



OPEN ACCESS

EDITED BY

Pasquale Cianci,
Azienda Sanitaria Localedella Provincia di
Barletta Andri Trani (ASL BT), Italy

REVIEWED BY

Sara Ahmadi Badi,
Pasteur Institute of Iran, Iran
Vincenzo Lizzi,
Azienda Ospedaliero-Universitaria Ospedali
Riuniti di Foggia, Italy

*CORRESPONDENCE

Yousheng Li
✉ guttx@hotmail.com

†These authors have contributed equally to
this work

SPECIALTY SECTION

This article was submitted to
Surgical Oncology,
a section of the journal
Frontiers in Oncology

RECEIVED 23 October 2022

ACCEPTED 16 February 2023

PUBLISHED 01 March 2023

CITATION

Yu C, Zhou Z, Liu B, Yao D, Huang Y,
Wang P and Li Y (2023) Investigation of
trends in gut microbiome associated with
colorectal cancer using machine learning.
Front. Oncol. 13:1077922.
doi: 10.3389/fonc.2023.1077922

COPYRIGHT

© 2023 Yu, Zhou, Liu, Yao, Huang, Wang
and Li. This is an open-access article
distributed under the terms of the [Creative
Commons Attribution License \(CC BY\)](#). The
use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Investigation of trends in gut microbiome associated with colorectal cancer using machine learning

Chaoran Yu[†], Zhiyuan Zhou[†], Bin Liu[†], Danhua Yao,
Yuhua Huang, Pengfei Wang and Yousheng Li*

Department of General Surgery, Shanghai Ninth People's Hospital, Shanghai Jiao Tong University
School of Medicine, Shanghai, China

Background: The rapid growth of publications on the gut microbiome and colorectal cancer (CRC) makes it feasible for text mining and bibliometric analysis.

Methods: Publications were retrieved from the Web of Science. Bioinformatics analysis was performed, and a machine learning-based Latent Dirichlet Allocation (LDA) model was used to identify the subfield research topics.

Results: A total of 5,696 publications related to the gut microbiome and CRC were retrieved from the Web of Science Core Collection from 2000 to 2022. China and the USA were the most productive countries. The top 25 references, institutions, and authors with the strongest citation bursts were identified. Abstracts from all 5,696 publications were extracted for a text mining analysis that identified the top 50 topics in this field with increasing interest. The colitis animal model, expression of cytokines, microbiome sequencing and 16s, microbiome composition and dysbiosis, and cell growth inhibition were increasingly noticed during the last two years. The 50 most intensively investigated topics were identified and further categorized into four clusters, including "microbiome sequencing and tumor," "microbiome compositions, interactions, and treatment," "microbiome molecular features and mechanisms," and "microbiome and metabolism."

Conclusion: This bibliometric analysis explores the historical research tendencies in the gut microbiome and CRC and identifies specific topics of increasing interest. The developmental trajectory, along with the noticeable research topics characterized by this analysis, will contribute to the future direction of research in CRC and its clinical translation.

KEYWORDS

colorectal cancer, microbiome, bibliometric, Latent Dirichlet Allocation, Web of Science

Abbreviations: LDA, Latent Dirichlet Allocation; CRC, colorectal cancer; AI, artificial intelligence; SCB, strongest citation bursts.

Introduction

Much of what makes colorectal cancer (CRC) studies cutting-edge has already been contributed by classic oncological approaches. But certain elements were lacking until the gut microbiome supplied them. What the gut microbiome contributes to carcinogenesis and tumor progression may be familiar to other research topics, but what the gut microbiome achieves in microbes' contribution to this field is even more exceptional. A large population of microorganisms accommodated by the gut constantly interacts with intestinal epithelial cells and influences the metabolome and immunity throughout the entire gastrointestinal tract (1–5). Unbiased microbiome profiling and relevant models have revealed mechanistic insights into the microbial features associated with CRC (1). Remarkable progress has been achieved in studies relating to gut microbiota and CRC, highlighting the distinguished value of diagnosis and therapeutic prediction of the gut microbiome (6, 7). Bacterial strains such as *Fusobacterium nucleatum*, *Escherichia coli*, and *Bacteroides fragilis* are known for their tumor activities (8–10). However, there are many uncertainties concerning the association between gut microbiome and CRC, such as the huge number of bacteria (approximately 100 million) and distinct microbiota signatures, bacterial interactions, as well as geographical and race differences (11). Up to now, the gut microbiome remains far from fully deciphered.

Artificial intelligence (AI) has been one of the mainstays of cancer research and opens a plethora of technical applications (12). Driven by algorithms such as convolutional neural networks, a large quantity of data was used for training and pattern identification. It is commonly used to extract digital information from medical images for accurate medical diagnosis, such as MRI and PET/CT (13, 14). Meanwhile, mucosal visualization and polyp detection in gastrointestinal endoscopy and hematoxylin–eosin-stained images in pathological diagnosis can also be facilitated by AI (15, 16).

What occurred was so promising that, until recently, researchers began to discuss AI and the microbiome for colorectal cancer (17, 18). It is possible, however, to understand the value of the gut microbiome in CRC *via* an AI-dependent approach, and it is well worth while to do so. Based on megagenomic data and the antibiotic resistance genomic database, a DeepARG model was developed for accurate antimicrobial resistance annotation (19). Another example was a machine-learning-based decision tree model for prediction of cancer therapeutic responsiveness by gut microbiota composition and functional repertoire (20). Increasing volumes of metagenomics data and communications propel AI research intensity as well as platform-based data management and reusability (21).

Of note, AI is currently enriched in data-sensitive scenarios, but rarely covers text-sensitive scenarios of the gut microbiome. The rapid growth of publications on this research topic makes it feasible for text mining and bibliometric analysis. However, most analysis of research trends is mainly performed by literature reviews or meta statistics, with most word information untouched. Since a huge amount of wordy data over the decades could be a formidable task for manual processing, AI techniques are therefore translated for

unbiased interpretation. Latent Dirichlet Allocation (LDA) is one of the most powerful machine learning-based approaches to text mining (22–24). It aids in the topics and relationship findings among publications and data. Therefore, this study was performed to characterize the research topics in the field of gut microbiome associated with CRC over the past twenty years. Topic modeling by LDA enables us to pinpoint each research topic and provide a more in-depth interpretation. Thus, the results of this study provide the developmental trajectory of this field, understandable sub-fields, or topic connections, as well as how each research topic branches out and multidisciplinary integration.

Materials and methods

Publications in the field of gut microbiome with CRC were screened and retrieved *via* the Web of Science Core Collection (<https://clarivate.com/webofsciencelibrary/solutions/web-of-science-core-collection/>) with search terms covering both CRC and microbiome from 2000 to 30 June 2022. All included publications were then processed for bibliometric extraction and citation analysis. Analysis software included R (4.1.1 version), CiteSpace (5.8 R3), and Gephi (0.9.5 version) (25–29). To identify specific research topics derived from all publications with insight, LDA, a machine learning-based algorithm, was used for text mining (30, 31).

