

UC San Diego

UC San Diego Previously Published Works

Title

Gut and oral microbial compositional differences in women with breast cancer, women with ductal carcinoma in situ, and healthy women.

Permalink

<https://escholarship.org/uc/item/9sv9w98w>

Journal

mSystems, 9(11)

Authors

McCune, Emma
Sharma, Anukriti
Johnson, Breanna
et al.

Publication Date

2024-11-19

DOI

10.1128/msystems.01237-24

Peer reviewed

Gut and oral microbial compositional differences in women with breast cancer, women with ductal carcinoma *in situ*, and healthy women

Emma McCune,¹ Anukriti Sharma,² Breanna Johnson,¹ Tess O'Meara,¹ Sarah Theiner,¹ Maribel Campos,¹ Diane Heditsian,¹ Susie Brain,¹ Jack A. Gilbert,³ Laura Esserman,¹ Michael J. Campbell¹

AUTHOR AFFILIATIONS See affiliation list on p. 15.

ABSTRACT This study characterized and compared the fecal and oral microbiota from women with early-stage breast cancer (BC), women with ductal carcinoma *in situ* (DCIS), and healthy women. Fecal and oral samples were collected from newly diagnosed patients prior to any therapy and characterized using 16S rRNA sequencing. Measures of gut microbial alpha diversity were significantly lower in the BC vs healthy cohort. Beta diversity differed significantly between the BC or DCIS and healthy groups, and several differentially abundant taxa were identified. Clustering (non-negative matrix factorization) of the gut microbiota identified five bacterial guilds dominated by *Prevotella*, Enterobacteriaceae, *Akkermansia*, Clostridiales, or *Bacteroides*. The *Bacteroides* and Enterobacteriaceae guilds were significantly more abundant in the BC cohort compared to healthy controls, whereas the Clostridiales guild was more abundant in the healthy group. Finally, prediction of functional pathways identified 23 pathways that differed between the BC and healthy gut microbiota including lipopolysaccharide biosynthesis, glycan biosynthesis and metabolism, lipid metabolism, and sphingolipid metabolism. In contrast to the gut microbiomes, there were no significant differences in alpha or beta diversity in the oral microbiomes, and very few differentially abundant taxa were observed. Non-negative matrix factorization analysis of the oral microbiota samples identified seven guilds dominated by *Veillonella*, *Prevotella*, Gemellaceae, *Haemophilus*, *Neisseria*, *Propionibacterium*, and *Streptococcus*; however, none of these guilds were differentially associated with the different cohorts. Our results suggest that alterations in the gut microbiota may provide the basis for interventions targeting the gut microbiome to improve treatment outcomes and long-term prognosis.

IMPORTANCE Emerging evidence suggests that the gut microbiota may play a role in breast cancer. Few studies have evaluated both the gut and oral microbiomes in women with breast cancer (BC), and none have characterized these microbiomes in women with ductal carcinoma *in situ* (DCIS). We surveyed the gut and oral microbiomes from women with BC or DCIS and healthy women and identified compositional and functional features of the gut microbiota that differed between these cohorts. In contrast, very few differential features were identified in the oral microbiota. Understanding the role of gut bacteria in BC and DCIS may open up new opportunities for the development of novel markers for early detection (or markers of susceptibility) as well as new strategies for prevention and/or treatment.

KEYWORDS gut microbiota, oral microbiota, breast cancer, DCIS

Breast cancer (BC) is a complex disease: its development has been linked to environmental, genetic, and biological risk factors, its progression can span decades,

Editor Lindsay Kalan, McMaster University, Hamilton, Ontario, Canada

Address correspondence to Michael J. Campbell, michael.campbell@ucsf.edu.

The authors declare no conflict of interest.

See the funding table on p. 16.

Received 10 September 2024

Accepted 5 October 2024

Published 29 October 2024

Copyright © 2024 McCune et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

and its clinical course is quite variable (1, 2). However, over half of the women who develop breast cancer have no known risk factors, and few of the women with a genetic predisposition to breast cancer or who have been exposed to known environmental risk factors go on to develop the disease (3, 4). Clearly, additional contributing factors need to be identified. One such factor that has gained recent attention is the human microbiome, the collection of commensal microbes living within and on the human body (5, 6).

Ductal carcinoma *in situ* (DCIS) is a heterogeneous proliferative condition with an increased risk of developing invasive breast cancer. While only 15%–45% of DCIS progresses to invasive cancer, this is still considered sufficient to warrant an aggressive surgical approach given the lack of criteria to distinguish indolent from aggressive disease. Here again, the human microbiome may play a role.

The gut microbiome harbors thousands of bacterial species that are affected by host genetics, lifestyle, and environmental factors. Alterations in these microbial communities have been linked to a variety of diseases, including inflammatory bowel disease (7, 8), obesity (9), diabetes (10, 11), rheumatoid arthritis (12), cardiovascular diseases (13), and cancer (14–20). Gut microbes have been shown to influence local and systemic immunity and metabolism (20–23), including systemic levels of estrogen and its metabolites (16, 18, 24, 25). Comparisons of the gut microbiota of women with breast cancer and healthy women have revealed differences in the quantity and diversity of various genera (26–31). In addition, antibiotic use has been linked to an increased risk of breast cancer, likely due to its effects on the gut microbiota (32).

A link between periodontal disease and cancer has been suggested through several studies looking both at specific types of cancers and at the overall total cancer rate and the relationship to periodontal disease (33). The scientific rationale behind this association is that inflammation is a major factor in both periodontal disease and cancer. With respect to breast cancer, Hujuel and coworkers (34) found a weak association between periodontal disease and breast cancer risk. Swedish researchers have reported a significantly increased risk of breast cancer associated with chronic periodontal disease (odd ratio = 2.36) (35, 36). Although these results suggest a link between breast cancer and the oral microbiota, two previous studies found no differences in microbial alpha or beta diversity and very few, if any, differentially abundant taxa in the oral microbiota from women with breast cancer compared to that from healthy women (37, 38).

Given the microbiome's potential involvement in carcinogenesis, we hypothesized that changes in the oral and/or gut microbiota might be associated with the development of breast cancer. Moreover, there have been no studies evaluating the microbiomes of women with DCIS. Therefore, in this study, we characterized the gut and oral microbiota of women with newly diagnosed invasive BC or DCIS and compared these to the gut and oral microbiota of healthy women recruited from an age-matched population.

RESULTS

Participant characteristics

A total of 185 women were enrolled and received home sample collection kits for this study. This included 85 women with breast cancer, 40 women with ductal carcinoma *in situ*, and 60 healthy women (Fig. S1A). Eighty-three percent ($n = 154$) of the sample kits were returned (73 BC, 32 DCIS, and 49 healthy). As shown in Table 1, women in each of the three groups were similar in age, body mass index (BMI), and breast density. Approximately 40% of oral samples failed to yield sequencing results, typically due to low DNA yield, whereas only 10% of the gut samples failed to yield sequencing results. Overall, we analyzed 137 gut samples (66 from women with BC, 29 from women with DCIS, and 42 from healthy women) and 93 oral samples (47 from women with BC, 15 from women with DCIS, and 31 from healthy women) (Fig. S1A). Paired gut and oral microbiome data from the same individual were obtained from 40 BC, 12 DCIS, and 24 healthy participants (Fig. S1B).

TABLE 1 Demographics of study participants^{a,b,c,d}

| Variable | Healthy (n = 49) | DCIS (n = 32) | BC (n = 73) | Healthy vs DCIS | Healthy vs BC | DCIS vs BC |
|----------------|------------------|---------------|-------------|-----------------|---------------|------------|
| Age (years) | 53 ± 10 | 54 ± 8 | 56 ± 11 | 0.762* | 0.191* | 0.368* |
| BMI | 25 ± 5 | 25 ± 5 | 25 ± 6 | 0.940* | 0.770* | 0.742* |
| Breast density | | | | | | |
| Low | 20 (41%) | 10 (31%) | 20 (27%) | 0.383** | 0.122** | 0.687** |
| High | 29 (59%) | 22 (69%) | 53 (73%) | | | |

^aAge and BMI values are presented as means ± standard deviations.

^bBreast density values are the number of cases (percentages).

^cMammographic breast density was combined into low (BI-RADS categories A and B) and high (categories C and D).

^d*P value Student's *t*-test; **P value chi-square.

Variations in gut and oral microbiota diversity in women with breast cancer, women with DCIS, and healthy women

The microbial α -diversity of gut and oral samples was compared between cohorts based on Shannon and Simpson diversity indices. As shown in Fig. 1A, the gut microbiota of women with BC had a statistically significant lower α -diversity compared to healthy women as indicated by the Simpson index ($P = 0.019$). No differences in gut microbial α -diversity were observed between women with DCIS and women with BC or healthy women. The Shannon and Simpson indices for the oral microbiome did not differ across the three cohorts (Fig. 1C).

Differences in β -diversity were evaluated using principal-coordinate analyses (PCoA) based on Bray-Curtis distance metrics and pairwise permutational multivariate analysis of variance (PERMANOVA) tests. Gut microbial β -diversity was significantly different between BC and healthy groups ($P = 0.002$), as well as between the DCIS and healthy groups ($P = 0.036$) (Fig. 1B). No differences in oral microbial β -diversity were observed between any group (Fig. 1D).

Comparison of gut and oral microbiota composition between women with breast cancer, women with DCIS, and healthy women

Taxa (phyla, families, and genera) with a relative abundance $\geq 1\%$ in the gut and oral microbiota in each cohort are presented in Fig. S2. Firmicutes and Bacteroidetes were the two most abundant phyla (together accounting for $\sim 90\%$) in the gut microbiota of all three cohorts (Fig. S2A). The fecal Firmicutes to Bacteroidetes ratio (F/B ratio) in the healthy group was significantly higher than in the BC group ($P = 0.027$; Fig. S3A) but similar to that in the DCIS group ($P = 0.55$). In the oral microbiota, Firmicutes, Proteobacteria, and Bacteroidetes were the three most abundant phyla (together accounting for 80%–90%) across all three cohorts (Fig. S2B), and the oral F/B ratios were similar in all groups (Fig. S3B).

Seventeen families were identified in the gut microbiota with a relative abundance of $\geq 1\%$ (Fig. S2C). Lachnospiraceae, Bacteroidaceae, and Ruminococcaceae together represented $\sim 70\%$ of the relative abundance of family taxa. In the oral microbiota, 13 families were identified with a relative abundance of $\geq 1\%$, with Streptococcaceae representing $\sim 40\%$ – 50% and Pasteurellaceae, Veillonellaceae, Gemellaceae, Prevotellaceae, and Neisseriaceae representing another 30%–40% (Fig. S2D).

