The core bacteria of Fucus distichus include apparent seaweed generalists

Jungsoo Park (✉ jungsoo.park@botany.ubc.ca )
University of British Columbia

Katherine Davis
University of British Columbia

Geneviève Lajoie
University of Montreal

Laura Wegener Parfrey
University of British Columbia

Research Article

Keywords: core microbiome, macroalgae, seaweed microbiome, microbial ecology, symbiosis, host specificity

Posted Date: May 23rd, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1620241/v2

License: ☒ ☀ This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Version of Record: A version of this preprint was published at Environmental Microbiome on November 16th, 2022. See the published version at https://doi.org/10.1186/s40793-022-00451-z.
Abstract

Background

Identifying the meaningful ecological associations between host and components of the microbiome is challenging. This is especially true for hosts where the taxonomic composition of the microbiome is highly diverse and variable in space and time such as marine macroalgae. Identifying core taxa is one way forward. This study leverages a large dataset of microbial communities associated with the widespread brown seaweed, *Fucus distichus*, at multiple spatial and temporal scales. We compare two different methodological approaches to identify core taxa at the amplicon sequence variant (ASV) level from this dataset: simple frequency analysis of *F. distichus* alone over the whole dataset, and an iterative indicator species analysis method identifies prevalent taxa that are consistently enriched on *F. distichus* in comparison to the local environment (IndVal). We then evaluate host-specificity of the identified *F. distichus* core ASVs using comparative data from 35 other seaweed species sampled at one of the sites.

Results

No bacterial taxa were at 100% prevalence in *F. distichus* samples, and IndVal identified a diverse array of *F. distichus* indicator taxa across sites and over time. Across the entire dataset, a set of eleven core ASVs were identified by both frequency analysis and IndVal. Frequency analysis captured a broader suit of core taxa, while IndVal was better at identifying host-specific microbes. Many *Fucus*-core ASVs, particularly within *Granulosicoccus* and *Litorimonas* are found on diverse seaweed species, while we find evidence of specificity to brown algae or *Fucus* for a few ASVs within *Blastopirellula* and *Rubritalea*.

Conclusions

We identified a suite of core taxa that are consistently associated with *F. distichus* and enriched in comparison to the environment, despite variation in the prevalent and predominant taxa on *F. distichus* over space and time. Moreover, we show that most of these core ASVs of *F. distichus* are found on diverse seaweed hosts, indicating that most occupy a seaweed generalist niche rather than forming highly specialized associations with *F. distichus*. Further studies should test whether seaweed generalists or specialists are more likely to engage in biologically important exchanges with host.

Background

There is an increasing recognition that bacteria closely associated with hosts play important roles in host development, survival, and fitness [1, 2]. With the advent of high-throughput sequencing techniques, the interest in understanding ecological and evolutionary relationships with bacteria in a variety of host systems has been explosive in multiple contexts, including epidemiology, species conservation given climate change, and microbial manipulation to improve crop yields [3–6]. What taxa constitute...
functionally important core microbiome of a host is more obvious in obligate symbioses where symbionts and their host are consistently engaging with each other, such as for bobtail squid and *Allivibrio fischeri* [7, 8] and for siboglinid tube worms and endosymbiotic bacteria [9]. However, other hosts have a much greater degree of variability in their microbial associates; the microbiota associated with plants and macroalgae is generally comprised of hundreds to thousands of species with spatial and temporal variation across environmental gradients [10–13]. The complex nature of the microbiota in these variable systems complicates the task of identifying meaningful relationships between host and microbe despite the large amount of microbiome data being generated for them.

In variable host-microbe systems, a core microbiome approach might be valuable to identify which bacterial taxa are potentially important for host biology [14, 15]. A core microbiome generally refers to a set of microbes consistently associated with a given host [15]. Their consistent presence, often with high abundance, is thought to be a product of evolutionary and ecological processes that govern host-microbe interactions [15]. Recent studies suggest that core taxa can predict animals’ health/disease states [16] and be used as targets for manipulation to improve crop yields or resilience in sustainable agroecosystems [17]. The rationale for spotlighting a small number of microbial taxa that are the most likely to be important symbionts is clear [18–20], but there are many and varied approaches for identifying the core. Across studies, core taxa have been identified cross-sectionally [21–23] and/or longitudinally [24] by establishing frequency thresholds. However, these thresholds also vary across studies, from 50% frequency in seagrasses [25] to 80% in corals [26] to 90% in amphibians [27]. Other studies require microbes to be specifically associated with a host, defined as enriched compared to the background environment, in addition to high frequency to qualify as core [28]. Such inconsistent methods and flexible parameters have called into question the robustness of a core microbiome approach [29]. Taking a more conservative approach to defining the core by considering enrichment compared to the environment and/or using broad sampling across populations and over time, will likely improve the utility of the core microbiota approach, particularly for host systems with highly variable microbiomes.