Results

A total of 5,696 publications related to the gut microbiome and CRC were retrieved from the Web of Science Core Collection from 2000 to 2022 (Figures 1A–C). The most productive countries included China, the USA, Italy, Japan, Germany, France, Korea, the UK, Spain, and India (Figure 1A). Although China published the most publications, the most multi-country publications were contributed by the USA. Meanwhile, a steady increase in annual publications was noticed (Figure 1B). Particularly since 2018, the increment of publications in each year has reached over 100.

To further demonstrate the most influential references, institutions, and authors, bibliometric analysis was performed on all 5,696 publications. The top 25 references with the strongest citation bursts (SCB) were identified from 2000 to 2022. The reference with the highest strength was published in 2012 by Arthur JC in SCIENCE, “Intestinal inflammation targets cancer-inducing activity of the microbiota” (Figure 2A). The top 25 institutions and authors with SCB were also demonstrated. Harvard University was the top institution with the highest strength (Figure 2B). Meanwhile, the French National Institute for Agricultural Research (INRA) displayed the longest and strongest citation burst period, ranging from 2001 to 2016. Jobin C was the top author with the highest strength (Figure 2C).

Next, to further characterize the changes in keywords, the annual counts of keywords, both author keywords and keywords plus, were calculated from 2000 to 2022 (Figures 3A, B). In the author keywords such as fatty acid, inflammation, probiotics, colitis, diet, immune,

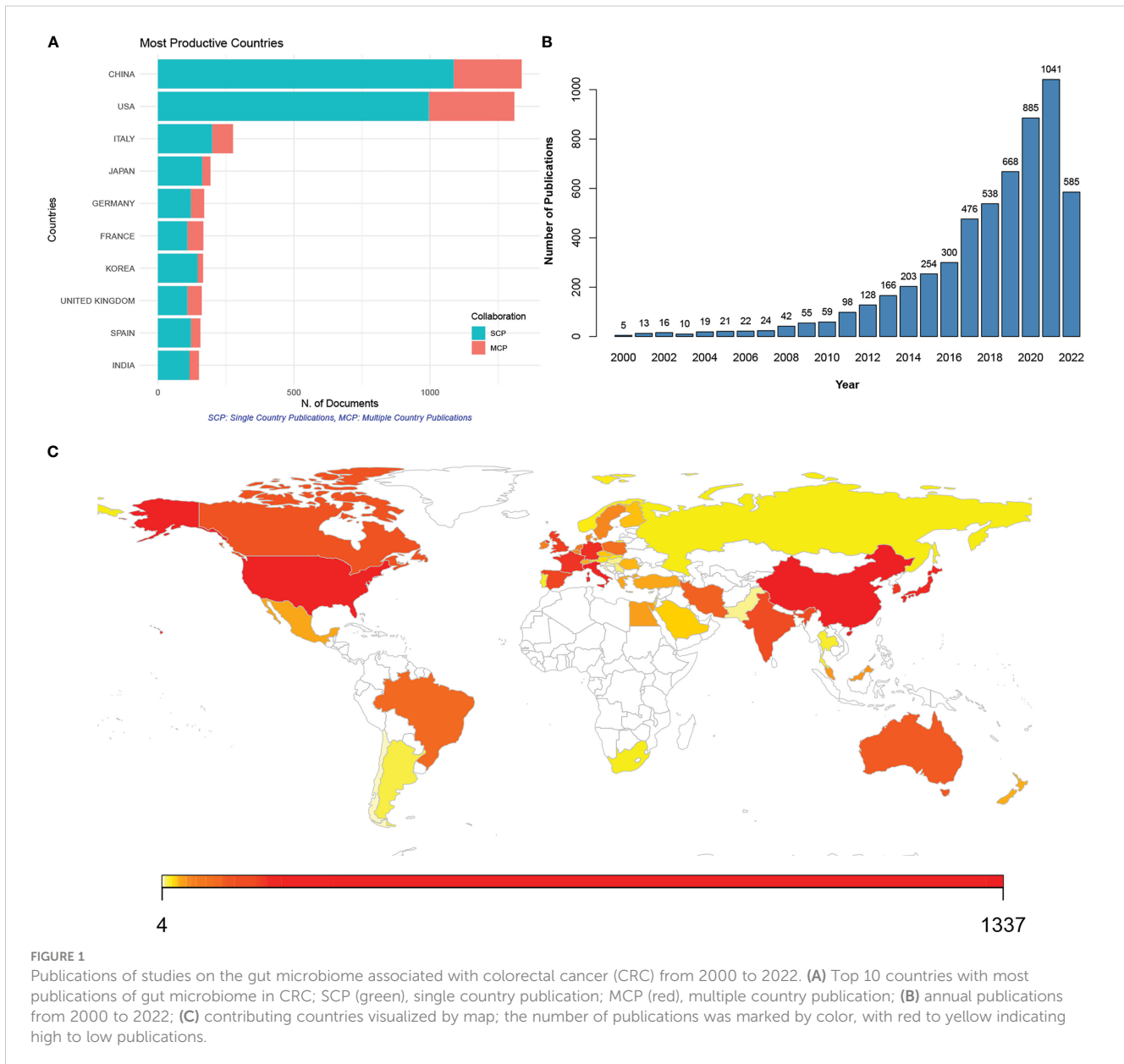


FIGURE 1 Publications of studies on the gut microbiome associated with colorectal cancer (CRC) from 2000 to 2022. (A) Top 10 countries with most publications of gut microbiome in CRC; SCP (green), single country publication; MCP (red), multiple country publication; (B) annual publications from 2000 to 2022; (C) contributing countries visualized by map; the number of publications was marked by color, with red to yellow indicating high to low publications.

metabolism, tumor, dysbiosis, and fusobacterium, biomarkers were among the top lists across the past twenty years. In keywords plus, not only similar words in author keywords were identified, but also keywords such as receptors, nucleatum, protein, and gene were added. In fact, the top key words can be categorized into four terms: microbiota composition (such as fusobacterium, nucleatum, and dysbiosis), microbiota and metabolism (such as metabolism, fatty acids, and protein), microbiota and treatment (such as prebiotics and probiotics, diet), and microbiota and disease course (such as tumor, biomarkers, activity, colitis, immune, infection, and risk).

To fully characterize the most investigated fields and to make research categorizations more precise within the gut microbiome and CRC, a LDA algorithm was employed. Abstracts from all 5,696 publications were extracted for a text-mining analysis. The results identified the top 50 topics in this field with increasing interest (Figure 4). In fact, topics such as colitis in mice, expression of cytokines, microbiome sequencing and 16s, gut microbiome composition and dysbiosis, and cell growth inhibition were dramatically increased during the last two years. Reasonably

presumed, increasing studies in those subfields have achieved remarkable progress relating to CRC and gut microbiota.