At the genus level in the gut microbiota, 17 genera were identified with relative abundance of $\geq 1\%$ (Fig. S2E). *Bacteroides* was the most prevalent in all groups (BC: 28%, DCIS: 23%, and healthy: 20%) followed by *Blautia* (BC: 7%, DCIS: 6%, and healthy: 8%), *Faecalibacterium* (BC: 5%, DCIS: 4%, and healthy: 6%), *Ruminococcus* (BC: 5%, DCIS: 5%, and healthy: 4%), *Prevotella* (BC: 3%, DCIS: 5%, and healthy: 7%), *Akkermansia* (BC: 3%, DCIS: 3%, and healthy: 3%), and *Coprococcus* (BC: 2%, DCIS: 3%, and healthy: 3%). The remaining genera each represented $\leq 2\%$ of the relative abundance in each cohort. Of the 13 genera identified with a relative abundance of $\geq 1\%$ in the oral microbiota (Fig. S2F), *Streptococcus* was the most prevalent genus in all groups (BC: 42%, DCIS: 54%, and

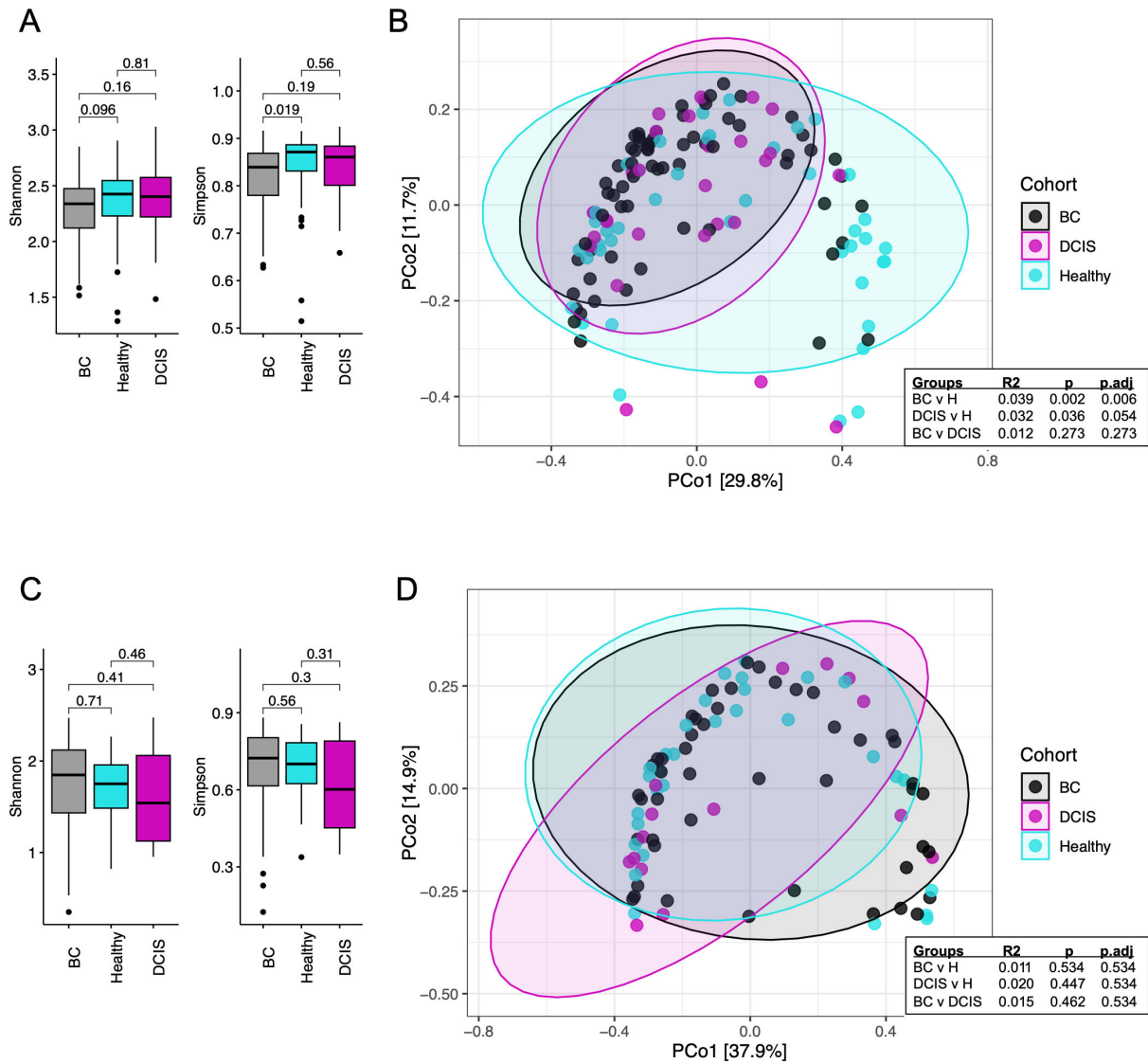


FIG 1 Diversity analysis of the oral and gut microbiota among the three cohorts. Shannon and Simpson alpha-diversity analyses at the genus level are shown for gut (A) and oral (C) samples. *P*-values were calculated using Wilcoxon tests. Principal-coordinate analysis (PCoA) based on Bray-Curtis distance at the genus level is shown for gut (B) and oral (D) samples. *R*² and *P*-values were calculated using pairwise permutational multivariate analysis of variance tests.

healthy: 48%) followed by *Veillonella* (BC: 11%, DCIS: 9%, and healthy: 9%), *Haemophilus* (BC: 10%, DCIS: 6%, and healthy: 8%), *Prevotella* (BC: 7%, DCIS: 7%, and healthy: 5%), *Neisseria* (BC: 5%, DCIS: 3%, and healthy: 5%), *Gemella* (BC: 2%, DCIS: 1%, and healthy: 3%), and *Rothia* (BC: 1%, DCIS: 3%, and healthy: 1%). The remaining genera each represented $\leq 2\%$ of the relative abundance in each cohort.

Identification of microbial taxa that differentiate DCIS, IDC, and healthy oral and gut microbiota

We used linear discriminant analysis effect size (LEfSe) analysis (with an LDA score cutoff ≥ 3) to identify key taxa responsible for the differences in the compositions of the oral and fecal microbiota between the three groups (Fig. 2). In the fecal microbiota, six taxa were more abundant in the BC group compared to the healthy group, including the family Bacteroidaceae and its corresponding genus *Bacteroides*; the class Actinobacteria; the family [Tissierellaceae] and corresponding genus *Fingoldia*; and the phylum Tenericutes (Fig. 2A). The phylum Firmicutes and its corresponding class Clostridia, order

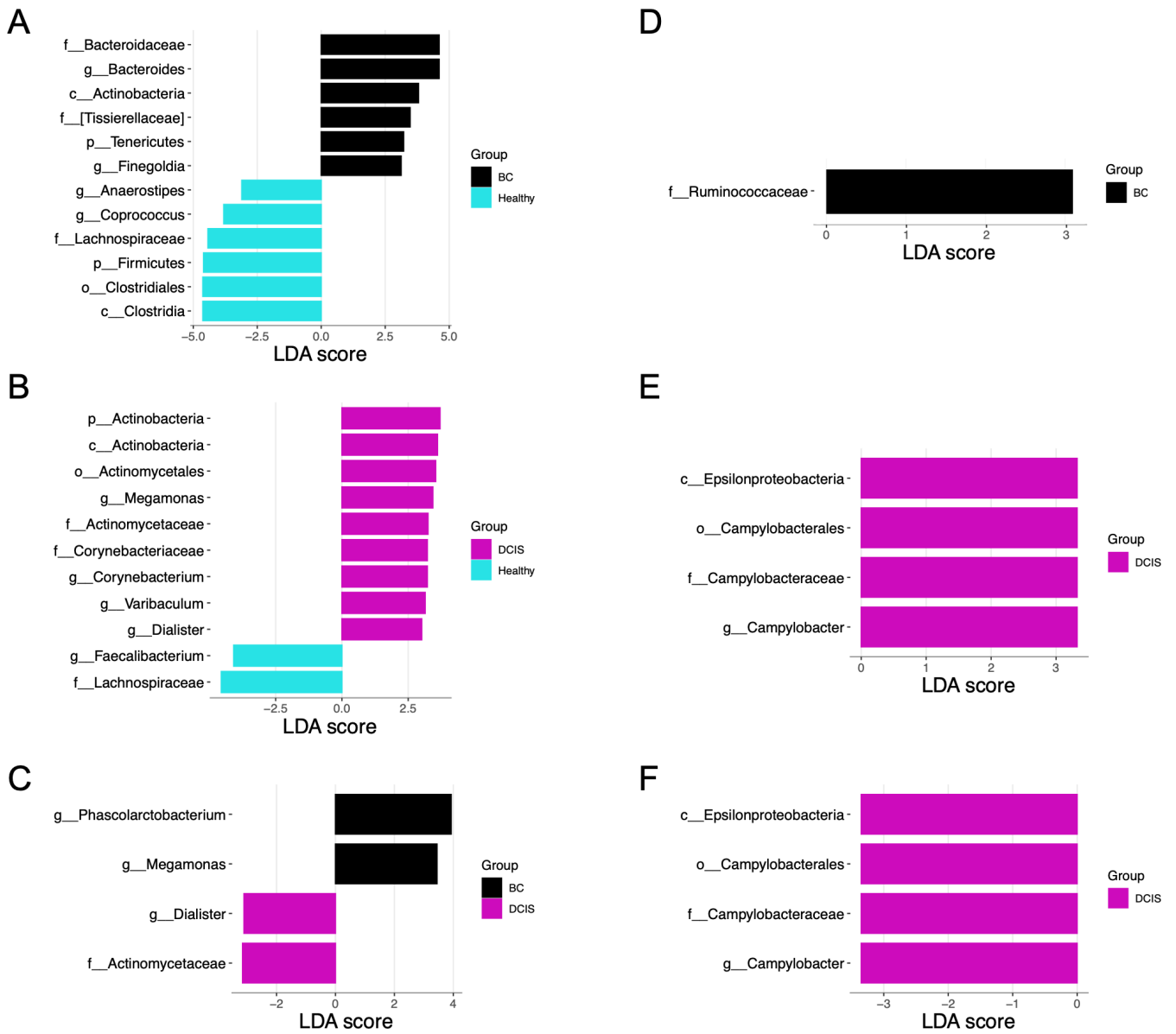


FIG 2 Taxonomic biomarkers identified by LEfSe. LDA bar graphs plotted from LEfSe analyses of gut (A–C) and oral (D–F) microbiota. The color of the bars represents the cohort. The logarithmic LDA score threshold was 3.0 for discriminative features. (A and D) BC vs healthy; (B and E) DCIS vs healthy; (C and F) BC vs DCIS.

