorithm. Absolute correlations below 0.3 were masked to show all signals that are exclusively significant. (A) Stronger co-occurrence ecological network in CRC than control (p<0.0001, Mann-Whitney U test). New correlations among Ascomycota, Basidiomycota and Mucoromycota phyla of the fungal kingdom emerged in CRC. (B) Ecological network among differentially abundant fungi and bacterial families in CRC compared with control.

fungal kingdom, extensive synergistic network formed among the CRC-enriched families, including Erysiphaceae, Malasseziaceae, Marasmiaceae, Pseudeurotiaceae, Chaetomiaceae and Pisolithaceae and several other undifferentiated fungal families in CRC compared with control ( $p < 0.0001$ ). In particular, Aspergillaceae was observed to have relatively strong positive ecological interactions with all CRC-enriched fungal families, a phenomenon that was non-existent in control (online supplementary figure 3). Thus, our analysis shows that co-occurrence interkingdom ecological relationships may be important for gut homeostasis in a healthy gut, while dysbiosis might provide a suitable environment for enhanced positive ecological interactions within the fungal kingdom, which, in turn, can play a role in colorectal carcinogenesis.

## DISCUSSION

In this study, faecal shotgun metagenomic sequences were analysed to investigate gut fungal compositional and ecological alterations in association with CRC. To avoid overestimation of fungal role in CRC, which is possible with targeted marker gene sequencing,<sup>23</sup> we assigned taxonomies simultaneously to bacterial and fungal DNA reads of metagenomic data<sup>24</sup> using exact k-mer alignment metagenomic sequence classifier. This enabled the elucidation of differences in the interkingdom interactions between patients with CRC and control subjects. For the first time, we show that the enteric mycobial profile in CRC cases is distinct from healthy controls. In accordance with reported human gut mycobiome profiles, both CRC and control subjects were dominated by Ascomycota and Basidiomycota phyla<sup>25–27</sup> with CRC being associated with a higher Basidiomycota:Ascomycota ratio, an index that defines fungal dysbiosis in an ecosystem.<sup>28 29</sup> Our analysis clearly underscores the disruption of enteric fungal community in CRC.

We identified CRC-specific shifts in fungal composition reflected by the enrichment of six genera, including *Rhodotorula* and *Malassezia* of the Basidiomycota phylum and *Acremonium* of the Ascomycota phylum. The pathogenic potentials of these genera in diseases have been reported. For instance, isolates of the *Acremonium* genus were recovered from human fluids, such as blood, sputum and tracheal aspirates, where they were recognised as opportunistic pathogens.<sup>30</sup> Also, *Rhodotorula* is an emerging pathogen capable of colonising, infecting and disrupting the essential microflora of human digestive system.<sup>31</sup> Moreover, *Malassezia*, a fungal genus commonly found on the human skin with the capability of gut colonisation,<sup>12 27 32</sup> encodes some secreted enzymes similar to the human pathogen *Candida albicans*.<sup>33</sup> The contribution of *Malassezia*-synthesised aryl hydrocarbon receptor ligands to basal cell carcinoma through UV radiation-induced carcinogenesis has also been suggested.<sup>34</sup> Notably, four species of *Aspergillus*, namely *A. flavus*, *A. rambellii*, *A. sydowii* and *A. ochraceoroseus*, were found to be enriched in CRC with the highest fold change of about four for *A. flavus*, a major producer of a highly toxic carcinogen aflatoxin.<sup>35</sup> While *Aspergillus* has been detected in the human gut mycobiome in different studies,<sup>7 36 37</sup> we particularly observed that the species whose abundance correlates with CRC in this study are known producers of carcinogens.<sup>38 39</sup> The higher abundance ratios of these genera and species in CRC therefore suggest their possible contribution towards CRC pathogenesis and warrants further investigations.

