

Article

Comparison of Three DNA Isolation Methods and Two Sequencing Techniques for the Study of the Human Microbiota

Julio Plaza-Díaz ^{1,2} , Mariana F. Fernández ^{1,3,4} , Federico García ^{1,5,6} , Natalia Chueca ^{1,5,6}, Luis Fontana ^{1,7,8,*}  and Ana I. Álvarez-Mercado ^{1,8,9,*} 

- ¹ Institute of Biosanitary Research (ibs.GRANADA), San Cecilio University Clinical Hospital, 18012 Granada, Spain; jrplaza@ugr.es (J.P.-D.); marieta@ugr.es (M.F.F.); fegarcia@ugr.es (F.G.); nchueca@ugr.es (N.C.)
- ² School of Health Sciences, International University of La Rioja, 26001 Logroño, Spain
- ³ Spanish Consortium for Research on Epidemiology and Public Health (CIBERESP), 28029 Madrid, Spain
- ⁴ Department of Radiology and Physical Medicine, School of Medicine, University of Granada, 18016 Granada, Spain
- ⁵ Microbiology Unit, San Cecilio University Clinical Hospital, 18016 Granada, Spain
- ⁶ Spanish Consortium for Research on Infectious Diseases (CIBERINFEC), 28029 Madrid, Spain
- ⁷ Department of Biochemistry and Molecular Biology II, School of Pharmacy, University of Granada, 18071 Granada, Spain
- ⁸ Institute of Nutrition and Food Technology “José Matáix”, Centre of Biomedical Research, University of Granada, 18016 Granada, Spain
- ⁹ Department Pharmacology, School of Pharmacy, 18071 Granada, Spain
- * Correspondence: fontana@ugr.es (L.F.); alvarezmercado@ugr.es (A.I.Á.-M.); Tel.: +95-82-41000 (ext. 41599) (L.F. & A.I.Á.-M.)

Abstract: Breast cancer is the most commonly diagnosed cancer in women and the second leading cause of female death. Altered interactions between the host and the gut microbiota appear to play an influential role in carcinogenesis. Several studies have shown different signatures of the gut microbiota in patients with breast cancer compared to healthy women. Currently, there is disagreement regarding the different DNA isolation and sequencing methodologies for studies on the human microbiota, given that they can influence the interpretation of the results obtained. The goal of this work was to compare (1) three different DNA extraction strategies to minimize the impact of human DNA, and (2) two sequencing strategies (16S rRNA and shotgun) to identify discrepancies in microbiome results. We made use of breast tissue and fecal samples from both healthy women and breast cancer patients who participated in the MICROMA study (reference NCT03885648). DNA was isolated by means of mechanical lysis, trypsin, or saponin. The amount of eukaryotic DNA isolated using the trypsin and saponin methods was lower compared to the mechanical lysis method (mechanical lysis, $89.11 \pm 2.32\%$; trypsin method, $82.63 \pm 1.23\%$; saponin method, $80.53 \pm 4.09\%$). In samples with a predominance of prokaryotic cells, such as feces, 16S rRNA sequencing was the most advantageous approach. For other tissues, which are expected to have a more complex microbial composition, the need for an in-depth evaluation of the multifactorial interaction between the various components of the microbiota makes shotgun sequencing the most appropriate method. As for the three extraction methods evaluated, when sequencing samples other than stool, the trypsin method is the most convenient. For fecal samples, where contamination by host DNA is low, no prior treatment is necessary.

Keywords: breast cancer; breast microbiota; gut microbiota; extraction methods; microbial identification; shotgun; 16S rRNA gene amplification



Academic Editor: Uri Gophna

Received: 27 February 2025

Revised: 21 March 2025

Accepted: 2 April 2025

Published: 4 April 2025

Citation: Plaza-Díaz, J.; Fernández, M.F.; García, F.; Chueca, N.; Fontana, L.; Álvarez-Mercado, A.I. Comparison of Three DNA Isolation Methods and Two Sequencing Techniques for the Study of the Human Microbiota. *Life* **2025**, *15*, 599. <https://doi.org/10.3390/life15040599>

Copyright: © 2025 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

1.1. Role of the Microbiome in Breast Cancer and the Importance of Microbiome Research

To maintain the homeostasis of the organism, humans need a stable and balanced microbiome or collection of microbes including bacteria, fungi, archaea, and viruses [1]. Alterations in the normal microbiota, referred to as dysbiosis, are related to a wide array of diseases including breast cancer [2]. As it occurs in the digestive tract, a specific microbiome is supposed to be beneficial to maintaining a healthy breast microenvironment. Evidence supports this hypothesis [3] since antibiotic treatment has resulted in breast tumor growth in mice [4,5] and humans [6,7]. Moreover, differences in gut microbiome profiles between benign and malignant tumors suggest that this profile is associated with the disease [2]. In a study reported by Ma et al. [8], 59 members of the microbiota differed in abundance between patients with breast cancer and those with benign breast disease. Patients with breast cancer exhibited lower relative abundances of *Firmicutes* (*Faecalibacterium*) and *Bacteroidetes* and increased abundances of *Verrucomicrobia*, *Proteobacteria*, *Actinobacteria*, *Bacillus*, *Enterobacter*, and *Staphylococcus*. In light of the fact that breast lesions can produce precursors of breast cancer or serve as risk markers for the disease, identifying microbial profiles associated with these lesions may contribute to early detection and prevent the development of cancer.

Interestingly, some studies suggest intestinal bacteria can reach the breast tissue through different pathways and promote tumorigenesis [9]. These observations are crucial given the fact that breast cancer is the most prevalent and the second cause of death for women [10]. However, the connection between the human microbiome and breast cancer has not yet been explored in detail [11].

1.2. Sequencing Strategies and Their Significance in Microbiome Studies

Metagenomic sequencing is an appropriate strategy for the study of the microbial community. The increasing use of metagenome next-generation sequencing (NGS) technologies presents both a challenge and an opportunity for expanding knowledge about the microbes that reside in the human body and their applicability to clinical practice [12]. Technological advances in NGS are resulting in a rapid increase in research aimed at analyzing complex microbiome environments. Such advances provide longer read lengths and paired-sequencing reads to identify microbial communities better [13]. To date, sequencing-based microbiome methods have greatly enriched our understanding of the microbe population of the human body in both health and disease and have contributed to the expansion of our view of microbes in polymicrobial communities and infections [14]. Currently, there are several NGS platforms: The Roche 454 Genome Sequencer FLX System, the Illumina HiSeq and MiSeq Systems, the Ion Torrent, the Applied Biosystems SOLiD sequencer, and Oxford Nanopore Technologies. However, some issues still need to be solved: (1) Existing sequencing technologies display platform-specific biases depending on the run mode and chemistry. These biases affect the read length, data throughput, guanine cytosine (GC) coverage bias, error rates, and the ability to resolve repetitive genomic elements [15]. (2) Due to the large diversity of generated data, the possibility to produce realistic benchmark datasets for particular experimental setups is challenging and essential. (3) Analyses are often hampered by overwhelming quantities of human DNA, yielding only a small proportion of microbial reads for analysis [14]. (4) Finally, the cost of metagenomic analysis per sample with high human-to-microbial DNA ratios could be high when considering the depths required for the thorough investigation of the microbiome in a large number of samples [16].

Nowadays, two extensively used metagenome sequencing strategies are shotgun and the small subunit (SSU) rRNA genes, namely the 16S rRNA. Both are used to catalog the human microbiome in health and disease and to study microbial communities of medical,

pharmaceutical, or biotechnological relevance [17]. 16S rRNA gene sequencing identifies and quantifies species or amplicon sequence variants (ASVs) [18]. This methodology has greatly expanded our understanding of microorganism diversity leading to the creation of some public databases like the National Center for Biotechnology info (NCBI) GenBank, SILVA [19], the Ribosomal Information Project (RDP) [20], and Greengenes [21]. Because the 16S rRNA gene is universal in bacteria, this method presents some benefits: (1) the sequencing of a short (~1500 base pairs) region is relatively economical and fast, and (2) the common use of the same target gene across multiple studies increases accuracy in meta-analysis.

By contrast, although the V3 and V4 hypervariable regions of the bacterial 16S rRNA can be used to distinguish taxa, this method provides uneven resolution of the taxonomic spectrum. In addition, variation in the copy number of this gene influences abundance estimates. Another disadvantage of this method is that the accuracy of PCR approaches is reduced because of primer bias and chimeras. Moreover, PCR-based databases and primer analysis tools overestimate primer coverage [22].

Shotgun is a method used for analyzing random DNA fragments ranging from 100 to 1000 base pairs. DNA is fragmented into numerous small segments and sequenced to obtain reads. Multiple overlapping reads for the target DNA are obtained by performing several rounds of this fragmentation and sequencing separately. Finally, the sequences are assembled to give an overall sequence [23]. Shotgun sequencing was one of the precursor technologies that led to whole genome sequencing. Metagenomic shotgun sequencing appears to be a powerful tool for identifying causative disease agents in human patients. This technique also allows the detection of different types of microbes other than bacteria. In addition, the limit of detection of this method is modulated by sequencing depth, read length, and data accuracy. This allows microbes to be detected to obtain more complete functional representations. However, one of its main limitations is its cost and the fact that the datasets obtained from microbial communities are very diverse, not only due to the natural variation of biological systems but also due to differences in laboratory protocols, the number of replications, and sequencing technologies [24].

1.3. Challenges in Microbiome Research and Rationale

Some challenges still need to be overcome: (1) Shotgun metagenomic datasets from microbial communities are highly variable, not only due to the natural variation of the underlying biological systems but also due to differences in laboratory protocols, replicate numbers, and sequencing technologies; (2) overwhelming amounts of human DNA lead to potential confounding and a small proportion of microbial reads for analysis.

