Cancer Causes & Control (2020) 31:1027–1038
https://doi.org/10.1007/s10552-020-01338-5

ORIGINAL PAPER

Microbiota composition in bilateral healthy breast tissue and breast tumors

Emily Klann[1,2] · Jessica M. Williamson[1] · Massimiliano S. Tagliamonte[2,3] · Maria Ukhanova[2] · Jaya Ruth Asirvatham[3] · Harvey Chim[4] · Lusine Yaghjyan[1] · Volker Mai[1,2]

Received: 15 November 2019 / Accepted: 13 August 2020 / Published online: 25 August 2020

Abstract

Purpose Previous reports suggest that a complex microbiome exists within the female human breast that might contribute to breast cancer etiology. The purpose of this pilot study was to assess the variation in microbiota composition by breast side (left versus right) within individual women and compare the microbiota of normal and breast tumor tissue between women. We aimed to determine whether microbiota composition differs between these groups and whether certain bacterial taxa may be associated with breast tumors.

Methods Bilateral normal breast tissue samples ($n=36$) were collected from ten women who received routine mammoplasty procedures. Archived breast tumor samples ($n=10$) were obtained from a biorepository. DNA was extracted, amplified, and sequenced. Microbiota data were analyzed using QIIME and RStudio.

Results The most abundant phyla in both tumor and normal tissues were Bacteroidetes, Firmicutes, Proteobacteria, and Actinobacteria. There were statistically significant differences in the relative abundance of various bacterial taxa between groups. Alpha diversity (Simpson's index) was significantly higher in normal compared to tumor samples (0.968 vs. 0.957, $p=0.022$). Based on unweighted UniFrac measures, breast tumor samples clustered distinctly from normal samples ($R^2=0.130$; $p=0.01$). Microbiota composition in normal samples clustered within women ($R^2=0.394$; $p=0.01$) and by breast side (left or right) within a woman ($R^2=0.189$; $p=0.03$).

Conclusion Significant differences in diversity between tumor and normal tissue and in composition between women and between breasts of the same woman were identified. These results warrant further research to investigate the relationship between microbiota and breast cancer.

Keywords Breast · Cancer · Microbiota · Tumor microenvironment

Background

Breast cancer is the most commonly diagnosed cancer and the leading cause of cancer-related mortality among women, accounting for more than 500,000 deaths worldwide each year [1]. Breast cancer is a highly heterogeneous disease with a complex etiology. Although the etiology is not fully understood, several well-established risk factors have been identified including age, race and ethnicity, reproductive characteristics, postmenopausal hormone use, genetic and epigenetic factors, and various environmental factors [2, 3]. However, these factors do not completely explain the risk of developing breast cancer. In fact, a recent study found that approximately 70% of women diagnosed with breast cancer under the age of 50 were not considered to be at a high lifetime risk [4]. Therefore, other potential factors, such as

Emily Klann and Jessica M. Williamson have contributed equally as first authors.

Lusine Yaghjyan and Volker Mai share senior authorship.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10552-020-01338-5) contains supplementary material, which is available to authorized users.

✉ Volker Mai
vmai@ufl.edu

Extended author information available on the last page of the article

the microbiome, have been of recent interest in regard to the development and progression of breast cancer.

The human microbiome consists of approximately 10–100 trillion symbiotic microorganisms that inhabit diverse anatomic sites within the body [5]. The organisms that constitute the microbiome, along with their respective metabolites, have been found to contribute to the risk of certain cancers such as colon [6], stomach [7], liver [8], and breast [9]. While the gut has received the most research interest in regard to the connection between the microbiome and cancer, other anatomic sites have also been explored. Although originally thought to be a sterile site, it has been suggested that breast tissue may be exposed to microorganisms and associated metabolites via bacterial translocation between the gut microbiome and the mammary glands [10] or direct nipple-oral contact [11]. The microbiome of the breast tissue itself has recently been hypothesized to influence the initiation and progression of breast cancer [11–18].

To date, only a few studies have explored the association between the microbial make-up of the local breast tissue environment and breast cancer. Most existing studies have found the breast microbiome composition to be dominated by a combination of Proteobacteria, Firmicutes, and Bacteroidetes [11–20]. Measures of diversity vary across these studies with the majority finding no significant differences in alpha diversity but observable differences in clustering between non-tumor and tumor tissues [11, 13, 17]. Further, a few of these studies have found differences in composition and abundance of the breast microbiome across tumor subtypes and stages of breast cancer. Although some studies found no difference across tumor subtypes [17] or stages [16, 17], other studies found differences in the relative abundance of microbiota across stages [19], observed an inverse relationship between total bacterial load and stage [18], and detected differences in microbial signatures between human epidermal growth factor receptor 2 (HER2) positive, endocrine receptor positive (estrogen and progesterone positive), triple negative (estrogen, progesterone, and HER2 negative), and triple positive (estrogen, progesterone, and HER2 positive) subtypes [12].

Breast cancer is typically diagnosed unilaterally, suggesting possible differences in the local tissue microenvironment between the two breasts of the same woman. One such local factor could be the microbiome. However, there are no existing studies that have described within-woman differences in the microbiomes of the left and right breasts. Therefore, differences in the composition and abundance of the tissue microbiome between the left and right breasts within a woman remain unknown. In this exploratory study, we aimed to (1) identify the presence of a distinct microbiome within normal breast tissue, (2) evaluate possible differences between the microbiome profiles of the left and right breasts within an individual woman, and (3) evaluate possible differences between the microbiome profiles of normal breast tissue and breast tumor tissue.

Materials and methods

Tissue collection

A total of ten archived breast cancer tumor tissue samples were obtained through the University of Florida (UF) Clinical and Translational Sciences Institute (CTSI) Biorepository. All ten samples were from white, non-Hispanic women. These samples were stored in liquid nitrogen and, after retrieval, were transported to the UF Emerging Pathogens Institute (EPI) for further processing.

Additionally, normal breast tissue samples were obtained from ten women undergoing routine reduction mammoplasty at the UF Health Department of Plastic Surgery. Out of the ten women who underwent reduction mammoplasty procedures, eight provided two samples from each of the breasts while the remaining two women provided two samples from only one breast. A total of 36 normal tissue samples were obtained. These fresh tissue samples were delivered to the UF Department of Pathology for routine clinical testing and histological assessment to ensure a cancer-free status. Tissue immediately adjacent to the section used for clinical testing was collected, de-identified, placed in a preservative solution (RNAlater), and transported to the UF EPI for further processing.

DNA extraction, amplification, and sequencing

DNA was extracted from homogenized tissue samples using Qiagen DNA Blood and Tissue kit (Qiagen) and manufacturer protocols with the addition of a mechanical lysis bead-beating step (Biospec Products). Following extraction, DNA concentration for each sample was measured using the Nanodrop platform. Genomic DNA from the tissue samples was amplified via polymerase chain reaction (PCR) using barcoded primers for the V1–V2 hypervariable regions of the bacterial 16S rRNA. Concentration of barcoded PCR products were measured (Quati-iT dsDNA assay kit, Invitrogen) and equimolar amounts were then pooled and purified (Mag-Bind TotalPure NGS beads, Omega Bio Tek). The Illumina Mi-Seq platform was used to pair-end (250×2) sequence the samples. A blank negative extraction control consisting of sterile water was run in conjunction with the samples during DNA extraction, amplification, and sequencing in order to identify possible contamination that could influence the results of microbiota analysis due to the low biomass nature of breast tissue.