The epiphytic bacteria on macroalgae, and surface bacterial communities in general, tend to be heavily influenced by environmental conditions [30, 31] and often composed of functionally redundant taxa [32]. Yet, within this variability a suite of taxa frequently found on seaweed surfaces is emerging across studies, including Saprospiraceae, *Granulosicoccus*, and Flavobacteria [33–35]. Earlier studies reported a handful of core bacterial taxa consistently associated with particular populations in a few seaweed species, including *Fucus vesiculosus* [28], *Ulva australis* [32, 36] and *Ascophyllum nodosum* [37]. However, these studies were limited in scope, focussing on one population at one point in time, often without corresponding sampling of the seaweeds’ environment. Thus, outstanding questions include whether core bacteria on seaweeds are specific to host species or macroalgal clades [38] and whether they are maintained over time in multiple populations against environmental variations in bacterial communities [31].

A wide range of beneficial and detrimental interactions occur between macroalgae and their associated bacteria [39, 40]. These interactions primarily occur at the seaweed surface, which is the physiological
and ecological interface with marine bacteria and is involved in the exchange of nutrients and chemical signals [41]. The microbes involved are known only in a few striking examples, such as *Ulva mutabilis* and growth-promoting Proteobacteria (i.e., *Roseobacter*, *Sulfitobacter*, and *Halomonas*) [42]. Several bacterial taxa that directly alter seaweed growth or development have also been identified by culture-based studies. (e.g., *Pseudomonas* [43], *Rhodopseudomonas* [44] and *Pseudoalteromonas* [44]), and these appear to be weedy species that are widely found on marine surfaces [45–47]. Because of important microbial roles in macroalgae, identifying core taxa, particularly among variable bacterial communities associated with macroalgae promises to improve our understanding of core microbiome and enable further investigation of bacterial roles in seaweed hosts, including determining whether the bacterial taxa most likely to influence seaweed biology and physiology are part of the core or not.

Marine macroalgae are extremely diverse globally and belong to three evolutionarily distinct macroalgal clades (Rhodophyta, Chlorophyta, Phaeophyceae) that have converged ecologically [48]. Earlier studies have observed distinct epiphytic bacterial community structures across diverse seaweed species [49]. This is likely to reflect the fact that differences in morphology, chemistry, and habitat result in host-species specific niches that shape bacterial communities. Seaweed species have an extensive chemical defense system against grazers and pathogens [50, 51]. For example, *Fucus* species (Phaeophyceae) produce a variety of defensive secondary metabolites, such as phlorotannins and fucoxanthin, that likely represent selective filters for the bacteria that colonize these macroalgae [52–55]. Seaweeds also exude various polysaccharides that bacteria feed upon [56], and in some cases function as antibiotics to protect the host against pathogens of fouling organisms [57, 58].

In this study, we investigate the core microbiome of a focal brown macroalgae, *Fucus distichus* by combining multiple datasets that encompass a wide range of spatial and temporal sampling schemes. We use a simple frequency threshold as well as indicator species analysis (hereinafter referred to as “IndVal”) as a tool to capture core bacterial taxa that are constantly prevalent and abundant on *F. distichus* across host populations and over time. While the frequency method is based on prevalence data from the seaweeds alone, the IndVal method identifies core taxa based on their enrichment compared to environmental samples (i.e. water and rock) acting as a control. We conduct these analyses at the amplicon sequence variant (ASV) level to better understand the distribution of common taxa at the finest resolution possible. We first ask 1) whether there are bacterial ASVs consistently associated with multiple *F. distichus* populations taken in different months and years. We then ask 2) how these conservatively defined core ASVs using IndVal for a particular time and place differ from the suite of core ASVs defined by a simple frequency threshold. We further asked 3) whether these core ASVs are specifically associated with *F. distichus* or are general colonizers on diverse macroalgal species and 4) if the core ASVs are part of unique macroalgal-associated clades or have been broadly characterized from other hosts and abiotic environments by building phylogenetic trees.