Several studies have shown that DNA extraction methods play a crucial role in bacterial community detection [25–29]. There are several steps involved in DNA extraction, including weighing, homogenization, lysis of bacteria, and purification of DNA, each of which still requires improvement and guidelines. The standard weighing procedure, for example, may be time consuming and tedious for collecting the same volume of fecal material for all samples. There is also a possibility that homogenization of the sample could impact the amount of detectable bacteria [30,31].

This work is part of an ongoing clinical case–control study (Identifier NCT03885648) in which the mammary and fecal microbiota from healthy women and breast cancer patients are being compared. Although it is nowadays known that breast tissue is not sterile, we expected the amount of microbial DNA to be low. In addition, the amount of breast tissue available was limited. Therefore, we decided that the extraction method should be optimized by comparing several protocols. Some current solutions to eliminate host contamination include the use of microbeads, but these methods require a larger sample

size and are more expensive. On the other hand, although the comparison of extraction methods has been reported in recent years, the available comparison between protocols other than stool are very limited.

In the present study, we aimed to compare (1) three different DNA extraction strategies to minimize the impact of human DNA, and (2) two sequencing strategies (16S rRNA and shotgun) to identify discrepancies in microbiome results. For this purpose, we took advantage of fecal and mammary tissue samples obtained in an ongoing case–control clinical study from both healthy and breast cancer-affected women (MICROMA study) [2,32].

2. Materials and Methods

2.1. Design

This work is part of an ongoing case–control study registered at [ClinicalTrials.gov](https://clinicaltrials.gov) (accessed on 30 September 2017) (Identifier NCT03885648) and described elsewhere [2,32]. Briefly, cases were women with stages I and II breast cancer. Controls were women who had breast augmentation or reduction surgery.

2.2. Biological Samples

Breast tissue and stool samples from cases and controls were collected as detailed in [2]. Samples were obtained from the Unit of Mammary Pathology, General Surgery Service at San Cecilio University Hospital (Granada, Spain). Breast tissue samples were obtained at the time of surgery from a marginal zone and distant from the tumor (3–5 cm), in the case of women with breast cancer, and any available breast tissue for women with breast reduction or enlargement. Stool samples were collected from healthy women and women with breast cancer in a time window of 1 to 3 days before the breast tissue collection [2].

2.3. DNA Isolation Methods

2.3.1. Sample Lysis

For breast tissue, frozen samples were first ground in a mortar. Twenty-five milligrams of sample (feces and breast tissue) were homogenized in PBS by vortex shaking at maximum speed. Later on, additional mechanical lysis with pathogen lysis tubes L (recommended for fungi and some resistant bacteria) and S (recommended for bacteria) (Qiagen, Barcelona, Spain) was carried out. Three different methodologies (described below) were tested for DNA isolation.

Method A. This method involves mechanical lysis with pathogen tubes L and S as described. No additional treatment was performed.

Method B. Selective removal of human DNA from metagenomic DNA samples was carried out following the methodology described by Hunter et al. (2011) [33]. Briefly, after the mechanical lysis described above, samples were centrifuged at $13,000 \times g$ for 2 min. The supernatant was discarded and the pellet containing all cells was suspended in 1 mL TrypZean (Sigma T3449, St. Louis, MO, USA) containing 0.05% Tween-20, and subsequently incubated at 37 °C for 70 min with gentle shaking in a water bath. Samples were again briefly vortexed and centrifuged at $5000 \times g$ for 2 min. The supernatant was transferred into a new tube and centrifuged at $13,000 \times g$ for 10 min. The pellet was then suspended in 0.4 mL PBS. Then, the described additional mechanical lysis was performed using pathogen lysis tubes L and S.

Method C. This is a modified saponin-based differential lysis method. Samples were centrifuged at $8000 \times g$ for 5 min, then the supernatant was carefully removed and the pellet suspended in 250 μ L of PBS. Saponin was applied to a final concentration of 2.5%, mixed well, and incubated at room temperature for 10 min to promote host cell lysis. Following this incubation, 350 μ L of water was added and incubation was continued

at room temperature for 30 s. After this, 12 μ L of 5 M NaCl was added to deliver an osmotic shock, lysing the damaged host cells. Samples were next centrifuged at $6000\times g$ for 5 min; the supernatant was removed and the pellet re-suspended in 100 μ L of PBS. Further mechanical lysis was performed with pathogen lysis tubes L and S. Figure 1 summarizes the aforementioned methods.

2.3.2. DNA Extraction

The QIAamp[®] cador[®] Pathogen Mini kit (Qiagen, Barcelona, Spain) was used to isolate genomic DNA according to the manufacturer's instructions. The concentration and purity of DNA were determined using a NanoDrop ND-1000 (Thermo Fisher Scientific, Waltham, MA, USA).

2.4. DNA Sequencing

Two different metagenome sequencing strategies were carried out in this study: the 16S rRNA method and shotgun sequencing.

16S rRNA method. The V3 and V4 hypervariable regions of the bacterial 16S rRNA gene were PCR amplified using the following primers: 16S Amplicon PCR Forward Primer (5'-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCTACGGGNGGCWGAG-3') and 16S Amplicon PCR Reverse Primer (5'-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGACTACHVGGGTATCTAATCC-3') [34]. A 25 μ L reaction volume was used to perform all PCR reactions containing 12.5 μ L 2X KAPA HiFi Hotstart ready mix (KAPA Biosystems, Woburn, MA, USA), 5 μ L of each forward and reverse primers (1 μ M), and 2.5 μ L of extracted DNA (10 ng) under the following cycling conditions: initial denaturation at 95 °C for 3 min, followed by cycles of denaturation at 95 °C for 30 s, annealing at 55 °C for 30 s, and elongation at 72 °C for 30 s, followed by a final extension at 72 °C for 5 min. The 16S V3 and V4 amplicons were purified with AMPure XP beads (Beckman Coulter, Indianapolis, IN, USA) from free primers and primer dimer species. Next, the index PCR was conducted, in which dual indices and Illumina sequencing adapters were attached with the Nextera XT Index Kit (Illumina, San Diego, CA, USA), PCR conditions were: 95 °C for 3 min; 8 cycles of 95 °C for 30 s, 55 °C, 72 °C for 30 s; 72 °C for 5 min, and hold at 4 °C. Before quantification, the pooled PCR products were purified with AMPure XP beads (Beckman Coulter, Indianapolis, IN, USA). Using a paired-end (2 \times 300 nt) Illumina MiSeq sequencing system (Illumina, San Diego, CA, USA), the amplicons were sequenced at MiSeq (Illumina, San Diego, CA, USA) [2,32].

Shotgun method. Metagenomic libraries were constructed with Nextera XT DNA Library Preparation Kit (Illumina, San Diego, CA, USA). For library construction, we used 100 pg input DNA and the amplicon tagment mix (ATM) in Nextera XT was diluted 1:10 in nuclease-free water. The tagmentation reaction mixture consisted of TD buffer (10 μ L), input DNA (5 μ L), and diluted ATM (5 μ L). Following the manufacturer's protocol, PCR cycles were conducted for 100 pg DNA for 17 cycles. The amplified libraries were purified with AMPure XP (Beckman Coulter, Indianapolis, IN, USA). The quality of the purified libraries was assessed using an Agilent High Sensitivity DNA Kit on an Agilent 2100 Bioanalyzer (Santa Clara, CA, USA). The sequencing libraries were further quantified using the KAPA Library Quantification Kit. The metagenomics libraries were mixed with PhiX Control v3 (Illumina, San Diego, CA, USA) at a ratio of 9:1 and sequenced with an Illumina MiSeq Reagent Kit v3 (600 cycles) (Illumina, San Diego, CA, USA) [2,32].

