AP&T Alimentary Pharmacology & Therapeutics WILEY

Novel microbiome signatures for non-invasive diagnosis of adenoma recurrence after colonoscopic polypectomy

Jessie Qiaoyi Liang[1,2,3,4,5] 🆔 | Yao Zeng[1,3,5] | Grace Kwok[2,4] | Chun Pan Cheung[2,3,4,5] | Bing Yee Suen[6] | Jessica Y. L. Ching[2,4,5] | Ka Fai To[7] | Jun Yu[2,3,4] | Francis K. L. Chan[2,3,4,5] 🆔 | Siew Chien Ng[2,3,4,5] 🆔

[1]Department of Microbiology, Faculty of Medicine, The Chinese University of Hong Kong, Hong Kong, China

[2]Department of Medicine and Therapeutics, Faculty of Medicine, The Chinese University of Hong Kong, Hong Kong, China

[3]Centre for Gut Microbiota Research, Faculty of Medicine, The Chinese University of Hong Kong, Hong Kong, China

[4]Institute of Digestive Disease, State Key Laboratory for Digestive Disease, Li Ka Shing Institute of Health Science, CUHK Shenzhen Research Institute, The Chinese University of Hong Kong, Hong Kong, China

[5]Microbiota I-Centre (MagIC) Limited, The Chinese University of Hong Kong, Hong Kong, China

[6]Department of Surgery, The Chinese University of Hong Kong, Hong Kong, China

[7]Department of Anatomy Chemical Pathology, The Chinese University of Hong Kong, Hong Kong, China

Correspondence
Francis K. L. Chan, Siew C. Ng and Jun Yu, Centre for Gut Microbiota Research, Faculty of Medicine, The Chinese University of Hong Kong. Shatin, NT, Hong Kong.
Email: fklchan@cuhk.edu.hk; siewchienng@cuhk.edu.hk; junyu@cuhk.edu.hk

Summary

Background: We previously reported a panel of novel faecal microbiome gene markers for diagnosis of colorectal adenoma and cancer.

Aim: To evaluate whether these markers are useful in detecting adenoma recurrence after polypectomy.

Methods: Subjects were enrolled in a polyp surveillance study from 2009 to 2019. Stool samples were collected before bowel preparation of index colonoscopy (baseline) and surveillance colonoscopy (follow-up). *Fusobacterium nucleatum* (*Fn*), *Lachnoclostridium* marker (*m3*), *Clostridium hathewayi* (*Ch*) and *Bacteroides clarus* were quantified in baseline and follow-up samples by quantitative polymerase chain reaction (qPCR) to correlate with adenoma recurrence. Recurrence was defined as new adenomas detected >6 months after polypectomy. Faecal immunochemical test (FIT) was performed for comparison.

Results: A total of 161 baseline and 104 follow-up samples were analysed. Among patients with adenoma recurrence, *Fn* and *m3* increased (both $P < 0.05$) while *Ch* were unchanged in follow-up versus baseline samples. Among patients without recurrence, *Fn* and *m3* were unchanged while *Ch* decreased ($P < 0.05$) in follow-up versus baseline samples. Logistic regression that included changes of *m3*, *Fn* and *Ch* at follow-up compared with baseline achieved an area under receiver operating characteristic curve (AUROC) of 0.95 (95%CI: 0.84-0.99) with 90.0% sensitivity and 87.0% specificity for detecting recurrent adenoma. Combination of *m3*, *Fn* and *Ch* at follow-up sample achieved AUROC of 0.74 (95%CI: 0.65-0.82) with 81.3% sensitivity and 55.4% specificity for detecting recurrent adenoma. FIT showed limited sensitivity (8.3%) in detecting recurrent adenomas.

Conclusion: Our combinations of faecal microbiome gene markers can be potentially useful non-invasive tools for detecting adenoma recurrence.

1 | INTRODUCTION

Colorectal cancer (CRC) is one of the most common cancers worldwide.[1] Most CRCs begin as adenomas. Adenomatous polyps can be detected in 20%-40% of patients undergoing screening colonoscopy, and their occurrence is associated with an increased risk of CRC. Although endoscopic removal of colorectal adenomas significantly reduces the risk of CRC, regular surveillance examination is needed as risk of recurrence after polypectomy ranges from 37% to 60%.[2] Recently, non-invasive biomarkers for CRC including stool and plasma tumour DNA have been approved by the U.S. Food and Drug Administration (FDA).[3,4] However, these DNA tests have low diagnostic accuracy for precancerous lesions, especially non-advanced adenomas, because genetic or epigenetic changes in cancerous cells are rarely present in small precancerous lesions.

Altered gut microbiota composition has been implicated in the initiation and progression of adenomas and CRC.[5-9] A direct causative role of gut microbiota for CRC development was demonstrated in germ-free animal models.[10] Specific bacterial pathogens, such as *Fusobacterium nucleatum* (*Fn*)[11-13] and *Peptostreptococcus anaerobius*,[14] have been proposed to promote colorectal tumorigenesis. We previously reported that certain bacterial gene markers in stool were useful as non-invasive tests for adenomas and CRC.[6,15-17] Using probe-based duplex quantitative polymerase chain reaction (qPCR) assays, we have identified a panel of bacterial gene markers of *Fn*, *Lachnoclostridium sp.* (*m3*), *Clostridium/Hungatella hathewayi* (*Ch*) and *Bacteroides clarus* (*Bc*) for the detection of adenoma and CRC.[15,16] *Fn* and *Ch* were found to be enriched in stool of patients with CRC.[15] *m3* is enriched in both adenoma and CRC, whereas *Bc* is enriched in normal subjects.[15,16] Our previous findings indicated that these bacterial gene markers could also detect both advanced and non-advanced adenomas.[16] We hypothesised that these bacterial gene markers would be effective in detecting recurrent adenomas. In this study, we evaluated the diagnostic accuracy of *Fn*, *m3*, *Ch* and *Bc* for adenoma recurrence following colonoscopic polypectomy.