Clostridiales, family Lachnospiraceae, and genera *Coprococcus* and *Anaerostipes* were more abundant in the healthy group compared to the BC group (Fig. 2A). Comparing the fecal microbiota of the healthy and DCIS groups, nine taxa were more abundant in the DCIS group (Fig. 2B). These included the phylum Actinobacteria and its corresponding class Actinobacteria, order Actinomycetales, family Actinomycetaceae, and genus *Varibaculum*; the family Corynebacteriaceae and the corresponding genus *Corynebacterium*; and the genera *Dialister* and *Megamonas*. Two taxa, the family Lachnospiraceae and the corresponding genus *Faecalibacterium*, were more abundant in the healthy group compared to the DCIS group (Fig. 2B). Finally, only four taxa were differentially abundant in the BC vs DCIS groups (Fig. 2C). The genera *Phascolarctobacterium* and *Megamonas* were enriched in the BC gut microbiota, whereas the family Actinomycetaceae and genus *Dialister* were enriched in the DCIS group.

Analysis of the oral microbiota identified the family Ruminococcaceae enriched in the BC group compared to the healthy cohort (Fig. 2D). The class Epsilonproteobacteria and corresponding order Campylobacterales, family Campylobacteraceae, and genus *Campylobacter* were more abundant in the DCIS gut microbiota compared to the BC and the healthy groups (Fig. 2E and F).

DESeq2 analyses were also used to identify differentially abundant taxa at the genus level. As shown in Fig. 3A, comparing the gut microbiota of the healthy cohort with the BC cohort, the genera *Coprobacillus*, *Parabacteroides*, *Streptococcus*, *WAL*, *Corynebacterium*, *Anaerococcus*, *Bacteroides*, *Acidaminococcus*, *Eggerthella*, *Peptoniphilus*, and *Finegoldia* were enriched in the BC cohort compared to the healthy cohort, while *Anaerostipes* was enriched in the healthy cohort. Comparing the healthy cohort with the DCIS cohort (Fig. 3B), DESeq2 analysis identified 16 genera enriched in the DCIS cohort, including seven genera that were also differentially enriched in BC vs healthy (*Finegoldia*, *Peptoniphilus*, *Eggerthella*, *Anaerococcus*, *Corynebacterium*, *WAL*, and *Streptococcus*). Only one genus, [*Prevotella*], was enriched in the healthy cohort compared to the DCIS group. Finally, comparing the DCIS cohort with the BC cohort, 13 genera were more abundant in the gut microbiota of the DCIS cohort compared to the BC cohort, and only 1, *Phascolarctobacterium*, was enriched in the BC group (Fig. 3C).

DESeq2 analyses of the oral microbiota found no genera significantly enriched when comparing the BC and healthy groups (Fig. 3D) or comparing the BC and DCIS groups (Fig. 3F). Comparing the DCIS and healthy oral microbiota, two genera, *Actinobacillus* and *Blautia*, were found to be differentially enriched (higher in healthy vs DCIS) (Fig. 3E).

Gut and oral microbial guilds associated with breast cancer

Since community structure may be more informative than abundance differences of individual taxa, we clustered the gut and oral microbiota into guilds using non-negative matrix factorization (NMF). We identified five guilds in the gut microbiota samples and named them according to the most notable taxa: G1-*Prevotella*, G2-Enterobacteriaceae, G3-*Akkermansia*, G4-Clostridiales, and G5-*Bacteroides* (Fig. 4A). The G4-Clostridiales guild is also characterized by higher abundances of the related families Ruminococcaceae and Lachnospiraceae and the related genera *Blautia*, *Faecalibacterium*, and *Ruminococcus*. The G2-Enterobacteriaceae and G5-*Bacteroides* guilds were significantly more abundant in the gut microbiota of women with breast cancer compared to healthy women, whereas the G4-Clostridiales guild was significantly more abundant in the healthy cohort compared to the BC group (Fig. 4B).

NMF applied to the oral microbiota samples identified seven guilds that were named according to the most notable taxa: G1-*Veillonella*, G2-*Prevotella*, G3-Gemellaceae, G4-*Haemophilus*, G5-*Neisseria*, G6-*Propionibacterium*, and G7-*Streptococcus* (Fig. 4C). None of these guilds were differentially abundant across the cohorts (Fig. 4D).

Predicted functional potential of the oral and gut microbial communities

While the taxonomic classifications of the bacterial communities of the three cohorts are useful, we were also interested in the functional potential of the oral and gut microbiota and how this might vary across the cohorts. Different methods are available that use 16S rRNA sequence data to infer the microbial genomes present and their potential functionality in the absence of whole-genome sequencing data (39–41). It should be noted that these methods are based on inference from known genomes and may not fully recapitulate the existing metagenomic content.

We performed a predictive functional analysis on the gut 16S sequence data using the Tax4Fun algorithm and LEfSe analysis to identify pathways differentially present in the three cohorts (Fig. 5). Twenty-three pathways differed between the BC and healthy cohorts, 14 associated with the BC group and 9 with the healthy group (Fig. 5A). Modules enriched in the BC group included lipopolysaccharide biosynthesis, glycan biosynthesis and metabolism, lipid metabolism, galactose metabolism, and sphingolipid metabolism. Modules enriched in the healthy controls included membrane transport, ABC

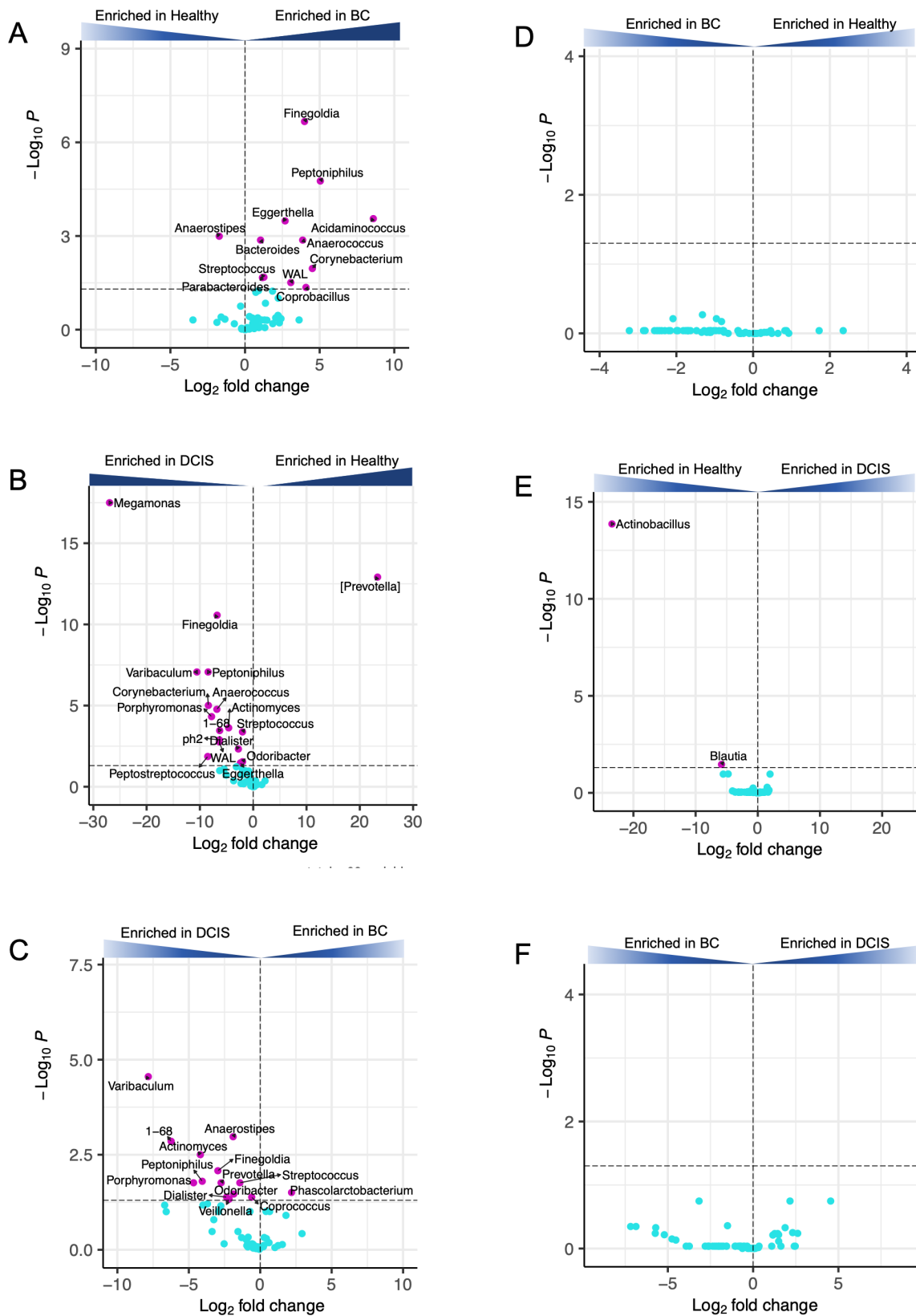


FIG 3 DESeq2 differential abundance analysis. Volcano plots of log₂ fold differences in genera abundance in the gut microbiota (A–C) or oral microbiota (D–F) between BC and healthy (A and D), DCIS and healthy (B and E), and BC and DCIS (C and F). Dashed horizontal lines reflect adjusted *P*-value = 0.05. Names of differentially abundant genera (*P*_{adj} > 0.05) are shown (magenta dots). Blue dots represent not significant genera.