To further complement the categorization of keywords aforementioned, the 50 topics were also categorized into several groups for analysis and developmental management. A total of four clusters were determined and colored, including “microbiome sequencing and tumor,” “microbiome compositions, interactions, and treatment,” “microbiome molecular features and mechanisms,” and “microbiome and metabolism” (Figure 5). The cluster “microbiome sequencing and tumor” was established by topics including microbiome sequencing and 16s, bacterial species, fecal sample collection, prediction, and identification of various tumors. The cluster “microbiome compositions, interactions, and treatment” was established by topics including human and animal models, treatment effects, immune responses to cancer immunotherapy, and others. The cluster “microbiome molecular features and mechanisms” was established by topics including expression of cytokines, carcinogenesis and cancer promotion and inflammation, regulation of pathway signals, and others. The cluster “microbiome and

A

Top 25 References

| References | Year | Strength | Begin | End | 2000 - 2022 |
|--|------|----------|-------|------|-------------|
| Sobhani I, 2011, PLOS ONE, V6, P0, DOI 10.1371/journal.pone.0016393, DOI | 2011 | 53.58 | 2011 | 2016 | |
| Qin JJ, 2010, NATURE, V464, P59, DOI 10.1038/nature08821, DOI | 2010 | 46.45 | 2011 | 2015 | |
| Wu SG, 2009, NAT MED, V15, P1016, DOI 10.1038/nm.2015, DOI | 2009 | 45.01 | 2011 | 2014 | |
| Castellarin M, 2012, GENOME RES, V22, P299, DOI 10.1101/gr.126516.111, DOI | 2012 | 84 | 2012 | 2017 | |
| Kostic AD, 2012, GENOME RES, V22, P292, DOI 10.1101/gr.126573.111, DOI | 2012 | 81.16 | 2012 | 2017 | |
| Arumugam M, 2011, NATURE, V473, P174, DOI 10.1038/nature09944, DOI | 2011 | 45.1 | 2012 | 2016 | |
| Marchesi JR, 2011, PLOS ONE, V6, P0, DOI 10.1371/journal.pone.0020447, DOI | 2011 | 43.11 | 2012 | 2016 | |
| Arthur JC, 2012, SCIENCE, V338, P120, DOI 10.1126/science.1224820, DOI | 2012 | 83.46 | 2013 | 2017 | |
| Wang TT, 2012, ISME J, V6, P320, DOI 10.1038/ismej.2011.109, DOI | 2012 | 49.33 | 2013 | 2017 | |
| Tjalsma H, 2012, NAT REV MICROBIOL, V10, P575, DOI 10.1038/nrmicro2819, DOI | 2012 | 42.19 | 2013 | 2017 | |
| Grivnenkov SI, 2012, NATURE, V491, P254, DOI 10.1038/nature11465, DOI | 2012 | 39.26 | 2013 | 2017 | |
| Huttenhower C, 2012, NATURE, V486, P207, DOI 10.1038/nature11234, DOI | 2012 | 37.58 | 2013 | 2017 | |
| Kostic AD, 2013, CELL HOST MICROBE, V14, P207, DOI 10.1016/j.chom.2013.07.007, DOI | 2013 | 77.98 | 2014 | 2018 | |
| Rubinstein MR, 2013, CELL HOST MICROBE, V14, P195, DOI 10.1016/j.chom.2013.07.012, DOI | 2013 | 65.89 | 2014 | 2018 | |
| Schwabe RF, 2013, NAT REV CANCER, V13, P800, DOI 10.1038/nrc3610, DOI | 2013 | 45.66 | 2014 | 2018 | |
| Ahn J, 2013, JNCI-J NATL CANCER I, V105, P1907, DOI 10.1093/jnci/djt300, DOI | 2013 | 40.72 | 2014 | 2018 | |
| Zackular JP, 2013, MBIO, V4, P0, DOI 10.1128/mBio.00692-13, DOI | 2013 | 37.2 | 2014 | 2018 | |
| Weir TL, 2013, PLOS ONE, V8, P0, DOI 10.1371/journal.pone.0070803, DOI | 2013 | 35.09 | 2014 | 2018 | |
| Sears CL, 2014, CELL HOST MICROBE, V15, P317, DOI 10.1016/j.chom.2014.02.007, DOI | 2014 | 40.08 | 2015 | 2019 | |
| David LA, 2014, NATURE, V505, P559, DOI 10.1038/nature12820, DOI | 2014 | 35.7 | 2015 | 2019 | |
| Iida N, 2013, SCIENCE, V342, P967, DOI 10.1126/science.1240527, DOI | 2013 | 34.47 | 2015 | 2018 | |
| Louis P, 2014, NAT REV MICROBIOL, V12, P661, DOI 10.1038/nrmicro3344, DOI | 2014 | 56.27 | 2016 | 2019 | |
| Zeller G, 2014, MOL SYST BIOL, V10, P0, DOI 10.15252/msb.20145645, DOI | 2014 | 40.72 | 2016 | 2019 | |
| Sivan A, 2015, SCIENCE, V350, P1084, DOI 10.1126/science.aac4255, DOI | 2015 | 35.5 | 2017 | 2020 | |
| Bray F, 2018, CA-CANCER J CLIN, V68, P394, DOI 10.3322/caac.21492, DOI | 2018 | 54.6 | 2020 | 2022 | |

B

Top 25 Institutions

| Institutions | Strength | Begin | End | 2000 - 2022 |
|--|----------|-------|------|-------------|
| German Inst Human Nutr Potsdam Rehbrucke | 4.4 | 2000 | 2011 | |
| INRA | 12.56 | 2001 | 2016 | |
| Natl Univ Ireland Univ Coll Cork | 6.03 | 2005 | 2014 | |
| Univ Reading | 6.76 | 2007 | 2014 | |
| Univ N Carolina | 12.95 | 2009 | 2015 | |
| CNRS | 4.84 | 2010 | 2016 | |
| Univ Chicago | 7.28 | 2011 | 2018 | |
| Harvard Univ | 15.18 | 2012 | 2016 | |
| Univ Florida | 9.66 | 2012 | 2017 | |
| Wageningen Univ | 8.55 | 2012 | 2015 | |
| INSERM | 6 | 2012 | 2016 | |
| Univ Aberdeen | 8.4 | 2013 | 2016 | |
| Univ Toronto | 6.49 | 2013 | 2017 | |
| Yale Univ | 5.7 | 2013 | 2017 | |
| Univ Calif Los Angeles | 4.38 | 2013 | 2017 | |
| Monash Univ | 4.46 | 2014 | 2016 | |
| Univ Penn | 4.59 | 2015 | 2017 | |
| Kyung Hee Univ | 4.37 | 2015 | 2016 | |
| Univ Calif Davis | 4.58 | 2017 | 2018 | |
| Univ Minnesota | 5.28 | 2018 | 2019 | |
| Univ South Carolina | 4.76 | 2018 | 2019 | |
| Harbin Med Univ | 5.44 | 2019 | 2020 | |
| Zhejiang Univ | 4.56 | 2019 | 2020 | |
| Huazhong Univ Sci & Technol | 4.56 | 2020 | 2022 | |
| Wuhan Univ | 4.51 | 2020 | 2022 | |

