ISSN 1007-9327 (print)
ISSN 2219-2840 (online)

World Journal of *Gastroenterology*

World J Gastroenterol 2022 July 7; 28(25): 2782-3007

Published by Baishideng Publishing Group Inc

World Journal of Gastroenterology

Contents

Weekly Volume 28 Number 25 July 7, 2022

REVIEW

MINIREVIEWS

ORIGINAL ARTICLE

Basic Study

World Journal of Gastroenterology

Weekly Volume 28 Number 25 July 7, 2022

Contents

ABOUT COVER

Editorial Board Member of *World Journal of Gastroenterology*, Hideyuki Chiba, MD, PhD, Director, Department of Gastroenterology, Omori Red Cross Hospital, 4-30-1, Chuo, Ota-Ku, Tokyo 143-8527, Japan. h.chiba04@gmail.com

AIMS AND SCOPE

The primary aim of *World Journal of Gastroenterology* (*WJG*, *World J Gastroenterol*) is to provide scholars and readers from various fields of gastroenterology and hepatology with a platform to publish high-quality basic and clinical research articles and communicate their research findings online. *WJG* mainly publishes articles reporting research results and findings obtained in the field of gastroenterology and hepatology and covering a wide range of topics including gastroenterology, hepatology, gastrointestinal endoscopy, gastrointestinal surgery, gastrointestinal oncology, and pediatric gastroenterology.

INDEXING/ABSTRACTING

The *WJG* is now indexed in Current Contents®/Clinical Medicine, Science Citation Index Expanded (also known as SciSearch®), Journal Citation Reports®, Index Medicus, MEDLINE, PubMed, PubMed Central, and Scopus. The 2021 edition of Journal Citation Report® cites the 2020 impact factor (IF) for *WJG* as 5.742; Journal Citation Indicator: 0.79; IF without journal self cites: 5.590; 5-year IF: 5.044; Ranking: 28 among 92 journals in gastroenterology and hepatology; and Quartile category: Q2. The *WJG*'s CiteScore for 2020 is 6.9 and Scopus CiteScore rank 2020: Gastroenterology is 19/136.

RESPONSIBLE EDITORS FOR THIS ISSUE

Production Editor: *Ying-Yi Yuan*; **Production Department Director:** *Xiang Li*; **Editorial Office Director:** *Jia-Ru Fan*.

NAME OF JOURNAL
World Journal of Gastroenterology

ISSN
ISSN 1007-9327 (print) ISSN 2219-2840 (online)

LAUNCH DATE
October 1, 1995

FREQUENCY
Weekly

EDITORS-IN-CHIEF
Andrzej S Tarnawski

EDITORIAL BOARD MEMBERS
http://www.wjgnet.com/1007-9327/editorialboard.htm

PUBLICATION DATE
July 7, 2022

INSTRUCTIONS TO AUTHORS
https://www.wjgnet.com/bpg/gerinfo/204

GUIDELINES FOR ETHICS DOCUMENTS
https://www.wjgnet.com/bpg/GerInfo/287

GUIDELINES FOR NON-NATIVE SPEAKERS OF ENGLISH
https://www.wjgnet.com/bpg/gerinfo/240

PUBLICATION ETHICS
https://www.wjgnet.com/bpg/GerInfo/288

PUBLICATION MISCONDUCT
https://www.wjgnet.com/bpg/gerinfo/208

ARTICLE PROCESSING CHARGE
https://www.wjgnet.com/bpg/gerinfo/242

STEPS FOR SUBMITTING MANUSCRIPTS
https://www.wjgnet.com/bpg/GerInfo/239

ONLINE SUBMISSION
https://www.f6publishing.com

World Journal of Gastroenterology

Submit a Manuscript: https://www.f6publishing.com

World J Gastroenterol 2022 July 7; 28(25): 2920-2936

DOI: 10.3748/wjg.v28.i25.2920

ISSN 1007-9327 (print) ISSN 2219-2840 (online)

ORIGINAL ARTICLE

Basic Study

Fecal gene detection based on next generation sequencing for colorectal cancer diagnosis

Si-Yu He, Ying-Chun Li, Yong Wang, Hai-Lin Peng, Cheng-Lin Zhou, Chuan-Meng Zhang, Sheng-Lan Chen, Jian-Feng Yin, Mei Lin

Specialty type: Gastroenterology and hepatology

Provenance and peer review: Invited article; Externally peer reviewed.

Peer-review model: Single blind

Peer-review report's scientific quality classification
Grade A (Excellent): A
Grade B (Very good): B
Grade C (Good): C
Grade D (Fair): 0
Grade E (Poor): 0

P-Reviewer: Beenet L, United States; Kołat D, Poland
A-Editor: Sahin TT, Turkey

Received: January 16, 2022
Peer-review started: January 16, 2022
First decision: March 8, 2022
Revised: March 18, 2022
Accepted: May 27, 2022
Article in press: May 27, 2022
Published online: July 7, 2022

Si-Yu He, Hai-Lin Peng, Cheng-Lin Zhou, Mei Lin, Department of Clinical Laboratory, Taizhou People's Hospital (Postgraduate Training Base of Dalian Medical University), Taizhou 225300, Jiangsu Province, China

Si-Yu He, Department of Clinical Laboratory, The First People's Hospital of Tianmen City, Tianmen 431700, Hubei Province, China

Ying-Chun Li, Yong Wang, Department of General Surgery, Taizhou People's Hospital (Postgraduate Training Base of Dalian Medical University), Taizhou 225300, Jiangsu Province, China

Chuan-Meng Zhang, Central Laboratory, Taizhou People's Hospital (Postgraduate training base of Dalian Medical University), Taizhou 225300, Jiangsu Province, China

Sheng-Lan Chen, Department of Laboratory, Taizhou Genewill Medical Laboratory Company Limited, Taizhou 225300, Jiangsu Province, China

Jian-Feng Yin, Department of Laboratory, Jiangsu CoWin Biotech Co., Ltd., Taizhou 225300, Jiangsu Province, China

Corresponding author: Mei Lin, MD, PhD, Senior Scientist, Department of Clinical Laboratory, Taizhou People's Hospital (Postgraduate Training Base of Dalian Medical University), No. 366 Taihu Road, Taizhou 225300, Jiangsu Province, China. l_mei@163.com

Abstract

BACKGROUND

Colorectal cancer (CRC) is one of the most common malignancies worldwide. Given its insidious onset, the condition often already progresses to advanced stage when symptoms occur. Thus, early diagnosis is of great significance for timely clinical intervention, efficacy enhancement, and prognostic improvement. Featuring high throughput, fastness, and rich information, next generation sequencing (NGS) can greatly shorten the detection time, which is a widely used detection technique at present.

AIM

To screen specific genes or gene combinations in fecal DNA that are suitable for diagnosis and prognostic prediction of CRC, and to establish a technological

platform for CRC screening, diagnosis, and efficacy monitoring through fecal DNA detection.

METHODS

NGS was used to sequence the stool DNA of patients with CRC, which were then compared with the genetic testing results of the stool samples of normal controls and patients with benign intestinal disease, as well as the tumor tissues of CRC patients. Specific genes or gene combinations in fecal DNA suitable for diagnosis and prognostic prediction of CRC were screened, and their significances in diagnosing CRC and predicting patients' prognosis were comprehensively evaluated.

RESULTS

High mutation frequencies of *TP53*, *APC,* and *KRAS* were detected in the stools and tumor tissues of CRC patients prior to surgery. Contrastively, no pathogenic mutations of the above three genes were noted in the postoperative stools, the normal controls, or the benign intestinal disease group. This indicates that tumor-specific DNA was detectable in the preoperative stools of CRC patients. The preoperative fecal expression of tumor-associated genes can reflect the gene mutations in tumor tissues to some extent. Compared to the postoperative stools and the stools in the two control groups, the pathogenic mutation frequencies of *TP53* and *KRAS* were significantly higher for the preoperative stools ($\chi^2 = 7.328$, $P < 0.05$; $\chi^2 = 4.219$, $P < 0.05$), suggesting that fecal *TP53* and *KRAS* genes can be used for CRC screening, diagnosis, and prognostic prediction. No significant difference in the pathogenic mutation frequency of the *APC* gene was found from the postoperative stools or the two control groups ($\chi^2 = 0.878$, $P > 0.05$), so further analysis with larger sample size is required. Among CRC patients, the pathogenic mutation sites of *TP53* occurred in 16 of 27 preoperative stools, with a true positive rate of 59.26%, while the pathogenic mutation sites of *KRAS* occurred in 10 stools, with a true positive rate of 37.04%. The sensitivity and negative predictive values of the combined genetic testing of *TP53* and *KRAS* were 66.67% (18/27) and 68.97%, respectively, both of which were higher than those of *TP53* or *KRAS* mutation detection alone, suggesting that the combined genetic testing can improve the CRC detection rate. The mutation sites *TP53* exon 4 A84G and *EGFR* exon 20 I821T (mutation start and stop positions were both 7579436 for the former, while 55249164 for the latter) were found in the preoperative stools and tumor tissues. These "undetected" mutation sites may be new types of mutations occurring during the CRC carcinogenesis and progression, which needs to be confirmed through further research. Some mutations of "unknown clinical significance" were found in such genes as *TP53*, *PTEN*, *KRAS*, *BRAF*, *AKT1*, and *PIK3CA*, whose clinical values is worthy of further exploration.

