Medical Oncology (2026) 43:72
https://doi.org/10.1007/s12032-025-03170-w

ORIGINAL PAPER

Microbiota relationship between breast, colorectal, and lung cancer types

Ebru Kanimdan[1,2,3] · Caspar Bundgaard-Nielsen[4] · Vildan Betul Yenigun[2,3] · Burcu Gul[5] · Zuhal Gucin[5] · Hasan Enis Komurcu[5] · Sahende Elagoz[5] · Ozge Pasin[6] · Suzette Sørensen[4,7] · Ercan Arican[8]

Received: 15 August 2025 / Accepted: 29 November 2025 / Published online: 26 December 2025

Abstract

Colorectal cancer (CRC), breast cancer (BC), and lung cancer (LC) are among the most common and deadly malignancies worldwide. In recent years, the role of tumor-associated microbiota in the initiation and progression of cancer has attracted increasing attention. However, studies examining different cancer types together with their clinical stages remain limited. The aim of this study was to compare the composition of tumor tissue microbiota across three cancer types and different disease stages. For this purpose, DNA was isolated from formalin-fixed paraffin-embedded (FFPE) tumor samples, and sequencing of the 16 S rRNA gene region was performed. Microbial profiles were analyzed at the phylum and genus levels, and diversity indices were compared between groups. In stage III LC samples, the phylum Cyanobacteria was found to be markedly more abundant compared to other stages. In CRC, the phylum Firmicutes was most abundant in stage III tumors, showing higher levels than in early stages. In metastatic cases (advanced stages), both the operational taxonomic unit (OTU) count and Shannon diversity index were found to be highly similar to those of LC control tissues but markedly lower than those of the CRC control group. Across all cancer types, *Actinobacteria*, *Cyanobacteria*, *Proteobacteria*, and *Firmicutes* were identified as the dominant phyla, though their relative abundances varied by cancer type and stage. The findings suggest that tumor-associated microbiota composition exhibits distinct signatures depending on tumor type and progression stage. Similarities observed between certain cancer types and stages may indicate the presence of shared microbial patterns within the tumor microenvironment. Such microbial signatures may serve as potential biomarkers for diagnosis, prognosis, or therapeutic targeting.

Keywords Colorectal cancer · Microbiota · Lung cancer · Breast cancer · Full-length 16S rRNA gene · FFPE

Introduction

The human microbiota is increasingly recognized as a key factor in maintaining health and contributing to disease through its roles in immune regulation, metabolism, and epithelial homeostasis [1]. This complex ecosystem comprising bacteria, viruses, fungi, and other microbes inhabits diverse body sites, including the gut, lungs, skin, and even breast tissue [2, 3].

✉ Ercan Arican
earican@istanbul.edu.tr

1 Institute of Graduate Studies in Science, Molecular Biology and Genetics Department, Istanbul University, Istanbul, Turkey

2 Vocational School of Health Services, Bezmialem Vakif University, Istanbul, Turkey

3 Faculty of Medicine, Department of Medical Biochemistry, Bezmialem Vakif University, Istanbul, Turkey

4 Centre for Clinical Research, North Denmark Regional Hospital, Aalborg, Denmark

5 Faculty of Medicine, Department of Medical Pathology, Bezmialem Vakif University, Istanbul, Turkey

6 Department of Biostatistics, Hamidiye Medical Faculty, Health Sciences University, Istanbul, Turkey

7 Department of Clinical Medicine, Aalborg University, Aalborg, Denmark

8 Science Faculty, Molecular Biology and Genetics Department, Istanbul University, Istanbul, Turkey

According to the International Agency for Research on Cancer (IARC), cancer is projected to become the leading cause of death globally by 2040, with an estimated 29.5 million new cases and 16.4 million deaths [4]. Factors such as late diagnosis, molecular heterogeneity, limited treatment accessibility, and therapeutic resistance complicate cancer management and outcomes [5]. This has prompted a growing interest in the tumor microenvironment, including the role of the microbiota.

Microbial dysbiosis, an imbalance in microbial composition has been linked to cancer development and progression via chronic inflammation, immune dysregulation, and production of genotoxic compounds [6–8]. While colorectal cancer (CRC) has the most well-established connection to microbiota alterations, emerging evidence suggests that lung and breast cancers also involve microbial components. In CRC, pro-inflammatory genera such as *Fusobacterium*, *Peptostreptococcus*, and *Bacteroides* are frequently overrepresented and have been implicated in tumorigenesis through immune modulation and barrier disruption [9]. In lung cancer (LC), changes in the pulmonary microbiota may promote tumor-supportive inflammation and alter immune responses [10]. Likewise, recent studies have demonstrated that breast tissue harbors a unique microbial signature, challenging the long held assumption of sterility and suggesting a potential role in oncogenesis.

Despite increasing research, few studies have systematically compared microbiota profiles across different cancer types or disease stages. Comparative analyses across malignancies remain limited, and how microbial communities evolve during tumor progression is not well understood. This knowledge gap restricts the clinical translation of microbial markers for cancer diagnosis, prognosis, or therapy.

In this study, we analyzed the microbiota composition of Formalin-fixed, paraffin-embedded (FFPE) tumor tissue from patients ($n=220$) with colorectal, lung, and breast cancers, aiming to identify cancer type specific and stage associated microbial signatures. This integrated approach was designed to provide novel insights into the tumor microenvironment and to evaluate the potential of microbial features as diagnostic or prognostic biomarkers across multiple cancer types.

