environmental microbiology

Environmental Microbiology (2022) **24**(8), 3500–3516 doi:10.1111/1462-2920.15993

Sarracenia pitcher plant-associated microbial communities differ primarily by host species across a longitudinal gradient

Jacob A. Heil,[1] Charles J. Wolock,[2] Naomi E. Pierce,[3] Anne Pringle[4,5] and Leonora S. Bittleston[1*]

[1]*Department of Biological Sciences, Boise State University, Boise, ID.*

[2]*Department of Biostatistics, University of Washington, Seattle, WA.*

[3]*Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA.*

[4]*Department of Botany, University of Wisconsin-Madison, Madison, WI.*

[5]*Department of Bacteriology, University of Wisconsin-Madison, Madison, WI.*

Summary

Plant-associated microbial communities can profoundly affect plant health and success, and research is still uncovering factors driving the assembly of these communities. Here, we examine how geography versus host species affects microbial community structure and differential abundances of individual taxa. We use metabarcoding to characterize the bacteria and eukaryotes associated with five, often co-occurring species of *Sarracenia* pitcher plants (Sarraceniaceae) and three natural hybrids along the longitudinal gradient of the U.S. Gulf Coast, as well as samples from *S. purpurea* in Massachusetts. To tease apart the effects of geography versus host species, we focus first on sites with co-occurring species and then on species located across different sites. Our analyses show that bacterial and eukaryotic community structures are clearly and consistently influenced by host species identity, with geographic factors also playing a role. Naturally occurring hybrids appear to also host unique communities, which are in some ways intermediate between their parent species. We see significant effects of geography (site and longitude), but these generally explain less of the variation among pitcher communities. Overall, in *Sarracenia* pitchers, host plant phenotype significantly affects the pitcher microbiomes and other associated organisms.

Introduction

Microbes are among the most abundant organisms on Earth and play key roles in nutrient and energy cycling; however, we are still learning about processes that govern microbial community assembly and biogeography (Shoemaker *et al*., 2017). Current research has examined the roles of factors such as host identity, abiotic constraints, dispersal and trophic interactions on microbial community structure and diversity (Nemergut *et al*., 2013; Prosser and Martiny, 2020). Community composition at any given time is a product of history and the development of the originally assembled community (Bittleston *et al*., 2020). Microbial communities assemble in ways similar to macrobial communities but vary due to microbe-specific characteristics in dispersal patterns, dormancy and phenotypic plasticity (Nemergut *et al*., 2013; Sorensen and Shade, 2020). Ecological patterns of bacterial communities can also differ strikingly from more familiar ones found in eukaryotic organisms (Logares *et al*., 2018; Gilbert *et al*., 2020a). Plant-associated microbial communities can be critical in promoting plant success; for example, through nitrogen fixation, plant growth promotion, plant-pathogen defence, and, in pitcher plants, through the breakdown of insect prey (Butler *et al*., 2008; Compant *et al*., 2019). In addition to abiotic environmental factors, plant-associated microbial communities are also influenced by phenotypic traits of the host plant (Leveau, 2019; Gilbert *et al*., 2020b).

Carnivorous pitcher plants present an excellent opportunity to study microbial community assembly, because the modified, pitcher-shaped leaves hold relatively small yet complex communities that can easily be sampled and analysed using techniques like environmental DNA metabarcoding (e.g. Bittleston *et al*., 2016). In this study, we aimed to characterize the communities from several pitcher plant species in the genus *Sarracenia* across a

Received 12 November, 2021; revised 18 March, 2022; accepted 31 March, 2022. *For correspondence. E-mail leonorabittleston@boisestate.edu; Tel. (208) 426-3530; Fax (208) 426-1040.

biogeographical gradient to determine how host species identity, spatial distance and pitcher characteristics influence the complete community structure of both bacteria and eukaryotes.

The pitchers of carnivorous pitcher plants hold 'phytotelmata' or 'plant-held waters' and trap prey, which are then digested and absorbed by the plant as a strategy for obtaining necessary nutrients (Schulze *et al*., 1997; Adlassnig *et al*., 2011). The microbiome of the pitcher fluid likely plays a key role in breaking down complex substrates, making nutrients available to the plant (Luciano and Newell, 2017; Koller-Peroutka *et al*., 2019). Beyond the bacterial microbiome, pitcher plant phytotelmata have been shown to host a multitrophic community of inquiline organisms including protozoa, rotifers, fungi and arthropods (Kitching, 2001).

The assembly history of the pitcher community and the life history of its host plant influence its present state (Koopman and Carstens, 2011; Bittleston *et al*., 2020). In the past 20 years, studies of the processes influencing community assembly and structure in carnivorous pitcher plants have increased dramatically. We refer to 'community structure' in this context to mean the richness and evenness of a sample (alpha diversity) and the composition across samples (beta diversity). One factor affecting community structure in pitcher microcosms is geographic variation (Buckley *et al*., 2003; Koopman and Carstens, 2011; Grothjan and Young, 2019; Gilbert *et al*., 2020a). Temperature fluctuation, specifically, has been shown to influence pitcher communities and can drive differences in inquiline identity (Zander *et al*., 2017). Other factors such as the fluid pH, fluid volume and pitcher morphology also affect community structure (Kanokratana *et al*., 2016; Bittleston *et al*., 2018; Gilbert *et al*., 2020b), as well as inquiline identity and associated ecological roles (Canter *et al*., 2018; Boynton *et al*., 2019).

Different findings emerge when analysing different scales. Geographic patterns are exhibited by the communities of a single plant host species on a single continent, while convergent patterns can be uncovered when comparing diverse communities from many host species on a global scale (Bittleston *et al*., 2018). Pitcher communities associated with the independently evolved *Nepenthes* and *Sarracenia* plants, separated by great geographic distance, house convergently assembled organisms from related groups such as the same bacterial families and specialized flies and mosquitos that appear to carry out similar functions within their respective pitcher habitats. Despite strong evidence that specific factors can influence pitcher community structure, the relative effects of interacting factors have not been systematically explored. Moreover, communities from different species in the genus *Sarracenia* have never been analysed in depth. Here, we expand on the comparisons among different *Sarracenia* species first assessed in Bittleston *et al*. (2018), and include communities from natural *Sarracenia* hybrids for the first time.

To understand factors affecting pitcher communities, we characterized samples from individual pitchers by identifying their constituents and quantifying community structure. DNA metabarcoding is an effective method that captures the relative abundances of different sequences (Taberlet *et al*., 2012; Deiner *et al*., 2017), and previous studies have outlined metabarcoding methods for characterizing pitcher communities (e.g. Bittleston *et al*., 2016, 2018). In this study, we used metabarcoding of the 16S ribosomal RNA gene to characterize the prokaryotes and of the 18S ribosomal RNA gene to characterize the eukaryotes. In most pitchers, virtually the entire community can be sampled by removing the liquid held in a pitcher and extracting DNA from it, although extraction issues and primer biases can limit the detection of certain groups (Vaginal Microbiome Consortium (additional members) *et al*., 2015). For each sample, we recorded host species identity, site where the plant was growing, longitude, latitude, and pitcher fluid pH and volume.

We compared communities from six different species of *Sarracenia*, a North American genus of pitcher plant composed of at least eight species that range from the Pacific Northwest, throughout Canada, the Eastern seaboard of the United States, and the Gulf Coast. *Sarracenia* is a relatively recent and rapid radiation (<3 million years old), with extensive hybridization among species (Ellison *et al*., 2012). Molecular phylogenies show discordance among different loci, making it difficult to identify species relationships (Ellison *et al*., 2012; Stephens *et al*., 2015). The most recent phylogeny of *Sarracenia*, using 199 nuclear genes, found that *S. alata* and *S. rubra* were part of a species complex, as were *S. purpurea* and *S. rosea* (Stephens *et al*., 2015, Fig. 1A). The study also found major disagreement with the placement of the *purpurea* complex when using concatenation versus coalescent approaches, and thus it is unclear whether the *purpurea* complex is more closely related to the *flava-minor-psittacina* clade or the *oreophila-leucophylla-alata-rubra* clade (Stephens *et al*., 2015).

Microbes associated with plant leaves and roots are known to be influenced by host plant identity (Edwards *et al*., 2015; Naylor *et al*., 2017; Leveau, 2019). We compared the microbiomes of five *Sarracenia* species (*S. alata*, *S. flava*, *S. leucophylla*, *S. rosea*, and *S. rubra*) and three hybrids (*flava-leucophylla*, *flava-rosea*, and *leucophylla-rosea*) along the Gulf Coast, and one species (*S. purpurea*) in Massachusetts. These samples enabled us to assess the relative roles of host species versus geographic location in shaping community structure of pitcher-associated bacteria and eukaryotes for