2 | MATERIALS AND METHODS

2.1 | Subject recruitment and stool sample collection

This retrospective study included subjects who were enrolled in a polyp surveillance programme conducted in a tertiary referral centre with a catchment population of 1.2 million between 2009 and 2019 (CREC Ref No: 2010.198). Subjects found to have adenoma on index colonoscopy underwent polypectomy and had regular surveillance colonoscopy according to international guidelines.[2,18,19] Consecutive subjects were eligible in this study if they provided stool samples within 1 month before bowel preparation of index colonoscopy (baseline) and/or surveillance colonoscopy (follow-up). Stool samples were excluded if subjects had taken antibiotics within 3 months before stool collection. Recurrence was defined as at least one adenoma found on surveillance colonoscopy. To minimise the

possibility that recurrent adenomas were due to missed lesions at index colonoscopy, we included only colonoscopy with satisfactory bowel preparation, adenomas identified more than 6 months after the index colonoscopy and colonoscopy performed by experienced colonoscopists with an adenoma detection rate of >30%. All polyps were evaluated by an experienced pathologist (TKF). Advanced adenomas were defined as adenomas with a diameter of ≥1 cm, with a tubulovillous or villous component, or with high-grade dysplasia. Subjects were asked to collect stool samples in standardised containers at home, and delivered to the hospitals in insulating polystyrene foam containers. Stool samples were then stored at −80°C immediately for further analysis. The study was approved by the Joint NTEC-CUHK Clinical Research Ethics Committee (CREC Ref No: 2021.104). All subjects provided written informed consent.

2.2 | Stool DNA extraction and quantification of bacterial gene markers

Stool DNA extraction was performed using Norgen Stool DNA Isolation Kit (Norgen Biotek Corp) manually following the manufacturer's instructions. DNA quality and quantity were determined using gel electrophoresis and a NanoDrop spectrophotometer. Faecal levels of four bacterial gene markers (*Fn*, *m3*, *Bc* and *Ch*) were quantified by qPCR. These markers have previously been shown to be enriched in CRC alone (*Fn* and *Ch*), both adenoma and CRC (*m3*), and healthy subjects (*Bc*). Primer and probe sequences targeting the markers and 16 s rDNA internal control have been verified for target specificity in our previous studies.[15,16] Each probe carried a 5′ reporter dye FAM (6-carboxy fluorescein) or VIC (4,7,2′-trichloro-7′-phenyl-6-carboxyfluorescein) and a 3′ quencher dye TAMRA (6-carboxytetramethyl-rhodamine). Primers and hydrolysis probes were synthesised by Invitrogen. qPCR amplifications were performed on an ABI QuantStudio sequence detection system as previously described, with thermal cycler parameters of 95°C 10 minutes and (95°C 15 seconds, 60°C 1 minute) × 45 cycles.[15,16] Positive controls of the markers and a negative control (H_2O as template) were included in every experiment. Measurements were performed in triplicates for each sample. Relative level of each marker was calculated using delta Cq method as compared to internal 16 s rDNA control (Power (2, −[Cq_{target}-$Cq_{control}$])) and shown as log value of "*10e6 + 1." All samples were processed together regardless of baseline/follow-up samples and the status of adenoma recurrence. The technicians involved were blinded to the identity of stool samples. Quantification of microbiome markers using our qPCR platform has been verified to be stable on a batch of 20 randomly selected stool samples between different laboratories in our institution (all $P > 0.1$ by paired t-tests). We included positive control samples in each experiment to monitor if deviation in quantification results occurred. Intra-assay coefficients of variability (CVs) for Cq values were all <0.7%. Intra-assay and inter-assay CVs for each marker were <5% and <9.5%, respectively.

2.3 | Faecal immunochemical test

Quantitative OC-Sensor test was performed on an automatic OCsensor instrument (Eiken Chemical) according to the manufacturer's instructions, using a positive cut-off value equivalent to a concentration of 100 ng of haemoglobin per millilitre (ng Hb/ml).

2.4 | Sample size calculation

In our pilot study of 40 subjects, the mean levels of m3 upon follow-up at 1-5 years were 5.924 (log of relative level of m3/16s rDNA control) and 2.569 in patients with and without recurrent adenomas, respectively. We estimated that a total of 70 subjects would be required to detect a difference of 2.7 with a SD of 4.0 to achieve 80% power at a two-sided significance level of 5%. To adjust for possible confounding effects of covariates, we further increased the total sample size by a variance inflation factor of 1.25 to a total of 88 subjects (44 subjects per group) with follow-up stools.

2.5 | Scoring algorithms and cut-off values

The combined score of four bacterial gene markers (4Bac) using a logistic regression model (4Bac score = $I_1 + \beta_1{}^*Fn + \beta_2{}^*m3 + \beta_3{}^*Bc + \beta_4{}^*Ch$) was determined in our previous study.[16] To discriminate patients with recurrent adenoma from those without recurrence as determined by colonoscopy and histological examination, the combined score of follow-up markers is: $I_2 + \beta_x{}^*m3_{followup} + \beta_y{}^*Ch_{followup} + \beta_z{}^*Fn_{followup}$, whereas the combined score of baseline and follow-up markers is: $I_3 + (\beta_i{}^*Fn_{followup} - \beta_{ii}{}^*Fn_{Baseline}) + (\beta_{iii}{}^*m3_{followup} - \beta_{iv}{}^*m3_{Baseline}) + (\beta_{iv}{}^*Ch_{followup} - \beta_{vi}{}^*Ch_{Baseline})$. I represents the intercept, β represents the regression coefficient and markers represent the corresponding relative levels. Cut-off values were determined by receiver operating characteristic (ROC) analyses that maximised the Youden index (J = Sensitivity + Specificity −1).[20]

