

REVIEW OPEN



Significance and challenges in dissecting cancer-bacteriome interactions

Ibraheem Alshareedah¹, James D. Brunner², Patrick S. G. Chain² and Anand Kumar¹✉

This is a U.S. Government work and not under copyright protection in the US; foreign copyright protection may apply 2026, modified publication 2026

Cancer is the leading cause of death around the world. While some types of cancer have become manageable due to advancements in medicine, most cancers still lack available cures and treatments. Recent studies have shown that changes in the human microbiome, especially in the bacteriome, are associated with some cancers. Certain bacterial strains have been reported to promote the initiation and progression of cancer in humans. Other studies have used sequencing to observe changes in the bacteriome of healthy and cancer patients. However, studies that investigate the interactions between cancer cells and the complex bacteriome as a whole remain scarce. This is due to the absence of experimental methods to study the interactions between cancer cells and complex bacterial populations, which has delayed the progress in identifying cancer-causing and cancer-inhibiting bacteria, and in understanding the bacterial interactions and their influence on host cells. Here, we review approaches to studying cancer cell interactions with complex bacteriomes and suggest possible routes to overcome this problem, highlighting the need for interdisciplinary studies that may help advance this field. We speculate that a good understanding of cancer-bacteriome interactions may open the door to new lines of holistic bacteriotherapy for cancer that is otherwise unavailable.

BJC Reports; <https://doi.org/10.1038/s44276-026-00229-7>

INTRODUCTION

The study of the human microbiome has attracted significant research interest in recent years [1–7]. This is due to accumulating evidence suggesting that the microbiome plays an important role in human health [8–10], which has led to many attempts to harness the microbiome for treating certain medical conditions, especially in the broader field of gastroenterology [11, 12]. Recently, it has been shown in multiple seminal studies that the microbiome is intimately connected with cancer [13–20]. Cancer patients have altered microbial populations compared to cancer-free patients [14, 21–23]. This suggests that there is a mode of interaction between cancer and the microbiome. Although we do not know whether all microbiome changes associated with cancer are causes for the latter [24], it is clear that cancer cells and microbial cells interact [25, 26]. Given the fact that many organs in the human body have distinct microbiome populations [3, 27, 28], understanding the interactions between specific organ cancer cells and the microbiome may lead to advances in the early detection and management of cancer. This opens a new frontier in the fight against cancer, one of the deadliest diseases in human history.

The human microbiome consists of three components: the virome (viruses), the bacteriome (bacteria), and the mycobiome (fungi). Of the three components, the bacteriome is the largest, accounting for more than 90% of all microbes hosted by the human body and ~50% of the human body cell content [29, 30]. However, even with many advancements in bacteriome studies, a significant portion of bacteria remain unclassified [31, 32], let alone the functions they perform. In the context of cancer,

the bacteriome is of special interest because studies have shown that some bacterial taxa have direct effects on cancer initiation and progression [33–37] (Table 1). For example, extensive evidence supports that *Helicobacter pylori* has a direct causal link with gastric cancer, which has made this the only bacterial species that has been listed as a carcinogen by the World Health Organization [38–40]. *F. nucleatum* is another bacterial species that is found in unusually high abundance in microbial communities in colorectal and lung cancers [41–44]. Furthermore, certain strains of *E. coli* have been shown to produce genotoxins that induce DNA double-strand breaks that are associated with various types of cancers [45–49].

Conversely, certain bacterial species have been associated with cancer-suppressive effects [50–53]. For example, *Bifidobacterium* has been shown to suppress tumor growth by stimulating a tumor-suppressive immune response [54]. Additionally, *Salmonella* has been shown to localize to multiple tumor types and induce apoptosis [55, 56]. *Listeria monocytogenes* also has cancer-suppressive properties as it activates the adaptive immune system against tumors and changes the tumor microenvironment [57–59]. This has spurred an exciting field of research that delves into engineering strategies to develop cancer bacteriotherapies [36, 50, 60, 61]. With that being said, the study of the interactions between bacterial populations and cancer cells remains in its infancy due to the complexity of the human bacteriome and the lack of suitable techniques to dissect these interactions at the cellular level.

¹Microbial and Biome Sciences group, Bioscience Division, Los Alamos National Laboratory, Los Alamos, NM, USA. ²Genomics and Bioanalytics group, Bioscience Division, Los Alamos National Laboratory, Los Alamos, NM, USA. ✉email: akumar@lanl.gov

Received: 4 October 2025 Revised: 10 March 2026 Accepted: 8 April 2026

Published online: 04 May 2026

Table 1. List of bacteria that are associated with cancer types.

Bacterial species	Cancer type	References
<i>H. pylori</i>	Gastric Cancer, liver cancer, MALT lymphoma	[39, 167, 168]
<i>F. nucleatum</i>	Colorectal, Cervical, Lung, Pancreas	[41, 169]
<i>P. gingivalis</i>	Colorectal, Pancreas	[170, 171]
<i>Enterococcus faecalis</i>	Liver	[172]
<i>E. coli</i>	Lung, Colorectal	[173, 174]
<i>Streptococcus</i>	Prostate	[175]
<i>Staphylococcus epidermidis</i>	Lung, Breast	[176, 177]
<i>Enterobacteriaceae</i>	Breast	[177]
<i>S. yanoikuyae</i>	Breast	[178]
<i>Salmonella typhi</i>	Gallbladder	[179]
<i>Enterotoxigenic B. fragilis</i>	Colorectal	[180]
<i>Campylobacter jejuni</i>	Colorectal	[181]
<i>Peptostreptococcus anaerobius</i>	Colorectal	[182]
<i>Morganella morganii</i>	Colorectal	[183]
<i>Chlamydia pneumoniae</i>	Lung	[184]
<i>Salmonella</i> Enteritidis	Colon	[185]

The link between the bacteriome and cancer is of immense interest to microbiologists and medical scientists [62]. One of the remarkable features of the bacteriome is that it exists in almost every organ and/or tissue that can develop tumors [63]. Furthermore, bacteriome populations change depending on the location within the human body and are dynamic [27, 64], which offers a route to developing tissue-specific treatments that can be tailored to the patient's bacteriome. This dichotomy between the general presence of the bacteriome and the specificity of the bacteriome population to different tissues offers a promise of personalized, yet equally effective, general cancer treatment. However, the path towards the development of such treatments remains elusive due to multiple technical challenges. Importantly, due to the complexity of bacteriomes, one needs to study the behavior of healthy and cancerous cells in the presence of complex bacterial communities and their combinatorial interactions, as opposed to the presence of a single bacterial species. This is because bacteria have species-/strain-specific effects [65–67], and are highly adaptive as their behavior is strongly influenced by the presence of other microbes within the population [68]. However, such investigations are obstructed by many challenges, including the difficulty of mimicking the natural host environment and the lack of high-throughput methods to study cancer cell-bacteriome interactions.

Here, we review the current state of the field with a focus on work that has been done to understand cancer progression in the presence of complex and naturally occurring human bacteriomes. We discuss currently available experimental approaches, challenges, and limitations to using these approaches, and possible routes to advance the study of cancer-bacteriome interactions. We further discuss the need for theoretical and mathematical frameworks to tackle the complexity of the human bacteriome. This problem needs strong interdisciplinary collaboration between microbiologists, geneticists, chemists, physicists, mathematicians, and data scientists. This paper is intended to highlight the need for developing new methods to study the interaction between cancer cells and the bacteriome and to present this problem in simple general terms to audiences across all scientific disciplines.

ROUTES OF BACTERIA-INDUCED TUMORIGENESIS

Several mechanisms have been reported in the literature through which bacteria can influence cancer development, including DNA damage, metabolic stimulation, inflammation, apoptosis regulation, or suppression of tumor suppressors [33, 34, 37, 69, 70]. In some cases, multiple mechanisms are responsible for cancer development [71]. Here, we discuss some of the pathways by which certain well-established “oncobacteria” induce cancer development. For example, *H. pylori* infection is believed to cause tumorigenesis by a complex cascade of interactions. These start by inducing macrophage and neutrophil accumulation in host cells, leading to increased production of reactive oxygen species (ROS), which in turn causes DNA damage and subsequently tumorigenesis [72, 73]. Additionally, *H. pylori* infection has been shown to increase the recruitment of cytokines, which locally induce chronic inflammation and the activation of tumor-promoting pathways such as the STAT3 pathway [70, 73, 74]. The STAT3 pathway affects many downstream processes such as proliferation and apoptosis inhibition [75, 76].

Another example includes certain *E. coli* strains that express polyketide synthases (PKs) [77], and produce a small molecule called colibactin, which is unstable and causes DNA damage via cross-linking [78]. Such DNA damage leads to mutations upon DNA repair, and the high rate of mutation of host cells in *E. coli*-rich environments leads to tumorigenesis via induced genetic instability [66]. *F. nucleatum* is another example that aids in secreting cytokines such as IL-1 β , IL-6, and TNF- α , which lead to chronic inflammation that fosters cancer progression [79, 80]. *F. nucleatum* has also been shown to dysregulate the Wnt/ β -catenin signaling pathway, which leads to abnormal increase in cell growth and inhibition of apoptosis [81]. Lastly, *F. nucleatum* stimulates DNA damage in many ways, including halting DNA repair processes [82], thereby increasing genetic instability and cancer initiation. *Bacteroides fragilis* is a species where some strains express a toxin known as Bacteroides fragilis toxin (BFT) that binds to E-cadherin and promotes the Wnt/ β -catenin signaling pathway [82]. It has been shown that BFT stimulates MYC signaling [83], which is an oncogene that promotes cell proliferation [84]. In this fashion, *B. fragilis* also induces cancer initiation and progression by upregulating cellular proliferation and downregulating apoptosis.

In sum, there are multiple mechanisms by which a bacterium can increase the risk of cancer initiation. However, the vast majority of human-inhabited bacterial species have not been mechanistically investigated in relation to cancer. This is partly because there are no current high-throughput platforms that can identify bacterial species that influence cancer cells in the first place. Furthermore, bacteria are known to be highly adaptive, and interbacterial networks can further complicate the mechanisms of cancer influence. It is therefore important to first appreciate the complexity of the bacteriome in connection with promoting or suppressing cancer and direct research towards developing high-throughput methods that can screen bacteria for influence on cancer. Doing so will enable us first to extract common mechanisms of bacteria-induced cancer (DNA damage, chronic inflammation, apoptosis dysregulation, etc.) and then to devise strategies to block these downstream mechanisms of cancer promotion.

