Received: 29 April 2021 | Accepted: 8 July 2021

DOI: 10.1002/aps3.11444

APPLICATION ARTICLE

Applications in Plant Sciences

Elucidating biofilm diversity on water lily leaves through 16S rRNA amplicon analysis: Comparison of four DNA extraction kits

Kathrin Janssen[1] | Shook Ling Low[2] | Yan Wang[3] | Qi-Yong Mu[3] | Gabriele Bierbaum[1] | Carole T. Gee[2,4]

[1]Institute of Medical Microbiology, Immunology and Parasitology, University Clinic of Bonn, Rheinische Friedrich-Wilhelms-University Bonn, Venusberg-Campus 1, 53127 Bonn, Germany

[2]Institute of Geosciences, Division of Paleontology, Rheinische Friedrich-Wilhelms-University Bonn, Nussallee 8, 53115 Bonn, Germany

[3]CAS Key Laboratory of Tropical Forest Ecology, Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Mengla 666303, China

[4]Huntington Botanical Gardens, 1151 Oxford Road, San Marino, California 91108, USA

Correspondence

Kathrin Janssen, Institute of Medical Microbiology, Immunology and Parasitology, University Clinic of Bonn, Rheinische Friedrich-Wilhelms-University Bonn, Venusberg-Campus 1, 53127 Bonn, Germany.
Email: s5kajans@uni-bonn.de

Carole T. Gee, Institute of Geosciences, Division of Paleontology, Rheinische Friedrich-Wilhelms-University Bonn, Nussallee 8, 53115 Bonn, Germany.
Email: cgee@uni-bonn.de

Abstract

Premise: Within a broader study on leaf fossilization in freshwater environments, a long-term study on the development and microbiome composition of biofilms on the foliage of aquatic plants has been initiated to understand how microbes and biofilms contribute to leaf decay and preservation. Here, water lily leaves are employed as a study model to investigate the relationship between bacterial microbiomes, biodegradation, and fossilization. We compare four DNA extraction kits to reduce biases in interpretation and to identify the most suitable kit for the extraction of DNA from bacteria associated with biofilms on decaying water lily leaves for 16S rRNA amplicon analysis.

Methods: We extracted surface-associated DNA from *Nymphaea* leaves in early stages of decay at two water depth levels using four commercially available kits to identify the most suitable protocol for bacterial extraction, applying a mock microbial community standard to enable a reliable comparison of the kits.

Results: Kit 4, the FastDNA Spin Kit for Soil, resulted in high DNA concentrations with better quality and yielded the most accurate depiction of the mock community. Comparison of the leaves at two water depths showed no significant differences in community composition.

Discussion: The success of Kit 4 may be attributed to its use of bead beating with a homogenizer, which was more efficient in the lysis of Gram-positive bacteria than the manual vortexing protocols used by the other kits. Our results show that microbial composition on leaves during early decay remains comparable and may change only in later stages of decomposition.

KEYWORDS

bacterial DNA sequencing, environmental microbiomics, leaf fossilization, leaf microbiome, *Nymphaea*, plant taphonomy

In many environmental settings, microorganisms form biofilms. Biofilms are surface layers that consist of bacteria and/or fungi, which in their mature form often contain a mixture of species. Microorganisms are embedded in carbohydrates, excreted proteins and lipids, and extracellular DNA (O'Toole et al., 2000; Sutherland, 2001; Jamal et al., 2018). This sessile lifestyle is advantageous for the microbial community, forming a protective barrier against dehydration, antimicrobial substances, and the grazing activity of protozoa, as well as facilitating adhesion to surfaces in the preferred biotopes (Roberson and Firestone, 1992; Decho, 2000; Mah and O'Toole, 2001; Flemming and

Kathrin Janssen and Shook Ling Low contributed equally to this work.

Wingender, 2010). Studies of biofilm formation on plants have mostly focused on the plant–microbe association in the rhizosphere, within xylem or phloem, or on aerial surfaces (Ramey et al., 2004; Danhorn and Fuqua, 2007; Bogino et al., 2013). Microbial communities, on the other hand, have been studied extensively in a wide variety of environments, including freshwater and marine systems, soil, and in mammalian tissue (e.g., Cox et al., 2013; Lee and Eom, 2016; Fierer, 2017; Thompson et al., 2017). While the microbiota associated with marine periphyton (Corcoll et al., 2017) and freshwater fauna (Xue et al., 2018) have also been investigated, data on biofilm formation associated with macrophytes in freshwater environments are limited (Hempel et al., 2009). Microbial biofilms on aquatic macrophytes are known to be involved in various processes, such as the microbial degradation of pollutants or the degradation of dead plant matter (Anesio et al., 2003; Gupta and Pancost, 2006; Reitner, 2011).

Several taphonomic studies on animals have shown that microorganisms are not only involved in tissue degradation, but also serve a protective function (Martin et al., 2005; Raff et al., 2013, 2014; Eagan et al., 2017). One of the most spectacular effects of bacterial biofilms is their putative involvement in the fossilization of soft tissues of animals, in which microbial biofilms and mats engulf the carcass and a mineralized biofilm forms a cast of the animal (e.g., Briggs and Kear, 1993; Darroch et al., 2012; Iniesto et al., 2013, 2017; Raff and Raff, 2014; Raff et al., 2014; Butler et al., 2015), thereby paving the way to fossilization. These studies emphasize the importance of the microbial impact on the decay processes during early fossilization and the shaping of the appearance of the fossil animal remains.

In contrast, there are only a handful of studies on leaf biofilms. Observations were first made by Spicer (1975, 1977, 1991), who noted that bacterial biofilms may lead to precipitation of minerals on plant matter before it is deposited and buried. Dunn et al. (1997) studied mineralization in the lab and in a freshwater stream on naturally occurring bacterial biofilms developing on *Platanus* leaves, while Iniesto et al. (2018) showed that fern pinnules laid on top of microbial mats, a complex type of biofilm, resulted in mineralization. It is thought that the mineral "veil" resulting from a biofilm protects the leaf surface from abrasion and damage caused by transport and invertebrate herbivory, slows down decay processes from further bacterial activity, and preserves fine-scale morphological features of the plant tissue (Spicer, 1991; Locatelli, 2014). The exopolymeric substance of the biofilm serves as a nucleation site due to its anionic characteristics and metal-binding capacity, which are absent on plant surfaces without a biofilm (Dunn et al., 1997; Klymiuk, 2018). Identifying the community assemblages in biofilms on plant leaves and shoots could contribute to a deeper understanding of the relationship between biofilm formation and mineralization in the early stages of fossilization.

The traditional approach for bacterial analysis in microbiology involves culturing, but most environmental bacteria are not readily cultured in the laboratory (Amann et al., 1995; Vartoukian et al., 2010; Stewart, 2012). Thus, for the analysis of complex microbial communities, the best method is to extract the total DNA and amplify the 16S rRNA gene sequences by PCR, followed by high-throughput sequencing and sequence comparison against available databases (McGovern et al., 2018). The 16S rRNA gene is ubiquitous throughout bacteria and archaea (Morgan and Huttenhower, 2012); it is composed of several conserved regions and nine hypervariable regions, and is therefore an ideal tool to identify bacteria (Neefs et al., 1993; Claesson et al., 2010). Moreover, universal primers applied for 16S rRNA gene amplification enable metabarcoding, which facilitates the analysis of the microbial community composition (Kim et al., 2011). The extraction of bacterial DNA from human or environmental samples, such as stools, soils, or water, has been broadly investigated and optimized (Peng et al., 2020; Yang et al., 2020), and various DNA extraction kits are commercially available for bacteria colonizing these niches.

Until now, comparative studies on the efficacy of DNA extraction kits for bacteria found on plants have been carried out mainly on terrestrial species, such as *Arabidopsis thaliana* (L.) Heynh., corn, soybean, or olives (Giangacomo et al., 2021; Haro et al., 2021). Bacterial community composition varies greatly among host species and is influenced by chemical and physical abiotic factors, such as the local environment, temperature, or water availability (Trevathan-Tackett et al., 2020; Giangacomo et al., 2021; Sehnal et al., 2021), as well as by changes occurring during decomposition (Ma et al., 2020a). It is likely that the microbial community composition of terrestrial plants is quite distinct from that of aquatic plants due to the influence of different environmental factors (Jackrel et al., 2017). For instance, the microbiome on freshwater macrophytes such as *Nymphaea* L. water lily leaves will certainly be shaped by microorganisms inhabiting the surrounding water, but the composition of such freshwater phyllosphere communities has rarely been explored (Zhao et al., 2017). Furthermore, the efficacy of extraction kits for microbial communities on freshwater plants has not yet been investigated. Thus, there is a basic need to test DNA extraction protocols prior to experimental studies on freshwater macrophytes to prevent biases in microbiome extraction and sequencing.

Within the context of a larger study on leaf fossilization in freshwater environments, we have embarked on a long-term study on the development and microbiome composition of biofilms on the foliage of aquatic plants, such as water lily leaves, to understand how microbes and biofilms contribute to leaf decay and preservation. Here, we compare four commercial kits originally designed for water or soil. With the exception of the DNeasy PowerWater Kit, which is mainly used for DNA extraction from large water samples, the kits have been used in other studies focusing on plant microbiomes (Tláskal et al., 2016; Zhao et al., 2017; Qiu et al., 2020; Sare et al., 2020). The goal of our study is to identify the most suitable kit for the extraction of DNA from bacteria associated with biofilms on decaying water lily leaves for 16S rRNA amplicon analysis.

METHODS

Study site and sampling

Leaves of the water lily *Nymphaea* sp. (Nymphaeaceae) were collected from a large pond along the southeastern boundary of the Botanical Garden of the University of Bonn, Germany. In October 2020, 24 replicates of floating and subsurface water lily leaves still attached to their parent plant were sampled: 12 floating leaves were taken from the surface of the water, while 12 subsurface leaves were gathered from a level roughly 0.5 m below the water's surface. The leaves all came from the same colony of plants in an area of ca. 2×2 m along the northwestern shore of the pond. Brown leaves that had started to decompose were preferentially selected for analysis.

The large pond supports a sizeable population of water lilies, most of which consist of various species of *Nymphaea*, although there are also *Nuphar* Sm. sp. present. Although the Botanical Garden is located in an urban environment, the pond's ecosystem is made up of a community of aquatic plants and animals in a relatively natural setting, which is replenished by rainfall. In October 2020, precipitation in the area amounted to 73.9 mm, while temperature was on average 11.5°C (https://www.wetterkontor.de).