2.6 | Statistical analysis

Values were expressed as median (IQR, interquartile range) or mean ± SD where appropriate. The differences in levels of bacterial gene markers were determined by Mann-Whitney U test or paired t-test. Mann-Whitney U test was used to assess trends of changes at follow-up compared with baseline whereby groups at different time points and different status were considered independent of each other. Paired t-test was performed for comparison between paired baseline and follow-up in the same patient. Continuous clinical and pathological variables were compared by t-test or one-way ANOVA. ROC curves were used to evaluate the diagnostic values of bacterial gene markers or models in distinguishing between patients with and without recurrent adenomas. Pairwise comparison of ROC curves was performed using a non-parametric approach.[21] All tests were done by Graphpad Prism 5.0 (Graphpad Software Inc.) or MedCalc Statistical Software V.18.5 (MedCalc Software bvba, Ostend, Belgium; http://www.medcalc.org; 2018). $P < 0.05$ was taken as statistical significance.

3 | RESULTS

3.1 | Clinical characteristics

A total of 222 consecutive eligible subjects who had adenomas at baseline colonoscopy were included in this study: Group I: 118 subjects had stool samples collected before index colonoscopy; Group II: 61 subjects had follow-up stool samples collected before surveillance colonoscopy; and Group III: 43 subjects had both baseline and follow-up stool samples. Stool samples collected from 154 subjects with normal index colonoscopy were included for marker comparison. A total of 161 baseline stool samples were collected before index colonoscopy from Group I and Group III, whereas 104 follow-up stools were collected before surveillance colonoscopy from Group II and Group III.

In all, 48 of the 104 post-polypectomy subjects were found to have adenomas at follow-up colonoscopy, seven of which were advanced adenomas. Detailed clinical characteristics are shown in Tables 1 and 2. In all, 42 of the 104 (40.4%) patients underwent surveillance colonoscopy in less than 3 years. Among them, 7 had endoscopic submucosal resection/endoscopic submucosal dissection for large or laterally spreading adenomas (surveillance interval: 0.8-2.4 years), 3 had sessile serrated adenomas (surveillance interval: 1-1.6 years), 1 had multiple (>10) adenomas (surveillance interval: 1 year), 17 had a family history of first-degree relatives with colorectal neoplasia, 3 had a family history of other cancers (surveillance interval: 1 year) and 11 requested shorter surveillance interval due to anxiety or other miscellaneous reasons (surveillance interval: 1-2.3 years). None of the patients had repeated colonoscopy due to incomplete polyp removal and all patients had clear resection margin at index colonoscopy. There was no difference in the mean surveillance intervals between patients with and without recurrent adenoma in the overall cohort (2.3 ± 1.2 years versus 2.5 ± 1.2 years; $P = 0.26$), or in the subgroup with surveillance interval < 3 years (1.3 ± 0.4 years versus 1.3 ± 0.5 years; $P = 0.57$).

3.2 | Stool bacterial gene markers increased with adenoma recurrence

Compared with normal controls, the levels of Fn, m3 and Ch were significantly higher while that of Bc was significantly lower in patients with adenoma at baseline (Figure 1A). We first compared the levels of four bacterial gene markers Fn, m3, Ch and Bc in baseline stool samples of subjects with adenomas and in follow-up stool samples with and without recurrent adenomas after polypectomy. In group III with paired samples, Fn ($P < 0.05$) and m3 ($P < 0.0001$) were significantly increased in follow-up samples with recurrent adenomas but not in samples without recurrence compared with baseline samples.

Variables	Patients with baseline adenoma (n = 222)	Patients with recurrent adenoma (n = 48)
Mean age ± SD year	60.0 ± 5.3	63.9 ± 4.7
Male gender (%)	139 (62.6)	30 (62.5)
Co-morbidity (%)	123 (55.4)	47 (97.9)
Adenoma[a]		
Non-advanced (%)	55 (24.8)	41 (85.4)
Advanced (%)	167 (75.2)	7 (14.6)
Mean number of adenomas ±SD	2.2 ± 1.8	1.8 ± 1.2
Number of patients with >5 adenomas (%)	16 (7.2)	0 (0)
Adenoma size ±SD mm	10.2 ± 8.5	5.1 ± 4.2
Location (%)		
Proximal	55 (24.8)	15 (31.2)
Distal	96 (43.2)	19 (39.6)
Both	66 (29.7)	14 (29.2)
Unknown	5 (2.3)	0 (0)
Histology (%)		
Tubular	156 (71.3)	44 (91.7)
Villous	66 (29.7)	4 (8.3)

[a]The advanced adenomas were adenomas 1 cm or larger in size, with a tubulovillous or villous component, or with high-grade or severe dysplasia.

Ch showed a non-significant decrease in follow-up samples with no recurrent adenoma ($P = 0.066$) but no change in those with recurrence, whereas *Bc* showed no change in follow-up stool samples regardless of adenoma recurrence status (Figure 1B). These findings were further validated in an enlarged dataset involving all samples from Groups I to III (Figure 1C).

3.3 | Paired stool samples showed a high accuracy in detecting adenoma recurrence

We found a significant increase in the level of *m3* and combined levels of four markers (*Fn, Ch, m3* and *Bc)* in the follow-up samples compared with baseline samples in subjects who developed adenoma recurrence ($P < 0.05$ by matched-pair tests) (Figure 2A). In contrast, there was no significant change in levels of bacterial gene markers in follow-up samples in subjects with no recurrence (Figure 2B). Using a logistic regression model that included changes in the markers at follow-up compared with baseline stools, we found that *m3* alone showed a good performance for detecting adenoma recurrence with an AUROC of 0.843 (95% CI: 0.700-0.936) ($P < 0.0001$), sensitivity of 85.0% and specificity of 87.0%. Using logistic regression to combine all the bacterial gene markers, the combination of *m3, Fn* and *Ch* achieved an AUROC of 0.950 (95%CI: 0.837-0.993) with sensitivity of 90.0% and specificity of 87.0% for detecting recurrent adenoma (Figure 3). FIT showed limited sensitivity (8.3%) in detecting recurrent adenomas, most of which were non-advanced adenomas. Adding FIT to this panel of combined bacterial gene markers did not improve the diagnostic performance.