Microbiota data processing and analysis

Quantitative Insights into Microbial Ecology (QIIME) tool (version 1.9.1; open source software) [21] and the Brazilian Microbiome Project (BMP) [22] pipelines were used to trim, quality filter, and cluster sequences into de novo operational taxonomic units (OTUs) at the 98% similarity level using UCLUST [23]. Rarefaction curves to assess species richness and comparability of the samples were also generated through these pipelines. Chimeras were removed and taxonomy was assigned to OTUs using the Ribosomal Database Project (RDP) [24] classifier and verified against the Greengenes 16S rRNA gene database [25]. A phylogenetic tree for diversity analyses was generated with the FastTree pipeline [26]. All OTUs present in the negative control and those not classified at the phylum level were removed prior to analysis. Further, phyloseq [27] was used to filter and remove all OTUs without at least ten total reads across samples and those not present in at least 10% of all samples prior to all downstream analyses. For each sample and taxonomic level, relative abundance was calculated as the number of counts for a specific taxonomic classification (i.e., particular genera such as Bacteroides) divided by the total number of counts for all classifications at that taxonomic level (i.e., all genera). Normalization, a procedure used to render all samples comparable by removing bias due to variable sequencing depths, was performed prior to conducting ordination or differential abundance analyses. The procedure chosen for this analysis involves the log transformation of OTU counts using the formula suggested by McCurdie and Holmes [27], $\log(1+x)$, which accounts for the high proportion of zero counts naturally present in microbiome datasets. Beta diversity (community diversity between samples) was assessed using both weighted and unweighted UniFrac measures of phylogenetic distance generated using Phyloseq [27]. Alpha diversity (diversity within a sample) was assessed using Shannon diversity and Simpson's index, also calculated with Phyloseq [27]. The DESeq function within the DESeq2 package, which uses built-in normalization procedures, was used to conduct differential abundance analyses between groups [28]. KEGG Functionality was predicted for all OTUs using the metagenomics inference tool Piphillin. Methods concerning the inference of metagenomic content and copy number normalization are described elsewhere [29]. Tumor samples were not analyzed by receptor status due to the small samples size ($n=10$).

Statistical analysis

All statistical analyses were conducted with RStudio (RStudio Team) [30] and QIIME (version 1.9.1; open source software) [21]. The DESeq2 package was used to conduct Wald's test to identify differentially abundant OTUs between groups. The Mann–Whitney U test, a non-parametric test for two independent samples, was used to assess differences in relative abundance of taxa, alpha diversity measures, and KEGG functional pathway abundance between tumor and normal breast tissue samples. Permutational Multivariate Analysis of Variance Using Distance Matrices (PERMANOVA) was performed with 99 iterations via the Adonis function within the R package vegan (version 2.5-4) [31] to evaluate the differences in UniFrac distance metrics between groups. The MaAsLin2 R package was used to run a generalized linear mixed effects model (with individual subject ID as the random effect and disease status as the fixed effect) to identify OTUs associated with tissue group while controlling for the paired nature of the normal tissue samples [32]. *p* values for relative abundance and differential abundances analyses were adjusted for multiple comparisons using the Benjamini–Hochberg method. A *p* value of less than 0.05 was considered to be statistically significant.

Results

In total, 46 breast tissue samples (10 tumor and 36 normal) were analyzed. From these samples, we obtained a total of 5,070,480 sequence reads with an average of 98,402 total reads per sample from tumor tissue and an average of 112,232 total reads per sample from normal tissue. A total of 2,366 unique OTUs were identified at the 98% similarity level. After filtering to remove all OTUs present in the negative control, those without a phylum level classification, and those not present in at least 10% of all samples, a total of 412 remained and were used in the analysis. The ten most abundant OTUs were of the family *Ruminococcaceae* and genera *Acidaminococcus, Acinetobacter, Akkermansia, Bacteroides,* and *Sutterella*. Of these OTUs, *Ruminococcaceae* and *Akkermanisa* were consistently in higher relative abundance in tumor compared to normal tissue. *Bacteroides* and two of the three OTUs classified as *Sutterella* were in lower relative abundance in tumor compared to normal tissue (Fig. 1).

Alpha diversity between groups

Simpson's index, on average, was significantly higher with more variation in range in normal breast tissue samples (0.968; range of 0.574–0.989) compared to tumor breast tissue samples (0.957; range of 0.741–0.989; *p* for difference = 0.022). Simpson's index is a measure of similarity and has a range of values between 0 and 1, with 1 equating to complete evenness of the community within a sample. This index accounts for both richness and evenness, with more weight placed on the presence of common or dominant species [33]. Similarly, the difference in the average Shannon

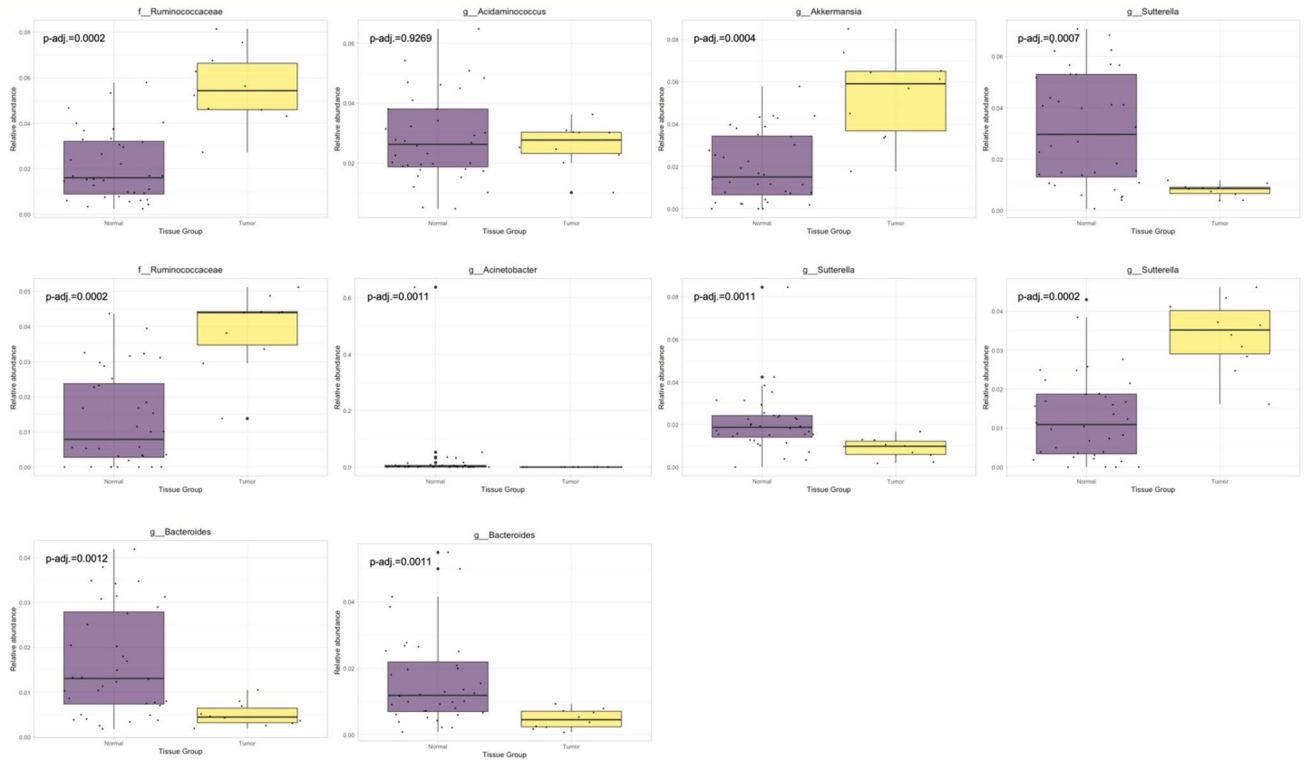

Fig. 1 Relative abundance of 10 most abundant OTUs between normal and tumor tissue groups. The average relative abundance of the top ten most abundant OTUs are displayed by group (normal and tumor tissue). The average relative abundance of each OTU is shown as a boxplot with the central line representing the median. The relative abundance values for each individual sample are included as dots around the boxplot. P values for the difference between groups are displayed for each OTU and were generated through non-parametric Mann–Whitney U test and adjusted using the Benjamini–Hochberg method

diversity value between groups was higher, although not reaching significance, in the normal samples (4.57; range of 1.82–4.98) compared to the tumor samples (4.51; range of 3.06–5.00; p for difference = 0.424). Shannon diversity can assume values greater than 1, with an increase in values indicating increases in both richness and evenness of the community within a sample. Neither Simpson's index ($p = 0.274$) nor Shannon diversity ($p = 0.706$) were found to be consistently different between the left and right breasts of healthy women.