Mock microbial community standard

To estimate the efficacy of the four different kits, a mock microbial community standard (MMC) was introduced and processed with the samples. The ZymoBIOMICS Microbial Community Standard (D6300; Zymo Research, Irvine, California, USA) is composed of eight bacterial strains (Table 1): *Pseudomonas aeruginosa*, *Escherichia coli*, *Salmonella enterica*, *Lactobacillus fermentum*, *Enterococcus faecalis*, *Staphylococcus aureus*, *Listeria monocytogenes*, and *Bacillus subtilis*. Two fungal species were also present in the mock community, but these were not considered in the 16S rRNA gene sequence analysis.

Preparation for DNA extraction

To dislodge the surface-associated microbial community, the leaves were first manually shaken in 200 mL of 1X TE buffer for 1 min and then sonicated in an ultrasound bath (Branson 1210 Ultrasonic Cleaner; Emerson, St. Louis, Missouri, USA) for 2 min. These treatments were repeated up to three times to maximize the yield of microbial material. Following sonication treatment, two consecutive filtration steps were carried out. The first step was coarse filtration with a folded filter (grade: 3hw; Binzer & Munktell Filter GmbH, Battenberg, Germany) to remove leaf debris and sediment. In the second step, vacuum filtration was used to pass the filtrate through a 0.25-μm pore filter (mixed cellulose ester membrane; Berrytec GmbH, Grünwald, Germany) to retain the bacteria on the filter membrane. Filters were stored in 1 mL of 1X DNA/RNA Shield (Zymo Research) at −20°C until DNA extraction. For mock community analysis, 75 μL of MMC were added to 1X TE buffer and processed in the same manner without the addition of plant material.

DNA extraction

The bacterial DNA from the MMC and from the floating and subsurface leaves was extracted in triplicate using four different kits (Table 2): DNeasy PowerWater Kit (referred to here as Kit 1; QIAGEN, Hilden, Germany), DNeasy PowerSoil Kit (referred to here as Kit 2; QIAGEN), E.Z.N.A. Water DNA Kit (referred to here as Kit 3; Omega Bio-Tek, Norcross, Georgia, USA), and FastDNA Spin Kit for Soil (referred to here as Kit 4; MP Biomedicals, Irvine, California, USA). The extractions were performed following manufacturer instructions, and the DNA was eluted in 50 μL of ultra-pure water (DNase/RNase-free water). DNA concentration and quality were checked using a NanoDrop One/OneC Microvolume-UV/VIS-spectrophotometer and a Qubit 4 Fluorometer (Thermo Fisher Scientific, Waltham, Massachusetts, USA).

TABLE 1 Composition of the ZymoBIOMICS Microbial Community Standard (D6300) used to evaluate the efficacy of the DNA extraction kits

Species	Genome size (Mbp)	Genomic DNA (%)	16S only (%)	Gram staining	Lysis behavior
Pseudomonas aeruginosa	6.77	12	4.2	G−	Easy to lyse
Escherichia coli	5.47	12	10.1	G−	Easy to lyse
Salmonella enterica	4.83	12	10.4	G−	Easy to lyse
Lactobacillus fermentum	2.08	12	18.4	G+	Difficult to lyse
Enterococcus faecalis	3.01	12	9.9	G+	Difficult to lyse
Staphylococcus aureus	2.93	12	15.5	G+	Difficult to lyse
Listeria monocytogenes	2.95	12	14.1	G+	Difficult to lyse
Bacillus subtilis	3.98	12	17.4	G+	Difficult to lyse

Abbreviations: G−, Gram negative; G+, Gram positive; Mbp, mega base pairs.

TABLE 2 The four DNA extraction kits selected for the study of microbial biofilms on *Nymphaea* leaves

Kit	Company	Bead beating	Average processing time	Use in previous studies
DNeasy PowerWater Kit (Kit 1)	QIAGEN	In vortex adapter	~1.5 h	Rocha and Manaia (2020)
DNeasy PowerSoil Kit (Kit 2)	QIAGEN	In vortex adapter	~1.5 h	Tláskal et al. (2016); Qiu et al. (2020)
E.Z.N.A. Water DNA Kit (Kit 3)	Omega Bio-Tek	Via vortexing	~4 h	Zhao et al. (2017)
FastDNA Spin Kit for Soil (Kit 4)	MP Biomedicals	With a homogenizer*	~1 h	Sare et al. (2020)

*Homogenization with a Precellys 24 Tissue Homogenizer.

16S rRNA gene amplicon sequencing

For 16S rRNA gene sequencing, the V4 variable region of the 16S rRNA gene sequence was amplified with the specific 16S primers of 16s-515F (GTGCCAGCMGCCGCGGTAA) and 16s-806R (GGACTACVSGGGTATCTAAT) (Caporaso et al., 2011). The PCR reaction was performed as a single-step PCR with the HotStarTaq Plus Master Mix Kit (QIAGEN) including an initial denaturation at 95°C for 5 min, followed by 30–35 cycles at 95°C for 30 s, 53°C for 40 s, and 72°C for 1 min, with a final elongation step at 72°C for 10 min. Paired-end sequencing (bTEFAP) was performed by MR DNA (http://www.mrdnalab.com; Shallowater, Texas, USA) on the MiSeq sequencing platform (Illumina, San Diego, California, USA) following the manufacturer's guidelines (Dowd et al., 2008). Raw sequence data were processed using QIIME2 (Bolyen et al., 2019) with default parameters unless otherwise noted. The DADA2 pipeline was used for sequence quality control, denoising, and chimeric filtering (Callahan et al., 2016). Taxonomy classification of the final zero-radius operational taxonomic units (zOTUs) was performed with BLASTN against a curated database from the National Center for Biotechnology Information.

All raw sequence data related to this study are deposited in the European Nucleotide Archive (European Bioinformatics Institute, EMBL-EBI) database (study accession no. 471 PRJEB43756).

Bioinformatics and statistics

Rarefaction curves for analysis of sequencing depth were conducted via QIIME2 (Bolyen et al., 2019). DNA concentrations were compared using the Kruskal–Wallis test in R (R Core Team, 2020). To evaluate the efficacy of the DNA extraction kits tested (e.g., DNA yield, genus frequencies), Kruskal–Wallis tests with the P value adjusted via the Benjamini–Hochberg method were performed in R, followed by a Dunn's post-hoc test when needed. For all the statistical methods, a P value of <0.05 was considered as statistically significant.

Analysis of alpha diversity metrics and the Bray–Curtis dissimilarity matrix were performed using the R package "vegan" (version 2.5.6; Oksanen et al., 2019). The significance analysis of the efficacy of the extraction kits applied in this study was determined following the estimation error analysis proposed by Yang et al. (2020) and non-metric multidimensional scaling (NMDS) distance dependency with Bray–Curtis dissimilarity dependencies. Heat trees were developed with the R packages "metacoder" (version 0.3.4; Foster et al., 2017) and "taxa" (version 0.3.4; Foster et al., 2018).

The microbiome shared by the four different microbial samples was compared and visualized through a Venn diagram plot (R package "VennDiagram," version: 1.6.20; Chen, 2018). The Wilcoxon unpaired t-test was used to compare the significance of characteristics of the bacterial microbiomes found in the two sets of leaves, such as oxygen tolerance, biofilm-forming ability, and Gram-staining properties.

Visualizations of the statistics and microbial community composition were performed with the R package "ggplot2" (Wickham, 2016).

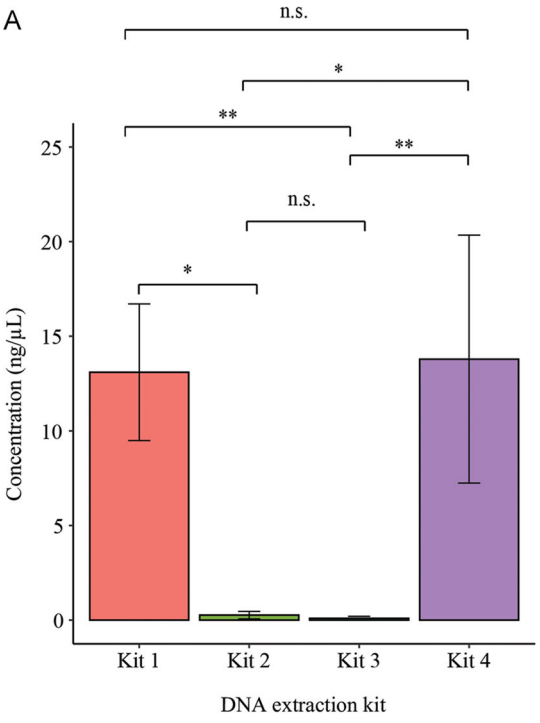
Allocation of phenotypes

Characteristics of bacteria were predicted via BugBase (Ward et al., 2017). The sequences processed were taxonomically classified for BugBase analysis using the Greengenes database (McDonald et al., 2012).

RESULTS

Processing time, handling, and DNA yield

The four kits tested in this study varied greatly in processing time and ease of handling (Table 2). While DNA extraction with Kits 1, 2, and 4 took between 1 to 2 h, the processing time of Kit 3 lasted up to 4 h and included several complicated steps. Furthermore, DNA concentrations extracted from leaf samples differed remarkably among the four extraction protocols (Figure 1). The highest DNA yield, detected via Qubit fluorometry, was obtained with Kit 4 with an average concentration of 13.79 ng/μL, followed by Kit 1 with an average yield of 13.1 ng/μL. Kits 2 and 3 resulted in significantly lower DNA concentrations (<1 ng/μL). The DNA yield of the MMC samples mirrored these results by showing a comparable trend in which Kits 1 and 4 yielded higher concentrations than Kits 2 and 3. All kits included a



B

Comparison	Adjusted *P* value	Significance
Kit 1 – Kit 2	0.021	*
Kit 1 – Kit 3	0.004	**
Kit 2 – Kit 3	0.555	n.s.
Kit 1 – Kit 4	0.806	n.s.
Kit 2 – Kit 4	0.014	*
Kit 3 – Kit 4	0.004	**

FIGURE 1 Concentrations of the DNA extracted by each of the four tested kits as measured by a Qubit 4 Fluorometer. (A) DNA concentration (in nanograms per microliter) yielded by each kit. Error bars show standard deviation. (B) Results of the Kruskal–Wallis test to calculate pairwise comparisons between the four extraction methods. *Benjamini-Hochberg–adjusted Dunn $P < 0.05$, **Benjamini-Hochberg–adjusted Dunn $P < 0.01$

bead-beating step, but varied in their recommended procedures. Kits 1, 2, and 3 implemented the bead beating via manual vortexing, whereas Kit 4 recommended the use of a homogenizer (Precellys 24 Tissue Homogenizer; Bertin Technologies, Montigny-le-Bretonneux, France).