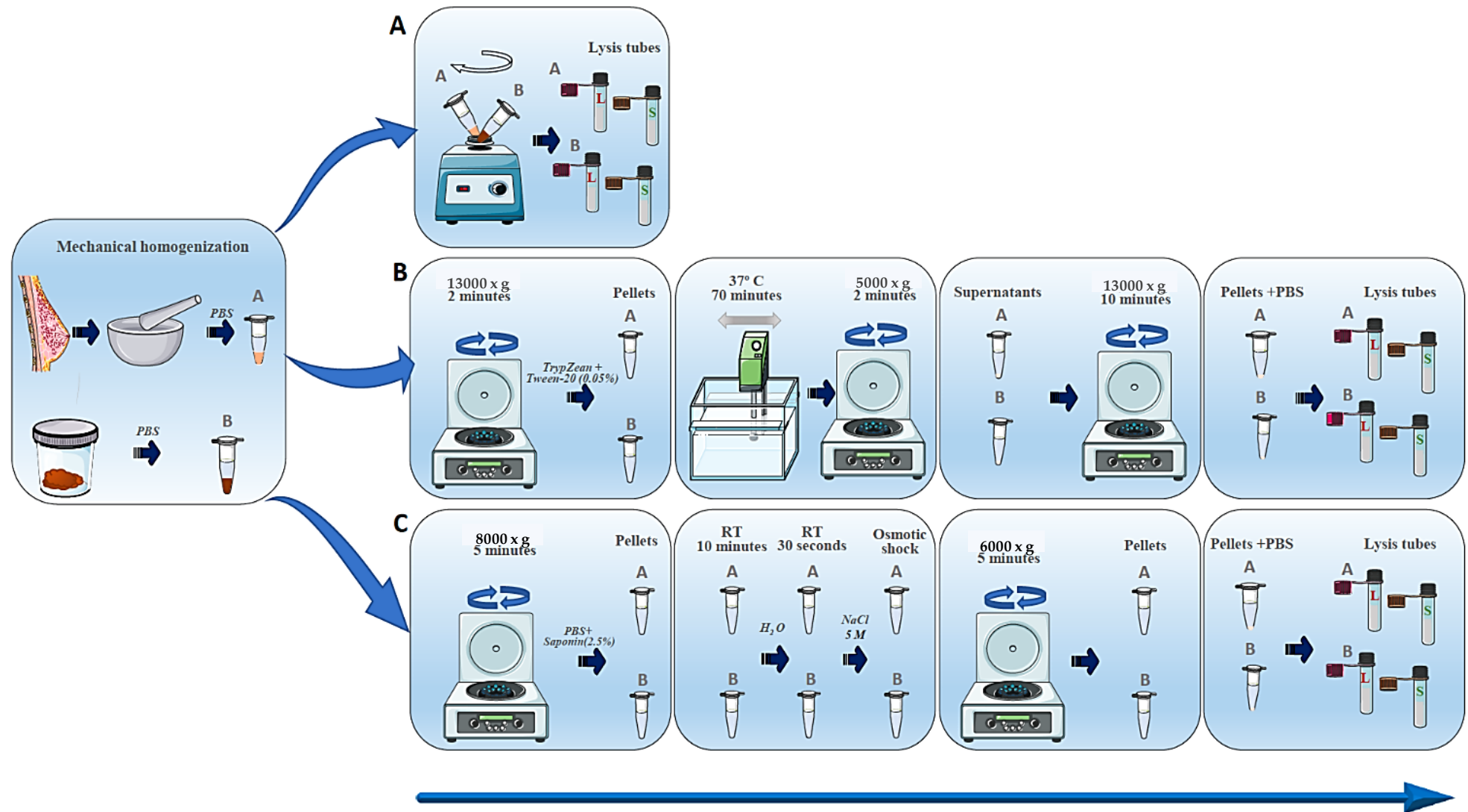


Figure 1. Schematic representation of different extraction methods used in this work. (A) Mechanical lysis; (B) trypsin method; (C) saponin method. PBS: phosphate-buffered saline; RT: room temperature.

2.5. Bioinformatic Analyses

2.5.1. 16S rRNA Method

Raw sequences were demultiplexed with Illumina's bcl2fastq2 Conversion Software v2.20 and raw data were imported into open-source software QIIME 2 2020.8 [35] with the q2-tools-import script with PairedEndFastqManifestPhred33 input format. Using the V3-V4 16S rRNA sequencing method, reads were generated. To obtain an abundance distribution of sequence variances that have a difference of one nucleotide, we applied DADA2 [18], which provides a quality-aware model of Illumina amplicon errors. Using the q2-dada2-denoise script, forward reads were truncated at position 288 and trimmed at position 6 after retrieving quality scores. The reverse reads were truncated at position 220 and trimmed at position 7. As part of the q2-dada2-denoise algorithm, chimeras were removed with the "consensus" filter, in which chimeras were detected in samples individually and sequences found chimeric in a sufficient fraction of samples were eliminated. As part of this process, forward and reverse reads were also merged. Using MAFFT [36] via q2-alignment, all amplicon sequence variants (ASVs) were aligned and a phylogeny was constructed with FASTTREE2 (via q2-phylogeny) [37]. Based on the SILVA 16S V3-V4 v132_99 [17], ASVs were assigned taxonomy using the Naive Bayes taxonomy classifier (via q2-feature-classifier) [38]. Samples with fewer than 10,000 reads were pruned during the data filtering process.

2.5.2. Shotgun Method

Raw sequences were analyzed with MetaPhlAn 3.0 (Metagenomic Phylogenetic Analysis), a computational tool for profiling the composition of microbial communities from metagenomic shotgun sequencing data [39].

2.6. Statistic Analysis

The median and range were used to express the results. $p < 0.05$ was considered statistically significant. Variables that were not normally distributed were log-transformed, and outliers were removed (without achieving a loss of value from samples of up to 15%). To ensure a clear understanding of the data, the values were presented untransformed. To assess differences in the relative abundance of bacteria (phylum and genus), the U Mann-Whitney test was applied. Unless otherwise specified, data are presented as mean \pm SEM. A Kruskal-Wallis test determined statistical significance. Significance was adjusted by the Bonferroni correction for multiple tests. For the alpha indices, two-tailed Student's *t*-tests were performed to compare each of the indices determined with the samples sequenced by 16S rRNA or shotgun.

Rivera-Pinto analysis identifies microbial signatures, i.e., groups of microbes that can predict particular phenotypes. Diagnosing, prognosticating, or predicting therapeutic responses is possible using the microbial signatures associated with an individual's microbiome. It is possible to identify microbial signatures by modeling the response variable and selecting taxa that yield the highest classification or prediction accuracy level. With the Rivera-Pinto method and the selbal algorithm, we evaluated specific signatures at the phylum and genus levels to select a sparse model that adequately explains the response variable [40].

3. Results

This study was carried out on stool and mammary tissue samples from patients with breast cancer (cases) and healthy women (controls) who participated in the MICROMA study [2,24].

3.1. Human DNA Depletion

To assess the efficiency of depleting human DNA from samples, we performed metagenomic sequencing of all DNA extracts, then calculated the proportion of eukaryotes to total sequence reads. Our results indicate a significant decrease in the amount of DNA from eukaryotic cells in the trypsin and saponin methods compared to the mechanical lysis method (Figure 2). We were able to detect archaea in our shotgun metagenomic sequencing.

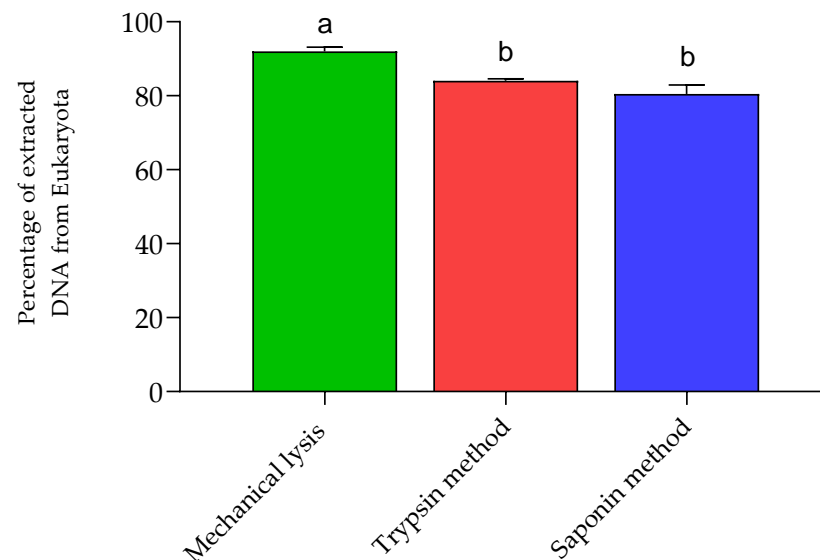


Figure 2. Percentage of DNA isolated from eukaryotic cells in breast tissue. Labeled medians with identical letters are not significant. Different letters mean significant differences ($p < 0.05$). Mechanical lysis vs. trypsin method, $p = 0.0031$; mechanical lysis vs. saponin method $p = 0.0006$; trypsin method vs. saponin method $p = 0.2571$.

3.2. 16S rRNA Sequencing

The abundance of the phylum *Actinomycetota* was greater in the fecal samples of healthy women treated with saponin compared to those treated with trypsin and mechanically. In comparison with the mechanical and saponin DNA isolation methods, the trypsin method resulted in a significant increase in the relative abundance of the phylum *Bacteroidota*. The phylum *Bacillota* was significantly more abundant when DNA was isolated by mechanical lysis or with saponin in comparison with the trypsin method. The *Pseudomonadota* and *Verrucomicrobiota* phyla were less abundant when the DNA was isolated using saponin compared to the mechanical lysis method. Furthermore, the DNA isolation method that resulted in the lowest yield in the phylum *Verrucomicrobiota* was trypsin (Table 1A).

The abundance of phyla *Bacteroidota*, *Pseudomonadota*, and *Bacillota* in breast tissue samples from healthy women differed significantly. *Bacteroidota* and *Pseudomonadota* were significantly higher with the mechanical lysis and trypsin methods compared with the saponin method, whereas *Bacillota* was significantly higher with the saponin method than with the mechanical lysis (Table 1A).

The methods of DNA isolation using mechanical lysis and trypsin produced similar results, while the results obtained with the saponin method were very different to the previous two. Due to this and the fact that the latter method includes more intermediate steps and is therefore longer, it was decided to continue with the first two methods (Table 1B).

Table 1. (A) Identification of bacterial DNA based on 16S rRNA gene using three extraction methods in healthy women. (B) Identification of bacterial DNA based on 16S rRNA gene using three extraction methods in patients with breast cancer.