Materıals and methods

Sample collection and grouping

FFPE tissue samples were obtained from the archives of the Department of Pathology at Bezmialem Vakif University, following ethical approval (Approval No: 2022/85). Archived samples from the last five years (2017–2022) were selected, representing patients with breast, colorectal, and lung cancers (Tables 1, 2 and 3). The total sample size was 233, comprising lung cancer (LC, $n=60$, control $n=9$), colorectal cancer (CRC, $n=70$, control $n=11$), breast cancer (BC, $n=60$, control $n=10$) cases, as well as metastatic tumor samples ($n=13$). Paraffin only blocks processed identically to tissue containing FFPE samples were periodically included as negative controls to assess reagent and environment derived contaminationa control paraffin block containing no embedded tissue for contamination assessment. Each FFPE block, 1 section of 4 µm thickness was obtained for hematoxylin-eosin (H&E) staining and 4 sections of 10 µm were collected for genomic DNA extraction. A fresh microtome blade was used for each sample, and the surface was decontaminated with RNase-away and alcohol to prevent cross-contamination. The study was structured into 16 distinct groups based on cancer type and stage, including control, non-tumor tissues and adjacent tissues.

Hematological parameters

In this retrospective study, complete blood count parameters of patients diagnosed with lung, breast, and colorectal cancer, as well as those without a cancer diagnosis, were obtained from archived laboratory records in the hospital information management system. All hematological data were derived from measurements performed using automated hematology analyzers routinely employed in clinical practice. Data were retrieved retrospectively from the

Table 1 Sex distribution by cancer type and disease stage in lung and colorectal cancer patients

	Female		Male		Total		
Group	*N*	%	*N*	%	*N*	%	*p* value
Control	4	50.0	4	50.0	8	100.0	*P*=0,842
LC Stage 1	5	33.3	10	66.7	15	100.0	
LC Stage 2	5	33.3	10	66.7	15	100.0	
LC Stage 3	6	33.3	12	66.7	18	100.0	
CRC Control	3	27.3	8	72.7	11	100.0	*P*=0,921
CRC Stage 1	4	40.0	6	60.0	10	100.0	
CRC Stage 2	11	36.7	19	63.3	30	100.0	
CRC Stage 3	12	40.0	18	60.0	30	100.0	

Table 2 Descriptive statistics of age across study groups

Group	Min	Max	Median	SD	Mean	*N*	Missing	*p* value
LC Control	47	68	59.5	7.4	60.5	8	0	*P*=0,460
LC Stage1	54	77	67.0	7.2	65.6	15	0	
LC Stage2	52	72	67.0	6.7	64.5	15	0	
LC Stage3	37	81	66.0	10.2	65.8	18	0	
CRC Control	40	82	60.0	12.7	64.0	11	0	*P*=0,656
CRC Stage1	45	83	68.5	11.8	66.9	10	0	
CRC Stage2	33	82	63.5	11.9	62.7	30	0	
CRC Stage3	35	88	59.5	13.1	61.3	30	0	
BC Control	18	48	30.0	8.5	31.2	9	0	*P*<0,001
BC Stage1	38	88	57.0	13.5	57.4	26	0	
BC Stage2	24	79	50.0	16.5	51.6	22	0	
BC Stage3	33	86	51.0	17.2	54.8	10	0	

electronic record system, with patient identifiers removed, and all records anonymized prior to analysis (Suppl. Tables 1–8).

DNA extraction and 16 S rRNA gene sequence

DNA extraction was performed with a QIAamp DNA FFPE Advanced (Qiagen) according to the manufacturer's instructions. Quantitation of DNA was measured using NanoDrop 2000 (Thermo Scientific). Full-length 16 S rRNA gene amplification was performed using the 16 S Barcoding Kit 1–24 (SQK-16S024, Oxford Nanopore Technologies) according to the manufacturer's protocol (Suppl. Table 1 and Suppl. Figure 1).

16 S rRNA gene sequence analysis

Full-length 16 S rRNA gene (V1-V9) amplification was performed using universal bacterial primers 27 F (5′-AGAGTTTGATCMTGGCTCAG-3′) and 1492R (5′-CGGTTACCTTGTTACGACTT-3′), targeting conserved regions flanking the entire 16 S gene. Library preparation was carried out using the 16 S Barcoding Kit 1–24 (SQK-16S024, Oxford Nanopore Technologies), and sequencing was conducted on a MinION platform utilizing R9 flow cell chemistry. Each sample was sequenced to a minimum read depth of 10,000 reads, ensuring sufficient taxonomic coverage for downstream microbiota profiling. Flow cells were run under default MinKNOW settings, and barcoded amplicons were demultiplexed prior to analysis. Raw sequencing data were processed using EMU (Efficient Microbial-Community Profiling Utility), an alignment-based taxonomic profiling tool optimized for full-length 16 S rRNA gene analysis. Reads were filtered based on a minimum quality threshold of 0 and a length range of 1250–1750 base pairs to ensure high-confidence full-length 16 S amplicons.

Taxonomic classification was performed using the SILVA reference database (version 138). EMU output provided genus-level abundance tables, which were used for downstream diversity analyses and taxonomic comparisons across cancer types and stages. Both the taxonomic classification table (equivalent to tax_table in the phyloseq framework) and the OTU abundance matrix (equivalent to otu_table, expressed as relative abundance) were generated during the data analysis using Python. Sequencing was performed through an external service provider. Only processed FASTQ files were provided, and raw per-read quality metrics (including Q-score distributions and read length statistics) were not available. Therefore, the quality control description reflects the standard preprocessing procedures reported by the provider.