Quality check sequencing

The sequencing was checked twice—once by the sequencing facility and once by the researchers—via QIIME2 (see Bolyen et al., 2019). In all samples, read quality did not diminish until a read length of 200 bp, which was taken as the threshold for trimming in the denoising step. After denoising and merging, over 80% of the input reads remained for samples of all four kits. Filtering of non-chimeric reads diminished the percentage of input reads to 50–60% (Appendix 1). The rarefaction curve of the samples generated by the four different kits revealed that all samples were sequenced to a sufficient depth to reliably estimate the Shannon diversity (Figure 2).

Evaluation of DNA extraction protocols on the mock community

The use of an MMC standard enabled a reliable comparison of the four extraction kits. The eight bacterial strains were equally frequent in the mock community, each comprising 12% of the microbiome. However, due to the variability in 16S rRNA gene copy number in each taxon, the abundance of 16S rRNA gene sequences differed among the bacteria (Table 1). In general, the relative abundance of bacterial species showed no trend in overestimation or underestimation of Gram-positive or Gram-negative bacteria (Figure 3A). Specifically, however, all four kits underestimated the relative abundance of *Enterococcus*, although Kit 4 exhibited the lowest deviations from the expected

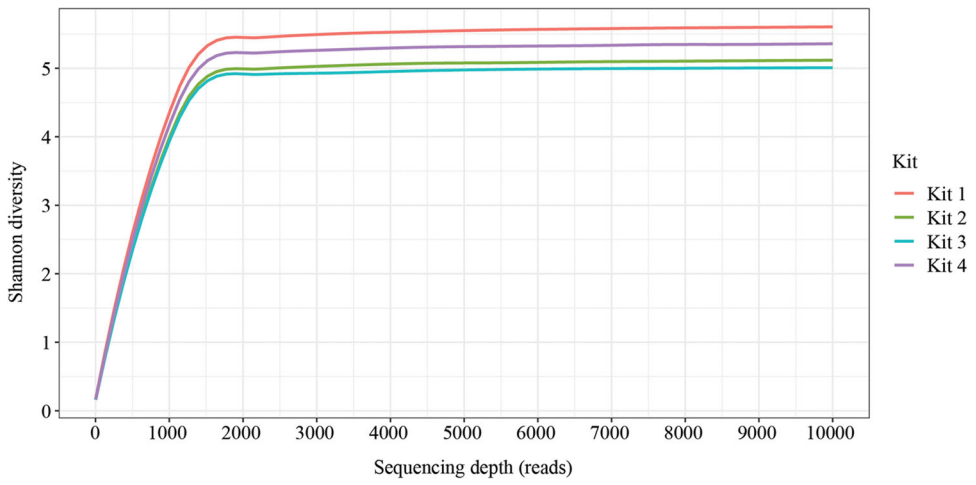


FIGURE 2 Rarefaction curve depicting the sequencing depth and alpha diversity of the samples

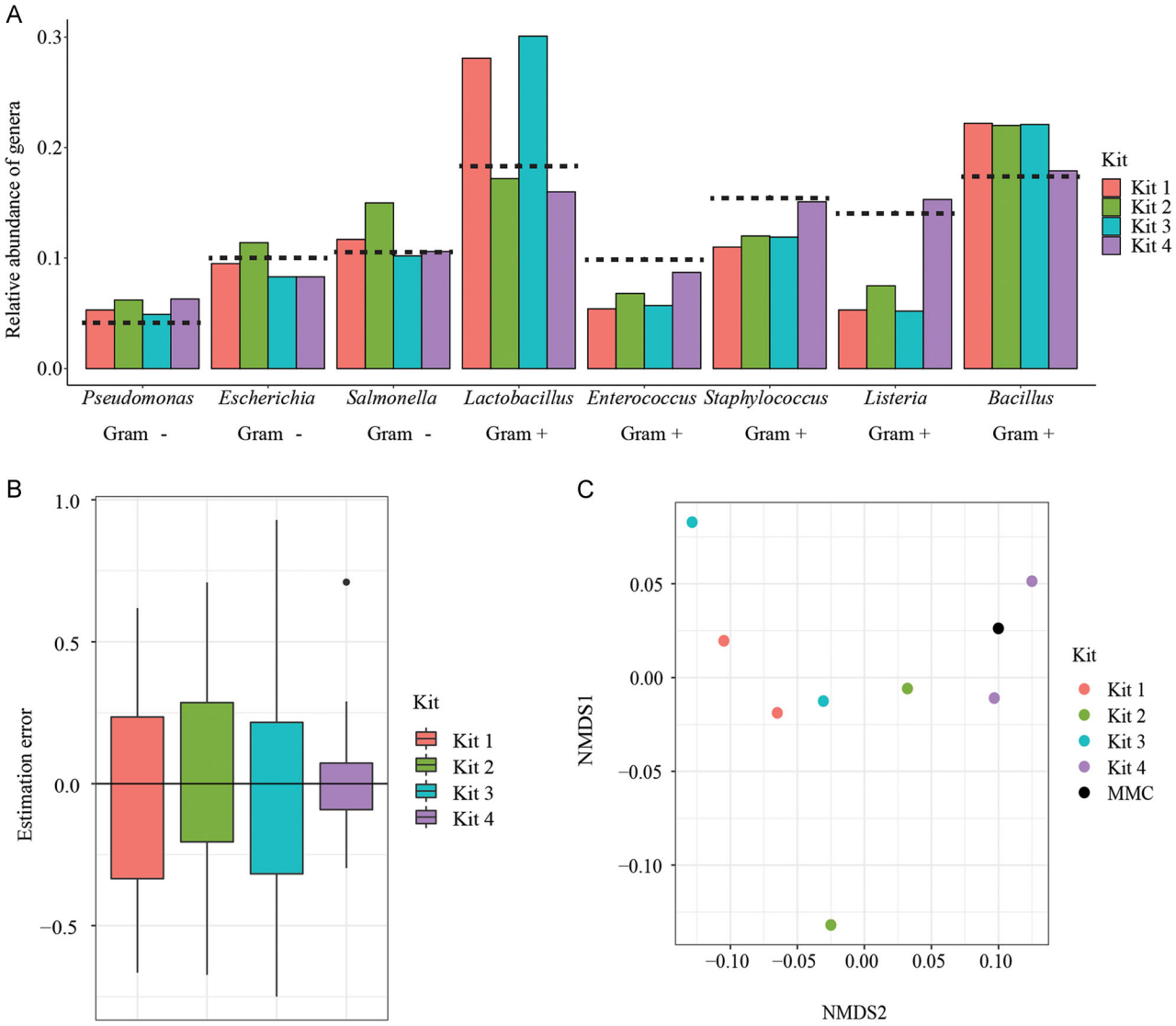


FIGURE 3 Accuracy of the four tested DNA extraction protocols on the mock microbial community standard (MMC). (A) Relative mean abundance of the eight bacterial genera in the MMC. Dashed lines indicate the ideal abundance of 16S sequences for each genus in the MMC. (B) Estimation error of the eight bacterial species in the MMC. (C) NMDS plot of Bray–Curtis dissimilarities between the microbial communities extracted by the four kits and the ideal microbiome of the MMC

values. Kits 1, 2, and 3 were not able to efficiently lyse *Listeria* and *Staphylococcus* and overestimated the amount of *Bacillus*. Additionally, Kits 1 and 2 greatly overestimated the frequency of *Lactobacillus*. In general, Kit 4 showed the greatest accuracy in estimating the relative abundance of bacteria, with a slightly lower mean estimation error and a smaller Bray–Curtis distance compared to the ideal mock abundances (mean estimation error: 0.022 ± 0.231; Figure 3B, C), erring only in an underestimation of *Lactobacillus*. The intra-protocol variability for the replicates of microbial abundance estimation showed differences between Kits 2 and 3, whereas Kits 1 and 4 exhibited comparable values and tendencies (Figure 3C).

In addition to the eight strains of the MMC, low-abundance genera that could introduce bias in the interpretation of diversity and microbial community composition were detected in all samples. To account for this, the threshold for further analyses of the samples was set to a relative abundance of 0.01 to minimize potential misinterpretation.

Evaluation of DNA extraction protocols on plant samples

The calculation of alpha diversity, as measured with the Shannon index and generic richness, did not reveal any significant inter-protocol differences between the four kits after threshold implementation (Figure 4A). However, there was a slight tendency toward lower alpha diversity in Kits 2 and 3 than in Kits 1 and 4, and the median of Kit 1 (Shannon index = 3.41) was slightly higher than in Kit 4

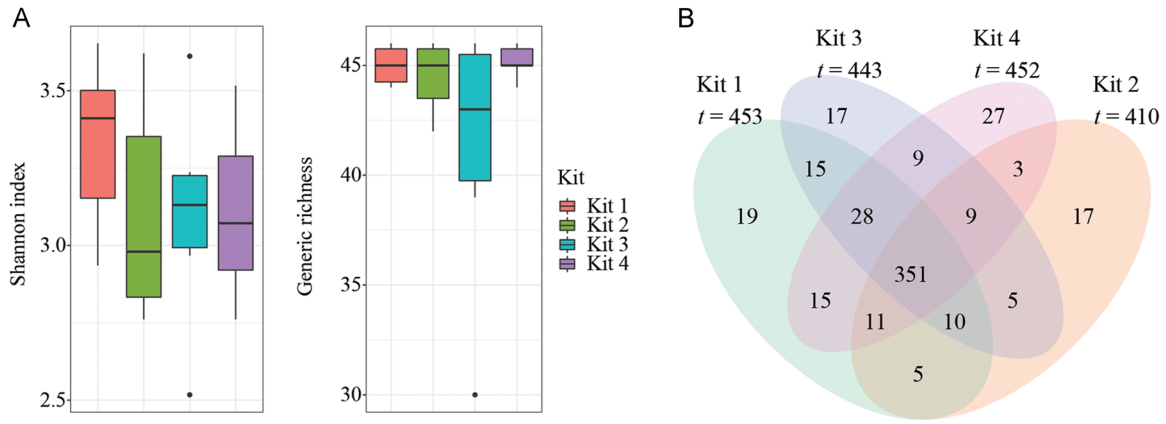


FIGURE 4 Comparison of alpha diversity in the bacterial microbiomes as extracted from all leaves by each of the four kits. (A) Comparisons using the Shannon index and generic richness. Taxa that were frequently under the threshold of 0.01 were excluded before calculation. (B) Venn diagram showing the taxa that all kits yielded in common, as well as one-to-one comparisons between each kit. Taxa with a relative abundance under 0.01 are included. *t* = total