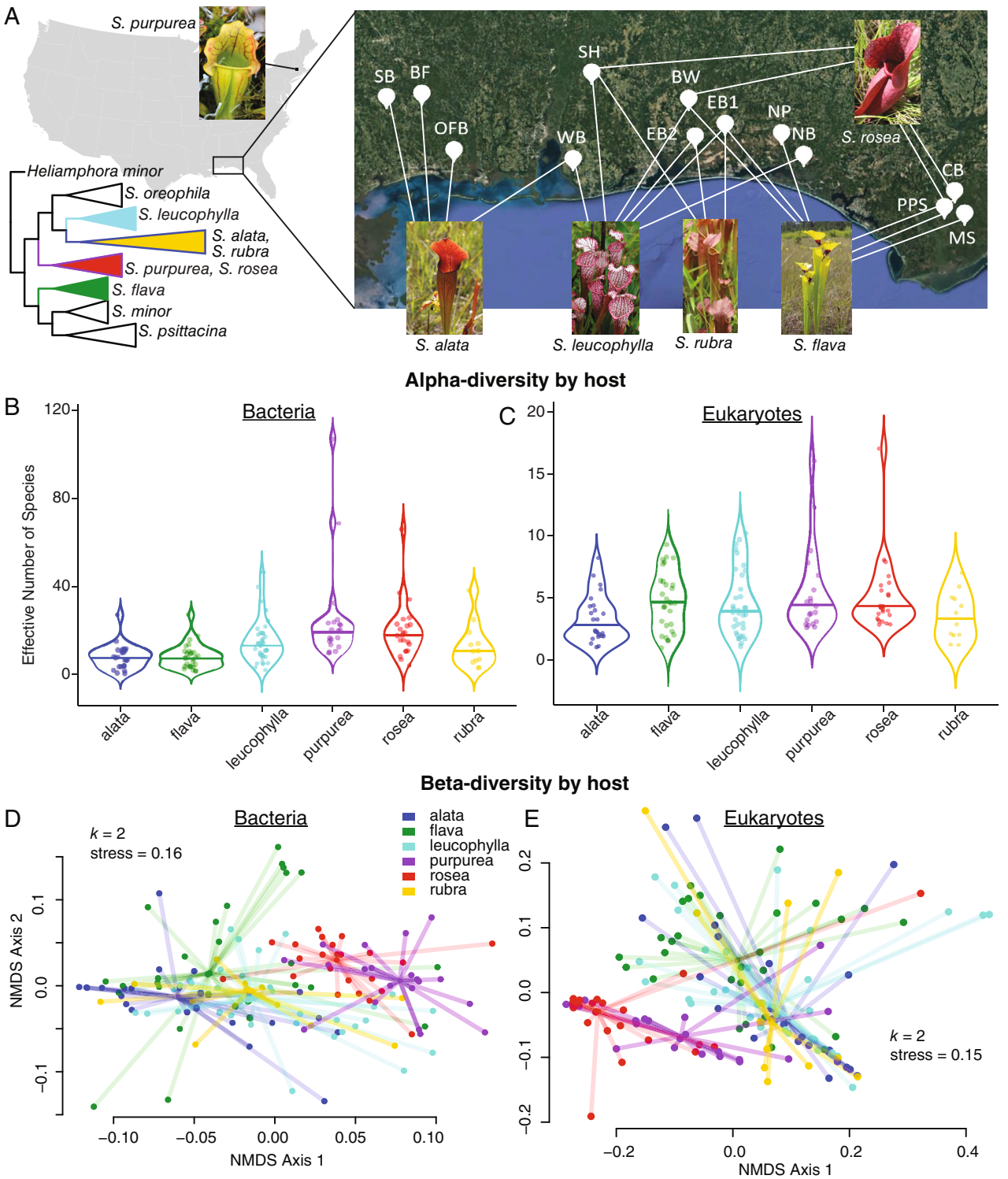

Fig. 1. Bacterial and eukaryotic communities differ across host species.
A. Locations of each host species across the sampling sites: Sweet Bay Bogs Preserve (SB), Buttercup Flats (BF), Old Fort Bayou (OFB), Week's Bay Pitcher Plant Bog (WB), Splinter Hill Bog (SH), Blackwater River State Forest (BW), Nokuse Plantation (NP), Nokuse Bog (NB), Mud Swamp New River Bog (MS), Crystal Bog (CB), Plea Phase Savanna (PPS), Eglin Airforce Base site 1 (EB1), Eglin Airforce Base site 2 (EB2), Tom Swamp (TS). The phylogeny of *Sarracenia* is modified from the RAxML concatenated tree from Stephens *et al*. (2015). Alpha diversity is shown as effective number of species for (B) bacteria and (C) eukaryotes. Beta diversity (using weighted UniFrac distances) is displayed in NMDS plots for (D) bacteria and (E) eukaryotes. Points are coloured according to host species, and come together at the mean of all samples from a single host species.

these six species and three hybrids. We hypothesize that host species is the main determinant of pitcher community structure, because host species likely influences relative fluid pH, volume and additional intrinsic factors (Gilbert *et al*., 2020b). A strong signature of host species would also support the corollary that plant phenotype is a strong determinant of pitcher community structure. Alternatively or additionally, we reason that if geographic location has a significant relationship with community structure, this would indicate that local environmental conditions and local dispersal from the existing regional species pool can also play an important role in shaping community structure. Geographic location can itself be parsed into the effects of longitude versus latitude, and any biogeographic gradients associated with longitude might further indicate whether relative differences in climate can influence these communities. Overall, our study provides insight into the assembly and structure of *Sarracenia* pitcher microbial communities and has broader implications for our general understanding of plant-associated communities.

Results

Community structure across host species

Bacterial community structure from our pitcher plant samples was primarily influenced by host species identity. We recorded 2720 unique bacterial amplicon sequence variants (ASVs) in the pitcher communities of six *Sarracenia* host species and three hybrids. Similar to previous *Sarracenia* studies (e.g. Koopman and Carstens, 2011; Grothjan and Young, 2019) we did not detect any archaea in our 16S results. We first investigated differences in alpha diversity, which was highest in the host species *S. purpurea* (measured as average effective number of species: 16.15, Fig. 1B). Because *S. purpurea* was only sampled in Tom Swamp in Massachusetts, and was discontinuous from the other samples, we left it out of our subsequent analyses of alpha diversity. Using *S. alata* and Blackwood Flats as references, we analysed variation in alpha diversity by host species, site, volume, pH, latitude and longitude using a generalized linear model (GLM). We found that, when controlling for all other factors, alpha diversity for *S. rosea* (z value $= 4.480$, $p < 0.001$) and *S. rosea x S. leucophylla* (z value $= 5.302$, $p < 0.001$) were significantly different from *S. alata* (Supplementary Table 3). Our omnibus tests also indicated that alpha diversity varied significantly among host species and among sites when controlling for other factors (both tests $p < 0.001$). All other sites and factors (pH, volume, latitude and longitude) significantly differed as well ($p < 0.05$), with the exception of the Weeks Bay site ($p = 0.178$). Among our 13 Gulf Coast sites, we found that latitude correlated most strongly with Annual Mean Temperature (linear model, $R^2 = 0.92$, $p < 0.001$) while longitude correlated most strongly with Annual Precipitation (linear model, $R^2 = 0.74$, $p < 0.001$).

Our bacterial beta diversity analyses showed a significant interaction between host species and site, which makes interpreting the individual effects difficult (Table 1). However, when measured alone, community composition varied significantly among host species (PERMANOVA; $R^2 = 0.227$, $p = 0.007$; Table 1); visualized with a non-metric multidimensional scaling (NMDS) ordination along two axes (Fig. 1D). The bacterial NMDS showed two distinct clusters around groups of host species: samples from *S. purpurea* and *S. rosea* clustered together more closely in the NMDS plot than samples from the other four species. Bacterial beta diversity also significantly differed by site (Table 1); however, the NMDS showed only minimal clustering by site, except for Tom Swamp in Massachusetts which had only a single species and was distant from the Gulf Coast sites (Supplementary Fig. 3). Using ANCOM, we identified 39 ASVs above the 0.8 threshold that were differentially abundant in communities from different host species. The bacterial ASVs exerting the strongest influence on this trend included: 16S2175 (Wstat = 98; family: Neisseriaceae, genus: *Aquitalea*), 16S0740 (ANCOM; Wstat = 98; family: Xanthomonadaceae, genus: *Rhodanobacter*), 16S1049 (Wstat = 98; family: Enterobacteriaceae), 16S2662 (Wstat = 98; family: Acetobacteraceae, genus: *Acidocella*). The significant ASVs showed high abundance associated with certain host species; for example, ASV 16S0740 assigned to the genus *Rhodanobacter* was present in at least one sample from every host with variable relative abundances among host species and the highest average relative abundance in *S. flava*.

Each sample consisted of both bacterial and eukaryotic communities. The interactions among these organisms likely influence their community structure, so we considered it essential to study both bacteria and eukaryotes simultaneously. We found 1413 unique eukaryotic ASVs in our samples. As in the bacterial results, eukaryotic alpha diversity was highest in *S. purpurea* (average effective number of species: 5.750, Fig. 1C). However, unlike the bacteria, GLM results showed no significant differences in alpha diversity among host species and sites, or for pH, volume, longitude and latitude (Supplementary Table 4). Our omnibus test for host species indicated that alpha diversity did not vary significantly among host species when controlling for other factors; however, the omnibus test for sites was significant ($p = 0.005$). In terms of eukaryotic beta diversity, we did not see a significant interaction between host species

Table 1. Beta diversity results for bacterial (16S) and eukaryotic (18S) pitcher communities across all host species (excluding hybrids) and sites using weighted UniFrac distances.

		Host species × site	Host species	Site		pH	Volume	Latitude	Longitude
16S	R^2	0.068	0.227	0.286	*r*	0.1	0.155	0.187	0.21
	p	0.007*	0.007*	0.007*	*p*	0.01*	0.008*	0.007*	0.007*
18S	R^2	0.046	0.294	0.224	*r*	−0	0.081	−0.032	−0.006
	p	0.185	0.007*	0.007*	*p*	1	0.185	1	1

The interaction of host species and site and then each factor alone were calculated with separate PERMANOVA tests. The continuous variables were calculated with separate Mantel tests. A star indicates significance at alpha = 0.05 for Holm-adjusted *p*-values.

and site, but each were significant alone (PERMANOVA; Host species $R^2 = 0.294$, $p = 0.007$; Site $R^2 = 0.224$, $p = 0.007$; Table 1). NMDS ordination showed a similar grouping as seen with the bacterial data, with samples from *S. purpurea* and *S. rosea* clustering together, separated from the samples from other host species (Fig. 1E). Also, like the bacterial data, the NMDS had minimal clustering by site (Supplementary Fig. 4). Using ANCOM, we identified 24 ASVs above the 0.8 threshold that significantly differed in abundance between host species. For example, ASV 18S0144 (Wstat = 23; class: Arachnida, family: Histiostomatidae) was identified as an astigmatid mite and was most abundant in *S. leucophylla* pitchers but was also found in *S. alata* and *S. rubra* pitchers; ASV 18S0204 (Wstat = 23; order: Diptera, family: Culicidae) was a mosquito only found in *S. rosea*; and ASV 18S0139 (Wstat = 23, order: Tremellales) was a fungus almost exclusively found in *S. alata*.

