



Food or just a free ride? A meta-analysis reveals the global diversity of the Plastisphere

Robyn J. Wright ^{1,2} · Morgan G. I. Langille^{2,3} · Tony R. Walker ¹

Received: 4 June 2020 / Revised: 13 October 2020 / Accepted: 15 October 2020 / Published online: 2 November 2020
© The Author(s) 2020. This article is published with open access

Abstract

It is now indisputable that plastics are ubiquitous and problematic in ecosystems globally. Many suggestions have been made about the role that biofilms colonizing plastics in the environment—termed the “Plastisphere”—may play in the transportation and ecological impact of these plastics. By collecting and re-analyzing all raw 16S rRNA gene sequencing and metadata from 2,229 samples within 35 studies, we have performed the first meta-analysis of the Plastisphere in marine, freshwater, other aquatic (e.g., brackish or aquaculture) and terrestrial environments. We show that random forest models can be trained to differentiate between groupings of environmental factors as well as aspects of study design, but—crucially—also between plastics when compared with control biofilms and between different plastic types and community successional stages. Our meta-analysis confirms that potentially biodegrading Plastisphere members, the hydrocarbonoclastic *Oceanospirillales* and *Alteromonadales* are consistently more abundant in plastic than control biofilm samples across multiple studies and environments. This indicates the predilection of these organisms for plastics and confirms the urgent need for their ability to biodegrade plastics to be comprehensively tested. We also identified key knowledge gaps that should be addressed by future studies.

Introduction

An estimated total of 7 billion metric tons of plastic waste has now been produced globally while approximately only 9% of this is recycled and 79% has been discarded in landfills or the environment [1]. In marine and other aquatic environments, plastics cause a range of negative environmental impacts: directly, through entanglement [2, 3] or ingestion [4–8], or indirectly, through the transfer of toxic chemicals [9, 10]. Despite that the majority of marine

plastic waste originates from land, less is known about the impacts of plastic waste in terrestrial ecosystems [11], and there are numerous factors that can affect the fate, transport, and impacts of plastics in all biospheres.

These factors include transport by air [12], rain [13], rivers [14] and currents [15], (de)sorption of chemicals [16], photo- or mechanical degradation or fragmentation [17] and microbial colonization and possibly degradation [18]. Deleterious impacts of plastics [6], their potential to transport invasive species across entire ocean basins [3, 19–21] and their biodegradation by isolated bacteria and fungi have been studied for over half a century [22]. The microbial communities colonizing plastics—commonly termed the “Plastisphere” [23]—however, have only been specifically investigated more recently. A call for research into the interactions between microorganisms and plastics at the beginning of 2011 marks almost a decade of Plastisphere research [24]. On 5 January, 2020 there were 50 publications that studied the Plastisphere using Illumina Next Generation Sequencing (NGS) methods. Although this is an emerging area of research, most Plastisphere studies have been focussed on the marine environment.

Microbial members of the Plastisphere have been found to be: (i) different from communities that colonize other

Supplementary information The online version of this article (<https://doi.org/10.1038/s41396-020-00814-9>) contains supplementary material, which is available to authorized users.

✉ Robyn J. Wright
robyn.wright@dal.ca

¹ School for Resource and Environmental Studies, Dalhousie University, Halifax, Canada

² Department of Pharmacology, Faculty of Medicine, Dalhousie University, Halifax, Canada

³ Department of Microbiology and Immunology, Dalhousie University, Halifax, Canada

surfaces [25–27]; (ii) not different from communities that colonize other surfaces [28]; (iii) only different from communities colonizing other surfaces under specific environmental conditions [29, 30] or at specific time points [31]; (iv) more diverse than other microbial communities [32]; (v) less diverse than other microbial communities [33, 34]; (vi) potentially degrading the plastics that they colonize [35, 36]; (vii) capable of degrading plastic additives [33, 37, 38]; and (viii) pathogenic and/or carrying antimicrobial resistance genes [39–42]. The marine Plasticsphere has been heavily reviewed within the last year, e.g., [18, 43–46], and there are also several recent reviews on plastic biodegradation, e.g., [47–49]. However, definitive answers on the metabolic capabilities of the Plasticsphere or the factors that drive its formation and composition are unknown. Indeed, a re-analysis of a small subset of Plasticsphere studies ($n = 5$) by Oberbeckmann and Labrenz [45] revealed that salinity along with other environmental factors appeared to have a larger effect on community composition than substrate.

To date, no large-scale meta-analysis of the Plasticsphere has been conducted, despite that all NGS data collected are theoretically comparable. This is likely because these data are not always deposited in publicly accessible databases (as best practice would dictate). In addition, the use of different methods for processing these data means that direct comparisons have not been possible without substantial bioinformatic analyses. In this study, we conducted a re-analysis of all studies from 2010 to 2019 (for which sequencing data were already accessible or were made available upon request) that use Illumina NGS to characterize the Plasticsphere using the 16S rRNA gene. We aimed to determine whether taxa identified as potential plastic biodegraders or potential pathogens -and that were higher in abundance on plastics than other samples- were significant across multiple studies and environments. We then investigated which environmental and methodological factors were shaping the Plasticsphere. We classified all sequences from these studies to amplicon sequence variants (ASVs) and used a phylogeny-based method [50] to overcome the problems presented by the use of different primer pairs. This allowed us to identify the common taxa between these studies and to use random forest models to draw conclusions on the overarching factors that shape the Plasticsphere.

Materials and methods

Experimental design

A literature search was performed on 5 January, 2020 using the search terms “Plastics plasticsphere”, “Plastics microbial community”, and “Plastics microbial degradation” in both the Web of Science Core Collection and Science Direct

(Supplementary Table 1A). The search was limited to studies that fit the following criteria: (i) were published between 2010 and 2019; (ii) had original data (iii) characterize the biofilm formed on nonbiodegradable plastics; and (iv) use Illumina NGS. This resulted in 50 studies for inclusion in this meta-analysis (Supplementary Table S1B); 41 studies that characterized only the 16S rRNA gene [25, 27–33, 35, 38, 41, 42, 51–79], two studies that characterized only the 18S rRNA gene [34, 80], two studies that characterized the 16S and 18S rRNA genes [26, 81], two studies that characterized the 16S and ITS2 rRNA genes [82, 83], two studies that used shotgun metagenomics [84, 85], and one study that used shotgun metagenomics and characterized the 16S rRNA gene [86].

Of these studies, 34 had sequencing data that were already publicly accessible. Requests for raw sequencing data and metadata were made to the corresponding and first authors of the remaining 16 studies, resulting in the provision of these data for a further five studies. This resulted in 39 studies with datasets that were available for inclusion: 31 that characterized only the 16S rRNA gene [25, 27, 28, 30–33, 35, 38, 52, 53, 55–59, 61, 62, 64–66, 68, 69, 71–76, 78, 79], two that characterized only the 18S rRNA gene [34, 80], two that characterized the 16S and ITS2 rRNA genes [82, 83], two that used shotgun metagenomics [84, 85], one that characterized the 16S and 18S rRNA genes [26], and one that used shotgun metagenomics and characterized the 16S rRNA gene [86] (Supplementary Table S1B). Relatively few studies focussed on the 18S and ITS2 rRNA genes and these used primers that targeted different regions and were from different environments. We therefore focussed on the 16S rRNA gene here, meaning that we included a total of 35 studies [25, 27, 28, 30–33, 35, 38, 52, 53, 55–59, 61, 62, 64–66, 68, 69, 71–76, 78, 79, 82, 83, 86] with 2,229 samples between them. Publicly accessible data containing sequencing reads for the remaining studies were downloaded (primarily from the NCBI SRA database) and, if necessary, files were converted to match a format that was compatible with QIIME2 (i.e., GNU zipped FASTQ files). Additional requests for raw sequencing data were made to the authors of studies where the forward and reverse reads were already joined, or the primers were already removed (full details can be found in Supplementary Table S1B). All available metadata were collected and supplemented with any additional information present in the supplementary information of published studies (Supplementary Table S2). Where details of salinity and temperature were not given, these were estimated based on typical characteristics in these areas at the time of year the samples were collected. If no light regime was specified, ambient light was assumed. If sample names were not given that could be matched between supporting information of the paper and metadata given to the NCBI SRA or plastic

type was not determined, these were classified as “unknown plastic”.

We added metadata categories that were not in any of the original studies: (i) environment—each study was classified as terrestrial, marine, freshwater, or aquatic (e.g., the environment was brackish or the experiment was carried out in a different system, such as in mariculture); (ii) whether the study was carried out in a laboratory or in the field; (iii) incubation or collection—whether the plastics used were incubated for a known length of time or were collected after an unknown residence time; (iv) general incubation time—samples were classified as early (≤ 7 days incubation), late (> 7 days incubation), or collection (samples were collected after an unknown residence time); (v) water or sediment—the plastics were from/were incubated either in the water column or on the sea floor/in soil in the terrestrial environment; (vi) source—the material used for microbial community characterization was classified to the general categories of plastic, water, sediment, organic, not plastic (i.e., an inert control surface, such as glass or metal), other or blanks and positive sequencing or methodological controls; and (vii) general plastic type—samples were classified as aliphatic (i.e., PE or PP), other plastic (i.e., plastics that contain other functional groups, e.g., PET, PS, and PVC), unknown plastic (the plastic type was not determined), biofilm (the sample was from a control substrate, such as glass or a leaf), planktonic, or blank (i.e., sequencing or methodological controls).

16S amplicon sequence processing (per study)

Data processing followed the standard operating procedure suggested in Comeau et al. [87] and used QIIME2 (2019.10 core distribution [88]). Raw forward and reverse read files were imported to the QIIME2 format after an initial visualization of read quality using the packages FastQC (Babraham Bioinformatics) and MultiQC [89]. Cutadapt [90] was used to remove primers from reads and VSEARCH [91] was used to join paired-end reads. These steps were omitted for samples where primers were already removed or reads were already joined, respectively, and for one study that used three different reverse primers [59]. All low-quality reads were then filtered using default quality thresholds before using Deblur [92] to denoise sequences and resolve ASVs. Deblur was run using trim lengths determined by the read quality for each study. For Frère et al. [59], the ends of reads corresponding to the lengths of primer sequences were also trimmed. The forward reads only were used for several studies where low reverse read quality led to too few reads remaining after running Deblur and for which the forward reads were of high enough quality to be used (full details of the processing steps carried out for each study can be found in Supplementary Table S2).

Combined processing

All studies were merged using QIIME2's merge and merge-seqs commands, then classified taxonomically using a classifier trained on the full-length 16S rRNA gene SILVA v132 database [93]. Classified sequences were filtered to remove mitochondria, chloroplasts, those that were unclassified at the kingdom level and those present at a cumulative abundance of ten or fewer. This left a median of 20,284 reads per sample (minimum 2 and maximum 995,391). Samples with $< 2,000$ reads were removed, leaving 2,056 samples and 34 studies, and phylogenetic trees were built using SEPP with a reference phylogeny created using the SILVA v128 database [50].