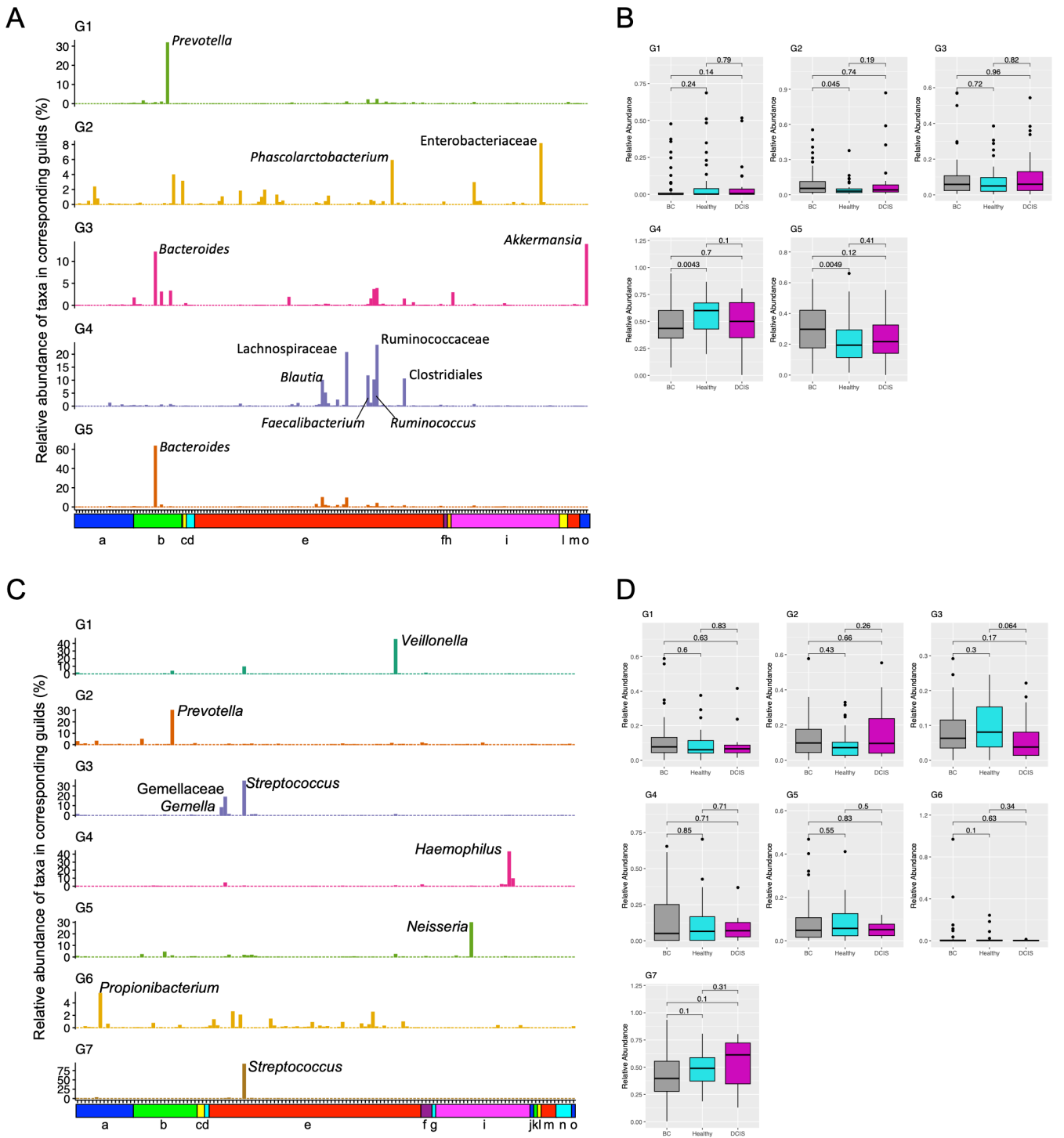


FIG 4 Microbiota guilds identified by non-negative matrix factorization analyses. (A and C) Each bar represents the relative abundance of a taxon in the guild (A: gut; C: oral). Color annotation bars indicate phyla for each taxon: (a) Actinobacteria, (b) Bacteroidetes, (c) Cyanobacteria, (d) Euryarchaeota, (e) Firmicutes, (f) Fusobacteria, (g) GN02, (h) Lentisphaerae, (i) Proteobacteria, (j) Spirochaetes, (k) SR1, (l) Synergistetes, (m) Tenericutes, (n) TM7, and (o) Verrucomicrobia. (B and D) Box plots showing the relative abundance of each guild by cohort (B: gut; D: oral). *P*-values were calculated using Wilcoxon tests.

transporters, cell motility, bacterial chemotaxis, flagellar assembly, and arginine and proline metabolism. Only three pathways differed between the DCIS and healthy cohorts, all three being enriched in the DCIS samples (Fig. 5B). Finally, 16 pathways were identified as differentially present in the BC and DCIS cohorts, 8 enriched in the BC group

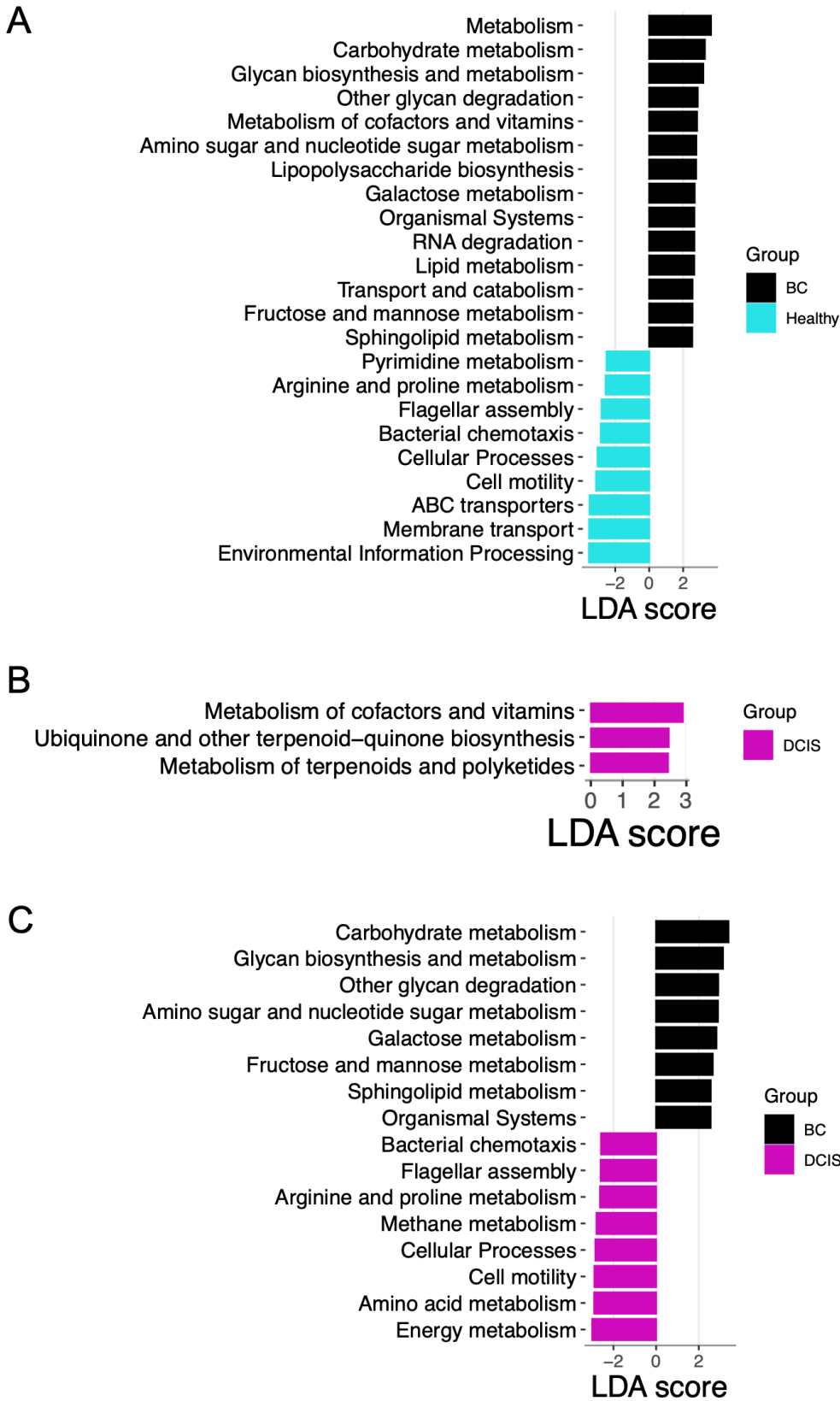


FIG 5 Functional analysis prediction. Tax4Fun was used to predict the functional potential of microbiota using 16S rRNA gene sequence data. Differentially enriched bacterial functions among groups were identified using LEfSe analysis (logarithmic LDA score threshold = 2.5; $P < 0.05$). The color of the bars represents the cohort. (A) BC vs healthy; (B) DCIS vs healthy; (C) BC vs DCIS.

and 8 enriched in the DCIS group (Fig. 5C). Interestingly, most of the modules that differentiated BC and DCIS also differentiated the BC and healthy groups.

Predictive functional analysis of the oral 16S sequence data demonstrated a lack of pathways discriminating the cohorts. No pathways were found to differentiate the BC and healthy or the DCIS and healthy cohorts, and one pathway (amino acid metabolism) was found to be enriched in the BC oral microbiota compared to the DCIS group (data not shown).

Associations of gut and oral microbiota with breast density

Mammographic breast density, categorized as high (BI-RADS A and B) or low (BI-RADS C and D), did not differ significantly across the three cohorts (Table 1). Alpha diversity of the fecal microbiota and the oral microbiota did not differ between women with high vs low mammographic density (Fig. 6A and B). There were no differences in the F/B ratios between the high and low breast density groups in either the gut or oral microbiota. LEfSe analysis identified seven taxa (LDA score ≥ 3) with differential abundance in the gut microbiota (Fig. 6C). Six taxa were more abundant in the high breast density group, including the family Bacteroidaceae and corresponding genus *Bacteroides*; the phylum Cyanobacteria and corresponding class 4C0d-2 and order YS2; and the genus *Phascolarctobacterium*. The family Christensenellaceae was enriched in the low-density group. No