Community structure across host species within single sites

In order to investigate the influence of host-associated factors on community composition and diversity without the confounding effect of site, we divided the data from individual sites into subsets for comparison. Factors considered in measuring intra-site community diversity included host species, volume and pH. One of the possible drivers of diversity at the site level was host species richness (the number of host species at a single site), which ranged from three species plus a hybrid at Blackwater and Splinter Hill to a single species at several other sites (Fig. 1A). The identity of host species at particular sites showed a distribution that corresponded to longitude, reflecting the ranges of different *Sarracenia* species. For example, *S. alata* populated more western sites and *S. flava* was generally found in more eastern sites. Beta diversity analysis indicated that both bacterial and eukaryotic community composition significantly varied with host species at most of the sites with more than two co-occurring *Sarracenia* species or hybrids (Table 2). Bacterial communities differed among host species at Crystal Bog, Eglin Airforce Base Site 1, Plea Phase Savanna and Splinter Hill, while eukaryotic communities differed at Blackwater, Plea Phase Savanna and Splinter Hill (Table 2). Four of these five sites have *S. rosea*, which tends to host communities that are distinct from those of the other Gulf Coast *Sarracenia* species.

We chose to focus on one of our sites with the highest host species richness to understand more about factors differentiating bacterial and eukaryotic communities living in co-occurring pitchers from different host species (Fig. 2). The Splinter Hill site hosted the species *S. leucophylla*, *S. rosea*, *S. rubra* and the hybrid *S. leucophylla x S.* rosea. We found that alpha diversity was highest for both bacterial and eukaryotic communities in the hybrid host species *S. leucophylla x S. rosea*; average effective number of species was 28.75 for bacteria and 10.5 for eukaryotes (Fig. 2A and B). Bacterial community beta diversity was significantly influenced by a single factor: host species (PERMANOVA, $R^2 = 0.297$, $p = 0.003$; Table 4). Eukaryotic community beta diversity was influenced by two factors: host species (PERMANOVA; $R^2 = 0.440$, $p = 0.003$; Table 4) and volume (Mantel test; $R = 0.321$, $p = 0.003$; Table 4). NMDS ordination illustrated the separation between the different communities by host species (Fig. 2C and D). For bacterial communities, the samples from *S. leucophylla x S. rosea* pitchers appeared to cluster near *S. rosea*, but this was not the case for eukaryotic communities. ANCOM analyses identified three bacterial ASVs (Fig. 2E) and 12 eukaryotic ASVs (top six are shown in Fig. 2F) as significantly differentially abundant among host species.

Community structure across sites within single host species

After investigating the effects of host species, we focused on the effects of site and biogeography on community richness and diversity. To eliminate the influence of host species as a factor, we analysed the samples within single host species. We chose to focus on samples from the host species *S. leucophylla* (Fig. 3) because this species was sampled at six different sites across a relatively broad longitudinal gradient: Blackwater (six pitchers), Eglin Airforce Base Sites 1 (three pitchers) and 2 (four

Table 2. Beta diversity results for bacterial (16S) and eukaryotic (18S) communities at all sites more than two co-occurring *Sarracenia* species or hybrids.

Bacteria		Host Species		Volume	pH
Blackwater	R^2	0.224	r	0.103	0.218
	p	0.138	p	0.17	0.138
Splinter Hill	R^2	0.297	r	0.127	0.122
	p	0.003*	p	0.158	0.158
Crystal Bog	R^2	0.52	r	−0.051	0.244
	p	0.003*	p	0.636	0.088
Plea Phase Savanna	R^2	0.372	r	0.276	0.158
	p	0.03*	p	0.106	0.112
Eglin Base 1	R^2	0.555	r	0.288	0.265
	p	0.015*	p	0.064	0.064
Eukaryotes		**Host Species**		**Volume**	**pH**
Blackwater	R^2	0.295	r	0.251	−0.169
	p	0.009*	p	0.062	0.867
Splinter Hill	R^2	0.440	r	0.321	−0.015
	p	0.003*	p	0.003*	0.537
Crystal Bog	R^2	0.292	r	−0.067	−0.204
	p	0.33	p	0.916	0.948
Plea Phase Savanna	R^2	0.47	r	−0.088	−0.005
	p	0.021*	p	0.936	0.936
Eglin Base 1	R^2	0.395	r	0.009	0.017
	p	0.144	p	0.802	0.802

A star indicates significance at alpha = 0.05 for Holm-adjusted p-values.

pitchers), Nokuse Bog (six pitchers), Splinter Hill (13 pitchers) and Weeks Bay (three pitchers; Fig. 1A). Alpha diversity was highest for the bacterial community at Eglin Airforce Base site 1 (average effective number of species: 27, Fig. 3A) and for the eukaryotic community at Splinter Hill (average effective number of species: 6.46, Fig. 3B). Beta diversity of the bacterial communities was significantly influenced by site (PERMANOVA; $R^2 = 0.210$, $p = 0.005$; Table 3) but not the eukaryotic communities (Table 4). For biogeography, we used longitude since our sites were spread more broadly across longitude than latitude. The *S. leucophylla* sites ranged on the longitudinal axis from −87.818567° (Weeks Bay) to −86.224828° (Nokuse Bog) as shown in Fig. 1A. We found significant correlations between the change in community composition and in the longitudinal gradient for both bacteria (Mantel test; $r = 0.220$, $p = 0.005$; Table 3) and the eukaryotes ($r = 0.171$, $p = 0.025$; Table 4). We visualized these trends using NMDS ordination along two axes (Fig. 3C and D) with the longitudinal gradient overlaid. In these figures, the size of the circles reflects the longitude of each sample and the lines represent a smooth surface fit to the longitude using a generalized additive model. Our ANCOM analysis found no ASVs significantly differing with longitude above a 0.8 threshold.

We investigated the influence of site, longitude and pitcher fluid characteristics on beta diversity for all other communities from single host species, even though only *S. flava* was distributed across as many sites. Beta diversity was significantly influenced by site in the bacterial communities of *S. alata* (PERMANOVA; $R^2 = 0.295$, $p = 0.020$) and *S. rosea* ($R^2 = 0.222$, $p = 0.020$). Beta diversity of the bacterial communities from *S. purpurea* was also influenced by pH (Mantel test; $r = 0.310$, $p = 0.028$) and volume (Mantel test; $r = 0.234$, $p = 0.044$), and just by pH for *S. rosea* (Mantel test, $R = 0.345$, $p = 0.010$; Table 3). The bacterial and eukaryotic communities associated with the species *S. alata*, *S. purpurea* and *S. rubra* were not significantly correlated with any other factors.

Hybrid microbiomes differ from parent species

When we normalized by sample size, we found numerous ASVs shared between parent species and their hybrids, but generally the majority were unique to each host (Fig. 4D–F and J–L). Holm-adjusted pairwise comparisons of beta-diversity among parent species and their hybrids indicated that all significantly differed from each other (for both bacteria and eukaryotes, $p < 0.05$), except for *S. flava x S. leucophylla* with either of its parent species. Bacterial communities from pitchers where *S. rosea* hybridized with another species (*S. leucophylla* or *S. flava*) tended to cluster more with *S. rosea* than the other parent in the NMDS plots, although more samples from hybrids would help to verify this observation (Fig. 4A and B). ANCOM analysis found significantly differentially