Custom scripts that wrapped all commands were used to carry out further analyses in R (version 3.6.1) and Python (version 3.8.3) using the data exported from QIIME2 and the packages Biopython [94], csv, itertools, lifelines [95], math, matplotlib [96], numpy, os, pandas [97], pdf2image, pickle, scipy [98], scikit-bio, scikit-learn [99], and sinfo for Python as well as ape [100], compositions [101], dplyr [102], exactRankTests [103], ggnewscale [104], ggplot2 [105], ggtree [106], knitr [107], metacoder [108], microbiome [109], nlme [110], phylr [111], phyloseq [112], reticulate [113], and vegan [114] for R. Several normalization methods were used to address the large disparity in the sequencing depths of different samples: (i) samples were rarefied to 2,000 and converted to relative abundance; (ii) samples were converted to relative abundance; (iii) samples were converted to a log scale (with a pseudo count of one); or (iv) samples were converted to a centered log ratio (CLR; with a pseudo count of half of the minimum nonzero count). ASVs with a maximum number of reads below 1% of the median number of reads per sample were removed—below 20 reads for (i) and 202 reads for (ii), (iii), and (iv)—and sequences were agglomerated at a height of 0.1 based on the SEPP insertion tree. This resulted in 4,469 ASVs remaining for (i) and 12,635 ASVs remaining for (ii), (iii), and (iv). For (i), (ii), and (iii), weighted and unweighted uniFrac [115] distances were calculated between all samples while for (iv), the log-ratio transformed data, Aitchison distances (i.e., Euclidean distances of CLR-transformed data) [116] as well as Phylogenetic Isometric Log-Ratio Transformation [111] distances were calculated between all samples. Unless otherwise mentioned, all analyses used the rarefied data.

Random forest model construction

To determine which taxa were most associated with groupings within each metadata category (e.g., environment, plastic type, salinity, incubation time; $n = 20$ categories), feature selection was performed using random forest models (either classification or regression models for discrete or

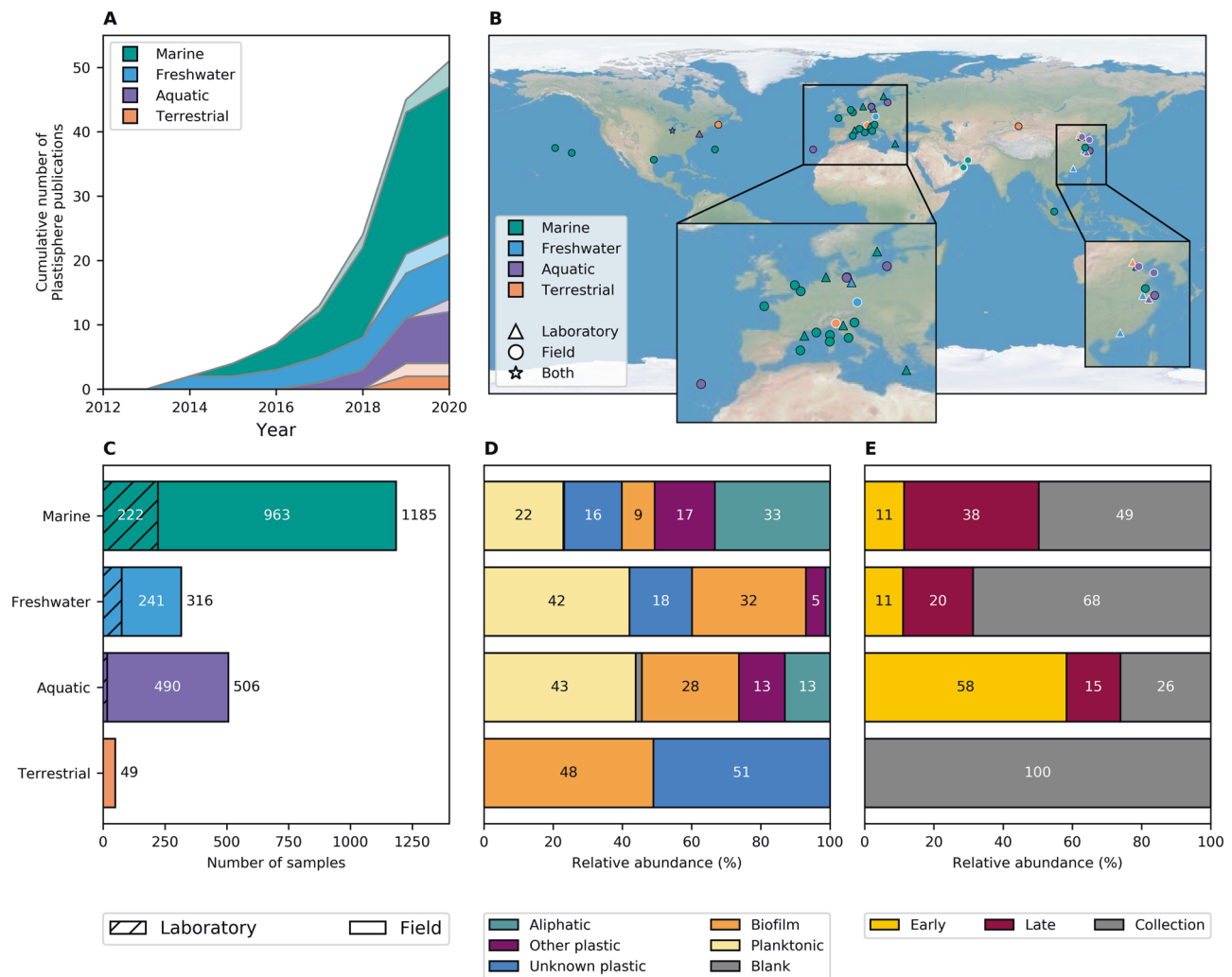


Fig. 1 Overview of the studies and samples included in the meta-analysis. Cumulative number of studies per year (A), study location (B), number of samples (C), relative abundance of sample type (D), and sample incubation time (E) for studies carried out in the marine, freshwater, other aquatic, and terrestrial environments. A, B Show all 50 studies whereas, C, D, and E show only studies/samples that were

included. Studies for which data were not provided (neither publicly available nor provided upon request) are shown with transparent colors in A and white marker edges in B. Note that those studies shown for 2020 were already in press and available online by 5 January, 2020. See Supplementary Tables S1 and S2 for full details of all studies and samples included.

continuous data, respectively) built using scikit-learn, each with 10,000 estimators and 80 and 20% of samples being used for training and testing, respectively. All taxa were scaled to the maximum value of taxon abundance prior to building the models. These were carried out separately for the taxonomic levels of phylum, class, order, family, genera, species, and ASV. Where groupings within metadata categories for samples were not known, these samples were removed from the model construction. To investigate the effect of normalization method on the selected features, these models were constructed for each of the four normalization methods described above. To determine the taxa that were most associated with different plastic types, random forest classification models (built using the same

parameters as above) were constructed separately for each environment (marine, aquatic, freshwater or terrestrial) for samples grouped to general plastic type (as described above). Samples were not further grouped into laboratory and field studies because not all environments included both. These were again constructed for the taxonomic levels of phylum, class, order, family, genus, species, and ASV and for each of the four normalization methods described above. The classification accuracy for each random forest model is defined as the percentage of the time that the model can classify a sample to the correct grouping (e.g., as marine within the environment metadata category) and the feature importance is defined as the proportion that the classification accuracy would decrease without that taxon (feature)

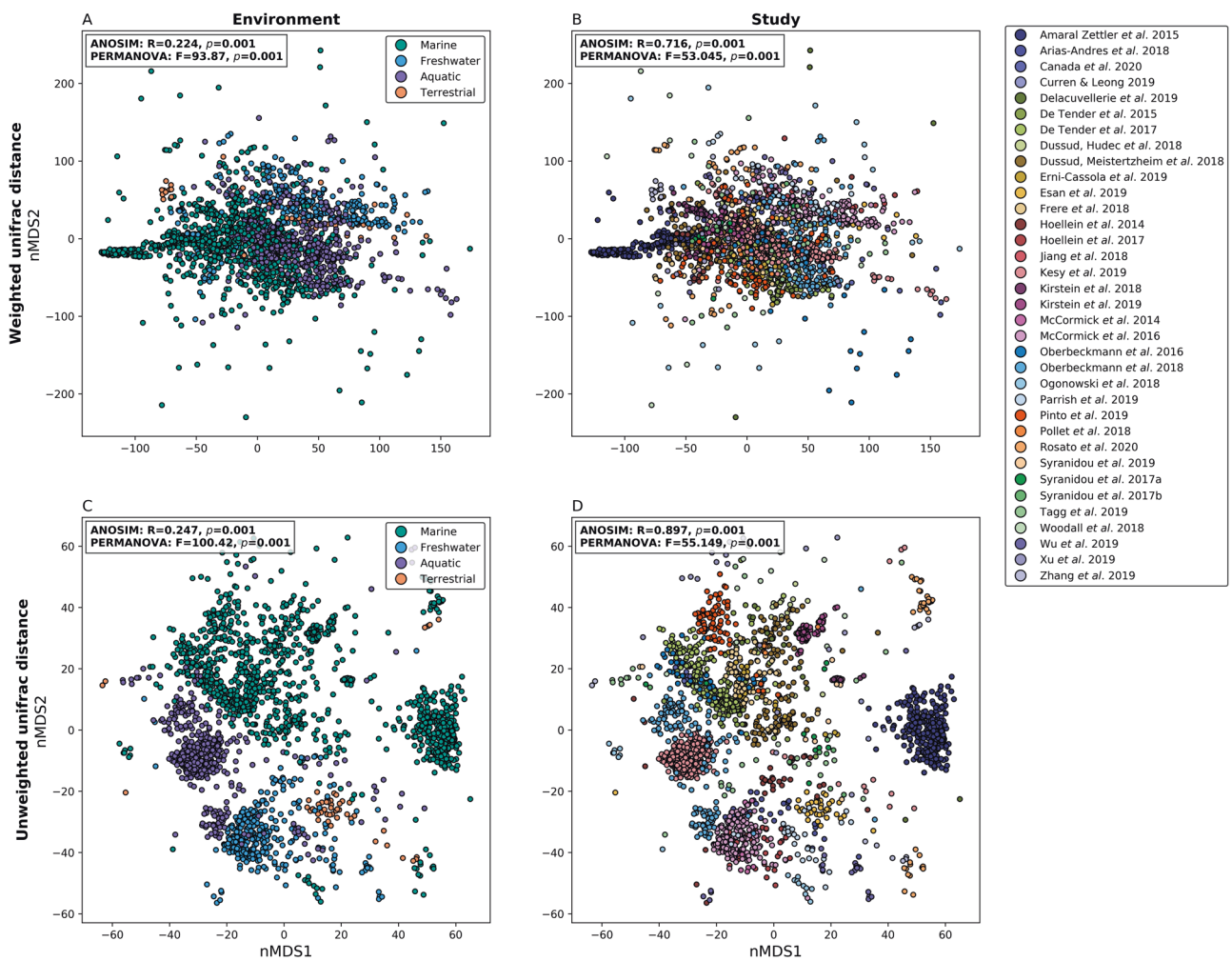


Fig. 2 nMDS plots showing uniFrac distance between samples. nMDS plots showing weighted (A, B) or unweighted (C, D) uniFrac distance (i.e., accounting for taxon phylogeny with or without taxon abundance, respectively) calculated between all samples and shown in

A and C as samples colored by environment and B and D as samples colored by study. Results of PERMANOVA and ANOSIM tests for significance between groups are shown on each plot.

present. We calculated concordance (using the Python life-lines package) [95] between the classification accuracies as well as between mean feature importance values to assess the similarity of the results obtained by the random forest models constructed using different normalization methods.