THE VASTNESS OF THE HUMAN BACTERIOME AND THE NEED FOR CANCER-BACTERIOME STUDIES

Not only does the total number of bacterial cells in humans equal that of human cells [29], but there are thousands of bacterial species in the human bacteriome. Importantly, the bacteriome composition is associated with the overall health of an individual [2]. The local bacteriome composition depends on the tissue type and is also not homogeneous within the human population [27].

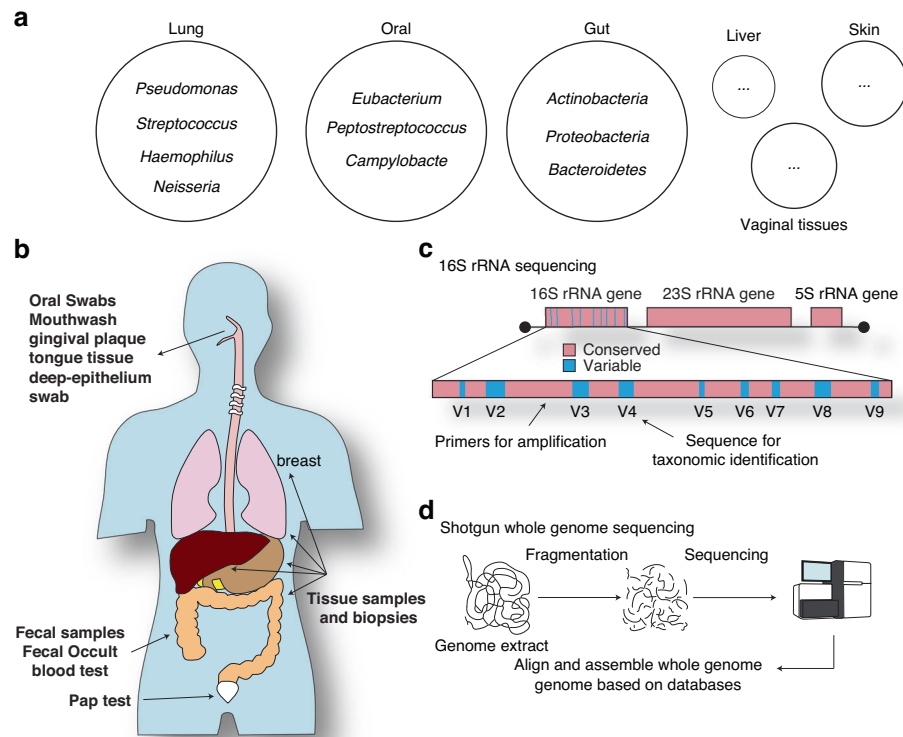


Fig. 1 Bacteriome variability and available methods of analysis. **a** The variability of the human bacteriome based on the organ. **b** Sample acquisition methods at different organs for bacteriome analysis of different cancers. **c** A scheme showing the 16S rRNA gene and its conserved and variable regions. **d** Diagram showing the basic steps of shotgun whole-genome sequencing of microbial samples for microbial diversity and functional analysis.

Distinct bacterial compositions are found in the lung [85–87], oral cavity [88], gut [89–91], and other organs [92–95] (Fig. 1a). These organs and other microenvironments represent complex ecological niches that influence bacterial adaptation and their interplay with cancer.

A recent review by Tekle and Garrett highlighted that the relationship between microbes and cancer is shaped by the broader microbial community, as coexisting species can modify the host's susceptibility to oncogenic bacteria [33]. The authors highlight a study that showed that the combination of *Faecalibaculum rodentium* and *Holdemania bififormis* has an anti-proliferative effect on tumors in mice, irrespective of the other microbiota [96]. This highlights several interesting possibilities. First, it confirms the possibility that some components of the microbiome possess tumor-suppressive capability irrespective of the residing microbial community. It also highlights ecological concepts such as cooperation and competition in dictating a cancer-promoting bacteriome. Accordingly, *one should not only pursue cancer-bacterium interactions but also look at the inter-bacterial interactions that can affect the susceptibility of the host to cancer development*. Consequently, there is a dire need for cancer-bacteriome studies that mechanistically probe the interaction of cancer cells with this complex cocktail of bacteria.

Given the diversity of the human bacteriome and its dependence on host organs, there is a multitude of methods that have been developed to acquire bacteriome samples from different organs. These methods provide a window into the organ-to-organ variation in the bacteriome. Importantly, this complicates comparisons between different studies, as the method of obtaining the bacteriome sample may influence the results [97]. This also highlights the need for mechanistic and co-culture studies as opposed to correlation in the context of cancer. *The*

abundance of certain bacterial species that are associated with a higher risk of cancer does not automatically entail a causal link that is general to other organs [98]. This may be why we see different bacteria associated with different cancers and not a single set of bacterial species that always correlates with cancer.

CURRENT APPROACHES TO STUDYING THE HUMAN CELL-BACTERIOME

There are several studies that have probed the relation between cancer and the bacteriome as a whole. The bulk of these studies focused on bacteriome changes associated with the occurrence of cancer. However, few of them probed cancer evolution as a function of an increased abundance of one bacterium in the whole bacteriome population. We will discuss the methods used in these studies and highlight their strengths and limitations. Most bacteriome studies have been done in the context of oral cancer or colorectal cancer. A typical study of this nature includes sample acquisition, sample processing, and sample analysis via various sequencing methods.

Sample acquisition

Acquiring samples from cancer patients usually includes tissue extraction and/or body fluids and decoding the bacteriome profile (Fig. 1b). Controls are generally healthy individuals or tumor-adjacent healthy tissue. In the context of oral cancer, sample acquisition can vary, such as tumor and tumor-adjacent tissues [99], biopsies and deep-epithelium swabs [100], and oral swabs and washes [101–104]. For colorectal cancer (CRC), most studies use fecal samples to analyze the gut bacteriome of cancer patients [105–112] in addition to tissue samples [113]. Caveats for sample acquisition include the variability of results based on collection

methods and based on the variance of the bacteriome itself within different locations [114, 115]. With that being said, tissue samples are superior and are needed in order to map complete relations between the bacteriome and cancer. This is because the tumor microenvironment has a significant role to play in cancer progression and proliferation [116]. The tumor microenvironment has also been shown to contain a unique microbiome and is therefore another important aspect of bacteriome-cancer interaction that needs to be studied [14]. We assert that conclusions drawn from bacteriome-cancer studies must be carefully assessed in relation to the sample type and how it may influence the generalizability of the obtained results.

Sample analysis

Sample analysis of bacteriome composition can be generally classified into two main methods: 16S rRNA sequencing and metagenomic shotgun sequencing. The 16S rRNA genes have several regions that are highly conserved throughout the entire bacterial tree of life. Interspersed among these regions of conservation are variable, species-specific regions (Fig. 1c). The conserved regions are used for primer design to amplify the variable regions using polymerase chain reaction methods. The variable regions are then sequenced for species identifications [117–121]. Metagenomic shotgun sequencing relies on sequencing the genetic material of the entire microbial population (bacteria, archaea, fungi, viruses, phages, etc.) and even human cells present within the sample without amplification [122, 123] (Fig. 1d). This is done by fragmenting the genetic content of a sample and then sequencing it all. The sequenced fragments are then either assembled or aligned directly to genomic databases to identify the species and their abundance in the population.

While both methods are useful, they differ in value. Legacy sequencing methods, such as amplicon sequencing (i.e., 16S rRNA), are cost-effective and quick to identify the composition of a bacteriome. However, these approaches lack the necessary resolution, and they do not allow for the isolation and/or culturing of the bacterial species of interest [124]. Metagenomic shotgun sequencing is significantly more expensive due to the need for more sequencing reads. However, it provides a more accurate determination of the bacteriome composition, especially concerning low-abundance species and all other microbial communities [124]. Metagenomics is therefore superior in most cases and should be used if available. The cost factor between the two methods can be mitigated if one can analyze a large number of samples at once. In other words, if a high-throughput method is developed where many bacteriome-cancer samples are collected and pooled together before analysis, one can increase the value of metagenomic sequencing by analyzing multiple samples in one sequencing run. This, of course, would need significant method development as we will discuss in the following sections.

DEFINING THE QUESTION OF CANCER-BACTERIOME INTERACTIONS

The study of bacteriome-cancer cell interactions represents a complex, multivariate problem. Consider the promotion or suppression of cancer as a multivariable function $F(x_1, x_2, \dots, x_n)$ that depends on thousands of variables x_n that include dependent and independent variables. There also exist interactions between those variables that can impact the overall outcome function. Each variable here represents a bacterial species within the bacteriome. The challenge is to map this function to predict the effect on cancer by looking at the bacteriome. To seek a solution for this problem, there are multiple approaches described in the literature. The simplest one is a brute-force method where we study the effect of each bacterial species on cancer cells (i.e., setting all the variables to zero except for one variable, Fig. 2a). Next, we study the effect of all possible two-component bacteriomes, and so on. While this

approach would yield a complete mapping of the cancer function F , it is not feasible due to the resources it requires and the time it would take (hundreds of years).

Another approach is to look at the asymptotic behavior of the bacteriome at the extremes of the cancer function F . We may do this by looking at individuals who have cancer and analyzing their bacteriome (Fig. 2b). We then look at the individuals who do not have cancer and analyze their bacteriomes (Fig. 2b). After accumulating sufficient data, we can extract common features of the bacteriomes associated with cancer and those not associated with cancer (Fig. 2b). This is the current approach in the field and has resulted in identifying many bacterial species that do have cancer-promoting activity. However, there are limitations to this approach. First, it does not detect bacterial species that are conditionally cancer-promoting (i.e., exert a cancer-promoting effect when another species is present). Second, isolating species of interest that are responsible for the observed effect from a complex bacteriome community is not always possible. Furthermore, this method does not work well in detecting cancer-suppressive bacterial species because the definition of a healthy individual lacks accuracy. A person who does not have cancer is not necessarily healthy, and there is no reason to think that such a person would not develop cancer in the near future.