Statistical analysis

Descriptive statistics for categorical variables were presented as frequency and percentage, while those for continuous variables were given as mean, standard deviation, median, minimum, and maximum. The normality of continuous variables was assessed using the Shapiro-Wilk test, and the homogeneity of variances was evaluated using the Levene test. One-way analysis of variance (ANOVA) and the Brown-Forsythe test were used for comparing means among more than two independent groups. For post hoc multiple comparisons, Tukey's and Dunnett's T3 tests were applied. The Kruskal-Wallis test was used for comparing medians across more than two independent groups, with the Dunn test employed for post hoc analysis. For the comparison of proportions in categorical variables between groups, Pearson's chi-square test and the Fisher-Freeman-Halton test were used. A *p-value* of <0.05 was considered statistically significant. All statistical analyses were performed using the SPSS software package (version 28).

Table 3 Distribution of histopathological tumor types across clinical stages of lung cancer (LC1–LC3), colorectal cancer (CRC1–CRC3), and breast cancer (BC1–BC3) and their corresponding control groups

Diagnosis	LC1	LC2	LC3	Control
NATs	0 (0.0%)	0 (0.0%)	0 (0.0%)	9 (100.0%)
Combined SCC + adenocarcinoma	0 (0.0%)	0 (0.0%)	1 (5.5%)	0 (0.0%)
NET	0 (0.0%)	1 (4.8%)	0 (0.0%)	0 (0.0%)
NSCLC	0 (0.0%)	0 (0.0%)	1 (5.6%)	0 (0.0%)
SCC	8 (38.1%)	7 (33.3%)	9 (50.0%)	0 (0.0%)
Adenocarcinoma	13 (61.9%)	13 (61.9%)	5 (27.7%)	0 (0.0%)
Carcinoma	0 (0.0%)	0 (0.0%)	1 (5.6%)	0 (0.0%)
Carcinosarcoma	0 (0.0%)	0 (0.0%)	1 (5.6%)	0 (0.0%)
Total (n)	21	21	18	9
Diagnosis	**CRC1**	**CRC2**	**CRC3**	**Control**
adenocarcinoma	10 (100%)	29 (96.7%)	27 (90%)	0 (0.0%)
diverticulosis	0 (0.0%)	0 (0.0%)	0 (0.0%)	8 (72.7%)
Inactive chronic antral gastritis	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (9.1%)
mucinous adenocarcinoma	0 (0.0%)	1 (3.3)	2 (6.7%)	0 (0.0%)
Rectosigmoid region wo residual tumor	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (9.1%)
SRCC	0 (0.0%)	0 (0.0%)	1 (3.3%)	0 (0.0%)
Subepithelial inflammation	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (9.1%)
Total	10 (100%)	30 (100%)	30 (100%)	11 (100%)
Diagnosis	**BC1**	**BC2**	**BC3**	**Control**
Invasive carcinoma NST + DCIS	1 (3.8%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Carcinoma in situ	2 (7.7%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
DCIS	1 (3.8%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
fibroadenoma	0 (0.0%)	0 (0.0%)	0 (0.0%)	10 (100%)
Invasive carcinoma	0 (0.0%)	1 (4.2%)	1 (10%)	0 (0.0%)
Invasive carcinoma NST	17 (65.4%)	16 (66.7%)	8 (80%)	0 (0.0%)
Invasive lobular carcinoma	4 (15.4%)	5 (20.8%)	1 (10%)	0 (0.0%)
Mixed type carcinoma	1 (3.8%)	1 (4.2%)	0 (0.0%)	0 (0.0%)
Mucinous carcinoma	0 (0.0%)	1 (4.2%)	0 (0.0%)	0 (0.0%)
Total (n)	26	24	10	10

Results

Across all analyses, statistically significant differences in hematological parameters were identified both within and between lung, colorectal, and breast cancer cohorts across clinical stages, as well as in comparisons with their respective control groups. These results underscore the presence of distinct, stage-dependent, and cancer type specific hematological profiles (Suppl. Tables 2–8). Significant differences in WBC, NEUT%, LYMPH%, MONO%, and NEUT# were observed when comparing the same stage across lung, colorectal, and breast cancers. Stage 1 lung cancer had higher WBC and NEUT# than colorectal and breast cancer. Breast cancer generally showed lower NEUT% and LYMPH% compared to the other cancers (Suppl. Table 8).

Table 4 Primary tumor origins and histological heterogeneity of the metastatic tissue samples

Primary Cancer Type	Histological Subtype	*N*	%
Breast	Breast Carcinoma	2	11.8
Colorectal	Colorectal Adenocarcinoma	3	17.6
Lung	Lung Adenocarcinoma	5	29.4
	Lung Carcinoma	1	5.9
	Lung Carcinosarcoma	1	5.9
	Lung NSCLC	1	5.9
	Lung SCC	2	11.8
Prostate	Prostate Adenocarcinoma	1	5.9
Unknown	-	1	5.9
Total		17	100.0

The distribution of age and sex across the study groups, excluding the breast cancer group was homogeneous ($p > 0.05$) (Table 1) and (Table 2). BC patients in the control group (30.0 ± 8.5 years) were significantly younger ($p < 0.05$) than the other groups In addition, the molecular subtype distribution of breast cancer cases was summarized and provided in the supplementary Fig. 2.

Survival analysis revealed that LC patients in Stage 1 and Stage 2 had identical survival rates of 90.5%, whereas Stage 3 patients exhibited a slightly higher survival rate of 94.4%. In the CRC cohort, the survival rate was 90.0% in Stage 2 and increased to 93.3% in Stage 3. Among BC patients, the survival rate was 96.2% in Stage 1 and reached 100% in Stage 3.

The histopathological diagnoses and their respective percentages for patients across all study groups are presented in Table 3. Due to insufficient sample sizes ($n = 17$) for stage 4 (metastatic cancer) in all groups, the available samples were pooled across groups. The metastatic tissue samples were included, originating from breast carcinoma, colorectal adenocarcinoma, prostate adenocarcinoma, and various lung cancer subtypes (adenocarcinoma, carcinoma NOS, carcinosarcoma, NSCLC, and SCC), which collectively represented the largest group (Table 4.).