Tests for differential abundance of taxa

ASVs shared between treatments were calculated for samples grouped by the environment. Differences between samples at early (≤ 7 days) or late (> 7 days) incubation times were determined using Wilcoxon Rank Sum tests within the metacoder R package with holm-bonferroni false discovery rate correction. ANCOM tests for differential abundance (with holm-bonferroni false discovery rate correction) were performed on all taxa (separately for each taxonomic level) for the taxa normalized using rarefaction.

Results

Summary of included studies and sequences

In this study, we reanalyzed the 16S rRNA gene amplicon sequencing data obtained from 35 studies (Supplementary Section 1). All Plastisphere studies to date have examined data collected in the Northern hemisphere and most of the included studies were conducted in and around Europe (Fig. 1). After removing all samples with below 2,000 reads there were 2,056 samples remaining; 1,185, 316, 506, and 49 samples in the marine, freshwater, aquatic, and terrestrial environments, respectively. One study [64] was removed from further analyses because all samples ($n = 9$) had below 2,000 reads. The abundance of different substratum types depended on the environment and most samples were collected from the field after unknown environmental residence times (Fig. 1).

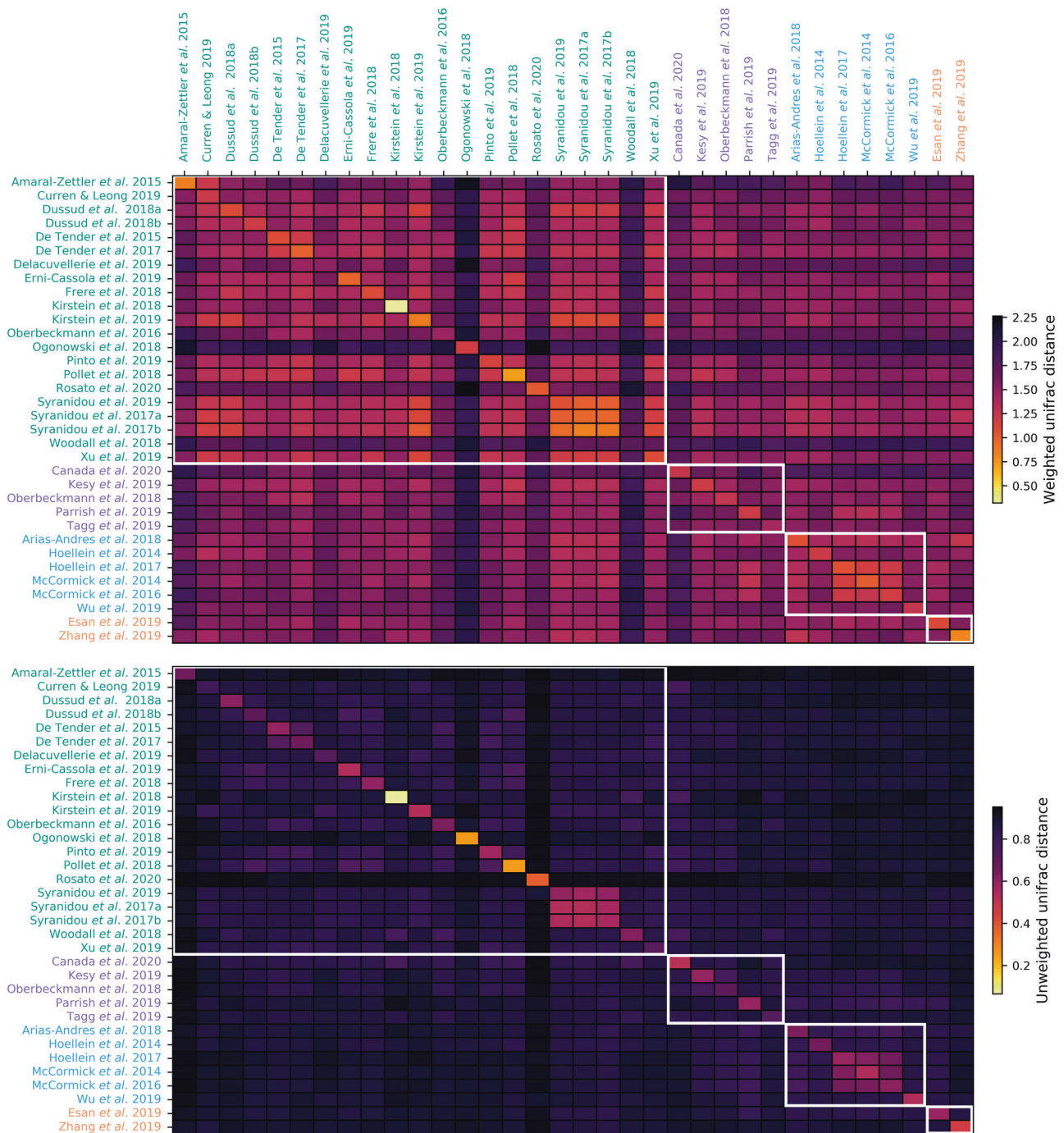


Fig. 3 Average similarity between samples within a study versus between studies. Average similarity (determined by weighted or unweighted uniFrac; i.e., accounting for taxon phylogeny with or without taxon abundance; top or bottom, respectively) between

samples within a study versus between studies, with white boxes showing samples grouped by environment. Study names are colored by environment, with green, purple, blue, and orange being for marine, aquatic, freshwater, and terrestrial, respectively (as in Figs. 1 and 2).

Normalization method affects the biological interpretation of results

There was a large disparity in the number of reads and taxonomic richness per sample between different studies (Fig. S1). We therefore investigated several different

methods for data normalization (further details are given in the methods section) on random forest classification accuracy and the concordance between taxa identified as important by each of these methods (Supplementary Section 2). These random forest models were constructed for: (i) all 20 metadata categories—including factors such as

environment, geographic location, temperature, primer pair, and plastic type to varying degrees of specificity—for each taxonomic level: phylum, class, order, family, genus, species, and ASV (i.e., 560 random forest models were constructed in total, 140 for each normalization method); and (ii) general plastic type within each of the four environments (marine, aquatic, freshwater, and terrestrial) for each taxonomic level: phylum, class, order, family, genus, species, and ASV (i.e., 112 random forest models were constructed in total, 28 for each normalization method). For both sets of random forest models, the models constructed using rarefied data were on average the least accurate (61–67% and 53–67% mean classification accuracy across all metadata categories and all environments for general plastic type, respectively, for the taxonomic levels of phylum-ASV; Figs. S2 and S3) and the models constructed on data transformed to relative abundance were on average the most accurate (82–90% and 72–86% mean classification accuracy across all metadata categories and all environments for general plastic type, respectively, for the taxonomic levels of phylum-ASV; Figs. S2 and S3). Models constructed using the compositionally aware centered log-ratio transformed data were on average 1.1 and 10% (for all metadata categories and all environments for general plastic type, respectively) less accurate than models constructed on data transformed to relative abundance. Feature importance values are the proportion that the classification accuracy decreases without that feature. The features that have importance values of either above 0.01 or 0.005 are the same between all normalization methods across all taxonomic levels. Concordance in feature importance values (whether the ranking of features by their values is the same) was on average 0.94 across all taxonomic levels between the relative abundance, log and CLR-transformed data and 0.78 for the relative abundance, log and CLR-transformed data against the rarefied data (Supplementary Section 2).

Rarefying has previously been found to most effectively account for large differences in library sizes, including lowering the false discovery rate when there are large differences in library sizes [117]. We therefore compromised on lower random forest classification accuracy and used the rarefied data for the remainder of analyses. Most analyses presented here have also been conducted on the data normalized using the other methods, but these results are presented in Supplementary Section 2.

Diversity within and between studies

Beta diversity analyses (weighted and unweighted uniFrac distance) showed that there were significant differences between samples when they were grouped by study or environment (PERMANOVA and ANOSIM $p = 0.001$; Fig. 2). The largest differences between studies and

environments were due to organisms that were present at lower abundances, as evidenced by higher pseudo-F and R statistics with the unweighted (pseudo-F = 100.42 and 55.149 and $R = 0.247$ and 0.897 for environment and study, respectively) than weighted (pseudo-F = 93.87 and 53.045 and $R = 0.224$ and 0.716 for environment and study, respectively) uniFrac distances (Figs. 2 and 3). Studies that were clearly less similar to other studies could be explained by them being laboratory-based [32, 71], focussed on anaerobic rather than aerobic communities [71], collecting samples from the deep sea [76] or sequencing amplicons that did not include the V4 16S rRNA gene region [52, 71]. Those that were particularly similar within [26] or between [72–74] studies could presumably be explained by very long incubation times (above 1 year), leading to community convergence, or having similar experimental setups and inoculums, respectively (Fig. 3).

When samples are split to environment and substratum type (i.e., the groupings shown in Fig. 1D), they have a tendency to group by environment, although this is not true for all cases (Fig. 4). The *Proteobacteria* dominate (above 50% relative abundance) in all but the terrestrial environment, freshwater planktonic, and aquatic control biofilm groups, which have larger proportions of *Actinobacteria* and *Planctomycetes*, respectively. Medians of Simpson's diversity indices were above 0.8 for all sample groups apart from the aquatic blank group (0.3), which had a high relative abundance of *Vibrionales* (i.e., above 15%) and was the only group with >1% of the extremophilic *Deinococcus-Thermus* phylum (i.e., almost 7%). The number of ASVs in each environment (only those that were >1% relative abundance are shown) was related to the number of samples from that environment, with there being the highest and lowest numbers of both ASVs and samples in the marine and terrestrial environments, respectively (Figs. 1 and 4), and each substratum type within each environment had ASVs that were unique (above 1% relative abundance in that treatment only) to it.