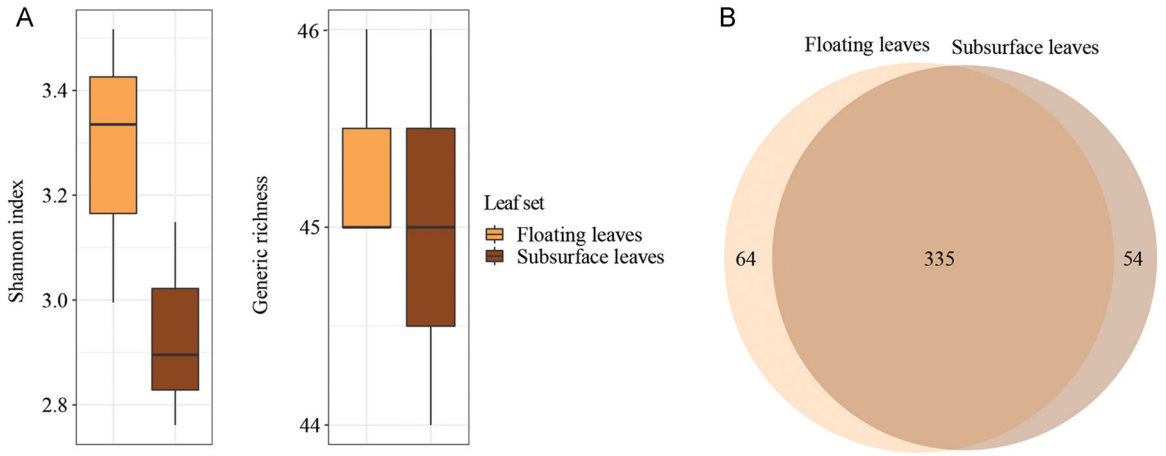


FIGURE 5 Comparison of alpha diversity in the bacterial microbiomes of the floating vs. the subsurface leaves. (A) Comparisons using the Shannon index and generic richness. Taxa that were frequently under the threshold of 0.01 were excluded before calculation. (B) Venn diagram showing the number of genera that each set of leaves shared in common, as well as the unique genera in each set. Taxa with a relative abundance under 0.01 are included. *t* = total

(Shannon index = 3.07). The Venn diagram for the genera detected by the four kits supports our previous observations (Figure 4B). The DNA of a core microbiome ($n = 351$) could be extracted using all four kits, and only genera that were abundant under the threshold were unique to the results produced with each kit.

16S rRNA amplicon analysis on the two sets of leaves (floating vs. subsurface) using the most suitable extraction kit (Kit 4) revealed negligible difference in the alpha diversity between the floating and subsurface leaves (Figure 5A). A Venn diagram of bacterial diversity supports the observation that the colonization between the physical stages is mostly comparable; the floating and subsurface leaves have 335 genera in common (Figure 5B). Only 64 and 54 genera, respectively, were uniquely present in either the set of floating or subsurface leaves, and these were below the predetermined threshold of relative abundance of 0.01.

The microbial community composition of both sets of leaves at the different water depths was dominated by members of the phylum Proteobacteria, followed by the phylum Bacteroidetes (Figures 6, 7). Only a small proportion of Gram-positive bacteria were present. In particular, β-Proteobacteria, such as *Rhodoferax* sp., were highly frequent in both sets of leaves.

There were, however, some compositional differences between the two sets of leaves. In the results yielded by Kit 4, in the set of floating leaves, α-Proteobacteria (e.g., *Novosphingobium* sp.) and Bacteroidetes (e.g., *Arcicella* sp. or *Flectobacillus* sp.) were more abundant than in the subsurface leaves, in addition to several other β-Proteobacteria (e.g., *Duganella* sp. or *Janthinobacterium* sp.). In contrast,

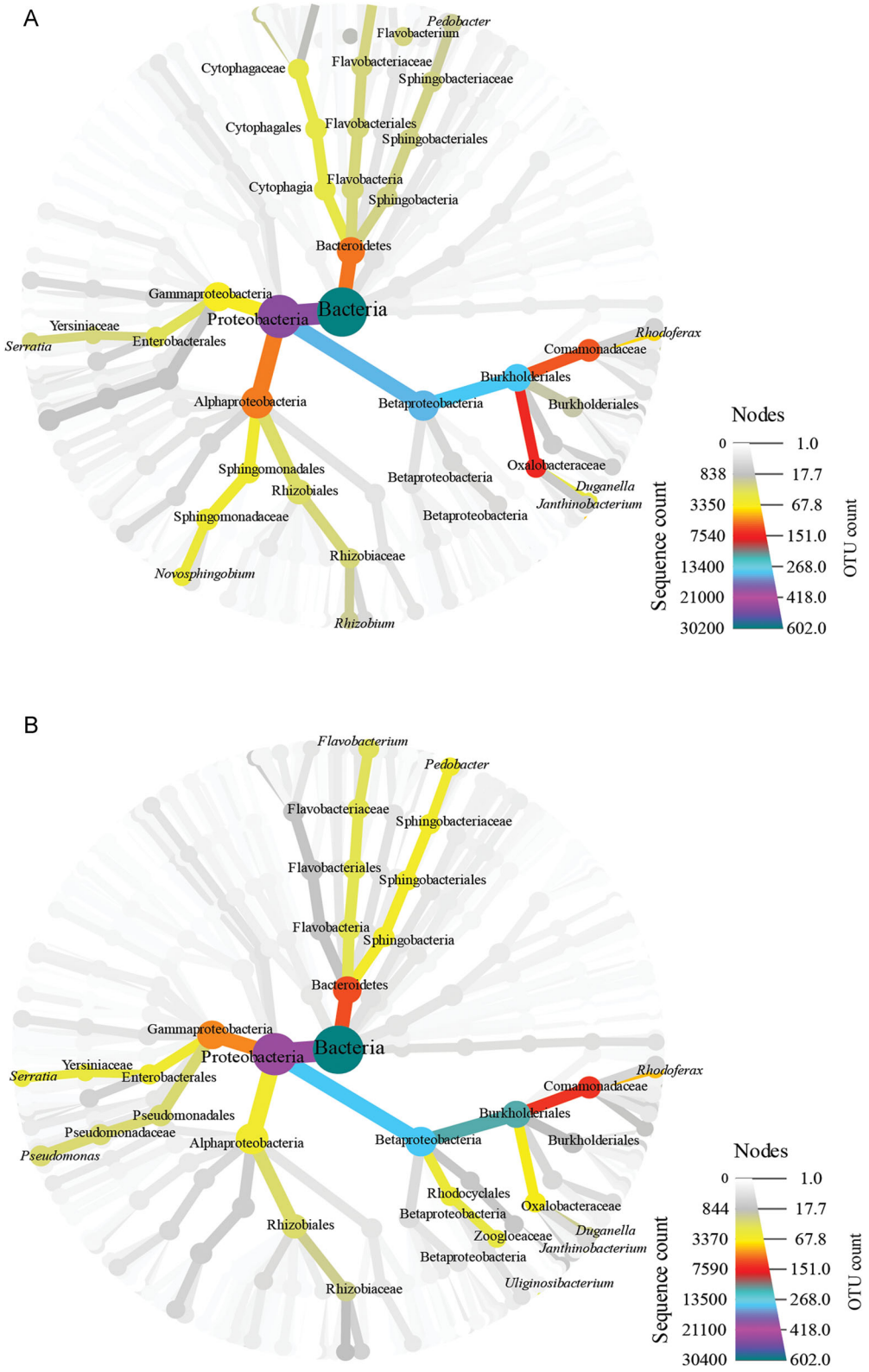


FIGURE 6 Comparison of microbial community composition by taxonomic rank on the two sets of leaves. Nodes with zero-radius operational taxonomic unit (zOTU) counts greater than 1000 are highlighted by colors. (A) Microbiome of the floating leaves. (B) Microbiome of the subsurface leaves

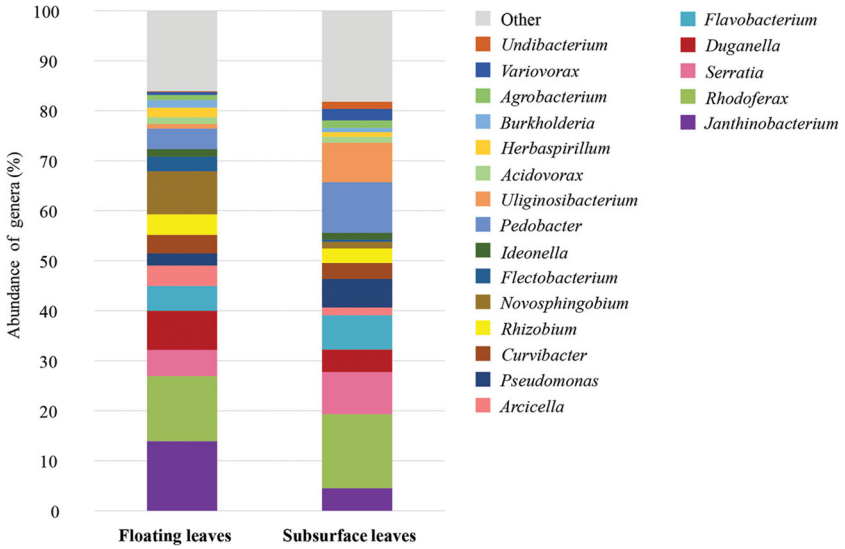


FIGURE 7 Abundance of genera (%) on the floating vs. the subsurface leaves

the subsurface leaves were more strongly colonized by γ-Proteobacteria (e.g., *Pseudomonas* sp. or *Serratia* sp.), as well as by Bacteroidetes (e.g., *Flavobacterium* sp. or *Pedobacter* sp.) and β-Proteobacteria (e.g., *Uliginosibacterium* sp. or *Variovorax* sp.). However, compositional differences in the microbial community between the two sets of leaves were not significant for any of the genera.

BugBase predictions on the taxonomic classifications of the bacteria revealed that there were no significant differences in the oxygen tolerance, biofilm formation ability, or Gram-staining properties between the microbial communities on the two sets of leaves (Figure 8). Biofilm-forming ability was a particularly strong characteristic in both microbiomes, and the majority of bacteria in both sets of leaves were aerobic bacteria.