C

Top 25 Authors

| Authors | Strength | Begin | End | 2000 - 2022 |
|------------------|----------|-------|------|-------------|
| BLAUT M | 4.39 | 2000 | 2011 | |
| GASKINS H | 5 | 2004 | 2012 | |
| O'KEEFE S | 6.98 | 2008 | 2014 | |
| MARCHESI J | 5.53 | 2008 | 2017 | |
| JOBIN C | 11.17 | 2009 | 2017 | |
| TOMAS-BARBERAN F | 4.69 | 2009 | 2017 | |
| KEKU T | 6.16 | 2010 | 2014 | |
| AHN J | 4.47 | 2012 | 2018 | |
| FUCHS C | 6.38 | 2013 | 2018 | |
| SEARS C | 5.69 | 2014 | 2016 | |
| CHAN A | 4.8 | 2014 | 2018 | |
| HOUSSEAU F | 4.46 | 2014 | 2018 | |
| YANG Y | 6.09 | 2015 | 2017 | |
| KIM D | 5.6 | 2015 | 2016 | |
| YU Y | 5.26 | 2015 | 2019 | |
| OGINO S | 5.26 | 2016 | 2018 | |
| WU J | 4.71 | 2016 | 2019 | |
| KINROSS J | 5.41 | 2017 | 2018 | |
| CHAN F | 5.08 | 2017 | 2019 | |
| SUNG J | 4.95 | 2017 | 2020 | |
| FU X | 4.59 | 2018 | 2019 | |
| ZHANG G | 4.55 | 2018 | 2020 | |
| WANG Q | 5.01 | 2019 | 2020 | |
| SUN T | 4.67 | 2019 | 2020 | |
| ZHANG S | 4.41 | 2019 | 2022 | |

FIGURE 2

Top-listed references, institutions, and authors with the strongest citation bursts (SCB). (A) Top 25 SCB references; (B) top 25 SCB institutions; (C) top 25 SCB authors; red bar to green bar indicates high frequent occurrence of citation period compared to common frequent citation occurrence.

metabolism” was established by topics including dietary intake and consumption, bile acids and butyrate, activities of oxidative stress, and others. Based on those categorizations, several study areas for future studies have also been listed (Table 1), including four major areas: investigation models, microbiome sequence techniques, clinical trials, and microbiota metabolism.

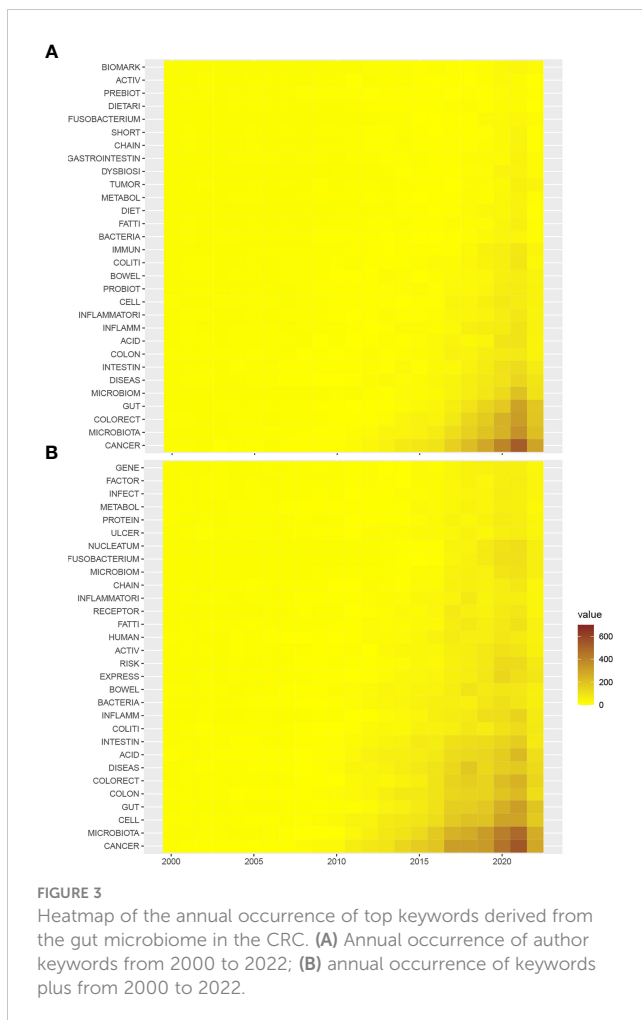
To visualize the citation pattern, a dual-map thematic overlay portfolio analysis was performed. The discipline distribution of publications associated with the gut microbiome and CRC was represented (Figure 6).

Discussion

Up to 2021, the publication of gut microbiome in CRC has significantly increased by 200 times compared to 2000. A faster

growth pace was found from 2015 to 2021, with significant increments in 2017 and 2020. The 50 predominant research topics identified in this study were clustered into four, including “microbiome sequencing and tumors,” “microbiome compositions, interactions, and treatment,” “microbiome molecular features and mechanisms,” and “microbiome and metabolism.”

Regarding microbiome sequencing and tumors, research is intensely focused on *F. nucleatum* in tissue, disease development and progress, microbiome sequencing and 16S, bacterial species, CRC, and colorectal adenoma. *F. nucleatum* had been reported to be closely associated with tumor subtypes as it induced diverse immune responses with respect to the microsatellite instability status of CRC (32). Interestingly, high levels of *F. nucleatum* improved the overall anti-tumor effects of PD-L1 blockade therapy with prolonged survival. Microbiome sequencing is another key topic identified in this analysis. Crucial alterations in