There is a third approach that can be used to understand the interaction between cancer cells and the bacteriome, which involves monitoring the co-evolution of cancer and bacteriomes in a co-culture setting. Briefly, one would track the growth and viability of cancer cells in a large number of distinct bacteriomes and select for bacteriomes that promote cancer cell proliferation as well as bacteriomes that suppress the growth of the cancer cell population (Fig. 2c). This approach is similar to brute force in principle, but it makes the entire process parallelized, which speeds up the discovery to a great extent. With that being said, in order for this approach to be feasible, it is required that we develop high-throughput methods for both the generation of bacteriome samples and screening them for cancer promotion or suppression. Without the high-throughput methods, this approach can only look at the relation between a cancer cell population and a few components, such as one or two bacterial species. We suggest that expanding research efforts into developing high-throughput screening methods for bacteria-cancer cell interactions is essential to develop a comprehensive understanding of the complex interaction hierarchy between cancer cells and the bacteriome. We speculate that such high-throughput methods will depend on the latest sequencing technologies, as well as significant advances in sample preparation, microscale culturing, and detection methods.

In sum, we can articulate three methods for probing cancer bacteriome interactions. The brute-force method looks at one species at a time and suffers from impracticality, high costs and labor, and limitations with regard to detecting complex and conditional bacterial networks. The second is large-scale asymptotic screening, where we take patients with clear cancer tendencies (i.e., developed cancers) and sequence their bacteriomes to extract common features of bacterial compositions. This is the most feasible approach currently, and it has been successful in detecting bacterial species with cancer-promoting activities [14, 23]. However, this approach is biased towards detecting only cancer-promoting bacteria and is oblivious to the mechanisms by which the said bacteriomes can increase cancer proliferation. A third approach that we suggest is high-throughput co-culturing of mammalian cells with distinct complex bacteriomes. This approach is currently not feasible due to the lack of experimental platforms that can generate a large number of bacteriomes and analyze their influence on cancer cells. However, this approach holds great promise since it can detect both cancer-promoting and cancer-suppressive species as well as conditional networks of bacteria that may promote or suppress cancer. We argue that this

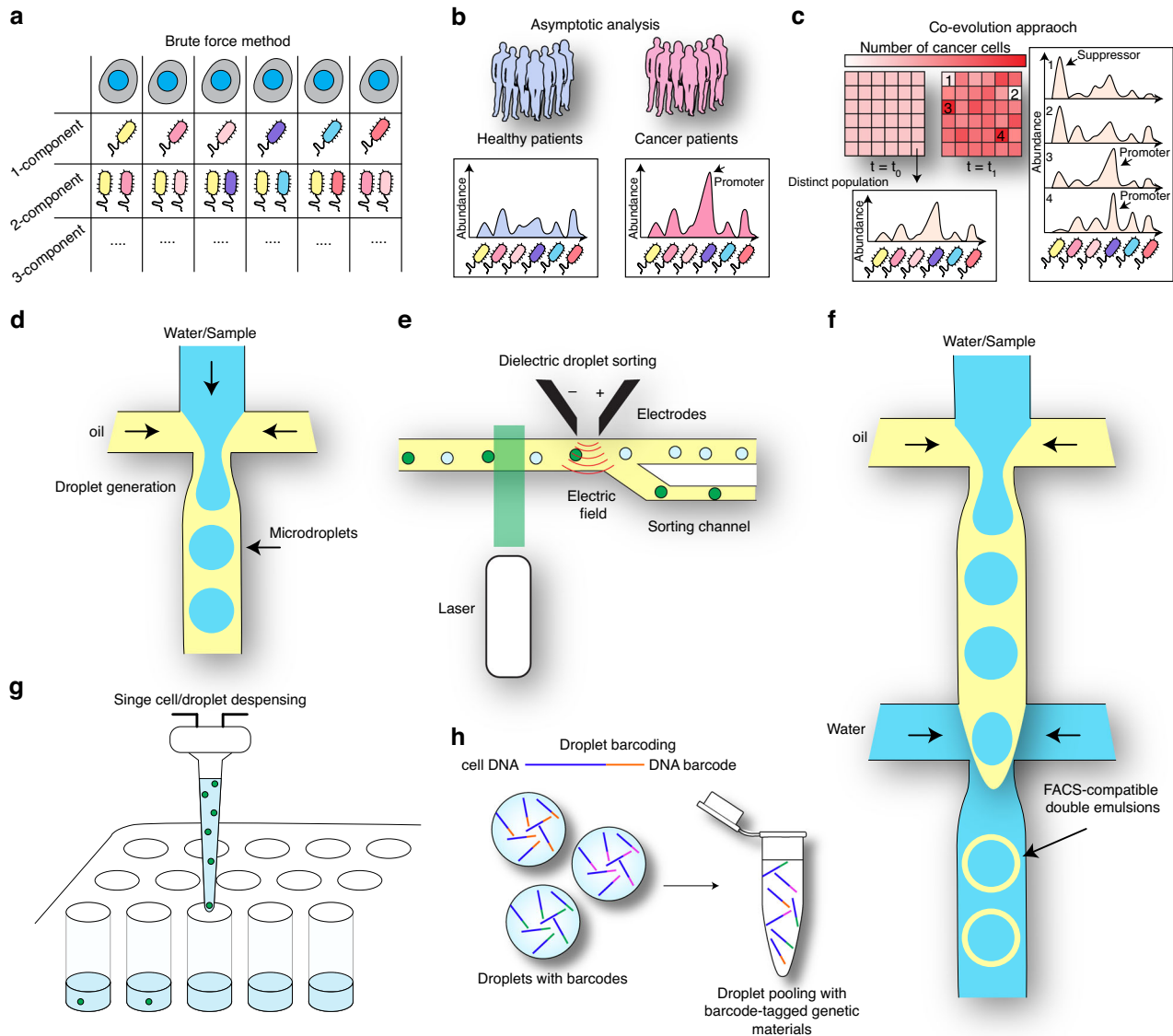


Fig. 2 Approaches to solving the cancer-bacteriome problem. **a** Brute-force method where the permutations of bacterial species are tested individually. **b** Cancer promoters and suppressors can be identified by looking at the bacteriome of asymptomatic cases of cancer patients and healthy patients. **c** Co-evolution approach where a large number of distinct bacteriomes are generated and the evolution of cancer cells within is tested. **d** A scheme of a flow-focusing microfluidic droplet generator to generate water-in-oil droplets. **e** An example of microfluidic droplet sorting using an electric field coming from electrical electrodes to deflect fluorescent droplets into the sorting channel based on fluorescence. **f** An example of how water/oil/water double emulsions can be generated. These double emulsions can be sorted using conventional FACS cell sorters. **g** A scheme illustrating the concept of single-cell/droplet dispensing. **h** Droplet barcodes allow for tracking the genetic material coming from each droplet via sequencing, even after the emulsion is broken and the droplet samples are pooled together.

approach may be feasible if researchers are able to develop high-throughput co-culturing platforms. In the next section, we highlight microfluidic technologies that may prove useful for developing such miniaturized co-culturing platforms.

HIGH-THROUGHPUT MICRODROPLETS FOR STUDYING THE BACTERIOME

High-throughput methods are necessary to effectively understand the complex interactions between the bacteriome and host cells. The difficulty in studying bacteriome-cancer interactions is that large volumes of data are needed for sample analysis and characterization and require high-throughput, multiplexed experimental techniques for sample and data generation and processing. Importantly, it is necessary to be able to generate a large

number of compositionally distinct bacteriomes in order to study bacteriome composition features and their effect on disease. Of the emerging technologies, microfluidics has shown incredible promise for high-throughput scaling of currently established experimental technologies and can already be readily integrated with other experimental techniques. Microfluidics relies on the manufacturing of microfluidic chips [125]. The microfluidic channel design allows the generation of equally sized aqueous microdroplets in an organic liquid phase, such as oil (Fig. 2d). Millions of droplets can be generated, representing millions of independent/isolated samples, in a relatively short time (1–2 h). The analysis of these samples, however, is not straightforward due to their pico- and femto-liter volume. In order for microfluidics to be truly high-throughput, one must be able to analyze each droplet independently.