**Methods**

**Dataset description and study design**
We analyzed four 16S rRNA gene amplicon datasets of the *F. distichus* microbiome, along with neighboring environmental samples from multiple intertidal locations on Calvert Island and Quadra Island, BC, Canada [Table. 1]. These four datasets were originally collected to address other research questions and present an opportunity to identify core bacterial taxa on *F. distichus* across multiple sites and timepoints. One dataset surveyed epiphytic microbiota associated with 36 sympatric seaweed species, including *F. distichus* on Calvert Island in 2015 [49] and two other datasets surveyed 5 and 4 populations of *F. distichus* on Calvert Island in 2018 [13] and 2019 (described herein) respectively and the last dataset is a longitudinal timeseries of *F. distichus* on Quadra Island from 2017 March through 2018 January (Davis submitted). The combined dataset comprised of a total of 1,106 samples.

**Sample collection**

Microbial DNA samples were obtained from the surface of apical tip (new growth of thallus; meristem tissue) of *F. distichus*. This area was targeted for sampling because the meristem tissues of macroalgae is younger and less subjected to fouling, and therefore represent a more selective microbial environment. In kelp (a related brown algae), the bacterial communities on meristem tissue have been shown to be more consistent over time than on older blade tissues [59]. *Fucus distichus* individuals were rinsed with 0.22 µm filtered sterile seawater for 10 seconds to remove loosely associated microbes and then swabbed with a Puritan® sterile swab for 15 seconds. Swabs were immediately stored in 2 mL cryovials (VWR). Bare rock substrates near where *F. distichus* were growing were sampled as a comparison to non-host associated microbial communities. Water column samples were also collected from adjacent seawater at each sampling site to characterize microbial source pool communities by filtering seawater onto a 0.22um Millipore Sterivex™ unit. DNA samples were stored at -80°C until DNA extraction. Microbial DNA on diverse sympatric seaweed species at West beach, Calvert Island, BC were also sampled using the surface swab method, typically from meristem tissue but meristem samples could not be taken for all species [49].

**Molecular methods**

DNA was extracted from swabs and water filter for 16S rRNA amplicon sequencing using MoBio PowerSoil Kit (QIAGEN), following the manufacturer’s recommended protocols. PCR amplification for bacterial DNA targeted the V4-V5 region of the 16S rRNA gene using primers, 515f: 5’–GTGYCAGCMGCCGCGGTAA–3’ and 806r: 5’ – GGACTACHVGGGTWTCTAAT – 3’ [60]. These primers included Illumina adapters and the forward included a 12 nucleotide Golay barcode. Then, we carried out amplicon library preparation, including PCR, quantification using Quant-IT Pico Green® ds DNA Assay Kit (Life Technologies) and pooled equal volumes (25 ng) of each sample, followed by purification using MoBio UltraClean® PCR clean-up kit. Sequencing with Illumina MiSeq using paired-end (2 x 300 bp) v3 chemistry were performed at the Integrated Microbiome Resource (IMR), Centre for Comparative Genomics and Evolutionary Bioinformatics (CGEB) at Dalhousie University according to published protocols [61].

**Bioinformatics**
Raw Illumina reads were demultiplexed in pairs using the idemp tool [62] without barcode error. Merging sequences from four datasets, quality filtering, trimming, dereplication, chimera removal, inference of true amplicon sequence variants (ASVs), and taxonomic assignment against the SILVA 128 [63] database clustered at 99% similarity [64] were processed with DADA2 pipeline [65] in R environment. In the process of filtering, we discarded ASVs if they were less than 0.1% of total number of reads or found in less than 5 samples. Overall, we obtained 6,348 amplicon sequence variants after filtering. The final products were then converted into phyloseq format in R for the downstream analysis. We then rarefied final 16S amplicon sequence products to 1,500 reads per sample prior to beta-diversity analysis. We used nonmetric multidimensional scaling (NMDS) to visualize all samples included in our study based on Bray-Curtis dissimilarity [66]. We also conducted permutational analysis of variance (PERMANOVA) using adonis2 by margin in the vegan package [67] in R to test for differences among bacterial communities between *F. distichus* and environmental samples as well as across sample sites. All PERMANOVA statistics were generated with 9,999 permutations.

**Identifying *Fucus distichus*-core bacteria**

To identify core bacterial taxa, we used a 50% frequency (prevalence) threshold and indicator species (IndVal) analysis using multipatt function within ‘indicspecies’ package [68] in R [69]. IndVal analysis captures bacterial ASVs based on both specificity (a measure of relative abundance compared to environmental rock and seawater samples) and fidelity (a measure of prevalence on *F. distichus*) of bacterial species to *F. distichus* samples [70, 71]. Permutation tests were then used to evaluate the statistical significance of the bacterial association for *F. distichus*. These IndVal analyses were separately performed for each location and time point to prevent biases from unequal sample sizes and to identify indicators of *F. distichus* at each site and time point while a 50% frequency threshold was simply applied to find bacterial ASVs that occur in greater than 50% of all the *Fucus* samples. Bacterial ASVs at greater than 0.7 IndVal value for each site and time point were considered as core candidates. We then determined *F. distichus*-core bacterial taxa by evaluating the core candidates of at least 7 of 11 sites/time points from four different datasets. We then visualized average relative abundance and prevalence of core ASVs using ggplot2 [72] in R to show their distribution and specificity to *F. distichus* hosts compared to environmental samples.