Environmental variables—and not plastic type—have the largest impact on microbial composition

To determine which metadata categories have the largest impact on microbial composition, we constructed random forest models for all 20 categories—including factors such as environment, geographic location, temperature, primer pair, and plastic type to varying degrees of specificity—for each taxonomic level: phylum, class, order, family, genus, species, and ASV (i.e., 140 random forest models were constructed on the rarefied data). We found that models constructed using the microbial composition at the ASV and phylum levels were on average the most (67%) and least (61%) accurate, respectively, and that the light regime

Fig. 5 Classification accuracy for random forest models.

Classification accuracy (%) for random forest models (classification or regression for discrete or continuous categories, respectively) constructed for (A) all samples grouped within different metadata categories or (B) samples within each environment grouped within plastic type (general) at different taxonomic levels. Random forest models are trained using a subset of 80% of samples (chosen randomly) and classification accuracy is based on testing using the remaining 20% of samples. Figure S4 shows the top most important features at the ASV level across all metadata categories while Supplementary Section 3 shows all taxonomic levels as well as metadata categories.

A Metadata categories							
Light	91	94	91	92	91	91	89
Laboratory or Field	90	89	91	90	91	88	89
Water or Sediment	88	87	88	89	88	88	90
Environment	77	76	86	81	86	86	84
Primer pair	75	79	85	83	83	85	83
Incubation or Collection	77	75	82	80	82	80	82
DNA extraction method	72	73	81	79	78	80	77
Incubation time (general)	70	74	74	79	80	73	77
Study	67	70	78	76	75	79	77
Longitude	63	62	67	71	70	75	72
Incubation time (specific)	66	59	47	63	56	66	77
Latitude	58	54	60	67	68	59	64
Season	54	60	59	63	64	64	61
Source	56	64	59	57	57	56	58
Salinity	46	52	54	49	57	58	60
Plastic type (specific)	43	45	42	46	48	48	48
Plastic type (general)	39	39	44	42	45	46	45
Material type	34	34	36	37	34	39	36
Temperature	31	34	38	31	36	32	40
Depth	17	7	-19	9	-15	10	21

B Plastic type (general)							
Marine	70	76	73	83	83	76	83
Freshwater	70	73	71	75	73	79	75
Terrestrial	40	60	30	50	20	60	60
Aquatic	32	41	40	39	42	44	49
	Phylum	Class	Order	Family	Genera	Species	ASV

(whether samples were incubated at ambient or modified lighting conditions, e.g., shaded or a laboratory 16:8 light: dark cycle) was the metadata category with the highest classification accuracy (maximum 94% at the class level; Fig. 5A). The other metadata categories that were successful more than 80% of the time were: (i) whether the samples came from experiments carried out in the laboratory or the field (maximum accuracy 91% at order or genera level); (ii) whether the samples were incubated/collected from sediment or the water column (maximum accuracy 90% at the ASV level); (iii) the environment that the sample came from (maximum accuracy 86% at order, genera or species level); (iv) the primer pair used for sequencing (maximum accuracy 85% at order or species levels); (v) whether the sample was collected (i.e., unknown environmental residence time) or incubated for a known length of time (maximum accuracy 82% at order, genera or ASV level); and (vi) the DNA extraction method used (maximum accuracy 81% at order level). Depth consistently produced the models with the lowest classification accuracy (maximum accuracy 21% at

the ASV level), with temperature also performing poorly (maximum accuracy 40% at the ASV level) and all other categories, such as plastic type, specific incubation time, and geographic location, having intermediate accuracy levels.

When we examine the taxa with the highest mean feature importance values (proportion that the classification accuracy decreases without that feature), we find that at all taxonomic levels besides ASV, it is a member of the *Bacteroidetes* with the highest values (maximum 0.159 at the phylum level; Fig. S4 and Supplementary Section 3); even though the *Bacteroidetes* are substantially less abundant than the *Proteobacteria* in all but the terrestrial environment (Fig. 4). The next highest feature importance values are generally from the *Proteobacteria*, and in particular the *Alphaproteobacteria* (e.g., the ASV with the highest mean value, ASV1197 *Erythrobacter*, 0.03), although the latitude and longitude usually have the highest individual importance values for both taxonomic groups. There are several metadata categories for which the taxa with high—or low—

feature importance values are strongly correlated, for example: latitude and longitude (e.g., ASV1954 *Lentimonas* and ASV1197 *Erythrobacter*); study, primer pair, and DNA extraction method (e.g., ASV0715 AEGEAN-169 marine group and ASV0496 SAR116 clade); and source, material type, plastic type and whether the sample comes from an incubation or collection experiment (e.g., ASV0808 *Sphingomonadaceae*, and ASV4333 *Alteromonas*; Fig. S4).

The Plastisphere includes potential plastic biodegraders and potentially pathogenic species

To determine the taxa that were potentially specific to plastics, samples were grouped by the environment they originated from and random forest models were trained on general plastic type (as in Fig. 1D). Random forest classification models showed that splitting samples by environment improved the classification accuracy of the plastic type metadata category (to a maximum of 83% at the family, genus or ASV level in the marine environment) and the accuracy remained at below 50% at all taxonomic levels in only the aquatic environment (Fig. 5B and Supplementary Section 3). We were able to identify taxa that were significantly differentially abundant between substratum types in all environments, although in the terrestrial environment we only found significant differences at the genus level (and not at any other taxonomic level; Fig. S5 and Supplementary Section 3). In the marine environment, these included large numbers of *Alphaproteobacteria*, that were more abundant in planktonic samples, while the *Bacteroidetes* and *Gammaproteobacteria* were more abundant in biofilm samples (either plastics or controls). Of particular interest were the taxonomic groups contained within the hydrocarbonoclastic *Oceanospirillales* (the families *Saccharospirillaceae* and *Halomonadaceae* and genera *Alcanivorax* and *Oleiphilus*) and *Alteromonadales* (the families *Alteromonadaceae*, *Marinobacteraceae* and *Pseudoalteromonadaceae*) [118] orders that were more abundant in plastic samples than control biofilms, with the *Oceanospirillales* generally being more abundant in the aliphatic plastic samples and the *Alteromonadales* generally being more abundant in the other plastic samples (Fig. S5 and Supplementary Section 3). The *Cyanobacteria* and the hydrocarbonoclastic *Halomonadaceae* were always more abundant in the unknown plastic samples (Fig. S5 and Supplementary Section 3), which were collected from the ocean after unknown residence times (Supplementary Table S2).

In the aquatic environment, there were also several potentially hydrocarbonoclastic taxonomic groups that were more abundant in plastic than other samples (the families *Thalassospiraceae*, *Alteromonadaceae*, *Pseudoalteromonadaceae*, *Saccharospirillaceae* and *Xanthomonadaceae*

and genera *Idiomarina* and *Alcanivorax*), however, this difference was only significant for the potentially PAH-degrading *Spirosomaceae* family [119]. In the freshwater environment, there were also several taxonomic groups that have either been isolated from hydrocarbon-contaminated environments (ASV0388 *Arcobacter* and ASV0394 *Arcobacter cryaerophilus* [120]; ASV3841 Unclassified *Xanthomonadaceae* [118]) or have been suggested to be capable of hydrocarbon (Unclassified *Immundisolibacteraceae* genera [121]) or biodegradable plastic (ASV2841 *Flavobacterium* [122]) degradation that were significantly more abundant in the unknown plastic samples. Those that were more abundant in other plastic samples, were classes (*Acidobacteria* and *Thermoanaerobaculia*) and orders (OPB56—*Ignavibacteria*—*Caulobacterales*, PB19—*Deltaproteobacteria*—and *Verrucomicrobiales*), and it is therefore much more difficult to speculate on the metabolic potential of these taxa. The same was true of the terrestrial environment, where many of the genera that were more abundant in the plastic samples remained unclassified at the order level (e.g., *Frankiales* and *Microtrichales*), although several genera known to be able to degrade hydrocarbons were also identified, e.g., *Arthrobacter*, *Acinetobacter* [118], *Methylocaldum* [123], and *Nitrosomonas* [124] as being significantly more abundant in the plastic (unknown plastics only) than control biofilm samples.

Previous studies have indicated the presence of potentially pathogenic hitchhikers in the Plastisphere, e.g., [30, 64, 125–128] however, other studies have found that potentially pathogenic species were actually higher in abundance on natural substrates, such as wood, e.g., [25]. Therefore, we investigated taxa that could potentially be pathogenic and that were more abundant on plastics than other samples. We found several taxonomic groups that: (i) were potential pathogens of animals—*Tenacibaculum* (abundant in marine aliphatic plastic samples) and unclassified *Pirellulaceae* (abundant in marine other plastic samples); (ii) contained lower taxonomic levels with human pathogens—*Clostridiales* and ASV0589 *Thalassospira* (both abundant in marine other plastic samples) and *Chlamydiae* (abundant in marine unknown plastic samples). However, the ability of a bacterium to be pathogenic is likely dependent on the presence of specific virulence factors which are often in mobile genetic elements [129]. The resolution of amplicon sequencing data is therefore insufficient to determine pathogenesis. There were also taxa, for example the genus *Vibrio*, of which members are potentially capable of degrading plastics [130] as well as being pathogens of humans and other organisms [129]. Curiously, we found that the *Vibrionales* were more abundant in plastic samples than control biofilm or planktonic samples in the marine environment, but in the aquatic environment they were more abundant in planktonic than any other samples. This highlights the need for other methods that will enable

differentiation between pathogens and non-pathogens, as well as between strains that are capable of plastic degradation and those that are not (as discussed in detail in [131]).

Microbial community succession on different material types

When we examined the taxa that discriminate between early and late incubation times (up to or above 7 days of incubation, respectively), we find some fairly consistent patterns across the marine and aquatic environments: (i) the *Bacteroidetes* were always significantly more abundant at later time points (where they differed in abundance between early and late time points); and (ii) the *Alphaproteobacteria* (or orders within the *Alphaproteobacteria*) are always more abundant at earlier time points (Fig. 6). Other taxonomic groups were more dependent upon the specific comparison. For example, the *Gammaproteobacteria*, that were more abundant at later time points in marine plastic samples (aliphatic and other plastics), but were generally more abundant at early time points in the aquatic environment, with the hydrocarbonoclastic *Oceanospirillales* and *Alteromonadales* in particular being more abundant at early time points. There were also a large number of phyla that were significantly ($p < 0.05$; Wilcoxon rank sum tests with holm-bonferroni false discovery rate correction) more abundant at early than late time points in the aquatic control biofilm samples, namely the *Chloroflexi*, *Verrucomicrobia*, *Actinobacteria*, *Cyanobacteria*, and *Deltaproteobacteria* (Fig. 6). Both other plastic and control biofilm samples in the freshwater environment showed different successional patterns than in either the marine or aquatic environments, although some members of the *Bacteroidia* were still more abundant at later time points. In the plastic samples, only the *Methylophilaceae* were more abundant at early than late incubation stages, while in the control biofilm samples there were no taxa that were significantly more abundant at late incubation stages.