Differences in relative abundance at various taxonomic levels

A total of ten phyla, 58 families, and 74 genera were identified. 25 OTUs were present in at least one normal tissue sample but none of the tumor tissue samples and only one was present in at least one tumor tissue sample but none of the normal tissue samples. Out of the ten total phyla, eight were observed in OTUs found only in normal tissue. The OTU exclusive to tumor tissue was of the *Aerococcaceae* family and was present in 50% of these samples. The four most predominant phyla which constitute over 94% of the average relative abundance per group demonstrated distinct clustering by weighted UniFrac distances, which take into account both presence and abundance, on individual PCoA plots (Fig. 2).

The relative abundance of five of the ten phyla were found to be significantly different between normal and tumor tissue with a higher relative abundance in normal tissue samples including Cyanobacteria (0.56 vs. 0.17%; $p = 0.021$), Proteobacteria (17.2 vs. 9.13%; $p < 0.001$), Synergistetes (4.70 vs. 0.06%; $p = 0.013$), and Tenericutes (0.82 vs. 0.21%; $p = 0.006$). Verrucomicrobia was significantly less abundant in normal compared to tumor samples (2.18 vs. 5.21%; $p < 0.001$) (Fig. 3a and Online Resource 1a).

Heterogeneity was observed at the phylum level between the left and right breasts within an individual woman. Of note, both samples for each breast side (two for left and two for right) were similar in their relative abundance measures for each phylum within each woman. The median percent difference observed for each phylum between the two breasts of each woman ranged from 15.65 to 180.25% (the change was larger than the original value of relative abundance in one breast) with an average of 53.62% across all phyla (Online Resource

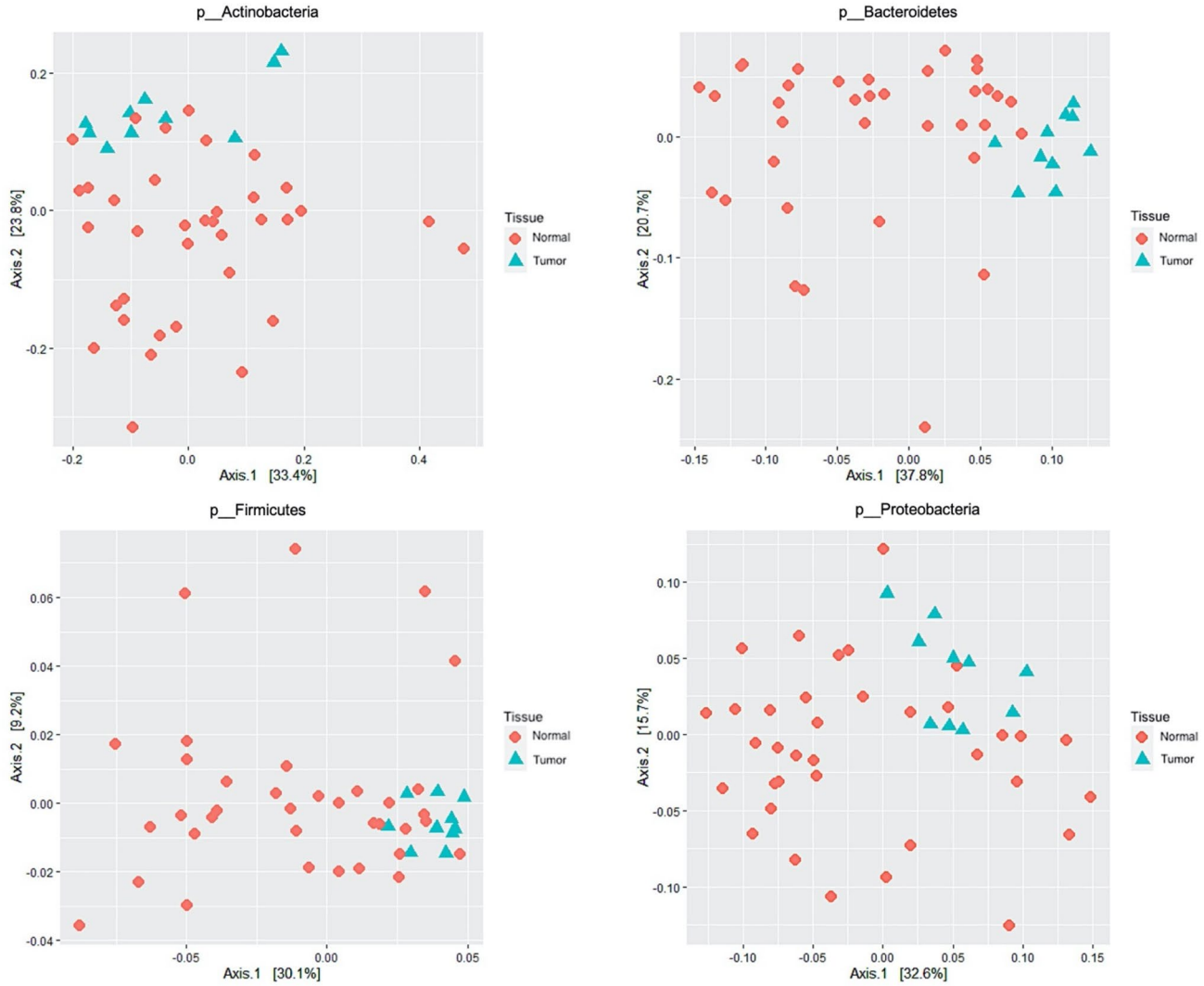

Fig. 2 Weighted UniFrac distances of most prominent phyla plotted via PCoA by group. Distinct clustering is observable between the normal and tumor tissue groups for each of the four most dominant phyla via individual PCoA plots of the weighted UniFrac distances. These four phyla (Actinobacteria, Bacteroidetes, Firmicutes, and Proteobacteria) constitute more than 94% of the average relative abundance per group. Each shape on the plot represents an individual sample. Data were normalized for equitable comparison prior to ordination. Normal tissue samples are red circles and tumor tissue samples are blue triangles

2). Synergistetes, Tenericutes, TM6, and Verrucomicrobia had a median difference of over 50% between the two breasts.

The relative abundance proportions of 36 bacterial families were also found to be statistically significantly different between normal and tumor tissue groups. Among the families constituting over 1% of the average total relative abundance, [*Odoribacteraceae*], *Porphyromonadaceae, Ruminococcaceae*, and *Verrucomicrobiaceae* were statistically significantly higher in relative abundance and *Alcaligenaceae* and *Flavobacteriaceae*, *Moraxellaceae*, and *Prevotellaceae* were significantly lower in relative abundance in tumor compared to normal tissue (Fig. 3b and Online Resource 1b).