**Phylogenetic analysis**

To understand the habitat distribution of core bacteria and closely related taxa, we placed core bacterial sequences into phylogenetic trees using QIIME2 [73] with closely related sequences from the SILVA database [63] and NCBI, identified using BLAST [74]. We reformatted our data to fasta files for the QIIME2 environment using the ShortRead and seqinr package in R. Sequences were aligned and the alignment masked to contain only alignment columns that are phylogenetically informative in q2-phylogeny pipeline in QIIME2 [73]. Phylogenetic tree files were then constructed by using RAxML rapid bootstrap method (replicates = 100) with GTRCAT model. We assessed the habitat distribution of close relatives of core bacteria. We used data from the GenBank records to annotate the habitat and the identity of the host
where they were isolated from. Phylogenetic tree visualization and annotation were performed in Interactive Tree of Life (iTOL) v4 [75].

**Results**

**Bacterial communities of* F. distichus across the entire dataset**

Our results from four datasets showed bacterial communities on *F. distichus* significantly differed from environmental samples (rock surface and seawater; Adonis2: Pseudo-F(2:740) = 66.009, R² = 0.15, p < 0.001) [Fig. 1A]. This indicates *F. distichus* host unique microbiomes, but within *F. distichus* samples there were significant differences by site and time point [Fig. 1B]. The dispersion within the largest dataset, “2017 Quadra” is a result of seasonal variation in bacterial communities across this one-year timeseries.

**Core bacterial taxa of* F. distichus**

To identify which taxa are tightly associated with *F. distichus*, we first used a two-step indicator species analysis (IndVal) approach to identify the core bacteria. In step 1 we identified bacterial ASVs that were enriched on *F. distichus* and prevalent across individuals at each site and time point (indicator taxa) at a threshold of > 0.7 index value. Then in step 2 we asked which of these indicator taxa were consistently identified across sites and time (core taxa), using the criteria that an ASV had to be an indicator taxon in all datasets and in at least 7 of 11 sampling events. In step 1 we identified a total of 263 ASVs within 76 bacterial genera and found that indicator taxa are highly variable across sites and time points [Supplementary Table 1]. In step 2, we identified 15 core bacterial ASVs. These 15 core ASVs belong to the genera *Granulosicoccus, Blastopirellula, Litorimonas, Rubidimonas, Hellea, Roseibacillus, Rubritalea,* and *Nonlabens* [Supplementary Table 2]. Interestingly, no ASVs were present in every sample of *F. distichus* across sites and over time, although these taxa meet our definition of core as they are significantly prevalent and predominant in *F. distichus* samples across all datasets.

We compared the IndVal method to the commonly used frequency threshold method, which does not consider the occurrence of host-associated bacteria in the surrounding environment. At a frequency threshold of 50% we identified 28 ASVs as *F. distichus*-core taxa across our datasets. The choice of a threshold is somewhat arbitrary across core microbiome studies, and we present the frequency (prevalence) of each ASV that was an indicator in at least one sampling event in Supplementary Table 2 for ease of comparison. For example, 80% frequency threshold captured 3 taxa (ASV1, ASV2 and ASV5) that are the most dominant taxa in our datasets (Supplementary Table 2). We found that the core identified using a simple frequency threshold of 50% resulted in overlap of 11 ASVs identified by our IndVal analysis. As expected, a simple frequency method also captured taxa (e.g., ASV10, ASV41, ASV43 and ASV52) that are also prevalent in environmental samples [Fig. 2]. On the other hand, our two-step IndVal approach detected 4 core ASVs that are highly specific to *F. distichus*, but less prevalent in the environment, overall [Fig. 2–4].
Core bacteria of *F. distichus* over time