CONCLUSION

NGS-based fecal genetic testing can be used as a complementary technique for the CRC diagnosis. Fecal *TP53* and *KRAS* can be used as specific genes for the screening, diagnosis, prognostic prediction, and recurrence monitoring of CRC. Moreover, the combined testing of *TP53* and *KRAS* genes can improve the CRC detection rate.

Key Words: Colorectal cancer; Feces; Next generation sequencing; Diagnosis; Gene

Core Tip: Colorectal cancer (CRC) is characterized by high morbidity and mortality, as well as low early diagnosis rate. The development of current gold standard for clinical diagnosis of CRC is restricted due to its invasiveness. The purpose of this study is to explore the potential value of fecal gene detection based on next generation sequencing in the diagnosis of CRC, to screen specific genes or gene combinations suitable for CRC diagnosis and prognosis prediction in fecal DNA, and to establish a technical platform for fecal DNA detection for CRC screening, diagnosis, and efficacy monitoring.

Citation: He SY, Li YC, Wang Y, Peng HL, Zhou CL, Zhang CM, Chen SL, Yin JF, Lin M. Fecal gene detection based on next generation sequencing for colorectal cancer diagnosis. *World J Gastroenterol* 2022; 28(25): 2920-2936

URL: https://www.wjgnet.com/1007-9327/full/v28/i25/2920.htm

DOI: https://dx.doi.org/10.3748/wjg.v28.i25.2920

INTRODUCTION

Colorectal cancer (CRC), as a health issue of great concern worldwide, poses a serious threat to human health. According to the latest data released by the International Agency for Research on Cancer, there were approximately 19.3 million new cancer cases and nearly 10 million cancer deaths globally in 2020. Among them, new CRC cases amounted to about 1.93 million (10%), ranking third, while CRC deaths amounted to about 0.94 million (9.4%), ranking second. For female patients, there were approximately 860000 new cases of CRC (9.4%), second only to breast cancer, and nearly 870000 CRC deaths (9.5%), ranking third. As for male population, the numbers of new cases and deaths both ranked third for CRC, which were 930000 (10.6%) and 50000 (9.3%), respectively[1].

In recent years, the morbidity and mortality of CRC have been on the rise in China, and the onset age of patients has gradually become younger. Early detection of CRC is difficult due to the insidious onset, long progressive course, and unobvious symptoms and signs of patients at the onset stage. As the tumor grows and gradually compresses the intestinal cavity, such symptoms as altered bowel habit and abnormal stool color or traits may occur. At this time, the condition often already progresses to advanced stage. Studies have shown that the 5-year survival of patients with early CRC can reach 90%, while is only about 12.5% for advanced stage patients[2]. Thus, early diagnosis is conducive to early intervention and treatment of CRC, which is of great significance for improving the efficacy and prognosis[3,4]. In response, the American Cancer Society recommended the CRC screening since the age of 45 for adults in 2018, which is 5 years earlier than the previously recommended age[5].

The occurrence and progression are a multi-factor, multi-stage complex process[6-8]. It is generally believed that CRC is caused by the mutations and malignant proliferation of colorectal mucosal epithelial cells. Other studies have shown that CRC can develop from inflammatory bowel disease (IBD). IBD is mainly characterized by intestinal mucosal barrier dysfunction, because the changes of cytokines aggravate tissue damage and eventually lead to tumorigenesis. It is reported that patients with ulcerative colitis have a higher risk of developing CRC than the general population[9]. In addition, microRNA and tumor microenvironment also have potential pathogenic effects on CRC[10,11]. Currently, the clinically recognized gold standard for CRC detection remains endoscopy combined with histopathological analysis. Given the heterogeneity of tumor tissues and the invasiveness of endoscopy, the same patient cannot be sampled multiple times, so that the genetic testing results cannot fully display the mutation spectrum of oncogenic genes[12]. According to the Chinese *Expert Consensus on Early Diagnosis and Screening Strategies for Colorectal Tumors* published in October 2018, fecal detection is a promising method for early diagnosis[13]. Its theoretical basis is that the tumor cells in the intestinal cavity of CRC patients adhere less to the base and can be continuously exfoliated from the colorectal mucosa. Studies have shown that tumor-associated DNA can stably exist in the stools, which has high specificity for the detection of CRC and is the most suitable biomarker in fecal detection[14-16]. By extracting the DNA of shed tumor cells in stools and analyzing their mutations, the traits of intestinal tumor cells can be reflected. Fecal DNA detection, as an emerging, noninvasive technique for screening CRC, has attracted great attention in recent years, which has the advantages of noninvasiveness, convenient access to materials, good patient compliance, and continuous dynamic monitoring. It is in line with the concept and conforms to the trend of "precision medicine", which is an exceptionally promising detection method.

With the continuous development of molecular biotechnology, next generation sequencing (NGS) has emerged to gradually become a hot research topic worldwide. It is also known as high-throughput sequencing, which operates on the following principles: Four different deoxynucleotide triphosphates (dNTPs) are labeled with different fluorescent colors. Conforming to the complementary base pairing rule, the bound base releases the corresponding fluorescence by the action of DNA polymerase every time a dNTP is added. Finally, the base is identified according to the fluorescent signal. Compared to the first-generation Sanger sequencing, NGS is a set of technologies that completely change the standard concept of nucleic acid sequencing. Capable of detecting millions of DNA molecules simultaneously, it achieves concurrent synthesis and sequencing, which can greatly shorten the detection time, and has the advantages of high throughput, high efficiency, and rich information[17,18]. NGS can more comprehensively display the full spectrum of tumor-associated genes, which can reflect the changes in tumor-specific genes, expand the understanding of tumor-specific genes, and improve the predictability of targeted therapy, suggesting its significance in achieving personalized tumor treatment and developing new therapeutic strategies[19-21]. Currently, NGS has been widely applied in the research of various solid tumors, including CRC. Kraus and his colleagues performed NGS analysis on the tissue paraffin sections and blood samples of 152 CRC patients based on 18 CRC-related genes. Their results showed that NGS had a high detection rate and could effectively identify the high-risk patients with hereditary CRC[22]. In a study by Li *et al*[23], the NGS (Ion Torrent PGM platform) technology was applied to analyze 22 tumor-associated mutation hotspots in 526 CRC patients, finding that *TP53* and *KRAS* mutations were the most common in tumor tissues, and confirming that NGS was a highly sensitive detection technique.

This study performed NGS-based genetic testing of 50 tumor-associated mutation hotspots in the stools of CRC patients, and compared them with the genetic testing results of corresponding tumor tissues, with an aim to find the fecal genes or gene combinations with high specificity and sensitivity,

and to establish a technological platform for CRC screening, diagnosis, and efficacy monitoring through fecal DNA detection. The findings herein will provide support and basis for personalized diagnosis and treatment of CRC.

MATERIALS AND METHODS

Subjects

Fifty tumor tissues, 27 preoperative stools, and 19 postoperative stools (stools formed over 7 d postoperatively) collected from patients with CRC were included in an experimental group. Twenty stool samples collected from healthy subjects were included in a normal control group. Seventeen stool samples of patients with benign intestinal polyps confirmed by digestive endoscopy were collected, as well as three stool samples from patients with ulcerative colitis, totaling 20 samples, were included in a benign control group. Postoperative pathological reports of polyp patients revealed inflammatory or adenomatous polyps, none of which had progressed to the CRC stage.

This study was approved by the Ethics Committee of Taizhou People's Hospital in Jiangsu, and the patients provided signed informed consent for sample collection.

Main reagents

Fecal DNA sample preservation tubes, tissue nucleic acid preservation tubes, fecal DNA extraction kit (immunomagnetic bead method), universal columnar genomic nucleic acid extraction kit, fast DNA library preparation kit were provided by Cowin Bio., Jiangsu, while high-throughput kit (MGISEQ-2000RS) was provided by MGI, Shenzhen.

Samples sources

The subjects of this project are patients who were diagnosed with CRC for the first time and underwent surgery at Taizhou People's Hospital from January 2019 to January 2021. A total of 54 tumor tissue samples, 45 preoperative stool samples, and 41 postoperative stool samples were collected. The normal control group consisted of healthy physical examination takers without malignancies (including CRC), intestinal polyps, or IBD, from whom 20 stool samples were collected. The benign control group comprised 20 patients who were diagnosed with intestinal polyps or IBD by digestive endoscopy, and admitted to the Department of Gastroenterology, Taizhou People's Hospital for polypectomy or anti-inflammatory treatment. Twenty formed stool samples were collected from patients who had not taken laxatives preoperatively or from patients before receiving anti-inflammatory therapies. The above samples were tested by NGS, unreliable results were filtered during analysis of genetic testing results based on biological information, and poor quality samples were eliminated (heterogeneous tumor tissues, atypical fecal specimens, and little or degraded DNA in stools). Finally, effectively detected samples included 50 tumor tissues, 27 preoperative stools, 19 postoperative stools, 20 stools in the normal control group, and 20 stools in the benign control group.