3.4 | Faecal bacterial gene markers as a standalone test for diagnosing adenoma recurrence

As shown in Figure 1B,C, the levels of *Fn, m3* and *Ch* were significantly increased in follow-up stools with recurrence as compared with those without recurrence ($P < 0.05$), whereas the level of *Bc* showed no difference. Using follow-up stool samples without baseline samples for comparison, combining three markers (*Fn, m3* and *Ch*) could detect recurrent adenoma with an AUROC of 0.741 (95%CI: 0.646-0.822) and sensitivity of 81.3% (Figure 4A). In addition, faecal levels of *Fn, m3* and *Ch* showed no difference in subjects with baseline or recurrent adenomas in the proximal colon versus distal colon (Figure 4B1 and Figure 4B2). Adding FIT to this panel of combined bacterial gene markers did not further improve the diagnostic performance. Correlation analysis showed no significant correlations between age and individual markers or the combined score of *m3, Fn* and *Ch* for recurrent adenoma in the follow-up cohort ($P > 0.4$). On multivariate analysis, the levels of *m3* and *Ch* ($P < 0.01$), and the combined score of *m3, Fn* and *Ch* ($P < 0.0001$) were independent risk factors for adenoma recurrence. As there were significant imbalances in some baseline characteristics including male gender, co-morbidity, number and size of adenomas between subjects with and without recurrent adenomas (Table 2), we further performed logistic regression analysis and found that none of these factors was a significant predictor of adenoma recurrence. *m3ChFn* appeared to be the only significant factor in the logistic regression model for recurrent adenomas. Pairwise comparison of ROC curves showed no significant difference between the model consisting of *m3ChFn* alone and

TABLE 2 Baseline clinical characteristics of patients with and without recurrence

Variables	Patients without recurrent adenoma (n = 56)	Patients with recurrent adenoma (n = 48)	P value
Mean age ± SD year	60.3 ± 6.3	61.6 ± 4.7	0.25
Male gender (%)	23 (41.1)	30 (62.5)	0.03
Co-morbidity (%)	19 (33.9)	30 (62.5)	0.006
Mean number of adenomas ± SD	1.7 ± 1.0	2.3 ± 1.6	0.04
Number of patients with >5 adenomas (%)	0 (0)	3 (6.3)	0.10
Adenoma size ± SD mm	9.7 ± 4.7	13.7 ± 10.3	0.01
Location (%)			
Proximal	11 (19.6)	13 (27.1)	0.61
Distal	29 (51.8)	21 (43.8)	
Both	16 (28.6)	14 (29.2)	
Histology (%)			
Tubular	37 (66.1)	34 (70.8)	0.68
Villous	19 (33.9)	14 (29.2)	

the model combining the four clinical characteristics and *m3ChFn* (P = 0.135; Figure 4C).