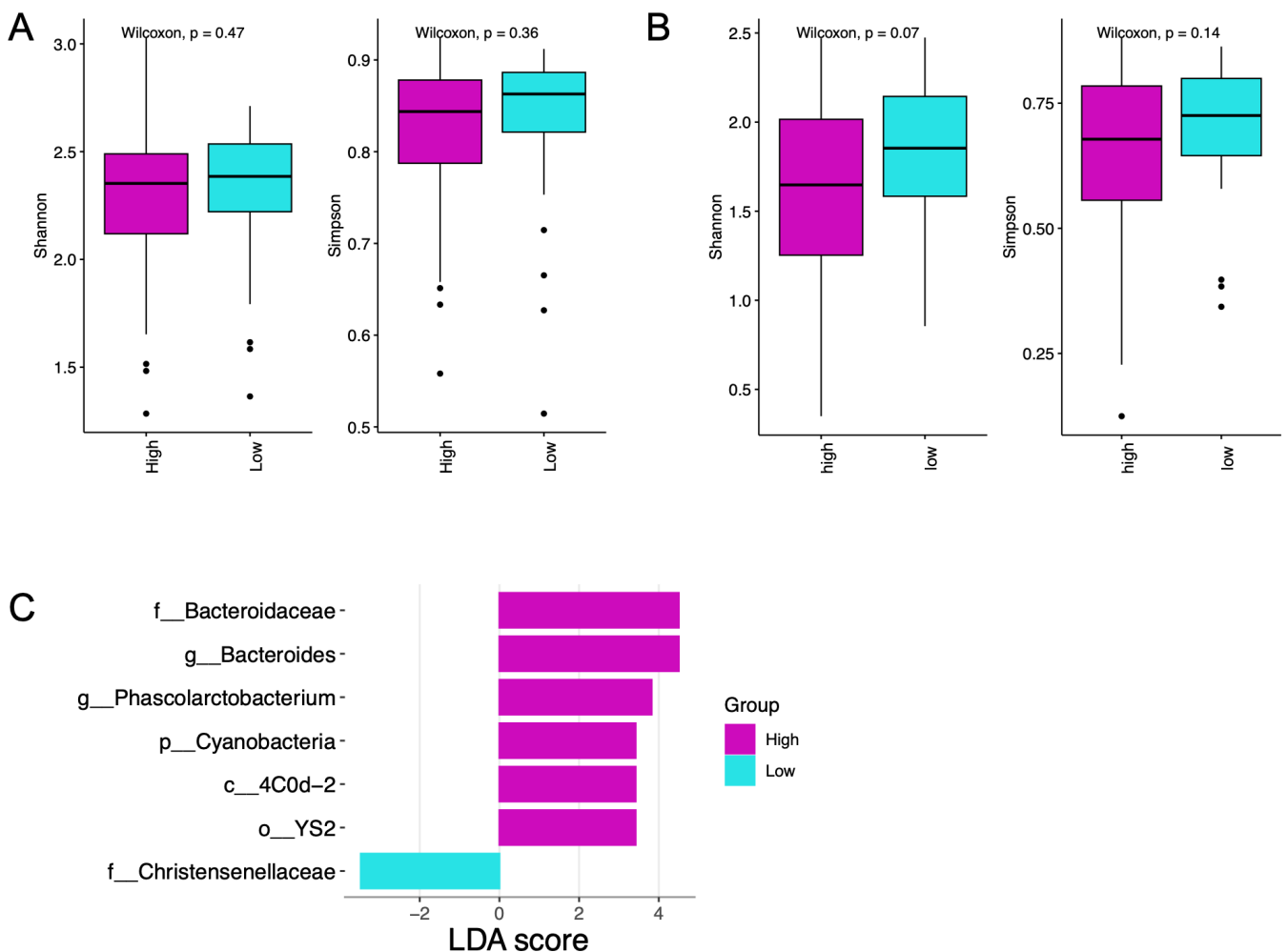


FIG 6 Oral and fecal microbiota features associated with mammographic breast density. Shannon and Simpson alpha-diversity analyses are shown for gut (A) and oral (B) samples comparing high and low breast density. (C) Differentially enriched taxa abundance in fecal samples between high and low breast density was identified using LEfSe analysis (logarithmic LDA score threshold = 3; $P < 0.05$).

taxa were differentially abundant in the oral microbiota when comparing high vs low breast density groups.

Overlapping taxa between the oral and gut microbiota

Because we characterized both gut and oral microbiota samples, we explored the similarity between these microbiota. As shown in Fig. S4A, the gut microbiota demonstrated significantly greater α -diversity compared to the oral microbiota. β -diversity, evaluated using PCoA based on Bray-Curtis distance metrics, significantly separated the gut and oral microbiota (Fig. S4B).

LEfSe analysis was used to identify differentially abundant taxa between the gut and oral microbiomes. Figure S5 shows the top 10 (by LDA score) differentially abundant taxa at the phylum, family, genus, and species levels. At the phylum level, 13 of the 15 phyla identified in the gut and oral microbiota were significantly different between the two sites (after P -value adjustment). Of the top 10 differentially abundant phyla, Bacteroidetes, Verrucomicrobia, Cyanobacteria, Tenericutes, and Euryarchaeota were more abundant in the gut microbiota, whereas Proteobacteria, Firmicutes, Actinobacteria, Fusobacteria, and TM7 were more abundant in the oral microbiota (Fig. S5A). At the family level, 55 of 77 identified families were differentially abundant in the gut vs oral microbiota. Of the top 10 differentially abundant families, Bacteroidaceae, Ruminococcaceae, Lachnospiraceae, and Verrucomicrobiaceae were enriched in the gut samples, while Streptococcaceae, Pasteurellaceae, Veillonellaceae, Gemellaceae, and Neisseriaceae were enriched in the oral samples (Fig. S5B). At the genus level, 102 of 143 identified genera were differentially abundant in the gut vs oral microbiota. Of the top 10 differentially abundant genera, *Bacteroides*, *Ruminococcus*, *Akkermansia*, and *Lachnospira* were enriched in the gut, and *Streptococcus*, *Haemophilus*, *Neisseria*, *Veillonella*, *Selenomonas*, and *Granulicatella* were enriched in the oral samples (Fig. S5C).

We next explored the similarity between the gut and oral microbiota within each cohort. Venn diagrams illustrate that in the BC group, 118 (43.4%) taxa were predominantly fecal, 74 (27.2%) were predominantly oral, and 80 (29.4%) overlapped between the gut and oral samples (Fig. S6A). In the DCIS group, 122 (52.1%) taxa were predominantly fecal, 79 (33.8%) were predominantly oral, and 85 (36.3%) overlapped between the two sites (Fig. S6B). Finally, of the healthy samples, 103 (44.2%) taxa were predominantly fecal, 68 (29.2%) were predominantly oral, and 62 (26.6%) overlapped (Fig. S6C).

DISCUSSION

In this study, we examined the oral and gut microbiomes of women with breast cancer, women with DCIS, and healthy women. Microbiota were analyzed in terms of diversity, taxonomic profiles, and predicted functions. We identified various compositional and functional features of the gut microbiota that differed between women with breast cancer, women with DCIS, and healthy women. We found that fecal microbial α diversity, as estimated by the Simpson index, was significantly lower in individuals with BC when compared to healthy individuals, whereas there was no statistically significant difference comparing women with DCIS and healthy women. Reduced α diversity of gut microbiota in women with BC has been reported in several previous studies (26–30).

Comparing the gut microbiota among the three cohorts, we found that the phylum Firmicutes showed a progressive decrease, while the phylum Bacteroidetes showed a progressive increase from healthy to DCIS to BC. This was reflected in the Firmicutes/Bacteroidetes ratio, which was significantly lower in the BC group compared to the healthy group. Consistent with the lower F/B ratio in the BC cohort, we also observed a significantly greater abundance of the family Bacteroidaceae and genus *Bacteroides* in the BC group vs the healthy cohort. A higher abundance of the genus *Bacteroides* in BC compared to healthy controls has been reported in previous studies (27, 28, 31). It has been observed that breast cancer risk increases with increasing levels of circulating parent estrogens and decreases with increasing ratios of estrogen metabolites to parent estrogens (24, 25, 42, 43). An association has also been observed between urinary

estrogens and estrogen metabolites, and the diversity and composition of the fecal microbiome. In postmenopausal women, high fecal microbial diversity is statistically significantly associated with a high ratio of estrogen metabolites to parent estrogens, while the abundance of *Bacteroides* is inversely associated with this ratio (25). These data suggest that women with high intestinal microbial diversity are at a lower risk for breast cancer, while women with a high intestinal abundance of *Bacteroides* may be at a higher risk for breast cancer. Our findings are consistent with this.

Specific microbial taxonomic differences vary widely across studies and a definitive disease-associated community structure has not been identified. This may be due to the large variation in gut microbial community composition among humans and/or to technical differences among the studies. However, several of the differentially abundant taxa in the gut microbiota we observed in our study are consistent with previous studies. In addition to the genus *Bacteroides* mentioned above, we also found the genera *Eggerthella* and *Peptoniphilis* and the class Actinobacteria enriched in the BC vs healthy cohort, consistent with previous studies (29, 44). In addition, Lachnospiraceae and *Coprococcus* were more abundant in the healthy cohort compared to BC, similar to what was previously reported (30, 44). Although there are no previous reports on the gut microbiome in women with DCIS, our findings that *Faecalibacterium* was more abundant in healthy vs DCIS gut microbiota and that *Peptoniphilis* and *Actinomyces* were enriched in the DCIS cohort are consistent with previous studies comparing BC with healthy controls (26, 28, 44).

Since community structure may be more informative than abundance differences of individual taxa, we clustered the fecal microbiota into guilds/enterotypes (ETs). Two clustering algorithms previously used to identify gut enterotypes are partitioning around medoids (PAMs) or Dirichlet multinomial mixture models. Previous studies have suggested that the human gut microbiome can be assigned to no ETs, two ETs (*Bacteroides* and *Prevotella* dominated), three ETs (enriched in *Bacteroides*, *Prevotella*, and *Ruminococcus*), or four ETs (same as the three ETs with an additional *Bacteroides*-dominated cluster) (45–48). Recently, non-negative matrix factorization was used to divide the human gut microbiome into five “enterosignatures” (ESs) dominated by *Bacteroides*, Firmicutes, *Prevotella*, *Bifidobacterium*, or *Escherichia* (49). The *Bacteroides*, Firmicutes, and *Prevotella* ESs were dominant in adult samples, while the *Bifidobacterium* and *Escherichia* ESs were more often observed in the infant samples. We applied NMF to our fecal microbiota samples and identified five guilds. The G5-*Bacteroides* and G1-*Prevotella* guilds were consistent with the previously described ETs/ESs, while the G4-Clostridiales guild was similar to the *Ruminococcus* and Firmicutes ETs/ESs.

We found that the G5-*Bacteroides* and G2-Enterobacteriaceae guilds were significantly enriched in the BC samples compared to healthy controls and the G4-Clostridiales guild was significantly enriched in the healthy samples compared to the BC group. Zhu and coworkers (50), using PAM clustering, identified two enterotypes (*Bacteroides* and *Prevotella* dominated) in fecal samples from women with breast cancer and healthy controls. However, they found no significant relationship between enterotype and disease status. The association of the G5-*Bacteroides* guild with breast cancer is consistent with our results and previous studies suggesting a greater proportion of the genus *Bacteroides* in BC compared to healthy controls (27, 28, 31). The G4-Clostridiales guild was more abundant in the healthy group compared to the BC group. This guild is characterized by a high proportion of the order Clostridiales, the families Lachnospiraceae and Ruminococcaceae, and the genera *Blautia*, *Faecalibacterium*, and *Ruminococcus*. Similar results have been reported in two other studies, where healthy controls had a greater proportion of Clostridiales, Ruminococcaceae, and *Faecalibacterium* (26, 28).

Breast cancer-associated alterations in the gut microbiota likely translate into alterations in gut microbial functions. We utilized Tax4Fun to predict functional pathways from the fecal 16S rRNA sequence data. Pathways enriched in the BC cohort compared to the healthy group included lipopolysaccharide biosynthesis, glycan biosynthesis and metabolism, lipid metabolism, galactose metabolism, and sphingolipid metabolism.

Lipopolysaccharide is a potent trigger of systemic inflammation (51) and has been suggested to play an important role in promoting tumor-associated inflammation (52–54). Sphingolipid metabolism also plays a role in intestinal inflammation (55). Enrichment of these pathways may induce systemic low-grade inflammation, thereby increasing the risk of developing or perhaps exacerbating breast cancer. We do not know if the dysbiosis observed in the BC cohort was present in these women prior to developing breast cancer. Interestingly, we only observed three differentially enriched functional pathways between the DCIS and healthy cohorts, and of the 16 pathways that differentiated DCIS and BC, most also differentiated BC and healthy gut microbiomes, indicating that the functionality of the gut microbiota from women with DCIS was more like that of healthy women than women with BC.

We found no differences in alpha or beta diversity in the oral microbiota of women with BC vs healthy controls. These results are consistent with the two other published reports on the oral microbiome and breast cancer (37, 38). While one previous report found no differences in the relative abundance of individual taxa (37), we identified one taxon from the family Ruminococcaceae that was in a significantly greater proportion in the BC cohort. We found no predicted metabolic pathways differentially enriched in the oral microbiota between the BC and healthy groups, which is consistent with the lack of differentially proportional taxa in the oral microbiota. Applying NMF to the oral microbiota, we identified seven guilds; however, there were no differences in guild memberships between the BC and healthy groups. We also compared the oral microbiota of women with DCIS and healthy women and found no differences in alpha diversity, very few differentially proportional taxa, no predicted pathway differences, and no differences in guild memberships. Overall, these results indicate that the oral microbiota are very similar between healthy women and women with DCIS or invasive breast cancer.