(A)						
Healthy Women (Controls)						
Phylum	Feces			Breast		
	A (n = 5)	B (n = 5)	C (n = 5)	A (n = 5)	B (n = 5)	C (n = 5)
<i>Actinomycetota</i>	4.7 (0.3–13.8) ^a	7.8 (2.7–10.7) ^a	14.4 (13.5–17.5) ^b	5.7 (4.8–6.2)	7.3 (5.6–9.9)	6.6 (5.0–8.0)
<i>Bacteroidota</i>	6.3 (3.0–22.8) ^a	27.2 (15.2–33.9) ^b	6.6 (2.6–22.7) ^a	16.6 (9.4–19.0) ^a	15.9 (6.2–19.2) ^a	8.9 (7.1–11.3) ^b
<i>Bacillota</i>	74.6 (69.7–87.0) ^a	62.5 (48.7–72.9) ^b	74.1 (63.1–79.6) ^{ab}	57.5 (46.2–63.7) ^a	64.6 (52.3–67.6) ^{ab}	71.7 (60.0–78.6) ^b
<i>Fusobacteria</i>	0 (0–0.03)	0 (0–0.2)	0 (0–0)	0 (0–0.04)	0 (0–0.1)	0 (0–0.01)
<i>Pseudomonadota</i>	2.0 (0.2–2.4) ^a	0.6 (0.3–1.9) ^{ab}	0.3 (0.1–1.3) ^b	4.7 (3.6–10.0) ^a	4.8 (3.7–12.6) ^a	2.6 (0.5–2.8) ^b
<i>Verrucomicrobiota</i>	2.3 (0–6.8) ^a	0 (0–26.5) ^b	0.02 (0–8.0) ^b	2.3 (0.8–3.7)	3.0 (0.2–4.1)	1.5 (1.1–3.1)
Unassigned	0 (0–0)	0 (0–0)	0 (0–0)	0.04 (0–0.3)	0 (0–0.5)	0.09 (0–0.2)
(B)						
Patients with Breast Cancer (Cases)						
Phylum	Feces		Breast			
	A (n = 5)	B (n = 5)	A (n = 5)	B (n = 5)		
<i>Actinomycetota</i>	1.6 (1.1–29.0)	5.0 (3.5–7.2)	4.9 (1.9–6.0)	3.1 (1.5–5.1)		
<i>Bacteroidota</i>	9.6 (0.8–32.0) ^a	17.3 (3.7–24.6) ^b	18.3 (16.1–18.9) ^a	19.4 (18.4–23.5) ^b		
<i>Bacillota</i>	68.2 (52.8–95.0)	72.4 (58.0–87.7)	63.0 (57.7–70.2)	66.9 (60.0–69.2)		
<i>Fusobacteria</i>	0 (0–0.01)	0 (0–0)	0 (0–0.1)	0.08 (0–0.3)		
<i>Pseudomonadota</i>	2.0 (0.9–7.7)	1.2 (0.3–7.6)	4.6 (3.1–8.2)	4.2 (2.5–9.1)		
<i>Verrucomicrobiota</i>	0.9 (0–6.3) ^a	5.2 (0.06–14.2) ^b	2.3 (1.5–3.0)	3.0 (1.9–3.5)		
Unassigned	0 (0–0)	0 (0–0)	0.04 (0.02–0.2)	0.01 (0–0.02)		

(A) A. Mechanical lysis. B. Trypsin method. C. Saponin method. Data are expressed as median and range. Labeled medians with identical letters are not significant. Different letters mean significant differences ($p < 0.05$) and were calculated using a Kruskal–Wallis test corrected by Bonferroni post hoc test. (B) A. Mechanical lysis. B. Trypsin method. Data are expressed as median and range. Labeled medians with identical letters are not significant. Different letters mean significant differences ($p < 0.05$) and were calculated using a Kruskal–Wallis test corrected by Bonferroni post hoc test.

We found significant differences in the *Bacteroidota* and *Verrucomicrobiota* phyla in the feces of patients with breast cancer. In these two phyla, the trypsin method resulted in significantly higher relative abundances compared with the mechanical lysis method (Table 1B). In breast tissue samples, *Bacteroidota* was the only phylum that changed, with a greater abundance with the trypsin method compared with the mechanical lysis (Table 1B).

3.3. Shotgun

Pseudomonadota, *Actinomycetota*, and *Bacteroidota* fecal abundances were significantly higher in trypsin-treated samples than in mechanically analyzed samples from control patients. In the shotgun analysis of breast tissue samples from breast cancer patients and control women, no identifications were found for the mechanical lysis method (Table 2). On the contrary, identifications were made for both control and breast cancer women using the trypsin method (Table 2). Significant statistical changes were observed in the feces of breast cancer patients: the abundance of the *Actinomycetota* and *Pseudomonadota* phyla was significantly greater with trypsin-treated samples than with mechanically analyzed samples (Table 2).

3.4. Diversity Index Comparison

When comparing different alpha diversity indices between samples sequenced by 16S rRNA or shotgun and extracted with mechanical lysis or trypsin, no significant differences were observed in the stool samples in four of the five indices determined (Figure 3A).

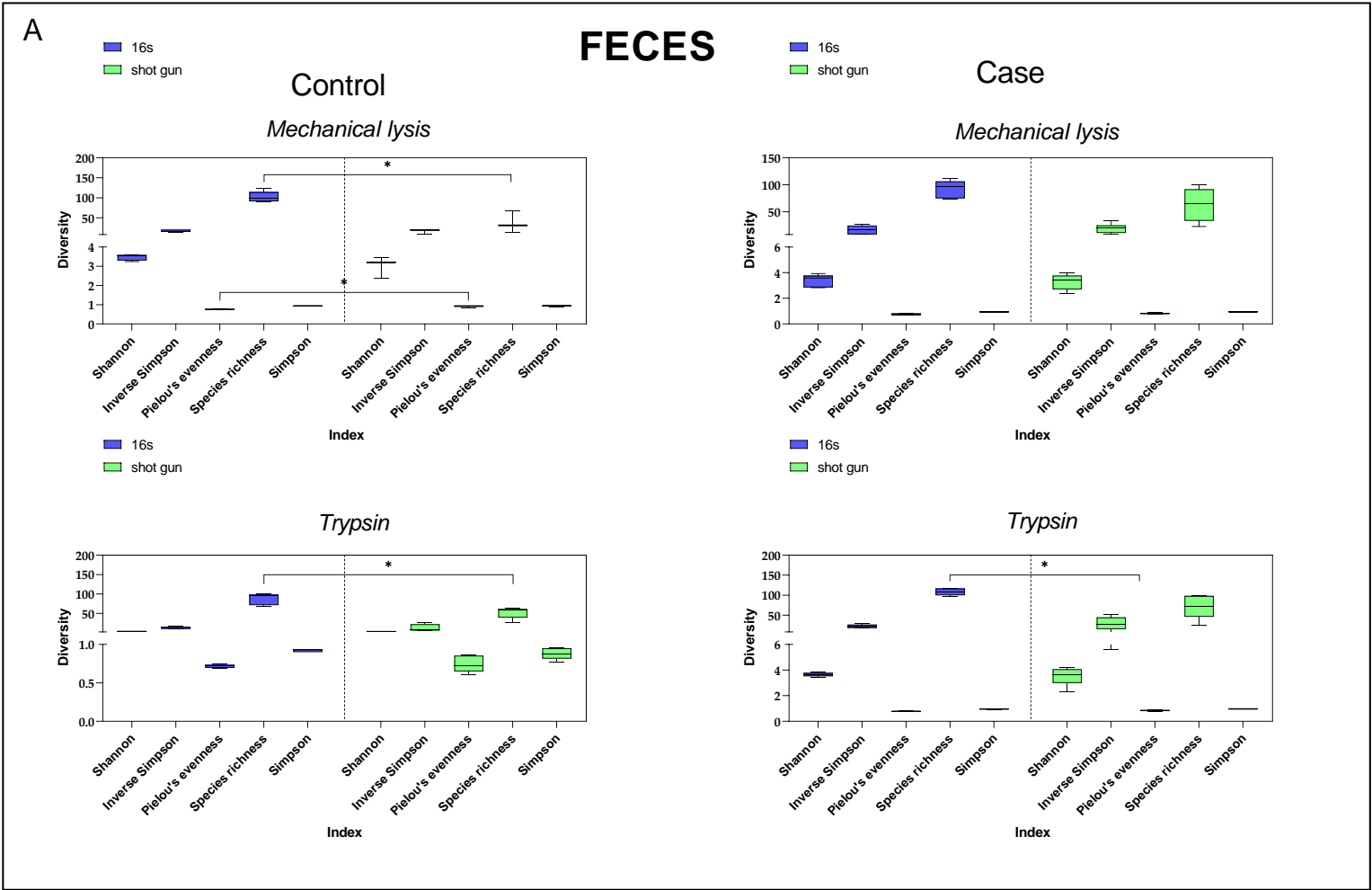


Figure 3. Cont.