We further investigated the “2017 Quadra” dataset longitudinally sampled between 2017 March and 2018 January to ask whether the core bacterial taxa are persistently present and stable in relative abundance over this seasonal timeseries. The full Quadra 2017 dataset was used to identify the core, so it was expected that our core bacteria are present in this dataset. Our results show that the relative abundances of most core ASVs fluctuate over time, but a few are stable [Fig. 3]. For example, *Blastopirellula* ASV2 and *Granulosicoccus* ASV4 were strikingly overrepresented on *F. distichus* between March and June but decreased in relative abundance during the summer season (July-October), while *Granulosicoccus* ASV1 was stable in relative abundance across all months. *Dokdonia* ASV6 identified by a simple frequency threshold of 50% was also relatively stable across all seasons, and commonly found on rocks. We also observed that a few of core taxa (ASV72, ASV84, ASV126, ASV130) by IndVal were nearly absent (< 0.1%) in *F. distichus* samples collected between July and January [Fig. 3]. Because the 16S amplicon data used in this study is compositional, it is not clear if changes in the prevalence of core taxa are driven by decreases in absolute abundance or influenced by blooms of other taxa on *F. distichus*.

Macroalgal host specificity of *F. distichus*-core bacteria

We then asked whether the *F. distichus*-core ASVs associate predominately with *Fucus* (*Fucus* specialists) or with diverse seaweed species (seaweed generalists). We assessed the specificity of the *Fucus* core bacterial ASVs by determining their distribution on *F. distichus* and 35 sympatric seaweed species sampled at West Beach, Calvert Island, BC [Fig. 4]. Bacterial communities were found to vary among the diverse sympatric seaweed species [49], presumably because these macroalgae differ in their morphology, production of secondary metabolites, and physiological functions which can act as selective filters for different bacterial communities [76, 77]. Our data showed that many of *F. distichus*-core bacterial taxa are still present on other seaweed species in sympatry, but a few were exclusively enriched on *F. distichus* [Fig. 4]. For instance, *Granulosicoccus* sp. (ASV1 and ASV4) and *Litorimonas* sp. (ASV5, ASV23 and ASV72) were found on most sympatric brown, green, and red macroalgae at West Beach [Fig. 4]. On the other hand, core bacterial taxa *Rubritalea* sp. (ASV3) and *Roseibacilus* sp. (ASV67 and ASV169), were specifically enriched on *F. distichus* and nearly absent on other seaweeds. *Blastopirellula* sp. (ASV2) was associated with seven macroalgal species and five of them were brown algae, suggesting a possible niche within brown macroalgae. Because the sample size for each seaweed species is uneven, we must be cautious in interpreting differences in their relative abundance and prevalence of the core ASVs. However, it is clear that many *F. distichus*-core taxa identified across our datasets associate with diverse macroalgae and are likely macroalgal generalists, while only *Rubritalea* sp. (ASV3) is a candidate for being a *Fucus* specialist.

Phylogenetic placement of core bacteria

We used broadly sampled phylogenetic trees to identify the closest relatives of the *F. distichus*-core ASVs. We annotated the tree with information on the isolation source from GenBank reporting the environment
or host or from which sequences were isolated. This allowed us to determine whether the *F. distichus* core fall within macroalgal-associated clades or clades that have been broadly characterized from other hosts and abiotic environments. In cases where we identified multiple core ASVs belonging to the same genus, this allowed us to determine whether the core ASVs are clustered together or fall within distinct clades.

**Granulosicoccus**

We found the three core ASVs within *Granulosicoccus* genus fall within distinct clades [Supplementary Fig. 1]. The closest relatives of ASV4 and ASV130 were previously detected on brown macroalgae. On the other hand, close relatives of ASV11 and ASV1 were isolated from the environment (i.e., marine sediment and ice) [Supplementary Fig. 1].

**Litorimonas**

Closely related bacteria of core ASVs within *Litorimonas* were mostly detected with brown and green macroalgae, although a few of them were still associated with environment (i.e., seawater and marine surface) [Supplementary Fig. 2]. Core ASVs within *Granulosicoccus* and *Litorimonas* genus are strong indicator bacteria of *F. distichus* and widespread on diverse seaweed species in our dataset. This is likely to reflect that they are generally well-adapted on seaweeds but facultative bacteria that can be acquired from surrounding environments.

The closest relatives of *Blastopirellula* (Planctomycetes), core ASV2 were particularly detected with *Fucus* species [Supplementary Fig. 3], supporting our results that ASV2 maybe specialist on *Fucus* or brown algal clade. Another specialist, *Rubritalea*, core ASV3 is one that enriched in *F. distichus* samples but nearly absent in other 34 seaweed samples. We found that a sequence that closely match core ASV3 and ASV38 was also found on *Fucus vesiculosus* [Supplementary Fig. 4]. Overall, the core taxa, ASV2 and ASV3 maybe specialists on *Fucus* or brown algae, respectively, but we could not find evidence of codiversification with seaweed hosts from the phylogenetic trees of the *F. distichus*-core bacteria because host and symbiont phylogenies are not concurrent.