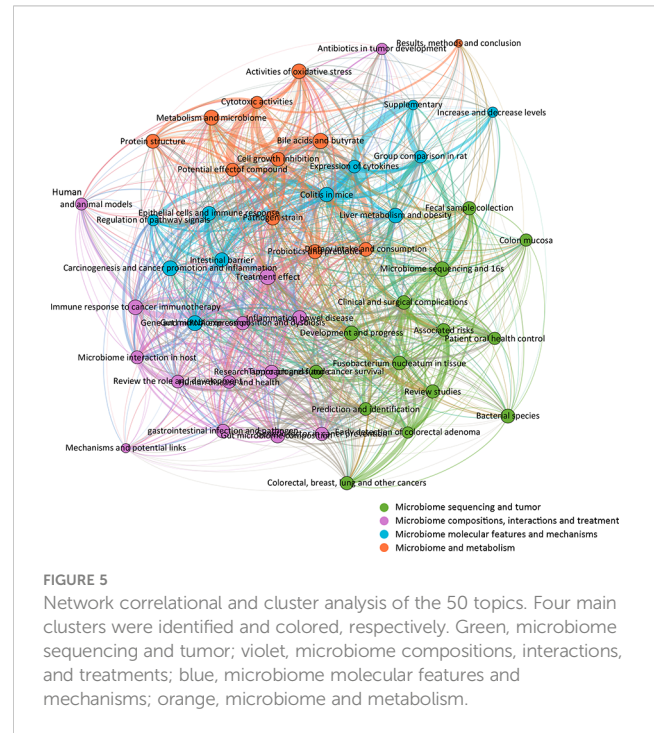
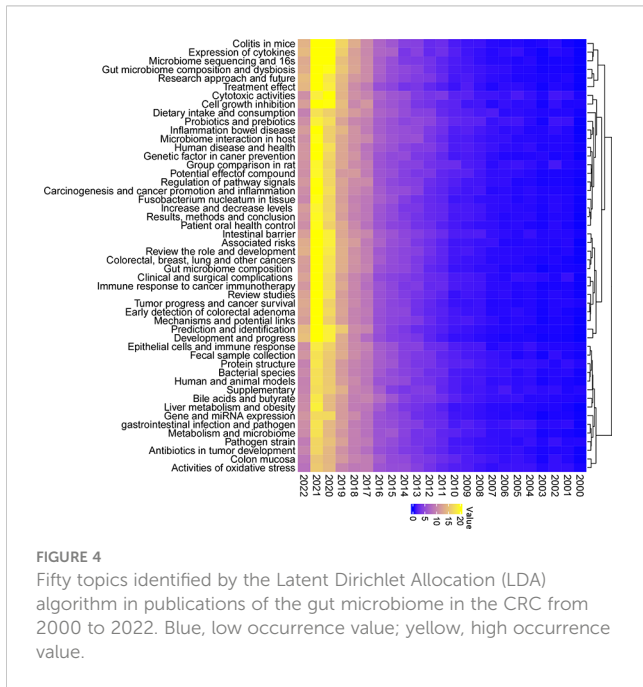
carcinogenesis and treatment effects could be monitored by microbiota metagenomic profiling (33). 16S rRNA sequencing techniques have contributed to the identification of several key bacterial strains, including *F. nucleatum*, *E. coli*, *Streptococcus intermedius*, *Gemella haemolysans*, and *B. fragilis* (34). Another 16S rRNA amplicon sequencing study from Korea observed three phyla were less found in tumor tissues, along with intensely enriched metabolic pathways of the bile acid section or bacterial motility proteins related to CRC (35). In addition, *F. nucleatum* and *B. fragilis* were also found more abundant in recurrent individuals compared to nonrecurrent ones (35). Besides, based on the keywords heatmap (Figure 3), *F. nucleatum*, a gram-negative, anaerobic opportunistic bacterium, is most likely to be associated with CRC. This kind of bacterium is common to the oral cavity and is associated with periodontal disease. Interestingly, one of the topics identified by LDA algorithms is “patient oral health control,” covering studies of the association between oral health and CRC risk. Accumulating evidence has also concluded that mucosa-associated *Escherichia coli* (*E. coli*) is involved with the tumorigenesis and progression of CRC, particularly some strains of *E. coli*, including enteropathogenic *E. coli* (EPEC) and cyclomodulin-positive *E. coli* (13). Based on the topic terms, both

E. coli and *Fusobacterium* are the main gut microbiota members associated with CRC.

Regarding microbiome compositions, interactions, and treatment, research is more likely to focus on human and animal models, treatment effects, inflammation bowel disease, immune responses to cancer immunotherapy, microbiome interactions in the host, microbiome dysbiosis, and infections. In fact, increasing experimental models have been developed to support rising research on microbiota in cancer studies, including a germ-free humanized microbiota sample transfer model and an antibiotic regiment-based animal model (Table 2). Effective animal models serve as an essential component of preclinical studies and critical evidence for mechanistic insights (36). Specific functional proteins such as FadA and Fap2 have been identified through successful experimental models (37–39). Interestingly, systemic, and mucosal immune status remains largely stable in the antibiotic treatment-based animal model, thereby making it a potential approach for immunotherapy evaluation. Soon, humanized microbiota animal models will be crucial to immunotherapy research (40–43) (Table 2). By far, studies are only beginning to scratch the surface of the nature of the association between the microbiome and CRC. It is reasonable to presume that microbiota-related treatment, immune response, and animal models will be some of the key developments in the future.

Regarding microbiome molecular features and mechanisms, research topics are focused on regulation of pathway signals, expression of cytokines, colitis in mice, intestinal barrier, and gene and miRNA expression. Several pathway signals relating to tumorigenesis have been involved in the gut microbiota, including the toll-like receptors (TLRs), KRAS, NF-kappa B, SARS-CoV-2, G protein-coupled receptors, and Wnt pathways (44–49). Chronic inflammation was taken as a major cause of tumorigenesis and progression. Therefore, the colitis model has been an opportunity to reveal the contribution of microbiota to colitis-associated CRC (50).

Regarding the microbiome and metabolism, topics include activities of oxidative stress, metabolism, bile acids and butyrate, cell growth inhibition and potential effects of compounds, probiotics and prebiotics, dietary intake, and consumption. From a metabolic point of view, the microbiome and metabolism demonstrated a dramatic research market with huge potential. Bile acid–gut microbiome interaction constitutes one of the most intensively investigated topics, with increasing publications over the years (51–54). The metabolism of bile acids and their symbiosis was commonly associated with a low-fiber diet. Moreover, researchers indicated bidirectional regulatory effects of bile acids on CRC along with its progression (55). Butyrate, one of the main short-chain fatty acids, serves as an effector for anti-inflammation and anti-tumor (56). Increasing levels of omega-3 polyunsaturated fatty acids may promote the bacteria that produce butyrate, lowering the risk of CRC (57). Although the overall trend in the gut microbiome has been positive, some specific topics indicated that research processes were lagging, including dietary intake and consumption, pathogen strains, and activities of oxidative stress, all of which belong to this cluster.



The most influential reference was identified as “Intestinal inflammation targets cancer-inducing activity of the microbiota” by Arthur et al., published in SCIENCE 2012. In this study, Arthur et al. reported that *Escherichia coli* NC101 (*E. coli*) was significantly enriched in inflammatory bowel disease and CRC mucosa. Intestinal microbiota was identified as a key target of intestinal inflammation, further affecting the disease course of CRC (58). This study for the first time answered the critical question of whether the gut microbiota was actively involved in the progress of carcinogenesis.