4 | DISCUSSION

To the best of our knowledge, this is the first study showing that faecal bacterial gene markers can effectively detect recurrent adenomas after polypectomy. By quantifying changes in faecal levels of our novel bacterial gene markers after index colonoscopy, we found a high accuracy of these markers in detecting adenoma recurrence. The combination of *m3*, *Fn* and *Ch* yielded the best AUROC and showed 90% sensitivity. The addition of FIT did not increase the diagnostic sensitivity for recurrent adenoma. Our findings have demonstrated the potential of using faecal bacterial gene markers in detecting early precancerous lesions.
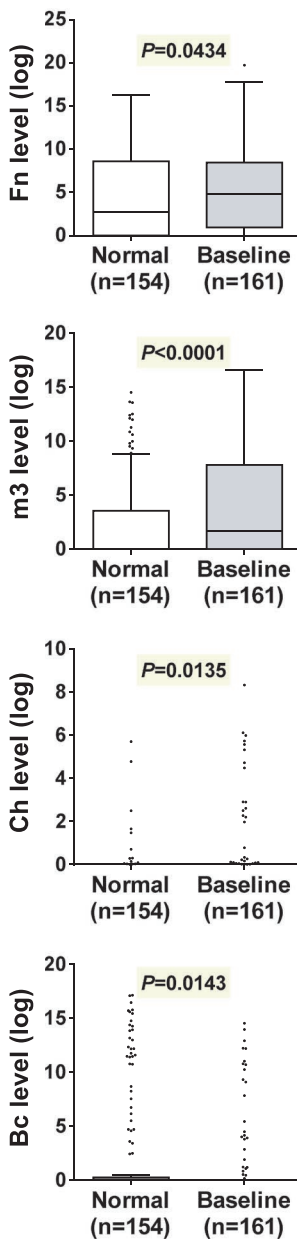
Different bacterial gene markers showed different activities in relation to adenoma recurrence. We found that our panel of bacterial gene markers remained unchanged 1 month after removal of adenoma,[16] which suggested that polypectomy does not lead to immediate change in the gut microbial environment. In the presence of recurrent adenoma, *m3* and *Fn* further increased, whereas *Ch* remained unchanged. In contrast, *m3* and *Fn* remained comparable to baseline levels, whereas *Ch* decreased further in the absence of recurrent adenoma. Therefore, a combination of *m3*, *Fn* and *Ch* is useful for detecting recurrent adenoma. Importantly, faecal levels of *m3*, *Fn* and *Ch* were not different between recurrent adenomas in the proximal and distal colon, indicating that the sensitivity of these markers is not affected by the location of lesions. Combination of these markers not only provides a non-invasive means of detecting recurrent adenoma, but also offers hope to lower the risk of colorectal neoplasms by modulation of gut microbiota in subjects who had high levels of unfavourable bacterial markers.
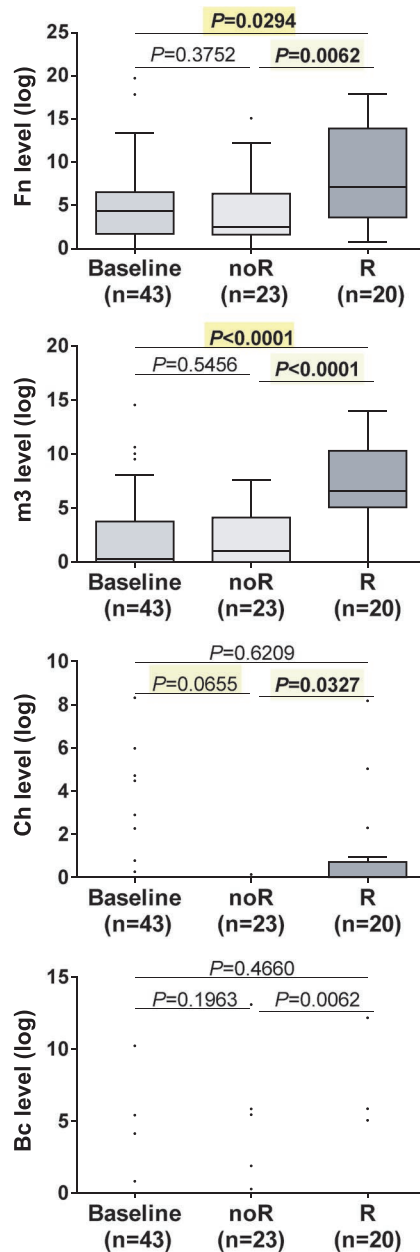
Several studies have implicated microbial dysbiosis in the aetiology of colorectal adenomas.[22] For instance, changes in the gut microbial community composition have been reported in faecal samples and tissues of adenoma.[7,23] Alteration in gut microbiota was reported at 3 months after adenoma resection but there was no significant change in the main phyla including Fusobacteria,[24] suggesting that changes in bacteria associated with adenoma recurrence may occur at a later time point. It has been demonstrated that *Fn* induced inflammation and modulated host immune response to promote tumour development.[11,12] *Ch* has been shown to promote colonic epithelial cell proliferation in mouse models.[25] The *Lachnoclostridium* species carrying *m3* is likely to play an important role in adenoma recurrence but its role in promoting colorectal tumorigenesis and adenoma recurrence remains unknown. We speculate that these unfavourable bacteria may trigger host immune responses to further promote the development of recurrent adenomas. Functional studies are required to evaluate the role of these bacteria in promoting adenoma recurrence.

There is a huge unmet need for non-invasive biomarkers to monitor adenoma recurrence. Current guidelines recommend colonoscopic surveillance at variable intervals after polypectomy, depending on the characteristics of the lesions.[2,19] Colonoscopy can be invasive and uptake is low due to suboptimal compliance. A recent national survey of CRC screening showed that stool-based tests were preferred to colonoscopy.[26] Accurate detection of early or small precancerous lesions is clinically relevant because these early lesions account for over 30% of adenomas found on surveillance colonoscopy. In the current study, our markers could detect adenoma recurrence with reasonable sensitivity. This non-invasive approach may inform the optimal timing of surveillance colonoscopy in the future.
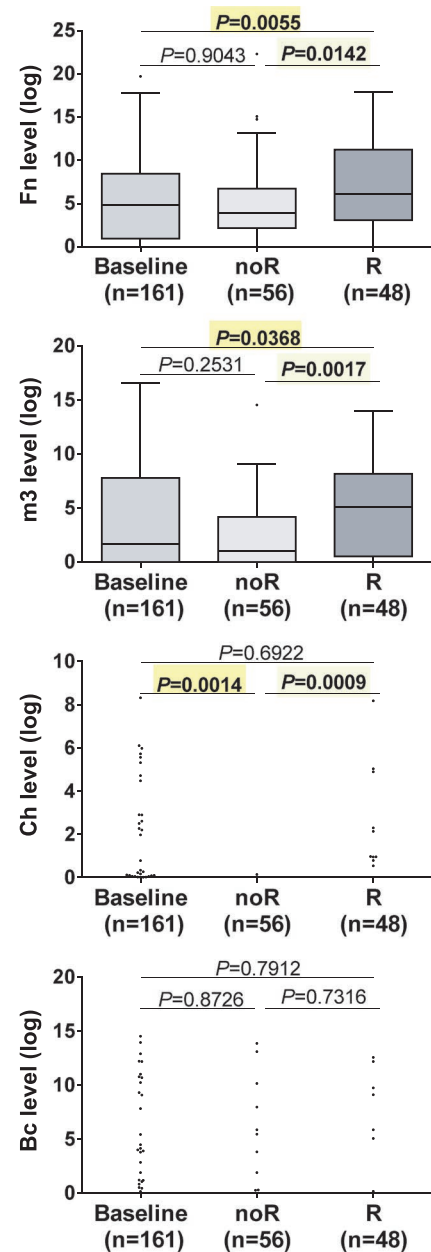
**(A)** *Normal vs Baseline*   **(B)** *Baseline vs Follow-up (Group 3)*   **(C)** *Baseline vs Follow-up (Groups 1-3)*