Further, 59 genera were statistically significantly different between normal and tumor tissue groups. Among the genera constituting over 1% of the average total relative abundance, *Acinetobacter, Flavobacterium, Prevotella, Staphylococcus*, and *Sutterella* were found in higher abundance and *Akkermansia, Butyricimonas*, and *Parabacteroides* were found in lower abundance in tumor compared to normal tissue (Fig. 3c and Online Resource 2c).

OTUs associated with normal and tumor tissue groups

Differential abundance calculations were also conducted in order to identify OTUs that are associated with tumor-status.

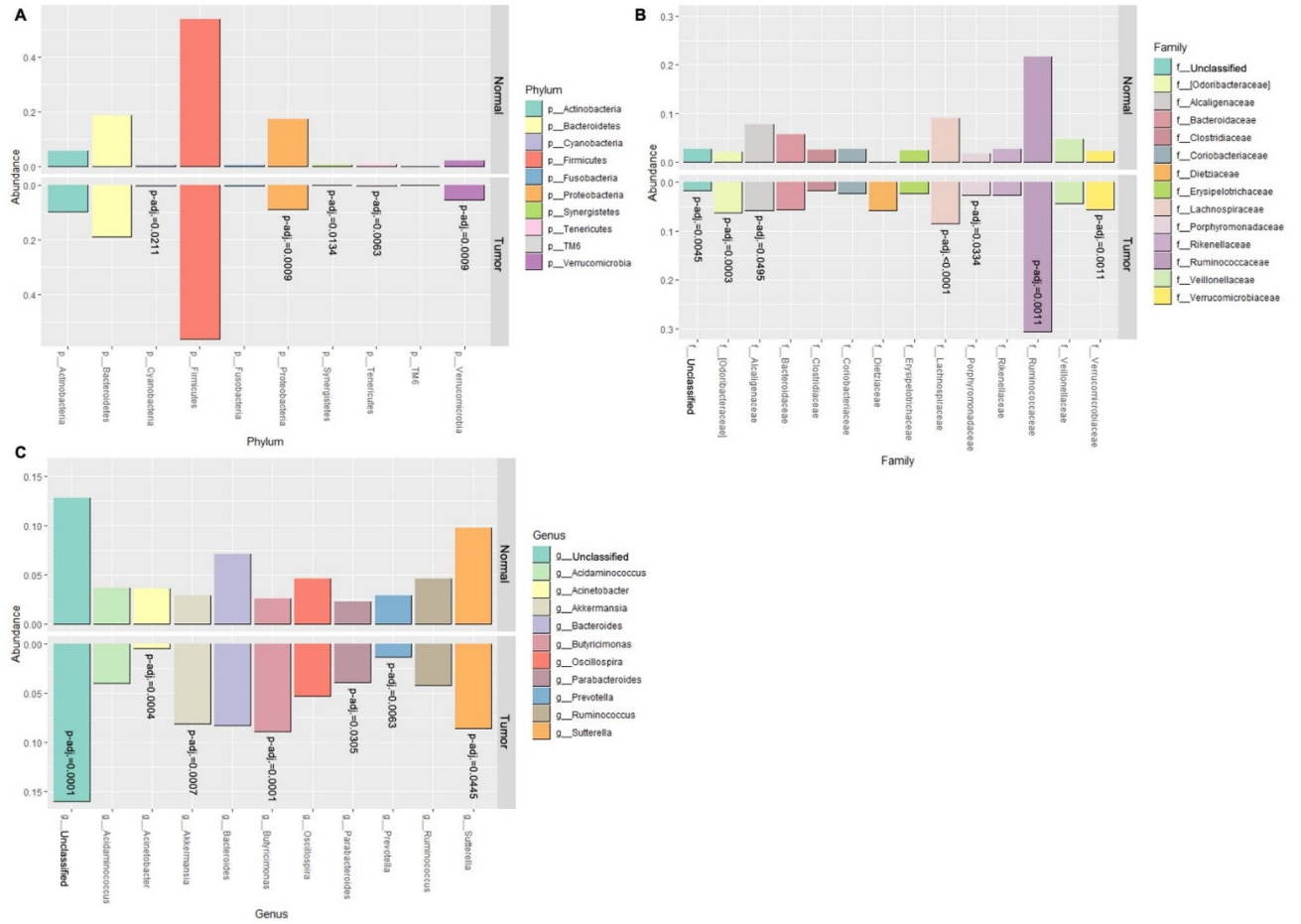

Fig. 3 Relative abundance of phyla, families, and genera in normal and tumor tissue. Average relative abundance of (**a**) phyla, (**b**) families, and **c** genera by group (normal and tumor tissues). *Indicates significant differences in relative abundance between groups ($p < 0.05$). Taxa that are significantly different between groups have a corresponding p value on each graph. p values were generated through a Mann–Whitney U test and adjusted for multiple comparisons using the Benjamini–Hochberg method. Only taxa with an average relative abundance of over 2% per group were included in the **b** family and **c** genus graphs

A total of 77 OTUs were found to be significantly different in terms of differential abundance between tumor and normal tissue samples (Fig. 4). The absolute value of the log2 fold changes for these differences per OTU ranged from 0.853 to 6.07. The OTUs with the highest magnitude of change associated with higher abundance in tumor tissue were of the *Pseudomonoadaceae* ($R^2 = 4.43$), *Dietziaceae* ($R^2 = 3.74$), *Gemellaceae* ($R^2 = 3.66$), and *Neisseriaceae* ($R^2 = 2.43$) families. Conversely, the OTUs with the highest magnitude of change associated with higher abundance in normal tissue were of [*Mogibacteriaceae*] family ($R^2 = 5.43$), and *Acinetobacter* ($R^2 = 6.07$), Brevibacillus ($R^2 = 5.15$), and *Flavobacterium* genera ($R^2 = 5.97$). Interestingly, only three genera were found to be associated with higher differential abundance for both tumor and normal tissue samples including *Bacteroides, Staphylococcus,* and *Sutterella* (Fig. 4).

A linear mixed effects model, controlling for the paired nature of the normal tissue samples by setting individual subject ID as a random effect, identified 131 OTUs as associated with tissue group with coefficients ranging from − 2.214 to 1.914. Of these, 59 were associated with normal tissue and 72 were associated with tumor tissue samples. The OTUs with the highest magnitude of association with tumor tissue were of families *Ruminococcaceae* and *Rikenellaceae* and genera *Butyricimonas, Sutterella,* and *Akkermansia.* Those with the highest magnitude of association with normal tissue were of Clostridiales order, *[Mogibacteriaceae]* family, and *Flavobacterium, Acinetobacter,* and *Brevibacillus* genera (Online Resource 3).

Differences in community similarity between groups

The overall similarity of the microbiota between the two groups was evaluated visually using a PCoA plot of unweighted UniFrac distances and statistically using PERMANOVA. The UniFrac distances differed significantly