**Discussion**

Identifying core taxa offers a way to simplify the complex microbial community associated with a host. In this study, we used an expansive dataset to define the core microbiome of *Fucus distichus* by indicator species analyses (IndVal) and using a simple frequency (prevalence) threshold of 50%. Because indicator taxa are identified for each sampling event, IndVal also captures variation of main components in host-associated community over space and time.

Comparison to the environment facilitates the detection of taxa that preferentially associate with the host and differentiates them from taxa that are widespread on a host and generally abundant in the environment. Our IndVal analysis showed that indicator bacterial ASVs of *F. distichus* varied across host populations. This suggests that local environmental factors have a large impact on which taxa are
prevalent and predominant on seaweed hosts. Hence, identifying core taxa based on frequency and abundance from one site at one point in time [28, 32, 36] provides limited information about the taxa that are truly widespread across host populations and over time. We showed that a simple frequency method with an arbitrary threshold (50%) captured a broad suite of bacteria associated with \textit{F. distichus}, including those that are also broadly found in the surrounding environment, whereas the IndVal analyses filtered out taxa that are commonly present in environmental samples (water and rock). Defining core bacteria using an approach that considers host specificity compared to the surrounding environment does not guarantee functional relevance or a specific relationship with host but can provide a more robust core membership within variable microbiota by excluding local environmental noise.

The method and threshold for including ASVs in the core microbiome will likely vary depending on the researcher’s aims and available data. It is not possible to determine which is the ‘best’ method because the core bacteria are effectively a suite of candidate bacteria that can be investigated to determine their functional roles in association with the host. Likewise, the distribution of core bacteria and their fidelity across time and space surely vary for different hosts. While a full exploration of this parameter space is outside the scope of this analysis, we present a comparison of methods and summary statistics in the supplement that can be used to explore the impact of different methods and thresholds. For instance, 13 additional bacterial ASVs are core of \textit{F. distichus} if we set a threshold to ASVs that are IndVal indicators in all 4 datasets (without considering the number of sampling events) [Supplementary Table 2]. On the other hand, no ASVs would meet a stringent threshold of > 90% frequency across the entire dataset [Supplementary Table 2], which includes samples from all seasons and five distinct sites. A more relaxed frequency threshold of > 50% captured many of the same taxa as the more conservative IndVal approach (11 of 15 IndVal core), but picked up a broader suit of bacterial taxa, particularly more those shared with the environment that are filtered out by IndVal. Both IndVal and frequency thresholds appear to be useful methods to identify core taxa that can be further tested for functional importance as symbionts. We argue that a broad survey on host populations over time is more critical than differences in methods and thresholds since highly prevalent taxa and abundant in host samples are often obvious.

We find that no core bacterial taxa are always present on \textit{F. distichus} and that indicator taxa vary across host populations. This suggests that local population dynamics among diverse but functionally redundant ASVs could allow different microbes adapted to different abiotic conditions to fill host-associated niches across sites and seasons. Functional redundancy is common in marine microbial systems [32, 78, 79], and in the plant microbiome system [80]. Indeed, many Proteobacteria (i.e., \textit{Roseobacter} sp., \textit{Sulfitobacter} sp., and \textit{Halomonas} sp.) and Bacteroidetes (i.e., \textit{Maribacter} sp.) have been shown to stimulate morphogenesis of the green seaweed \textit{Ulva}, and are substitutable [42, 81, 82]. These bacterial genera associated with \textit{Ulva} morphogenesis are commonly found in various marine environments [83, 84], although they are unlikely the same strain. Overall, the \textit{F. distichus}-core taxa may be substitutable component of microbiomes if they are biologically functional. Next research steps should include cultivating core and non-core members of the seaweed microbiome to test the prevailing hypothesis that core bacteria play crucial roles in biology and functions [15] within this context of many functionally redundant seaweed-associated bacteria.
Host species specificity of microbial symbionts has been a long interest in many host microbiome systems [85–87], including seaweeds [88]. It is commonly thought that symbionts evolve to compete within the host ecosystem and microbial specificity to a particular host species is likely the product of evolutionary processes of host microbiomes [89]. Our core microbiome study provides a great opportunity to ask whether the core ASVs identified here are host species specific. There are bacterial genera commonly found on diverse seaweed species across studies, yet it is unclear whether they are the same strain. For example, the most obvious *Fucus* core genera in this study were *Granulosicoccus* and *Litorimonas*, which are also dominantly found in the broad range of brown, green and red seaweed species; *Fucus* species [90, 91], *Nereocystis luetkeana* [92], *Laminaria setchellii* [93], *Macrocystis pyrifera* [34], *Ecklonia cava* [94], *Caulerpa cylindracea* [95], *Ulva rigida* [96], *Porphyra umbilicalis* [97] and *Gelidium lingulatum* [98]. Our study shows a large portion of identifiable core bacteria of *F. distichus* are seaweed generalists at the ASV level, contradicting the prediction that bacterial genera frequently associated with seaweeds are host-specific at the strain level. For instance, most sympatric seaweed species shares the *F. distichus*-core ASVs within *Granulosicoccus* and *Litorimonas* [Fig. 4]. This suggests the generality at the ASV level is common and also the bacterial genus repeatedly found in different seaweed species are possibly identical strain.