FIGURE 1 Changes in faecal bacterial markers in post-polypectomy subjects as compared to patients with baseline advanced adenoma. A, Comparison of the levels of four faecal bacterial markers in baseline stool samples collected before index colonoscopy between patients diagnosed with advanced adenoma and subjects confirmed with normal colon. B, Comparison of the levels of four faecal bacterial markers between baseline stool samples collected before index colonoscopic diagnosis of advanced adenoma and follow-up stools before surveillance colonoscopy from subjects who developed recurrent adenoma and those without recurrence in group III subjects with paired baseline and follow-up stools. (C) Comparison of the levels of four faecal bacterial markers between baseline stool samples and follow-up samples from all subjects recruited in this study. *Fn*, *fusobacterium nucleatum*; *m3*, *Lachnoclostridium* marker *m3*; *Ch*, *Clostridium hathewayi*; *Bc*, *Bacteroides clarus*; no-R, no-recurrence; R, recurrence

Our study has a number of strengths. First, we have prospectively collected serial stool samples for up to 10 years. Polyps removed were assessed by an experienced pathologist to eliminate inter-observer variation. Second, we previously showed that our bacterial gene markers were independent risk factors for CRC,[15,16] and these markers were not affected by age or gender in the current study cohort (data not shown). Third, unlike stool DNA or plasma DNA tests that target genetic/epigenetic changes from cancerous cells, our bacterial gene markers can detect CRC, advanced and non-advanced adenomas. With paired baseline and follow-up stools, our markers (*Fn*, *m3* and *Ch*) could diagnose recurrent adenoma with an AUROC of 0.950 and sensitivity of 90%. We hypothesised that if the gut microbiome becomes

(A) _Recurrence group (n=20) (follow-up vs. baseline)_
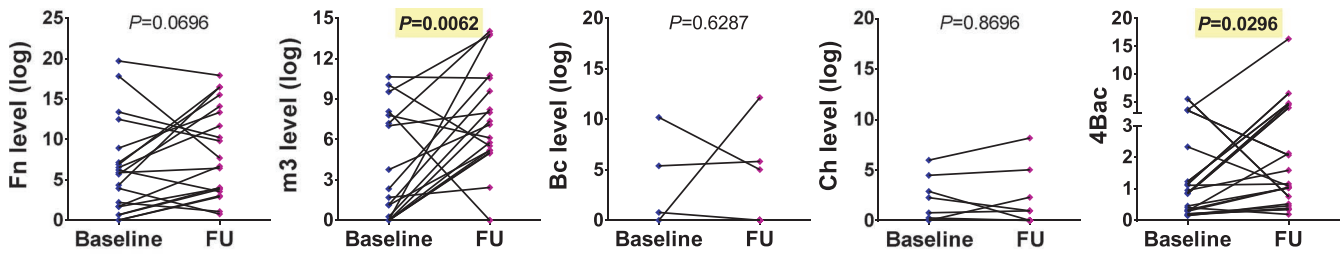
(B) _No-recurrence group (n=23) (follow-up vs. baseline)_
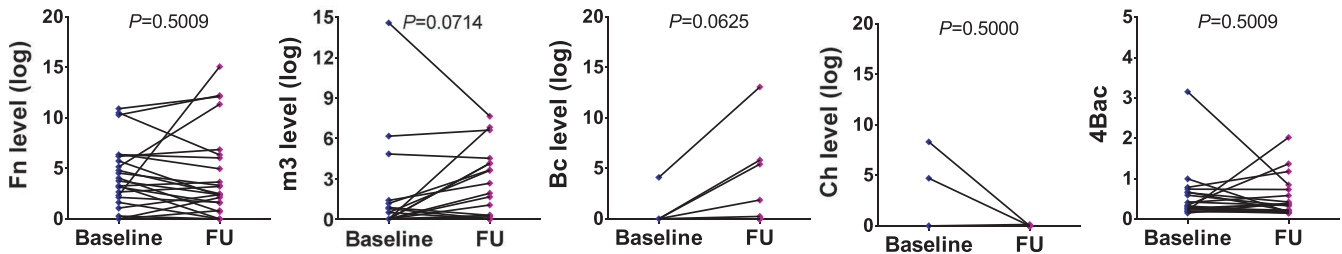
FIGURE 2 Changes in bacterial markers at follow-up (FU) versus baseline for diagnosis of recurrent adenoma. A, The four bacterial markers and their combined score 4Bac showed no significant difference between baseline and FU stools in patients without recurrence. B, Significant increases in _m3_ and _4Bac_ were detected in FU stools as compared to baseline stools in patients with recurrence. _Fn, fusobacterium nucleatum; m3, Lachnoclostridium_ marker _m3; Ch, Clostridium hathewayi; Bc, Bacteroides clarus;_ 4Bac: Combined score of _Fn, m3, Ch_ and _Bc_ previously devised for diagnosis of CRC and adenoma

Logistic regression involving changes of marker levels at follow-up _vs._ baseline ➜

| Variable	AUROC	SE	95% CI	P	Sen. %	Spe. %
△m3	0.843	0.067	0.700 to 0.936	<0.0001	85.0	87.0
△'4Bac'	0.826	0.062	0.680 to 0.924	<0.0001	70.0	82.6
△m3Fn	0.857	0.059	0.716 to 0.945	<0.0001	85.0	78.3
△m3FnCh	0.950	0.029	0.837 to 0.993	<0.0001	90.0	87.0
△m3FnChBc	0.946	0.031	0.831 to 0.992	<0.0001	90.0	87.0
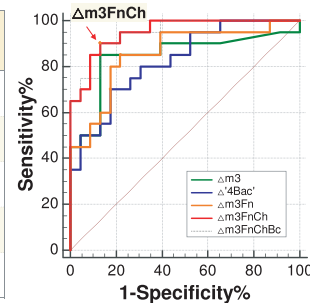
FIGURE 3 Performances of logistic regression models involving changes in FU stools versus baseline stools in discriminating patients with recurrence from those without recurrence. The model combining changes in _Fn, m3_ and _Ch_ showed the best performance. _Fn, fusobacterium nucleatum; m3, Lachnoclostridium_ marker _m3; Ch, Clostridium hathewayi; Bc, Bacteroides clarus;_ 4Bac: Combined score of _Fn, m3, Ch_ and _Bc_

less healthy after polypectomy, such unfavourable gut environment would promote adenoma recurrence, whereas the risk of recurrence will be low if the gut microbiome remains stable or improves.