The differences between substratum types were also compared at different time points, revealing different patterns between the marine and aquatic environments; there were more differences between sample pairings at late time points in the marine and at early time points in the aquatic environment, with a greater number of differences overall in the aquatic environment (Supplementary Sections 4 and 5). These again included the hydrocarbonoclastic *Oceanospirillales* and *Alteromonadales*, that were more abundant in aliphatic plastic samples than control biofilms at early time points, but these differences had largely disappeared by later time points. In the freshwater environment there were not enough samples for comparisons to be made for most sample pairings (only the other plastic and control biofilm

samples had both early and late time points), however, differences between sample types were only observed at late incubation times (between other and unknown plastics as well as other plastics and control biofilms).

Discussion

The re-analysis of 35 Plastisphere amplicon sequencing studies that we present here has allowed for a comprehensive characterization of the 16S rRNA gene community that colonizes plastics in marine, freshwater, other aquatic and terrestrial environments. We trained random forest models to classify samples within metadata categories, revealing that, overall, there were a number of variables—both relating to study design and environmental factors—that were better able to differentiate between sample groupings than plastic type (Fig. 5A). This agrees with the small re-analysis of studies presented by Oberbeckmann and Labrenz (2020; $n = 5$) [45], who reported that geographic location and salinity, but not substrate type, significantly discriminated between different samples. They did not, however, investigate different aspects of study design, such as environment, light regime or DNA extraction methods and only included studies that used the same primer pair, all of which are categories that we found to have a higher classification accuracy than location (latitude and longitude) or salinity (Fig. 5A). We tried to control for primer pair and DNA extraction technique by agglomerating sequences based on phylogenetic placement, which is known to reduce the variation between different 16S rRNA gene regions [50], although there are some taxa that will not be amplified by particular primer pairs, including the hydrocarbonoclastic groups targeted by many Plastisphere studies [132].

When we split our analyses to different environments (marine, aquatic, freshwater, and terrestrial), we find that plastic type could be used to train random forest models that were up to 83% accurate (Fig. 5B). We were, therefore, able to show that the differences identified by individual studies between microbial biofilms on control substrates and the Plastisphere were the same across the Plastisphere studies that we re-analysed here, regardless of the methodological differences between studies. For example, we find that the hydrocarbonoclastic *Oceanospirillales* and *Alteromonadales*, both previously suggested to be plastic biodegraders, e.g., [35, 73], were both important for differentiating between different substrate types in the random forest models (i.e., between aliphatic plastics, other plastics, control biofilms, and planktonic communities). They were also identified as significantly differentially abundant between these substrates. Furthermore, we find that the *Oceanospirillales* are higher in abundance in aliphatic plastics while the *Alteromonadales* are higher in abundance

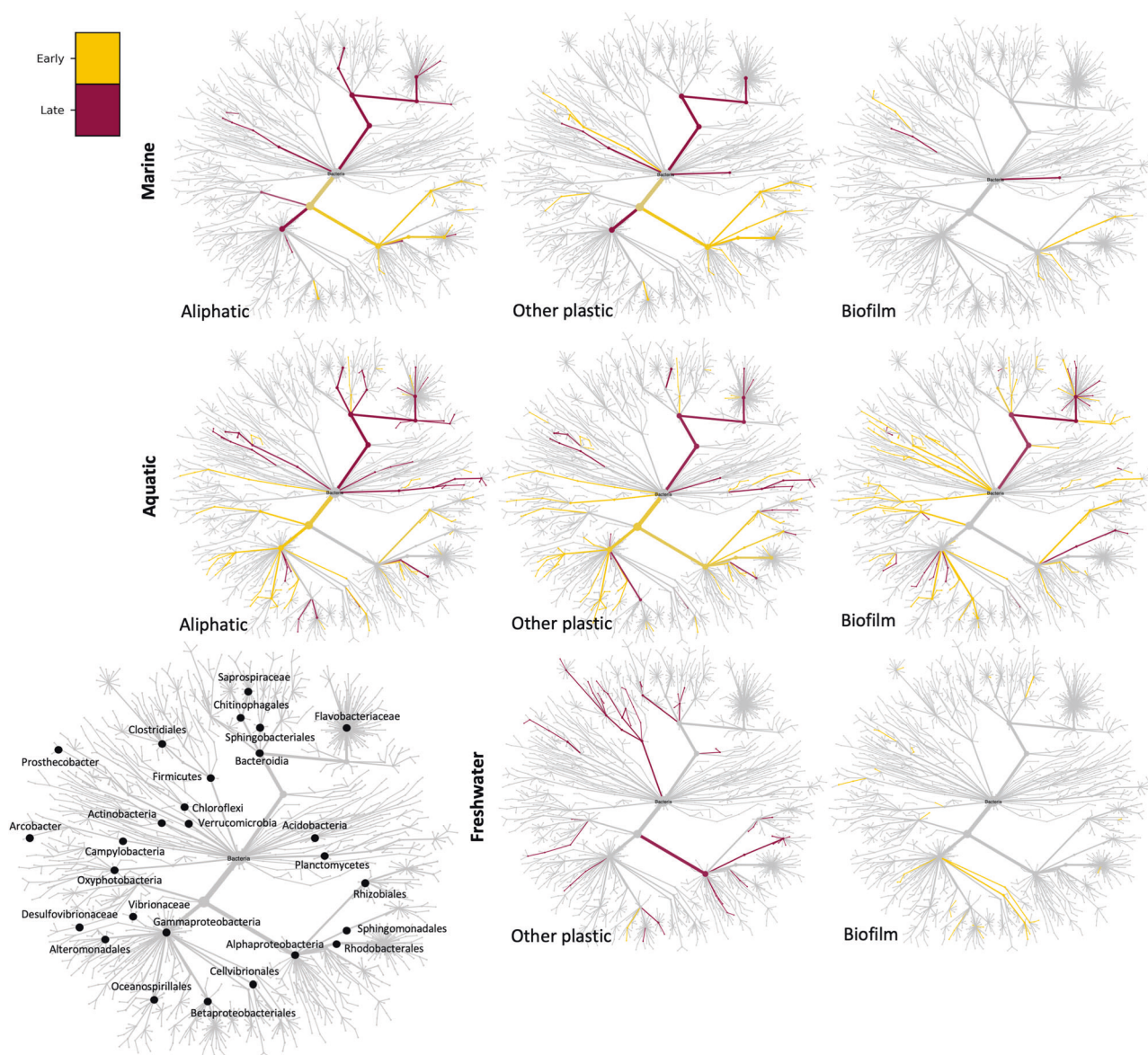


Fig. 6 Differential abundance of taxa between early and late incubation times. Heat trees showing differential abundance of taxa between early and late incubation times (up to or above 7 days, respectively) within substratum types (aliphatic, other plastic, and control biofilms) for each of the marine, aquatic and freshwater environments where samples were taken at different time points. Strength of colors indicates differential abundance, with gray indicating no significant difference ($p > 0.05$; Wilcoxon rank sum tests with holm-bonferroni false discovery rate correction), and strong

yellow or red colors indicating that \log_2 fold change is at least threefold higher in the early or late samples, respectively. Some key taxa are indicated in the empty, larger tree on the left and all are shown in Supplementary Section 4. Tests between different substrata at the same time point as well as between different substrata at all time points are also shown in Supplementary Sections 4 and 5. All terrestrial samples were collected after an unknown environmental residence time and could therefore not be included.

in the other plastic types. This is an observation that, to our knowledge, has not been previously acknowledged. Members of both of these orders, e.g., *Alcanivorax* sp. (*Oceanospirillales*) [35] and *Alteromonas* sp. (*Alteromonadales*) [133], have been reported to be able to degrade PE (aliphatic plastic) or PET (other plastic), respectively. However, these studies do not incorporate measurements of the plastic carbon assimilation into biomass (or of respiration)

and therefore fall short of being definitive measurements of plastic biodegradation [134].

Almost 1,000 toxic chemicals are known to be associated with plastics [135]. These include manufacturing additives as well as weathering sub-products, most of which are not chemically bound and leach from plastics on environmental exposure [136]. They are more labile and are therefore more likely to be biodegraded by the Plastisphere than plastics

themselves. However, there are only two studies included here that consider the effect that these have on community composition [31, 33]. For many plastics, these chemicals are not only additives but are also likely intermediates of both biotic and abiotic degradation. Cyclic and noncyclic dicarboxylic acids, for example, are the most commonly identified intermediates for the abiotic degradation of PE, PP, PS, and PET [137] and are also intermediates of biotic PET [138] and polyester [48] as well as plastic additive (plasticizer) [37] degradation. It has been suggested that the microbial community that is likely able to use these more labile chemicals is only present at earlier time points [31]. Almost all surfaces that enter the environment go through distinct and well-characterized stages of microbial community succession, and it is well known that microbial communities on different surfaces converge at later time points [139–141]. Here we have defined early incubation stages as before and later as after 1 week of incubation. However, due to the relatively small number of studies that characterize the Plastisphere across a wide range of different incubation times, we do not actually know what early or late incubation stages are within the Plastisphere. Long residence times in the order of weeks to months have generally been favored by Plastisphere studies—the mean incubation time for all samples included here was 100 days—rather than the days that it typically takes for microbial communities to diverge from organisms that are efficient at degrading the surfaces they colonize to cheaters, cross-feeders, and grazers [139, 142]. This makes sense when we know that plastics may take longer than hundreds of years to be completely mineralized [143], but it also explains why several studies have found that potential degraders are only present at very low relative abundances [26] or are abundant only at early time points [31].

There are also several other areas that we have identified as needing further research before definitive conclusions may be made. For instance, there are currently only four studies performed in the terrestrial environment (and only two that could be included), all current Plastisphere studies were performed in the Northern Hemisphere and the majority were performed in temperate environments (Fig. 1). Whilst this hinders our ability to draw large-scale conclusions, it offers an opportunity for researchers to ensure that future data collected in these areas are comparable and address these knowledge gaps. We were unable to confirm or reject the suggestions that plastics carry higher abundances of pathogenic hitchhikers, e.g., [23, 30, 64, 126–128], and/or antimicrobial resistance genes, e.g., [39], than control biofilms. We do note, however, that due to their recalcitrance and buoyancy plastics present a higher chance of carrying pathogens or antimicrobial resistance genes across greater distances than most natural surfaces. We, along with many of the studies

included here, have identified the presence of potential plastic biodegraders and potentially pathogenic organisms on plastics. However, the amplicon sequencing used does not usually give the resolution required to differentiate between closely related strains of the same species [144], for example between pathogenic and non-pathogenic *Vibrio* spp. [129, 145] or hydrocarbon degrading and nondegrading *Pseudomonas putida* (previously *P. oleovorans*) spp. [146]. The presence of genes encoding the production of enzymes used in hydrocarbon biodegradation [118], or specific virulence factors required for pathogenicity are often conferred by mobile genetic elements [129] and their presence is therefore not necessarily based in phylogeny. The use of metagenomics rather than amplicon sequencing in future studies would aid in determining whether these *potential* plastic biodegraders and *potential* pathogens are indeed biodegraders or pathogenic [131]. As discussed further in [131], it may even generate candidate enzymes that could be tested for their ability to degrade plastics, such as in Danso et al. [147] for PET hydrolases.