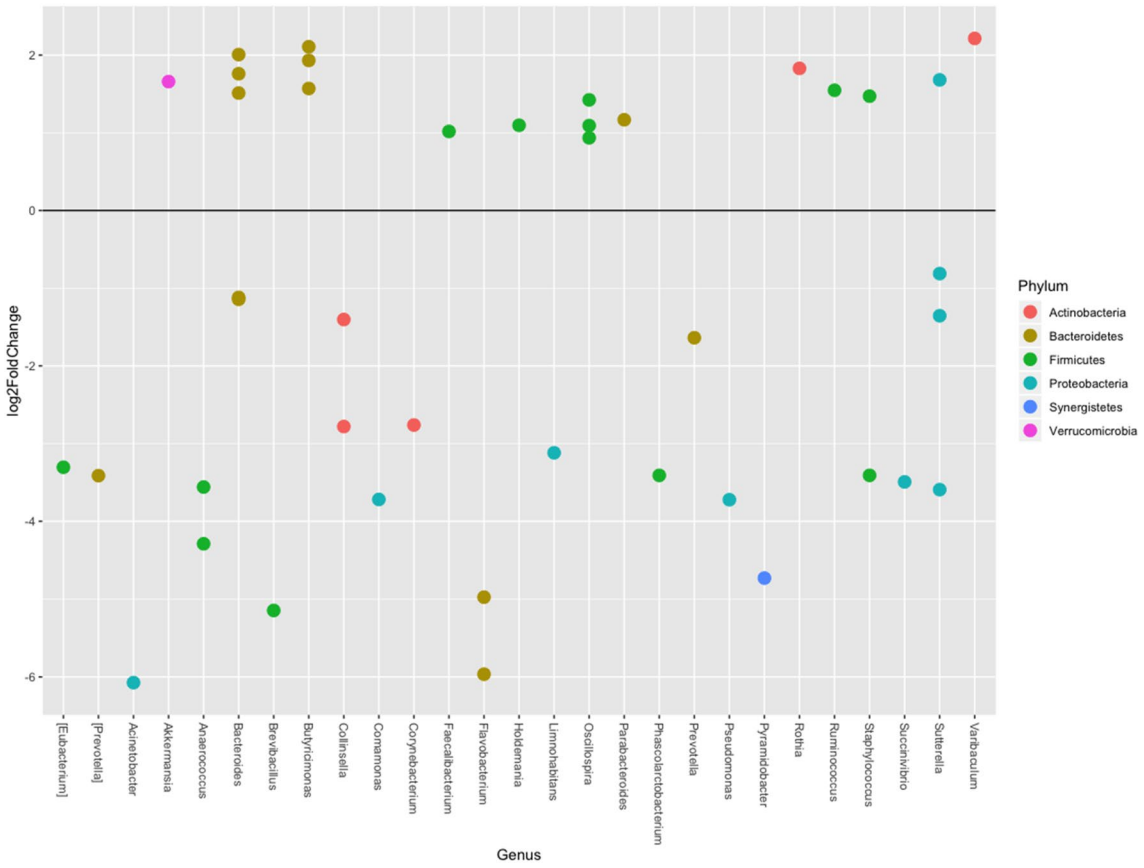

Fig. 4 Significantly different taxa identified by differential abundance between normal and tumor tissue. This graph depicts the magnitude of change (log2 FoldChange) in terms of abundance between tumor and normal tissue. The black horizontal line indicates no difference ($R^2=0$). All circles above this line represent OTUs more abundant in tumor samples and all circles below this line represent OTUs more abundant in normal samples. The color of the circle indicates phylum and the location on the horizontal axis indicates genus. All OTUs shown were found to be statistically significant in terms of differential abundance using Wald's test with Benjamini–Hochberg adjustment in the DESeq2 package (DESeq function). Data were normalized within the DESeq function for equitable comparison prior to analysis

between tumor and normal tissue samples ($R^2=0.130$; $p=0.01$) and the breast tumor tissue samples demonstrated distinct clustering on the PCoA plot (Fig. 5). Although normal tissue samples were distributed somewhat heterogeneously across the plot overall, we found that the samples clustered significantly by individual women ($R^2=0.394$; $p=0.01$) (Fig. 6). Further, we detected significant clustering by breast side (left or right) within each woman ($R^2=0.189$; $p=0.03$) (Fig. 6). As an additional control measure, the tissue samples were compared to soil and human stool samples, both of which are commonly investigated in the lab used in this study. We determined that breast tissue, soil, and stool samples clustered distinctly differently from one another, suggesting limited cross-contamination between these sources (Online Resource 4). Although swabs from the breast skin prior to each surgical procedure would have also served as informative comparisons, collection of such samples was outside of the scope of the current project. However, a similar study investigating the breast tissue microbiome utilized swabs of disinfected areas of the breast skin prior to surgery as controls. These authors concluded that the breast tissue samples clustered separately from the skin swab controls, suggesting a distinct microbial profile associated with breast tissue [16].

Predicted functionality of OTUs by group

A total of 288 predicted KEGG Orthology functional pathways were identified across all microbiota samples. The predicted abundance of 205 of these functional pathways were significantly different between the microbiota from normal and tumor tissue groups. Of these, 16 pathways predicted to be in higher abundance in microbiota from tumor tissue share similarities to pathways associated with carcinogenesis in breast tissue including those with predicted functions related to base excision repair, Th17 cell differentiation, choline and central carbon metabolism, necroptosis, microRNAs and proteoglycans involved in carcinogenesis,

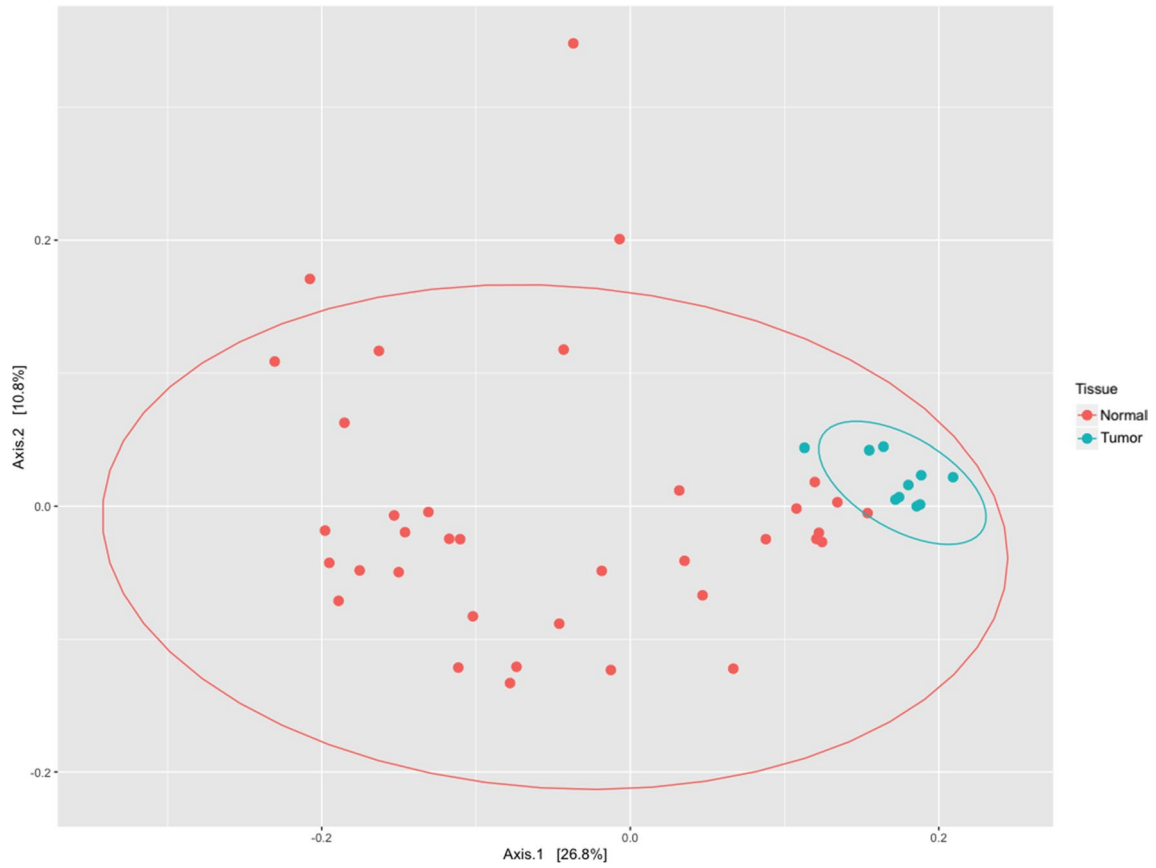

Fig. 5 Principal coordinate analysis of unweighted UniFrac measures of normal and tumor breast tissue. Distinct clustering is observable between the normal and tumor tissue groups on a PCoA plot of the unweighted UniFrac distances. The distances between normal and tumor tissue samples were determined to be significantly different ($R^2 = 0.130$; $p = 0.01$) through PERMANOVA using the Adonis function in R package vegan. Data were normalized for equitable comparison prior to ordination. Each shape on the plot represents an individual sample. Normal tissue samples are red circles and tumor tissue samples are blue circles

and various signaling pathways including IL-17, PI3K-Akt, HIF-1, and AMPK.