Breast density, as estimated by mammography, is a strong risk factor for breast cancer. In this study, we found that alpha diversity of fecal and oral microbiota did not differ between women with high vs low breast density. Mammographic breast density also was not associated with fecal or oral microbiota beta diversity. In addition, few taxa were differentially associated with high vs low breast density. These results are consistent with a previously published study on mammographic breast density associations with fecal microbiota (56).

The oral and gut microbiomes are well segregated due to the presence of an oral–gut barrier, represented by their physical distance and the presence of gastric acid and bile (57–60). Consistent with previous reports (6, 61, 62), the compositions of the oral and gut microbiota were found to differ greatly in the present study. However, impairment of the oral–gut barrier can allow inter-organ translocation and communication. While inflammatory bowel disease and colon cancer are both conditions where oral taxa are found to be enriched in the gut, recent improvements in quantifying the absolute bacterial load in oral and gut samples have revealed that oral taxa enrichment may be more a reflection of reduced gut bacterial loads (63). We analyzed the oral-gut microbiota overlap and found from Venn diagram analysis that 29%, 36%, and 27% of taxa overlapped across the three cohorts (BC, DCIS, and healthy), respectively. These results suggest that oral-gut inter-organ translocation was similar across these cohorts. Intriguingly, we did find that the cariogenic dental pathogen *Streptococcus* was enriched in the gut microbiota of women with breast cancer and DCIS compared to healthy women.

There are several limitations within our study that should be noted when reviewing our results. The first is that we did not control for various factors that are known to affect the microbiome, such as diet, probiotic use, physical activity, and environmental exposures. However, we did control for treatment exposure by collecting samples from the BC and DCIS cohorts prior to any therapy. Second, the sample size of each cohort is small, although still large enough to allow us to identify differences in community composition and individual taxon abundances, particularly in the gut samples. The lack

of differences seen in the oral samples may be due to the low biomass of microbes at this site, as well as the limited sample size. A methodological limitation is that 16S sequencing provides limited information compared to metagenomic sequencing (additional studies using metagenomic sequencing are underway). Finally, our cohorts were relatively homogeneous with respect to ethnic diversity, limiting our ability to identify microbial features in groups that are disproportionately affected by breast cancer.

In conclusion, our study is the first to analyze and compare both the oral and gut microbiota of women with DCIS, women with BC, and healthy women. We identified several distinguishing features in the gut microbiota associated with BC and DCIS, and studies might now be designed to deepen our understanding of these associations. In particular, guild-based analysis may provide an ecologically meaningful approach to understanding the relationship between the gut microbiota and breast disease. Lower gut microbial alpha diversity in women with breast cancer suggests that there is room for interventions targeting the microbiome to improve treatment outcomes and long-term prognosis. Longitudinal studies, for example, might focus on the association between the gut microbiome and chemotherapy, or breast cancer mutations such as BRCA. Interventional studies could be designed to explore the manipulation of the microbiome to determine its role in breast cancer prevention or treatment response. Not least, investigation into specific microbes and the elucidation of the metabolic, biochemical, and immunological processes in which they are involved may help us to better understand how the microbiome can influence the initiation and/or progression of breast cancer.

MATERIALS AND METHODS

Participant selection and enrollment

Participants eligible for enrollment in this cohort study had a recent diagnosis of invasive ductal carcinoma (IDC) or DCIS. Women who had started chemotherapy or hormone treatment, or who had a previous diagnosis of breast cancer, were excluded. Healthy women without a diagnosis of breast cancer or a history of breast cancer served as the control group. Participants with IDC or DCIS were identified at the UCSF Breast Care Center, while healthy participants were identified via the ATHENA network (www.athena-carenetwork.org). Participants were enrolled and samples were collected between May 2015 and January 2017.

Sample collection, processing, and sequencing

Kits containing material for collecting saliva and stool swab samples were distributed to participants for self-collection. Participants were asked to collect oral samples immediately upon awakening to avoid any immediate effects of daily activities (e.g., eating and brushing teeth) on the oral microbiome. Samples were collected prior to any systemic therapy (including chemotherapy, hormone therapy, and radiation) to avoid therapy-associated effects on the gut and oral microbiomes. Sample swabs were placed in a vial containing a lysis and stabilization buffer that preserves the DNA for transport at ambient temperatures and were stored at -80°C prior to batch DNA extraction. Samples were lysed using bead beating, and DNA was extracted by a guanidine thiocyanate silica column-based purification method using a liquid-handling robot (64, 65). PCR amplification of the 16S rRNA gene was performed with primers targeting the V4 variable region (515F: GTGCCAGCMGCCGCGGTAA and 806R: GGACTACHVGGGTWTC TAAT) (66). In addition, the primers contained Illumina tags and barcodes. Samples were barcoded with a unique combination of forward and reverse indexes allowing for simultaneous processing of multiple samples. PCR products were pooled, column-purified, and size-selected through microfluidic DNA fractionation (67). Consolidated libraries were quantified by quantitative real-time PCR using the Kapa Bio-Rad iCycler qPCR

kit on a BioRad MyiQ before loading into the sequencer. Sequencing was performed in a pair-end modality on the Illumina NextSeq 500 platform rendering 2 × 150 bp paired-end sequences.

Data analysis

For the 16S rRNA analysis, the raw reads were joined using `join_paired_ends.py` script followed by quality filtering and demultiplexing using `split_libraries_fastq.py` script in QIIME 1.9.1. The final set of demultiplexed sequences was then selected for amplicon exact sequence variant (ESV) picking using the DeBlur pipeline. In the pipeline, *de novo* chimeras and artifacts were removed, and ESVs present in less than 10 samples were removed using the *phyloseq* package. The final BIOM file comprising unique 12,931 ESVs with an average of 35,378 reads per sample was then used for further analyses.

We used R software packages *phyloseq* (68) and *microeco* (69) to conduct analyses of the gut and oral microbiomes. All plots were generated in R using *phyloseq*, *microeco*, *ggplot2*, and *ggpubr* packages. Data were not rarefied. Shannon and Simpson indices were used to estimate α diversity, and the variation between groups (β diversity) was tested using PERMANOVA. PCoA plots based on Bray-Curtis distance at the genus level were employed to illustrate β diversity variations.

Key taxa responsible for the differences in the oral and gut microbiota between the DCIS, breast cancer, and healthy cohorts were identified using the LefSe algorithm for biomarker discovery (70). We defined a significant α of 0.05 and an effect size threshold of 3 for these analyses. DESeq2 analyses (as implemented in the *microeco* R package) were also used to identify differentially abundant bacterial genera between the groups based on an FDR-adjusted *P*-value cutoff of 0.05.

Functional predictions of the gut and oral microbial communities were performed using Tax4Fun (39) as implemented in the *microeco* R package, using the SILVA123 database. LefSe analyses were used to analyze the differences in KEGG pathways between the three cohorts, with a significant α of 0.05 and an effect size threshold of 2.5.

Guild analyses

Non-negative matrix factorization was used to identify bacterial communities or guilds in the gut and oral microbiota samples. NMF was implemented using the *NMF* package in R (71). An optimal rank (number of guilds) was determined by performing a rank survey analysis on relative abundance matrices of taxa with ranks = 2–15. Unit Invariant Knee methodology, implemented with the R package *inflection* (72), was used to identify the knee/elbow point (optimal rank) of the NMF rank survey plot of residual sum of squares (73). NMF analyses were then run using the optimal rank values, and the resulting guilds were named according to the most notable taxa in each. The relative abundance of guilds in each sample was determined, and the sample was assigned to the guild with the highest relative abundance. The relative abundances of each guild in the three cohorts were compared using Wilcoxon tests.

ACKNOWLEDGMENTS

We would like to thank the women who participated in this study. We also thank Dr. Katrine Whiteson for her feedback on the manuscript.

This work was supported by the Breast Cancer Research Foundation and the California Breast Cancer Research Program. While engaged in this study, no authors had conflicts of interest to report.

AUTHOR AFFILIATIONS

¹Department of Surgery, University of California, San Francisco, California, USA

²Department of Quantitative Health Sciences, Cleveland Clinic, Cleveland, Ohio, USA

³Department of Pediatrics, University of California, San Diego, California, USA

AUTHOR ORCID*s*

Michael J. Campbell  <http://orcid.org/0000-0002-3788-9240>

FUNDING

| Funder | Grant(s) | Author(s) |
|---|----------|---------------------|
| Breast Cancer Research Foundation (BCRF) | | Laura Esserman |
| California Breast Cancer Research Program (CBCRP) | | Michael J. Campbell |

AUTHOR CONTRIBUTIONS

Emma McCune, Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review and editing | Anukriti Sharma, Data curation, Formal analysis, Investigation, Software, Visualization, Writing – original draft, Writing – review and editing | Breanna Johnson, Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review and editing | Tess O'Meara, Investigation, Visualization | Sarah Theiner, Investigation, Visualization | Maribel Campos, Investigation | Diane Heditsian, Investigation, Visualization, Writing – review and editing | Susie Brain, Visualization, Writing – review and editing | Jack A. Gilbert, Validation, Visualization, Writing – review and editing | Laura Esserman, Conceptualization, Funding acquisition, Resources, Validation, Writing – review and editing | Michael J. Campbell, Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review and editing

DATA AVAILABILITY

All fastq files for 16S sequencing were deposited in the NCBI Sequence Read Archive (SRA) (accession number [PRJNA1136994](https://doi.org/10.5281/zenodo.12775458)). Participant metadata and taxa counts for the gut and oral samples are available for download from <https://doi.org/10.5281/zenodo.12775458>. A STORMS (Strengthening The Organization and Reporting of Microbiome Studies) checklist (74) is available at <https://doi.org/10.5281/zenodo.12775468>.

ETHICS APPROVAL

Ethical approval was obtained from the UCSF Institutional Review Board, and subjects provided written consent for sample collection and subsequent analyses.

ADDITIONAL FILES

The following material is available [online](#).

Supplemental Material

Supplemental Figures (mSystems01237-24-s0001.pdf). Figures S1 to S6.