DISCUSSION

Comparison of the four kits

Biases in the interpretation of microbial community composition may be introduced during various steps in the microbiome methodology. DNA extraction protocols in particular can play a major role in shifting results and influencing analyses. Therefore, it is essential to investigate the efficacy of DNA extraction protocols prior to the start of the main experiments. The four DNA extraction kits tested in this study were chosen on the basis of previous studies that had yielded good sequencing results. Two of the kits were designed especially for soil samples (Kits 2 and 4) and had been previously used for extracting DNA from decomposing leaf litter (Kit 2; Tláskal et al., 2016) or in other studies that compared extraction procedures to minimize the introduction of bias (Kit 2, Qiu et al., 2020; Kit 4, Sare et al., 2020). Kit 3 was used to extract DNA from macrophyte leaf litter decomposition (Zhao et al., 2017). Prior to our study, Kit 1 had been specialized and optimized for the analysis of water microbiomes (Rocha and Manaia, 2020), but had not yet been applied to the DNA extraction of plant microbiomes. Dislodging the surface-associated microbiome by washing and ultrasound treatment prevents the co-extraction of chloroplast or mitochondrial DNA that might be introduced by grinding or the physico-chemical lysis of the plant material (Lutz et al., 2011; Fitzpatrick et al., 2018). This issue can also be overcome, however, by the choice of specific primers that reduce the detection of plastids and mitochondria that are descended from bacteria (Sagan, 1967; Sakai et al., 2004; Beckers et al., 2016).

Of the four DNA extraction protocols tested, Kit 4 is the most suitable for DNA extraction and microbiome sequencing. It not only resulted in higher DNA concentrations with better quality, but also yielded the most accurate depiction of the microbial community using the MMC standard. This was because the bead-beating step performed with a homogenizer was far more efficient in the lysis of the strong cell wall in Gram-positive bacteria (*Enterococcus*, *Staphylococcus*, *Listeria*) than the manual vortexing protocols. In addition, it may be assumed that the beads cannot access the entire filter surface during manual vortexing. Such incomplete lysis could lead to misinterpretations and false conclusions on the microbial community composition as it underestimates or overestimates bacterial groups (Pollock et al., 2018; Ma et al., 2020b). Whereas the depiction of the frequencies of Gram-negative bacteria was accurate among the four kits, there was some overestimation, especially with the strains of *Bacillus* and *Lactobacillus*. Previous studies (Yuan et al., 2012; Hughes et al., 2017; Ojo-Okunola et al., 2020; Yang et al., 2020) had also observed a considerable overestimation of *Lactobacillus* sp. and *Bacillus* sp. with the MMC provided by ZymoResearch. The growth phase of the bacteria may also influence lysis efficiency, as *Bacillus* sp. show remodeling of their cell wall during different growth phases (Li et al., 2018). In this context, Song et al. (2021) detected decreased lysis efficacy of stationary *Lactobacillus* sp. cultures.

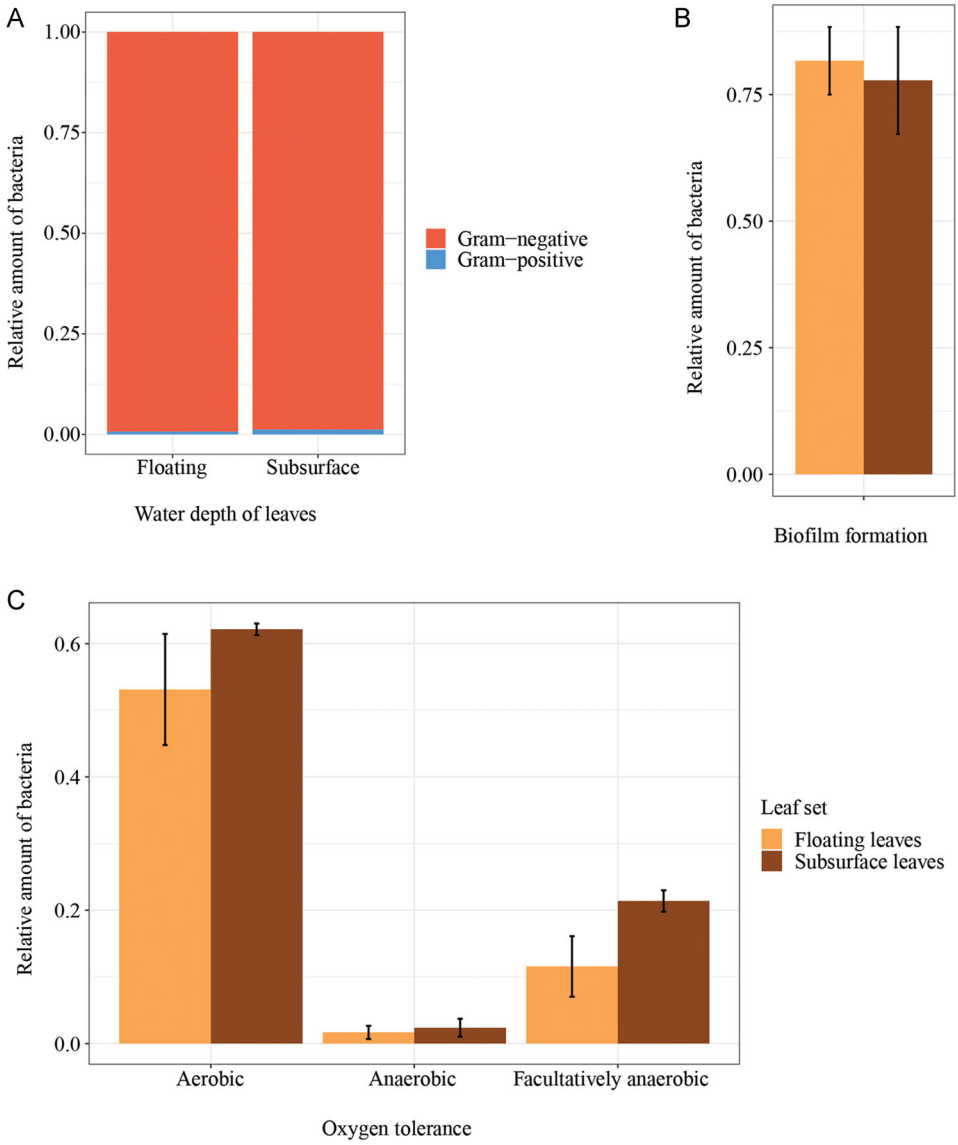


FIGURE 8 Comparison of predicted bacterial characteristics of the microbiomes on the two sets of leaves. Error bars show standard deviation. (A) Gram-staining properties. (B) Biofilm-forming ability. (C) Oxygen tolerance

Furthermore, other steps in the DNA extraction protocols, such as ultrasound treatment and washing procedure, should be closely scrutinized. Other studies have investigated the effect of multiple washing steps and concluded that this could have a substantial impact on community composition analysis (Sare et al., 2020).

Our sequencing results of the MMC highlighted another source of bias. In addition to the eight major bacterial species in the MMC, up to 49 additional genera could be detected, albeit in low frequencies (<0.01 relative abundance). Only *Enterobacter* sp. was present in higher frequencies, composing up to 5% of the detected zOTUs of the entire sample. Contamination with this genus has been reported previously (Salter et al., 2014). Contaminant sequences could be introduced by chemicals or equipment during DNA extraction, library preparation, or sequencing (McFeters et al., 1993; McAlister et al., 2002; Grahn et al., 2003; Shen et al., 2006), and these impurities might critically skew community analysis, especially if samples exhibit a low microbial biomass (Salter et al., 2014; Glassing et al., 2016; Quince et al., 2017). Therefore, caution should be exercised concerning taxa occurring in low frequencies, and thresholds should be set to exclude low-level zOTU reads from the evaluation. In our case, the threshold for the relative abundance was set to 0.01. All zOTUs present below the threshold were excluded from the analyses, except for their depiction in the Venn diagrams, to highlight the relevance of the core microbiome. Additional controls, such as extraction and reagent blanks, should be introduced in future experiments to check for contamination.

Microbial community composition on macrophytes

It has been previously shown that the bacterial communities on subsurface macrophytes are influenced by the macrophyte species as well as by the body of water (Zeng et al., 2012; Xian et al., 2020). The microbial community on the water lily leaves at our study site in Bonn mainly consisted of Proteobacteria, which made up the majority of the microbiome (76% of the zOTUs). Similarly, a high proportion of δ-Proteobacteria, β-Proteobacteria, and γ-Proteobacteria totaling up to 72% of the entire microbiome was found in a freshwater lake with water lilies in Lake Kolon, Hungary (Mentes et al., 2018). In contrast, water samples taken in the water lily stands of the eutrophic Lake Taihu in China yielded only 40% Proteobacteria (Zeng et al., 2012). Another study by Zhao et al. (2017) detected a decrease of γ-Proteobacteria and an increase of Firmicutes, β-, δ-, and α-Proteobacteria in decomposing macrophyte leaf litter over time. In our study, it was observed that the amount of γ-Proteobacteria on subsurface leaves was slightly higher than on floating leaves.

Whereas some studies on the decay of leaf litter have shown an increase in diversity throughout decomposition time (up to 24 months; Tláskal et al., 2016; Zhao et al., 2017), our study yielded a slightly lower diversity in subsurface leaves. However, the differences in diversity were not significant, and it is obvious that subsurface leaves represent an early stage in decomposition soon after the floating leaves have turned brown and begun to sink. In most freshwater macrophytes, the decomposition rate is highest during the first 10 days. During this time, fungi dominate and start the initial decay of plant matter (Zhao et al., 2017; Ma et al., 2020a). This highlights the importance of the additional study of fungal composition in future experiments in order to understand all aspects of leaf decomposition. Shortly after the onset of decomposition in the leaf cuticle, bacteria can start metabolizing the other plant cell polymers, and the microbial frequencies in the leaf-associated community may begin to shift in accordance with the availability of resources (Tláskal et al., 2016; Zhao et al., 2017).

Extracts from various *Nymphaea* species have been shown to exhibit antibacterial effects that may, along with phenolic compounds that are shed from the leaves into the biofilm, inhibit the growth of specific microbial genera (Hempel et al., 2009; Parimala and Shoba, 2014; Mechesso et al., 2019). However, some of the bacterial genera encountered on the water lily leaves sampled in our study, such as *Novosphingobium*, *Burkholderia*, and *Pseudomonas*, are also able to decompose complex organic compounds (Sohn et al., 2004; Liu et al., 2005; Sharma et al., 2017). Furthermore, the genera *Pseudomonas* and *Burkholderia* are known for their capacity to excrete cellulase (Liang et al., 2014; Zhang et al., 2016) and have been detected in similar studies on *Nymphaea* leaf litter (Zhao et al., 2017). In concert with earlier studies, cyanobacteria that are characteristic for a site without macrophyte stands, such as in Lake Taihu (Zeng et al., 2012), were present only in very low concentrations in the Bonn samples (0.3% of signals).