Following DNA extraction from each group, DNA concentrations were measured using a Qubit fluorometer (Suppl. Table 1). The Qubit measurement for the paraffin block was found to be very low (< 0.1 ng/µL) and samples were electrophoresed on a 2% agarose gel to obtain gel images (Suppl. Figure 1).

Microbial alpha diversity was calculated using the number of unique OTUs and the Shannon diversity index

to evaluate the richness and diversity of the microbiota within each group of samples. In lung cancer, OTU richness appeared slightly elevated in stages 1 and 3 compared to stage 2 and control tissues. Interestingly, Shannon diversity remained relatively consistent across stages 1 and 2 but showed a marked decrease in stage 3, suggesting a potential shift toward microbial dominance or dysbiosis in later stages. In colorectal cancer, OTU richness was slightly reduced in both stage 2 and stage 3 compared to control samples. However, a contrasting pattern was observed in Shannon diversity: both stage 2 and stage 3 exhibited markedly higher Shannon indices than the control group. For breast cancer, both OTU richness and Shannon diversity were notably higher in stage 3 compared to stage 1 Microbial diversity analysis of the metastasis group, in comparison to control groups of primary tumor types, demonstrated distinct patterns. Specifically, the OTU richness and Shannon diversity indices in metastatic samples were found to be comparable to those of the lung cancer control group, indicating a similar level of microbial composition in terms of both richness and evenness. In contrast, both diversity metrics were markedly lower than those observed in the colorectal cancer control group, suggesting a reduced microbial complexity in metastatic tissues relative to colorectal tumor-associated environments. The paraffin only block samples, used as a negative control in this study, exhibited a notably reduced microbial richness and diversity compared to all cancer and control tissue groups (Fig. 1).

The pCoA plot illustrates the beta diversity of microbial communities across different cancer types and stages. The analysis is based on Bray-Curtis dissimilarity metrics. The pCoA revealed clear differentiation in microbial community composition across cancer types and stages. Lung cancer samples demonstrated high intra-group variability, with stage 3 positioned distinctly in the upper right quadrant, indicating a divergent microbial profile compared to stages 1 and 2. Colorectal cancer samples also showed considerable dispersion, with stage 2 located in the lower right quadrant and stage 3 placed in the lower left, indicating substantial microbiota heterogeneity between tumor stages. In breast cancer, stage 1 and stage 3 were positioned far apart along both axes, suggesting significant microbial shifts during disease progression. The metastasis sample (stage 4) was located centrally, near the lung and breast early-stage tumor samples, and clearly separated from colorectal tumor and control samples. In this analysis, paraffin-embedded tissue was included as a negative control to evaluate the extent of potential background contamination and processing-related microbial signatures. The pCoA plot shows that the paraffin sample is positioned closely to the lung cancer control sample and is distinctly separated from all tumor-derived samples, suggesting minimal overlap in microbial composition with actual tumor tissues (Fig. 2).

The relative abundance of microbial taxa at the genus level varied markedly across cancer types, disease stages, and control samples. In lung cancer, early-stage samples (stage 1 and 2) exhibited high proportions of *Proteobacteria* (notably *Escherichia-Shigella* and *Aquabacterium*) and *Actinobacteria*, while stage 3 showed an increased dominance of Firmicutes, particularly *Ruminococcus*, along with

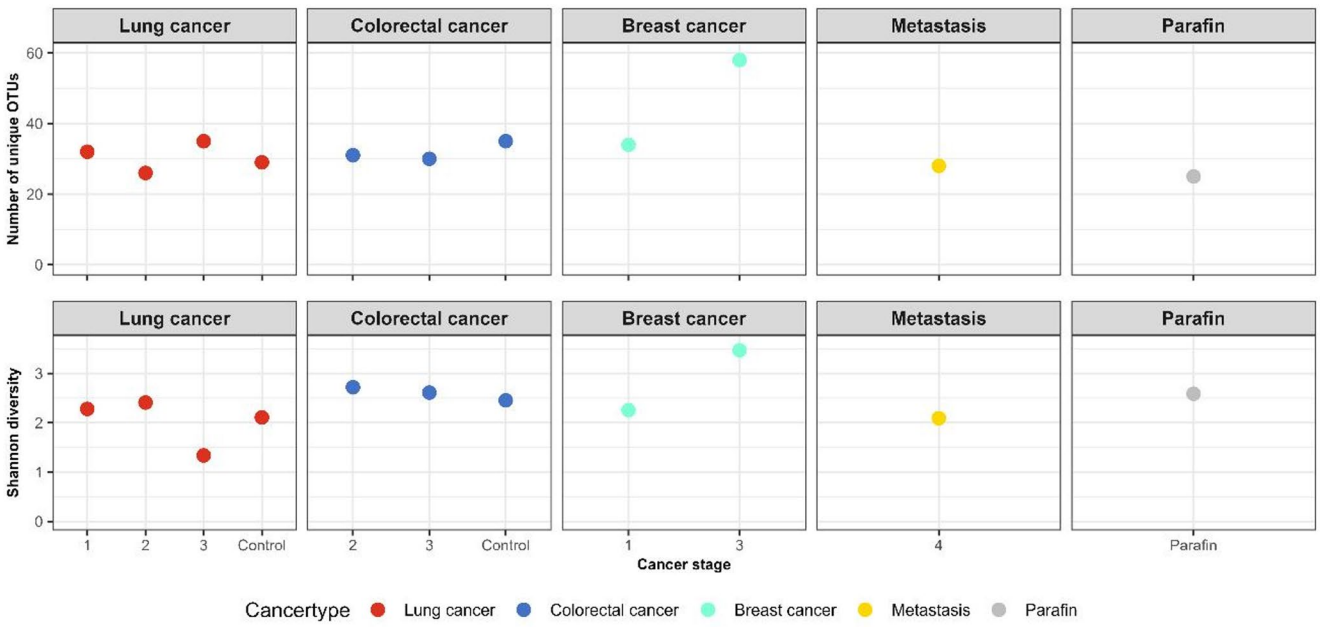

Fig. 1 Alpha diversity Analysis plot showing microbial community composition across different cancer types and stages. Colors represent cancer types, and shapes indicate cancer stages. Control and paraffin samples are included as references