This study presents, for the first time, a comprehensive re-analysis of all Plastisphere studies that utilize amplicon sequencing of the 16S rRNA gene. We have revealed through machine learning methods that environmental factors, such as environment and light availability as well as aspects of study design, such as primer pair and incubation time play a large role in shaping Plastisphere community composition. Notably, we have identified members of the microbial community that are consistently more abundant in biofilms formed on plastics than control biofilms across multiple studies and environments. This highlights the urgent need to determine whether these microbes are capable of plastic biodegradation or are pathogens of humans or other organisms. We have also identified a number of other key areas in which we are lacking even basic knowledge and where future research should be directed. It is clear that plastic pollution is a key indicator of the Anthropocene and we must focus future research on gaps in our knowledge that we highlight here if we wish to mitigate its effects.

Data availability

All sequencing data used here were either obtained from publicly accessible databases or directly from the authors of the studies (full details are in Table S1B) and all scripts used during these analyses can be found in File S1 and at <https://github.com/R-Wright-1/Plastisphere-MetaAnalysis>. This includes a pipeline with step-by-step instructions and code for reproducing the analyses carried out here, with or without the inclusion of additional studies. All files used and produced during these analyses (e.g., QIIME2 objects, random forest output files and figures) are available on Figshare: raw read files in QIIME2 zipped format, <https://>

doi.org/10.6084/m9.figshare.12931372; QIIME2 merged input files, including processed individual study and all merged study reads, <https://doi.org/10.6084/m9.figshare.12217682>; QIIME2 intermediate and output files <https://doi.org/10.6084/m9.figshare.12227522>; and all files created during the analysis <https://doi.org/10.6084/m9.figshare.12227303>.

Acknowledgements RJW gratefully acknowledges funding from Waitrose & Partners as part of the Association of Commonwealth Universities Blue Charter Program and thanks members of the Christie-Oleza (University of Warwick, UK, and University of the Balearic Islands, Spain) and Langille labs (Dalhousie University, Canada) for insightful discussions and advice. MGIL and TRW acknowledge support from NSERC Discovery Grants 2016–05039 and 2018–04119, respectively. We would also like to thank the authors of all studies that are included in this analysis for making their data accessible, as well as those that responded to requests for additional data or clarification on aspects of their methodology.

Author contributions RJW designed and carried out all practical aspects of the study with guidance and technical support from MGIL and TRW. RJW wrote the first draft of the manuscript and all authors contributed substantially to revisions. RJW and TRW acquired funding for this project.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Geyer R, Jambeck JR, Law KL. Production, use, and fate of all plastics ever made. *Sci Adv*. 2017;3:e1700782.
- Carr A. Impact of nondegradable marine debris on the ecology and survival outlook of sea turtles. *Mar Pollut Bull*. 1987;18:352–6.
- Gregory MR. Environmental implications of plastic debris in marine settings—entanglement, ingestion, smothering, hangers-on, hitch-hiking and alien invasions. *Philos Trans R Soc*. 2009; B364:2013–2025.
- Browne MA, Dissanayake A, Galloway TS, Lowe DM, Thompson RC. Ingested microscopic plastic translocates to the circulatory system of the mussel, *Mytilus edulis* (L.). *Environ Sci Technol*. 2008;42:5026–31.
- Colabuono FI, Taniguchi S, Montone RC. Polychlorinated biphenyls and organochlorine pesticides in plastics ingested by seabirds. *Mar Pollut Bull*. 2010;60:630–4.
- Ryan PG, Connell AD, Gardner BD. Plastic ingestion and PCBs in seabirds: is there a relationship? *Mar Pollut Bull*. 1988;19:174–6.
- Vered G, Kaplan A, Avisar D, Shenkar N. Using solitary ascidians to assess microplastic and phthalate plasticizers pollution among marine biota: a case study of the Eastern Mediterranean and Red Sea. *Mar Pollut Bull*. 2019;138:618–25.
- Karbalaei S, Golieskardi A, Hamzah HB, Abdulwahid S, Hanachi P, Walker TR, et al. Abundance and characteristics of microplastics in commercial marine fish from Malaysia. *Mar Pollut Bull*. 2019;148:5–15.
- Rochman CM, Hoh E, Kurobe T, Teh SJ. Ingested plastic transfers hazardous chemicals to fish and induces hepatic stress. *Sci Rep*. 2013;3:1–7.
- Teuten EL, Saquing JM, Knappe DRU, Barlaz MA, Jonsson S, Bjorn A, et al. Transport and release of chemicals from plastics to the environment and to wildlife. *Philos Trans R Soc B Biol Sci*. 2009;364:2027–45.
- de Souza Machado AA, Kloas W, Zarfl C, Hempel S, Rillig MC. Microplastics as an emerging threat to terrestrial ecosystems. *Glob Chang Biol*. 2018;24:1405–16.
- Allen S, Allen D, Phoenix VR, Le Roux G, Durántez Jiménez P, Simonneau A, et al. Atmospheric transport and deposition of microplastics in a remote mountain catchment. *Nat Geosci*. 2019;12:339–44.
- Bergmann M, Mützel S, Primpke S, Tekman MB, Trachsel J, Gerdt G. White and wonderful? Microplastics prevail in snow from the Alps to the Arctic. *Sci Adv*. 2019;5:1–11.
- Horton AA, Dixon SJ. Microplastics: an introduction to environmental transport processes. *Wiley Interdiscip Rev Water*. 2018;5:e1268.
- Van Sebille E, Wilcox C, Sherman P, Hardesty BD, Law KL. Modelling global distribution, risk and mitigation strategies of floating plastic pollution. *Geophys Research Abstract*. EGU General Assembly. Vienna, Austria. 2016.
- Bakir A, Rowland SJ, Thompson RC. Competitive sorption of persistent organic pollutants onto microplastics in the marine environment. *Mar Pollut Bull*. 2012;64:2782–9.
- Andrady AL. The plastic in microplastics: a review. *Mar Pollut Bull*. 2017;119:12–22.
- Amaral-Zettler LA, Zettler ER, Mincer TJ. Ecology of the plastisphere. *Nat Rev Microbiol*. 2020;18:139–51.
- Derraik JGB. The pollution of the marine environment by plastic debris: a review. *Mar Pollut Bull*. 2002;44:842–52.
- Gregory MR. Accumulation and distribution of virgin plastic granules on New Zealand beaches. *N Zeal J Mar Freshw Res*. 1978;12:399–414.
- Barnes DKA. Invasions by marine life on plastic debris. *Nature*. 2002;416:808–9.
- Darby RT, Kaplan AM. Fungal susceptibility of polyurethanes. *Appl Microbiol*. 1968;16:900–5.
- Zettler ER, Mincer TJ, Amaral-Zettler LA. Life in the “Plastisphere”: Microbial communities on plastic marine debris. *Environ Sci Technol*. 2013;47:7137–7146.
- Harrison JP, Sapp M, Schratzberger M, Osborn AM. Interactions between microorganisms and marine microplastics: a call for research. *Mar Technol Soc J*. 2011;45:12–20.
- Kesy K, Oberbeckmann S, Kreikemeyer B, Labrenz M. Spatial environmental heterogeneity determines young biofilm assemblages on microplastics in Baltic Sea Mesocosms. *Front Microbiol*. 2019;10:1–18.
- Kirstein IV, Wichels A, Krohne G, Gerdt G. Mature biofilm communities on synthetic polymers in seawater - Specific or general? *Mar Environ Res*. 2018;142:147–54.