Fossilization of plant material can be induced by different chemical processes, including authigenic preservation, which is dependent on encrustation with minerals. It has been shown that the biofilm-forming activity of bacteria plays an important role in this process (Dunn et al., 1997; Iniesto et al., 2018). In experiments with dicot leaves that were incubated in mineral solutions containing 10 mM $FeCl_3$, only leaves with preformed biofilms were able to adsorb metal ions (Dunn et al., 1997). This adsorption was caused by the negative surface charges of bacteria and the biofilm polymer, which resulted in iron encrustations that subsequently inhibited microbial activity and decay. In long-term experiments with fern leaves (Iniesto et al., 2018), microbial mats entombed the plant material and supported maintenance of tissue integrity as well as mineralization processes. In this sense, the presence of biofilm-forming bacteria on leaves in a freshwater setting, as predicted by BugBase, would be the first prerequisite for fossilization.

CONCLUSIONS

Through comparative statistical analysis of microbial communities extracted from plant surfaces using four DNA extraction protocols, we identified the kit that offers the best yield and most accurate depiction of bacterial communities of biofilms on decaying water lily leaves and thereby discovered a diverse bacterial community composed mainly of biofilm-forming proteobacteria.

In general, the use of diverse protocols—varying in the choice of commercially available kits, diverse methods for the harvesting of bacteria, and the use of different sequencing primers—introduces biases during extraction, library preparation, and sequencing, and makes comparisons between different studies on plant microbiomes difficult. It would be very useful if, in analogy to the National Institutes of Health's Human Microbiome Project, recommendations and protocols to enhance the comparability and reproducibility of DNA extractions were established for environmental studies.

ACKNOWLEDGMENTS

The authors sincerely thank Dr. Cornelia Löhne and Stefan Giefer (Botanical Garden at the University of Bonn) for access to and assistance with collecting water lily leaves, as well as *APPS* Managing Editor Beth Parada and Senior Associate Editor Michael J. Moore, and two anonymous reviewers for helpful comments. This is contribution number 35 of the Deutsche Forschungsgemeinschaft (DFG) Research Unit FOR 2685 "The Limits of the Fossil Record: Analytical and Experimental Approaches to Fossilization" (grant numbers 396706817 to C.T.G. and 396704301 to G.B.).

AUTHOR CONTRIBUTIONS

K.J. and S.L.L. performed sample preparation, DNA extraction, and data analysis. K.J. wrote the manuscript with contributions from C.T.G., G.B., and S.L.L. S.L.L. collected the samples. Y.W, Q.-Y.M., and G.B. contributed to data analysis. K.J. and G.B. interpreted the findings. C.T.G conceived of and designed the study, as well as provided funding. All authors approved the final manuscript.

DATA AVAILABILITY STATEMENT

All raw sequence data related to this study are deposited in the European Nucleotide Archive (European Bioinformatics Institute, EMBL-EBI) database (study accession no. 471 PRJEB43756).

REFERENCES

Amann, R. I., W. Ludwig, and K.-H. Schleifer. 1995. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiological Reviews* 59: 143–169.

Anesio, A. M., P. C. Abreu, and B. A. Biddanda. 2003. The role of free and attached microorganisms in the decomposition of estuarine macrophyte. *Estuarine, Coastal and Shelf Science* 56: 197–201.

Beckers, B., M. Op De Beeck, S. Thijs, S. Truyens, N. Weyens, W. Boerjan, and J. Vangronsveld. 2016. Performance of 16s rDNA primer pairs in the study of rhizosphere and endosphere bacterial microbiomes in metabarcoding studies. *Frontiers in Microbiology* 7: 650.

Bogino, P. C., M. de las Mercedes Oliva, F. G. Sorroche, and W. Giordano. 2013. The role of bacterial biofilms and surface components in plant–bacterial associations. *International Journal of Molecular Sciences* 14: 15838–15859.

Bolyen, E., J. R. Rideout, M. R. Dillon, N. A. Bokulich, C. C. Abnet, G. A. Al-Ghalith, H. Alexander, et al. 2019. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nature Biotechnology* 37: 852–857.

Briggs, D. E. G., and A. J. Kear. 1993. Fossilization of soft tissue in the laboratory. *Science* 259: 1439–1442.

Butler, A. D., J. A. Cunningham, G. E. Budd, and P. C. J. Donoghue. 2015. Experimental taphonomy of *Artemia* reveals the role of endogenous microbes in mediating decay and fossilization. *Proceedings of the Royal Society B: Biological Sciences* 282: 1–10.

Callahan, B. J., P. J. McMurdie, M. J. Rosen, A. W. Han, A. J. A. Johnson, and S. P. Holmes. 2016. DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods* 13: 581–583.

Caporaso, J. G., C. L. Lauber, W. A. Walters, D. Berg-Lyons, C. A. Lozupone, P. J. Turnbaugh, N. Fierer, and R. Knight. 2011. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proceedings of the National Academy of Sciences, USA* 108: 4516–4522.

Chen, H. 2018. VennDiagram: Generate high-resolution Venn and Euler plots. R package version 1.6.20, 1(1). Website: https://rdrr.io/cran/VennDiagram/ [accessed 28 June 2021].

Claesson, M. J., Q. Wang, O. O'Sullivan, R. Greene-Diniz, J. R. Cole, R. P. Ross, and P. W. O'Toole. 2010. Comparison of two next-generation sequencing technologies for resolving highly complex microbiota composition using tandem variable 16S rRNA gene regions. *Nucleic Acids Research* 38: e200.

Corcoll, N., T., Österlund, L. Sinclair, A. Eiler, E. Kristiansson, T. Backhaus, and K. M. Eriksson. 2017. Comparison of four DNA extraction methods for comprehensive assessment of 16S rRNA bacterial diversity in marine biofilms using high-throughput sequencing. *FEMS Microbiology Letters* 364: fnx139.

Cox, M. J., W. O. Cookson, and M. F. Moffatt. 2013. Sequencing the human microbiome in health and disease. *Human Molecular Genetics* 22(R1): R88–R94.

Danhorn, T., and C. Fuqua. 2007. Biofilm formation by plant-associated bacteria. *Annual Reviews of Microbiology* 61: 401–422.

Darroch, S. A. F., M. Laflamme, J. D. Schiffbauer, and D. E. G. Briggs. 2012. Experimental formation of a microbial death mask. *Palaios* 27: 293–303.

Decho, A. W. 2000. Exopolymer microdomains as a structuring agent for heterogeneity within microbial biofilms, pp. 9–15. *In* R. E. Riding and S. M. Awramik [eds.], Microbial sediments. Springer-Verlag, Berlin, Germany.

Dowd, S. E., T. R. Callaway, R. D. Wolcott, Y. Sun, T. McKeehan, R. G. Hagevoort, and T. S. Edrington. 2008. Evaluation of the bacterial diversity in the feces of cattle using 16S rDNA bacterial tag-encoded FLX amplicon pyrosequencing (bTEFAP). *BMC Microbiology* 8: 125.

Dunn, K. A., R. J. C. McLean, G. Upchurch, and R. L. Folk. 1997. Enhancement of leaf fossilization potential by bacterial biofilms. *Geology* 25: 1119–1122.

Eagan, J. L., M. E. Andrews, R. L. Pearson, F. R. Turner, E. C. Raff, and R. A. Raff. 2017. Identification and modes of action of endogenous bacteria in taphonomy of embryos and larvae. *Palaios* 32(4): 206–217. https://doi.org/10.2110/palo.2016.071

Fierer, N. 2017. Embracing the unknown: Disentangling the complexities of the soil microbiome. *Nature Reviews Microbiology* 15: 579–590.

Fitzpatrick, C. R., P. Lu-Irving, J. Copeland, D. S. Guttman, P. W. Wang, D. A. Baltrus, K. M. Dlugosch, and M. T. J. Johnson. 2018. Chloroplast sequence variation and the efficacy of peptide nucleic acids for blocking host amplification in plant microbiome studies. *Microbiome* 6: 144.

Flemming, H. C., and J. Wingender. 2010. The biofilm matrix. *Nature Reviews Microbiology* 8: 623–633.

Foster, Z. S. L., T. J. Sharpton, and N. J. Grünwald. 2017. Metacoder: An R package for visualization and manipulation of community taxonomic diversity data. R package version 0.3.4. *PLoS Computational Biology* 13: e1005404.

Foster, Z. S. L., S. Chamberlain, and N. J. Grünwald. 2018. Taxa: An R package implementing data standards and methods for taxonomic data [version 2; peer review: 4 approved]. *F1000Research* 7: 272.

Giangacomo, C., M. Mohseni, L. L. Kovar, and J. Wallace. 2021. Comparing DNA extraction and 16S rRNA gene amplification methods for plant-associated bacterial communities. *Phytobiomes Journal* 5(2): 190–201.

Glassing, A., S. E. Dowd, S. Galandiuk, B. Davis, and R. J. Chiodini. 2016. Inherent bacterial DNA contamination of extraction and sequencing reagents may affect interpretation of microbiota in low bacterial biomass samples. *Gut Pathogens* 8: 24.

Grahn, N., M. Olofsson, K. Ellnebo-Svedlund, H. J. Monstein, and J. Jonasson. 2003. Identification of mixed bacterial DNA contamination in broad-range PCR amplification of 16S rDNA V1 and V3 variable regions by pyrosequencing of cloned amplicons. *FEMS Microbiology Letters* 219: 87–91.

Gupta, N. S., and R. D. Pancost. 2006. Biomolecular and physical taphonomy of angiosperm leaf during early decay: Implications for fossilization. *Palaios* 19: 428–440.

Haro, C., M. Anguita-Maeso, M. Metsis, J. A. Navas-Cortés, and B. B. Landa. 2021. Evaluation of established methods for DNA extraction and primer pairs targeting 16S rRNA gene for bacterial microbiome profiling of olive xylem sap. *Frontiers in Plant Science* 12: 640829.

Hempel, M., H. P. Grossart, and E. M. Gross. 2009. Community composition of bacterial biofilms on two submerged macrophytes and an artificial substrate in a pre-alpine lake. *Aquatic Microbial Ecology* 58: 79–94.