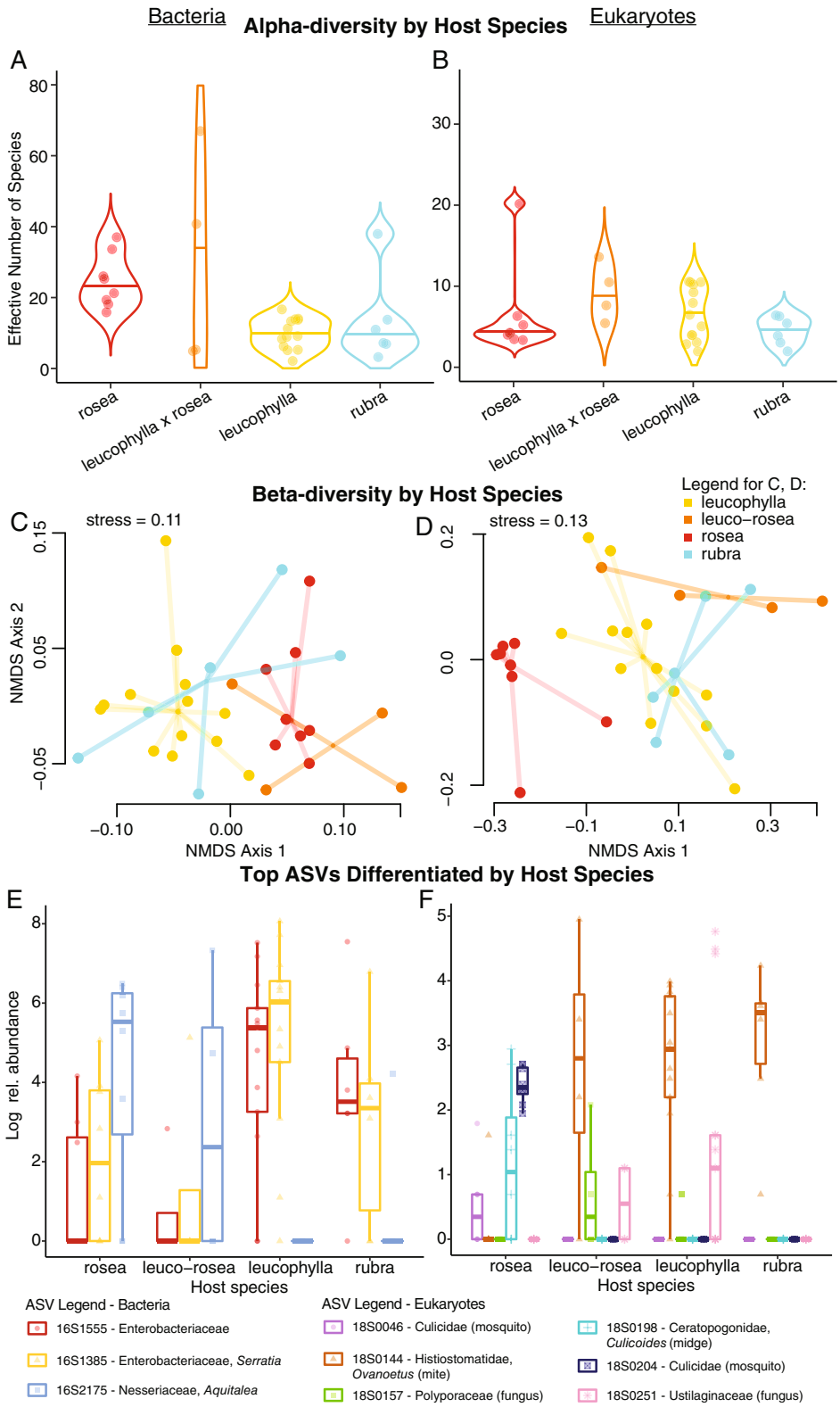

Fig. 2. The Splinter Hill bacterial and eukaryotic communities differ across co-occurring host species.
A. Bacterial and (B) eukaryotic alpha diversity (effective number of species, the exponential of the Shannon index) across three species and one hybrid.
C. Bacterial and (D) eukaryotic NMDS plots using weighted UniFrac distances show beta diversity by host species.
E. Bacterial and (F) eukaryotic ASVs that significantly differ by host species above a 0.8 (E) and 0.9 (F) threshold.

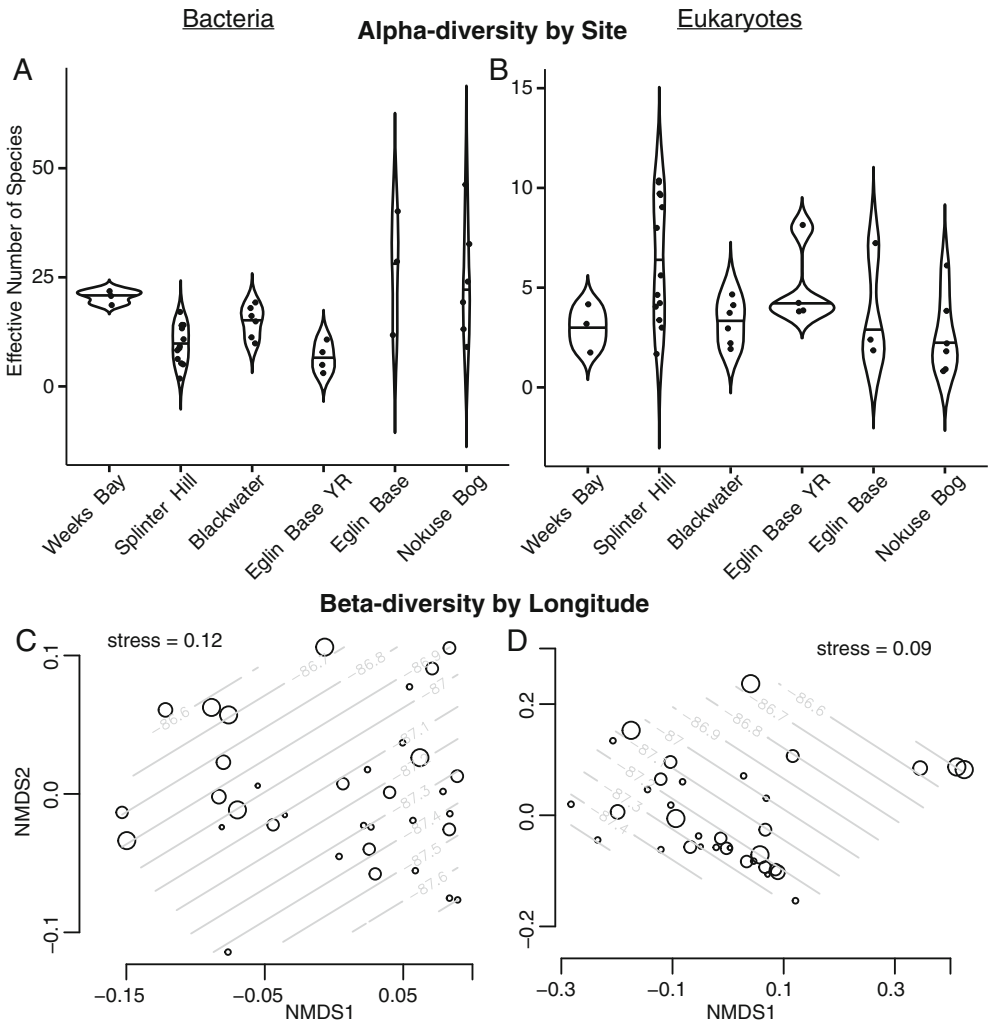

Fig. 3. *S. leucophylla*-associated bacterial and eukaryotic communities differ across a longitudinal gradient.
A. Bacterial and (B) eukaryotic alpha diversity (effective number of species) across six sites.
C. Bacterial and (D) eukaryotic NMDS plots with weighted UniFrac distances show beta diversity with point size scaled by longitude and with longitude overlaid.

Table 3. Beta diversity results for bacterial (16S) communities subset by individual host species.

Host species		Site		Volume	pH	Latitude	Longitude
S. alata	R^2	0.295	*r*	−0.006	−0.039	0.031	0.386
	p	0.020*	*p*	0.924	0.924	0.618	0.048*
S. flava	R^2	0.379	*r*	0.070	0.167	−0.003	0.025
	p	0.005*	*p*	0.687	0.088	0.687	0.687
S. leucophylla	R^2	0.361	*r*	0.103	0.172	0.167	0.214
	p	0.005*	*p*	0.112	0.060	0.018*	0.005*
S. purpurea	R^2	NA	*r*	0.234	0.310	NA	NA
	p	NA	*p*	0.044*	0.028*	NA	NA
S. rosea	R^2	0.222	*r*	−0.015	0.345	−0.010	−0.039
	p	0.020*	*p*	1.000	0.010*	1.000	1.000
S. rubra	R^2	0.148	*r*	0.263	−0.019	−0.075	−0.092
	p	1.000	*p*	0.925	1.000	1.000	1.000

A star indicates significance at alpha = 0.05 for Holm-adjusted *p*-values.

Table 4. Beta diversity results for eukaryotic (18S) communities subset by individual host species.

Host species		Site		Volume	pH	Latitude	Longitude
S. alata	R^2	0.1843	*r*	−0.119	−0.160	0.081	−0.012
	p	0.468	*p*	1.000	1.000	0.325	1.000
S. flava	R^2	0.206	*r*	0.310	−0.1129	0.066	0.0620
	p	0.496	*p*	0.095	0.956	0.304	0.304
S. leucophylla	R^2	0.214	*r*	−0.1015	−0.051	−0.007	0.190
	p	0.252	*p*	1.000	1.000	1.000	0.020*
S. purpurea	R^2	NA	*r*	0.367	−0.110	NA	NA
	p	NA	*p*	0.076	0.736	NA	NA
S. rosea	R^2	0.145	*r*	−0.148	−0.144	−0.043	−0.039
	p	1.000	*p*	1.000	1.000	1.000	1.000
S. rubra	R^2	0.3265	*r*	0.3136	0.175	0.133	0.171
	p	0.210	*p*	0.228	0.228	0.228	0.210

A star indicates significance at alpha = 0.05 for Holm-adjusted *p*-values.

abundant bacterial families (at or above a 0.6 threshold) in *S. leucophylla*–*S. rosea* comparisons (seven families) and in *S. flava*–*S. rosea* comparisons (eight families). These families, on average, were almost always present in hybrids at levels that were intermediate between the two parent species (Supplementary Fig. 5).

Discussion

Microbial community structure in pitchers is primarily influenced by host species

The assembly of host-associated microbial communities is conditioned by factors determined by both the host organism and its physical environment. Host organisms regulate the assembly of their microbiomes through various mechanisms such as the topography, chemistry and immune responses of the colonized surface (Schlechter *et al*., 2019). Previous research with *Sarracenia* species has examined the influence of various factors on microbiome assembly, but little has been done to compare environmental effects to those of the host species, particularly among different species within the genus. In this study we were able to tease apart the influence of host species from that of location by sampling multiple, co-occurring *Sarracenia* species along a continuous geographic range. We found that host species identity more consistently explains a higher portion of the variation in bacterial and eukaryotic pitcher community structure than latitude, longitude or specific sampling sites. We also found that many ASVs significantly differed across host species, even within a single site and thus a common environment.