Sample collection and processing

Stool samples: Patients were instructed to use the fecal DNA preservation tube correctly (tubes contained protective solution to prevent the degradation of nucleic acid, while steel beads played a mixing role). They were asked to take samples immediately after natural defecation. Using sampling spoon located on the tube cap, typical formed stool about the size of soybean was excavated while avoiding spillover of protective solution and steel beads in the tube. After tightening the tube cap, the patients were asked to shake the tubes for approximately 30 s for uniform mixing. The samples were stored in a -80 °C refrigerator for subsequent testing.

Tissue samples: The freshly isolated tumor tissues were quickly cut into pieces (< 0.5 cm in thickness) using sterile tissue scissors, and then the tissue pieces were soaked in the protective solution of preservation tubes (if the volume was excessively large, the protective solution would not easily penetrate the tissues, which could not protect the nucleic acid from degradation). After tightening the tube caps, the samples were stored in a 4–8 °C refrigerator for subsequent testing.

DNA extraction from stool samples

The DNA in stool samples should be extracted strictly in accordance with the kit instructions (immunomagnetic bead method). Initially, the samples were homogenized, and then the precipitates (fecal particles) were removed, and the supernatants were retained for subsequent use. Different reagents were added to the corresponding wells of deep 96-well plates as *per* the instruction procedures, and then the deep well plates and magnetic sleeves containing the reagents were placed at the corresponding positions of nucleic acid extractor, followed by running of the stool extraction program. Under high salinity condition, DNA can bind to the surfaces of silica-coated Magbeads. After rinsing, high-purity DNA was eluted and stored in Buffer GE. The DNA-containing Buffer GE was transferred to an EP tube and stored under refrigeration for subsequent testing.

Table 1 details the reaction system. Relevant reagents were added to the corresponding wells of deep 96-well plates (wells 1–6 corresponded to one sample, wells 7–12 corresponded to another sample, and so forth).

The deep 96-well plates and corresponding magnetic sleeves were placed into the corresponding positions of an automatic nucleic acid extractor, and then the fecal DNA extraction program was initiated. After running for about 20 min, the program was paused, and each 220 μL of pre-prepared mixture (Magbeads PN 20 μL, isopropanol 200 μL) was added to the first well of each sample, followed by continuation of the program. The product in the last well of each sample was transferred to the centrifuge tube and stored at low temperature for subsequent testing.

DNA extraction from tissue samples

DNA was extracted from tissue samples strictly following the instructions of universal columnar genomic nucleic acid extraction kit. Initially, fresh tumor tissues were ground to small pieces and added with lysis buffer and proteinase K for complete digestion and lysis. Then, RNase was added to remove RNA. Since DNA was insoluble in absolute ethanol, the samples added with absolute ethanol were centrifuged to adsorb the DNA on the column. The residual impurities dissolved in absolute ethanol were discarded along with the liquid waste in the collection tube. Next, the collection tube was replaced with a new one, and the centrifugation was repeated several times. The DNA-containing adsorption column was placed at room temperature, and let stand to dry the absolute ethanol. The elution buffer was suspended, added to the adsorption column with a new collection tube, and centrifuged, and then the DNA-containing eluate in the collection tube was collected and refrigerated for subsequent testing.

The nucleic acid extraction was carried out in Cowin Bio., Jiangsu.

Library construction and NGS detection

In accordance with the instructions of fast DNA library preparation kit, the hotspot mutation panels of 50 tumor-associated genes designed for specific genomic regions or targets were added initially based on the multiplex polymerase chain reaction (PCR) library construction technology, and then multiplex PCR enrichment was performed on specific gene region fragments. Next, the obtained target gene fragments were subjected to end repair, barcode adapter ligation, and PCR amplification (Eastwin Scientific Equipment, Suzhou) to prepare a library. Afterwards, the library was quantified with Qubit 3.0 fluorescence quantifier (Thermo Fisher Scientific, Shanghai), and the fragment size of the library was analyzed with the Bioptic Qsep 100 automatic nucleic acid analyzer (Bioptic Inc., Jiangsu) to evaluate the library quality. Finally, a high-throughput sequencer (MGI, Shenzhen) was used for sequencing, and the data splitting, denoising, and comparison were accomplished in the Linux system to obtain the gene mutation results of various samples.

The library construction and NGS were carried out in Cowin Bio., Jiangsu.

Statistical analysis

Data were analyzed using SPSS 26.0 software. The gene mutation frequencies are expressed as the numbers of cases or percentages. Inter-group differences were examined by χ^2 and Fisher exact tests. $P < 0.05$ was considered statistically significant.

RESULTS

Clinical data analysis of CRC patients

Among the CRC patients corresponding to 50 tumor tissue samples, 24 were female (48.00%) and 26 were male (52.00%). The age span was large (35–84 years), and the median age was 59.5 years. According to the histopathological results of tumor tissues, there were 29 cases of rectal cancer (58.00%), 11 cases of left colon cancer + sigmoid colon cancer (22.00%), and 10 cases of right colon cancer (20.00%). Twelve cases (24.00%) were well-differentiated + well-to-moderately differentiated, 25 (50.00%) were moderately differentiated, and 12 (24.00%) were moderately-to-poorly differentiated + poorly differentiated (the degree of differentiation was not determined in 1 of 50 colorectal cancer tumor tissues). Among 19 cases (38.00%) of lymph node metastasis, only 1 had distant metastasis (metastasis to the liver). Regarding pathological staging (according to eighth edition of AJCC TNM staging system), 33 cases (66.00%) were at stages I + II, and 17 (34.00%) were at stages III + IV (Table 2).

Gene mutations in various samples

As is clear from the NGS testing results of CRC specimens in Table 3, genes like *TP53*, *APC*, *KRAS*, *PTEN*, *MET*, *HRAS*, and *PDGFRA* had high incidences of mutations in tumor tissues and preoperative stools, while in postoperative stools, the mutation frequencies of *TP53*, *HRAS*, and *PDGFRA* were high. Despite high mutation frequencies of *HRAS* and *PDGFRA* in the three groups of samples, their mutations were mostly same sense mutations, meaning that there were only base mutations, without changes in amino acids, which had no clinical significance.

Table 1 The reaction system

Position	Reagent
Wells 1 and 7	Proteinase K: 20 μL
	Lysate: 200 μL
Wells 2 and 8	Buffer KCL: 750 μL
Wells 3 and 9	Buffer GW1: 750 μL
Wells 4 and 10	Buffer GW2: 750 μL
Wells 5 and 11	Buffer MW3: 750 μL
Wells 6 and 12	Buffer GE: 100 μL

Table 2 Clinical data analysis of colorectal cancer patients

Clinical feature	Stratification	Number of people (%)
Age (years old)	≥ 60	30 (60.00)
	< 60	20 (40.00)
Sex	Male	26 (52.00)
	Female	24 (48.00)
Tumor location	Sigmoid colon + left hemicolon	11 (22.00)
	Right hemicolon	10 (20.00)
	Rectum	29 (58.00)
Tumor size (cm)	≥ 6	12 (24.00)
	< 6	38 (76.00)
Degree of tumor differentiation	High + medium high	12 (24.00)
	Moderately	25 (50.00)
	Low medium + low	12 (24.00)
TNM classification	I + II	33 (66.00)
	III + IV	17 (34.00)

As shown in Table 4, the *PDGFRA*, *HRAS*, and *KIT* mutations occurred in only 3 of 20 stool samples in the normal control group. In the benign control group, only 4 stool samples had *PDGFRA*, *HRAS*, *KIT*, or *STK11* mutations.

Since every gene has numerous mutation sites, different site mutations have different meanings. We classified and summarized the results of gene sequencing, and queried them online (https://www.ncbi.nlm.nih.gov/clinvar/). After combining the start and stop positions of each gene mutation site, various gene mutation sites were classified into four types depending on their pathogenicity: "Pathogenic", "benign", "unknown clinical significance", and "undetected in the system" (hereinafter referred to as "undetected"). "Benign" gene mutation sites imply a large number of mutations that can occur multiple times in the normal population, albeit less than 1% incidences. Mutation sites of "unknown clinical significance" indicate that such mutations are neither pathogenic nor benign, on whom the experimental conclusions in the literature reports are inconsistent. There is insufficient evidence to clearly classify them, and further exploration is required. "Undetected" mutation sites are identified by comprehensively checking multiple items such as the exons, the types of amino acid mutations, and the start and end positions of mutations. In the case of non-conformity of any item, the corresponding mutation site is regarded as "undetected". In Table 3, the mutations of the *PTEN* gene in tumor tissues were quite frequent, all of which were mutation sites of "unknown clinical significance". Among the 27 cases of preoperative stools, 1 had pathogenic mutation. Given the small case number, the possible cause of tumor heterogeneity was not ruled out. The *MET* gene exhibited high mutation frequencies in both tumor tissues and preoperative stools, all of which were "benign" mutations upon query. The four *MET* mutations in the postoperative stools were all "undetected" mutations. In Table 4, the mutation sites of *KIT* and *STK11* in the benign control group were benign mutation sites, while the rest were all same sense mutations and had no clinical significance.