REFERENCES

- Rice MS, Eliassen AH, Hankinson SE, Lenart EB, Willett WC, Tamimi RM. 2016. Breast cancer research in the nurses' health studies: exposures across the life course. *Am J Public Health* 106:1592–1598. <https://doi.org/10.2105/AJPH.2016.303325>
- Anderson WF, Rosenberg PS, Prat A, Perou CM, Sherman ME. 2014. How many etiological subtypes of breast cancer: two, three, four, or more? *J Natl Cancer Inst* 106:dju165. <https://doi.org/10.1093/jnci/dju165>
- Lacey JV Jr, Kreimer AR, Buys SS, Marcus PM, Chang S-C, Leitzmann MF, Hoover RN, Prorok PC, Berg CD, Hartge P, Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial Project Team. 2009. Breast cancer epidemiology according to recognized breast cancer risk factors in the prostate, lung, colorectal and ovarian (PLCO) cancer screening trial cohort. *BMC Cancer* 9:84. <https://doi.org/10.1186/1471-2407-9-84>
- Madigan MP, Ziegler RG, Benichou J, Byrne C, Hoover RN. 1995. Proportion of breast cancer cases in the United States explained by well-established risk factors. *J Natl Cancer Inst* 87:1681–1685. <https://doi.org/10.1093/jnci/87.22.1681>
- Peterson J, Garges S, Giovanni M, McInnes P, Wang L, Schloss JA, Bonazzi V, McEwen JE, Wetterstrand KA, Deal C, et al. 2009. The NIH human

- microbiome project. *Genome Res* 19:2317–2323. <https://doi.org/10.1101/gr.096651.109>
6. Human Microbiome Project Consortium. 2012. Structure, function and diversity of the healthy human microbiome. *Nature* 486:207–214. <https://doi.org/10.1038/nature11234>
 7. Frank DN, St Amand AL, Feldman RA, Boedeker EC, Harpaz N, Pace NR. 2007. Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases. *Proc Natl Acad Sci U S A* 104:13780–13785. <https://doi.org/10.1073/pnas.0706625104>
 8. Ott SJ, Musfeldt M, Wenderoth DF, Hampe J, Brant O, Fölsch UR, Timmis KN, Schreiber S. 2004. Reduction in diversity of the colonic mucosa associated bacterial microflora in patients with active inflammatory bowel disease. *Gut* 53:685–693. <https://doi.org/10.1136/gut.2003.025403>
 9. Turnbaugh PJ, Hamady M, Yatsunenko T, Cantarel BL, Duncan A, Ley RE, Sogin ML, Jones WJ, Roe BA, Affourtit JP, Egholm M, Henrissat B, Heath AC, Knight R, Gordon JI. 2009. A core gut microbiome in obese and lean twins. *Nature* 457:480–484. <https://doi.org/10.1038/nature07540>
 10. Karlsson FH, Tremaroli V, Nookaew I, Bergström G, Behre CJ, Fagerberg B, Nielsen J, Bäckhed F. 2013. Gut metagenome in European women with normal, impaired and diabetic glucose control. *Nature* 498:99–103. <https://doi.org/10.1038/nature12198>
 11. Brown CT, Davis-Richardson AG, Giongo A, Gano KA, Crabb DB, Mukherjee N, Casella G, Drew JC, Ilonen J, Knip M, Hyöty H, Veijola R, Simell T, Simell O, Neu J, Wasserfall CH, Schatz D, Atkinson MA, Triplet EW. 2011. Gut microbiome metagenomics analysis suggests a functional model for the development of autoimmunity for type 1 diabetes. *PLoS One* 6:e25792. <https://doi.org/10.1371/journal.pone.0025792>
 12. Vahtovuori J, Munukka E, Korkeamäki M, Luukkainen R, Toivanen P. 2008. Fecal microbiota in early rheumatoid arthritis. *J Rheumatol* 35:1500–1505.
 13. Wang Z, Klipfell E, Bennett BJ, Koeth R, Levison BS, Dugar B, Feldstein AE, Britt EB, Fu X, Chung Y-M, Wu Y, Schauer P, Smith JD, Allayee H, Tang WHW, DiDonato JA, Lusis AJ, Hazen SL. 2011. Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. *Nature* 472:57–63. <https://doi.org/10.1038/nature09922>
 14. Scanlan PD, Shanahan F, Clune Y, Collins JK, O’Sullivan GC, O’Riordan M, Holmes E, Wang Y, Marchesi JR. 2008. Culture-independent analysis of the gut microbiota in colorectal cancer and polyposis. *Environ Microbiol* 10:789–798. <https://doi.org/10.1111/j.1462-2920.2007.01503.x>
 15. Sobhani I, Tap J, Roudot-Thoraval F, Roperch JP, Letulle S, Langella P, Corthier G, Tran Van Nhieu J, Furet JP. 2011. Microbial dysbiosis in colorectal cancer (CRC) patients. *PLoS One* 6:e16393. <https://doi.org/10.1371/journal.pone.0016393>
 16. Bultman SJ. 2014. Emerging roles of the microbiome in cancer. *Carcinogenesis* 35:249–255. <https://doi.org/10.1093/carcin/bgt392>
 17. Gargano LM, Hughes JM. 2014. Microbial origins of chronic diseases. *Annu Rev Public Health* 35:65–82. <https://doi.org/10.1146/annurev-publhealth-032013-182426>
 18. Plottel CS, Blaser MJ. 2011. Microbiome and malignancy. *Cell Host Microbe* 10:324–335. <https://doi.org/10.1016/j.chom.2011.10.003>
 19. Rea D, Coppola G, Palma G, Barbieri A, Luciano A, Del Prete P, Rossetti S, Berretta M, Facchini G, Perdonà S, Turco MC, Arra C. 2018. Microbiota effects on cancer: from risks to therapies. *Oncotarget* 9:17915–17927. <https://doi.org/10.18632/oncotarget.24681>
 20. Tilg H, Schimiderer A, Djanani A. 2018. Gut microbiome-immune crosstalk affects progression of cancer. *Transl Gastroenterol Hepatol* 3:34. <https://doi.org/10.21037/tgh.2018.06.02>
 21. Maynard CL, Elson CO, Hatton RD, Weaver CT. 2012. Reciprocal interactions of the intestinal microbiota and immune system. *Nature* 489:231–241. <https://doi.org/10.1038/nature11551>
 22. Belkaid Y, Hand TW. 2014. Role of the microbiota in immunity and inflammation. *Cell* 157:121–141. <https://doi.org/10.1016/j.cell.2014.03.011>
 23. Jain T, Sharma P, Are AC, Vickers SM, Dudeja V. 2021. New insights into the cancer-microbiome-immune axis: decrypting a decade of discoveries. *Front Immunol* 12:622064. <https://doi.org/10.3389/fimmu.2021.622064>
 24. Flores R, Shi J, Fuhrman B, Xu X, Veenstra TD, Gail MH, Gajer P, Ravel J, Goedert JJ. 2012. Fecal microbial determinants of fecal and systemic estrogens and estrogen metabolites: a cross-sectional study. *J Transl Med* 10:253. <https://doi.org/10.1186/1479-5876-10-253>
 25. Fuhrman BJ, Feigelson HS, Flores R, Gail MH, Xu X, Ravel J, Goedert JJ. 2014. Associations of the fecal microbiome with urinary estrogens and estrogen metabolites in postmenopausal women. *J Clin Endocrinol Metab* 99:4632–4640. <https://doi.org/10.1210/jc.2014-2222>
 26. Goedert JJ, Jones G, Hua X, Xu X, Yu G, Flores R, Falk RT, Gail MH, Shi J, Ravel J, Feigelson HS. 2015. Investigation of the association between the fecal microbiota and breast cancer in postmenopausal women: a population-based case-control pilot study. *J Natl Cancer Inst* 107:djv147. <https://doi.org/10.1093/jnci/djv147>
 27. Byrd DA, Vogtmann E, Wu Z, Han Y, Wan Y, Clegg-Lamperty J-N, Yarney J, Wiafe-Addai B, Wiafe S, Awuah B, Ansong D, Nyarko K, Hullings AG, Hua X, Ahearn T, Goedert JJ, Shi J, Knight R, Figueroa JD, Brinton LA, Garcia-Closas M, Sinha R. 2021. Associations of fecal microbial profiles with breast cancer and nonmalignant breast disease in the Ghana breast health study. *Int J Cancer* 148:2712–2723. <https://doi.org/10.1002/ijc.33473>
 28. Ma J, Sun L, Liu Y, Ren H, Shen Y, Bi F, Zhang T, Wang X. 2020. Alter between gut bacteria and blood metabolites and the anti-tumor effects of *Faecalibacterium prausnitzii* in breast cancer. *BMC Microbiol* 20:82. <https://doi.org/10.1186/s12866-020-01739-1>
 29. Ma Z, Qu M, Wang X. 2022. Analysis of gut microbiota in patients with breast cancer and benign breast lesions. *Pol J Microbiol* 71:217–226. <https://doi.org/10.33073/pjm-2022-019>
 30. Bobin-Dubigeon C, Luu HT, Leuillet S, Lavergne SN, Carton T, Le Vacon F, Michel C, Nazih H, Bard J-M. 2021. Faecal microbiota composition varies between patients with breast cancer and healthy women: a comparative case-control study. *Nutrients* 13:2705. <https://doi.org/10.3390/nu13082705>
 31. Benini A, Bertazzoni Minelli E, Vesentini S, Marchiori L, Nardo G. 1992. Intestinal ecosystem and metabolic capacity in women with breast cancer. *Pharmacol Res* 25:184–185. [https://doi.org/10.1016/1043-6618\(92\)90352-C](https://doi.org/10.1016/1043-6618(92)90352-C)
 32. Velicer CM, Heckbert SR, Lampe JW, Potter JD, Robertson CA, Taplin SH. 2004. Antibiotic use in relation to the risk of breast cancer. *JAMA* 291:827–835. <https://doi.org/10.1001/jama.291.7.827>
 33. Fitzpatrick SG, Katz J. 2010. The association between periodontal disease and cancer: a review of the literature. *J Dent* 38:83–95. <https://doi.org/10.1016/j.jdent.2009.10.007>
 34. Hujoel PP, Drangsholt M, Spiekerman C, Weiss NS. 2003. An exploration of the periodontitis-cancer association. *Ann Epidemiol* 13:312–316. [https://doi.org/10.1016/s1047-2797\(02\)00425-8](https://doi.org/10.1016/s1047-2797(02)00425-8)
 35. Söder B, Yakob M, Meurman JH, Andersson LC, Klinge B, Söder P-Ö. 2011. Periodontal disease may associate with breast cancer. *Breast Cancer Res Treat* 127:497–502. <https://doi.org/10.1007/s10549-010-1221-4>
 36. Söder B, Yakob M, Meurman JH, Andersson LC, Söder P-Ö. 2012. The association of dental plaque with cancer mortality in Sweden. A longitudinal study. *BMJ Open* 2:e001083. <https://doi.org/10.1136/bmjopen-2012-001083>
 37. Wang H, Altamir J, Niazi F, Green H, Calhoun BC, Sturgis C, Grobmyer SR, Eng C. 2017. Breast tissue, oral and urinary microbiomes in breast cancer. *Oncotarget* 8:88122–88138. <https://doi.org/10.18632/oncotarget.21490>
 38. Nearing JT, DeClercq V, Langille MGI. 2023. Investigating the oral microbiome in retrospective and prospective cases of prostate, colon, and breast cancer. *NPJ Biofilms Microbiomes* 9:23. <https://doi.org/10.1038/s41522-023-00391-7>
 39. Aßhauer KP, Wemheuer B, Daniel R, Meinicke P. 2015. Tax4Fun: predicting functional profiles from metagenomic 16S rRNA data. *Bioinformatics* 31:2882–2884. <https://doi.org/10.1093/bioinformatics/btv287>
 40. Iwai S, Weinmaier T, Schmidt BL, Albertson DG, Poloso NJ, Dabbagh K, DeSantis TZ. 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>
 41. Langille MGI, Zaneveld J, Caporaso JG, McDonald D, Knights D, Reyes JA, Clemente JC, Burkpile DE, Vega Thurber RL, Knight R, Beiko RG, Huttenhower C. 2013. Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nat Biotechnol* 31:814–821. <https://doi.org/10.1038/nbt.2676>