Hughes, R., Z. Alkan, N. L. Keim, and M. E. Kable. 2017. Impact of sequence variant detection and bacterial DNA extraction methods on the measurement of microbial community composition in human stool. *BioRxiv* 212134 [Preprint] [posted 1 November 2017]. Available at https://doi.org/10.1101/212134 [accessed 29 July 2021].

Iniesto, M., A. I. Lopez-Archilla, M. Fregenal-Martínez, A. D. Buscalioni, and M. C. Guerrero. 2013. Involvement of microbial mats in delayed decay: An experimental essay on fish preservation. *Palaios* 28: 56–66.

Iniesto, M., I. Villalba, A. D. Buscalioni, M. C. Guerrerom, and A. I. López-Archilla. 2017. The effect of microbial mats in the decay of anurans with implications for understanding taphonomic processes in the fossil record. *Scientific Reports* 7: 45160.

Iniesto, M., C. Blanco-Moreno, A. Villabla, A. D. Buscalioni, M. C. Guerrero, and A. I. López-Archilla. 2018. Plant tissue decay in long-term experiments with microbial mats. *Geosciences* 8: 387.

Jackrel, S. L., S. M. Owens, J. A. Gilbert, and C. A. Pfister. 2017. Identifying the plant-associated microbiome across aquatic and terrestrial environments: The effects of amplification method on taxa discovery. *Molecular Ecology Resources* 17: 931–942.

Jamal, M., W. Ahmad, S. Andleeb, F. Jalil, M. Imran, M. A. Nawaz, T. Hussain, et al. 2018. Bacterial biofilm and associated infections. *Journal of the Chinese Medical Association* 81: 7–11.

Kim, M., M. Morrison, and Z. Yu. 2011. Evaluation of different partial 16S rRNA gene sequence regions for phylogenetic analysis of microbiomes. *Journal of Microbiological Methods* 84: 81–87.

Klymiuk, A. A. 2018. Microbiological insights into ecology and taphonomy of prehistoric wetlands. Doctoral dissertation, University of Kansas, Lawrence, Kansas, USA.

Lee, S. Y., and Y. B. Eom. 2016. Analysis of microbial composition associated with freshwater and seawater. *Biomedical Science Letters* 22: 150–159.

Li, K., X. X. Yuan, H. M. Sun, L. S. Zhao, R. Tang, Z. H. Chen, Y.-Z. Zhang, and H. N. Su. 2018. Atomic force microscopy of side wall and septa peptidoglycan from *Bacillus subtilis* reveals an architectural remodeling during growth. *Frontiers in Microbiology* 9: 620.

Liang, Y. L., Z. Zhang, M. Wu, Y. Wu, and J. X. Feng. 2014. Isolation, screening, and identification of cellulolytic bacteria from natural reserves in the subtropical region of China and optimization of cellulase production by *Paenibacillus terrae* ME27-1. *BioMed Research International* 2014: 512497.

Liu, Z.-P., B.-J. Wang, Y.-H. Liu, and S.-J. Liu. 2005. *Novosphingobium taihuense* sp. nov., a novel aromatic-compound-degrading bacterium isolated from Taihu Lake, China. *International Journal of Systematic and Evolutionary Microbiology* 55: 1229–1232.

Locatelli, E. R. 2014. The exceptional preservation of plant fossils: A review of taphonomic pathways and biases in the fossil record. *In* M. Laflamme, J. D. Schiffbauer, and S. A. F. Darroch [eds.], *Reading and writing of the fossil record: Preservational pathways to exceptional fossilization. The Paleontological Society Papers 20*: 237–257.

Lutz, K. A., W. Wang, A. Zdepski, and T. P. Michael. 2011. Isolation and analysis of high quality nuclear DNA with reduced organellar DNA for plant genome sequencing and resequencing. *BMC Biotechnology* 11: 54.

Ma, Y., S. Huang, Z. Gan, Y. Xiong, R. Cai, Y. Liu, L. Wu, and G. Ge. 2020a. The succession of bacterial and fungal communities during decomposition of two hygrophytes in a freshwater lake wetland. *Ecosphere* 11: e03242.

Ma, Z. Y., X. M. Zhang, R. Wang, M. Wang, T. Liu, and Z. L. Tan. 2020b. Effects of chemical and mechanical lysis on microbial DNA yield, integrity, and downstream amplicon sequencing of rumen bacteria and protozoa. *Frontiers in Microbiology* 11: 581227.

Mah, T. F. C., and G. A. O'Toole. 2001. Mechanisms of biofilm resistance to antimicrobial agents. *Trends in Microbiology* 9: 34–39.

Martin, D., D. E. Briggs, and R. J. Parkes. 2005. Decay and mineralization of invertebrate eggs. *Palaios* 20: 562–572.

McAlister, M. B., L. A. Kulakov, J. F. O'Hanlon, M. J. Larkin, and K. L Ogden. 2002. Survival and nutritional requirements of three bacteria isolated from ultrapure water. *Journal of Industrial Microbiology and Biotechnology* 29: 75–82.

McDonald, D., M. N. Price, J. Goodrich, E. P. Nawrocki, T. Z. Desantis, A. Probst, G. L. Andersen, et al. 2012. An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *ISME Journal* 6: 610–618.

McFeters, G. A., S. C. Broadaway, B. H. Pyle, and Y. Egozy. 1993. Distribution of bacteria within operating laboratory water purification systems. *Applied and Environmental Microbiology* 59: 1410–1415.

McGovern, E., S. M. Waters, G. Blackshields, and M. S. McCabe. 2018. Evaluating established methods for rumens 16S rRNA amplicon sequencing with mock microbial populations. *Frontiers in Microbiology* 9: 1365.

Mechesso, A. F., Q. Yixian, and S. C. Park. 2019. Methyl gallate and tylosin synergistically reduce the membrane integrity and intracellular survival of *Salmonella typhimurium*. *PLoS ONE* 14: e0221386.

Mentes, A., A. Szabó, B. Somogyi, B. Vajna, N. Tugyi, B. Csitári, L. Vörös, and T. Felföldi. 2018. Differences in planktonic microbial communities associated with three types of macrophyte stands in a shallow lake. *FEMS Microbiology Ecology* 94: fix164.

Morgan, X. C., and C. Huttenhower. 2012. Human microbiome analysis. *PLoS Computational Biology* 8: e1002808.

Neefs, J. M., Y. Van de Peer, P. De Rijk, S. Chapelle, S., and R. De Wachter. 1993. Compilation of small ribosomal subunit RNA structures. *Nucleic Acids Research* 21: 3025–3049.

Ojo-Okunola, A., S. Claassen-Weitz, K. S. Mwaikono, S. Gardner-Lubbe, H. J. Zar, M. P. Nicol, and E. du Toit. 2020. The influence of DNA extraction and lipid removal on human milk bacterial profiles. *Methods and Protocols* 3: 39. https://doi.org/10.3390/mps3020039

Oksanen, J., F. G. Blanchet, M. Friendly, R. Kindt, P. Legendre, D. McGlinn, P. R. Minchin, et al. 2019. vegan: Community Ecology Package. R package version 2.56. Website: https://CRAN.R-project.org/package=vegan [accessed 28 June 2021].

O'Toole, G., H. B. Kaplan, and R. Kolter. 2000. Biofilm formation as microbial development. *Annual Review of Microbiology* 54: 49–79.

Parimala, M., and F. G. Shoba. 2014. In vitro antimicrobial activity and HPTLC analysis of hydroalcoholic seed extract of *Nymphaea nouchali* Burm. f. *BMC Complementary and Alternative Medicine* 14: 361.

Peng, Z., X. Zhu, Z. Wang, X. Yan, G. Wang, M. Tang, A. Jiang, and K. Kristiansen. 2020. Comparative analysis of sample extraction and library construction for shotgun metagenomics. *Bioinformatics and Biology Insights* 14. https://doi.org/10.1177/1177932220915459

Pollock, J., L. Glendinning, T. Wisedchanwet, and M. Watson. 2018. The madness of microbiome: Attempting to find consensus "best practice" for 16S microbiome studies. *Applied and Environmental Microbiology* 84. https://doi.org/10.1128/AEM.02627-17

Qiu, Z., J. Wang, M. Delgado-Baquerizo, P. Trivedi, E. Egidi, Y.-M. Chen, H. Zhang, and B. K. Singh. 2020. Plant microbiomes: Do different preservation approaches and primer sets alter our capacity to assess microbial diversity and community composition? *Frontiers in Plant Science* 11: 993.

Quince, C., A. W. Walker, J. T. Simpson, N. J. Loman, and N. Segata. 2017. Shotgun metagenomics, from sampling to analysis. *Nature Biotechnology* 35: 833–844.

R Core Team. 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Website: http://www.r-project.org/ [accessed 28 June 2021].

Raff, E. C., M. E. Andrews, F. R. Turner, E. Toh, D. E. Nelson, and R. A. Raff. 2013. Contingent interactions among biofilm-forming bacteria determine preservation or decay in the first steps toward fossilization of marine embryos. *Evolution & Development* 15: 243–256.

Raff, R. A., and E. C. Raff. 2014. The role of biology in the fossilization of embryos and other soft-bodied organisms: Microbial films and Lagerstätten. In M. Laflamme, J. D. Schiffbauer, and S. F. Darroch [eds.], Reading and writing of the fossil record: Preservational pathways to exceptional fossilization. The Palaeontological *Society Papers* 20: 83–100.

Raff, R. A., M. E. Andrews, R. L. Pearson, F. R. Turner, S. T. Saur, D. C. Thomas, J. L. Eagan, and E. C. Raff. 2014. Microbial ecology and biofilms in the taphonomy of soft tissues. *Palaios* 29: 560–569.

Ramey, B. E., M. Koutsoudis, S. B. von Bodman, and C. Fuqua. 2004. Biofilm formation in plant–microbe associations. *Current Opinion in Microbiology* 7: 602–609.

Reitner, J. 2011. Biofilms and fossilization. *In* J. Reitner and V. Thiel [eds.], Encyclopedia of geobiology, 136–137. Springer Netherlands, Dordrecht, The Netherlands.

Roberson, E. B., and M. K. Firestone. 1992. Relationship between desiccation and exopolysaccharide production in a soil *Pseudomonas* sp. *Applied and Environmental Microbiology* 58: 1284–1291.

Rocha, J., and C. M. Manaia. 2020. Cell-based internal standard for qPCR determinations of antibiotic resistance indicators in environmental water samples. *Ecological Indicators* 113: 106194.

Sagan, L. 1967. On the origin of mitosing cells. *Journal of Theoretical Biology* 14: 255–274.