Table 3 Summary of gene mutations in colorectal cancer patients

Gene	Tumor tissue		Preoperative feces		Postoperative stools	
	Total number of mutations	Mutation frequency	Total number of mutations	Mutation frequency	Total number of mutations	Mutation frequency
TP53	40	62.00% (31/50)	27	59.26% (16/27)	7	31.58% (6/19)
APC	15	24.00% (12/50)	3	11.11% (3/27)	0	0
KRAS	30	58.00% (29/50)	11	37.04% (10/27)	0	0
PIK3CA	13	22.00% (11/50)	1	3.70% (1/27)	2	10.53% (2/19)
FBXW7	5	10.00% (5/50)	1	3.70% (1/27)	0	0
GNAS	2	4.00% (2/50)	1	3.70% (1/27)	0	0
PTEN	11	22.00% (11/50)	1	3.70% (1/27)	0	0
ABL1	0	0	1	3.70% (1/27)	0	0
PDGFRA	12	24.00% (12/50)	8	29.63% (8/27)	6	31.58% (6/19)
ATM	1	2.00% (1/50)	0	0	0	0
SMAD4	1	2.00% (1/50)	0	0	0	0
BRAF	2	4.00% (2/50)	0	0	2	10.53% (2/19)
PTPN11	1	2.00% (1/50)	0	0	0	0
NRAS	1	2.00% (1/50)	0	0	0	0
CTNNB1	3	6.00% (3/50)	1	3.70% (1/27)	0	0
STK11	2	4.00% (2/50)	1	3.70% (1/27)	0	0
AKT1	1	2.00% (1/50)	1	3.70% (1/27)	0	0
CDKN2A	3	6.00% (3/50)	4	7.41% (2/27)	3	5.26% (1/19)
HRAS	21	42.00% (21/50)	12	40.74% (11/27)	8	42.11% (8/19)
EGFR	2	4.00% (2/50)	2	7.41% (2/27)	1	5.26% (1/19)
IDH1	2	4.00% (2/50)	2	7.41% (2/27)	1	5.26% (1/19)
NOTCH1	2	4.00% (2/50)	0	0	2	5.26% (1/19)
VHL	0	0	1	3.70% (1/27)	0	0
KIT	7	14.00% (7/50)	2	7.41% (2/27)	4	21.05% (4/19)
MET	9	18.00% (9/50)	5	18.52% (5/27)	4	21.05% (4/19)
MLH1	3	6.00% (3/50)	2	7.41% (2/27)	2	10.53% (2/19)
MPL	0	0	0	0	1	5.26% (1/19)

Total number of mutations: The overall number of mutations of any gene in various samples. Mutation frequency: The number of people with mutations in various samples/total number of people.

This study focused on analyzing the pathogenic mutation sites of various genes.

Comparison between preoperative stool and tumor tissue results in CRC patients

As shown in Table 5, the *TP53*, *APC*, and *KRAS* genes had high incidences of pathogenic mutations in tumor tissues and preoperative stools among CRC patients. Despite high mutation frequency of *PIK3CA* in tumor tissues, its pathogenic mutation frequency in preoperative stools was only 3.70% (1/27). The possible causes of tumor heterogeneity or individual disparity could not be ruled out. Although other genes had pathogenic mutations, their mutation frequencies were rather low.

Comparison between preoperative/postoperative stool results of CRC patients and control stool results

For CRC patients, the pathogenic mutation incidences of *APC*, *TP53*, and *KRAS* in the preoperative stools were 11.11% (3/27), 37.04% (10/27), and 25.93% (7/27), respectively. There were mutations of *PIK3CA*, *BRAF*, and *MPL* genes in the postoperative stools. Although they were pathogenic, only one

Table 4 Fecal gene mutation results in the control groups

Group	Genes	Location	Amino acid mutation	Mutation frequency
Normal control	*PDGFRA*	Exon 19	V824V	2
	KIT	Exon 10	K546K	1
	HRAS	Exon 2	H27H	1
Intestinal benign disease	*PDGFRA*	Exon 19	V824V	3
	HRAS	Exon 2	H27H	1
	KIT	Exon 10	M541L	1
	STK11	Exon 8	F354L	1

Table 5 Comparison of pathogenic mutation sites in preoperative stools *vs* tumor tissues

Gene	Positive rate of pathogenic gene mutation sites	
	Preoperative feces	Tumor tissues
TP53	37.04% (10/27)	46.00% (23/50)
APC	11.11% (3/27)	18.00% (9/50)
KRAS	25.93% (7/27)	54.00% (27/50)
PIK3CA	3.70% (1/27)	22.00% (11/50)
FBXW7	3.70% (1/27)	6.00% (3/50)
GNAS	3.70% (1/27)	4.00% (2/50)
PTEN	3.70% (1/27)	0
ABL1	3.70% (1/27)	0
PDGFRA	3.70% (1/27)	0
ATM	0	2.00% (1/50)
SMAD4	0	2.00% (1/50)
BRAF	0	2.00% (1/50)
PTPN11	0	2.00% (1/50)
NRAS	0	2.00% (1/50)
CTNNB1	0	2.00% (1/50)
VHL	3.70% (1/27)	0

Positive rate of pathogenic gene mutation site: The number of people with mutations corresponding to the pathogenic mutation sites of any gene/total number of samples.

case was found for each gene. The possible cause of tumor heterogeneity could not be ruled out. No pathogenic gene mutation sites were detected in stool samples from the normal and benign control groups (Tables 6–8).

According to a combination of Tables 5–8, the *TP53*, *APC*, and *KRAS* genes all had high mutation frequencies in the preoperative stools and tumor tissues of CRC patients, while exhibiting no pathogenic mutations in the postoperative stools, or in the stool samples of the normal or benign control group. This indicates that tumor-specific DNA can be detected in the preoperative stools of CRC patients, and that the preoperative fecal expression of tumor-associated genes can reflect the gene mutations in tumor tissue to some extent. Compared to the postoperative stools and the stool samples of the two control groups, the pathogenic mutation frequencies of *TP53* and *KRAS* were significantly higher in the preoperative stools ($P < 0.05$), suggesting that fecal *TP53* and *KRAS* genes can be used for the screening, diagnosis, and prognostic prediction of CRC. Contrastively, the pathogenic mutation frequency of the *APC* gene in the preoperative stools differed insignificantly from that in the postoperative stool or the two control groups ($P > 0.05$), which was probably associated with the excessively small sample size.

Table 6 Comparison of pathogenic mutation sites in preoperative stools *vs* normal control stools

Gene	Positive rate of pathogenic gene mutation sites		χ^2	*P* value
	Preoperative feces	Normal control group		
TP53	37.04% (10/27)	0	7.328	0.007
APC	11.11% (3/27)	0	0.878	0.349
KRAS	25.93% (7/27)	0	4.219	0.040
PIK3CA	3.70% (1/27)	0	Fisher exact test	1
FBXW7	3.70% (1/27)	0	Fisher exact test	1
GNAS	3.70% (1/27)	0	Fisher exact test	1
PTEN	3.70% (1/27)	0	Fisher exact test	1
ABL1	3.70% (1/27)	0	Fisher exact test	1
PDGFRA	3.70% (1/27)	0	Fisher exact test	1

Positive rate of pathogenic gene mutation site: The number of people with mutations corresponding to the pathogenic mutation sites of any gene/total number of samples.

Mutation site analysis for TP53, KRAS, and APC genes

As shown in Table 9, *TP53* mutations occurred in 31 of 50 tumor tissues, and the total mutation frequency was 40 times, of which pathogenic mutation sites accounted for 60.00% (24/40), with exon 5 R175H, exon 7 R248Q/W, and exon 8 R273H/C being the most common. The proportion of pathogenic mutation sites in 29 tumor tissues with *KRAS* mutations was 90.00% (27/30), most of which were located in exon 2, with G12D/V/S being predominant, followed by G13D/C. Additionally, there was 1 case of exon 3 Q61H mutation. As for the *APC* gene, all its pathogenic mutation sites were located in exon 17.

The 12 of 27 *TP53* mutations in preoperative stools were pathogenic mutations, which were mostly exon 7 R248Q and exon 8 R273H/C. The pathogenic mutation sites of the *KRAS* gene were primarily exon 2 G12D/V. As for the *APC* gene, all its pathogenic mutation sites were located in exon 17 (Table 10).

As are clear from Tables 9 and 10, the aforementioned pathogenic mutation sites occurred in both preoperative stools and tumor tissues, suggesting the possible correlation between preoperative stool and tumor tissue in CRC patients.

Combined TP53–KRAS detection

Among the 27 preoperative stools of CRC patients, 16 had pathogenic mutations of the *TP53* gene, with a true positive rate of 59.26%, and 10 had pathogenic mutations of the *KRAS* gene, with a true positive rate of 37.04%. Pathogenic mutations of *TP53* or *KRAS* occurred in 18 preoperative stools, showing a true positive rate of 66.67% (18/27). According to Table 11, the sensitivity and negative predictive value of the combined TP53–KRAS detection were 66.67% and 68.97%, respectively, which were higher than that those of *TP53* or *KRAS* mutation detection alone. This suggests that the combined testing of *TP53* and *KRAS* genes can improve the detection rate.