42. Falk RT, Brinton LA, Dorgan JF, Fuhrman BJ, Veenstra TD, Xu X, Gierach GL. 2013. Relationship of serum estrogens and estrogen metabolites to postmenopausal breast cancer risk: a nested case-control study. *Breast Cancer Res* 15:R34. <https://doi.org/10.1186/bcr3416>
43. Dallal CM, Tice JA, Buist DSM, Bauer DC, Lacey JV Jr, Cauley JA, Hue TF, Lacroix A, Falk RT, Pfeiffer RM, Fuhrman BJ, Veenstra TD, Xu X, Brinton LA, B-FIT Research Group. 2014. Estrogen metabolism and breast cancer risk among postmenopausal women: a case-cohort study within B-FIT. *Carcinogenesis* 35:346–355. <https://doi.org/10.1093/carcin/bgt367>
44. Shrode RL, Knobbe JE, Cady N, Yadav M, Hoang J, Cherwin C, Curry M, Garje R, Vikas P, Sugg S, et al. 2023. Breast cancer patients from the Midwest region of the United States have reduced levels of short-chain fatty acid-producing gut bacteria. *Sci Rep* 13:526. <https://doi.org/10.1038/s41598-023-27436-3>
45. Arumugam M, Raes J, Pelletier E, Le Paslier D, Yamada T, Mende DR, Fernandes GR, Tap J, Bruls T, Batto J-M, et al. 2011. Enterotypes of the human gut microbiome. *Nature* 473:174–180. <https://doi.org/10.1038/nature09944>
46. Holmes I, Harris K, Quince C. 2012. Dirichlet multinomial mixtures: generative models for microbial metagenomics. *PLoS One* 7:e30126. <https://doi.org/10.1371/journal.pone.0030126>
47. Wu GD, Chen J, Hoffmann C, Bittinger K, Chen Y-Y, Keilbaugh SA, Bewtra M, Knights D, Walters WA, Knight R, Sinha R, Gilroy E, Gupta K, Baldassano R, Nessel L, Li H, Bushman FD, Lewis JD. 2011. Linking long-term dietary patterns with gut microbial enterotypes. *Science* 334:105–108. <https://doi.org/10.1126/science.1208344>
48. Costea PI, Hildebrand F, Arumugam M, Bäckhed F, Blaser MJ, Bushman FD, de Vos WM, Ehrlich SD, Fraser CM, Hattori M, et al. 2018. Enterotypes in the landscape of gut microbial community composition. *Nat Microbiol* 3:388. <https://doi.org/10.1038/s41564-018-0114-x>
49. Frioux C, Ansoerge R, Özkurt E, Ghassemi Nedjad C, Fritscher J, Quince C, Waszak SM, Hildebrand F. 2023. Enterosignatures define common bacterial guilds in the human gut microbiome. *Cell Host Microbe* 31:1111–1125. <https://doi.org/10.1016/j.chom.2023.05.024>
50. Zhu J, Liao M, Yao Z, Liang W, Li Q, Liu J, Yang H, Ji Y, Wei W, Tan A, Liang S, Chen Y, Lin H, Zhu X, Huang S, Tian J, Tang R, Wang Q, Mo Z. 2018. Breast cancer in postmenopausal women is associated with an altered gut metagenome. *Microbiome* 6:136. <https://doi.org/10.1186/s40168-018-0515-3>
51. Seeley JJ, Ghosh S. 2017. Molecular mechanisms of innate memory and tolerance to LPS. *J Leukoc Biol* 101:107–119. <https://doi.org/10.1189/jlb.3MR0316-118RR>
52. Zeller G, Tap J, Voigt AY, Sunagawa S, Kultima JR, Costea PI, Amiot A, Böhm J, Brunetti F, Habermann N, et al. 2014. Potential of fecal microbiota for early-stage detection of colorectal cancer. *Mol Syst Biol* 10:766. <https://doi.org/10.15252/msb.20145645>
53. Tang X-Y, Zhu Y-Q, Wei B, Wang H. 2010. Expression and functional research of TLR4 in human colon carcinoma. *Am J Med Sci* 339:319–326. <https://doi.org/10.1097/MAJ.0b013e3181cef1b7>
54. Cario E, Rosenberg IM, Brandwein SL, Beck PL, Reinecker HC, Podolsky DK. 2000. Lipopolysaccharide activates distinct signaling pathways in intestinal epithelial cell lines expressing Toll-like receptors. *J Immunol* 164:966–972. <https://doi.org/10.4049/jimmunol.164.2.966>
55. Duan R-D, Nilsson A. 2009. Metabolism of sphingolipids in the gut and its relation to inflammation and cancer development. *Prog Lipid Res* 48:62–72. <https://doi.org/10.1016/j.plipres.2008.04.003>
56. Jones GS, Spencer Feigelson H, Falk RT, Hua X, Ravel J, Yu G, Flores R, Gail MH, Shi J, Xu X, et al. 2019. Mammographic breast density and its association with urinary estrogens and the fecal microbiota in postmenopausal women. *PLoS One* 14:e0216114. <https://doi.org/10.1371/journal.pone.0216114>
57. Tennant SM, Hartland EL, Phumoonna T, Lyras D, Rood JJ, Robins-Browne RM, van Driel IR. 2008. Influence of gastric acid on susceptibility to infection with ingested bacterial pathogens. *Infect Immun* 76:639–645. <https://doi.org/10.1128/IAI.01138-07>
58. Segata N, Haake SK, Mannon P, Lemon KP, Waldron L, Gevers D, Huttenhower C, Izard J. 2012. Composition of the adult digestive tract bacterial microbiome based on seven mouth surfaces, tonsils, throat and stool samples. *Genome Biol* 13:R42. <https://doi.org/10.1186/gb-2012-13-6-r42>
59. Ridlon JM, Kang DJ, Hylemon PB, Bajaj JS. 2014. Bile acids and the gut microbiome. *Curr Opin Gastroenterol* 30:332–338. <https://doi.org/10.1097/MOG.0000000000000057>
60. Schmidt TS, Hayward MR, Coelho LP, Li SS, Costea PI, Voigt AY, Wirbel J, Maistrenko OM, Alves RJ, Bergsten E, de Beaufort C, Sobhani I, Heintz-Buschart A, Sunagawa S, Zeller G, Wilmes P, Bork P. 2019. Extensive transmission of microbes along the gastrointestinal tract. *Elife* 8:e42693. <https://doi.org/10.7554/eLife.42693>
61. Russo E, Bacci G, Chiellini C, Fagorzi C, Niccolai E, Taddei A, Ricci F, Ringressi MN, Borrelli R, Melli F, Miloeva M, Bechi P, Mengoni A, Fani R, Amedei A. 2017. Preliminary comparison of oral and intestinal human microbiota in patients with colorectal cancer: a pilot study. *Front Microbiol* 8:2699. <https://doi.org/10.3389/fmicb.2017.02699>
62. Uchino Y, Goto Y, Konishi Y, Tanabe K, Toda H, Wada M, Kita Y, Beppu M, Mori S, Hijioka H, Otsuka T, Natsugoe S, Hara E, Sugiura T. 2021. Colorectal cancer patients have four specific bacterial species in oral and gut microbiota in common—a metagenomic comparison with healthy subjects. *Cancers (Basel)* 13:3332. <https://doi.org/10.3390/cancers13133332>
63. Liao C, Rolling T, Djukovic A, Fei T, Mishra V, Liu H, Lindberg C, Dai L, Zhai B, Peled JU, van den Brink MRM, Hohl TM, Xavier JB. 2023. A high fraction of oral bacteria in the feces indicates gut microbiota depletion with implications for human health. *bioRxiv*. <https://doi.org/10.1101/2022.10.24.513595>
64. Hummel W, Kula MR. 1989. Simple method for small-scale disruption of bacteria and yeasts. *J Microbiol Methods* 9:201–209. [https://doi.org/10.1016/0167-7012\(89\)90037-7](https://doi.org/10.1016/0167-7012(89)90037-7)
65. Cady NC, Stelick S, Batt CA. 2003. Nucleic acid purification using microfabricated silicon structures. *Biosens Bioelectron* 19:59–66. [https://doi.org/10.1016/s0956-5663\(03\)00123-4](https://doi.org/10.1016/s0956-5663(03)00123-4)
66. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Lozupone CA, Turnbaugh PJ, Fierer N, Knight R. 2011. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc Natl Acad Sci U S A* 108:4516–4522. <https://doi.org/10.1073/pnas.1000080107>
67. Minalla AR, Dubrow R, Bousse LJ. 2001. Feasibility of high-resolution oligonucleotide separation on a microchip, p 90–97. In *Proc. SPIE 4560, Microfluidics and BioMEMS*. <https://doi.org/10.1117/12.443044>
68. 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>
69. Liu C, Cui Y, Li X, Yao M. 2021. microeco: an R package for data mining in microbial community ecology. *FEMS Microbiol Ecol* 97:faa255. <https://doi.org/10.1093/femsec/faa255>
70. Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, Huttenhower C. 2011. Metagenomic biomarker discovery and explanation. *Genome Biol* 12:R60. <https://doi.org/10.1186/gb-2011-12-6-r60>
71. Gaujoux R, Seoighe C. 2010. A flexible R package for nonnegative matrix factorization. *BMC Bioinformatics* 11:367. <https://doi.org/10.1186/1471-2105-11-367>
72. Christopoulos DT. 2016. On the efficient identification of an inflection point. *Int J Math Sci Comput* 6:13–20.
73. Guven E. 2023. Decision of the optimal rank of a nonnegative matrix factorization model for gene expression data sets utilizing the unit invariant knee method: development and evaluation of the elbow method for rank selection. *JMIR Bioinform Biotechnol* 4:e43665. <https://doi.org/10.2196/43665>
74. Mirzayi C, Renson A, Zohra F, Elsafoury S, Geistlinger L, Kasselmann LJ, Eckenrode K, van de Wijgert J, Loughman A, Marques FZ, et al. 2021. Reporting guidelines for human microbiome research: the STORMS checklist. *Nat Med* 27:1885–1892. <https://doi.org/10.1038/s41591-021-01552-x>