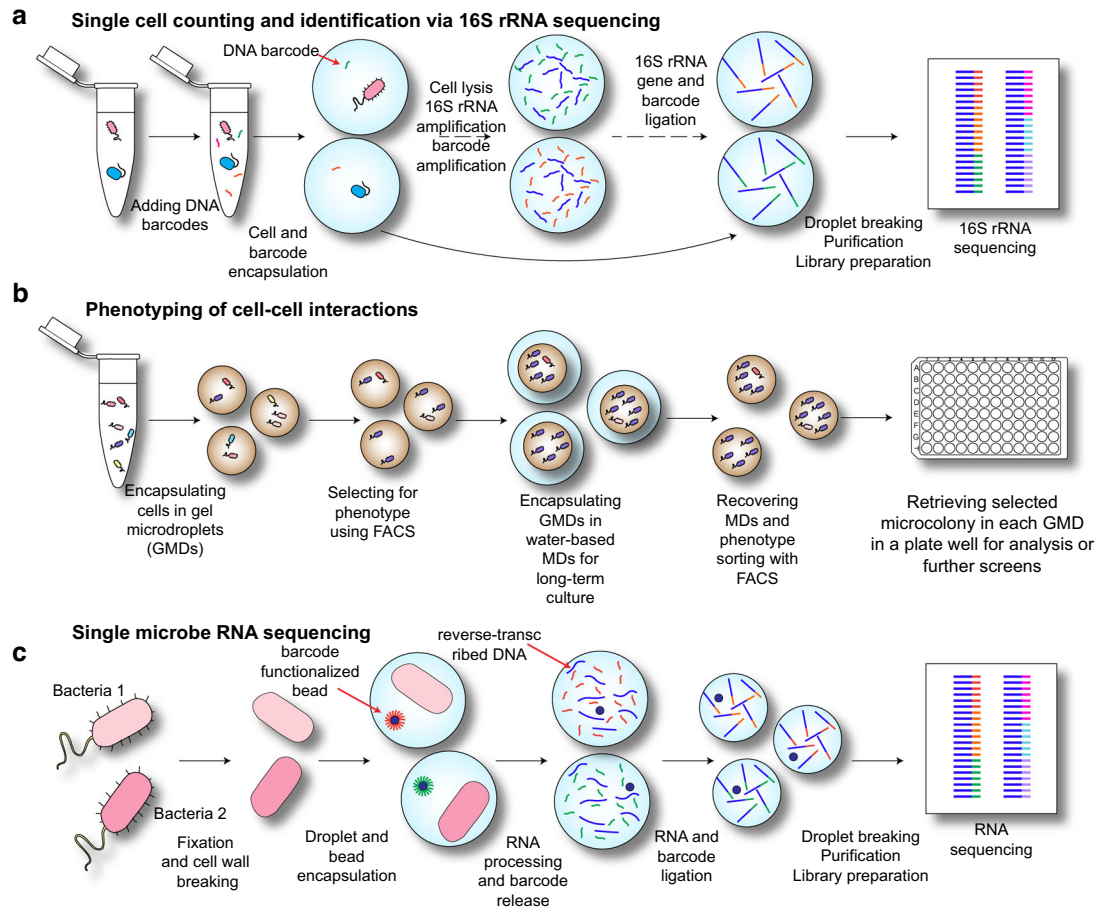


Fig. 3 Microfluidics-based high-throughput methods for studying host cell and bacteriome interactions. a Single-cell counting and identification via BARBIQ-seq [133]. **b** cell-cell interaction phenotyping by HiSCI [137]. **c** smRandom-seq workflow for single-microbe RNA sequencing [139].

The first mode of analysis of microdroplets is droplet sorting [126], which can split the droplet population into two populations depending on a certain readout, such as fluorescence, light scattering, or other measurements. This allows the selection of a small percentage of the droplet population that exhibits a high trait (detected by fluorescence) [127] (Fig. 2e). Other microfluidic variations can be made to allow droplets to be sorted with conventional cell sorting flow cytometry (double emulsions, Fig. 2f) [128–130]. While droplet sorting can lead to bulk splitting of the sample population into smaller subpopulations, it does not achieve single droplet resolution, which significantly limits the applications of this technology. An advancement came with the advent of single-cell dispensing technology, which allows droplets/cells to be dispensed into 96-well or 384-well plates [131] (Fig. 2g). This improves the throughput, yet the size of the well plate poses a limit on the number of droplets that can be analyzed individually. A breakthrough advancement was achieved with the invention of DNA barcoding [132–134], which allows tagging the genetic material in each droplet with a short DNA sequence that is unique to the droplet (Fig. 2h). This DNA barcode allows for the identification of genetic material originating from each droplet individually via sequencing experiments without the need for unique sorting instruments, which significantly improves the throughput.

These developments represent a paradigm shift in high-throughput studies and have been used with immeasurable success in the study of the human microbiome. They can be used

to detect and quantify bacteriome populations, study bacteria-host interactions via transcriptomics and RNA sequencing, and identify strains and genetic variants associated with cancer phenotypes. Below, we provide examples of the combined use of microfluidics with genomics methods that have been successfully applied to high-throughput microbiome characterization studies that are relevant to the study of bacteriome-cancer interactions.

Microfluidic high-throughput methods can be used for species detection. Zhang et al. sequenced ~20000 single-microbe genomes with strain resolution from the human gut microbiome [135]. They revealed 100 bacterial species, along with their sub-strain identification, and detected horizontal gene transfers and host-phage interactions from a single donor's microbiome sample. In terms of microbial composition and counting, Jin et al. developed a method based on droplet microfluidics to identify and count single bacterial cells within the microbiome using a DNA-barcoding method called BarBIQ [136] (Fig. 3a). This method overcomes the limitations associated with quantifying bacterial species and features enormous throughput, in the range of 500,000 cells, as the authors reported. Ohan et al. used gel microdroplets (GMDs) as picoliter microenvironments to perform high-throughput screening of cell-to-bacteria interactions (HiSCI) using fluorescence sorting methods [137] (Fig. 3b). Other droplet-based techniques have also been deployed to enrich rare and slow-growing species in natural microbiomes via encapsulation, culturing, and sorting of micro-organisms [138].

Microfluidic high-throughput methods can also be used to perform RNA sequencing and transcriptomics to observe bacterial interactions and genetic profiles. There are multiple studies that successfully used high-throughput RNA sequencing to study bacterial transcriptomics with single-cell resolution. Xu et al. proposed a droplet-based high-throughput single microbe RNA sequencing method called smRandom-seq [139] (Fig. 3c). It is also possible to carry out complete biochemical reactions within microfluidic droplets. The same group has shown the possibility of doing single-cell transcriptomics using droplet technology by storing the PCR primers and cell barcodes into a hydrogel particle and mixing them with the cell suspension [133]. In such a case, the sequencing data could be split based on the barcodes, and hence droplet-specific transcriptomics are obtained. These and other examples illustrate the great potential for using microdroplet techniques to probe host-bacteria interactions. We believe that now is a great opportunity to harness the immense progress in droplet microfluidics and microbiology to address the problem of sampling the complexity of the microbiome, and couple this technique with high-throughput 'omics data generation and analysis. These directions could advance the field forward by enabling high-throughput co-culture of cancer cells and distinct bacteriomes using droplet microfluidics as a platform.

ANALYTICAL TOOLS TO ANALYZE BACTERIOME-CANCER INTERACTION DATA

The tremendous volume and high-dimensional nature of the data obtained from high-throughput bacteriome-cancer interaction experiments have led to the development of a myriad of analysis techniques ranging from statistical ecological analysis to dynamical modeling. Empirical techniques like statistical analysis and some machine-learning algorithms attempt to sift through data to reveal underlying relationships. In contrast, "physics-based" or "mechanistic" modeling encodes a biological hypothesis and asks how well the data fit that hypothesis. This seeming dichotomy of "top-down" vs. "bottom-up" is actually a spectrum of strategies, with information often able to flow between modeling paradigms, such as when statistical analysis suggests relationships that can be encoded in dynamical models that then provide predictions to be verified by statistics. As with all computational biology, the right tool for answering any question about the bacteriome-cancer interaction will be highly dependent on the question being asked. In this section, we highlight a sampling of analysis tools that may prove useful for discovering patterns or revealing the interactions between the bacteriome and cancer.

Dimensionality reduction

Dimensionality reduction techniques, such as principal component analysis (PCA), identify dependencies in a data set to remove noise and enable further analysis. Reducing the dimensionality of the data allows researchers to visualize data and is a key step to efficient downstream analysis [140]. Dimensionality reduction alone cannot reveal causal relationships within data, but it can suggest relationships between data features. For example, PCA has been used to identify biomarkers that best describe disease conditions, as well as distinguish whether a highly complex microbial diversity is closer to those of healthy patients or cancer patients [141]. Several other dimensionality reduction techniques, which are reviewed elsewhere [140], include both linear and nonlinear methods for projecting the data onto lower-dimensional spaces or manifolds in order to reduce noise and enable analysis.

Statistical modeling and network inference

Statistical models can be used to draw conclusions and generate new hypotheses from data by comparing the data to the expected outcome from some random process. For example, multifactor analysis of variance (ANOVA) is used to determine the extent to

which a predictor variable is able to affect the outcome variable [142, 143] by comparing the data to the expectation and variance of a general linear predictor model. It is mostly used when a single outcome variable depends on many factors. Multi-factor ANOVA computes not only the one-to-one relationship between each predictor factor and the outcome variable but also computes the interactions between the factors themselves, making it a potentially useful tool for analyzing large sets of data coming from high-throughput bacteriome-cancer cell interaction studies. Another class of statistical methods compares the data to a model that asserts that the bacteriome consists of several communities. Non-negative matrix factorization [144] or parameter fitting to more complex probabilistic models [145] can reveal the composition of these communities and their contribution to each sample of the data. These approaches can be especially powerful in identifying communities, rather than individual microbes, that may be associated with different cancer outcomes.

Machine learning and artificial intelligence

The use of AI in microbiome research is vast and has recently been reviewed elsewhere [146–149]. AI methods have been used in many aspects of microbiome research, including taxonomic profiling [150], genome assembly and classification from sequencing data [151, 152], antibiotic resistance prediction [153, 154], and more. AI has also been used to identify microbial markers associated with cancers such as CRC [155–157]. These studies rely on existing data sets from cancer patients to construct models that can predict cancer based on the microbiome composition obtained from tissues or other sample sources. Another relevant application of AI includes predicting protein-protein interactions between bacteria and host cells [158, 159] and predicting microbe-disease associations [160, 161]. In fact, even large language models such as GPT-3 and BERT have been used to survey the literature for microbe-disease associations [162].

CONCLUSION

Cancer has put a huge toll on human health and our economy [163]. Although there has been tremendous progress in our understanding of the genetic and epigenetic factors that govern cancer initiation and progression [164, 165], the rise in the number of cancer patients and the mortality rate remains high. The emerging link between cancer and the microbiome has recently allowed us to utilize and engineer microbes to detect and even fight cancer [13, 17, 18, 166]. Yet, our understanding of the interplay between cancer and the microbiome in the complex ecological system that is the human body is still limited. Here, we highlight the challenge of investigating cancer-bacteriome interactions and emphasize the importance of studying the bacteriome as a whole entity and probing its interactions with cancer cells. While some progress has been made in genetic engineering of individual bacterial strains to fight cancer, we propose that engineering the bacteriome as a whole, i.e., manipulating the population of natural species, may offer a more robust strategy to combat cancer. Apart from highlighting the challenges, we also review some of the approaches taken to better understand the impact of the bacteriome on cancer, speculating on the most appropriate and effective approaches, such as high-throughput co-culturing platforms. Consequently, we argue that high-throughput methods are essential to effectively probe and map cancer-bacteriome interactions and highlight some of the promising methods that can be used to study cancer-bacteriome interactions. To be able to effectively understand and then harness the interactions of our bacteriome with cancer, we also need to develop effective experimental and analytical tools that are capable of handling large volumes of complex, multidimensional data. To this end, we also review some of the multivariable statistical methods as well as the emerging use of AI

in microbiology. Overall, there is significant potential for progress that could be made towards microbe-based cancer therapy by studying and understanding the relation between cancer and complex multi-species bacterial populations that constitute the bacteriome.