Discussion

The results of this pilot study suggest distinct bacterial profiles associated with breast tumors and normal tissues from healthy controls. Additionally, we found that normal breast tissue samples differed in their microbiota composition across individual women. Further, the microbial profiles of the left and right breasts within a woman were found to be distinctly different in terms of both relative abundance and composition. These differences between the breasts may play a role in the unilateral nature of breast cancer development. These findings are noteworthy as the current evidence suggests a link between certain microbiota and the presence of cancers, with differences in relative abundance and composition observed between diseased and non-diseased individuals. Considering this, the differences in microbiota detected between tumor and normal breast tissue, as well as between the left and right sides of the breasts within an individual woman, may partially contribute to the risk for developing breast cancer.

Consistent with other similar studies, the most abundant phyla found in our samples included Firmicutes, Bacteroidetes, Proteobacteria, Actinobacteria, and Verrucomicrobia [11, 14, 15, 18]. Further, a recent study by Smith and colleagues also found Pseudomonadaceae to be enriched in breast tumor tissue compared normal breast tissue samples [19] and other studies found this family to be consistently present across all breast tissue samples [15, 20]. Similar to our findings, two recent studies [13, 19] found non-tumor tissue to have slightly higher alpha diversity when compared to tumor tissue; however, this result was not statistically significant. Other similar studies that did not find a difference in alpha diversity between normal and tumor tissue might be explained by the sample retrieval approaches: the "normal" tissue in these studies was represented by the tissue adjacent to the tumor of the same woman [11, 17], unlike

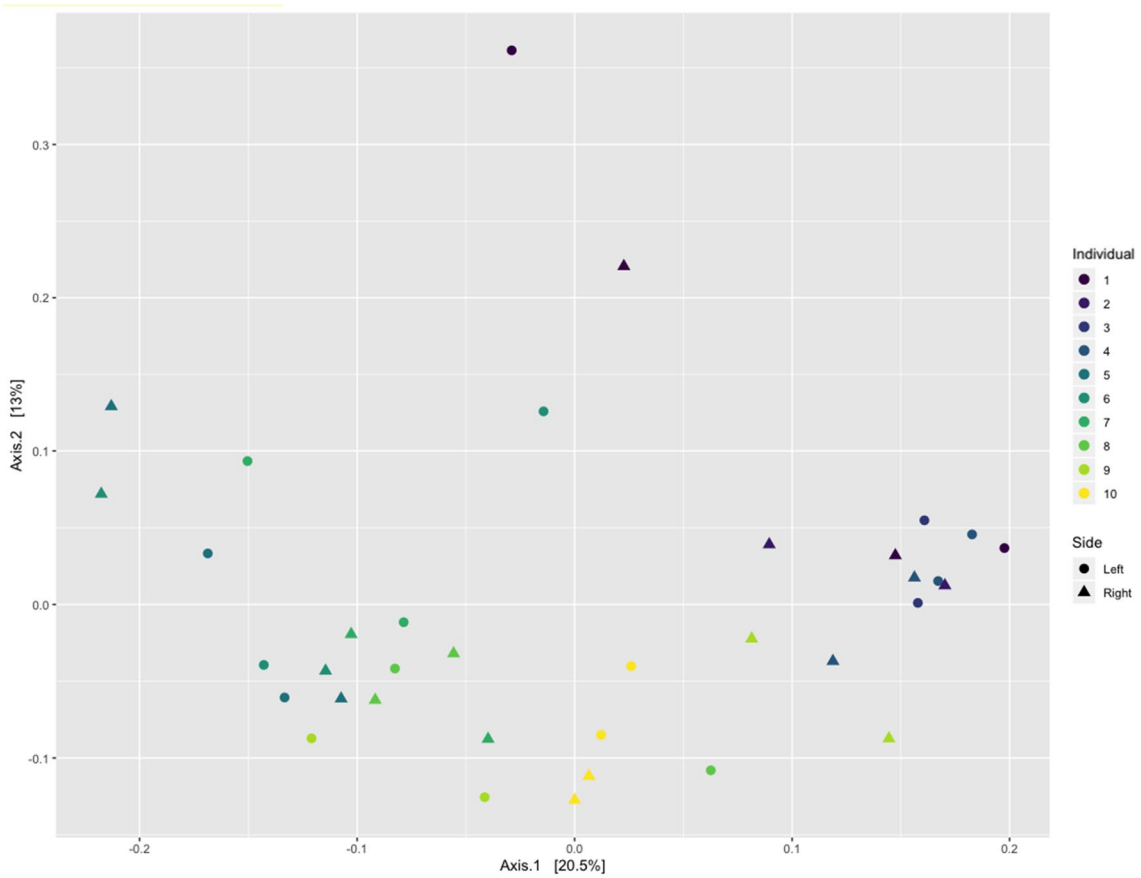

Fig. 6 Principal coordinate analysis plot of unweighted UniFrac measures of normal breast tissue by individual and breast side. This PCoA plot displays the unweighted UniFrac distances between all normal tissue samples. Samples from each individual woman are shown by a unique color. Breast side is depicted by shape (circle = left; triangle = right). The distances between individual women ($R^2 = 0.394$; $p = 0.01$) and between the left and right breast within one woman ($R^2 = 0.189$; $p = 0.03$) were determined to be significantly different through PERMANOVA using the Adonis function in R package vegan. Data were normalized for equitable comparison prior to ordination

our study where "normal" breast tissue was obtained from other cancer-free women. Several similar studies also had comparable findings to our study in terms of distinct composition-dependent clustering differences between cancerous and noncancerous samples [11, 14, 17, 19].

Notably, some bacterial taxa with the largest effect sizes associated with greater differential abundance, as well as those found in higher relative abundance in breast tumor tissue compared to normal tissue, have been found to be pro-inflammatory and immunogenic through increased levels of chemokine ligand 13 (CCL13), pentraxin 3 (PTX3) [34], interleukin-8 (IL-8), and C-reactive protein (CRP) [35]. Inflammation is known to play a role in breast cancer etiology and several recent studies have found increased levels of similar pro-inflammatory biomarkers, including TNF-α, IL-8, C-reactive protein (CRP), and fibrinogen [36], in women with breast cancer compared to women without breast cancer. Further, many of these bacteria have been associated with other cancers (genera within the *Pseudomonadaceae* family with colorectal adenoma [37], both *Pseudomonadacea* and *Dietziaceae* families with bile duct tissue involved in cholangiocarcinoma [38], *Rikenellaceae and Ruminococcaceae* families with colorectal cancer [39, 40], and *Neisseriaceae* family with esophageal adenocarcinoma and squamous cell carcinoma [41]), autoimmune disorders [42], and opportunistic infections [43, 44]. Further, the sole OTU unique to tumor tissue (family *Aerococcaceae*) has been found to be enriched in prostate tumors [45] and significantly associated with the presence of leukemia [46]. Interestingly, members of the *Gemellaceae* family have been found in the infant gut and oral microbiome as well as in the breast milk of their respective mothers [47, 48]. Additionally, the *Gemellaceae* family has been detected within the endometrial fluid and vaginal aspirates of fertile women [49], suggesting a possible maternal vaginal etiology of these microbiota in the context of the infant microbiome. This offers a possible explanation for the presence of these bacteria in breast

tissue with a possible nipple-oral or mammary gland etiology of this microbiome. The above-described biological mechanisms may suggest a potential link between the presence and relative abundance of *Pseudomonoadaceae, Dietziaceae, Gemellaceae*, and *Neisseriaceae* families and the occurrence of breast cancer. That some functional pathways related to cancer pathophysiology were predicted to be enriched in microbiota from tumor samples is intriguing but needs to be confirmed by direct determination of microbiota gene and protein content. However, it is not yet known whether the detected microbiota influence the development of breast cancer or if they are conversely a product of the diseased microenvironment (reverse causality).