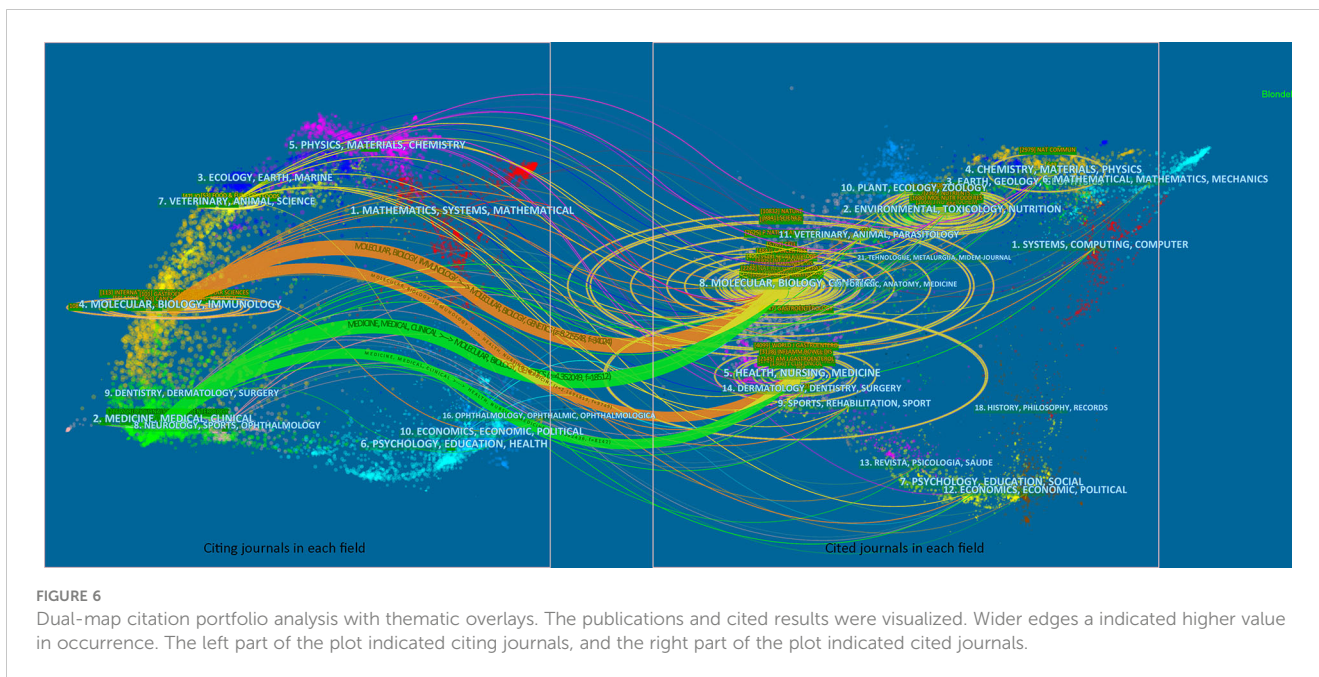
Among all the leading institutions, INRA showed the longest period in SCB. As one of the top research institutions in France, INRA has made a considerable contribution to the field of gut microbiota. To achieve deep knowledge of health-related challenges affected by the gut microbiota, INRA also launched a nationwide collaborative project called Le French Gut on 15 September 2022, continuing to support the “Million Microbiome of Humans

Project,” an international project to build up a world-class human microbiota database for public use.

Of all the 50 topics, some showed remarkable progress during the last few years, for example, the role of colitis in a mouse model associated with gut microbiota and CRC. A few noticeable achievements had been made by several studies (59, 60). Zhu et al. reported that by precisely editing the microbiota composition in mouse models, the risk of tumor development in colitis-associated CRC could be reduced (59). Particularly, Enterobacteriaceae was identified as the key target in this intestinal inflammation course. By using a similar colitis cancer mouse model, Ji et al. highlighted the modulatory role of jujube polysaccharides in ameliorating colitis-related cancer and microbiota dysbiosis. Firmicutes and Bacteroidetes were also significantly reduced in this model (60).

TABLE 1 Gut microbiome research for future guidance.

| Study area | Research topics |
|-------------------------------------|--|
| Human or animal models | <ol style="list-style-type: none"> 1. Optimize the humanized microbiota-associated animal models 2. Establish animal model study on response to immunotherapy 3. Improving imaging techniques, such as animal endoscopy or MRI or tumor-specific markers |
| Microbiome sequencing | <ol style="list-style-type: none"> 1. Machine learning application in next generation sequencing for gut microbiome 2. Microbiota dysbiosis and causes 3. Sample collections for sequencing |
| Clinical trials with gut microbiota | <ol style="list-style-type: none"> 1. Specify the contribution of fecal microbiota transplantation (FMT) 2. Specify the association between FMT and immunotherapy 3. Application of microbial ecosystem therapeutics and its mechanisms 4. Association between probiotics and microbiota |
| Microbiota metabolism | <ol style="list-style-type: none"> 1. Gut microbiota-derived short-chain fatty acids 2. Diet researches and tumorigenesis 3. Chemical-microbiota interaction for precancerous study |



Colitis-associated colorectal cancer (CAC) is a critical complication of inflammatory bowel disease, accounting for around 15% of mortality. However, the molecular mechanisms underlying the carcinogenesis of CAC remain largely unclear and are deemed different from other types of CRC (61). In fact, previous clinical and experimental clues have indicated that inflammation serves as a key to initiating CAC, while it may not be a decisive trigger for common CRC. Thus, investigation into the role of the gut microbiome in CRC and CAC is highly valuable. Research progress relating to gut microbiota and CAC is mostly represented by two types of studies. Type I is the mechanistic interaction between members of the microbial community and the intestinal tract. Microbial community members produce and release genotoxins in the gut to exasperate inflammation or induce carcinogenesis. Type II is the well-established animal model used for CAC. Of note, there is a high-value unanswered question: could an inflammatory bowel disease-driven microbe be able to automatically promote

CAC and, in a broader sense, CRC? In fact, there are several types of microorganisms that are both linked to the disease courses of inflammatory bowel disease and CRC, including *Fusobacterium* species and *Streptococcus bovis*. Other types of microorganisms, such as the enterotoxigenic strain of *B. fragilis*, are only associated with CRC. Up until now, a dispute remains as to whether microbes are vital to the carcinogenesis of CAC and other types of CRC or just innocent bystanders. Animal models of CAC may be the key to that question.

Conclusion

This bibliometric analysis explores the historical research trends in gut microbiome and CRC and identifies specific topics with increasing interests. The developmental trajectory, along with the noticeable research topics characterized by this analysis, will contribute to the future direction of research in CRC and clinical translation.

TABLE 2 Humanized microbiota-associated animal models.

| Animal models | Stability | Features |
|---|---------------------|--|
| Human microbiota samples transferring to germ-free animal model | Short term | <ol style="list-style-type: none"> 1. Germ-free condition with specialized equipment 2. Expensive maintenance 3. Experimental genotyping 4. Developmental defects 5. Targeting specified microbes |
| Antibiotics treatment-based animal model | Potential long-term | <ol style="list-style-type: none"> 1. Cost effective 2. Feasible to alternative types of antibiotics 3. Applicable to multiple genotypes 4. Bacterial or other microorganism bias to antibiotics treatment 5. Host health concerns in response to antibiotics treatment |

Data availability statement

The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to the corresponding author.