Our study also had some limitations. First, the sample size is small which may result in overfitting of the diagnostic model without cross-validation and preliminary data from the paired sample group needs to be confirmed in larger prospective studies. Second, diagnostic performance of the biomarkers for stool samples collected at a single time point was lower than diagnostic performance of paired stool samples. Further effort should focus on improving performance of these biomarkers based on a single time point stool test given that it is more user-friendly in clinical practice than testing paired stool samples. Third, we cannot exclude the possibility of missed lesions at index colonoscopy which could be misclassified

as recurrent adenoma. However, measures have been taken to minimise the likelihood of missed lesions. Fourth, the small number of subjects with advanced adenoma at surveillance precluded the use of these markers in predicting recurrent advanced versus non-advanced adenomas. It is logical to believe that higher levels of bacterial gene markers may indicate a higher likelihood of recurrence of advanced adenoma. However, we previously showed that the sensitivity of our bacterial gene markers was similar for both advanced and non-advanced adenoma.[16] Therefore, "more abnormal" levels of markers are likely to be more indicative of recurrence of adenoma but may not differentiate advanced and non-advanced lesions. Future larger studies would be needed to investigate the correlation between marker levels and recurrence of advanced adenoma.
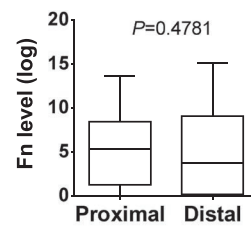
(A) *Logistic regression of m3, Ch and Fn at follow-up*

(B1) *Baseline adenoma* **(B2)** *Recurrent adenoma*
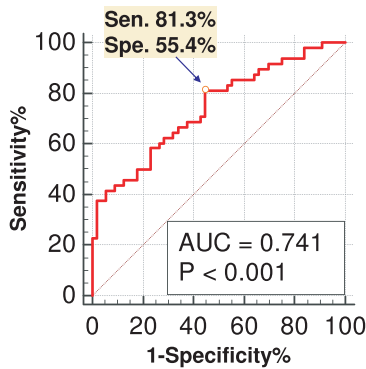
Model	m3ChFn
AUROC	0.741
SE	0.048
95% CI	0.646 to 0.822
P	<0.0001
cutoff (Yoden)	0
Sensitivity %	81.3
Specificity %	55.4
PPV %	60.9
NPV %	77.5
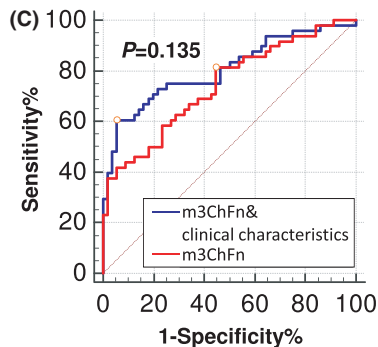
Variable	Coeff.	P value
Number	0.28487	0.1434
Size	0.02239	0.4867
Gender	0.90453	0.0670
Co-morbidity	0.60074	0.2281
m3ChFn	0.34710	0.0014
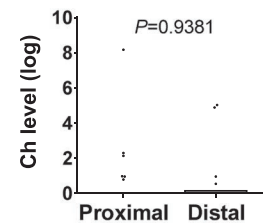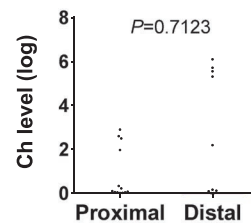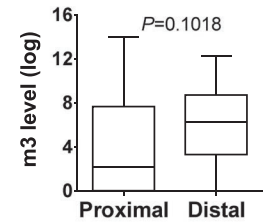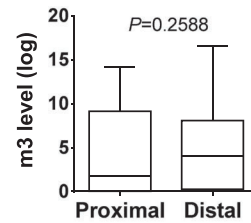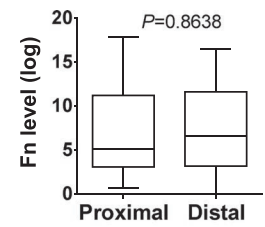
FIGURE 4   A, Performance of the logistic regression model involving follow-up levels of *m3*, *Ch* and *Fn* (*m3ChFn*) in discriminating patients with recurrence from those without recurrence. B, In faecal samples of patients with baseline adenoma (B1) and those with recurrent adenoma (B2), levels of *m3*, *Fn* and *Ch* showed no difference between patients with proximal lesions and those with distal lesions. C, Pairwise comparison of ROC curves showed no significant difference between *m3ChFn* and the logistic regression model adding clinical characteristics (number and size of adenoma, gender, and co-morbidity at baseline). In the logistic regression involving clinical characteristics and *m3ChFn*, only *m3ChFn* was a significant factor in the model. *Fn*, *fusobacterium nucleatum*; *m3*, *Lachnoclostridium* marker m3; *Ch*, *Clostridium hathewayi*; *Bc*, *Bacteroides clarus*; no-R, no-recurrence; R, recurrence; AUROC, area under ROC; CI, confidence interval; PPV, positive predictive value; NPV, negative predictive value

In summary, we have identified a panel of novel bacterial gene markers in stool that is a potentially useful non-invasive tool for detecting adenoma recurrence. Our findings offer hope to inform health authorities and guideline committees to review current strategies for CRC screening.

AUTHORSHIP

Guarantor of the article: S.C.N.