It is not surprising that seaweed generalists exist because several cell wall polysaccharides are widely distributed across green, red and brown seaweeds [99, 100]. Specific marine bacterial strains may colonize on the surface of diverse seaweeds for the degradation of the polysaccharides from macroalgae [101, 102]. In addition, a possible explanation for seaweed generalists is generalized signal molecule, such as dimethylsulfoniopropionate (DMSP) [103–105] produced by most macroalgal species. Many marine bacteria can be attracted by DMSP as a reliable signal indicating a food source [106]. DMSP can also be used directly as a food source. The ability to catabolize DMSP in bacteria, such as *Granulosicoccus* and *Litorimonas* [107, 108] could promote their ability to colonize diverse seaweed species.

Our study also showed that a few core ASVs are somewhat specific to *F. distichus* or the brown algal clade. For instance, *Blastopirellula* (ASV2) and *Rubritalea* (ASV3) were not commonly associated with other seaweed species. Interestingly, the closest relates of ASV2 were isolated only from macroalgae; 3 *Fucus* species [109, 110] and 1 kelp [34] shown in the phylogenetic tree [Supplementary Fig. 3]. This suggests that this clade of *Blastopirellula* may have a distinct evolutionary history of association with *F. distichus* and/or brown macroalgae. The core taxon within *Rubritalea*, ASV3 was particularly abundant on *F. distichus* in this dataset but we could not find previous reference falling within the same clade in the phylogenetic tree [Supplementary Fig. 4]. Although another *Rubritalea* taxon, ASV38 identified by a simple frequency threshold was closer to previous isolate from *Fucus vesiculosus*, it was not so much specific to *F. distichus* in our dataset.

**Conclusion**
Using both cross-sectional and longitudinal datasets, we identified a suite of core taxa that are frequently associated with the seaweed *Fucus distichus*, but none of them are always present. Different approaches to identify a core yield an overlapping suite of core taxa, those that are most prevalent and abundant. Strong variation across populations and time points indicates that local and seasonal environmental factors have a large impact on which taxa are prevalent and predominant on seaweed hosts. This study shows a large portion of identifiable core bacteria of *F. distichus* are seaweed generalists at the ASV level, contradicting the prediction that bacterial genera frequently associated with seaweeds are host-specific at the strain level. Still, we do not know if these core bacteria are functionally important for the seaweed host, and further studies should test if these seaweed generalists are biologically important symbionts.

**List Of Abbreviations**

ASV
Amplicon sequence variants

IndVal
Indicator species analysis

DMSP
Dimethylsulfoniopropionate

NMDS
Non-metric multidimensional scaling

PERMANOVA
Permutational analysis of variance

NB
North Beach site on Calvert Island

PB
Pruth Bay site on Calvert Island

WB-L
West Beach at low tide on Calvert Island

WB-H
West Beach at high tide on Calvert Island

**Declarations**

**Ethics approval and consent to participate**

Not Applicable

**Consent for publication**

Not Applicable

**Availability of data and materials**
All datasets within phyloseq format and bioinformatic scripts from this study are available at [https://github.com/Jungsooparkubc/Fucus-Core-Microbiome](https://github.com/Jungsooparkubc/Fucus-Core-Microbiome). All relevant data for Indicator Species Analysis are within the main paper and its supplementary files.

**Competing interests**

The authors declare that they have no competing interests.

**Funding**

J.P was supported by the UBC International Doctoral Fellowship (IDF). K.M was supported by a UBC Four Year Fellowship. This work was funded by an NSERC Discovery grant to L.W.P and a grant from the Tula Foundation to L.W.P.