Author contributions

CY, ZZ, BL, DY, YH, and PW carried out literature and data analysis. CY, ZZ, BL, DY, YH, and YL drafted the manuscript. CY, ZZ, BL, DY, YH, PW, and YL participated in study design and data collection. All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Funding

This work was funded by the Construction and application of biobank of Crohn's disease in Chinese Research Hospital

References

- Clay SL, Fonseca-Pereira D, Garrett WS. Colorectal cancer: the facts in the case of the microbiota. *J Clin Invest* (2022) 132(4):e155101. doi: 10.1172/JCI155101
- Wong SH, Yu J. Gut microbiota in colorectal cancer: mechanisms of action and clinical applications. *Nat Rev Gastroenterol Hepatol* (2019) 16(11):690–704. doi: 10.1038/s41575-019-0209-8
- Sánchez-Alcoholado L, Ramos-Molina B, Otero A, Laborda-Illanes A, Ordóñez R, Medina JA, et al. The role of the gut microbiome in colorectal cancer development and therapy response. *Cancers (Basel)* (2020) 12(6):1406. doi: 10.3390/cancers12061406
- Rebersek M. Gut microbiome and its role in colorectal cancer. *BMC Cancer* (2021) 21(1):1325. doi: 10.1186/s12885-021-09054-2
- Privitera G, Rana N, Scaldaferrri F, Armuzzi A, Pizarro TT. Novel insights into the interactions between the gut microbiome, inflammasomes, and gasdermins during colorectal cancer. *Front Cell Infect Microbiol* (2022) 11:806680. doi: 10.3389/fcimb.2021.806680
- Liu J, Qi M, Qiu C, Wang F, Xie S, Zhao J, et al. Integrative analysis of the mouse fecal microbiome and metabolome reveal dynamic phenotypes in the development of colorectal cancer. *Front Microbiol* (2022) 13:1021325. doi: 10.3389/fmicb.2022.1021325
- Cavallucci V, Palucci I, Fidaleo M, Mercuri A, Masi L, Emoli V, et al. Proinflammatory and cancer-promoting pathobiont fusobacterium nucleatum directly targets colorectal cancer stem cells. *Biomolecules* (2022) 12(9):1256. doi: 10.3390/biom12091256
- Mesa F, Mesa-López MJ, Egea-Valenzuela J, Benavides-Reyes C, Nibali L, Ide M, et al. A new comorbidity in periodontitis: Fusobacterium nucleatum and colorectal cancer. *Medicina* (2022) 58(4):546. doi: 10.3390/medicina58040546
- Veziat J, Villéger R, Barnich N, Bonnet M. Gut microbiota as potential biomarker and/or therapeutic target to improve the management of cancer: focus on colibactin-producing *Escherichia coli* in colorectal cancer[J]. *Cancers* (2021) 13(9):2215. doi: 10.3390/cancers13092215
- Scott N, Whittle E, Jeraldo P, Chia N. A systemic review of the role of enterotoxigenic *Bacteroides fragilis* in colorectal cancer. *Neoplasia* (2022) 29:100797. doi: 10.1016/j.neo.2022.100797
- Ahmad S, Ashktorab H, Brim H, Housseau F. Inflammation, microbiome and colorectal cancer disparity in African-americans: Are there bugs in the genetics? *World J Gastroenterol* (2022) 28(25):2782–801. doi: 10.3748/wjg.v28.i25.2782
- Yu C, Helwig EJ. The role of AI technology in prediction, diagnosis and treatment of colorectal cancer. *Artif Intell Rev* (2022) 55(1):323–43. doi: 10.1007/s10462-021-10034-y
- Wong C, Fu Y, Li M, Mu S, Chu X, Fu J, et al. MRI-Based artificial intelligence in rectal cancer. *J Magnetic Resonance Imaging* (2022) 57(1):45–56. doi: 10.1002/jmri.28381
- Association (YBKB201916) and the Fundamental Research Program Funding of Ninth People's Hospital Affiliated to Shanghai Jiao Tong University School of Medicine (JYZZ189).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher. Abbreviation