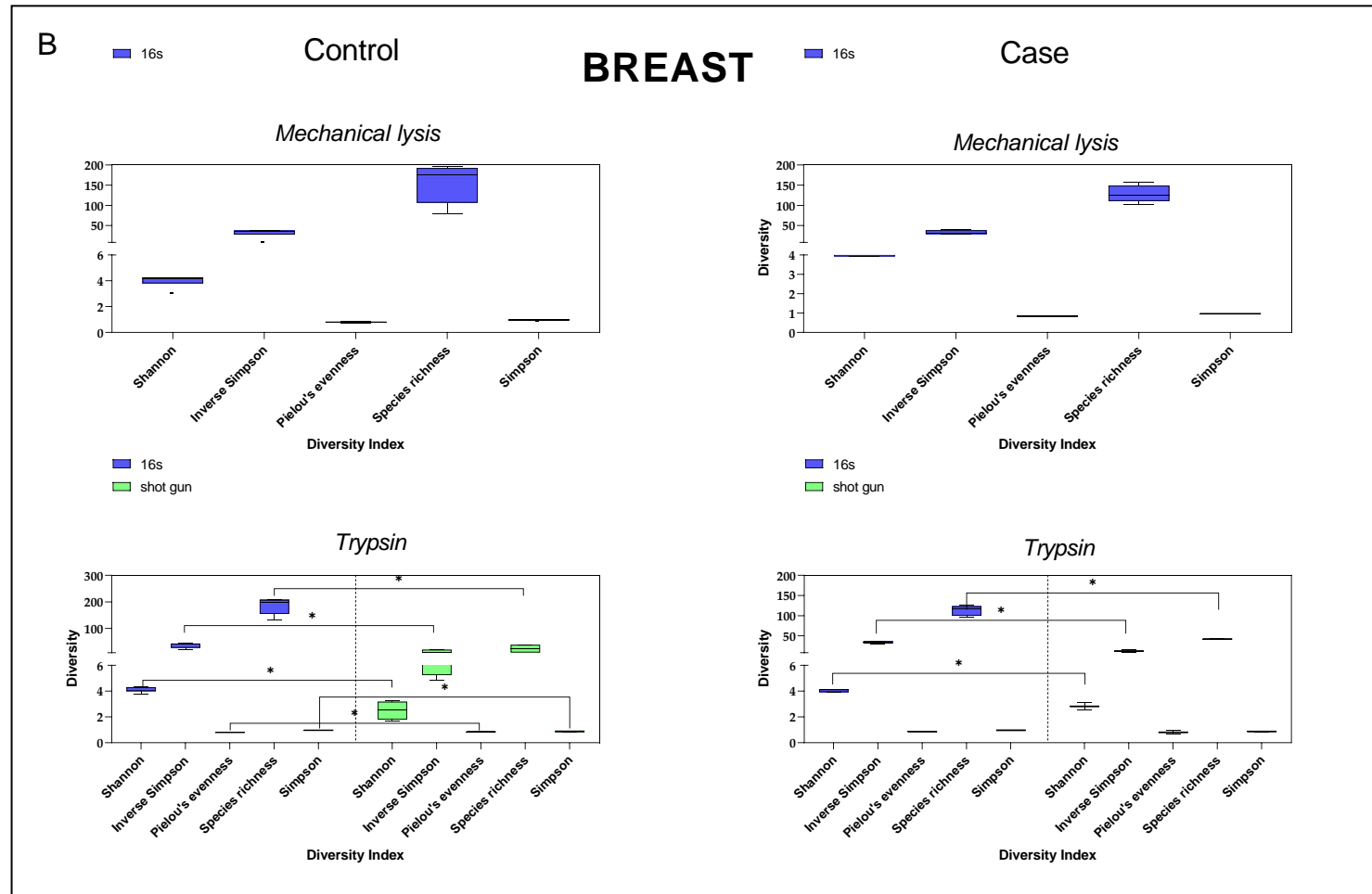


Figure 3. Alpha diversity indices for stool samples (A) and breast tissue samples (B). Figure shows various alpha diversity indices for samples whose DNA was isolated using mechanical lysis or trypsin methods, and sequenced by 16S rRNA or shotgun methods. * $p < 0.05$. Cases ($n = 5$), controls ($n = 5$).

Table 2. Identification of bacterial DNA based on shotgun using two extraction methods.

Variables	Control				Patients with Breast Cancer			
	Feces		Breast		Feces		Breast	
	A (n = 5)	B (n = 5)	A (n = 5)	B (n = 5)	A (n = 5)	B (n = 5)	A (n = 5)	B (n = 5)
<i>Actinomycetota</i>	0 (0–21.8)	1.9 (0–16.0) *	0 (0–0)	1.2 (1.0–1.0)	0.3 (0–26.3)	6.7 (2.0–14.0) *	0 (0–0)	0.7 (1.0–1.0)
<i>Bacteroidota</i>	0 (0–5.5)	4.2 (0–21.0) *	0 (0–0)	10.4 (10.0–10.0)	5.5 (0–39.2)	7.7 (1.0–28.0)	0 (0–0)	13.4 (12–15)
<i>Bacillota</i>	88.3 (72.6–100)	83.9 (63–100)	0 (0–0)	94.2 (88.0–100.0)	77.5 (46.8–100)	83.2 (56–100)	0 (0–0)	83.9 (80–88)
<i>Pseudomonadota</i>	0 (0–0)	1.7 (0–3.0) *	0 (0–0)	0 (0–0)	0 (0–6.7)	2.7 (1.0–5.0) *	0 (0–0)	4.7 (5.0–5.0)

Data are expressed as median and range. Labeled medians with * are significantly different ($p < 0.05$) and were calculated using a Kruskal–Wallis test corrected by Bonferroni post hoc test. A. Mechanical lysis. B. Trypsin method.

On the contrary, in the case of breast tissue samples, we were unable to determine any of these indices when the DNA was extracted by mechanical lysis and sequenced by shotgun. Significant differences in the diversity indices were observed in both cases and controls for both 16S rRNA and shotgun sequencing methods when DNA isolation was carried out with trypsin (Figure 3B).

3.5. Microbiome Balance

An overview of microbial balances for stool and mammary tissue samples is provided in Figures 4 and 5. These figures depict the distribution of balance scores obtained with both sequencing techniques. As a result of the 16S rRNA analysis and using mechanical lysis in feces samples, *Collinsella* had lower balance scores than *Clostridium sensu stricto* 1 (Figure 4A).

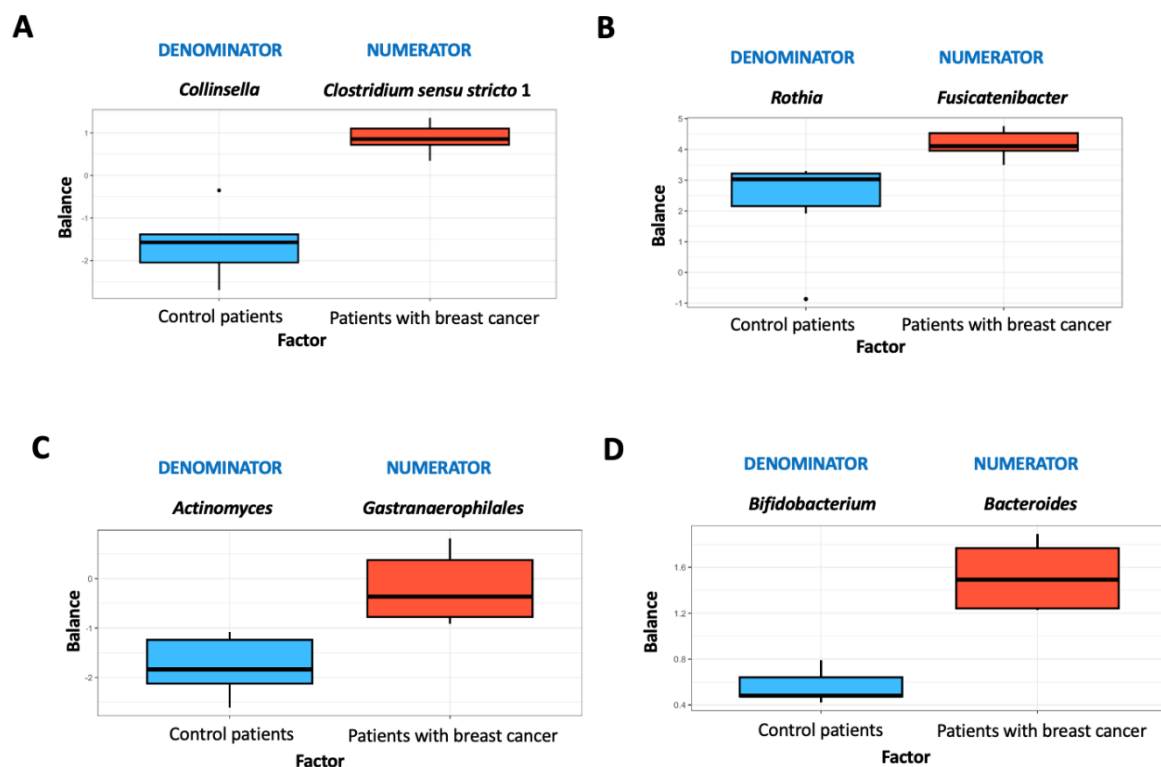


Figure 4. Microbiome balance for 16S rRNA method. Distribution of balance scores obtained with 16S rRNA analysis using mechanical lysis in feces samples (A) and breast samples (B), and with 16S rRNA analysis using trypsin method in feces samples (C) and breast samples (D). Patients with breast cancer (n = 5), controls (n = 5).

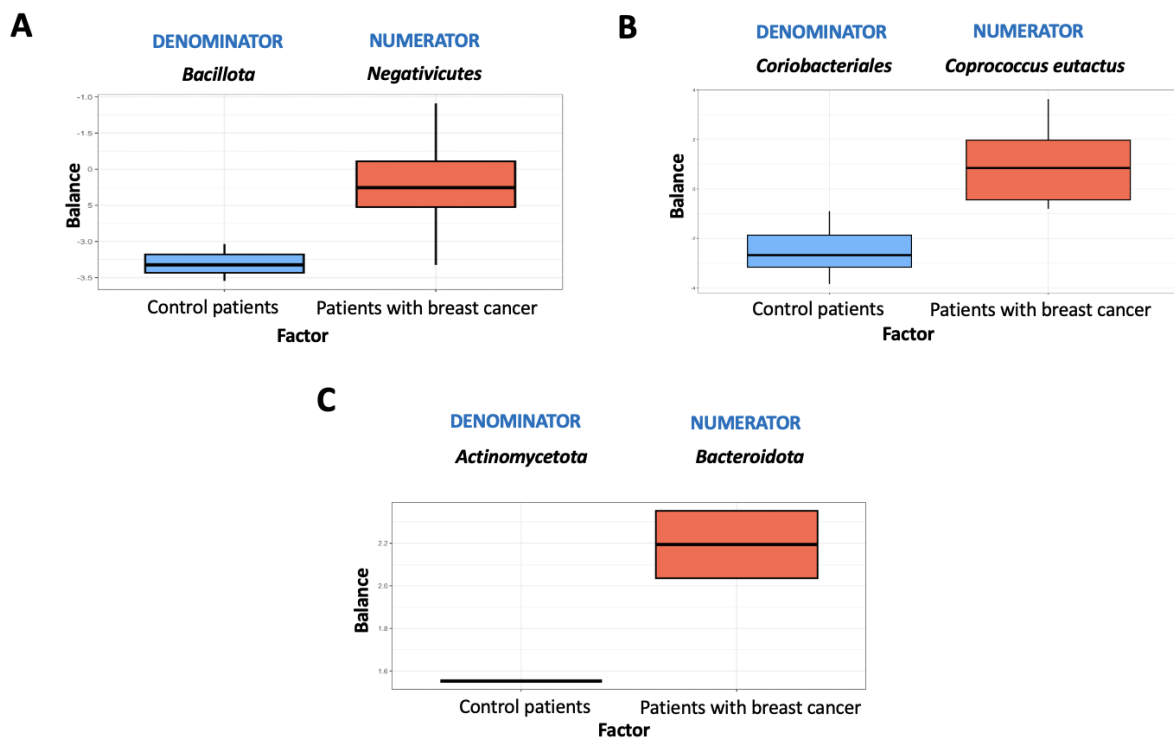


Figure 5. Microbiome balance for shotgun method. Distribution of balance scores obtained with shotgun analysis using mechanical lysis in feces samples (A) and with shotgun analysis using trypsin method in fecal samples (B) and breast samples (C). Patients with breast cancer (n = 5), controls (n = 5).