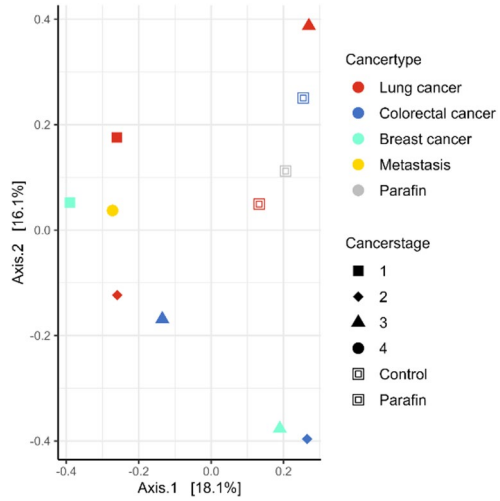

Fig. 2 Principal Coordinates Analysis (PCoA) illustrating microbial beta diversity across three cancer types. The plot is based on Bray-Curtis dissimilarity calculated from full-lenght 16 S rRNA sequencing data. Each point represents a pooled sample for a specific cancer group

a substantial reduction in *Actinobacteria*. Microbial composition was evaluated at the genus level across all cancer groups. In colorectal cancer (stages II and III) and the control group, prominent genera included *Abiotrophia*, *Altererythrobacter*, and *Anaerostipes*, with *Abiotrophia* being notably abundant in the control sample. In lung cancer samples (stages I–III), *Acinetobacter* and *Aquabacterium* were among the most dominant taxa, with *Aquabacterium* comprising 9.8% in stage II and 6.5% in stage I. Breast cancer samples (stages I and III) also showed high levels of *Aquabacterium*, particularly in stage I (21.5%). Anaerostipes was detected in the breast cancer stage III sample. These findings suggest that distinct microbiota profiles may exist for each cancer type, and certain bacterial genera may be preferentially associated with specific cancer types (Fig. 3).

Discussion

In lung cancer samples, OTU counts in stage 1 and stage 3 tissues were observed to be higher than those in stage 2 and control tissues. However, a marked decrease in Shannon diversity was noted in stage 3 samples, suggesting a state of dysbiosis in advanced stage lung cancer, characterized by reduced microbial evenness and dominance of certain taxa. The distinct clustering of stage 3 samples in the PCoA analysis further supports the presence of this dysbiotic profile. For LC patients, the survival rates for both Stage 1 and Stage 2 were 90.5%, showing almost identical outcomes, whereas LC Stage 3 patients demonstrated a slightly higher survival rate of 94.4%. *Cyanobacteria*-positive tissues, *Cyanobacteria* modulates the inflammatory response through microcystin production [11], promoting cell proliferation and carcinogenesis by suppressing CD36 levels and increasing PARP1 expression [12, 13]. The positive correlation between *Cyanobacteria* and $CD8^+$ and $PD\text{-}1^+$ T cell infiltration, as reported by Shen et al. (2024), suggests that this phylum may modulate the tumor immune microenvironment and, consequently, influence the potential responsiveness to PD-1/PD-L1-based immunotherapies [14]. When considered alongside this dominance observed in stage 3, our findings suggest that *Cyanobacteria* may not be merely a passive colonizer in advanced lung tumors, but rather an active microbial player potentially influencing tumor progression and immune dynamics. In the LC group, *Sphingomonas* (Proteobacteria) has been associated with modulation of the tumor immune microenvironment, particularly by enhancing IFN-γ mediated macrophage engagement and Th1 responses, which may contribute to both pro tumor and anti-tumor effects depending on the context [15]. Similarly, unclassified *Neisseriaceae* has been shown to exhibit an inverse correlation with PD-L1 expression in lung tumors, which may indicate the presence of a more favorable immune environment when this taxon is present [16]. These

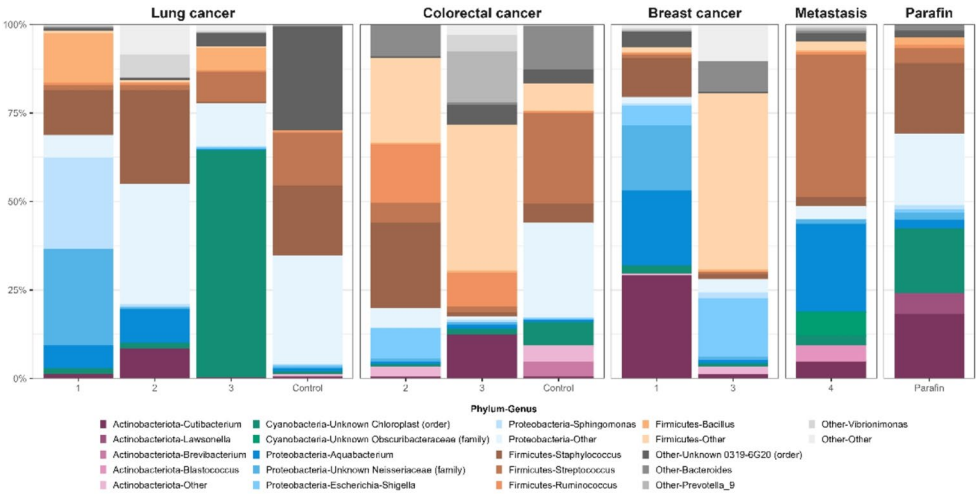

Fig. 3 Relative abundance of dominant bacterial genera identified in pooled samples from three different cancer types. Each bar represents a pooled sample from a specific cancer group/control. Distinct microbial compositions were observed across colorectal, lung, and breast cancer samples, with certain genera showing prominent abundance in specific cancer types