In this study, the fungal class Saccharomycetes and, at the species level, *Lipomyces starkeyi* and *S. cerevisiae*, were found to be depleted in CRC. *S. cerevisiae* is a major component of

the human gut microbiota,<sup>32 40 41</sup> which had been shown to reduce adherent-invasive *E. coli*-induced ileal colitis in a mouse model.<sup>42</sup> *S. cerevisiae* was also found to exhibit regulatory and anti-inflammatory effects on the host by inducing interleukin-10 production.<sup>8 43</sup> The depletion of *S. cerevisiae* in CRC as identified in this study highlights the potential beneficial role of *S. cerevisiae* in the gut. This phenomenon can be adapted as a therapeutic approach for CRC prevention or treatment. Interestingly, we demonstrated that the faecal fungi species could discriminate healthy subjects from patients with CRC with validation in independent and ethnically different cohorts, suggesting that fungal markers might help to improve CRC detection.

Commensal fungi and bacteria coexist in the gut.<sup>44</sup> As such, their mode of interactions in the colon may be altered in disease state and could reflect their potential roles in CRC. Our analysis shows that there exists complex ecological network between fungi and bacteria in healthy gut, which is substantially disrupted in CRC. We show that both intrafungal and interkingdom bacteria–fungi correlation networks were altered in CRC. Intriguingly, while correlation networks within fungi became more positive in CRC, those between fungi and bacteria were largely negative. We also observed the disappearance of some co-occurring bacteria–fungi correlations in CRC. The CRC-associated co-occurring intrafungal network suggests that the roles of fungi in CRC may be interdependent, while the observed co-exclusive bacteria–fungi interkingdom ecological network suggests that antagonistic interkingdom interactions may contribute to colorectal tumourigenesis and that altered bacterial composition in CRC may provide fungi with a favourable condition for intrafungi interaction, which may mediate their effect on CRC development.

The gut mycobiota has been shown to be significantly influenced by food colonisation, with same fungal species detected in consumed food and faecal samples of subjects.<sup>26 45</sup> As such, diet is an important factor to be considered in associating specific fungi with disease and may affect their universal application as diagnostic markers. Although we did not take diet into account in this study, the demonstrated diagnostic potential of fungal markers in two ethnicities with different dietary habits indicate the potential importance of our identified markers in CRC.

In conclusion, our study elucidates the disruption of gut mycobiome homeostasis in colorectal carcinogenesis and demonstrates the potential usefulness of faecal fungal markers in detection of CRC. Moreover, we unravel CRC-specific intrakingdom and interkingdom ecological network alterations, indicating that synergistic intrafungal and antagonistic bacteria–fungi interactions might contribute to CRC. Further investigations to identify the functional consequences of the disrupted fungal composition and ecology are warranted for deeper understanding of the roles of fungi in colon tumourigenesis.

**Contributors** OOC analysed data and drafted the manuscript. GN and RZD organised and analysed data. WKKW, SHW, SCN, FKLC and JJYS commented on the study and revised the manuscript. JY designed, supervised the study and wrote the paper.

**Funding** This project was supported by Science and Technology Program Grant Shenzhen (JCYJ20170413161534162), HMRF Hong Kong (17160862), grant from Faculty of Medicine CUHK on Microbiota Research, RGC-GRF Hong Kong (14106415, 14111216, 14163817), Vice-Chancellor's Discretionary Fund CUHK and CUHK direct grant, Shenzhen Virtual University Park Support Scheme to CUHK Shenzhen Research Institute.