The influence of the host plant on microbial community structure can be strong even at great geographic distances. Previous research has demonstrated that within a single host species, climatic variables that vary with geographic distance influence pitcher communities (Buckley *et al*., 2003; Zander *et al*., 2017). However, we found that closely related pitcher plant species, *S. purpurea* and *S. rosea*, developed more structurally similar microbial communities despite living in distant and different environments: a Massachusetts bog versus Florida longleaf pine savannas. Our results suggest that host phenotype, here described by general pitcher shape and fluid characteristics, has a stronger influence than climatic variables across spatial gradients. For both the bacterial and the eukaryotic communities, NMDS ordinations showed one cluster including communities from *S. rosea* and *S. purpurea*, and the other group included communities from the other host species (Fig. 1). As mentioned in the introduction, *S. rosea* may not be a separate species; phylogenetic evidence shows mixed results, and *S. rosea* is also known as *S. purpurea* ssp. *venosa* var. *burkii* (Ellison *et al*., 2012; Stephens *et al*., 2015). The close phylogenetic relatedness of *S. rosea* and *S. purpurea* is reflected in strong phenotypic similarities; for example, pitchers from *S. rosea* and *S. purpurea* are short and round and generally contain larger volumes of relatively clear fluid, while the other species have tall, narrow pitchers with smaller volumes of more cloudy fluid. Pitcher volume also explained a significant portion of the variation in total bacterial community beta diversity (Table 1); however, its impact was less clear within individual host species (Table 3) and only significant for *S. purpurea* communities. Because of the discrepancies in the current *Sarracenia* phylogeny (e.g. disagreement in clade placement when using different methods) and the variability in samples from a single host species (as shown by the NMDS plots in Fig. 1), we did not test for phylosymbiosis (Lim and Bordenstein, 2020) to see if the microbial community relationships recapitulated the phylogeny of the *Sarracenia* hosts.

Our results for the hybrid *Sarracenia* species indicate that the microbial and inquiline communities hosted by

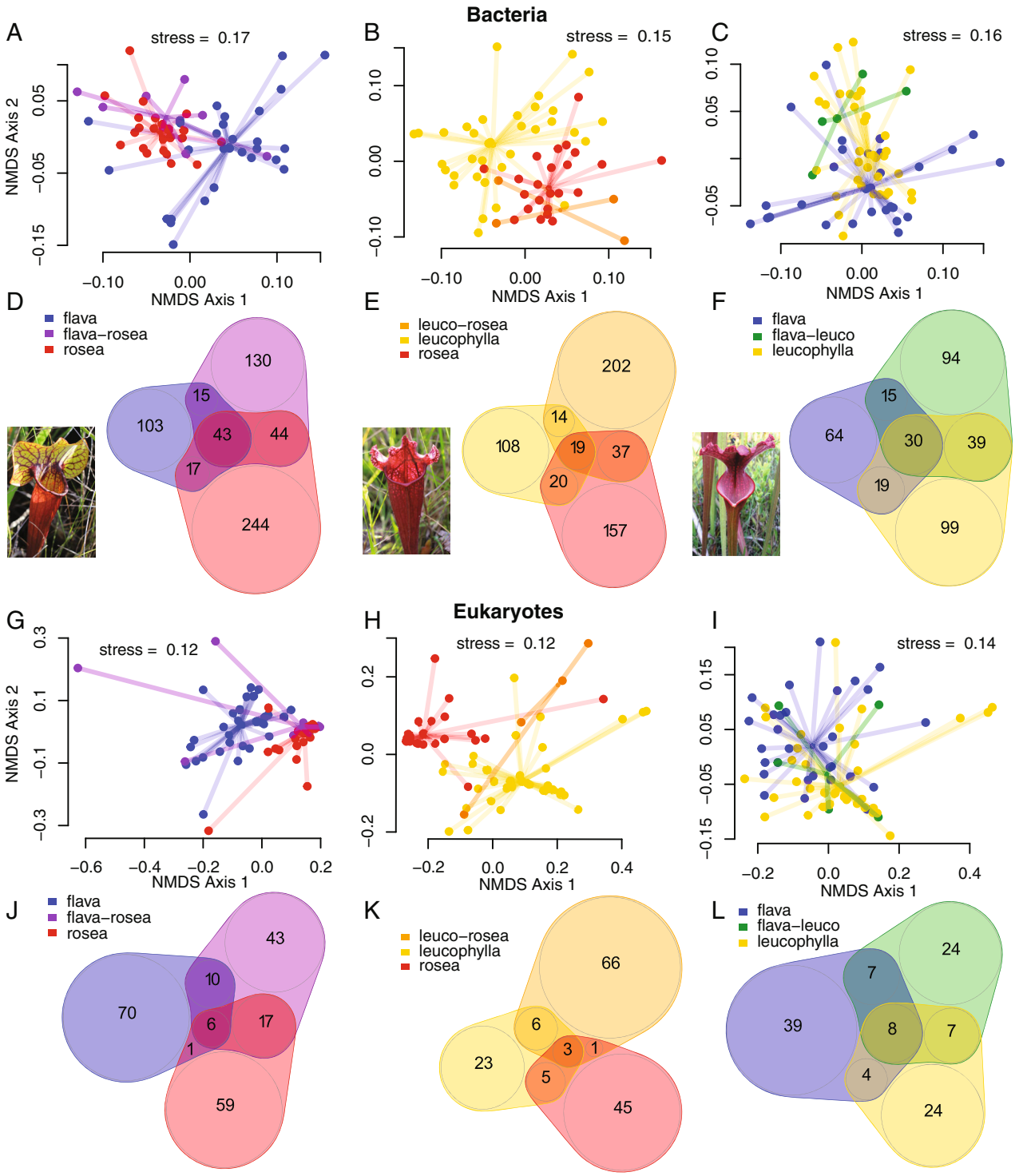

Fig. 4. Beta diversity and Venn diagrams of bacterial and eukaryotic communities associated with hybrids (see inset photos for examples). NMDS plots using weighted UniFrac distances of bacterial (A–C) and eukaryotic (G–I) communities associated with natural hybrids and their parent species: *S. rosea*, *S. leucophylla* and *S. flava*. Venn diagrams show bacterial (D–F) and eukaryotic (J–L) ASVs normalized by sample size that are shared between natural hybrids and their parent species.

hybrids are often also unique, despite sometimes showing more similarity to one or the other parent species (Fig. 4). Morphologically, the shape of the hybrid pitchers appears intermediate between that of their parent species (see Bittleston *et al.*, 2018 for morphological measurements of these species). Bacterial families significantly

differing across hybrids and their parent species were present at intermediate abundance levels in hybrids, suggesting that the hybrids had intermediate phenotypes, leading to recruitment of microbes at intermediate levels. Although only limited data are available about how hybridization affects plant microbiomes (Sahu and Mishra, 2021), previous studies have shown that rhizosphere and leaf-associated communities can differ in hybrids compared with parent species (Adam *et al*., 2018; Wagner *et al*., 2020). Our current study contains only small numbers of natural hybrid samples due to their comparative rarity at our sites; however, our results suggest interesting patterns that future studies could investigate further in more controlled contexts.

Although limited, our hybrid results also suggest that pitcher shape may be more important than phylogenetic relatedness. Both *S*. *flava* and *S*. *leucophylla* have tall, fluted pitchers despite being in different clades (Fig. 1A) and *post hoc* tests of the hybrid PERMANOVAs did not find significant differences between the communities of the *S*. *flava x leucophylla* hybrid and either of the parent species, while differences were found for all communities in hybrids involving the more squat, round *S*. *rosea*. Furthermore, no bacterial families were found to be significantly differentially abundant between pitchers of *S*. *flava*, *S*. *leucophylla* and their hybrid, while many differed in comparisons involving *S*. *rosea* (Supplementary Fig. 5).

We used the opportunity presented by different host species co-occurring within the Splinter Hill site to examine the influence of host species within the same environment (Fig. 2). Host species strongly influenced beta diversity of bacterial and eukaryotic pitcher communities at Splinter Hill and most other sites with more than two co-occurring species or hybrids (Table 2). The only other influential factor we measured in this analysis was volume, which was only significant for eukaryotes at Splinter hill and could be confounded by the presence of different host species that generally contain different volumes (e.g. *S*. *rosea* pitchers generally have higher fluid volumes than those of *S*. *leucophylla*). Our results suggest that there are other, unmeasured aspects of host species that influence bacterial and eukaryotic community diversity. Gilbert *et al*. (2020b) demonstrated that the Asian pitcher plant lineage *Nepenthes* regulates pitcher microbial community through differentiation in pitcher characteristics, such as fluid pH and viscosity. We detected a similar pattern in *Sarracenia* and further research is needed to parse the influence of other pitcher characteristics on microbial communities. One additional factor that was not captured in this study is the effect of microbial community succession within individual pitchers, because we sampled at single time points. Previous research has found strong signals of succession on organisms living in *S*. *purpurea* (Fish and Hall, 1978; Gray, 2012; Miller and terHorst, 2012). In this study, we targeted communities living in healthy leaves of wild pitcher plants across different ages and successional stages to get a broad survey of community diversity for each host and site.