Authors' contributions: The study was conceived by J.Q.L., S.C.N., F.K.C. and J.Y., with methodology developed by J.Q.L., patient sample material and clinical information provided by S.C.N., B.Y.S. and J.Y.C., experimental data collection performed by Y.Z. and C.P.C, and clinical data collected by G.K. and K.F.T. The data were analysed and manuscript drafted by J.Q.L., revised by S.C.N, F.K.C. and J.Y., and all authors reviewed and approved the final version.

ORCID

Jessie Qiaoyi Liang https://orcid.org/0000-0001-8496-3442
Francis K. L. Chan https://orcid.org/0000-0001-7388-2436
Siew Chien Ng https://orcid.org/0000-0002-6850-4454

REFERENCES

1. Allemani C, Matsuda T, Di Carlo V, et al. Global surveillance of trends in cancer survival 2000-14 (CONCORD-3): analysis of individual records for 37 513 025 patients diagnosed with one of 18 cancers from 322 population-based registries in 71 countries. *Lancet.* 2018;391:1023-1075.
2. Lieberman DA, Rex DK, Winawer SJ, Giardiello FM, Johnson DA, Levin TR. Guidelines for colonoscopy surveillance after screening and polypectomy: a consensus update by the US multi-society task force on colorectal cancer. *Gastroenterology.* 2012;143:844-857.
3. Imperiale TF, Ransohoff DF, Itzkowitz SH, et al. Multitarget stool DNA testing for colorectal-cancer screening. *N Engl J Med.* 2014;370:1287-1297.
4. de Vos T, Tetzner R, Model F, et al. Circulating methylated SEPT9 DNA in plasma is a biomarker for colorectal cancer. *Clin Chem.* 2009;55:1337-1346.

5. Irrazabal T, Belcheva A, Girardin SE, Martin A, Philpott DJ. The multifaceted role of the intestinal microbiota in colon cancer. *Mol Cell.* 2014;54:309-320.

6. 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;66:70-78.

7. Nakatsu G, Li X, Zhou H, et al. Gut mucosal microbiome across stages of colorectal carcinogenesis. *Nature Commun.* 2015;6:8727.

8. Dai Z, Coker OO, Nakatsu G, et al. Multi-cohort analysis of colorectal cancer metagenome identified altered bacteria across populations and universal bacterial markers. *Microbiome.* 2018;6:70.

9. Tilg H, Adolph TE, Gerner RR, Moschen AR. The intestinal microbiota in colorectal cancer. *Cancer Cell.* 2018;33:954-964.

10. Wong SH, Zhao L, Zhang X, et al. Gavage of fecal samples from patients with colorectal cancer promotes intestinal carcinogenesis in germ-free and conventional mice. *Gastroenterology.* 2017;153:1621-1633 e6.

11. Kostic AD, Chun E, Robertson L, et al. Fusobacterium nucleatum potentiates intestinal tumorigenesis and modulates the tumor-immune microenvironment. *Cell Host Microbe.* 2013;14:207-215.

12. Rubinstein MR, Wang X, Liu W, Hao Y, Cai G, Han YW. Fusobacterium nucleatum promotes colorectal carcinogenesis by modulating E-cadherin/beta-catenin signaling via its FadA adhesin. *Cell Host Microbe.* 2013;14:195-206.

13. Yu T, Guo F, Yu Y, et al. Fusobacterium nucleatum promotes Chemoresistance to colorectal cancer by modulating autophagy. *Cell.* 2017;170:548-563 e16.

14. Tsoi H, Chu ESH, Zhang X, et al. Peptostreptococcus anaerobius induces intracellular cholesterol biosynthesis in colon cells to induce proliferation and causes dysplasia in mice. *Gastroenterology.* 2017;152:1419-1433 e5.

15. Liang Q, Chiu J, Chen Y, et al. Fecal bacteria act as novel biomarkers for noninvasive diagnosis of colorectal cancer. *Clin Cancer Res.* 2017;23:2061-2070.

16. Liang JQ, Li T, Nakatsu G, et al. A novel faecal Lachnoclostridium marker for the non-invasive diagnosis of colorectal adenoma and cancer. *Gut.* 2020;69:1248-1257.

17. Liang JQ, Wong SH, Szeto CH, et al. Fecal microbial DNA markers serve for screening colorectal neoplasm in asymptomatic subjects. *J Gastroenterol Hepatol.* 2021;36:1035-1043.

18. Rutter MD, East J, Rees CJ, et al. British Society of Gastroenterology/ Association of Coloproctology of Great Britain and Ireland/Public Health England post-polypectomy and post-colorectal cancer resection surveillance guidelines. *Gut.* 2020;69:201-223.

19. Gupta S, Lieberman D, Anderson JC, et al. Recommendations for follow-up after colonoscopy and polypectomy: a consensus update by the US multi-society task force on colorectal cancer. *Gastrointest Endos.* 2020;91:463-485 e5.

20. Youden WJ. Index for rating diagnostic tests. *Cancer.* 1950;3:32-35.

21. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics.* 1988;44:837-845.

22. Dulal S, Keku TO. Gut microbiome and colorectal adenomas. *Cancer J.* 2014;20:225-231.

23. Dadkhah E, Sikaroodi M, Korman L, et al. Gut microbiome identifies risk for colorectal polyps. *BMJ Open Gastroenterol.* 2019;6:e000297.

24. Yu SY, Xie YH, Qiu YW, Chen YX, Fang JY. Moderate alteration to gut microbiota brought by colorectal adenoma resection. *J Gastroenterol Hepatol.* 2019;34:1758-1765.

25. Xia X, Wu WKK, Wong SH, et al. Bacteria pathogens drive host colonic epithelial cell promoter hypermethylation of tumor suppressor genes in colorectal cancer. *Microbiome.* 2020;8:108.

26. Zhu X, Parks PD, Weiser E, et al. National Survey of patient factors associated with colorectal cancer screening preferences. *Cancer Prevent Res.* 2021;14:603-614.