**Author's Contributions**

J.P and L.W.P conceived the study. K.D and J.P collected samples on Calvert Island in 2018 and 2019 respectively and performed molecular methods. K.D contributed to Quadra 2017 dataset used for this core microbiome study. J.P combined datasets and designed the figures following computational data analyses. J.P wrote the manuscript with input from all authors. L.W.P supervised the project. All the authors reviewed and approved the manuscript.

**Acknowledgments**

We acknowledge that this research was carried out on the traditional, ancestral, and unceded territory of the Coast Salish peoples – Sḵwx̱wú7mesh (Squamish), Stó:lō and S̱əl̓ílwəta /Selilwitulh (Tsleil-Waututh) and x�kosəm (Musqueam) Nations. The authors would like to thank M. Lemay for Calvert WestBeach 2015 dataset.

**References**


90. Quigley CT, Capistrant-Fossa KA, Morrison HG, Johnson LE, Morozov A, Hertzberg VS, et al. Bacterial Communities Show Algal Host (Fucus spp.)/Zone Differentiation Across the Stress Gradient of the Intertidal Zone. Frontiers in microbiology. 2020;11:563118-.


Tables

**TABLE 1 | Overview of samples included in multiple datasets.** Microbial samples that are included in our study sorted by sample type and geographical location.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Data type</th>
<th>Sample type</th>
<th>North Beach (n)</th>
<th>Pruth Bay (n)</th>
<th>West Beach high-tide zone (n)</th>
<th>West Beach low-tide zone (n)</th>
<th>West Beach wall (n)</th>
<th>Quadra (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2015 Calvert</td>
<td>Cross-sectional</td>
<td>Fucus distichus</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>15</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Diverse seaweeds*</td>
<td>0</td>
<td>0</td>
<td>38</td>
<td>208</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rock and seawater</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>50</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2017 Quadra</td>
<td>Longitudinal</td>
<td>Fucus distichus</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>247</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rock and seawater</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2018 Calvert</td>
<td>Cross-sectional</td>
<td>Fucus distichus</td>
<td>24</td>
<td>28</td>
<td>22</td>
<td>24</td>
<td>23</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rock and seawater</td>
<td>3</td>
<td>25</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>2019 Calvert</td>
<td>Cross-sectional</td>
<td>Fucus distichus</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rock and seawater</td>
<td>6</td>
<td>6</td>
<td>3</td>
<td>6</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

*Include 38 other seaweed species

Figures

**Figure 1**

NMDS plots constructed from Bray-Curtis dissimilarities of microbiota on *Fucus distichus* and from the neighbouring environments (rock substrate and seawater). (A) Bacterial community composition differs between *F. distichus* and environmental samples. (B) *F. distichus* microbiota varies across sites and different time points.
Figure 2

Average relative abundance and prevalence of 32 core ASVs determined in this study across sites in different time points. Core ASVs identified by our two-step IndVal analysis with a set threshold; >0.8 index and ≥ 7 times of IndVal detection from each sampling event are colored blue, while core ASVs identified by a simple frequency threshold of 50% are colored green. Core ASVs are identified by both methods are colored black. IndVal method detected core taxa that are highly specific to *F. distichus*, while a simple frequency threshold detect a broader suite of bacteria, including more weedy taxa that are found in environmental samples.
Average relative abundance and prevalence of *F. distichus*-core bacteria over time in the 2017 Quadra dataset. Core ASVs identified by our two-step IndVal analysis with a set threshold; >0.8 index and $\geq 7$ times of IndVal detection from each sampling event are colored blue, while core ASVs identified by a simple frequency threshold of 50% are colored green. Core ASVs are identified by both methods are colored black. Most core taxa are repeatedly associated with *F. distichus* but dynamic in relative abundance over seasonal timeseries.
Figure 4

Average relative abundance and prevalence of *F. distichus*-core bacteria on 36 seaweed species in the 2015 WestBeach dataset. Core ASVs identified by our two-step IndVal analysis with a set threshold; >0.8 index and ≥ 7 times of IndVal detection from each sampling event are colored blue, while core ASVs identified by a simple frequency threshold of 50% are colored green. Core ASVs are identified by both methods are colored black. In general, *Fucus* core taxa within the dataset are commonly found on diverse seaweeds (e.g., *Granulosicoccus* ASV1 and ASV4, *Litorimonas* ASV5, ASV23 and ASV72 etc.) but a few of taxa are somewhat specific to *F. distichus* (e.g., *Blastopirellula* ASV2 and *Rubritalea* ASV3 etc.).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- FucusCoreSupplementaryTable1.xlsx
- SupplementaryFigures.docx
- FucusCoreSupplementaryTable2.xlsx