Sakai, M., A. Matsuka, T. Komura, and S. Kanazawa. 2004. Application of a new PCR primer for terminal restriction fragment length polymorphism analysis of the bacterial communities in plant roots. *Journal of Microbiological Methods* 59: 81–89.

Salter, S. J., M. J. Cox, E. M. Turek, S. T. Calus, W. O. Cookson, M. F. Moffatt, P. Turner, et al. 2014. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biology* 12: 87.

Sare, A. R., G. Stouvenakers, M. Eck, A. Lampens, S. Goormachtig, M. H. Jijakli, and S. Massart. 2020. Standardization of plant microbiome studies: Which proportion of the microbiota is really harvested? *Microorganisms* 8: 342.

Sehnal, L., E. Brammer-Robbins, A. M. Wormington, L. Blaha, J. Bisesi, I. Larkin, C. J. Martyniuk, et al. 2021. Microbiome composition and function in aquatic vertebrates: Small organisms making big impacts on aquatic animal health. *Frontiers in Microbiology* 12: 567408.

Sharma, J. K., R. K. Gautam, S. V. Nanekar, R. Weber, B. K. Singh, S. K. Singh, and A. A. Juwarkar. 2017. Advances and perspective in bioremediation of polychlorinated biphenyl-contaminated soils. *Environmental Science and Pollution Research* 25: 16355–16375.

Shen, H., S. Rogelj, and T. L. Kieft. 2006. Sensitive, real-time PCR detects low-levels of contamination by *Legionella pneumophila* in commercial reagents. *Molecular and Cellular Probes* 20: 147–153.

Sohn, J. H., K. K. Kwon, J.-H. Kang, H.-B. Jung, and S.-J. Kim. 2004. *Novosphingobium pentaromativorans* sp. nov., a high-molecular-mass polycyclic aromatic hydrocarbon-degrading bacterium isolated from estuarine sediment. *International Journal of Systematic and Evolutionary Microbiology* 54: 1438–1487.

Song, F., J. V. Kuehl, A. Chandran, and A. P. Arkin. 2021. A simple, cost-effective and automation-friendly direct PCR approach for bacterial community analysis. *BioRxiv* 433496 [Preprint] [posted 2 March 2021]. Available at https://doi.org/10.1101/2021.03.01.433496 [accessed 29 July 2021].

Spicer, R. A. 1975. The sorting of plant remains in a recent depositional environment. Ph.D. thesis, University of London, London, United Kingdom.

Spicer, R. A. 1977. The pre-depositional formation of some leaf impressions. *Palaeontology* 20: 907–912.

Spicer, R. A. 1991. Plant taphonomic processes. *In* P. A. Allison and B. E. G. Briggs [eds.], Taphonomy: Releasing the data locked in the fossil record, 72–113. Plenum, New York, New York, USA.

Stewart, E. J. 2012. Growing unculturable bacteria. *Journal of Bacteriology* 194: 4151–4160.

Sutherland, I. W. 2001. Biofilm exopolysaccharides: A strong and sticky framework. *Microbiology* 147: 3–9.

Thompson, L. R., J. G. Sanders, D. McDonald, A. Amir, J. Ladau, K. J. Locey, R. J. Prill, et al. 2017. A communal catalogue reveals Earth's multiscale microbial diversity. *Nature* 551: 457–463.

Tláskal, V., J. Voříšková, and P. Baldrian. 2016. Bacterial succession on decomposing leaf litter exhibits a specific occurrence pattern of cellulolytic taxa and potential decomposers of fungal mycelia. *FEMS Microbiology Ecology 92: fiw177*.

Trevathan-Tackett, S. M., T. R. Allnutt, C. D. Sherman, M. F. Richardson, T. M. Crowley, and P. I. Macreadie. 2020. Spatial variation of bacterial and fungal communities of estuarine seagrass leaf microbiomes. *Aquatic Microbial Ecology* 84: 59–74.

Vartoukian, S. R., R. M. Palmer, and W. G. Wade. 2010. Strategies for culture of 'unculturable' bacteria. *FEMS Microbiology Letters* 309: 1–7.

Ward, T., J. Larson, J. Meulemans, B. Hillmann, J. Lynch, D. Sidiropoulos, J. R. Spear, et al. 2017. BugBase predicts organism-level microbiome phenotypes. *BioRxiv* 133462 [Preprint] [posted 2 May 2017]. Available at https://doi.org/10.1101/133462 [accessed 29 July 2021].

Wickham, H. 2016. ggplot2: Elegant graphics for data analysis. Springer-Verlag, New York, New York, USA.

Xian, L., T. Wan, Y. Cao, J. Sun, T. Wu, A. A. Apudo, W. Li, and F. Liu. 2020. Structural variability and functional prediction in the epiphytic bacteria assemblies of *Myriophyllum spicatum*. *Current Microbiology* 77: 3582–3594.

Xue, M., L. Wu, Y. He, H. Liang, and C. Wen. 2018. Biases during DNA extraction affect characterization of the microbiota associated with larvae of the Pacific white shrimp, Litopenaeus vannamei. *PeerJ* 6: e5257.

Yang, F., J. Sun, H. Luo, H. Ren, H. Zhou, Y. Lin, M. Han, et al. 2020. Assessment of fecal DNA extraction protocols for metagenomic studies. *GigaScience* 9: giaa071.

Yuan, S., D. B. Cohen, J. Ravel, Z. Abdo, and L. J. Forney. 2012. Evaluation of methods for the extraction and purification of DNA from the human microbiome. *PLoS ONE* 7: e33865.

Zeng, J., Y. Bian, P. Xing, and Q. L. Wu. 2012. Macrophyte species drive the variation of bacterioplankton community composition in a shallow freshwater lake. *Applied and Environmental Microbiology* 78: 177–184.

Zhang, J., H. Hou, G. Chen, S. Wang, and J. Zhang. 2016. The isolation and functional identification on producing cellulase of *Pseudomonas mendocina*. *Bioengineered* 7: 382–391.

Zhao, B., P. Xing, and Q. L. Wu. 2017. Microbes participated in macrophyte leaf litters decomposition in freshwater habitat. *FEMS Microbiology Ecology* 93: fix108.

How to cite this article: Janssen, K., S. L. Low, Y. Wang, Q.-Y. Mu, G. Bierbaum, and C. T. Gee. 2021. Elucidating biofilm diversity on water lily leaves through 16S rRNA amplicon analysis: Comparison of four DNA extraction kits. *Applications in Plant Sciences* 9(8): e11444. https://doi.org/10.1002/aps3.11444

APPENDIX 1: Overview of sample composition after modeling and correction of sequenced amplicon errors with the DADA2 pipeline

SampleID	Input	Filtered	Percentage of input passed filter	Denoised	Merged	Percentage of input merged	Non-chimeric	Percentage of input non-chimeric	Kit
FL-Kit1-1	462,855	434,320	93.84	430,344	413,824	89.41	393,213	84.95	Kit 1
FL-Kit1-2	528,277	493,610	93.44	490,190	476,127	90.13	461,367	87.33	Kit 1
FL-Kit1-3	376,857	355,347	94.29	347,172	298,515	79.21	187,077	49.64	Kit 1
FL-Kit2-1	447,247	426,316	95.32	421,229	383,919	85.84	298,026	66.64	Kit 2
FL-Kit2-2	642,817	601,550	93.58	595,347	551,978	85.87	391,214	60.86	Kit 2
FL-Kit2-3	478,479	441,328	92.24	436,412	395,641	82.69	274,475	57.36	Kit 2
FL-Kit3-1	773,274	730,187	94.43	724,400	690,366	89.28	633,094	81.87	Kit 3
FL-Kit3-2	743,584	707,421	95.14	699,648	644,819	86.72	427,578	57.5	Kit 3
FL-Kit3-3	302,487	276,281	91.34	275,059	261,517	86.46	216,422	71.55	Kit 3
FL-Kit4-1	670,347	636,629	94.97	628,012	568,381	84.79	327,862	48.91	Kit 4
FL-Kit4-2	457,728	433,196	94.64	428,492	387,280	84.61	221,897	48.48	Kit 4
FL-Kit4-3	649,033	596,513	91.91	587,461	518,698	79.92	341,830	52.67	Kit 4
SL-Kit1-1	448,591	424,648	94.66	418,948	382,451	85.26	252,280	56.24	Kit 1
SL-Kit1-2	1,030,513	981,661	95.26	972,002	893,577	86.71	486,500	47.21	Kit 1
SL-Kit1-3	657,554	612,346	93.12	605,041	558,937	85	370,825	56.39	Kit 1
SL-Kit2-1	520,130	490,836	94.37	487,321	461,120	88.65	314,964	60.55	Kit 2
SL-Kit2-2	452,533	430,019	95.02	425,900	398,000	87.95	253,985	56.13	Kit 2
SL-Kit2-3	456,990	434,433	95.06	431,584	409,477	89.6	281,342	61.56	Kit 2
SL-Kit3-1	347,913	321,721	92.47	319,850	301,923	86.78	223,265	64.17	Kit 3
SL-Kit3-2	408,249	383,920	94.04	380,430	350,360	85.82	197,062	48.27	Kit 3
SL-Kit3-3	302,389	215,615	71.3	215,093	205,745	68.04	168,903	55.86	Kit 3
SL-Kit4-1	586,523	554,452	94.53	547,809	506,870	86.42	382,642	65.24	Kit 4
SL-Kit4-2	476,451	453,923	95.27	446,063	388,815	81.61	228,346	47.93	Kit 4
SL-Kit4-3	180,539	170,024	94.18	167,418	153,372	84.95	111,359	61.68	Kit 4
MMC-Kit1-1	519,190	492,252	94.81	490,503	466,449	89.84	261,951	50.45	Kit 1
MMC-Kit1-2	603,379	560,103	92.83	558,085	525,034	87.02	261,228	43.29	Kit 1
MMC-Kit2-1	604,482	551,650	91.26	549,673	512,429	84.77	239,385	39.6	Kit 2
MMC-Kit2-2	620,411	591,397	95.32	589,700	561,243	90.46	286,449	46.17	Kit 2
MMC-Kit3-1	601,753	551,892	91.71	550,067	514,273	85.46	254,461	42.29	Kit 3
MMC-Kit3-2	781,894	740,791	94.74	738,614	699,270	89.43	385,380	49.29	Kit 3
MMC-Kit4-1	581,630	534,020	91.81	532,439	502,901	86.46	268,491	46.16	Kit 4
MMC-Kit4-2	515,250	488,617	94.83	486,683	455,036	88.31	178,495	34.64	Kit 4