To our knowledge, this is the first study to evaluate characteristics of the tissue microbiome by breast side. Our study has several strengths including routine clinical ascertainment of cancer-free breast tissue for normal tissue samples, use of validated sequencing methods [50] and bioinformatic pipelines to analyze the microbiome data [51], and use of methods to control for and remove sequences likely to have originated from environmental contamination. Further, our tumor tissue samples are comprised of a combination of receptor-status subtypes, representing a majority of the subtypes found in the general population [52]. This study also has a few limitations which warrant discussion. First, due to the exploratory nature of this pilot study, the sample size was small, which may have attenuated the ability to detect significant differences in composition and relative abundance between tumor and normal tissue samples. Additionally, information related to demographics and individual characteristics, such race and ethnicity or medical history, was not available. Further, the two types of tissue samples (archived tumor tissue and normal tissue from reduction mammoplasty) were collected and stored differently and at different points in time. This may have introduced contamination distinct from that identified by the control from the laboratory setting. The tumor and normal samples were also obtained from different areas of the breast. While the breast cancer samples were comprised strictly of tumor tissue, the normal samples were collected from multiple areas of the breast. Although thorough measures were taken to prevent and control for possible environmental contamination, the results should still be interpreted with caution due to the ability for minor contamination to appreciably skew microbiota data originating from low biomass samples, such as breast tissue. Lastly, functional predictions from 16S data are dependent on accurate gene annotation and do not reflect whether these genes are transcribed or translated in the environment from which the sequences originated and should therefore also be interpreted with caution. Despite these limitations, similar studies have produced comparable findings in terms of relative abundance and beta diversity of the breast tissue.

Conclusion

In this pilot study, we found significant differences in both composition and relative abundance between breast tumor tissue and normal tissue from reduction mammoplasty procedures as well as across women and between the breasts (left versus right) within each woman. Further research is needed to elucidate the potential role of breast microbiota in breast cancer etiology and to further evaluate the potential utility of the breast tissue microbiome as a potential diagnostic tool and/or a target for therapeutic interventions.

Funding The study was funded by the University of Florida Health Cancer Center and National Science Foundation.

Compliance with ethical standards

Conflict of interest The authors declare that they have no competing interests.

Ethical approval This study was reviewed and approved as exempt by the UF Institutional Review Board (IRB) (Protocol Number IRB201600709).

References

1. Bray F, Ferlay J, Soerjomataram I et al (2018) Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: Cancer J Clin 68:394–424. https://doi.org/10.3322/caac.21492
2. Sun Y-S, Zhao Z, Yang Z-N et al (2017) Risk factors and preventions of breast cancer. Int J Biol Sci 13:1387–1397. https://doi.org/10.7150/ijbs.21635
3. Tao Z, Shi A, Lu C et al (2015) Breast cancer: epidemiology and etiology. Cell Biochem Biophys 72:333–338. https://doi.org/10.1007/s12013-014-0459-6
4. Sutton T, Reilly P, Johnson N, Garreau JR (2018) Breast cancer in women under 50: most are not high risk. Am J Surg 215:848–851. https://doi.org/10.1016/j.amjsurg.2018.01.003
5. Ursell LK, Metcalf JL, Parfrey LW, Knight R (2012) Defining the human microbiome. Nutr Rev 70:S38–S44. https://doi.org/10.1111/j.1753-4887.2012.00493.x
6. Dahmus JD, Kotler DL, Kastenberg DM, Kistler CA (2018) The gut microbiome and colorectal cancer: a review of bacterial pathogenesis. J Gastrointest Oncol 9:769–777. https://doi.org/10.21037/jgo.2018.04.07
7. Meng C, Bai C, Brown TD et al (2018) Human gut microbiota and gastrointestinal cancer. Genomics Proteomics Bioinformatics 16:33–49. https://doi.org/10.1016/j.gpb.2017.06.002
8. Tilg H, Cani PD, Mayer EA (2016) Gut microbiome and liver diseases. Gut. https://doi.org/10.1136/gutjnl-2016-312729
9. Fernández M, Reina-Pérez I, Astorga J et al (2018) Breast cancer and its relationship with the microbiota. Int J Environ Res Public Health 15:1747. https://doi.org/10.3390/ijerph15081747
10. Shively CA, Register TC, Appt SE et al (2018) Consumption of mediterranean versus western diet leads to distinct mammary