Specific ASVs were significantly associated with distinct host species, indicating possible specialization. Some of these ASVs corroborate previous observations of specificity in dipterans and mites (Dahlem and Naczi, 2006; Satler and Carstens 2016). For example, mosquitos were only found in *S*. *rosea* (Fig. 2F), while an aquatic mite was found in *S*. *leucophylla*, *S*. *rubra* and *S*. *alata* but not *S*. *rosea*. Additionally, two fungal ASVs were enriched in *S*. *leucophylla* (Fig. 2F) and *S*. *alata*. When organisms are specific to particular host species it suggests that inquilines and hosts are co-diversifying, as has been found for certain species associated with *S*. *alata* (Satler and Carstens, 2016, 2019), and have a persistent symbiotic (and potentially mutualistic) relationship with their host. For bacteria, an *Aquitalea* ASV was enriched in *S*. *rosea* (Fig. 2E) that has independently been recorded as an abundant ASV in *S*. *purpurea* samples from Massachusetts (ASV698 from Bittleston *et al*., 2020), while two Enterobacteriaceae ASVs were particularly abundant in *S*. *leucophylla* and *S*. *rubra* (Fig. 2E). Pitchers from tall, narrow *Sarracenia* species often have prey filling a large portion of the height of the pitcher, as opposed to just at the bottom. They are also more likely to have cloudy fluid and a putrefied scent, suggesting lower oxygen levels. Because of the shape differences, *S*. *leucophylla* and *S*. *rubra* are more likely to have less oxygen than *S*. *rosea*, which could facilitate the facultative anaerobes in the family Enterobacteriaceae.

Site explains variation in community structure more frequently than latitude or longitude

Environmental factors significantly influenced bacterial pitcher communities, despite generally explaining less of the variation than host species identity. Interestingly, alpha diversity showed more differences by site than beta diversity. Bacterial alpha diversity differed across almost all sites in our GLM, and, although no specific sites differed for eukaryotic alpha diversity, the omnibus test isolating the effect of site was significant. By focusing on single host species in our beta diversity analyses, we were able to more effectively isolate environmental and site effects on the pitcher communities. Site, and, to a lesser extent, latitude or longitude, often explained variation in bacterial beta diversity (Table 3); however, this was not true for eukaryotic communities (Table 4). Only *S*. *leucophylla* eukaryotic communities showed significant differences across space, and it was along the

longitudinal gradient (Fig. 3D; Table 4). Environmental factors vary along geographic scales; for example, for our Gulf Coast sites we found that latitude was strongly correlated with annual mean temperature and longitude was strongly correlated with annual precipitation. However, environmental variables can also vary more stochastically on a site-by-site basis. Within the geographic range of the Gulf Coast sites, site proved to more frequently structure pitcher bacterial communities than latitude or longitude. Specifically, site explained significant variation in bacterial beta diversity in every host species (Table 3) except *S. rubra*, which was sampled at only three sites vs. four to six sites for the other species.

Previous research has demonstrated that *S. alata* pitcher communities (particularly the arthropods) co-diversify with their host plants when geographically isolated (Koopman and Carstens, 2011; Satler and Carstens, 2016, 2019). Thus, biogeographic effects on community structure cannot be completely separated from host effects. Host phenotype may change across spatial gradients along with genotypic changes in a way that affects associated communities, or co-dispersal may drive phylogeographic congruence of pitcher plants and their associated organisms (Koopman and Carstens, 2011; Satler and Carstens, 2017).

Conclusions

Overall, host species identity was the primary factor structuring pitcher communities, generally explaining a large portion of the variation in both bacterial and eukaryotic community composition, even when different species were growing together within the same habitat. Characteristics of host species, such as their pitcher shape, attributes of the pitcher fluid and other unmeasured variables likely drive this relationship. Although it was less consistent, we also found differences in community composition across sampling sites on our longitudinal gradient. When comparing bacteria to eukaryotes, our measured factors more often explained the diversity in bacterial than in eukaryotic communities. This supports the theory that community assembly of bacteria is relatively deterministic compared with that of micro-eukaryotes, which is affected more by stochastic events (Ragon *et al*., 2012; Li *et al*., 2017; Logares *et al*., 2018), although contrasting patterns have also been found (e.g. Wu *et al*., 2018). Despite bacterial communities more often showing significant associations with measured factors, host species identity explained a significant portion of the diversity for both bacterial and eukaryotic communities, and drove differences in multiple specific ASVs. This was true even when the different hosts co-occurred within the same sites and the same environmental conditions. In *Sarracenia* pitcher plants, the host plant phenotype strongly affects its microbiome and inquilines.

Experimental procedures

Sampling and DNA extraction of pitcher fluid

Samples were collected and processed as described in Bittleston *et al*. (2018). Briefly, we sampled from pitchers of five *Sarracenia* species (*S. alata* 24 samples, *S. flava* 30 samples, *S. leucophylla* 35 samples, *S. rosea* 24 samples and *S. rubra* 12 samples) and their natural hybrids from 13 sites along the U.S. Gulf Coast (Fig. 1A, Supplementary Table 1) in June 2014, and from a sixth species (*S. purpurea* 20 samples) from Harvard Forest in Massachusetts (Fig. 1A) in July 2014. All pitcher plants, including hybrids, were identified in the field by sight. All recognized species of pitcher plant hybridize with each other, and hybrid pitcher plants can be identified in the field by their characteristics that are intermediate between two non-hybrid parent species growing sympatrically (Bell, 1952). This long-practised identification method has more recently been validated by genetic microsatellite analysis (Furches *et al*., 2013).

To collect the pitcher communities, we mixed the contents of each pitcher and collected the fluid using sterile, single-use plastic transfer pipettes and deposited each sample into a sterile plastic tube. We recorded the fluid volume and measured pH with colorpHast strips (EMD Millipore). To preserve DNA, we added cetyl trimethyl ammonium bromide and salt solution (hereafter 'CTAB'; final concentrations: 2% CTAB, 1.4 M NaCl, 20 mM EDTA, 100 mM Tris pH 8) to each sample in the same volume as the collected fluid. All samples were processed the same day as collection and were transported at room temperature to Harvard University, where they were frozen at −80°C until DNA extraction.

Details of DNA extraction and sequencing are described in Bittleston *et al*. (2018). Briefly, we used bead-beating and then a phenol-chloroform extraction (Sambrook *et al*., 2001) and each set of extractions included a negative control. We quantified DNA with a Quant-iT High-Sensitivity dsDNA Assay Kit (Invitrogen) before sending samples to the Environmental Sample Preparation and Sequencing Facility (ESPSF) at Argonne National Laboratory for amplicon sequencing.

Amplicon sequencing of 16S and 18S rRNA genes

ESPSF used the Earth Microbiome Project's original barcoded 16S (515F-806R, V4 region) and 18S (1391f-1510r, V9 region) primers that were adapted for the Illumina MiSeq (Amaral-Zettler *et al*., 2009; Stoeck *et al*., 2010; Caporaso *et al*., 2012) to amplify prokaryotic

and eukaryotic DNA respectively. Each 25 μl PCR reaction contained 9.5 μl of MO BIO PCR Water (Certified DNA-Free), 12.5 μl of QuantaBio's AccuStart II PCR ToughMix (2× concentration, 1× final), 1 μl Golay barcode tagged Forward Primer (5 μM concentration, 200 pM final), 1 μl Reverse Primer (5 μM concentration, 200 pM final) and 1 μl of template DNA. The conditions for PCR were as follows: 94°C for 3 min to denature the DNA, with 35 cycles at 94°C for 45 s, 50°C for 60 s and 72°C for 90 s; with a final extension of 10 min at 72°C to ensure complete amplification. Amplicons were then quantified using PicoGreen (Invitrogen) and a plate reader (Infinite 200 PRO, Tecan) before being pooled in equimolar amounts. This pool was then cleaned with AMPure XP Beads (Beckman Coulter), and quantified using a fluorometer (Qubit, Invitrogen). After quantification, the molarity of the pool was determined and it was diluted down to 2 nM, denatured, and then diluted to a final concentration of 6.75 pM with a 10% PhiX spike for sequencing on the Illumina MiSeq. The 16S and 18S amplicons were sequenced on separate 151 bp × 12 bp × 151 bp MiSeq runs.

Amplicon sequence data for most of the samples are available in the Sequence Read Archive NCBI BioProject PRJNA448553. Sequences from hybrid pitchers were not included in the Bittleston *et al*. (2018) study, and were deposited independently in NCBI BioProject PRJNA777590.

Amplicon sequence analysis

On a computing cluster, we demultiplexed and processed our sequences using QIIME2 versions 2018.4 and 2019.10 (Bolyen *et al*., 2019), using the DADA2 (Callahan *et al*., 2016) plugin to denoise sequences and to generate ASVs of ∼250 base pairs in length. We assigned taxonomy for 16S sequences using the classify-sklearn method which is a Naive Bayes classifier (Pedregosa *et al*., 2011), and a pre-trained classifier made with the Greengenes database, version 13_8 (McDonald *et al*., 2012; Bokulich *et al*., 2018). For 18S sequences we used the classify-consensus-vsearch method and the SILVA 128 database. We built phylogenetic trees with the QIIME2 SEPP plugin (Janssen *et al*., 2018). For ease of communication, we abbreviated the ASV names, and provide a supplementary table with the original and abbreviated names for those mentioned in this study (Supplementary Table 2).