27. Kirstein IV, Wichels A, Gullans E, Krohne G, Gerdt G. The plastisphere – Uncovering tightly attached plastic “specific” microorganisms. *PLoS One*. 2019;14:1–17.
28. Oberbeckmann S, Osborn AM, Duhaime MB. Microbes on a bottle: substrate, season and geography influence community composition of microbes colonizing marine plastic debris. *PLoS One*. 2016;11:1–24.
29. Muthukrishnan T, Khaburi MAL, Abed RMM. Fouling microbial communities on plastics compared with wood and steel: Are they substrate- or location-specific? *Micro Ecol*. 2018;78:361–74.
30. Oberbeckmann S, Kreikemeyer B, Labrenz M. Environmental factors support the formation of specific bacterial assemblages on microplastics. *Front Microbiol*. 2018;8:1–12.
31. Erni-Cassola G, Wright RJ, Gibson MI, Christie-Oleza JA. Early colonization of weathered polyethylene by distinct bacteria in marine coastal seawater. *Micro Ecol*. 2019;79:517–26.
32. Ogonowski M, Motiei A, Ininbergs K, Hell E, Gerdes Z, Udekwu KI, et al. Evidence for selective bacterial community structuring on microplastics. *Environ Microbiol*. 2018;20:2796–808.
33. Pinto M, Langer TM, Huffer T, Hofmann T, Herndl GJ, Hüffer T, et al. The composition of bacterial communities associated with plastic biofilms differs between different polymers and stages of biofilm succession. *PLoS One*. 2019;14:1–20.
34. Kettner MT, Oberbeckmann S, Labrenz M, Grossart HP. The eukaryotic life on microplastics in brackish ecosystems. *Front Microbiol*. 2019;10:538.
35. Delacuvellerie A, Cyriaque V, Gobert S, Benali S, Wattiez R. The plastisphere in marine ecosystem hosts potential specific microbial degraders including *Alcanivorax borkumensis* as a key player for the low-density polyethylene degradation. *J Hazard Mater*. 2019;380:120899.
36. Devi RS, Ramya R, Kannan K, Antony AR, Kannan VR. Investigation of biodegradation potentials of high density polyethylene degrading marine bacteria isolated from the coastal regions of Tamil Nadu, India. *Mar Pollut Bull*. 2019;138:549–60.
37. Wright RJ, Bosch R, Gibson MI, Christie-Oleza JA. Plasticizer degradation by marine bacterial isolates: a proteogenomic and metabolomic characterization. *Environ Sci Technol*. 2020;54:2244–56.
38. De Tender CA, Devriese LI, Haegeman A, Maes S, Ruttink T, Dawyndt P. Bacterial community profiling of plastic litter in the Belgian part of the North Sea. *Environ Sci Technol*. 2015;49:9629–38.
39. Laganà P, Caruso G, Corsi I, Bergami E, Venuti V, Majolino D, et al. Do plastics serve as a possible vector for the spread of antibiotic resistance? First insights from bacteria associated to a polystyrene piece from King George Island (Antarctica). *Int J Hyg Environ Health*. 2019;222:89–100.
40. Yang Y, Liu G, Song W, Ye C, Lin H, Li Z, et al. Plastics in the marine environment are reservoirs for antibiotic and metal resistance genes. *Environ Int*. 2019;123:79–86.
41. Zhang Y, Lu J, Wu J, Wang J, Luo Y. Potential risks of microplastics combined with superbugs: Enrichment of antibiotic resistant bacteria on the surface of microplastics in mariculture system. *Ecotoxicol Environ Saf*. 2020;187:109852.
42. Wang S, Xue N, Li W, Zhang D, Pan X, Luo Y. Selectively enrichment of antibiotics and ARGs by microplastics in river, estuary and marine waters. *Science Total Environment*. 2020;708:134594.
43. Jacquin J, Cheng J, Odobel C, Pandin C, Conan P, Pujo-Pay M, et al. Microbial ecotoxicology of marine plastic debris: a review on colonization and biodegradation by the “Plastisphere”. *Front Microbiol*. 2019;10:1–16.
44. Roager L, Sonnenschein EC. Bacterial candidates for colonization and degradation of marine plastic debris. *Environ Sci Technol*. 2019;53:11636–43.
45. Oberbeckmann S, Labrenz M. Marine microbial assemblages on microplastics: diversity, adaptation, and role in degradation. *Ann Rev Mar Sci*. 2020;12:1–24.
46. Rogers KL, Carreres-calabuig JA, Gorokhova E, Posth NR. Microby-micro interactions: how microorganisms influence the fate of marine microplastics. *Limnol Oceanogr Lett*. 2020;5:1–19.
47. Wei R, Zimmermann W. Microbial enzymes for the recycling of recalcitrant petroleum-based plastics: how far are we? *Micro Biotechnol*. 2017;10:1308–22.
48. Danso D, Chow J, Streit WR. Plastics: microbial degradation, environmental and biotechnological perspectives. *Appl Environ Microbiol*. 2019;85:1–14.
49. Salvador de Lara M, Abdulmutalib U, Gonzalez J, Kim J, Smith AA, Faulon J-L, et al. Genes for a circular and sustainable bio-PET economy. *Genes (Basel)*. 2019;10:1–15.
50. Janssen S, McDonald D, Gonzalez A, Navas-molina JA, Jiang L, Xu Z. Phylogenetic placement of exact amplicon sequences improves associations with clinical information. *mSystems*. 2018;3:e00021–18.
51. Abed RMM, Muthukrishnan T, Al Khaburi M, Al-Senafi F, Munam A, Mahmoud H. Degradability and biofouling of oxo-biodegradable polyethylene in the planktonic and benthic zones of the Arabian Gulf. *Mar Pollut Bull*. 2020;150:110639.
52. Amaral-Zettler LA, Zettler ER, Slikas B, Boyd GD, Melvin DW, Morrall CE, et al. The biogeography of the Plastisphere: implications for policy. *Front Ecol Environ*. 2015;13:541–6.
53. Arias-Andres M, Klümper U, Rojas-Jimenez K, Grossart H-P. Microplastic pollution increases gene exchange in aquatic ecosystems. *Environ Pollut*. 2018;237:253–61.
54. Bakal T, Janata J, Sabova L, Grabic R, Zlabek V, Najmanova L. Suitability and setup of next-generation sequencing-based method for taxonomic characterization of aquatic microbial biofilm. *Folia Microbiol*. 2019;64:9–17.
55. Canada P, Pereira A, Nogueira N, Png-Gonzalez L, Andrade C, Xavier R. Analysis of bacterial microbiome associated with nylon and copper nets in an aquaculture context. *Aquaculture*. 2020;516:734540.
56. Curren E, Leong SCY. Profiles of bacterial assemblages from microplastics of tropical coastal environments. *Sci Total Environ*. 2019;655:313–20.
57. Dussud C, Hudec C, George M, Fabre P, Higgs P, Bruzaud S, et al. Colonization of non-biodegradable and biodegradable plastics by marine microorganisms. *Front Microbiol*. 2018;9:1–13.
58. Dussud C, Meistertzheim AL, Conan P, Pujo-pay M, George M, Fabre P, et al. Evidence of niche partitioning among bacteria living on plastics, organic particles and surrounding seawaters. *Environ Pollut*. 2018;236:807–16.
59. Frère L, Maignien L, Chalopin M, Huvet A, Rinnert E, Morrison H, et al. Microplastic bacterial communities in the bay of brest: influence of polymer type and size. *Environ Pollut*. 2018;242:614–25.
60. Gong M, Yang G, Zhuang L, Zeng EY. Microbial biofilm formation and community structure on low-density polyethylene microparticles in lake water microcosms. *Environ Pollut*. 2019;252:94–102.
61. Hoellein T, Rojas M, Pink A, Gasior J, Kelly J. Anthropogenic litter in urban freshwater ecosystems: distribution and microbial interactions. *PLoS One*. 2014;9:e98485.
62. Hoellein TJ, McCormick AR, Hittie J, London MG, Scott JW, Kelly JJ. Longitudinal patterns of microplastic concentration and bacterial assemblages in surface and benthic habitats of an urban river. *Freshw Sci*. 2017;36:491–507.

63. Huang Y, Zhao Y, Wang J, Zhang M, Jia W, Qin X. LDPE microplastic films alter microbial community composition and enzymatic activities in soil. *Environ Pollut.* 2019;254:112983.
64. Jiang P, Zhao S, Zhu L, Li D. Microplastic-associated bacterial assemblages in the intertidal zone of the Yangtze Estuary. *Sci Total Environ.* 2018;624:48–54.
65. McCormick A, Hoellein TJ, Mason SA, Schlupe J, Kelly JJ. Microplastic is an abundant and distinct microbial habitat in an Urban river. *Environ Sci Technol.* 2014;48:11863–71.
66. McCormick AR, Hoellein TJ, London MG, Hittie J, Scott JW, J. KJ. Microplastic in surface waters of urban rivers: concentration, sources, and associated bacterial assemblages. *Ecosphere.* 2016;7:1–22.
67. Miao L, Wang P, Hou J, Yao Y, Liu Z, Liu S, et al. Distinct community structure and microbial functions of biofilms colonizing microplastics. *Sci Total Environ.* 2019;650:2395–402.
68. Parrish K, Fahrenfeld NL. Microplastic biofilm in fresh- and wastewater as a function of microparticle type and size class. *Environ Sci Water Res Technol.* 2019;5:495–505.
69. Pollet T, Berdjeb L, Garnier CC, Durrieu GG, Le Poupon C, Misson B, et al. Prokaryotic community successions and interactions in marine biofilms: the key role of Flavobacteriia. *FEMS Micro Ecol.* 2018;94:1–13.
70. Puglisi E, Romaniello F, Galletti S, Boccaleri E, Frache A, Cocconcelli PS. Selective bacterial colonization processes on polyethylene waste samples in an abandoned landfill site. *Sci Rep.* 2019;9:1–13.
71. Rosato A, Barone M, Negroni A, Brigidi P, Fava F, Xu P, et al. Microbial colonization of different microplastic types and biotransformation of sorbed PCBs by a marine anaerobic bacterial community. *Sci Total Environ.* 2019;705:135790.
72. Syranidou E, Karkanorachaki K, Amorotti F, Repouskou E, Kroll K, Kolvenbach B, et al. Development of tailored indigenous marine consortia for the degradation of naturally weathered polyethylene films. *PLoS One.* 2017;12:1–21.
73. Syranidou E, Karkanorachaki K, Amorotti F, Franchini M, Repouskou E, Kaliva M, et al. Biodegradation of weathered polystyrene films in seawater microcosms. *Sci Rep.* 2017;7:1–12.
74. Syranidou E, Karkanorachaki K, Amorotti F, Aygeropoulos A, Kolvenbach B, Zhou N, et al. Biodegradation of mixture of plastic films by tailored marine consortia. *J Hazard Mater.* 2019;375:33–42.
75. Tagg AS, Oberbeckmann S, Fischer D, Kreikemeyer B, Labrenz M. Paint particles are a distinct and variable substrate for marine bacteria. *Mar Pollut Bull.* 2019;146:117–24.
76. Woodall LC, Jungblut AD, Hopkins K, Id AH, Robinson F, Gwinnett C, et al. Deep-sea anthropogenic macrodebris harbours rich and diverse communities of bacteria and archaea. *PLoS ONE.* 2018;13:1–11.
77. Wu N, Zhang Y, Zhao Z, He J, Li W, Li J, et al. Colonization characteristics of bacterial communities on microplastics compared with ambient environments (water and sediment) in Haihe Estuary. *Sci Total Environ.* 2019;708:134876.
78. Xu X, Wang S, Gao F, Li J, Zheng L, Sun C, et al. Marine microplastic-associated bacterial community succession in response to geography, exposure time, and plastic type in China's coastal seawaters. *Mar Pollut Bull.* 2019;145:278–86.
79. Zhang M, Zhao Y, Qin X, Jia W, Chai L, Huang M, et al. Microplastics from mulching film is a distinct habitat for bacteria in farmland soil. *Sci Total Environ.* 2019;688:470–8.
80. Kettner MT, Rojas-Jimenez K, Oberbeckmann S, Labrenz M, Grossart H-P. Microplastics alter composition of fungal communities in aquatic ecosystems. *Environ Microbiol.* 2017;19:4447–59.
81. Didier D, Anne M, Alexandra TH. Plastics in the North Atlantic garbage patch: a boat-microbe for hitchhikers and plastic degraders. *Sci Total Environ.* 2017;599–600:1222–32.
82. Esan EO, Abbey, Lord, Yurgel S. Exploring the long-term effect of plastic on compost microbiome. *PLoS One.* 2019;14:1–17.
83. De Tender C, Devriese LI, Haegeman A, Maes S, Vangeyte J, Cattrijsse A, et al. Temporal dynamics of bacterial and fungal colonization on plastic debris in the north sea. *Environ Sci Technol.* 2017;51:7350–60.
84. Pinnell LJ, Turner JW. Shotgun metagenomics reveals the benthic microbial community response to plastic and bioplastic in a coastal marine environment. *Front Microbiol.* 2019;10:1–13.
85. Bryant JA, Clemente TM, Viviani DA, Fong AA, Thomas KA, Kemp P, et al. Diversity and activity of communities inhabiting plastic debris in the North Pacific Gyre. *mSystems.* 2016;1:e00024–16.
86. Wu X, Pan J, Li M, Li Y, Bartlam M, Wang Y. Selective enrichment of bacterial pathogens by microplastic biofilm. *Water Res.* 2019;165:114979.
87. Comeau AM, Douglas GM, Langille MGI. Microbiome helper: a custom and streamlined workflow for microbiome research. *mSystems.* 2017;2:1–11.
88. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol.* 2019;37:852–7.
89. Ewels P, Magnusson M, Lundin S, Källner M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics.* 2016;32:3047–8.
90. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* 2011;17:10–12.
91. Rognes T, Flouri T, Nichols B, Quince C, Mahé F. VSEARCH: a versatile open source tool for metagenomics. *PeerJ.* 2016;4:e2584.
92. Amir A, Daniel M, Navas-Molina J, Kopylova E, Morton J, Xu ZZ, et al. Deblur rapidly resolves single-nucleotide community sequence patterns. *Am Soc Microbiol.* 2017;2:1–7.
93. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Glo FO, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* 2013;41:590–6.
94. Cock PJA, Antao T, Chang JT, Chapman BA, Cox CJ, Dalke A, et al. Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics.* 2009;25:1422–3.
95. Harrell FE, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med.* 2005;15:361–87.
96. Hunter JD. Matplotlib: a 2D graphics environment. *Comput Sci Eng.* 2007;9:90–95.
97. McKinney W. Data structures for statistical computing in Python. In: van der Walt S, Millman J, editors. *Proc. 9th Python Sci. Conf.* 2010.
98. Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat Methods.* 2020;17:261–72.
99. Pedregosa F, Weiss R, Brucher M. Scikit-learn: machine learning in Python. *J Mach Learn Res.* 2011;12:2825–30.
100. Paradis E, Schliep K. Ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics.* 2019;35:526–8.
101. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2013. <http://www.R-project.org/>.
102. Hadley W, François R, Henry L, Müller K. dplyr: a grammar of data manipulation. 2020. <https://dplyr.tidyverse.org/>.