taxa are rarely dominant in CRC or breast cancer, further suggesting a lung specific microbiota profile with potential implications for tumor immune interactions. *Staphylococcus* and *Bacillus* (Firmicutes), which were enriched in early stage LC, have been associated with processes that contribute to tumorigenesis. For example, *Staphylococcus aureus* can promote HIF-1α–driven metastasis by inducing lactate production and pseudo-hypoxia in tumor cells [17].

Recent studies have demonstrated that microbial balance within the tumor microenvironment of CRC changes markedly in a stage dependent manner, and that these alterations can have significant impacts on disease progression, immune response, and clinical prognosis. In our study, among CRC patients, the survival rate was 90% in Stage 2 with a mortality rate of 10%, while Stage 3 showed a survival rate of 93.3% and a mortality rate of 6.7%. In addition, stage 2 and stage 3 CRC samples differed in microbial composition, with the phylum *Firmicutes* being notably more dominant in stage 3 tumors compared to other stages.

Recent studies have shown that microbial balance within the CRC tumor microenvironment is disrupted, and that certain members of the phylum Firmicutes may have notable effects on tumor development, immune responses, and clinical outcomes [18]. In the study conducted by Cai et al., it was demonstrated that α-diversity was significantly reduced and β-diversity differed markedly in stage IV CRC patients, and that as early lesions progressed to advanced-stage tumors, microbial diversity (alpha diversity) decreased, while the microbiome composition changed substantially compared to that of healthy individuals [19]. According to the survival analysis by Xu et al., microbiome profiles of 533 colorectal tumor tissues were examined, and two distinct prognostic microbial clusters were identified. In Cluster 1, tumors exhibited significantly lower Firmicutes and higher *Proteobacteria* and *Bacteroidetes* ratios compared to the other cluster, whereas Cluster 2 showed higher Firmicutes abundance and consequently a higher F/B ratio. In colorectal tissues, a high proportion of Proteobacteria was associated with significantly poorer patient survival ($p = 0.0025$), while patients with higher Firmicutes abundance had significantly better survival than those with lower Firmicutes levels ($p = 0.035$). Only Actinobacteria showed no effect on CRC patient survival ($p = 0.83$) [20]. These findings suggest that the presence of Firmicutes in the tumor microenvironment may be associated with a beneficial milieu that supports anti-tumor immunity and suppresses tumor growth. In addition, at the genus level within the microbiome, certain members of Firmicutes may have prognostic value. In a genus-level study by Li et al., co-culture with *Streptococcus* thermophilus or its conditioned medium was shown to reduce the proliferation of CRC cells in culture, and oral gavage with *S. thermophilus* significantly reduced tumorigenesis. It has been shown that patients with tumors rich in Firmicutes genera such as *Streptococcus* and *Lactobacillus* have better outcomes compared to those with lower levels of these bacteria [21]. In another study conducted by Li et al., β-galactosidase secreted by *S. thermophilus* was shown to inhibit cell proliferation, reduce colony formation, induce cell cycle arrest, and promote apoptosis of cultured CRC cells, thereby significantly reducing CRC cell proliferation [22]. In light of these findings from the literature, the detection of a higher relative abundance of the phylum Firmicutes in stage 3 colorectal tumor samples in our study, compared to other stages, may be considered an indicator of microbial restructuring associated with disease progression. Firmicutes (*Bacillus*) also plays a similar role in CRC, where it has been associated with epithelial barrier disruption and inflammation-driven carcinogenesis [17], underscoring shared microbial influences across different cancers.

The survival rate of BC Stage 1 patients was 96.2%, while Stage 3 showed a survival rate of 100%. Consistent with previous reports, our breast cancer analysis also revealed higher OTU counts and greater Shannon diversity in the stage 3 tumor sample compared to stage 1. While Proteobacteria (*Aquabacterium*, *Unknown Neisseriaceae*), *Actinobacteria (Cutibacterium)*, and Firmicutes (*Staphylococcus*) were more abundant in stage 1, stage 3 showed an increase in Firmicutes *(other)*, Proteobacteria *(Escherichia–Shigella)*, and other Bacteroides, along with a decrease in the Proteobacteria and Actinobacteria phyla that predominated in stage 1. Although the lower diversity observed in our early-stage breast cancer sample may initially appear contradictory to the literature, the marked increase in diversity in stage 3 suggests that, in general, breast tumors at advanced stages can harbor a wide variety of distinct bacterial taxa. In a comprehensive study conducted in 2020 that included 33 cancer types, breast cancer was reported to be among the cancer types with the richest and most diverse bacterial communities in terms of tumor microbiome [23]. German et al. (2023) demonstrated that the microbiota composition in tissue samples obtained from 76 breast cancer patients (primary tumor, adjacent normal tissue, and distant metastasis) varied according to tissue type [24]. In tumor-adjacent normal tissue, Cyanobacteria and Corynebacteriaceae increased compared to healthy tissue, while Lactobacillaceae, Acetobacteraceae, and Xanthomonadaceae decreased. Primary tumor tissues were enriched in Staphylococcaceae (Firmicutes) and Corynebacteriaceae (Actinobacteria), whereas families such as Lactobacillaceae, which were abundant in healthy tissue, were markedly reduced in tumors [24, 25]. Kartti et al. (2023) analyzed 94 fresh breast tumors and adjacent normal tissue samples obtained from 47 patients, examining breast cancer according to four molecular subtypes. In both tissues, the most