Distribution of mutation sites of "unknown clinical significance"

Mutation sites of "unknown clinical significance" occurred for such genes as TP53, *PTEN*, *KRAS*, *BRAF*, and *AKT1* in the tumor tissues of CRC patients. The five mutation sites of "unknown clinical significance" for *TP53* were respectively exon 5 A159V/V172G/G154D/P153S and exon 7 S241T. In the case of *PTEN*, 9 of 11 mutations were of "unknown clinical significance", all of which were exon 6 V133I. For the *KRAS* gene, the mutation sites of "unknown clinical significance" accounted for 10.00% (3/30), all of which were exon 4 A146T. In preoperative stools, the mutation sites of "unknown clinical significance" were scattered in *TP53*, *KRAS*, *AKT1*, and *STK11*. In the case of *TP53*, such mutation sites were exon 5 V172G and exon 7 I251T/S241T, which were identical to the mutation locations in tumor tissues. For the *KRAS* gene, the mutation site of "unknown clinical significance" was exon 4 A146T. As for *AKT1* and *STK11*, such mutation sites were located separately in exon 3 E17K and exon 4 E199K.

Distribution of "undetected" mutation sites

As shown in Table 12, genes like *TP53*, *APC*, *EGFR*, *FBXW7*, and *NOTCH1* in tumor tissues all had "undetected" mutation sites, among which exon 4 A84G of *TP53* occurred more frequently (4/40). For the *APC* gene, "undetected" mutation sites were all found in exon 17. The pathogenic mutation sites of *APC* in Tables 9 and 10 were also located in exon 17, suggesting that the "undetected" mutation sites of

Table 7 Comparison of pathogenic mutation sites in preoperative stools *vs* benign control group

Gene	Positive rate of pathogenic gene mutation sites		χ^2	*P* value
	Preoperative feces	Benign control group		
TP53	37.04% (10/27)	0	7.328	0.007
APC	11.11% (3/27)	0	0.878	0.349
KRAS	25.93% (7/27)	0	4.219	0.040
PIK3CA	3.70% (1/27)	0	Fisher exact test	1
FBXW7	3.70% (1/27)	0	Fisher exact test	1
GNAS	3.70% (1/27)	0	Fisher exact test	1
PTEN	3.70% (1/27)	0	Fisher exact test	1
ABL1	3.70% (1/27)	0	Fisher exact test	1
PDGFRA	3.70% (1/27)	0	Fisher exact test	1

Positive rate of pathogenic gene mutation site: The number of people with mutations corresponding to the pathogenic mutation sites of any gene/total number of samples.

Table 8 Comparison of pathogenic mutation sites in preoperative stools *vs* postoperative stools

Gene	Positive rate of pathogenic gene mutation sites		χ^2	*P* value
	Preoperative feces	Postoperative stools		
TP53	37.04% (10/27)	0	6.947	0.008
APC	11.11% (3/27)	0	0.804	0.370
KRAS	25.93% (7/27)	0	3.974	0.046
PIK3CA	3.70% (1/27)	5.26% (1/19)	Fisher exact test	1
FBXW7	3.70% (1/27)	0	Fisher exact test	1
GNAS	3.70% (1/27)	0	Fisher exact test	1
PTEN	3.70% (1/27)	0	Fisher exact test	1
ABL1	3.70% (1/27)	0	Fisher exact test	1
PDGFRA	3.70% (1/27)	0	Fisher exact test	1
BRAF	0	5.26% (1/19)	Fisher exact test	1
MPL	0	5.26% (1/19)	Fisher exact test	1

Positive rate of pathogenic gene mutation site: The number of people with mutations corresponding to the pathogenic mutation sites of any gene/total number of sample.

APC in Table 12 may be unreported new sites. They are probably associated with the carcinogenesis and progression of CRC, which require further exploration. Table 13 shows that 6 of 27 *TP53* mutation sites in the preoperative stools were "undetected" sites. They were distributed in exon 4, exon 5, exon 7, and exon 8, among which there were three exon 7 N247D (mutation start and stop positions were both 7577542). Additionally, genes like *CDKN2A*, *HRAS*, *EGFR*, *IDH1*, and *PDGFRA* also had "undetected" mutations sites.

DISCUSSION

Studies have shown that the genetic factors coupled with prolonged unfavorable external factors induce the disturbance of homeostasis, which leads to a series of changes including angiogenesis, cell hyperproliferation, apoptosis evasion, and enhanced invasion capacity, ultimately resulting in the occurrence of malignancies. It is generally believed that CRC often develops slowly from adenomas, during which multiple steps are involved, including the activation of proto-oncogenes and the

Table 9 Analysis of pathogenic mutation sites in tumor tissues

Gene	Location	Amino acid mutation	Mutation frequency	N	M
TP53	Exon 4	R110L	1		
	Exon 5	R175H	4		
		Y163C	1		
		C176Y	1		
	Exon 6	R196[1]	2		
		Y220C	2		
	Exon 7	R248Q/W	3		
		G245S/C	2		
	Exon 8	R273H/C	4		
		R282W	2		
		R306[1]	1		
	Exon 10	R342[1]	1	24/40	23/31
KRAS	Exon 2	G12D/V/S	22		
		G13D/C	4		
	Exon 3	Q61H	1	27/30	27/29
APC	Exon 17	Q886[1]	1		
		S1483fs	1		
		R876[1]	2		
		R1450[1]	2		
		E1306[1]	1		
		Q1294[1]	1		
		G1312[1]	1	9/15	9/12

[1]Indicates stop codon.

N: Pathogenic sites/total mutation sites; M: Number of pathogenic cases/total number of mutation cases.

inactivation of tumor suppressor genes. The mutation statuses of genes like *TP53*, *APC*, and *KRAS* have been reported to be closely associated with the carcinogenesis and progression of CRC[24,25]. With the introduction of the concept of "precision medicine", molecular targeted therapy has become a hot research topic in recent years. This NGS-based study investigated the value of fecal genetic testing in the CRC diagnosis.

As a tumor suppressor gene, *TP53* regulates downstream genes when cells are under stress, which induces apoptosis and cell cycle arrest *via* a series of processes, thereby preventing the growth of tumor cells[26,27]. Its mutation is a key step in facilitating the adenoma–adenocarcinoma transition. In malignancies, the p53 protein encoded by *TP53* is mutant. The mutant p53 protein can lose its tumor suppressor effect and acquire new oncogenic functions to promote the invasion and metastasis of tumor cells[28]. In this study, *TP53* had 60.00% (24/40) and 59.26% (16/27) of mutations in the tumor tissues and preoperative stools of CRC patients, respectively, among which pathogenic mutation sites accounted for 60.00% (24/40) and 40.74% (11/27) of total mutation sites. Such pathogenic mutation sites as exon 5 C176Y, exon 6 R196*, exon 7 R248Q, exon 7 G245C/S, and exon 8 R273H/C existed in both types of samples, showing consistency. Such mutation hotspots as *R273H*, *R248Q*, and *R282W* have been reported as typical "hotspots" in the DNA binding domain. Over 95% of missense *TP53* mutations occur in the DNA binding domain, which can affect the ability of proteins to bind to their target DNA sequences[29-31]. Lo *et al*[32] screened a patient with CRC lymphatic metastasis for tumor infiltrating lymphocytes. Since some tumor patients had TP53 R175H mutations during genetic testing, and HLA-A *0201 could limit the recognition of mutant TP53 R175H, they identified the T cell receptors capable of recognizing the TP53 R175H mutation hotspots, thereby finding patients with different tumor types who shared the same mutation hotspots. In their opinion, the mutant *TP53* could serve as an ideal common antigen target, and these specific hotspot mutations had potential as valuable therapeutic targets for a variety of tumor cells. A study by Olszewski *et al*[33] suggested that changes in the

Table 10 Analysis of pathogenic mutation sites in preoperative feces

Gene	Location	Amino acid mutation	Mutation frequency	N	M
TP53	Exon 5	C176Y	1		
		F134V	1		
	Exon 6	R196[1]	1		
	Exon 7	N235D	1		
		R248Q	3		
		G245C/S	2		
	Exon 8	R273H/C	3	12/27	10/16
KRAS	Exon 2	G12D/V	5		
		G13D	1		
	Exon 3	Q61H	1		
		A59T	1	8/11	7/10
APC	Exon 17	R1450[1]	1		
		Q886[1]	1		
		Q1294[1]	1	3/3	3/3

[1]Indicates stop codon.
N: Pathogenic sites/total mutation sites; M: Number of pathogenic cases/total number of mutation cases.

Table 11 Colorectal cancer diagnosis results by *TP53* and *KRAS* mutations in preoperative stools, *n* (%)

Gene	TP	FP	FN	TN	Sensitivity	Specificity	PPV	NPV
TP53	16	0	11	20	59.26	100.00	100.00	64.52
KRAS	10	0	17	20	37.04	100.00	100.00	54.05
TP53 or *KRAS*	18	0	9	20	66.67	100.00	100.00	68.97

TP: True positive; FP: False positive; FN: False negative; TN: True negative; PPV: Positive predictive value; NPV: Negative predictive value.

oncogenic activity of TP53 led to acquisition of new functions for the mutated tumor suppressor p53 protein, and that the overexpression of such mutation hotspots as codons 175, 248, or 273 might be associated with the chemotherapeutic resistance or invasive changes. Their analysis of *TP53* mutation database in human somatic cells revealed that patients carrying TP53 R248Q allele had earlier tumor onsets, but lower frequencies of metastasis.