DATA AVAILABILITY

No datasets were generated or analyzed during the current study.

REFERENCES

- Durack J, Lynch SV. The gut microbiome: Relationships with disease and opportunities for therapy. *J Exp Med*. 2019;216:20–40.
- Shreiner AB, Kao JY, Young VB. The gut microbiome in health and in disease. *Curr Opin Gastroenterol*. 2015;31:69–75.
- Ursell LK, Metcalf JL, Parfrey LW, Knight R. Defining the human microbiome. *Nutr Rev*. 2012;70:S38–S44. (suppl_1).
- Blaser MJ. The microbiome revolution. *J Clin Invest*. 2014;124:4162–5.
- Knight R, Callewaert C, Marotz C, Hyde ER, Debelius JW, McDonald D, et al. The microbiome and human biology. *Annu Rev Genom Hum Genet*. 2017;18:65–86.
- Dominguez-Bello MG, Godoy-Vitorino F, Knight R, Blaser MJ. Role of the microbiome in human development. *Gut*. 2019;68:1108–14.
- Gerber GK. The dynamic microbiome. *FEBS Lett*. 2014;588:4131–9.
- de Vos WM, de Vos EA. Role of the intestinal microbiome in health and disease: from correlation to causation. *Nutr Rev*. 2012;70:S45–S56.
- Moloney RD, Desbonnet L, Clarke G, Dinan TG, Cryan JF. The microbiome: stress, health and disease. *Mamm Genome*. 2014;25:49–74.
- Sharma N, Bhatia S, Sodhi AS, Batra N. Oral microbiome and health. *AIMS Microbiol*. 2018;4:42.
- Sorbara MT, Pamer EG. Microbiome-based therapeutics. *Nat Rev Microbiol*. 2022;20:365–80.
- Mimee M, Citorik RJ, Lu TK. Microbiome therapeutics—advances and challenges. *Adv Drug Deliv Rev*. 2016;105:44–54.
- Sepich-Poore GD, Zitvogel L, Straussman R, Hasty J, Wargo JA, Knight R. The microbiome and human cancer. *Science*. 2021;371:eabc4552.
- Nejman D, Livyatan I, Fuks G, Gavert N, Zwang Y, Geller LT, et al. The human tumor microbiome is composed of tumor type-specific intracellular bacteria. *Science*. 2020;368:973–80.
- Chen J, Douglass J, Prasath V, Neace M, Atrichian S, Manjili MH, et al. The microbiome and breast cancer: a review. *Breast Cancer Res Treat*. 2019;178:493–6.
- Ghaddar B, Biswas A, Harris C, Omary MB, Carpizo DR, Blaser MJ, et al. Tumor microbiome links cellular programs and immunity in pancreatic cancer. *Cancer Cell*. 2022;40:1240–53. e5.
- Gopalakrishnan V, Helmink BA, Spencer CN, Reuben A, Wargo JA. The influence of the gut microbiome on cancer, immunity, and cancer immunotherapy. *Cancer cell*. 2018;33:570–80.
- Helmink BA, Khan MAW, Hermann A, Gopalakrishnan V, Wargo JA. The microbiome, cancer, and cancer therapy. *Nat Med*. 2019;25:377–88.
- Irfan M, Delgado RZR, Frias-Lopez J. The oral microbiome and cancer. *Front Immunol*. 2020;11:591088.
- Moskal K, Khurana N, Siegert L, Lee YS, Clevers H, Elinav E, et al. Modeling cancer-microbiome interactions in vitro: A guide to co-culture platforms. *Int J Cancer*. 2024;156:2053–67.
- Sobhani I, Tap J, Roudot-Thoraval F, Roperch JP, Letulle S, Langella P, et al. Microbial dysbiosis in colorectal cancer (CRC) patients. *PLoS One*. 2011;6:e16393.
- Liu N-N, Ma Q, Ge Y, Yi C-X, Wei L-Q, Tan J-C, et al. Microbiome dysbiosis in lung cancer: from composition to therapy. *NPJ Precis Oncol*. 2020;4:33.
- Sepich-Poore GD, McDonald D, Kopylova E, Guccione C, Zhu Q, Austin G, et al. Robustness of cancer microbiome signals over a broad range of methodological variation. *Oncogene*. 2024;43:1127–48.
- Xavier JB, Young VB, Skufca J, Ginty F, Testerman T, Pearson AT, et al. The cancer microbiome: distinguishing direct and indirect effects requires a systemic view. *Trends Cancer*. 2020;6:192–204.
- Dutta D, Lim SH. Bidirectional interaction between intestinal microbiome and cancer: opportunities for therapeutic interventions. *Biomark Res*. 2020;8:31.
- Zong Y, Zhou Y, Liao B, Liao M, Shi Y, Wei Y, et al. The interaction between the microbiome and tumors. *Front Cell Infect Microbiol*. 2021;11:673724.
- She J-J, Liu W-X, Ding X-M, Guo G, Han J, Shi F-Y, et al. Defining the biogeographical map and potential bacterial translocation of microbiome in human 'surface organs'. *Nat Commun*. 2024;15:427.
- Kennedy MS, Chang EB. The microbiome: composition and locations. *Prog Mol Biol Transl Sci*. 2020;176:1–42.
- Sender R, Fuchs S, Milo R. Revised estimates for the number of human and bacteria cells in the body. *PLoS Biol*. 2016;14:e1002533.
- Li X, Bi Y. How many human and bacteria cells are in the human body. *Infect Dis Transl Med*. 2017;3:1–2.
- Thomas AM, Segata N. Multiple levels of the unknown in microbiome research. *BMC Biol*. 2019;17:48.
- Kowarsky M, Camunas-Soler J, Kertesz M, De Vlaminck I, Koh W, Pan W, et al. Numerous uncharacterized and highly divergent microbes which colonize humans are revealed by circulating cell-free DNA. *Proc Natl Acad Sci USA*. 2017;114:9623–8.
- El Tekle G, Garrett WS. Bacteria in cancer initiation, promotion and progression. *Nat Rev Cancer*. 2023;23:600–18.
- Li S, Liu J, Zheng X, Ren L, Yang Y, Li W, et al. Tumorigenic bacteria in colorectal cancer: mechanisms and treatments. *Cancer Biol Med*. 2022;19:147–62.
- Okumura S, Konishi Y, Narukawa M, Sugiyama Y, Yoshimoto S, Arai Y, et al. Gut bacteria identified in colorectal cancer patients promote tumorigenesis via butyrate secretion. *Nat Commun*. 2021;12:5674.
- Duong MT-Q, Qin Y, You S-H, Min J-J. Bacteria-cancer interactions: bacteria-based cancer therapy. *Exp Mol Med*. 2019;51:1–15.
- Lax AJ, Thomas W. How bacteria could cause cancer: one step at a time. *Trends Microbiol*. 2002;10:293–9.
- Polk DB, Peek Jr RM. *Helicobacter pylori*: gastric cancer and beyond. *Nat Rev Cancer*. 2010;10:403–14.
- Parsonnet J. *Helicobacter pylori* and gastric cancer. *Gastroenterol Clin North Am*. 1993;22:89–104.
- Venerito M, Vasapolli R, Rokkas T, Delchier JC, Malfertheiner P. *Helicobacter pylori*, gastric cancer and other gastrointestinal malignancies. *Helicobacter*. 2017;22:e12413.
- Alon-Maimon T, Mandelboim O, Bachrach G. *Fusobacterium nucleatum* and cancer. *Periodontology* 2000. 2022;89:166–80.
- Shang F-M, Liu H-L. *Fusobacterium nucleatum* and colorectal cancer: a review. *World J Gastrointest Oncol*. 2018;10:71.
- Hashemi Goradel N, Heidarzadeh S, Jahangiri S, Farhood B, Mortezaee K, Khanlarkhani N, et al. *Fusobacterium nucleatum* and colorectal cancer: a mechanistic overview. *J Cell Physiol*. 2019;234:2337–44.
- McIlvanna E, Linden GJ, Craig SG, Lundy FT, James JA. *Fusobacterium nucleatum* and oral cancer: a critical review. *BMC Cancer*. 2021;21:1–11.
- Wassenaar TM. *E. coli* and colorectal cancer: a complex relationship that deserves a critical mindset. *Crit Rev Microbiol*. 2018;44:619–32.
- Bonnet M, Buc E, Sauvanet P, Darcha C, Dubois D, Pereira B, et al. Colonization of the human gut by *E. coli* and colorectal cancer risk. *Clin Cancer Res*. 2014;20:859–67.
- Nouri R, Hasani A, Shirazi KM, Alivand MR, Sepehri B, Sotoodeh S, et al. *Escherichia coli* and colorectal cancer: unfolding the enigmatic relationship. *Curr Pharm Biotechnol*. 2022;23:1257–68.
- Jian C, Yinhang W, Jing Z, Zhanbo Q, Zefeng W, Shuwen H. *Escherichia coli* on colorectal cancer: a two-edged sword. *Microb Biotechnol*. 2024;17:e70029.
- Swidsinski A, Khilkin M, Kerjaschki D, Schreiber S, Ortner M, Weber J, et al. Association between intraepithelial *Escherichia coli* and colorectal cancer. *Gastroenterology*. 1998;115:281–6.
- Sedighi M, Zahedi Bialvaei A, Hamblin MR, Ohadi E, Asadi A, Halajzadeh M, et al. Therapeutic bacteria to combat cancer: current advances, challenges, and opportunities. *Cancer Med*. 2019;8:3167–81.
- Fan JY, Huang Y, Li Y, Muluh TA, Fu SZ, Wu JB. Bacteria in cancer therapy: a new generation of weapons. *Cancer Med*. 2022;11:4457–68.
- Song S, Vuai MS, Zhong M. The role of bacteria in cancer therapy—enemies in the past, but allies at present. *Infect agents cancer*. 2018;13:9.
- Heimann DM, Rosenberg SA. Continuous intravenous administration of live genetically modified *Salmonella typhimurium* in patients with metastatic melanoma. *J Immunother (Hagerstown, Md)*. 1997. 2003;26:179.
- Sivan A, Corrales L, Hubert N, Williams JB, Aquino-Michaels K, Earley ZM, et al. Commensal *Bifidobacterium* promotes antitumor immunity and facilitates anti-PD-L1 efficacy. *Science*. 2015;350:1084–9.
- Chang W-W, Lee C-H. *Salmonella* as an innovative therapeutic antitumor agent. *Int J Mol Sci*. 2014;15:14546–54.
- Aganija RP, Sivasankar C, Senevirathne A, Lee JH. *Salmonella* as a promising curative tool against cancer. *Pharmaceutics*. 2022;14:2100.
- Ding Y-D, Shu L-Z, He R-S, Chen K-Y, Deng Y-J, Zhou Z-B, et al. *Listeria monocytogenes*: a promising vector for tumor immunotherapy. *Front Immunol*. 2023;14:1278011.
- Tangney M, Gahan CG. *Listeria monocytogenes* as a vector for anti-cancer therapies. *Curr Gene Ther*. 2010;10:46–55.
- Van Pijkeren JP, Morrissey D, Monk IR, Cronin M, Rajendran S, O'Sullivan GC, et al. A novel *Listeria monocytogenes*-based DNA delivery system for cancer gene therapy. *Hum Gene Ther*. 2010;21:405–16.