27. Perianes-Rodriguez A, Waltman L, Van Eck NJ. Constructing bibliometric networks: A comparison between full and fractional counting. *J Informetrics* (2016) 10(4):1178–95. doi: 10.1016/j.joi.2016.10.006
28. Chen C, Song M. Visualizing a field of research: A methodology of systematic scientometric reviews. *PLoS One* (2019) 14(10):e0223994. doi: 10.1371/journal.pone.0223994
29. Chen C, Leydesdorff L. Patterns of connections and movements in dual-map overlays: A new method of publication portfolio analysis. *J Assoc Inf Sci Technol* (2014) 65(2):334–51. doi: 10.1002/asi.22968
30. Jelodar H, Wang Y, Yuan C, Feng X, Jiang X, Li Y, et al. Latent dirichlet allocation (LDA) and topic modeling: models, applications, a survey. *Multimedia Tools Appl* (2019) 78(11):15169–211. doi: 10.1007/s11042-018-6894-4
31. Hoffman M, Bach F, Blei D. Online learning for latent dirichlet allocation. *Adv Neural Inf Process Syst* (2010) 23. doi: 10.5555/2997189.2997285
32. Gao Y, Bi D, Xie R, et al. Fusobacterium nucleatum enhances the efficacy of PD-L1 blockade in colorectal cancer. *Signal Transduction Targeted Ther* (2021) 6(1):1–10. doi: 10.1038/s41392-021-00795-x
33. Young C, Wood HM, Seshadri RA, Van Nang P, Vaccaro C, Melendez LC, et al. The colorectal cancer-associated faecal microbiome of developing countries resembles that of developed countries. *Genome Med* (2021) 13(1):1–13. doi: 10.1186/s13073-021-00844-8
34. Nardelli C, Granata I, Nunziato M, Setaro M, Carbone F, Zulli C, et al. 16S rRNA of mucosal colon microbiome and CCL2 circulating levels are potential biomarkers in colorectal cancer. *Int J Mol Sci* (2021) 22(19):10747. doi: 10.3390/ijms221910747
35. Choi S, Chung J, Cho ML, Park D, Choi SS. Analysis of changes in microbiome compositions related to the prognosis of colorectal cancer patients based on tissue-derived 16S rRNA sequences. *J Trans Med* (2021) 19(1):1–12. doi: 10.1186/s12967-021-03154-0
36. Staley C, Kaiser T, Beura LK, Hamilton MJ, Weingarden AR, Bobr A, et al. Stable engraftment of human microbiota into mice with a single oral gavage following antibiotic conditioning. *Microbiome* (2017) 5(1):87. doi: 10.1186/s40168-017-0306-2
37. Abed J, Emgård JE, Zamir G, Faroja M, Almogy G, Grenov A, et al. Fap2 mediates fusobacterium nucleatum colorectal adenocarcinoma enrichment by binding to tumor-expressed gal-GalNAc. *Cell Host Microbe* (2016) 20(2):215–25. doi: 10.1016/j.chom.2016.07.006
38. Rubinstein MR, Wang X, Liu W, Hao Y, Cai G, Han YW. Fusobacterium nucleatum promotes colorectal carcinogenesis by modulating e-cadherin/ β -catenin signaling via its FadA adhesin. *Cell Host Microbe* (2013) 14(2):195–206. doi: 10.1016/j.chom.2013.07.012
39. Yuan C, Zhao X, Wangmo D, Alshareef D, Gates TJ, Subramanian S, et al. Tumor models to assess immune response and tumor-microbiome interactions in colorectal cancer. *Pharmacol Ther* (2021) 231:107981. doi: 10.1016/j.pharmthera.2021.107981
40. Zhang C, Franklin CL, Ericsson AC. Consideration of gut microbiome in murine models of diseases. *Microorganisms* (2021) 9(5):1062. doi: 10.3390/microorganisms9051062
41. Chen G, Cao Z, Shi Z, Lei H, Chen C, Yuan P, et al. Microbiome analysis combined with targeted metabolomics reveal immunological anti-tumor activity of icaricide I in a melanoma mouse model[J]. *Biomedicine pharmacotherapy* (2021) 140:111542. doi: 10.1016/j.biopha.2021.111542
42. Shen L, Ye Y, Sun H, Su B. ILC3 plasticity in microbiome-mediated tumor progression and immunotherapy. *Cancer Cell* (2021) 39(10):1308–10. doi: 10.1016/j.ccell.2021.08.002
43. Kalasabail S, Engelman J, Zhang LY, El-Omar E, Yim HCH. A perspective on the role of microbiome for colorectal cancer treatment. *Cancers* (2021) 13(18):4623. doi: 10.3390/cancers13184623
44. Sheikh A, Taube J, Greathouse KL. Contribution of the microbiota and their secretory products to inflammation and colorectal cancer pathogenesis: the role of toll-like receptors. *Carcinogenesis* (2021) 42(9):1133–42. doi: 10.1093/carcin/bgab060
45. Ternet C, Kiel C. Signaling pathways in intestinal homeostasis and colorectal cancer: KRAS at centre stage. *Cell Communication Signaling* (2021) 19(1):1–22. doi: 10.1186/s12964-021-00712-3
46. Karpiński TM, Ożarowski M, Stasiewicz M. *Carcinogenic microbiota and its role in colorectal cancer development*[C]//Seminars in cancer biology. Academic Press (2022)
47. Howell MC, Green R, McGill AR, Dutta R, Mohapatra S, Mohapatra SS, et al. SARS-CoV-2-induced gut microbiome dysbiosis: Implications for colorectal cancer. *Cancers* (2021) 13(11):2676. doi: 10.3390/cancers13112676
48. Peng Y, Nie Y, Yu J, Wong CC. Microbial metabolites in colorectal cancer: basic and clinical implications. *Metabolites* (2021) 11(3):159. doi: 10.3390/metabo11030159
49. Priya S, Burns MB, Ward T, Mars RAT, Adamowicz B, Lock EF, et al. Identification of shared and disease-specific host gene–microbiome associations across human diseases using multi-omic integration. *Nat Microbiol* (2022), 1–16. doi: 10.1038/s41564-022-01121-z
50. Lin Q, Liu M, Yue GGL, Cheung MK, Lai Z, Kwok FH, et al. Anti-inflammatory activities of natural cyclopeptide RA-XII in colitis-associated colon cancer mouse model and its effect on gut microbiome. *Phytotherapy Res* (2022) 36(6):2641–59. doi: 10.1002/ptr.7482
51. Roberts DC, Chidambaram S, Kinross JM. The role of the colonic microbiota and bile acids in colorectal cancer. *Curr Opin Gastroenterol* (2022) 38(2):179–88. doi: 10.1097/MOG.0000000000000816
52. Ocvirk S, O’Keefe SJD. Dietary fat, bile acid metabolism and colorectal cancer [C]//Seminars in cancer biology. *Acad Press* (2021) 73:347–55. doi: 10.1016/j.semcancer.2020.10.003.
53. Lavelle A, Nancey S, Reimund JM, Laharie D, Marteau P, Treton X, et al. Fecal microbiota and bile acids in IBD patients undergoing screening for colorectal cancer. *Gut Microbes* (2022) 14(1):2078620. doi: 10.1080/19490976.2022.2078620
54. Wu L, Feng J, Li J, Yu Q, Ji J, Wu J, et al. The gut microbiome-bile acid axis in hepatocarcinogenesis. *Biomedicine Pharmacotherapy* (2021) 133:111036. doi: 10.1016/j.biopha.2020.111036
55. Yu S, Shao X, Zhou Y, Yu Y, Kuai X, Zhou C. Bidirectional regulation of bile acid on colorectal cancer through bile acid-gut microbiota interaction. *Am J Transl Res* (2021) 13(10):10994–1003
56. Gheorghe AS, Negru ŞM, Preda M, Mihăilă RI, Komporaly IA, Dumitrescu EA, et al. Biochemical and metabolic pathways associated with microbiota-derived butyrate in colorectal cancer and omega-3 fatty acids implications: A narrative review. *Nutrients* (2022) 14(6):1152. doi: 10.3390/nu14061152
57. Fu Y, Wang Y, Gao H, Li D, Jiang R, Ge L, et al. Associations among dietary omega-3 polyunsaturated fatty acids, the gut microbiota, and intestinal immunity. *Mediat Inflamm* (2021) 2021:8879227. doi: 10.1155/2021/8879227
58. Arthur JC, Perez-Chanona E, Mühlbauer M, Tomkovich S, Uronis JM, Fan TJ, et al. Intestinal inflammation targets cancer-inducing activity of the microbiota. *Science* (2012) 338(6103):120–3. doi: 10.1126/science.1224820
59. Zhu W, Miyata N, Winter MG, Arenales A, Hughes ER, Spiga L, et al. Editing of the gut microbiota reduces carcinogenesis in mouse models of colitis-associated colorectal cancer. *J Exp Med* (2019) 216(10):2378–93. doi: 10.1084/jem.20181939
60. Ji X, Hou C, Gao Y, Xue Y, Yan Y, Guo X, et al. Metagenomic analysis of gut microbiota modulatory effects of jujube (*Ziziphus jujuba* mill.) polysaccharides in a colorectal cancer mouse model. *Food Funct* (2020) 11(1):163–73. doi: 10.1039/c9fo02171j
61. Grivennikov SI. Inflammation and colorectal cancer: colitis-associated neoplasia. *Semin Immunopathol* (2013) 35(2):229–44. doi: 10.1007/s00281-012-0352-6