The 21 kDa protein encoded by *KRAS*, a member of the ras proto-oncogene family, can participate in the G protein signaling pathway and regulate the proliferation and differentiation of cells. *KRAS* mutations can continuously activate the RAS-RAF-MAPK pathway, causing uncontrolled cellular proliferation and differentiation, thereby inducing CRC[14]. Research has demonstrated that the patients' resistance to EGFR-targeted drugs was closely associated with the activation of the *KRAS* gene[34]. In the present study, the incidences of *KRAS* mutations in tumor tissue and preoperative stools of CRC patients were 58.00% (29/50) and 37.04% (10/27), respectively. Among them, pathogenic mutation sites existed in 27 tissues and 7 preoperative stools, with the most common ones being G12D/V and G13D located in exon 2. Armengol *et al*[35] explored the gene mutations in stool samples of 52 CRC patients. They found 12 cases of *KRAS* gene mutations, of which 58.3% (7/12) were located in codons 12 and 13 of exon 2. Knight argued that when there was a KRAS G12D carcinogenic mutation in the body, the formation or growth of intestinal adenomas was not inhibited by rapamycin[36]. Compared to KRAS G12D, patients with KRAS A146T mutation could have stronger resistance to EGFR tolerance, and better survival rates[37]. Mulla *et al*[38] analyzed the RAS mutations in 51 CRC patients, finding that 43% of them had RAS mutations, 91% of which were *KRAS* mutations. The majority of *KRAS* mutations were located in G12D, followed by G13D. They also found that the incidence of *KRAS* mutations was slightly higher among individuals under 50 years of age than that of the elderly (> 50 years). Ottaiano *et al*[39] reported that mutant *KRAS* was a negative influencing factor of prognosis in CRC patients. Among these patients, those with G12C/S mutation showed the shortest survival and the worst prognosis.

Table 12 "Undetected" gene mutation sites in tumor tissues

Gene	Location	Amino acid mutation	Mutation start position	Mutation end position	Number of cases	Undetected sites/total mutation sites
TP53	Exon 4	A84G	7579436	7579436	4	
	Exon 5	P152A	7578476	7578476	1	
	Exon 8	L289P	7577072	7577072	1	6/40
APC	Exon 17	S1346[1]	112175328	112175328	1	
		K1573fs	112175953	112175954	1	
		E1327fs	112175213	112175217	1	3/15
FBXW7	Exon 7	R278[1]	153258983	153258983	1	
	Exon 12	R266C	153247289	153247289	1	2/5
NOTCH1	Exon 26	R1599P	139399350	139399350	2	2/2
EGFR	Exon 20	I821T	55249164	55249164	1	1/2

[1]Indicates stop codon.
fs: Frameshift mutation.

Table 13 "Undetected" gene mutation sites in preoperative feces

Gene	Location	Amino acid mutation	Mutation start position	Mutation end position	Number of cases	Undetected sites/total mutation sites
TP53	Exon 4	A84G	7579436	7579436	1	
	Exon 5	S166P	7578434	7578434	1	
	Exon 7	N247D	7577542	7577542	3	
	Exon 8	L289P	7577072	7577072	1	6/27
CDKN2A	Exon 2	V51A	21971206	21971206	2	
		L63P	21971170	21971170	1	3/4
HRAS	Exon 2	G12R	534289	534289	1	1/10
EGFR	Exon 20	I821T	55249164	55249164	1	1/2
IDH1	Exon 4	R119Q	209113151	209113151	1	1/2
PDGFRA	Exon 13	G594fs	55141059	55141059	1	1/8

fs: Frameshift mutation.

Varshavi *et al*[40] investigated how *KRAS* mutations in different codons affected the metabolic pathways. According to their findings, metabolic differences existed between various *KRAS* mutation sites, which might be one of the reasons for the varying efficacies of anti-cancer therapies across patients. Hence, more effective, individualized molecular targeted drugs can be searched based on such metabolic differences.

Mutations of *APC* were linked to familial adenomatous polyposis, and its inactivation and the activation of Wnt signaling pathway also played pivotal roles in the occurrence of CRC, which could affect the adhesion between tumor cells to accelerate the cellular migration[41,42]. It has been reported that in sporadic CRC, most of the *APC* mutations occurred around the mutation cluster region, especially between codons 1286 and 1513[43]. In this study, the incidences of *APC* mutations in tumor tissues and preoperative stools were 24.00% (12/50) and 11.11% (3/27), respectively. As shown in Tables 5 and 6, the pathogenic mutation frequency of the *APC* gene in preoperative stools differed insignificantly from that in control stools (P = 0.370), which thus cannot serve as a specific gene for CRC screening and diagnosis. This may be attributed to the small sample size of this study, which remains further exploration. We found that all the pathogenic mutation sites of *APC* were located in exon 17, most of which used the mutations of certain amino acids as stop codons, thus allowing premature termination of mRNA translation (Tables 8 and 9). Luo's team performed NGS-based gene sequencing

on the blood and tissue samples of 22 CRC patients and 21 patients with intestinal polyps, in order to understand gene mutations. Through screening, they found that the *APC* mutations were most common in both types of samples, which were mostly pathogenic mutation sites. These pathogenic mutation sites of the *APC* gene accounted for 77.3% (17/22) of the total pathogenic mutation sites in all samples[44].

In addition to the pathogenic mutation sites of various samples, this study also found multiple mutation sites of "unknown clinical significance", which were scattered in such genes as *TP53*, *PTEN*, *KRAS*, *BRAF*, *AKT1*, and *PIK3CA*. There were also some gene mutation sites that had not been found in the Clinvar database. *TP53* exon4 A84G and *EGFR* exon20 I821T were present in both the preoperative stools and tumor tissues of CRC patients. The mutation start and stop positions were both 7579436 for the former, while 55249164 for the latter.

CONCLUSION

Conclusively, NGS-based fecal DNA detection can be used as a complementary technique for the CRC diagnosis. Fecal *TP53* and *KRAS* can serve as a specific gene combination for the screening, diagnosis, prognostic prediction, and recurrence monitoring of CRC. Moreover, the combined TP53–KRAS testing can improve the CRC detection rate, which provides a molecular reference for developing personalized treatment programs. The mutation sites of "unknown clinical significance" and the "undetected" mutation sites may be unreported new mutation sites associated with CRC. Whether these genes are linked to the CRC diagnosis and prognosis remains to be further explored and researched.

ARTICLE HIGHLIGHTS

Research background

Colorectal cancer (CRC) is currently a health problem of global concern. In recent years, the incidence of CRC presents a trend of gradual increase. Most patients have unobvious early symptoms, and they are commonly in mid and advanced stages when the symptoms become evident, with rather high mortalities. Featuring high throughput, fastness, and rich information, next generation sequencing (NGS) can greatly shorten the detection time, which is a research hotspot at home and abroad at present.

Research motivation

As we all know, histopathological examination is the gold standard of diagnosis, but its invasiveness limits its development. Therefore, it is imperative to explore the screening, diagnosis, and prognosis of CRC by strong specificity, high sensitivity, and non-invasive methods.

Research objectives

In this study, NGS technology was used to conduct genetic testing on stool samples of CRC patients, and the results were compared with the corresponding tumor tissue genetic testing results. The aim was to find genes or gene combinations with high specificity and sensitivity in the stool and establish a technical platform for CRC screening and diagnosis and curative effect monitoring through fecal DNA detection, providing a strong basis and support for personalized diagnosis and treatment of CRC.

Research methods

NGS was used to sequence the DNA in stools of patients with CRC, which were then compared with the genetic testing results of the stool samples of normal control and benign intestinal disease groups, as well as the tumor tissues of CRC patients. Specific genes or gene combinations in fecal DNA suitable for diagnosis and prognostic prediction of CRC were screened, and their significance in diagnosing CRC and predicting patients' prognosis was comprehensively evaluated.

Research results

High mutation frequencies of *TP53*, *APC,* and *KRAS* were detected in the stools and tumor tissues of CRC patients prior to surgery. Contrastively, no pathogenic mutations of the above three genes were noted in the postoperative stools, or two control groups. This indicates that the tumor-specific DNA was detectable in the preoperative stools of CRC patients. Compared to the postoperative stools and the stools in the two control groups, the pathogenic mutation frequencies of *TP53* and *KRAS* were significantly higher for the preoperative stools ($P < 0.05$), suggesting that fecal *TP53* and *KRAS* genes can be used for CRC screening, diagnosis, and prognostic prediction. No significant difference in the pathogenic mutation frequency of the *APC* gene was found from the postoperative stools or the two control groups ($P > 0.05$), so further analysis with larger sample size is required. In 27 preoperative stools of CRC patients, the sensitivity and negative predictive value of *TP53*- *KRAS* gene combination detection were higher than those of *TP53* mutation or *KRAS* mutation alone, suggesting that *TP53*-*KRAS*

gene combination detection can improve the detection rate of CRC. The "undetected" mutation sites found in preoperative stools and tumor tissues may be new mutation types in the occurrence and development of CRC, which need to be further studied. In addition, some mutations of "unknown clinical significance" were found, and their clinical value is worth further study.

Research conclusions

NGS-based fecal genetic testing can be used as a complementary technique for the CRC diagnosis. Fecal *TP53* and *KRAS* can be used as specific genes for the screening, diagnosis, prognostic prediction, and recurrence monitoring of CRC. Moreover, the combined testing of *TP53* and *KRAS* genes can improve the CRC detection rate.