On the contrary, *Clostridium sensu stricto* 1 was more prevalent in patients with breast cancer than in control women (Figure 4A). Using the mechanical lysis method and 16S rRNA sequencing, the breast samples were characterized by lower *Rothia* scores than *Fusicatenuibacter* scores (Figure 4B). *Rothia* was the most commonly associated genus in healthy women (Figure 4B). In the case of the trypsin method in feces samples, *Gastranaerophilales* had higher scores than *Actinomyces*, the latter being more associated with control women (Figure 4C). The levels of *Bacteroides* were higher in breast samples treated with trypsin compared to *Bifidobacterium*. The presence of *Bacteroides* was associated with breast cancer patients (Figure 4D).

For mechanical lysis in feces samples, *Bacillota* and *Negativicutes* were the main taxa detected (Figure 5A). Using the trypsin method in feces samples, the order *Coriobacteriales* had lower scores than *Coprococcus eutactus*, the latter being more associated with patients with breast cancer (Figure 5B). Breast samples treated with trypsin contained higher levels of *Bacteroidota* than *Actinomycetota*. Patients with breast cancer were associated with the presence of the phylum *Bacteroidota* (Figure 5C).

4. Discussion

Technological advancements in NGS are rapidly expanding research into analyzing complex microbiome environments and revolutionizing microbiology [41]. Bioinformatics will inevitably be applied in this transition from a purely laboratory task to a computational one. Conversely, the possibility of conducting metagenomics studies quickly and at a reduced cost may facilitate faster disease diagnosis [42]. In order to establish a relationship between health and disease and develop an effective treatment for infectious diseases, it is essential to have a comprehensive understanding of epidemiology, virulence mechanisms,

and host–organism interactions [43], with the human microbiota playing a key role in this scenario.

The present work makes use of breast tissue and fecal samples from both healthy women and breast cancer patients who participated in the MICROMA study (Identifier NCT03885648) [2,32]. We compared the effectiveness of various DNA extraction and genome sequencing strategies to identify possible discrepancies in the microbiome profile, as well as evaluated the precision and accuracy of microbial identification to improve diagnostic and treatment tools. Since human tissue analysis is often hampered by overwhelming quantities of human DNA [14], we also aimed to minimize the impact of human DNA in the samples. With this purpose, we compared the outcomes from three distinct extraction methods—namely, mechanical lysis, mechanical lysis accompanied by trypsin digestion, and a modified saponin-based differential lysis method. Saponin was selected since detergents, including saponin, are frequently used in cell lysis before DNA extraction [44,45]. Our results showed that the saponin method yielded the most divergent results. This finding, in conjunction with the observation that this method involves a greater number of intermediate steps and consequently exposes the process to a heightened risk of experimental inaccuracy, led to the decision to exclude it from subsequent experiments and continue the study without this method.

Concerning the trypsin method, it is known that this enzyme catalyzes the hydrolysis of peptide bonds, breaking down proteins into smaller peptides [46]. The use of trypsin to clear DNA is not new and has been previously reported, for instance, in forensic DNA analysis of bone samples before DNA isolation [47] and for preparing DNA from formalin-fixed paraffin-embedded samples for methylation analysis [48].

We are aware of the fact that alternative extraction methodologies may be employed for the optimization of DNA isolation by, for example, bead beating [49]. However, the implementation of this process produces a considerable increase in the expense of the sequencing process, whilst concomitantly necessitating a greater volume of samples, a requirement that is not universally met, particularly in the context of samples of human origin.

Although each person harbors a distinctive gut microbiota composition, the overall structure conforms to patterns that are repeated in different individuals (enterotypes). The enterotype concept suggests that the microbial ecosystem in the human gut forms internal states of symbiosis between the different members of the microbial community, probably determined by the very metabolic or social networks in which they are integrated. These interactions explain the stability and resilience of an ecosystem subject to fluctuations [50] and make necessary consider all the components of the microbiota. Having this in mind, in stool samples we did not find a significant influence of trypsin in the overall picture of microbial diversity, where more than 90% of the microbial DNA is from bacteria. In contrast, our results for breast tissue samples were different since human DNA is predominant. By eliminating human DNA, one can better discern the microbial diversity although this could result in the loss of valuable sequencing material.

Overall, in the present study, we confirmed that the extraction method used to identify microbial DNA sequences significantly influences the results. We also found that in human samples such as breast tissue, where microbial DNA is not predominant, the trypsin method was useful because it decreased the amount of human DNA and therefore facilitated sequencing.

On the other hand, the results of Rivera-Pinto's analysis demonstrated different phyla- and genera-specific signatures. This finding is of great interest, since predicting particular phenotypes (based on the microbial signatures associated with an individual's microbiome) is crucial in the diagnosis, prognosis, and therapeutic responses. Indeed, our results corroborate the fact that strategies to obtain and analyze metagenomic data still need a

wide effort of standardization and validation to get comparable results and guarantee that the extensive amount of data being produced is reliable.

Given that the 16S rRNA method is specific to bacteria, it presents a reduced risk of host contamination [41]. Due to extensive reference databases and computational error correction tools, 16S rRNA sequencing also reduces the risk of false positives; however, the likelihood of false positives increases as sample biomass decreases [51]. In contrast, shotgun metagenomic sequencing represents an evolutionary precursor to whole genome sequencing. The technology is a highly effective means of identifying the causative agents of human diseases. However, a significant limitation of this method is the fact that the presence of a vast quantity of human DNA may result in the potential for confounding and a comparatively low proportion of microbial reads for analysis. This may explain the lack of data obtained for shotgun sequencing in breast tissue samples. With the 16S rRNA sequencing, however, the bacterial population is amplified and, consequently, human DNA is eliminated and microbial identification is facilitated. To the best of our knowledge, the majority of studies that use mammary tissue utilize 16S rRNA sequencing [52], and therefore there is no clear picture of the mammary microbiome.

In summary, this study confirms that DNA extraction and sequencing methods are critical in microbiota studies, as they can influence the interpretation of results. There are many aspects that remain to be resolved, so some recommendations that can be made for future research in this field would be as follows.

With regard to DNA extraction, it is necessary to standardize protocols to guarantee reproducibility. Variability in extraction methods can affect microbial composition, as our work suggests. Possible biases would also have to be evaluated (for example, comparing different commercial kits and homemade methods to determine which ones best preserve microbial diversity, avoiding the over- or under-representation of certain taxa) and protocols adapted to different types of samples to maximize the recovery of quality DNA. The elimination of contaminants is also a pending issue: it is necessary to implement negative controls and the use of DNA-free reagents to avoid contamination.

As for sequencing, the most common bioinformatic tools used to study microbial ecology include QIIME pipelines [53], Mothur software [54], and R packages, such as vegan [55]; phyloseq [56], DADA2 [18], and RAM [57]. Alpha/beta diversity metrics are calculated using the vegan R package [55]. It is of course important to make a comparison of the various existing platforms, but we think that these studies would benefit from multi-omic approaches to obtain a more complete picture of microbial function. Finally, although we are aware that it is a work in progress, there is a need to continue refining reference databases, bioinformatics tools and algorithms for genome reconstruction to minimize errors in species assignment.

The resolution of these pending issues is crucial for clinical research related to the human microbiota as they influence the accuracy of disease diagnosis, the personalization of and response to treatments, and the development of new therapies, among others.

5. Conclusions

In conclusion, in samples with a predominance of prokaryotic cells, such as feces, 16S rRNA sequencing is the most advantageous approach. For other tissues, which are expected to have a more complex microbial composition, the need for an in-depth evaluation of the multifactorial interaction between the various components of the microbiota makes shotgun sequencing the most appropriate sequencing method. As for the three extraction methods evaluated, when sequencing samples other than stool, the trypsin method is the most convenient. For fecal samples, where contamination by host DNA is low, no prior treatment is necessary.

Author Contributions: Conceptualization, M.F.F. and L.F.; methodology, A.I.Á.-M., J.P.-D., F.G. and N.C.; formal analysis, A.I.Á.-M. and J.P.-D.; investigation, A.I.Á.-M.; resources, M.F.F. and L.F.; data curation, A.I.Á.-M. and J.P.-D.; writing—original draft preparation, A.I.Á.-M., J.P.-D. and L.F.; writing—review and editing, all authors; funding acquisition, M.F.F. and L.F. All authors have read and agreed to the published version of the manuscript.

Funding: The authors acknowledge the financial support received from the European Union's Horizon 2020 Research and Innovation Program HBM4EU under Grant Agreement #733032 and Innovation Program Partnership for the Assessment of Risks from Chemicals (PARC) for the financial support under Grant Agreement #101057014. The study was further supported by the Instituto de Salud Carlos III, Madrid, Spain, through grant FIS_PI23_01884 (to M.F.F.), and by Junta de Andalucía, Spain, through grant PI-0538-2017 (to L.F.).