**Competing interests** None declared.

**Patient consent** Not required.

**Provenance and peer review** Not commissioned; externally peer reviewed.



**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

## REFERENCES

- Ferlay J, Soerjomataram I, Dikshit R, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer* 2015;136:E359–86.
- Gao R, Gao Z, Huang L, et al. Gut microbiota and colorectal cancer. *Eur J Clin Microbiol Infect Dis* 2017;36:757–69.
- Parrify LW, Walters WA, Knight R. Microbial eukaryotes in the human microbiome: ecology, evolution, and future directions. *Front Microbiol* 2011;2:153.
- Underhill DM, Iliev ID. The mycobiota: interactions between commensal fungi and the host immune system. *Nat Rev Immunol* 2014;14:405–16.
- Nakatsu G, Zhou H, Wu WKK, et al. Alterations in enteric virome are associated with colorectal cancer and survival outcomes. *Gastroenterology* 2018;155:529–41.
- Luan C, Xie L, Yang X, et al. Dysbiosis of fungal microbiota in the intestinal mucosa of patients with colorectal adenomas. *Sci Rep* 2015;5:7980.
- Li Q, Wang C, Tang C, et al. Dysbiosis of gut fungal microbiota is associated with mucosal inflammation in Crohn's disease. *J Clin Gastroenterol* 2014;48:513–23.
- Bajaj JS, Liu EJ, Kheradman R, et al. Fungal dysbiosis in cirrhosis. *Gut* 2018;67.
- Botschuijver S, Roeselers G, Levin E, et al. Intestinal fungal dysbiosis is associated with visceral hypersensitivity in patients with irritable bowel syndrome and rats. *Gastroenterology* 2017;153:1026–39.
- Iliev ID, Leonardi I. Fungal dysbiosis: immunity and interactions at mucosal barriers. *Nat Rev Immunol* 2017;17:635–46.
- Paterson MJ, Oh S, Underhill DM. Host–microbe interactions: commensal fungi in the gut. *Curr Opin Microbiol* 2017;40:131–7.
- Sokol H, Leducq V, Aschard H, et al. Fungal microbiota dysbiosis in IBD. *Gut* 2017;66:1039–48.
- Jiang TT, Shao TY, Ang WXG, et al. Commensal fungi recapitulate the protective benefits of intestinal bacteria. *Cell Host Microbe* 2017;22:809–16.
- Yu J, Feng Q, Wong SH, et al. Metagenomic analysis of faecal microbiome as a tool towards targeted non-invasive biomarkers for colorectal cancer. *Gut* 2017;66:70–8.
- Zeller G, Tap J, Voigt AY, et al. Potential of fecal microbiota for early-stage detection of colorectal cancer. *Mol Syst Biol* 2014;10:766.
- Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 2014;30:2114–20.
- Rosenbloom KR, Armstrong J, Barber GP, et al. The UCSC Genome Browser database: 2015 update. *Nucleic Acids Res* 2015;43:D670–81.
- Wood DE, Salzberg SL. Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biol* 2014;15:R46.
- Paulson JN, Stine OC, Bravo HC, et al. Differential abundance analysis for microbial marker-gene surveys. *Nat Methods* 2013;10:1200–2.
- Standaert-Vitse A, Sendid B, Joossens M, et al. *Candida albicans* colonization and ASCA in familial Crohn's disease. *Am J Gastroenterol* 2009;104:1745–53.
- Friedman J, Alm EJ. Inferring correlation networks from genomic survey data. *PLoS Comput Biol* 2012;8:e1002687.
- Sovran B, Planchais J, Jegou S, et al. Enterobacteriaceae are essential for the modulation of colitis severity by fungi. *Microbiome* 2018;6:152.
- Tedersoo L, Bahram M, Pölme S, et al. Fungal biogeography. Global diversity and geography of soil fungi. *Science* 2014;346:1256688.
- Oh J, Byrd AL, Deming C, et al. Biogeography and individuality shape function in the human skin metagenome. *Nature* 2014;514:59–64.