103. Hothorn T, Hornik K. exactRankTests: exact distributions for rank and permutation tests. 2019. <https://rdrr.io/cran/exactRankTests/>.
104. Campitelli E. ggnewscale: multiple fill and colour scales in 'ggplot2'. <https://rdrr.io/cran/ggnewscale/>.
105. Wickham H. ggplot2: elegant graphics for data analysis. New York: Springer-Verlag; 2016.
106. Yu G, Smith DK, Zhu H, Guan Y, Lam TT. GGTREE: an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods Ecol Evol.* 2017;8:28–36.
107. Xie Y. Knitr: a comprehensive tool for reproducible research in R. In: Stodden V, Leisch F, Peng RD, editors. *Implementing reproducible research*. New York: Taylor and Francis; 2014.
108. Foster ZSL, Sharpton TJ, Grünwald NJ. Metacoder: an R package for visualization and manipulation of community taxonomic diversity data. *PLoS Comput Biol.* 2017;13:1–15.
109. Lahti L, Shetty S. Microbiome R package. 2019. <https://microbiome.github.io/tutorials/>.
110. Pinheiro J, Bates D, DebRoy S, Sarkar D. nlme: linear and nonlinear mixed effects models. 2020. https://www.researchgate.net/publication/303803175_Nlme_Linear_and_Nonlinear_Mixed_Effects_Models.
111. Silverman JD, Washburne AD, Mukherjee S, David LA. A phylogenetic transform enhances analysis of compositional microbiota data. *Elife.* 2017;6:1–20.
112. McMurdie PJ, Holmes S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One.* 2013;8:1–11.
113. Ushey K, Allaire J, Tang Y. Reticulate: interface to 'Python'. 2020.
114. Dixon P. VEGAN, a package of R functions for community ecology. *J Veg Sci.* 2003;14:927–30.
115. Lozupone C, Knight R. UniFrac: a new phylogenetic method for comparing microbial communities. *Appl Environ Microbiol.* 2005;71:8228–35.
116. Gloor GB, Macklaim JM, Pawlowsky-Glahn V, Egozcue JJ. Microbiome datasets are compositional: and this is not optional. *Front Microbiol.* 2017;8:1–6.
117. Weiss S, Xu ZZ, Peddada S, Amir A, Bittinger K, Gonzalez A, et al. Normalization and microbial differential abundance strategies depend upon data characteristics. *Microbiome.* 2017;5:1–18.
118. Prince RC, Gramain A, McGenity TJ. Prokaryotic hydrocarbon degraders. In: Timmis K, editor. *Handbook of hydrocarbon and lipid microbiology*. Berlin, Heidelberg: Springer; 2010. p. 1669–92.
119. Zhou HW, Luan TG, Zou F, Tam NFY. Different bacterial groups for biodegradation of three- and four-ring PAHs isolated from a Hong Kong mangrove sediment. *J Hazard Mater.* 2008;152:1179–85.
120. Evans MV, Panescu J, Hanson AJ, Welch SA, Sheets JM, Nastasi N, et al. Members of marinobacter and arcobacter influence system biogeochemistry during early production of hydraulically fractured natural gas wells in the appalachian basin. *Front Microbiol.* 2018;9:1–17.
121. Corteselli EM, Aitken MD, Singleton DR. Description of *Immundisolibacter cernigliae* gen. Nov., sp. nov., a high-molecular-weight polycyclic aromatic hydrocarbon-degrading bacterium within the class Gammaproteobacteria, and proposal of *Immundisolibacterales* ord. nov. and *Immundisolibacteraceae* fa. *Int J Syst Evol Microbiol.* 2017;67:925–31.
122. Feng XM, Tan X, Jia L, Long PP, Han L, Lv J. *Flavobacterium butense* sp. nov., isolated from freshwater. *Arch Microbiol.* 2015;197:1109–15.
123. Rochman FF, Sheremet A, Tamas I, Saidi-Mehrabad A, Kim JJ, Dong X, et al. Benzene and naphthalene degrading bacterial communities in an oil sands tailings pond. *Front Microbiol.* 2017;8:1–12.
124. John RC, Okpokwasili GC. Crude oil-degradation and plasmid profile of nitrifying bacteria isolated from oil-impacted mangrove sediment in the Niger Delta of Nigeria. *Bull Environ Contam Toxicol.* 2012;88:1020–6.
125. Zettler ER, Mincer TJ, Amaral-Zettler LA. Life in the 'plastisphere': microbial communities on plastic marine debris. *Environ Sci Technol.* 2013;47:7137–46.
126. Keswani A, Oliver DM, Gutierrez T, Quilliam RS. Microbial hitchhikers on marine plastic debris: human exposure risks at bathing waters and beach environments. *Mar Environ Res.* 2016;118:10–19.
127. Kirstein IV, Kirmizi S, Wichels A, Garin-Fernandez A, Erler R, Löder M, et al. Dangerous hitchhikers? Evidence for potentially pathogenic *Vibrio* spp. on microplastic particles. *Mar Environ Res.* 2016;120:1–8.
128. Silva MM, Maldonado GC, Castro RO, de Sá Felizardo J, Cardoso RP, Anjos RMDos, et al. Dispersal of potentially pathogenic bacteria by plastic debris in Guanabara Bay, RJ, Brazil. *Mar Pollut Bull.* 2019;141:561–8.
129. Sakib SN, Reddi G, Almagro-Moreno S. Environmental role of pathogenic traits in *Vibrio cholerae*. *J Bacteriol.* 2018;200:1–12.
130. Raghu SS, Bhat SG, Chandrasekaran M, Francis V, Thachil ET. Biodegradation of polyvinyl alcohol-low linear density polyethylene-blended plastic film by consortium of marine benthic vibrios. *Int J Environ Sci Technol.* 2014;11:1827–34.
131. Wright RJ, Erni-Cassola G, Zadjelovic V, Latva M, Christie-Oleza J. Marine plastic debris – a new surface for microbial colonization. *Environ Sci Technol.* 2020;54:11657–72.
132. Berry D, Gutierrez T, Waite DW. Evaluating the detection of hydrocarbon-degrading bacteria in 16S rRNA gene sequencing surveys. *Front Microbiol.* 2017;8:1–6.
133. Webb H. Biodegradation of poly(ethylene terephthalate) by marine bacteria, and strategies for its enhancement. Swinburne University of Technology: Melbourne, Australia. 2012.
134. Zumstein MT, Narayan R, Kohler H-PE, McNeill K, Sander M. Do and do nots when assessing the biodegradation of plastics. *Environ Sci Technol.* 2019;53:9967–9.
135. Groh KJ, Backhaus T, Carney-Almroth B, Guecke B, Inostroza PA, Lennquist A, et al. Overview of known plastic packaging-associated chemicals and their hazards. *Sci Total Environ.* 2019;651:3253–68.
136. Hahladakis JN, Velis CA, Weber R, Iacovidou E, Purnell P. An overview of chemical additives present in plastics: migration, release, fate and environmental impact during their use, disposal and recycling. *J Hazard Mater.* 2018;344:179–99.
137. Gewert B, Plassmann M, Sandblom O, Macleod M. Identification of chain scission products released to water by plastic exposed to ultraviolet light. *Environ Sci Technol Lett.* 2018;5:272–6.
138. Yoshida S, Hiraga K, Takehana T, Taniguchi I, Yanaji H, Maeda Y, et al. A bacterium that degrades and assimilates poly(ethylene terephthalate). *Science.* 2016;351:1196–9.
139. Datta MS, Sliwerska E, Gore J, Polz MF, Cordero OX. Microbial interactions lead to rapid micro-scale successions on model marine particles. *Nat Commun.* 2016;7:1–7.
140. Dang H, Li T, Chen M, Huang G. Cross-ocean distribution of Rhodobacterales bacteria as primary surface colonizers in temperate coastal marine waters. *Appl Environ Microbiol.* 2008;74:52–60.

141. Wu X, Wu L, Liu Y, Zhang P, Li Q, Zhou J, et al. Microbial interactions with dissolved organic matter drive carbon dynamics and community succession. *Front Microbiol.* 2018;9:1–12.
142. Wright RJ, Gibson MI, Christie-Oleza JA. Understanding microbial community dynamics to improve optimal microbiome selection. *Microbiome.* 2019;7:1–14.
143. Andrady AL. Persistence of plastic litter in the oceans. In: Bergmann M, Gutow L, Klages M, editors. *Marine Anthropogenic Litter.* Springer Open: Springer Nature Switzerland. 2015. p. 57–72.
144. Johnson JS, Spakowicz DJ, Hong BY, Petersen LM, Demkowicz P, Chen L, et al. Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. *Nat Commun.* 2019;10:1–11.
145. Horseman MA, Surani S. A comprehensive review of *Vibrio vulnificus*: an important cause of severe sepsis and skin and soft-tissue infection. *Int J Infect Dis.* 2011;15:e157–e166.
146. Van Hamme JD, Singh A, Ward OP. Recent advances in petroleum microbiology. *Microbiol Mol Biol Rev.* 2003;67:503–49.
147. Danso D, Zimmermann, Schmeisser C, Chow J, Zimmermann W, Wei R, et al. New insights into the function and global distribution of polyethylene terephthalate (PET) degrading bacteria and enzymes in marine and terrestrial metagenomes. *Appl Environ Microbiol.* 2018;84:e02773–17.