Institutional Review Board Statement: The study was conducted under the Declaration of Helsinki, and approved by the Ethics Committee of Biomedical Research of Andalucía (reference 9/2017. Date of approval: 2 February 2018).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Data are available on request from the authors. The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

- Scarpellini, E.; Ianiro, G.; Attili, F.; Bassanelli, C.; De Santis, A.; Gasbarrini, A. The Human Gut Microbiota and Virome: Potential Therapeutic Implications. *Dig. Liver Dis.* **2015**, *47*, 1007–1012. [\[CrossRef\]](#) [\[PubMed\]](#)
- Plaza-Díaz, J.; Álvarez-Mercado, A.I.; Ruiz-Marín, C.M.; Reina-Pérez, I.; Pérez-Alonso, A.J.; Sánchez-Andujar, M.B.; Torné, P.; Gallart-Aragón, T.; Sánchez-Barrón, M.T.; Reyes Lartategui, S.; et al. Association of Breast and Gut Microbiota Dysbiosis and the Risk of Breast Cancer: A Case-Control Clinical Study. *BMC Cancer* **2019**, *19*, 495. [\[CrossRef\]](#) [\[PubMed\]](#)
- Kwa, M.; Plottel, C.S.; Blaser, M.J.; Adams, S. The Intestinal Microbiome and Estrogen Receptor-Positive Female Breast Cancer. *JNCI J. Natl. Cancer Inst.* **2016**, *108*, djw029. [\[CrossRef\]](#)
- Rossini, A.; Rumio, C.; Sfondrini, L.; Tagliabue, E.; Morelli, D.; Miceli, R.; Mariani, L.; Palazzo, M.; Ménard, S.; Balsari, A. Influence of Antibiotic Treatment on Breast Carcinoma Development in Proto- Neu Transgenic Mice. *Cancer Res.* **2006**, *66*, 6219–6224. [\[CrossRef\]](#)
- Velicer, C.M.; Lampe, J.W.; Heckbert, S.R.; Potter, J.D.; Taplin, S.H. Hypothesis: Is Antibiotic Use Associated with Breast Cancer. *Cancer Causes Control* **2003**, *14*, 739–747. [\[CrossRef\]](#)
- Reid, G. Can Breast Microbiota Provide Protective Effects Against Cancer? *Future Microbiol.* **2016**, *11*, 987–989. [\[CrossRef\]](#)
- Roderburg, C.; Loosen, S.H.; Joerdens, M.S.; Demir, M.; Luedde, T.; Kostev, K. Antibiotic Therapy Is Associated with an Increased Incidence of Cancer. *J. Cancer Res. Clin. Oncol.* **2023**, *149*, 1285–1293. [\[CrossRef\]](#) [\[PubMed\]](#)
- Ma, J.; Sun, L.; Liu, Y.; Ren, H.; Shen, Y.; Bi, F.; Zhang, T.; Wang, X. Alter between Gut Bacteria and Blood Metabolites and the Anti-Tumor Effects of Faecalibacterium Prausnitzii in Breast Cancer. *BMC Microbiol.* **2020**, *20*, 82. [\[CrossRef\]](#)
- Zhang, J.; Xia, Y.; Sun, J. Breast and Gut Microbiome in Health and Cancer. *Genes Dis.* **2021**, *8*, 581–589. [\[CrossRef\]](#)
- Arnold, M.; Morgan, E.; Rumgay, H.; Mafra, A.; Singh, D.; Laversanne, M.; Vignat, J.; Gralow, J.R.; Cardoso, F.; Siesling, S.; et al. Current and Future Burden of Breast Cancer: Global Statistics for 2020 and 2040. *Breast* **2022**, *66*, 15–23. [\[CrossRef\]](#)
- Chadha, J.; Nandi, D.; Atri, Y.; Nag, A. Significance of Human Microbiome in Breast Cancer: Tale of an Invisible and an Invincible. *Semin. Cancer Biol.* **2021**, *70*, 112–127. [\[CrossRef\]](#) [\[PubMed\]](#)
- Ames, S.K.; Gardner, S.N.; Marti, J.M.; Slezak, T.R.; Gokhale, M.B.; Allen, J.E. Using Populations of Human and Microbial Genomes for Organism Detection in Metagenomes. *Genome Res.* **2015**, *25*, 1056–1067. [\[CrossRef\]](#)
- Zhu, J.; Liao, M.; Yao, Z.; Liang, W.; Li, Q.; Liu, J.; Yang, H.; Ji, Y.; Wei, W.; Tan, A.; et al. Breast Cancer in Postmenopausal Women Is Associated with an Altered Gut Metagenome. *Microbiome* **2018**, *6*, 136. [\[CrossRef\]](#)
- Langille, M.G.I.; Zaneveld, J.; Caporaso, J.G.; McDonald, D.; Knights, D.; Reyes, J.A.; Clemente, J.C.; Burkpile, D.E.; Vega Thurber, R.L.; Knight, R.; et al. Predictive Functional Profiling of Microbial Communities Using 16S rRNA Marker Gene Sequences. *Nat. Biotechnol.* **2013**, *31*, 814–821. [\[CrossRef\]](#)
- Sevim, V.; Lee, J.; Egan, R.; Clum, A.; Hundley, H.; Lee, J.; Everroad, R.C.; Detweiler, A.M.; Bebout, B.M.; Pett-Ridge, J.; et al. Shotgun Metagenome Data of a Defined Mock Community Using Oxford Nanopore, PacBio and Illumina Technologies. *Sci. Data* **2019**, *6*, 285. [\[CrossRef\]](#)

16. Nelson, M.T.; Pope, C.E.; Marsh, R.L.; Wolter, D.J.; Weiss, E.J.; Hager, K.R.; Vo, A.T.; Brittnacher, M.J.; Radey, M.C.; Hayden, H.S.; et al. Human and Extracellular DNA Depletion for Metagenomic Analysis of Complex Clinical Infection Samples Yields Optimized Viable Microbiome Profiles. *Cell Rep.* **2019**, *26*, 2227–2240.e5. [[CrossRef](#)] [[PubMed](#)]
17. Fritz, A.; Hofmann, P.; Majda, S.; Dahms, E.; Dröge, J.; Fiedler, J.; Lesker, T.R.; Belmann, P.; DeMaere, M.Z.; Darling, A.E.; et al. CAMISIM: Simulating Metagenomes and Microbial Communities. *Microbiome* **2019**, *7*, 17. [[CrossRef](#)]
18. Callahan, B.J.; McMurdie, P.J.; Rosen, M.J.; Han, A.W.; Johnson, A.J.A.; Holmes, S.P. DADA2: High-Resolution Sample Inference from Illumina Amplicon Data. *Nat. Methods* **2016**, *13*, 581–583. [[CrossRef](#)]
19. Quast, C.; Pruesse, E.; Yilmaz, P.; Gerken, J.; Schweer, T.; Yarza, P.; Peplies, J.; Glöckner, F.O. The SILVA Ribosomal RNA Gene Database Project: Improved Data Processing and Web-Based Tools. *Nucleic Acids Res.* **2012**, *41*, D590–D596. [[CrossRef](#)]
20. Maidak, B.L. The RDP (Ribosomal Database Project) Continues. *Nucleic Acids Res.* **2000**, *28*, 173–174. [[CrossRef](#)]
21. DeSantis, T.Z.; Hugenholtz, P.; Larsen, N.; Rojas, M.; Brodie, E.L.; Keller, K.; Huber, T.; Dalevi, D.; Hu, P.; Andersen, G.L. Greengenes, a Chimera-Checked 16S rRNA Gene Database and Workbench Compatible with ARB. *Appl. Environ. Microbiol.* **2006**, *72*, 5069–5072. [[CrossRef](#)] [[PubMed](#)]
22. Zou, B.; Li, J.; Zhou, Q.; Quan, Z.-X. MIPE: A Metagenome-Based Community Structure Explorer and SSU Primer Evaluation Tool. *PLoS ONE* **2017**, *12*, e0174609. [[CrossRef](#)] [[PubMed](#)]
23. Anderson, S. Shotgun DNA Sequencing Using Cloned DNase I-Generated Fragments. *Nucleic Acids Res.* **1981**, *9*, 3015–3027. [[CrossRef](#)]
24. Bars-Cortina, D.; Ramon, E.; Rius-Sansalvador, B.; Guinó, E.; Garcia-Serrano, A.; Mach, N.; Khannous-Lleiffe, O.; Saus, E.; Gabaldón, T.; Ibáñez-Sanz, G.; et al. Comparison between 16S rRNA and Shotgun Sequencing in Colorectal Cancer, Advanced Colorectal Lesions, and Healthy Human Gut Microbiota. *BMC Genom.* **2024**, *25*, 730. [[CrossRef](#)]
25. Costea, P.I.; Zeller, G.; Sunagawa, S.; Pelletier, E.; Alberti, A.; Levenez, F.; Tramontano, M.; Driessen, M.; Hercog, R.; Jung, F.-E.; et al. Towards Standards for Human Fecal Sample Processing in Metagenomic Studies. *Nat. Biotechnol.* **2017**, *35*, 1069–1076. [[CrossRef](#)]
26. Knudsen, B.E.; Bergmark, L.; Munk, P.; Lukjancenko, O.; Priemé, A.; Aarestrup, F.M.; Pamp, S.J. Impact of Sample Type and DNA Isolation Procedure on Genomic Inference of Microbiome Composition. *mSystems* **2016**, *1*, e00095-16. [[CrossRef](#)] [[PubMed](#)]
27. Fouhy, F.; Clooney, A.G.; Stanton, C.; Claesson, M.J.; Cotter, P.D. 16S rRNA Gene Sequencing of Mock Microbial Populations—Impact of DNA Extraction Method, Primer Choice and Sequencing Platform. *BMC Microbiol.* **2016**, *16*, 123. [[CrossRef](#)]
28. Albertsen, M.; Karst, S.M.; Ziegler, A.S.; Kirkegaard, R.H.; Nielsen, P.H. Back to Basics—The Influence of DNA Extraction and Primer Choice on Phylogenetic Analysis of Activated Sludge Communities. *PLoS ONE* **2015**, *10*, e0132783. [[CrossRef](#)]
29. Wesolowska-Andersen, A.; Bahl, M.I.; Carvalho, V.; Kristiansen, K.; Sicheritz-Pontén, T.; Gupta, R.; Licht, T.R. Choice of Bacterial DNA Extraction Method from Fecal Material Influences Community Structure as Evaluated by Metagenomic Analysis. *Microbiome* **2014**, *2*, 19. [[CrossRef](#)]
30. Hsieh, Y.-H.; Peterson, C.M.; Raggio, A.; Keenan, M.J.; Martin, R.J.; Ravussin, E.; Marco, M.L. Impact of Different Fecal Processing Methods on Assessments of Bacterial Diversity in the Human Intestine. *Front. Microbiol.* **2016**, *7*, 1643. [[CrossRef](#)]
31. Gorzelak, M.A.; Gill, S.K.; Tasnim, N.; Ahmadi-Vand, Z.; Jay, M.; Gibson, D.L. Methods for Improving Human Gut Microbiome Data by Reducing Variability through Sample Processing and Storage of Stool. *PLoS ONE* **2015**, *10*, e0134802. [[CrossRef](#)] [[PubMed](#)]
32. Ruiz-Marín, C.M.; Isabel Álvarez-Mercado, A.; Plaza-Díaz, J.; Rodríguez-Lara, A.; Gallart-Aragón, T.; Sánchez-Barrón, M.T.; Lartategui, S.D.R.; Alcaide-Lucena, M.; Fernández, M.F.; Fontana, L. A Clustering Study of Sociodemographic Data, Dietary Patterns, and Gut Microbiota in Healthy and Breast Cancer Women Participating in the MICROMA Study. *Mol. Nutr. Food Res.* **2024**, *68*, 2400253. [[CrossRef](#)]
33. Hunter, S.J.; Easton, S.; Booth, V.; Henderson, B.; Wade, W.G.; Ward, J.M. Selective Removal of Human DNA from Metagenomic DNA Samples Extracted from Dental Plaque. *J. Basic Microbiol.* **2011**, *51*, 442–446. [[CrossRef](#)] [[PubMed](#)]
34. Herlemann, D.P.R.; Labrenz, M.; Jürgens, K.; Bertilsson, S.; Waniek, J.J.; Andersson, A.F. Transitions in Bacterial Communities along the 2000 Km Salinity Gradient of the Baltic Sea. *ISME J.* **2011**, *5*, 1571–1579. [[CrossRef](#)]
35. Bolyen, E.; Rideout, J.R.; Dillon, M.R.; Bokulich, N.A.; Abnet, C.C.; Al-Ghalith, G.A.; Alexander, H.; Alm, E.J.; Arumugam, M.; Asnicar, F.; et al. Reproducible, Interactive, Scalable and Extensible Microbiome Data Science Using QIIME 2. *Nat. Biotechnol.* **2019**, *37*, 852–857. [[CrossRef](#)]
36. Katoh, K.; Standley, D.M. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Mol. Biol. Evol.* **2013**, *30*, 772–780. [[CrossRef](#)]
37. Price, M.N.; Dehal, P.S.; Arkin, A.P. FastTree 2—Approximately Maximum-Likelihood Trees for Large Alignments. *PLoS ONE* **2010**, *5*, e9490. [[CrossRef](#)]
38. Bokulich, N.A.; Kaehler, B.D.; Rideout, J.R.; Dillon, M.; Bolyen, E.; Knight, R.; Huttley, G.A.; Gregory Caporaso, J. Optimizing Taxonomic Classification of Marker-Gene Amplicon Sequences with QIIME 2's Q2-Feature-Classifer Plugin. *Microbiome* **2018**, *6*, 90. [[CrossRef](#)]