Community diversity

We first examined how various factors affected the richness and evenness of a sample (alpha diversity) as well as the composition across samples (beta diversity) of the bacteria and eukaryotes in our pitcher samples. All of our analyses were conducted in R (version 4.0.2, R Core Team, 2020). In order to properly compare across samples with different sequencing depths, we rarefied (subsampled) each sample to match the sample with the fewest sequences, 3910 sequences for 16S samples and 411 sequences for 18S, using the functions rrarefy() and rarecurve() from the vegan package (Oksanen *et al*., 2018, see Supplementary Fig. 1 for rarefaction curves). We used a relatively low sequence cutoff for 18S in order to retain important samples, and because results were very similar when we subsampled to a higher sequencing depth. Shannon diversity was calculated for each sample using the diversity() function from the vegan package, and the effective number of species for each sample was produced by calculating the exponent of the Shannon values. To test the effect of subsampling on alpha diversity, we used the function estimateD() from the iNEXT package (Hsieh *et al*., 2016) to generate confidence intervals around our measures of effective number of species when subsampling, and found that the 0.95 confidence intervals were very small (Supplementary Fig. 2). We determined alpha diversity for microbial communities associated with different categorical factors (Host Species, Site) and along continuous gradients (Longitude, pH, Volume) as the average effective number of species for each sample in a category or at different points on a gradient. The range of alpha diversity for all samples within categorical variables was visualized using violin plots overlaid with jitter plots with the package ggplot2 (Wickham, 2016). We determined the relationships between alpha diversity and all factors using GLMs with function glm() and a Poisson distribution, and also used omnibus tests to see if alpha diversity varies significantly among species or among sites when controlling for other factors.

To calculate beta diversity, we generated weighted UniFrac dissimilarity matrices for the ASV data using the distance() function from the phyloseq package (McMurdie and Holmes, 2013). We used a PERMANOVA for categorical factors with the adonis() function from the vegan package, and *post hoc* tests were done with the pairwise.adonis() function. Mantel tests were used to test the relationships between beta diversity and continuous factors using the mantel() function from the vegan package. *p*-values were adjusted for multiple comparisons using the Holm method. We examined beta diversity further by visualizing the similarity of communities with NMDS. NMDS was calculated using the metaMDS() function from the vegan package, using two dimensions ($k = 2$) for all analyses. Ordination plots were generated using the plot() function from the base package in R and overlaying factors in the ordination space using the functions ordispider() and ordiplot() from the vegan package

to show the alignment of closely related samples in comparison to their grouping with categorical and continuous factors.

We discerned individual ASVs significantly correlated with our environmental factors using Analysis of Composition of Microbiomes, ANCOM (Mandal *et al*., 2015). ANCOM was calculated using an updated version of the R code published on the personal website of Siddharta Mandal, Ph.D. To learn more about the identity of ASVs with significant correlations that had minimal taxonomic information, we used NCBI's BLASTn megablast tool (Supplementary Dataset 1). Differential abundance of significant ASVs across factors was visualized using ggplot2 violin and jitter plots.

For our 13 Gulf Coast sites, we extracted WorldClim data (Fick and Hijmans, 2017) for Annual Mean Temperature, Temperature Annual Range and Annual Precipitation (Supplementary Table 1), and used linear models in R to measure the correlations of latitude and longitude with these climatic variables at our sites. We did not include data for the Harvard Forest site in the WorldClim analysis due to the large geographic distance between this site and the Gulf Coast sites. Our R code with all necessary data input files has been deposited in the Harvard Dataverse. The URL to access these files is https://doi.org/10.7910/DVN/DB6WIN.

Shared ASVs across hybrid and parent species

We visualized the similarity of microbiomes from pairs of host species and their hybrids. Hybrids were found at Blackwater, Crystal Bog, Eglin Airforce Base Site 1, Mud Swamp, Plea Phase Savanna and Splinter Hill sites and were identified in the field based on visual assessment. Beta diversity NMDS plots and PERMANOVA tests were analysed as previously described. Presence/absence of ASVs was determined for sets of host species (e.g. *S. flava* and *S. leucophylla*) and their hybrid (e.g. *S. flava x S. leucophylla*) to investigate how hybrids housed ASVs in common with or different from their parent species. However, due to their rarity, we had very few samples from hybrids, and when one host species is represented by many more samples than another, it is more likely to have more ASVs. To better compare across the different host species, we chose to normalize to the lowest number of samples for a single host species in the set. A number of average sets of ASVs were constructed for each host species equal to the lowest number of samples for a single host species. For the average sets, each ASV was selected based on a probability equal to its percent presence in the total number of samples collected from each host species. The number of ASVs selected for each set was equal to the mean ASV richness in the samples from the host species. All generated ASV sets were compiled into a single normalized set. We used Venn diagrams to visualize how the normalized ASVs were either unique to a host species or shared across multiple hosts. Venn diagrams were generated using the nVennR package (Perez-Silva *et al*., 2018). We also used ANCOM for bacteria at the family level, to investigate differentially abundant bacterial families across hybrid and parent host species.

Acknowledgements

We appreciate the insightful suggestions from our anonymous reviewers, which improved the manuscript. We thank the many people and organizations for giving us support and permission to sample at their sites, including Harvard Forest, the Nature Conservancy, Eglin Air Force Base, Matthew Aresco at Nokuse Plantation, Tom Miller from FSU and local State Forests. L.S.B. and J.A.H were supported by the NSF Idaho EPSCoR Program and by the NSF under award number OIA-1757324.

References

Adam, E., Bernhart, M., Müller, H., Winkler, J., and Berg, G. (2018) The Cucurbita pepo seed microbiome: genotype-specific composition and implications for breeding. *Plant Soil* **422**: 35–49.

Adlassnig, W., Peroutka, M., and Lendl, T. (2011) Traps of carnivorous pitcher plants as a habitat: composition of the fluid, biodiversity and mutualistic activities. *Ann Bot* **107**: 181–194.

Amaral-Zettler, L.A., McCliment, E.A., Ducklow, H.W., and Huse, S.M. (2009) A method for studying protistan diversity using massively parallel sequencing of V9 hypervariable regions of small-subunit ribosomal RNA genes. *PLoS One* **4**: e6372.

Bell, C.R. (1952) Natural hybrids in the genus Sarracenia: I. history, distribution and taxonomy. *J Elisha Mitchell Sci Soc* **68**: 55–80.

Bittleston, L.S., Baker, C.C.M., Strominger, L.B., Pringle, A., and Pierce, N.E. (2016) Metabarcoding as a tool for investigating arthropod diversity in Nepenthes pitcher plants. *Austral Ecol* **41**: 120–132.

Bittleston, L.S., Gralka, M., Leventhal, G.E., Mizrahi, I., and Cordero, O.X. (2020) Context-dependent dynamics lead to the assembly of functionally distinct microbial communities. *Nat Commun* **11**: 1440.

Bittleston, L.S., Wolock, C.J., Yahya, B.E., Chan, X.Y., Chan, K.G., Pierce, N.E., and Pringle, A. (2018) Convergence between the microcosms of Southeast Asian and North American pitcher plants. *Elife* **7**: e36741.

Bokulich, N.A., Kaehler, B.D., Rideout, J.R., Dillon, M., Bolyen, E., Knight, R., *et al*. (2018) Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. *Microbiome* **6**: 90.

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

Boynton, P.J., Peterson, C.N., and Pringle, A. (2019) Superior dispersal ability can lead to persistent ecological dominance throughout succession. *Appl Environ Microbiol* **85**: e02421-18.

Buckley, H.L., Miller, T.E., Ellison, A.M., and Gotelli, N.J. (2003) Reverse latitudinal trends in species richness of pitcher-plant food webs. *Ecol Lett* **6**: 825–829.

Butler, J.L., Gotelli, N.J., and Ellison, A.M. (2008) Linking the Brown and Green: nutrient transformation and fate in the Sarracenia microecosystem. *Ecology* **89**: 898–904.

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

Canter, E.J., Cuellar-Gempeler, C., Pastore, A.I., Miller, T. E., and Mason, O.U. (2018) Predator identity more than predator richness structures aquatic microbial assemblages in *Sarracenia purpurea* leaves. *Ecology* **99**: 652–660.

Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Huntley, J., Fierer, N., *et al*. (2012) Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J* **6**: 1621–1624.

Compant, S., Samad, A., Faist, H., and Sessitsch, A. (2019) A review on the plant microbiome: ecology, functions, and emerging trends in microbial application. *J Adv Res* **19**: 29–37.

Dahlem, G.A., and Naczi, R.F.C. (2006) Flesh flies (Diptera: Sarcophagidae) associated with north American pitcher plants (Sarraceniaceae), with descriptions of three new species. *Ann Entomol Soc Am* **99**: 218–240.

Deiner, K., Bik, H.M., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt, F., *et al*. (2017) Environmental DNA metabarcoding: transforming how we survey animal and plant communities. *Mol Ecol* **26**: 5872–5895.

Edwards, J., Johnson, C., Santos-Medellín, C., Lurie, E., Podishetty, N.K., Bhatnagar, S., *et al*. (2015) Structure, variation, and assembly of the root-associated microbiomes of rice. *Proc Natl Acad Sci U S A* **112**: E911–E920.

Ellison, A.M., Butler, E.D., Hicks, E.J., Naczi, R.F.C., Calie, P.J., Bell, C.D., and Davis, C.C. (2012) Phylogeny and biogeography of the Carnivorous plant family Sarraceniaceae. *PLoS One* **7**: e39291.

Fick, S.E., and Hijmans, R.J. (2017) WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *Int J Climatol* **37**: 4302–4315.

Fish, D., and Hall, D.W. (1978) Succession and stratification of aquatic insects inhabiting the leaves of the insectivorous pitcher plant, Sarracenia purpurea. *Am Midl Nat* **99**: 172–183.

Furches, M.S., Small, R.L., and Furches, A. (2013) Hybridization leads to interspecific gene flow in Sarracenia (Sarraceniaceae). *Am J Bot* **100**: 2085–2091.

Gilbert, K.J., Bittleston, L.S., Naive, M.A.K., Kiszewski, A.E., Buenavente, P.A.C., Lohman, D.J., and Pierce, N.E. (2020a) Investigation of an elevational gradient reveals strong differences between bacterial and eukaryotic communities coinhabiting nepenthes Phytotelmata. *Microb Ecol* **80**: 334–349.