Research perspectives

Fecal genetic detection is a new method for CRC diagnosis, which has the advantages of non-invasiveness, convenient sampling, and dynamic monitoring. Although the sensitivity of fecal genetic test in CRC screening is low, it is certain that it has great potential and broad prospects in the diagnosis and prognosis assessment of CRC. In addition, the "undetected" mutation sites in preoperative stools of CRC patients and the "unknown clinical significance" mutation sites are related to the occurrence and development of CRC, which requires further research and exploration.

FOOTNOTES

Author contributions: He SY performed most of the experiments, analyzed some data, and drafted the manuscript; Li YC and Wang Y provided specimens for the study; Peng HL and Zhou CL gave constructive guidance on the study; Zhang CM contributed to the statistics; Chen SL and Yin JF performed some of the experiments; Lin M designed and supervised the study and edited the manuscript; all authors approved the final version of the article.

Supported by Taizhou Social Development Plan, No. TS202004; Natural Science Foundation of Nanjing University of Chinese Medicine China, No. XZR2020093; and Taizhou People's Hospital Medical Innovation Team Foundation, No. CXTDA201901.

Institutional review board statement: This study was reviewed and approved by the Ethics Committee of Taizhou people's Hospital in Jiangsu Province (No. KY201912501).

Conflict-of-interest statement: All the authors report no relevant conflicts of interest for this article.

Data sharing statement: No additional data are available.

Country/Territory of origin: China

ORCID number: Si-Yu He 0000-0003-3411-8616; Ying-Chun Li 0000-0003-0650-531X; Yong Wang 0000-0002-6348-4562; Hai-Lin Peng 0000-0001-8332-1704; Cheng-Lin Zhou 0000-0003-2661-2067; Chuan-Meng Zhang 0000-0002-5642-1512; Sheng-Lan Chen 0000-0002-8366-451X; Jian-Feng Yin 0000-0002-5638-263X; Mei Lin 0000-0001-9815-2966.

S-Editor: Fan JR
L-Editor: Wang TQ
P-Editor: Fan JR

REFERENCES

1 **Sung H**, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* 2021; **71**: 209-249 [PMID: 33538338 DOI: 10.3322/caac.21660]
2 **Siegel R**, Desantis C, Jemal A. Colorectal cancer statistics, 2014. *CA Cancer J Clin* 2014; **64**: 104-117 [PMID: 24639052 DOI: 10.3322/caac.21220]
3 **Huang GW**. Clinical Value of Combined Examination of Fecal Occult Blood and Tumor Markers in the Diagnosis of Colorectal Cancer. [cited 10 January 2022]. Available from: http://en.cnki.com.cn/Article_en/CJFDTotal-HZZZ201907062.htm
4 **Chen WQ**, Li XQ, Zhang XM, Wu JN, Zhao XH, Gu HG, Ren J, Wang DQ. Expression and clinical significance of C-

terminal tensin-like protein in colorectal cancer. *Zhonghua Xiaohua Waike Zazhi* 2016; **15**: 809-814 [DOI: 10.21767/2172-0479.100116]
5 **Croke L**. Colorectal Cancer Screening: ACS Updates Guideline for Adults with Average Risk. *Am Fam Physician* 2019; **99**: 129-130 [PMID: 30633474]
6 **van Lanschot MC**, Carvalho B, Coupé VM, van Engeland M, Dekker E, Meijer GA. Molecular stool testing as an alternative for surveillance colonoscopy: a cross-sectional cohort study. *BMC Cancer* 2017; **17**: 116 [PMID: 28173852 DOI: 10.1186/s12885-017-3078-y]
7 **Sun Y**, Yang B, Lin M, Yu H, Chen H, Zhang Z. Identification of serum miR-30a-5p as a diagnostic and prognostic biomarker in colorectal cancer. *Cancer Biomark* 2019; **24**: 299-305 [PMID: 30829615 DOI: 10.3233/CBM-182129]
8 **Loktionov A**. Biomarkers for detecting colorectal cancer non-invasively: DNA, RNA or proteins? *World J Gastrointest Oncol* 2020; **12**: 124-148 [PMID: 32104546 DOI: 10.4251/wjgo.v12.i2.124]
9 **Al Buthi SA**, Bin Gheshayan SF, Al Samaani I, Al Ahmadi B, Al Selaim NA. A Mixed Neuroendocrine/Non-Neuroendocrine Neoplasm Arising in the Background of Ulcerative Colitis: A Case Report and Review of the Literature. *Am J Case Rep* 2022; **23**: e934396 [PMID: 35236818 DOI: 10.12659/AJCR.934396]
10 **Yan F**, Tu Z, Duan L, Wang D, Lin F. MicroRNA-383 suppresses cell proliferation and invasion in colorectal cancer by directly targeting paired box 6. *Mol Med Rep* 2018; **17**: 6893-6901 [PMID: 29512711 DOI: 10.3892/mmr.2018.8682]
11 **Yu CT**, Chen T, Lu S, Hu W, Zhang Q, Tan J, Sun D, Li L, Sun X, Xu C, Lai Y, Fan M, Shen Z, Shen W, Cheng H. Identification of Significant Modules and Targets of Xian-Lian-Jie-Du Decoction Based on the Analysis of Transcriptomics, Proteomics and Single-Cell Transcriptomics in Colorectal Tumor. *J Inflamm Res* 2022; **15**: 1483-1499 [PMID: 35256851 DOI: 10.2147/JIR.S344861]
12 **Wang B**, Wu S, Huang F, Shen M, Jiang H, Yu Y, Yu Q, Yang Y, Zhao Y, Zhou Y, Pan B, Liu T, Guo W. Analytical and clinical validation of a novel amplicon-based NGS assay for the evaluation of circulating tumor DNA in metastatic colorectal cancer patients. *Clin Chem Lab Med* 2019; **57**: 1501-1510 [PMID: 31339850 DOI: 10.1515/cclm-2019-0142]
13 **Ma CX**, Guan X, Wang S, Liu Z, Jiang Z, Wang XS. [Application and prospect of fecal DNA test in colorectal cancer screening]. *Zhonghua Wei Chang Wai Ke Za Zhi* 2019; **22**: 491-494 [PMID: 31104434 DOI: 10.3760/cma.j.issn.1671-0274.2019.05.018]
14 **Mahasneh A**, Al-Shaheri F, Jamal E. Molecular biomarkers for an early diagnosis, effective treatment and prognosis of colorectal cancer: Current updates. *Exp Mol Pathol* 2017; **102**: 475-483 [PMID: 28506769 DOI: 10.1016/j.yexmp.2017.05.005]
15 **Carethers JM**. Fecal DNA Testing for Colorectal Cancer Screening. *Annu Rev Med* 2020; **71**: 59-69 [PMID: 31451044 DOI: 10.1146/annurev-med-103018-123125]
16 **Anderson BW**, Ahlquist DA. Molecular Detection of Gastrointestinal Neoplasia: Innovations in Early Detection and Screening. *Gastroenterol Clin North Am* 2016; **45**: 529-542 [PMID: 27546847 DOI: 10.1016/j.gtc.2016.04.009]
17 **Wakai T**, Prasoon P, Hirose Y, Shimada Y, Ichikawa H, Nagahashi M. Next-generation sequencing-based clinical sequencing: toward precision medicine in solid tumors. *Int J Clin Oncol* 2019; **24**: 115-122 [PMID: 30515675 DOI: 10.1007/s10147-018-1375-3]
18 **Del Vecchio F**, Mastroiaco V, Di Marco A, Compagnoni C, Capece D, Zazzeroni F, Capalbo C, Alesse E, Tessitore A. Next-generation sequencing: recent applications to the analysis of colorectal cancer. *J Transl Med* 2017; **15**: 246 [PMID: 29221448 DOI: 10.1186/s12967-017-1353-y]
19 **Hinrichs JW**, van Blokland WT, Moons MJ, Radersma RD, Radersma-van Loon JH, de Voijs CM, Rappel SB, Koudijs MJ, Besselink NJ, Willems SM, de Weger RA. Comparison of next-generation sequencing and mutation-specific platforms in clinical practice. *Am J Clin Pathol* 2015; **143**: 573-578 [PMID: 25780010 DOI: 10.1309/AJCP40XETVYAMJPY]
20 **Nagahashi M**, Shimada Y, Ichikawa H, Kameyama H, Takabe K, Okuda S, Wakai T. Next generation sequencing-based gene panel tests for the management of solid tumors. *Cancer Sci* 2019; **110**: 6-15 [PMID: 30338623 DOI: 10.1111/cas.13837]
21 **Xuan J**, Yu Y, Qing T, Guo L, Shi L. Next-generation sequencing in the clinic: promises and challenges. *Cancer Lett* 2013; **340**: 284-295 [PMID: 23174106 DOI: 10.1016/j.canlet.2012.11.025]
22 **Kraus C**, Rau TT, Lux P, Erlenbach-Wünsch K, Löhr S, Krumbiegel M, Thiel CT, Stöhr R, Agaimy A, Croner RS, Stürzl M, Hohenberger W, Hartmann A, Reis A. Comprehensive screening for mutations associated with colorectal cancer in unselected cases reveals penetrant and nonpenetrant mutations. *Int J Cancer* 2015; **136**: E559-E568 [PMID: 25142776 DOI: 10.1002/ijc.29149]
23 **Li W**, Qiu T, Guo L, Ying J, Zhou A. NGS-based oncogenic mutations analysis in advanced colorectal cancer patients improves targeted therapy prediction. *Pathol Res Pract* 2019; **215**: 483-489 [PMID: 30611619 DOI: 10.1016/j.prp.2018.12.037]
24 **Fearon ER**, Vogelstein B. A genetic model for colorectal tumorigenesis. *Cell* 1990; **61**: 759-767 [PMID: 2188735 DOI: 10.1016/0092-8674(90)90186-i]
25 **Kronborg O**, Fenger C, Olsen J, Jørgensen OD, Søndergaard O. Randomised study of screening for colorectal cancer with faecal-occult-blood test. *Lancet* 1996; **348**: 1467-1471 [PMID: 8942774 DOI: 10.1016/S0140-6736(96)03430-7]
26 **Soussi T**, Wiman KG. Shaping genetic alterations in human cancer: the p53 mutation paradigm. *Cancer Cell* 2007; **12**: 303-312 [PMID: 17936556 DOI: 10.1016/j.ccr.2007.10.001]
27 **Li XL**, Zhou J, Chen ZR, Chng WJ. P53 mutations in colorectal cancer - molecular pathogenesis and pharmacological reactivation. *World J Gastroenterol* 2015; **21**: 84-93 [PMID: 25574081 DOI: 10.3748/wjg.v21.i1.84]
28 **Tong X**, Xu D, Mishra RK, Jones RD, Sun L, Schiltz GE, Liao J, Yang GY. Identification of a druggable protein-protein interaction site between mutant p53 and its stabilizing chaperone DNAJA1. *J Biol Chem* 2021; **296**: 100098 [PMID: 33208462 DOI: 10.1074/jbc.RA120.014749]
29 **Shah HD**, Saranath D, Murthy V. A molecular dynamics and docking study to screen anti-cancer compounds targeting mutated p53. *J Biomol Struct Dyn* 2022; **40**: 2407-2416 [PMID: 33111621 DOI: 10.1080/07391102.2020.1839559]
30 **Navalkar A**, Ghosh S, Pandey S, Paul A, Datta D, Maji SK. Prion-like p53 Amyloids in Cancer. *Biochemistry* 2020; **59**: 146-155 [PMID: 31603660 DOI: 10.1021/acs.biochem.9b00796]