60. Ballister ER, Michels A, Vincent RL, Kreindler L, Chowdhury S, Upadhyaya S, et al. The emerging landscape of engineered bacteria cancer therapies. *Nat Biotechnol.* 2025;43:1–5.
61. Patyar S, Joshi R, Byrav DP, Prakash A, Medhi B, Das B. Bacteria in cancer therapy: a novel experimental strategy. *J Biomed Sci.* 2010;17:1–9.
62. Liu L, Shah K. The potential of the gut microbiome to reshape the cancer therapy paradigm: a review. *JAMA Oncol.* 2022;8:1059–67.
63. Gilbert JA, Blaser MJ, Caporaso JG, Jansson JK, Lynch SV, Knight R. Current understanding of the human microbiome. *Nat Med.* 2018;24:392–400.
64. Caporaso JG, Lauber CL, Costello EK, Berg-Lyons D, Gonzalez A, Stombaugh J, et al. Moving pictures of the human microbiome. *Genome Biol.* 2011;12:R50.
65. Mäklin T, Taira A, Arredondo-Alonso S, Shao Y, Stratton MR, Lawley TD, et al. Geographical variation in the incidence of colorectal cancer and urinary tract cancer is associated with population exposure to colibactin-producing *Escherichia coli*. *Lancet Microbe.* 2025;6:101015.
66. Wang Y, Fu K. Genotoxins: the mechanistic links between *Escherichia coli* and colorectal cancer. *Cancers.* 2023;15:1152.
67. Komiya Y, Shimomura Y, Higurashi T, Sugi Y, Arimoto J, Umezawa S, et al. Patients with colorectal cancer have identical strains of *Fusobacterium nucleatum* in their colorectal cancer and oral cavity. *Gut.* 2019;68:1335–7.
68. Voolstra CR, Ziegler M. Adapting with microbial help: microbiome flexibility facilitates rapid responses to environmental change. *BioEssays.* 2020;42:2000004.
69. Ilyeva EA, Grivennikov SI. Microbiota-driven mechanisms at different stages of cancer development. *Neoplasia.* 2022;32:100829.
70. Sheweita SA, Alsamghan AS. Molecular mechanisms contributing bacterial infections to the incidence of various types of cancer. *Mediators Inflamm.* 2020;2020:4070419.
71. Duan Y, Xu Y, Dou Y, Xu D. *Helicobacter pylori* and gastric cancer: mechanisms and new perspectives. *J Hematol Oncol.* 2025;18:10.
72. O'Hara AM, Bhattacharyya A, Bai J, Mifflin RC, Ernst PB, Mitra S, et al. Tumor necrosis factor (TNF)- α -induced IL-8 expression in gastric epithelial cells: role of reactive oxygen species and AP endonuclease-1/redox factor (Ref)-1. *Cytokine.* 2009;46:359–69.
73. Krzysiek-Maczka G, Targosz A, Szczyrk U, Strzalka M, Brzozowski T, Ptak-Belowska A. Involvement of epithelial-mesenchymal transition-inducing transcription factors in the mechanism of *Helicobacter pylori*-induced fibroblasts activation. *J Physiol Pharmacol.* 2019;70:727–36.
74. Bockerstett KA, DiPaolo RJ. Regulation of gastric carcinogenesis by inflammatory cytokines. *Cell Mol Gastroenterol Hepatol.* 2017;4:47–53.
75. Fan Y, Mao R, Yang J. NF- κ B and STAT3 signaling pathways collaboratively link inflammation to cancer. *Protein Cell.* 2013;4:176–85.
76. Zou S, Tong Q, Liu B, Huang W, Tian Y, Fu X. Targeting STAT3 in cancer immunotherapy. *Mol Cancer.* 2020;19:1–19.
77. Auvray F, Perrat A, Arimizu Y, Chagneau CV, Bossuet-Greif N, Massip C, et al. Insights into the acquisition of the pks island and production of colibactin in the *Escherichia coli* population. *Microb Genom.* 2021;7:000579.
78. Dziubańska-Kusibab PJ, Berger H, Battistini F, Bouwman BA, Iftekhhar A, Katainen R, et al. Colibactin DNA-damage signature indicates mutational impact in colorectal cancer. *Nat Med.* 2020;26:1063–9.
79. Galasso L, Termite F, Mignini I, Esposito G, Borriello R, Vitale F, et al. Unraveling the role of *Fusobacterium nucleatum* in colorectal cancer: molecular mechanisms and pathogenic insights. *Cancers.* 2025;17:368.
80. Wu Z, Ma Q, Guo Y, You F. The role of *Fusobacterium nucleatum* in colorectal cancer cell proliferation and migration. *Cancers.* 2022;14:5350.
81. Rubinstein MR, Baik JE, Lagana SM, Han RP, Raab WJ, Sahoo D, et al. *Fusobacterium nucleatum* promotes colorectal cancer by inducing Wnt/ β -catenin modulator Annexin A1. *EMBO Rep.* 2019;20:e47638.
82. Sayed IM, Chakraborty A, Abd El-Hafeez AA, Sharma A, Sahan AZ, Huang WJM, et al. The DNA glycosylase NEIL2 suppresses *fusobacterium*-infection-induced inflammation and DNA damage in colonic epithelial cells. *Cells.* 2020;9:1980.
83. Wu S, Morin PJ, Maouyo D, Sears CL. *Bacteroides fragilis* enterotoxin induces c-Myc expression and cellular proliferation. *Gastroenterology.* 2003;124:392–400.
84. Miller DM, Thomas SD, Islam A, Muench D, Sedoris K. c-Myc and cancer metabolism. *Clin Cancer Res.* 2012;18:5546–53.
85. Li R, Li J, Zhou X. Lung microbiome: new insights into the pathogenesis of respiratory diseases. *Signal Transduct Target Ther.* 2024;9:19.
86. Dickson RP, Erb-Downward JR, Freeman CM, McCloskey L, Falkowski NR, Huffnagle GB, et al. Bacterial topography of the healthy human lower respiratory tract. *MBio.* 2017;8:e02287–16.
87. Hilty M, Burke C, Pedro H, Cardenas P, Bush A, Bossley C, et al. Disordered microbial communities in asthmatic airways. *PLoS one.* 2010;5:e8578.
88. Marsh PD. Role of the oral microflora in health. *Microb Ecol Health Dis.* 2000;12:130–7.
89. Qin J, Li R, Raes J, Arumugam M, Burgdorf KS, Manichanh C, et al. A human gut microbial gene catalogue established by metagenomic sequencing. *nature.* 2010;464:59–65.
90. Rajilić-Stojanović M, De Vos WM. The first 1000 cultured species of the human gastrointestinal microbiota. *FEMS Microbiol Rev.* 2014;38:996–1047.
91. Thursby E, Juge N. Introduction to the human gut microbiota. *Biochem J.* 2017;474:1823–36.
92. Leinwand JC, Paul B, Chen R, Xu F, Sierra MA, Paluru MM, et al. Intrahepatic microbes govern liver immunity by programming NKT cells. *J Clin Investig.* 2022;132:e151725.
93. Sookoian S, Salatino A, Castaño GO, Landa MS, Fijalkowky C, Garaycochea M, et al. Intrahepatic bacterial metatranscriptomic signature in non-alcoholic fatty liver disease. *Gut.* 2020;69:1483–91.
94. Urbaniak C, Cummins J, Brackstone M, Macklaim JM, Gloor GB, Baban CK, et al. Microbiota of human breast tissue. *Appl Environ Microbiol.* 2014;80:3007–14.
95. Voigt AY, Emiola A, Johnson JS, Fleming ES, Nguyen H, Zhou W, et al. Skin microbiome variation with cancer progression in human cutaneous squamous cell carcinoma. *J Invest Dermatol.* 2022;142:2773–82. e16.
96. Zagato E, Pozzi C, Bertocchi A, Schioppa T, Saccheri F, Guglietta S, et al. Endogenous murine microbiota member *Faecalibaculum rodentium* and its human homologue protect from intestinal tumour growth. *Nat Microbiol.* 2020;5:511–24.
97. Dohlmán AB, Mendoza DA, Ding S, Gao M, Dressman H, Iliev ID, et al. The cancer microbiome atlas: a pan-cancer comparative analysis to distinguish tissue-resident microbiota from contaminants. *Cell Host Microbe.* 2021;29:281–98. e5.
98. Connolly JP, Kelly L. The physical biogeography of *Fusobacterium nucleatum* in health and disease. *mBio.* 2025;16:e02989–24.
99. Mukherjee PK, Wang H, Retuerto M, Zhang H, Burkey B, Ghannoum MA, et al. Bacteriome and mycobiome associations in oral tongue cancer. *Oncotarget.* 2017;8:127273.
100. Al-Hebshi NN, Nasher AT, Maryoud MY, Homeida HE, Chen T, Idris AM, et al. Inflammatory bacteriome featuring *Fusobacterium nucleatum* and *Pseudomonas aeruginosa* identified in association with oral squamous cell carcinoma. *Sci Rep.* 2017;7:1834.
101. Sajid M, Sharma P, Srivastava S, Hariprasad R, Singh H, Bharadwaj M. Alteration of oral bacteriome of smokeless tobacco users and their association with oral cancer. *Appl Microbiol Biotechnol.* 2023;107:4009–24.
102. Shay E, Sangwan N, Padmanabhan R, Lundy S, Burkey B, Eng C. Bacteriome and mycobiome and bacteriome-mycobiome interactions in head and neck squamous cell carcinoma. *Oncotarget.* 2020;11:2375.
103. Srivastava A, Mishra S, Garg PK, Dubey AK, Deo S, Verma D. Comparative and analytical characterization of the oral bacteriome of smokeless tobacco users with oral squamous cell carcinoma. *Appl Microbiol Biotechnol.* 2022;106:4115–28.
104. Gopinath D, Kunnath Menon R, Chun Wie C, Banerjee M, Panda S, Mandal D, et al. Salivary bacterial shifts in oral leukoplakia resemble the dysbiotic oral cancer bacteriome. *J Oral Microbiol.* 2021;13:1857998.
105. Chowdhury MR, Hone KGMS, Prevost K, Balthazar P, Avino M, Arguin M, et al. Optimizing fecal occult blood test (FOBT) colorectal cancer screening using gut bacteriome as a biomarker. *Clin Colorectal Cancer.* 2024;23:22–34. e2.
106. Grion BAR, Fonseca PLC, Kato RB, Garcia GJY, Vaz ABM, Jiménez BN, et al. Identification of taxonomic changes in the fecal bacteriome associated with colorectal polyps and cancer: potential biomarkers for early diagnosis. *Front Microbiol.* 2024;14:1292490.
107. Ha MV, McCormick TS, Salem I, Al-Shakhshir H, Ghannoum MA, Carroll BT. Skin and gut microbial associations with squamous cell carcinoma in solid organ transplant recipients. *Arch Dermatol Res.* 2023;315:2709–13.
108. Ho SX, Law J-H, Png C-W, Alberts R, Zhang Y, Chu JH, et al. Alterations in colorectal cancer virome and its persistence after surgery. *Sci Rep.* 2024;14:2819.
109. Kaur K, Saxena A, Debnath I, O'Brien JL, Ajami NJ, Auchtung TA, et al. Antibiotic-mediated bacteriome depletion in *ApcMin/+* mice is associated with reduction in mucus-producing goblet cells and increased colorectal cancer progression. *Cancer Med.* 2018;7:2003–12.
110. Png C-W, Chua Y-K, Law J-H, Zhang Y, Tan K-K. Alterations in co-abundant bacteriome in colorectal cancer and its persistence after surgery: a pilot study. *Sci Rep.* 2022;12:9829.
111. Ramaboli M, Ocivirk S, Khan Mirzaei M, Eberhart B, Valdivia-García M, Metwaly A, et al. Diet changes due to urbanization in South Africa are linked to microbiome and metabolome signatures of Westernization and colorectal cancer. *Nat Commun.* 2024;15:3379.
112. Shastry RP, Ghate SD, Hameed A, Rao RSP, Bhandary YP, Shetty R. Emergence of rare and low abundant anaerobic gut Firmicutes is associated with a significant downfall of *Klebsiella* in human colon cancer. *Microb Pathogenesis.* 2024;193:106726.
113. Marongiu L, Landry JJ, Rausch T, Abba ML, Delecluse S, Delecluse HJ, et al. Metagenomic analysis of primary colorectal carcinomas and their metastases identifies potential microbial risk factors. *Mol Oncol.* 2021;15:3363–84.
114. Martínez-Guyn K, Leone V, Chang EB. Regional diversity of the gastrointestinal microbiome. *Cell host microbe.* 2019;26:314–24.