- Chehoud C, Albenberg LG, Judge C, et al. Fungal signature in the gut microbiota of pediatric patients with inflammatory bowel disease. *Inflamm Bowel Dis* 2015;21:1948–56.
- Hoffmann C, Dollive S, Grunberg S, et al. Archaea and fungi of the human gut microbiome: correlations with diet and bacterial residents. *PLoS One* 2013;8:e66019.
- Richard ML, Lamas B, Liguori G, et al. Gut fungal microbiota: the Yin and Yang of inflammatory bowel disease. *Inflamm Bowel Dis* 2015;21:656–65.
- Kuramae EE, Hillekens RH, de Hollander M, et al. Structural and functional variation in soil fungal communities associated with litter bags containing maize leaf. *FEMS Microbiol Ecol* 2013;84:519–31.
- Malik AA, Chowdhury S, Schlager V, et al. Soil fungal:bacterial ratios are linked to altered carbon cycling. *Front Microbiol* 2016;7:1247.
- Perdomo H, Sutton DA, Garcia D, et al. Spectrum of clinically relevant *Acremonium* species in the United States. *J Clin Microbiol* 2011;49:243–56.
- Wirth F, Goldani LZ. Epidemiology of *Rhodotorula*: an emerging pathogen. *Interdiscip Perspect Infect Dis* 2012;2012:1–7.
- Suhr MJ, Banjara N, Hallen-Adams HE. Sequence-based methods for detecting and evaluating the human gut mycobiome. *Lett Appl Microbiol* 2016;62:209–15.
- Buentke E, Heffler LC, Wilson JL, et al. Natural killer and dendritic cell contact in lesional atopic dermatitis skin--*Malassezia*-influenced cell interaction. *J Invest Dermatol* 2002;119:850–7.
- Gaitanis G, Magiatis P, Hantschke M, et al. The *Malassezia* genus in skin and systemic diseases. *Clin Microbiol Rev* 2012;25:106–41.
- Arrieta MC, Arévalo A, Stiemsma L, et al. Associations between infant fungal and bacterial dysbiosis and childhood atopic wheeze in a nonindustrialized setting. *J Allergy Clin Immunol* 2018;142.
- Gouba N, Raoult D, Drancourt M. Eukaryote culturomics of the gut reveals new species. *PLoS One* 2014;9:e106994.
- Ukhanova M, Wang X, Baer DJ, et al. Effects of almond and pistachio consumption on gut microbiota composition in a randomised cross-over human feeding study. *Br J Nutr* 2014;111:2146–52.
- Cary JW, Ehrlich KC, Beltz SB, et al. Characterization of the *Aspergillus ochraceoroseus* aflatoxin/sterigmatocystin biosynthetic gene cluster. *Mycologia* 2009;101:352–62.
- Hedayati MT, Pasqualotto AC, Warn PA, et al. *Aspergillus flavus*: human pathogen, allergen and mycotoxin producer. *Microbiology* 2007;153:1677–92.
- Scanlan PD, Marchesi JR. Micro-eukaryotic diversity of the human distal gut microbiota: qualitative assessment using culture-dependent and -independent analysis of faeces. *ISME J* 2008;2:1183–93.
- Angebault C, Djossou F, Abélanet S, et al. *Candida albicans* is not always the preferential yeast colonizing humans: a study in Wayampi Amerindians. *J Infect Dis* 2013;208:1705–16.
- Sivignon A, de Vallée A, Barnich N, et al. *Saccharomyces cerevisiae* CNCM I-3856 prevents colitis induced by AIEC bacteria in the transgenic mouse model mimicking Crohn's disease. *Inflamm Bowel Dis* 2015;21:276–86.
- Jawahara S, Habib K, Maggiotto F, et al. Modulation of intestinal inflammation by yeasts and cell wall extracts: strain dependence and unexpected anti-inflammatory role of glucan fractions. *PLoS One* 2012;7:e40648.
- Iliev ID, Funari VA, Taylor KD, et al. Interactions between commensal fungi and the C-type lectin receptor Dectin-1 influence colitis. *Science* 2012;336:1314–7.
- David LA, Maurice CF, Carmody RN, et al. Diet rapidly and reproducibly alters the human gut microbiome. *Nature* 2014;505:559–63.