prevalent phyla were identified as Proteobacteria (*Gammaproteobacteria*), Firmicutes (*Bacillus*), and Actinobacteria. While normal breast tissue was more densely colonized by Gammaproteobacteria, the tumor tissue was dominated by the classes Bacilli (18.8%) and Actinobacteria (17.2%). In addition, normal tissue samples were enriched with *Escherichia–Shigella* [26, 27]. Ciernikova et al. (2022) reported in their analysis of seven solid tumor types that breast tumors had the richest microbial diversity, with an average of 16 different bacterial species identified per sample. Nejman et al. (2020) suggested that these bacteria are localized in both tumor cells and immune cells, and may influence tumor development and the immune response [18, 23]. Intriguing findings have been obtained regarding the clinical prognostic significance of *Bacillus* presence. Banerjee et al. (2021) investigated the tumor microbiota across different breast cancer subtypes and evaluated the association between microbiota and prognosis in formalin-fixed tumor tissues [28]. In TNBC patients, high levels of Bacillus species detected in tumor tissues have been associated with longer survival times. In the triple-positive patient group, high levels of certain bacterial genera such as Orientia, Klebsiella, Fusobacterium, and Yersinia were found to be significantly associated with poorer prognosis (shorter survival and shorter disease-free period). These findings also suggest that the tumor microbial profile may vary in a subtype specific manner and that the presence of certain microbes could serve as either positive or negative prognostic indicators [28, 29]. It has been reported that, particularly in tumor tissues, the overall abundance of Fusobacteria (e.g., *Fusobacterium nucleatum*) together with Firmicutes is increased [18].

In our heterogeneous metastatic cohort, pooled analysis revealed five taxa *Streptococcus*, *Aquabacterium*, an unclassified Obscuribacteraceae member, *Cutibacterium*, and *Brevibacterium* with Firmicutes and Proteobacteria as the dominant phyla. Kovaleva et al. similarly reported a predominance of Actinobacteria, Proteobacteria, Firmicutes, and Bacteroidetes in NSCLC tumor tissue and adjacent mucosa, and showed that, in advanced tumors, *Pseudomonas*, *Burkholderia*, and *Aquabacterium* decreased while genera such as *Corynebacterium*, *Sphingomonas*, and *Streptococcus*became more abundant [30]. *Staphylococcus* was found to be more abundant compared to the other groups. The presence of *Staphylococcus* in the tumor microenvironment has been reported to be associated with inflammation, immune response, and metastasis [26]. Since metastases generally spread through the bloodstream, their potential to carry microorganisms from the original tumor is limited; they are more likely influenced by the flora of the new organ in which they reside or by circulating microbes. We recognize that such heterogeneity may pose potential confounding factors, particularly when evaluating taxa that tend to cluster in relation to specific tumor biological features or anatomical niches. Nevertheless, this pattern aligns with observations from the SHIVA01 metastatic microbiome study [31], which demonstrated that, in metastatic lesions, microbiome diversity is predominantly shaped by the metastatic site itself rather than by the primary tumor origin (Chao1 $p = 0.52$, Shannon $p = 0.54$; PCoA primary tumor $p = 0.82$). Taken together, these comparative insights align with recent studies showing that the microbiota of LC, CRC, and breast cancer differ not only in their composition but also in their immunomodulatory and metabolic capacities, which may influence disease progression and therapeutic response [6, 23].

This study has several methodological constraints that should be considered when interpreting the findings. The low bacterial biomass and limited DNA yield of FFPE tissues required sequencing of pooled material for each clinical group, which restricted patient level resolution and limited the use of statistical, distribution-based visualizations such as violin plots or differential abundance analyses. Because sequencing was conducted externally, only processed FASTQ files were available, and raw quality metrics could not be retrieved. The retrospective nature of the cohort also resulted in incomplete clinical metadata, including BMI and antibiotic exposure, limiting the ability to explore microbiome clinical associations. Furthermore, contamination monitoring was restricted to paraffin only controls, and additional validation with fresh or matched specimens was not feasible. Together, these factors frame the study as an exploratory assessment of tumor-associated microbiota. Sequencing of individual FFPE or fresh-frozen specimens in future studies will be important for enabling a more detailed assessment of intratumoral microbial heterogeneity and for linking microbial signatures to patient-level clinicopathological features.

Conclusion

In conclusion, our study reveals distinct and stage-dependent microbial signatures across lung, colorectal, and breast cancers using FFPE tumor tissues. The Cyanobacteria driven dysbiosis observed in advanced lung cancer, the enrichment of Firmicutes in later stage colorectal tumors, and the high microbial diversity in breast cancer collectively highlight both tissue-specific and shared microbial patterns. Integrating microbial abundances with survival outcomes further suggests potential links between tumor-associated bacteria and clinical behavior across cancer types. Although the use of pooled sequencing limits statistical inference, this exploratory work provides preliminary insights into the

biological and immunological roles of intratumoral microbiota and offers guidance for future, more comprehensive investigations.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s12032-025-03170-w.

Acknowledgements E. Kanimdan gratefully acknowledges the valuable support provided by the Scientific and Technological Research Council of Türkiye (TUBİTAK) through the 2214 A International Research Fellowship Program for PhD Students [Grant number: 1059B142200683], which enabled the overseas component of this research.