115. Valciukiene J, Strupas K, Poskus T. Tissue vs. fecal-derived bacterial dysbiosis in precancerous colorectal lesions: a systematic review. *Cancers*. 2023;15:1602.
116. Wang Q, Shao X, Zhang Y, Zhu M, Wang FX, Mu J, et al. Role of tumor micro-environment in cancer progression and therapeutic strategy. *Cancer Med*. 2023;12:11149–65.
117. Johnson JS, Spakowicz DJ, Hong B-Y, Petersen LM, Demkowicz P, Chen L, et al. Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. *Nat Commun*. 2019;10:5029.
118. Tran Q, Pham D-T, Phan V. Using 16S rRNA gene as marker to detect unknown bacteria in microbial communities. *BMC Bioinforma*. 2017;18:155–61.
119. Patel JB. 16S rRNA gene sequencing for bacterial pathogen identification in the clinical laboratory. *Mol Diagnosis*. 2001;6:313–21.
120. Woo PC, Lau SK, Teng JL, Tse H, Yuen K-Y. Then and now: use of 16S rDNA gene sequencing for bacterial identification and discovery of novel bacteria in clinical microbiology laboratories. *Clin Microbiol Infect*. 2008;14:908–34.
121. Drancourt M, Berger P, Raoult D. Systematic 16S rRNA gene sequencing of atypical clinical isolates identified 27 new bacterial species associated with humans. *J Clin Microbiol*. 2004;42:2197–202.
122. Pérez-Cobas AE, Gomez-Valero L, Buchrieser C. Metagenomic approaches in microbial ecology: an update on whole-genome and marker gene sequencing analyses. *Microb Genom*. 2020;6:e000409.
123. Quince C, Walker AW, Simpson JT, Loman NJ, Segata N. Shotgun metagenomics, from sampling to analysis. *Nat Biotechnol*. 2017;35:833–44.
124. Durazzi F, Sala C, Castellani G, Manfreda G, Remondini D, De Cesare A. Comparison between 16S rRNA and shotgun sequencing data for the taxonomic characterization of the gut microbiota. *Sci Rep*. 2021;11:3030.
125. Lashkaripour A, McIntyre DP, Calhoun SG, Krauth K, Densmore DM, Fordyce PM. Design automation of microfluidic single and double emulsion droplets with machine learning. *Nat Commun*. 2024;15:83.
126. Xi H-D, Zheng H, Guo W, Gañán-Calvo AM, Ai Y, Tsao C-W, et al. Active droplet sorting in microfluidics: a review. *Lab a Chip*. 2017;17:751–71.
127. Baret J-C, Miller OJ, Taly V, Ryckelynck M, El-Harrak A, Frenz L, et al. Fluorescence-activated droplet sorting (FADS): efficient microfluidic cell sorting based on enzymatic activity. *Lab a Chip*. 2009;9:1850–8.
128. Brower KK, Carswell-Crumpton C, Klemm S, Cruz B, Kim G, Calhoun SG, et al. Double emulsion flow cytometry with high-throughput single droplet isolation and nucleic acid recovery. *Lab Chip*. 2020;20:2062–74.
129. Sukovich DJ, Kim SC, Ahmed N, Abate AR. Bulk double emulsification for flow cytometric analysis of microfluidic droplets. *Analyst*. 2017;142:4618–22.
130. Brower KK, Khariton M, Suzuki PH, Still C, Kim G, Calhoun SG, et al. Double emulsion picoreactors for high-throughput single-cell encapsulation and phenotyping via FACS. *Anal Chem*. 2020;92:13262–70.
131. Cole RH, Tang S-Y, Siltanen CA, Shahi P, Zhang JQ, Poust S, et al. Printed droplet microfluidics for on demand dispensing of picoliter droplets and cells. *Proc Natl Acad Sci*. 2017;114:8728–33.
132. Lan F, Haliburton JR, Yuan A, Abate AR. Droplet barcoding for massively parallel single-molecule deep sequencing. *Nat Commun*. 2016;7:11784.
133. Klein AM, Mazutis L, Akartuna I, Tallapragada N, Veres A, Li V, et al. Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell*. 2015;161:1187–201.
134. Guo MT, Rotem A, Heyman JA, Weitz DA. Droplet microfluidics for high-throughput biological assays. *Lab a Chip*. 2012;12:2146–55.
135. Zheng W, Zhao S, Yin Y, Zhang H, Needham DM, Evans ED, et al. High-throughput, single-microbe genomics with strain resolution, applied to a human gut microbiome. *Science*. 2022;376:eabm1483.
136. Jin J, Yamamoto R, Takeuchi T, Cui G, Miyauchi E, Hojo N, et al. High-throughput identification and quantification of single bacterial cells in the microbiota. *Nat Commun*. 2022;13:863.
137. Ohan J, Pelle B, Nath P, Huang J, Hovde B, Vuyisich M, et al. High-throughput phenotyping of cell-to-cell interactions in gel microdroplet pico-cultures. *Bio-Techniques*. 2019;66:218–24.
138. Watterson WJ, Tanyeri M, Watson AR, Cham CM, Shan Y, Chang EB, et al. Droplet-based high-throughput cultivation for accurate screening of antibiotic resistant gut microbes. *Elife*. 2020;9:e56998.
139. Xu Z, Wang Y, Sheng K, Rosenthal R, Liu N, Hua X, et al. Droplet-based high-throughput single microbe RNA sequencing by smRandom-seq. *Nat Commun*. 2023;14:5130.
140. Armstrong G, Rahman G, Martino C, McDonald D, Gonzalez A, Mishne G, et al. Applications and comparison of dimensionality reduction methods for microbiome data. *Front Bioinforma*. 2022;2:821861.
141. Liu W, Zhang R, Shu R, Yu J, Li H, Long H, et al. Study of the relationship between microbiome and colorectal cancer susceptibility using 16SrRNA sequencing. *BioMed Res Int*. 2020;2020:7828392.
142. Anderson MJ. A new method for non-parametric multivariate analysis of variance. *Austral Ecol*. 2001;26:32–46.
143. McArdle BH, Anderson MJ. Fitting multivariate models to community data: a comment on distance-based redundancy analysis. *Ecology*. 2001;82:290–7.
144. Abe K, Hirayama M, Ohno K, Shimamura T. Hierarchical non-negative matrix factorization using clinical information for microbial communities. *BMC Genom*. 2021;22:1–17.
145. Hosoda S, Nishijima S, Fukunaga T, Hattori M, Hamada M. Revealing the microbial assemblage structure in the human gut microbiome using latent Dirichlet allocation. *Microbiome*. 2020;8:1–12.
146. Zeng T, Yu X, Chen Z. Applying artificial intelligence in the microbiome for gastrointestinal diseases: a review. *J Gastroenterol Hepatol*. 2021;36:832–40.
147. Kumar P, Sinha R, Shukla P. Artificial intelligence and synthetic biology approaches for human gut microbiome. *Crit Rev Food Sci Nutr*. 2022;62:2103–21.
148. Lim H, Cankara F, Tsai C-J, Keskin O, Nussinov R, Gursay A. Artificial intelligence approaches to human-microbiome protein-protein interactions. *Curr Opin Struct Biol*. 2022;73:102328.
149. Fonseca DC, da Rocha Fernandes G, Waitzberg DL. Artificial intelligence and human microbiome: a brief narrative review. *Clin Nutr Open Sci*. 2025;159:134–42.
150. Liang Q, Bible PW, Liu Y, Zou B, Wei L. DeepMicrobes: taxonomic classification for metagenomics with deep learning. *NAR Genom Bioinforma*. 2020;2:lqaa009.
151. Lin Y, Wang G, Yu J, Sung JJ. Artificial intelligence and metagenomics in intestinal diseases. *J Gastroenterol Hepatol*. 2021;36:841–7.
152. Mineeva O, Rojas-Carulla M, Ley RE, Schölkopf B, Youngblut ND. DeepMAED: evaluating the quality of metagenomic assemblies. *Bioinformatics*. 2020;36:3011–7.
153. Arango-Argoty G, Garner E, Pruden A, Heath LS, Vikesland P, Zhang L. DeepARG: a deep learning approach for predicting antibiotic resistance genes from metagenomic data. *Microbiome*. 2018;6:1–15.
154. Tang R, Luo R, Tang S, Song H, Chen X. Machine learning in predicting antimicrobial resistance: a systematic review and meta-analysis. *Int J Antimicrob Agents*. 2022;60:106684.
155. Radhakrishnan SK, Nath D, Russ D, Merodio LB, Lad P, Daisi FK, et al. Machine learning-based identification of proteomic markers in colorectal cancer using UK Biobank data. *Front Oncol*. 2025;14:1505675.
156. Novielli P, Romano D, Magarelli M, Bitonto PD, Diacono D, Chiatante A, et al. Explainable artificial intelligence for microbiome data analysis in colorectal cancer biomarker identification. *Front Microbiol*. 2024;15:1348974.
157. Xu W, Wang T, Wang N, Zhang H, Zha Y, Ji L, et al. Artificial intelligence-enabled microbiome-based diagnosis models for a broad spectrum of cancer types. *Brief Bioinforma*. 2023;24:bbad178.
158. Zeng M, Zhang F, Wu F-X, Li Y, Wang J, Li M. Protein-protein interaction site prediction through combining local and global features with deep neural networks. *Bioinformatics*. 2020;36:1114–20.
159. Chen H, Shen J, Wang L, Song J. A framework towards data analytics on host-pathogen protein-protein interactions. *J Ambient Intell Humanized Comput*. 2020;11:4667–79.
160. Ma Y, Jiang H. NinimHMDA: neural integration of neighborhood information on a multiplex heterogeneous network for multiple types of human microbe-disease association. *Bioinformatics*. 2020;36:5665–71.
161. Jiang C, Feng J, Shan B, Chen Q, Yang J, Wang G, et al. Predicting microbe-disease associations via graph neural network and contrastive learning. *Front Microbiol*. 2024;15:1483983.
162. Karkera N, Acharya S, Palaniappan SK. Leveraging pre-trained language models for mining microbiome-disease relationships. *BMC Bioinforma*. 2023;24:290.
163. Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer J Clin*. 2024;74:229–63.
164. You JS, Jones PA. Cancer genetics and epigenetics: two sides of the same coin?. *Cancer Cell*. 2012;22:9–20.
165. Esteller M. Epigenetics in cancer. *N Engl J Med*. 2008;358:1148–59.
166. Cullin N, Azevedo Antunes C, Straussman R, Stein-Thoeringer CK, Elinav E. Microbiome and cancer. *Cancer Cell*. 2021;39:1317–41.
167. Farinha P, Gascoyne RD. *Helicobacter pylori* and MALT lymphoma. *Gastroenterology*. 2005;128:1579–605.
168. Xuan S-Y, Xin Y-N, Chen A-J, Dong Q-J, Qiang X, Li N, et al. Association between the presence of *H pylori* in the liver and hepatocellular carcinoma: a meta-analysis. *World J Gastroenterol*. 2008;14:307.
169. Huang S-T, Chen J, Lian L-Y, Cai H-H, Zeng H-S, Zheng M, et al. Intratumoral levels and prognostic significance of *Fusobacterium nucleatum* in cervical carcinoma. *Aging (Albany NY)*. 2020;12:23337.
170. Wang X, Jia Y, Wen L, Mu W, Wu X, Liu T, et al. *Porphyromonas gingivalis* promotes colorectal carcinoma by activating the hematopoietic NLRP3 inflammasome. *Cancer Res*. 2021;81:2745–59.
171. Gnanasekaran J, Binder Gallimidi A, Saba E, Pandi K, Eli Berchoer L, Hermano E, et al. Intracellular *Porphyromonas gingivalis* promotes the tumorigenic behavior of pancreatic carcinoma cells. *Cancers*. 2020;12:2331.