31 **Cho Y**, Gorina S, Jeffrey PD, Pavletich NP. Crystal structure of a p53 tumor suppressor-DNA complex: understanding tumorigenic mutations. *Science* 1994; **265**: 346-355 [PMID: 8023157 DOI: 10.1126/science.8023157]

32 **Lo W**, Parkhurst M, Robbins PF, Tran E, Lu YC, Jia L, Gartner JJ, Pasetto A, Deniger D, Malekzadeh P, Shelton TE, Prickett T, Ray S, Kivitz S, Paria BC, Kriley I, Schrump DS, Rosenberg SA. Immunologic Recognition of a Shared p53 Mutated Neoantigen in a Patient with Metastatic Colorectal Cancer. *Cancer Immunol Res* 2019; **7**: 534-543 [PMID: 30709841 DOI: 10.1158/2326-6066.CIR-18-0686]

33 **Olszewski MB**, Pruszko M, Snaar-Jagalska E, Zylicz A, Zylicz M. Diverse and cancer typespecific roles of the p53 R248Q gainoffunction mutation in cancer migration and invasiveness. *Int J Oncol* 2019; **54**: 1168-1182 [PMID: 30968154 DOI: 10.3892/ijo.2019.4723]

34 **Yun J**, Mullarky E, Lu C, Bosch KN, Kavalier A, Rivera K, Roper J, Chio II, Giannopoulou EG, Rago C, Muley A, Asara JM, Paik J, Elemento O, Chen Z, Pappin DJ, Dow LE, Papadopoulos N, Gross SS, Cantley LC. Vitamin C selectively kills KRAS and BRAF mutant colorectal cancer cells by targeting GAPDH. *Science* 2015; **350**: 1391-1396 [PMID: 26541605 DOI: 10.1126/science.aaa5004]

35 **Armengol G**, Sarhadi VK, Ghanbari R, Doghaei-Moghaddam M, Ansari R, Sotoudeh M, Puolakkainen P, Kokkola A, Malekzadeh R, Knuutila S. Driver Gene Mutations in Stools of Colorectal Carcinoma Patients Detected by Targeted Next-Generation Sequencing. *J Mol Diagn* 2016; **18**: 471-479 [PMID: 27155048 DOI: 10.1016/j.jmoldx.2016.01.008]

36 **Knight JRP**, Alexandrou C, Skalka GL, Vlahov N, Pennel K, Officer L, Teodosio A, Kanellos G, Gay DM, May-Wilson S, Smith EM, Najumudeen AK, Gilroy K, Ridgway RA, Flanagan DJ, Smith RCL, McDonald L, MacKay C, Cheasty A, McArthur K, Stanway E, Leach JD, Jackstadt R, Waldron JA, Campbell AD, Vlachogiannis G, Valeri N, Haigis KM, Sonenberg N, Proud CG, Jones NP, Swarbrick ME, McKinnon HJ, Faller WJ, Le Quesne J, Edwards J, Willis AE, Bushell M, Sansom OJ. MNK Inhibition Sensitizes *KRAS*-Mutant Colorectal Cancer to mTORC1 Inhibition by Reducing eIF4E Phosphorylation and c-MYC Expression. *Cancer Discov* 2021; **11**: 1228-1247 [PMID: 33328217 DOI: 10.1158/2159-8290.CD-20-0652]

37 **Poulin EJ**, Bera AK, Lu J, Lin YJ, Strasser SD, Paulo JA, Huang TQ, Morales C, Yan W, Cook J, Nowak JA, Brubaker DK, Joughin BA, Johnson CW, DeStefanis RA, Ghazi PC, Gondi S, Wales TE, Iacob RE, Bogdanova L, Gierut JJ, Li Y, Engen JR, Perez-Mancera PA, Braun BS, Gygi SP, Lauffenburger DA, Westover KD, Haigis KM. Tissue-Specific Oncogenic Activity of $KRAS^{A146T}$. *Cancer Discov* 2019; **9**: 738-755 [PMID: 30952657 DOI: 10.1158/2159-8290.CD-18-1220]

38 **Mulla N**, Alshareef A, Syed AR, Al-Jahel M. Clinico-Pathological Study of K-ras Mutations in Colorectal Tumors: A Single-Center Retrospective Study of 51 Patients in Madinah, Saudi Arabia. *Cureus* 2020; **12**: e9978 [PMID: 32983680 DOI: 10.7759/cureus.9978]

39 **Ottaiano A**, Normanno N, Facchini S, Cassata A, Nappi A, Romano C, Silvestro L, De Stefano A, Rachiglio AM, Roma C, Maiello MR, Scala S, Delrio P, Tatangelo F, Di Mauro A, Botti G, Avallone A, Nasti G. Study of Ras Mutations' Prognostic Value in Metastatic Colorectal Cancer: STORIA Analysis. *Cancers (Basel)* 2020; **12** [PMID: 32708575 DOI: 10.3390/cancers12071919]

40 **Varshavi D**, Varshavi D, McCarthy N, Veselkov K, Keun HC, Everett JR. Metabolic characterization of colorectal cancer cells harbouring different KRAS mutations in codon 12, 13, 61 and 146 using human SW48 isogenic cell lines. *Metabolomics* 2020; **16**: 51 [PMID: 32300895 DOI: 10.1007/s11306-020-01674-2]

41 **Phull MS**, Jadav SS, Gundla R, Mainkar PS. A perspective on medicinal chemistry approaches towards adenomatous polyposis coli and Wnt signal based colorectal cancer inhibitors. *Eur J Med Chem* 2021; **212**: 113149 [PMID: 33445154 DOI: 10.1016/j.ejmech.2020.113149]

42 **Caspi M**, Wittenstein A, Kazelnik M, Shor-Nareznoy Y, Rosin-Arbesfeld R. Therapeutic targeting of the oncogenic Wnt signaling pathway for treating colorectal cancer and other colonic disorders. *Adv Drug Deliv Rev* 2021; **169**: 118-136 [PMID: 33346022 DOI: 10.1016/j.addr.2020.12.010]

43 **Aitchison A**, Hakkaart C, Day RC, Morrin HR, Frizelle FA, Keenan JI. *APC* Mutations Are Not Confined to Hotspot Regions in Early-Onset Colorectal Cancer. *Cancers (Basel)* 2020; **12** [PMID: 33352971 DOI: 10.3390/cancers12123829]

44 **Luo J**. Mutation analysis of sporadic colorectal cancer and polypby next-generation of sequencing. M.M. Thesis, Fujian Medical University. 2015. [cited 10 January 2022]. Available from: https://kns.cnki.net/kns8/defaultresult/index

Published by **Baishideng Publishing Group Inc**
7041 Koll Center Parkway, Suite 160, Pleasanton, CA 94566, USA
Telephone: +1-925-3991568
E-mail: bpgoffice@wjgnet.com
Help Desk: https://www.f6publishing.com/helpdesk
https://www.wjgnet.com