Author contributions E.K. and E.A. contributed the study design and conceptualization, E.K. performed experiments and data analysis, wrote the main manuscript text, E.A. provided funding and technical Support, S.S. provided technical support, B.G., Z.G., E.H.K., S.E. provided FFPE samples, O.P. performed statistical analysis, C.B. N. performed bioinformatic analysis, V.B.Y critically revised the manuscript.

Funding This work was supported by the Istanbul University Scientific Research Projects Units under Grant number 39019.

Data availability All available OTU tables and processed outputs are available from the corresponding author upon reasonable request.

Declarations

Competing interests The authors declare no competing interests.

References

1. Clemente JC, et al. The impact of the gut microbiota on human health: an integrative view. Cell. 2012;148(6):1258–70.
2. Matson V, et al. The commensal microbiome is associated with anti-PD-1 efficacy in metastatic melanoma patients. Science. 2018;359(6371):104–8.
3. Tanoue T, et al. A defined commensal consortium elicits CD8 T cells and anti-cancer immunity. Nature. 2019;565(7741):600–5.
4. Cancer IA. f.R.o., *Putting an end to cancer before it can begin.* 2024.
5. Zhang Y, Liang Y, He C. Anticancer activities and mechanisms of heat-clearing and detoxicating traditional Chinese herbal medicine. Chin Med. 2017;12:1–15.
6. Helmink BA, et al. The microbiome, cancer, and cancer therapy. Nat Med. 2019;25(3):377–88.
7. Mao Q, Hu J. Differential microbiota features in lung tumor and adjacent normal tissues in lung cancer patients. J Thorac Oncol. 2018;13(10):S911–2.
8. Mao Q, et al. Differential flora in the microenvironment of lung tumor and paired adjacent normal tissues. Carcinogenesis. 2020;41(8):1094–103.
9. Sánchez-Alcoholado L, et al. The role of the gut microbiome in colorectal cancer development and therapy response. Cancers. 2020;12(6):1406.
10. Goto T. Airway microbiota as a modulator of lung cancer. Int J Mol Sci. 2020. https://doi.org/10.3390/ijms21093044.
11. Apopa PL, et al. PARP1 is up-regulated in non-small cell lung cancer tissues in the presence of the cyanobacterial toxin microcystin. Front Microbiol. 2018;9:1757.
12. Choi J, et al. A common intronic variant of PARP1 confers melanoma risk and mediates melanocyte growth via regulation of MITF. Nat Genet. 2017;49(9):1326–35.
13. Luo Y, et al. MiR-335 regulates the chemo-radioresistance of small cell lung cancer cells by targeting PARP-1. Gene. 2017;600:9–15.
14. Shen S, et al. Special tissue microbiota such as cyanobacteria are associated with the immune microenvironment of lung adenocarcinoma. Transl Cancer Res. 2024;13(8):4408–19.
15. Tsay JJ, et al. Airway microbiota is associated with upregulation of the PI3K pathway in lung cancer. Am J Respir Crit Care Med. 2018;198(9):1188–98.
16. Huffnagle G, Dickson R, Lukacs N. The respiratory tract microbiome and lung inflammation: a two-way street. Mucosal Immunol. 2017;10(2):299–306.
17. Yu H, et al. Lactate production by tumor-resident Staphylococcus promotes metastatic colonization in lung adenocarcinoma. Cell Host Microbe. 2025;33(7):1089–105. .e7.
18. Ciernikova S, et al. Tumor microbiome – an integral part of the tumor microenvironment. Front Oncol. 2022;12:1063100.
19. Cai P, et al. Fecal bacterial biomarkers and blood biochemical indicators as potential key factors in the development of colorectal cancer. mSystems. 2025;10(3):e0004325.
20. Xu Y, et al. The microbiome types of colorectal tissue are potentially associated with the prognosis of patients with colorectal cancer. Front Microbiol. 2023;14:1100873.
21. Li Z, et al. Altered actinobacteria and firmicutes phylum associated epitopes in patients with parkinson's disease. Front Immunol. 2021;12:632482.
22. Li Q, et al. Streptococcus thermophilus inhibits colorectal tumorigenesis through secreting β-Galactosidase. Gastroenterology. 2021;160(4):1179–e119314.
23. Nejman D, et al. The human tumor microbiome is composed of tumor type-specific intracellular bacteria. Science. 2020;368(6494):973–80.
24. German R, et al. Exploring breast tissue microbial composition and the association with breast cancer risk factors. Breast Cancer Res. 2023;25(1):82.
25. Costantini L, et al. Characterization of human breast tissue microbiota from core needle biopsies through the analysis of multi hypervariable 16S-rRNA gene regions. Sci Rep. 2018;8(1):16893.
26. Kartti S, et al. Metagenomics analysis of breast microbiome highlights the abundance of *Rothia* genus in tumor tissues. J Pers Med. 2023. https://doi.org/10.3390/jpm13030450.
27. Tsang JYS, Tse GM. Molecular classification of breast cancer. Adv Anat Pathol. 2020;27(1):27–35.
28. Banerjee S, et al. Prognostic correlations with the microbiome of breast cancer subtypes. Cell Death Dis. 2021. https://doi.org/10.1038/s41419-021-04092-x.
29. Banerjee S, et al. Distinct microbial signatures associated with different breast cancer types. Front Microbiol. 2018;9:951.
30. Kovaleva O, et al. Lung microbiome differentially impacts survival of patients with non-small cell lung cancer depending on tumor stroma phenotype. Biomedicines. 2020;8(9):349.
31. Hilmi M, et al. Intratumoral Microbiome is driven by metastatic site and associated with immune histopathological parameters: an ancillary study of the SHIVA clinical trial. Eur J Cancer. 2023;183:152–61.