gland microbiome populations. Cell Rep 25:47–56.e3. https://doi.org/10.1016/j.celrep.2018.08.078
11. Hieken TJ, Chen J, Hoskin TL et al (2016) The microbiome of aseptically collected human breast tissue in benign and malignant disease. Sci Rep 6:30751. https://doi.org/10.1038/srep30751
12. Banerjee S, Tian T, Wei Z et al (2018) Distinct microbial signatures associated with different breast cancer types. Front Microbiol. https://doi.org/10.3389/fmicb.2018.00951
13. Meng S, Chen B, Yang J et al (2018) Study of microbiomes in aseptically collected samples of human breast tissue using needle biopsy and the potential role of in situ tissue microbiomes for promoting malignancy. Front Oncol. https://doi.org/10.3389/fonc.2018.00318
14. Thompson KJ, Ingle JN, Tang X et al (2017) A comprehensive analysis of breast cancer microbiota and host gene expression. PLoS ONE 12:e0188873. https://doi.org/10.1371/journal.pone.0188873
15. Urbaniak C, Cummins J, Brackstone M et al (2014) Microbiota of human breast tissue. Appl Environ Microbiol 80:3007–3014. https://doi.org/10.1128/AEM.00242-14
16. Urbaniak C, Gloor GB, Brackstone M et al (2016) The microbiota of breast tissue and its association with breast cancer. Appl Environ Microbiol 82:5039–5048. https://doi.org/10.1128/AEM.01235-16
17. Wang H, Altemus J, Niazi F et al (2017) Breast tissue, oral and urinary microbiomes in breast cancer. Oncotarget 8:88122–88138. https://doi.org/10.18632/oncotarget.21490
18. Xuan C, Shamonki JM, Chung A et al (2014) Microbial dysbiosis is associated with human breast cancer. PLoS ONE 9:e83744. https://doi.org/10.1371/journal.pone.0083744
19. Smith A, Pierre JF, Makowski L et al (2019) Distinct microbial communities that differ by race, stage, or breast-tumor subtype in breast tissues of non-Hispanic Black and non-Hispanic White women. Sci Rep 9:1–10. https://doi.org/10.1038/s41598-019-48348-1
20. Costantini L, Magno S, Albanese D et al (2018) Characterization of human breast tissue microbiota from core needle biopsies through the analysis of multi hypervariable 16S-rRNA gene regions. Sci Rep 8:1–9. https://doi.org/10.1038/s41598-018-35329-z
21. Caporaso JG, Kuczynski J, Stombaugh J et al (2010) QIIME allows analysis of high-throughput community sequencing data. Nat Methods 7:335–336. https://doi.org/10.1038/nmeth.f.303
22. Pylro VS, Roesch LFW, Ortega JM et al (2014) Brazilian microbiome project: revealing the unexplored microbial diversity—challenges and prospects. Microb Ecol 67:237–241. https://doi.org/10.1007/s00248-013-0302-4
23. Edgar RC (2010) Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26:2460–2461. https://doi.org/10.1093/bioinformatics/btq461
24. Wang Q, Garrity GM, Tiedje JM, Cole JR (2007) Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Appl Environ Microbiol 73:5261–5267. https://doi.org/10.1128/AEM.00062-07
25. DeSantis TZ, Hugenholtz P, Larsen N et al (2006) Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. Appl Environ Microbiol 72:5069–5072. https://doi.org/10.1128/AEM.03006-05
26. Price MN, Dehal PS, Arkin AP (2010) FastTree 2—approximately maximum-likelihood trees for large alignments. PLoS ONE 5:e9490. https://doi.org/10.1371/journal.pone.0009490
27. McMurdie PJ, Holmes S (2013) phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. PLoS ONE 8:e61217. https://doi.org/10.1371/journal.pone.0061217
28. Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. https://doi.org/10.1186/s13059-014-0550-8
29. Iwai S, Weinmaier T, Schmidt BL et al (2016) Piphillin: improved prediction of metagenomic content by direct inference from human microbiomes. PLoS ONE 11:e0166104. https://doi.org/10.1371/journal.pone.0166104
30. RStudio Team (2018) RStudio: integrated development for R. RStudio Inc, Boston, MA
31. Oksanen J, Guillaume FB, Friendly M et al (2019) vegan: Community Ecology Package. R package
32. Mallick H, McIver LJ, Rahnavard A et al (2020) Multivariable association in population-scale meta-omics studies
33. Morris EK, Caruso T, Buscot F et al (2014) Choosing and using diversity indices: insights for ecological applications from the German Biodiversity Exploratories. Ecol Evol 4:3514–3524. https://doi.org/10.1002/ece3.1155
34. Pellegrini S, Sordi V, Bolla AM et al (2017) Duodenal mucosa of patients with type 1 diabetes shows distinctive inflammatory profile and microbiota. J Clin Endocrinol Metab 102:1468–1477. https://doi.org/10.1210/jc.2016-3222
35. Zhou J, Yao Y, Jiao K et al (2017) Relationship between gingival crevicular fluid microbiota and cytokine profile in periodontal host homeostasis. Front Microbiol. https://doi.org/10.3389/fmicb.2017.02144
36. Tobias DK, Akinkuolie AO, Chandler PD et al (2018) Markers of inflammation and incident breast cancer risk in the women's health study. Am J Epidemiol 187:705–716. https://doi.org/10.1093/aje/kwx250
37. Sanapareddy N, Legge RM, Jovov B et al (2012) Increased rectal microbial richness is associated with the presence of colorectal adenomas in humans. ISME J 6:1858–1868. https://doi.org/10.1038/ismej.2012.43
38. Chng KR, Chan SH, Ng AHQ et al (2016) Tissue microbiome profiling identifies an enrichment of specific enteric bacteria in Opisthorchis viverrini associated Cholangiocarcinoma. EBioMedicine 8:195–202. https://doi.org/10.1016/j.ebiom.2016.04.034
39. Yang J, McDowell A, Kim EK et al (2019) Development of a colorectal cancer diagnostic model and dietary risk assessment through gut microbiome analysis. Exp Mol Med 51:1–15. https://doi.org/10.1038/s12276-019-0313-4
40. Sun T, Liu S, Zhou Y et al (2016) Evolutionary biologic changes of gut microbiota in an 'adenoma-carcinoma sequence' mouse colorectal cancer model induced by 1,2-dimethylhydrazine. Oncotarget 8:444–457. https://doi.org/10.18632/oncotarget.13443
41. Peters BA, Wu J, Pei Z et al (2017) Oral microbiome composition reflects prospective risk for esophageal cancers. Cancer Res 77:6777–6787. https://doi.org/10.1158/0008-5472.CAN-17-1296
42. Shaw KA, Bertha M, Hofmekler T et al (2016) Dysbiosis, inflammation, and response to treatment: a longitudinal study of pediatric subjects with newly diagnosed inflammatory bowel disease. Genome Med. https://doi.org/10.1186/s13073-016-0331-y
43. Liu G, Tang CM, Exley RM (2015) Non-pathogenic Neisseria: members of an abundant, multi-habitat, diverse genus. Microbiology 161:1297–1312. https://doi.org/10.1099/mic.0.000086
44. Iglewski BH (1996) Pseudomonas. In: Baron S (ed) Medical microbiology, 4th ed. University of Texas Medical Branch at Galveston, Galveston (TX)
45. Cavarretta I, Ferrarese R, Cazzaniga W et al (2017) The microbiome of the prostate tumor microenvironment. Eur Urol 72:625–631. https://doi.org/10.1016/j.eururo.2017.03.029
46. Wang Y, Xue J, Zhou X et al (2014) Oral microbiota distinguishes acute lymphoblastic leukemia pediatric hosts from healthy populations. PLoS ONE. https://doi.org/10.1371/journal.pone.0102116

47. Biagi E, Quercia S, Aceti A et al (2017) The bacterial ecosystem of mother's milk and infant's mouth and gut. Front Microbiol. https://doi.org/10.3389/fmicb.2017.01214
48. Ruiz L, Bacigalupe R, García-Carral C et al (2019) Microbiota of human precolostrum and its potential role as a source of bacteria to the infant mouth. Sci Rep 9:1–13. https://doi.org/10.1038/s41598-019-42514-1
49. Moreno I, Codoñer FM, Vilella F et al (2016) Evidence that the endometrial microbiota has an effect on implantation success or failure. Am J Obstet Gynecol 215:684–703. https://doi.org/10.1016/j.ajog.2016.09.075
50. McGovern E, Waters SM, Blackshields G, McCabe MS (2018) Evaluating established methods for Rumen 16S rRNA amplicon sequencing with mock microbial populations. Front Microbiol. https://doi.org/10.3389/fmicb.2018.01365
51. Siegwald L, Touzet H, Lemoine Y et al (2017) Assessment of common and emerging bioinformatics pipelines for targeted metagenomics. PLoS ONE. https://doi.org/10.1371/journal.pone.0169563
52. Female Breast Cancer Subtypes - Cancer Stat Facts. In: SEER. https://seer.cancer.gov/statfacts/html/breast-subtypes.html. Accessed 8 Oct 2019

Affiliations

Emily Klann[1,2] · Jessica M. Williamson[1] · Massimiliano S. Tagliamonte[2,3] · Maria Ukhanova[2] · Jaya Ruth Asirvatham[3] · Harvey Chim[4] · Lusine Yaghjyan[1] · Volker Mai[1,2]

Emily Klann
emilyklann@ufl.edu

Jessica M. Williamson
jwilliamson1@ufl.edu

Massimiliano S. Tagliamonte
mstagliamonte@ufl.edu

Maria Ukhanova
mukhanova@epi.ufl.edu

Jaya Ruth Asirvatham
ruthasirvatham@ufl.edu

Harvey Chim
harvey.chim@surgery.ufl.edu

Lusine Yaghjyan
lyaghjyan@ufl.edu

[1] Department of Epidemiology, College of Public Health and Health Professions and College of Medicine, University of Florida, Gainesville, USA

[2] Emerging Pathogens Institute, University of Florida, Gainesville, USA

[3] Department of Pathology, Immunology and Laboratory Medicine, College of Medicine, University of Florida, Gainesville, USA

[4] Division of Plastic and Reconstructive Surgery, University of Florida, Gainesville, USA