39. Beghini, F.; McIver, L.J.; Blanco-Míguez, A.; Dubois, L.; Asnicar, F.; Maharjan, S.; Mailyan, A.; Manghi, P.; Scholz, M.; Thomas, A.M.; et al. Integrating Taxonomic, Functional, and Strain-Level Profiling of Diverse Microbial Communities with bioBakery 3. *eLife* **2021**, *10*, e65088. [\[CrossRef\]](#)
40. Rivera-Pinto, J.; Egozcue, J.J.; Pawlowsky-Glahn, V.; Paredes, R.; Noguera-Julian, M.; Calle, M.L. Balances: A New Perspective for Microbiome Analysis. *mSystems* **2018**, *3*, e00053-18. [\[CrossRef\]](#)
41. Satam, H.; Joshi, K.; Mangrolia, U.; Waghoo, S.; Zaidi, G.; Rawool, S.; Thakare, R.P.; Banday, S.; Mishra, A.K.; Das, G.; et al. Next-Generation Sequencing Technology: Current Trends and Advancements. *Biology* **2023**, *12*, 997. [\[CrossRef\]](#)
42. Maljkovic Berry, I.; Melendrez, M.C.; Bishop-Lilly, K.A.; Rutvisuttinunt, W.; Pollett, S.; Talundzic, E.; Morton, L.; Jarman, R.G. Next Generation Sequencing and Bioinformatics Methodologies for Infectious Disease Research and Public Health: Approaches, Applications, and Considerations for Development of Laboratory Capacity. *J. Infect. Dis.* **2019**, *221*, S292–S307. [\[CrossRef\]](#)
43. Soni, J.; Sinha, S.; Pandey, R. Understanding Bacterial Pathogenicity: A Closer Look at the Journey of Harmful Microbes. *Front. Microbiol.* **2024**, *15*, 1370818. [\[CrossRef\]](#)
44. Yang, L.; Haidar, G.; Zia, H.; Nettles, R.; Qin, S.; Wang, X.; Shah, F.; Rapport, S.F.; Charalampous, T.; Methé, B.; et al. Metagenomic Identification of Severe Pneumonia Pathogens in Mechanically-Ventilated Patients: A Feasibility and Clinical Validity Study. *Respir. Res.* **2019**, *20*, 265. [\[CrossRef\]](#) [\[PubMed\]](#)
45. Moragues-Solanas, L.; Le-Viet, T.; McSorley, E.; Halford, C.; Lockhart, D.S.; Aydin, A.; Kay, G.L.; Elumogo, N.; Mullen, W.; O'Grady, J.; et al. Development and Proof-of-Concept Demonstration of a Clinical Metagenomics Method for the Rapid Detection of Bloodstream Infection. *BMC Med. Genom.* **2024**, *17*, 71. [\[CrossRef\]](#)
46. Mansuri, M.S.; Bathla, S.; Lam, T.T.; Nairn, A.C.; Williams, K.R. Optimal Conditions for Carrying out Trypsin Digestions on Complex Proteomes: From Bulk Samples to Single Cells. *J. Proteom.* **2024**, *297*, 105109. [\[CrossRef\]](#)
47. Li, R.; Liriano, L. A Bone Sample Cleaning Method Using Trypsin for the Isolation of DNA. *Leg. Med.* **2011**, *13*, 304–308. [\[CrossRef\]](#)
48. Han, Y.; Yu, Y.P.; Tseng, G.C.; Luo, J. Trypsin and Reduction Method to Prepare DNA from Formalin-fixed Paraffin-embedded Samples for Methylation Analysis. *Histopathology* **2009**, *54*, 773–775. [\[CrossRef\]](#) [\[PubMed\]](#)
49. Anderson, S.R.; Thompson, L.R. Optimizing an Enclosed Bead Beating Extraction Method for Microbial and Fish Environmental DNA. *Environ. DNA* **2022**, *4*, 291–303. [\[CrossRef\]](#)
50. MetaHIT Consortium (Additional Members); Arumugam, M.; Raes, J.; Pelletier, E.; Le Paslier, D.; Yamada, T.; Mende, D.R.; Fernandes, G.R.; Tap, J.; Bruls, T.; et al. Enterotypes of the Human Gut Microbiome. *Nature* **2011**, *473*, 174–180. [\[CrossRef\]](#)
51. Wensel, C.R.; Pluznick, J.L.; Salzberg, S.L.; Sears, C.L. Next-Generation Sequencing: Insights to Advance Clinical Investigations of the Microbiome. *J. Clin. Investig.* **2022**, *132*, e154944. [\[CrossRef\]](#) [\[PubMed\]](#)
52. Álvarez-Mercado, A.I.; Cano, A.D.V.; Fernández, M.F.; Fontana, L. Gut Microbiota and Breast Cancer: The Dual Role of Microbes. *Cancers* **2023**, *15*, 443. [\[CrossRef\]](#) [\[PubMed\]](#)
53. Caporaso, J.G.; Kuczynski, J.; Stombaugh, J.; Bittinger, K.; Bushman, F.D.; Costello, E.K.; Fierer, N.; Peña, A.G.; Goodrich, J.K.; Gordon, J.I.; et al. QIIME Allows Analysis of High-Throughput Community Sequencing Data. *Nat. Methods* **2010**, *7*, 335–336. [\[CrossRef\]](#)
54. Schloss, P.D.; Westcott, S.L.; Ryabin, T.; Hall, J.R.; Hartmann, M.; Hollister, E.B.; Lesniewski, R.A.; Oakley, B.B.; Parks, D.H.; Robinson, C.J.; et al. Introducing Mothur: Open-Source, Platform-Independent, Community-Supported Software for Describing and Comparing Microbial Communities. *Appl. Environ. Microbiol.* **2009**, *75*, 7537–7541. [\[CrossRef\]](#) [\[PubMed\]](#)
55. Oksanen, J.; Kindt, R.; Legendre, P.; O'Hara, B.; Stevens, M.H.H.; The Vegan Package. Community Ecology Package. Available online: <https://cran.r-project.org/web/packages/vegan/vegan.pdf> (accessed on 3 June 2024).
56. McMurdie, P.J.; Holmes, S. Phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS ONE* **2013**, *8*, e61217. [\[CrossRef\]](#)
57. Wen Chen, J.S. RAM-Package: Analysis of Amplicon-Based Metagenomic Data. Available online: <https://rdrr.io/cran/RAM/man/RAM-package.html> (accessed on 18 June 2020).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.