172. Iida N, Mizukoshi E, Yamashita T, Yutani M, Seishima J, Wang Z, et al. Chronic liver disease enables gut *Enterococcus faecalis* colonization to promote liver carcinogenesis. *Nat cancer*. 2021;2:1039–54.
173. Veziant J, Gagnière J, Jouberton E, Bonnin V, Sauvanet P, Pezet D, et al. Association of colorectal cancer with pathogenic *Escherichia coli*: focus on mechanisms using optical imaging. *World J Clin Oncol*. 2016;7:293.
174. Laroumagne S, Salinas-Pineda A, Hermant C, Murris M, Gourraud P, Do C, et al. Incidence and characteristics of bronchial colonisation in patient with lung cancer: a retrospective study of 388 cases. *Rev Mal Respir*. 2011;28:328–35.
175. Liss MA, White JR, Goros M, Gelfond J, Leach R, Johnson-Pais T, et al. Metabolic biosynthesis pathways identified from fecal microbiome associated with prostate cancer. *Eur Urol*. 2018;74:575–82.
176. Apostolou P, Tsantsaridou A, Papasotiriou I, Toloudi M, Chatziioannou M, Giannoulis G. Bacterial and fungal microflora in surgically removed lung cancer samples. *J Cardiothorac Surg*. 2011;6:1–5.
177. Urbaniak C, Gloor GB, Brackstone M, Scott L, Tangney M, Reid G. The microbiota of breast tissue and its association with breast cancer. *Appl Environ Microbiol*. 2016;82:5039–48.
178. Xuan C, Shamonki JM, Chung A, DiNome ML, Chung M, Sieling PA, et al. Microbial dysbiosis is associated with human breast cancer. *PLoS One*. 2014;9:e83744.
179. Tewari M, Mishra RR, Shukla HS. *Salmonella typhi* and gallbladder cancer: report from an endemic region. *Hepatobiliary Pancreat Dis Int*. 2010;9:524–30.
180. Haghi F, Goli E, Mirzaei B, Zeighami H. The association between fecal enterotoxigenic *B. fragilis* with colorectal cancer. *BMC Cancer*. 2019;19:1–4.
181. He Z, Gharaibeh RZ, Newsome RC, Pope JL, Dougherty MW, Tomkovich S, et al. *Campylobacter jejuni* promotes colorectal tumorigenesis through the action of cytolethal distending toxin. *Gut*. 2019;68:289–300.
182. Zhang Y, Fan K, Li L, He J, Sun Y. *Peptostreptococcus anaerobius* is a potential diagnostic biomarker of colorectal cancer. *Cell Mol Biol*. 2023;69:244–7.
183. Cao Y, Oh J, Xue M, Huh WJ, Wang J, Gonzalez-Hernandez JA, et al. Commensal microbiota from patients with inflammatory bowel disease produce genotoxic metabolites. *Science*. 2022;378:eabm3233.
184. Chaturvedi AK, Gaydos CA, Agreda P, Holden JP, Chatterjee N, Goedert JJ, et al. *Chlamydia pneumoniae* infection and risk for lung cancer. *Cancer Epidemiol Biomark Prev*. 2010;19:1498–505.
185. Mughini-Gras L, Schaapveld M, Kramers J, Mooij S, Neeffjes-Borst EA, Pelt WV, et al. Increased colon cancer risk after severe *Salmonella* infection. *PLoS One*. 2018;13:e0189721.

AUTHOR CONTRIBUTIONS

I.A. and A.K. conceptualized and designed the manuscript. Both I.A. and A.K. conducted the literature review. The manuscript draft, including figures and tables,

was prepared by I.A., J.B., P.C., and A.K. J.B. and P.C. made significant contributions by reviewing, editing, and providing suggestions to enhance the quality of the manuscript. All authors have read and approved the final version of the manuscript for publication.

FUNDING

This work was primarily supported by internal funding from Los Alamos National Laboratory, specifically the Laboratory Directed Research and Development (LDRD) grant 20240092DR awarded to A.K. I.A. is supported by the LDRD Frederick Reines Fellowship.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Correspondence and requests for materials should be addressed to Anand Kumar.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

This is a U.S. Government work and not under copyright protection in the US; foreign copyright protection may apply 2026, modified publication 2026