Gilbert, K.J., Bittleston, L.S., Tong, W., and Pierce, N.E. (2020b) Tropical pitcher plants (nepenthes) act as ecological filters by altering properties of their fluid microenvironments. *Sci Rep* **10**: 4431.

Gray, S.M. (2012) Succession in the aquatic Sarracenia purpurea community: deterministic or driven by contingency? *Aquat Ecol* **46**: 487–499.

Grothjan, J.J., and Young, E.B. (2019) Diverse microbial communities hosted by the model carnivorous pitcher plant Sarracenia purpurea: analysis of both bacterial and eukaryotic composition across distinct host plant populations. *PeerJ* **7**: e6392.

Hsieh, T.C., Ma, K.H., and Chao, A. (2016) iNEXT: an R package for rarefaction and extrapolation of species diversity (hill numbers). *Methods Ecol Evol* **7**: 1451–1456.

Janssen, S., McDonald, D., Gonzalez, A., Navas-Molina, J. A., Jiang, L., Xu, Z.Z., *et al*. (2018) Phylogenetic placement of exact amplicon sequences improves associations with clinical information. *mSystems* **3**: e00021-18.

Kanokratana, P., Mhuanthong, W., Laothanachareon, T., Tangphatsornruang, S., Eurwilaichitr, L., Kruetreepradit, T., *et al*. (2016) Comparative study of bacterial communities in nepenthes pitchers and their correlation to species and fluid acidity. *Microb Ecol* **72**: 381–393.

Kitching, R.L. (2001) Food webs in Phytotelmata: "bottom-up" and "top-down" explanations for community structure. *Annu Rev Entomol* **46**: 729–760.

Koller-Peroutka, M., Krammer, S., Pavlik, A., Edlinger, M., Lang, I., and Adlassnig, W. (2019) Endocytosis and digestion in carnivorous pitcher plants of the family Sarraceniaceae. *Plan Theory* **8**: 367.

Koopman, M.M., and Carstens, B.C. (2011) The microbial phyllogeography of the carnivorous plant Sarracenia alata. *Microb Ecol* **61**: 750–758.

Leveau, J.H. (2019) A brief from the leaf: latest research to inform our understanding of the phyllosphere microbiome. *Curr Opin Microbiol* **49**: 41–49.

Li, Y., Adams, J., Shi, Y., Wang, H., He, J.-S., and Chu, H. (2017) Distinct soil microbial communities in habitats of differing soil water balance on the Tibetan plateau. *Sci Rep* **7**: 46407.

Lim, S.J., and Bordenstein, S.R. (2020) An introduction to phylosymbiosis. *Proc R Soc B: Biol Sci* **287**: 20192900.

Logares, R., Tesson, S.V.M., Canbäck, B., Pontarp, M., Hedlund, K., and Rengefors, K. (2018) Contrasting prevalence of selection and drift in the community structuring of bacteria and microbial eukaryotes. *Environ Microbiol* **20**: 2231–2240.

Luciano, C.S., and Newell, S.J. (2017) Effects of prey, pitcher age, and microbes on acid phosphatase activity in fluid from pitchers of Sarracenia purpurea (Sarraceniaceae). *PLoS One* **12**: e0181252.

Mandal, S., Van Treuren, W., White, R.A., Eggesbø, M., Knight, R., and Peddada, S.D. (2015) Analysis of composition of microbiomes: a novel method for studying microbial composition. *Microb Ecol Health Dis* **26**: 27663.

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

McMurdie, P.J., and Holmes, S. (2013) Phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* **8**: e61217.

Miller, T.E., and terHorst, C.P. (2012) Testing successional hypotheses of stability, heterogeneity, and diversity in pitcher-plant inquiline communities. *Oecologia* **170**: 243–251.

Naylor, D., DeGraaf, S., Purdom, E., and Coleman-Derr, D. (2017) Drought and host selection influence bacterial community dynamics in the grass root microbiome. *ISME J* **11**: 2691–2704.

Nemergut, D.R., Schmidt, S.K., Fukami, T., O'Neill, S.P., Bilinski, T.M., Stanish, L.F., *et al*. (2013) Patterns and processes of microbial community assembly. *Microbiol Mol Biol Rev* **77**: 342–356.

Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H., Szoecs, E., and Wagner, H.. (2018) Vegan: community ecology package.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., *et al*. (2011) Scikit-learn: machine learning in python. *J Mach Learn Res* **12**: 2825–2830.

Perez-Silva, J.G., Araujo-Voces, M., and Quesada, V. (2018) nVenn: generalized, quasi-proportional Venn and Euler diagrams. *Bioinformatics* **34**: 2322–2324.

Prosser, J.I., and Martiny, J.B.H. (2020) Conceptual challenges in microbial community ecology. *Philos Trans R Soc B: Biol Sci* **375**: 20190241.

R Core Team. (2020) *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Ragon, M., Fontaine, M.C., Moreira, D., and López-García, P. (2012) Different biogeographic patterns of prokaryotes and microbial eukaryotes in epilithic biofilms. *Mol Ecol* **21**: 3852–3868.

Sahu, P.K., and Mishra, S. (2021) Effect of hybridization on endophytes: the endo-microbiome dynamics. *Symbiosis* **84**: 369–377.

Sambrook, J., Russell, D.W., and Laboratory, C.S.H. (2001) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Satler, J.D., and Carstens, B.C. (2016) Phylogeographic concordance factors quantify phylogeographic congruence among co-distributed species in the Sarracenia alata pitcher plant system. *Evolution* **70**: 1105–1119.

Satler, J.D., and Carstens, B.C. (2017) Do ecological communities disperse across biogeographic barriers as a unit? *Mol Ecol* **26**: 3533–3545.

Satler, J.D., and Carstens, B.C. (2019) The Sarracenia alata pitcher plant system and obligate arthropod inquilines should be considered an evolutionary community. *J Biogeogr* **46**: 485–496.

Schlechter, R.O., Miebach, M., and Remus-Emsermann, M. N.P. (2019) Driving factors of epiphytic bacterial communities: a review. *J Adv Res* **19**: 57–65.

Schulze, W., Schulze, E.D., Pate, J.S., and Gillison, A.N. (1997) The nitrogen supply from soils and insects during growth of the pitcher plants Nepenthes mirabilis, Cephalotus follicularis and Darlingtonia californica. *Oecologia* **112**: 464–471.

Shoemaker, W., Locey, K., and Lennon, J. (2017). A macroecological theory of microbial biodiversity. *Nat Ecol Evol*, **1**, 0107. https://doi.org/10.1038/s41559-017-0107

Sorensen, J.W., and Shade, A. (2020) Dormancy dynamics and dispersal contribute to soil microbiome resilience. *Philos Trans R Soc B: Biol Sci* **375**: 20190255.

Stephens, J.D., Rogers, W.L., Heyduk, K., Cruse-Sanders, J.M., Determann, R.O., Glenn, T.C., and Malmberg, R.L. (2015) Resolving phylogenetic relationships of the recently radiated carnivorous plant genus Sarracenia using target enrichment. *Mol Phylogenet Evol* **85**: 76–87.

Stoeck, T., Bass, D., Nebel, M., Christen, R., Jones, M.D.M., Breiner, H.-W., and Richards, T.A. (2010) Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. *Mol Ecol* **19**: 21–31.

Taberlet, P., Coissac, E., Pompanon, F., Brochmann, C., and Willerslev, E. (2012) Towards next-generation biodiversity assessment using DNA metabarcoding: next-generation DNA metabarcoding. *Mol Ecol* **21**: 2045–2050.

Vaginal Microbiome Consortium (additional members), Brooks, J.P., Edwards, D.J., Harwich, M.D., Rivera, M.C., Fettweis, J.M., *et al*. (2015) The truth about metagenomics: quantifying and counteracting bias in 16S rRNA studies. *BMC Microbiol* **15**: 66.

Wagner, M.R., Roberts, J.H., Balint-Kurti, P., and Holland, J. B. (2020) Heterosis of leaf and rhizosphere microbiomes in field-grown maize. *New Phytol* **228**: 1055–1069.

Wickham, H. (2016) *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer-Verlag.

Wu, W., Lu, H.-P., Sastri, A., Yeh, Y.-C., Gong, G.-C., Chou, W.-C., and Hsieh, C.-H. (2018) Contrasting the relative importance of species sorting and dispersal limitation in shaping marine bacterial versus protist communities. *ISME J* **12**: 485–494.

Zander, A., Bersier, L.-F., and Gray, S.M. (2017) Effects of temperature variability on community structure in a natural microbial food web. *Glob Chang Biol* **23**: 56–67.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Supplementary Fig. 1. Rarefaction curves of all our samples, with the subsampling cutoffs as red lines.

Supplementary Fig. 2. Effective number of species (the exponential of the Shannon index) for 16S (a) and 18S (b) across all samples, with 0.95 confidence intervals to show the effect of subsampling on alpha diversity.

Supplementary Fig. 3. NMDS of 16S communities coloured by site with vectors added for volume, pH, latitude, and longitude.

Supplementary Fig. 4. NMDS of 18S communities coloured by site.

Supplementary Fig. 5. Bacterial families associated with hybrid and parental pitcher plant species, with significant differential abundances above a 0.6 threshold as analysed with ANCOM.

Supplementary Table 1. The Gulf Coast sites sampled in our study along with their latitude, longitude, and WorldClim climatic variables. Latitude and longitude are only reported to 0.1 accuracy here to help protect plants from potential poaching.

Supplementary Table 2. Original and abbreviated ASV names for those mentioned in the main text.

Supplementary Table 3. GLM results for alpha diversity, using 16S effective number of species as a response variable.

Supplementary Table 4. GLM results for alpha diversity, using 18S effective number of species as a response variable.

Appendix S1. Supporting Information.