10-1-2017

Environmental proteomics reveals taxonomic and functional changes in an enriched aquatic ecosystem

Amanda C. Northrop
University of Vermont

Rachel K. Brooks
University of Vermont

Aaron M. Ellison
Harvard Forest

Nicholas J. Gotelli
University of Vermont

Bryan A. Ballif
University of Vermont

Recommended Citation

Northrop AC, Brooks RK, Ellison AM, Gotelli NJ, Ballif BA. Environmental proteomics reveals taxonomic and functional changes in an enriched aquatic ecosystem. Ecosphere. 2017 Oct;8(10):e01954.

esa

ECOSPHERE

Environmental proteomics reveals taxonomic and functional changes in an enriched aquatic ecosystem

AMANDA C. NORTHROP,[1] RACHEL K. BROOKS,[1] AARON M. ELLISON,[2] NICHOLAS J. GOTELLI,[1] AND BRYAN A. BALLIF[1,†]

[1]*Department of Biology, University of Vermont, Burlington, Vermont 05405 USA*
[2]*Harvard Forest, Harvard University, Petersham, Massachusetts 01366 USA*

Citation: Northrop, A. C., R. K. Brooks, A. M. Ellison, N. J. Gotelli, and B. A. Ballif. 2017. Environmental proteomics reveals taxonomic and functional changes in an enriched aquatic ecosystem. Ecosphere 8(10):e01954. 10.1002/ecs2.1954

Abstract. Aquatic ecosystem enrichment can lead to distinct and irreversible changes to undesirable states. Understanding changes in active microbial community function and composition following organic matter loading in enriched ecosystems can help identify biomarkers of such state changes. In a field experiment, we enriched replicate aquatic ecosystems in the pitchers of the northern pitcher plant, *Sarracenia purpurea*. Shotgun metaproteomics using a custom metagenomic database identified proteins, molecular pathways, and contributing microbial taxa that differentiated control ecosystems from those that were enriched. The number of microbial taxa contributing to protein expression was comparable between treatments; however, taxonomic evenness was higher in controls. Functionally active bacterial composition differed significantly among treatments and was more divergent in control pitchers than in enriched pitchers. Aerobic and facultative anaerobic bacteria contributed most to identified proteins in control and enriched ecosystems, respectively. The molecular pathways and contributing taxa in enriched pitcher ecosystems were similar to those found in larger enriched aquatic ecosystems and are consistent with microbial processes occurring at the base of detrital food webs. Detectable differences between protein profiles of enriched and control ecosystems suggest that a time series of environmental proteomics data may identify protein biomarkers of impending state changes to enriched states.

Key words: aquatic ecosystems; bacterial communities; environmental proteomics; model ecosystem; organic matter enrichment; *Sarracenia purpurea.*

Received 19 April 2017; revised 27 July 2017; accepted 31 July 2017. Corresponding Editor: Debra P. C. Peters.

† **E-mail:** bballif@uvm.edu

INTRODUCTION

Chronic and directional environmental drivers such as nutrient and organic matter enrichment are causing state changes in many ecosystems (Rabalais et al. 2009, Scheffer 2009). Mitigating or preventing these state changes requires predicting them with sufficient lead-time (Biggs et al. 2009). Current prediction methods rely on the statistical signature of "critical slowing down" (Scheffer et al. 2009), an increase in the variance, or temporal autocorrelation of a state variable (Dakos et al. 2015). However, such indicators usually require long time series of data with frequent sampling of an appropriate state variable (Bestelmeyer et al. 2011, Levin and Mollmann 2015). Even when such data are available, the signature of critical slowing down may not provide enough lead-time for intervention (Biggs et al. 2009, Contamin and Ellison 2009).

In aquatic systems, water quality indicators such as total suspended solids (Hargeby et al. 2007), submersed macrophyte vegetation cover (Dennison et al. 1993, Sondergaard et al. 2010), diatom composition (Pan et al. 1996), and phytoplankton biomass (Carpenter et al. 2008) often are used as state variables. However, whether

top-down or bottom-up forces initiate the change, the proximate cause of eutrophication in many freshwater aquatic ecosystems is microbial processes associated with the breakdown of detritus (Chrost and Siuda 2006). A primary reason that it has been difficult to forecast shifts with sufficient lead-time may be that changes in monitored variables lag behind the microbial processes that underlie state changes. We hypothesize that biomarkers linked closely to microbial function, such as proteins, may serve as better early warning signals of impending state changes than traditional aquatic ecosystem biomarkers.

One of the challenges to studying aquatic ecosystem state changes is the lack of replicable natural ecosystems that can be ethically manipulated. Recently, we have identified the aquatic ecosystem that assembles in the cup-shaped leaves of the northern pitcher plant *Sarracenia purpurea* as a model system for identifying whole-ecosystem microbial processes associated with detrital enrichment. Each leaf functions as an independent ecosystem that can be experimentally enriched and monitored through time in the field or laboratory (Srivastava et al. 2004). Arthropod prey, mostly ants and flies, form the base of a "brown" food web that includes dipteran larvae, protozoa, mites, rotifers, and a diverse assemblage of bacteria that decompose and mineralize nearly all the captured prey biomass (Ellison et al. 2003, Butler et al. 2008, Koopman and Carstens 2011, Gray et al. 2012). Even in the absence of macroinvertebrates, the dominant transfer of nutrients to the plant occurs via microbial activity (Butler et al. 2008). With excess organic matter loading, microbial activity increases, pitcher fluid becomes turbid, and oxygen levels collapse to hypoxic conditions even during daytime photosynthesis (Sirota et al. 2013). Such consequences are similar to those seen in larger aquatic ecosystems that have switched from a green to a brown food web dominated by detritivores, as an initial increase in primary production leads to internal organic matter loading and increasing biological oxygen demand as primary producers decompose (Correll 1998).

In the last decade, environmental proteomics has emerged as a powerful tool to measure microbial community function in a variety of aquatic habitats, including contaminated groundwater (Benndorf et al. 2007), coastal upwelling systems (Sowell et al. 2011), estuaries (Colatriano et al. 2015), and meromictic lakes (Lauro et al. 2011). Additionally, environmental proteomics has promise as a tool for identifying biomarkers of changing environmental conditions, including aquatic pollution (Campos et al. 2012, Ullrich et al. 2016). Environmental proteomics looks at the complete set of proteins expressed in an ecosystem at a single time point and gives insight into the function of a community. While metatranscriptomics also serves as an important tool for understanding community function, mRNA and protein levels are generally not strongly correlated (Vogel and Marcotte 2012); this is especially true for bacteria in perturbed systems (Jayapal et al. 2008). Therefore, metaproteomics may provide a more accurate picture of bacterial community function in enriched aquatic habitats.

As a first step toward determining the utility of microbial protein biomarkers as early warning signals of state changes, we conducted an environmental proteomics screen of the aquatic ecosystem in *S. purpurea* pitchers enriched with organic matter to determine whether there are detectable differences between the proteins, associated molecular pathways, and taxa contributing to expressed proteins in microbial (nonviral organisms <30 μm) communities in enriched vs. control ecosystems. We hypothesized that an environmental proteomics survey would reveal detectable differences in taxa contributing to protein expression, proteins, and functional pathways between enriched and control ecosystems. We expected to find differences between control and enriched pitchers in pathways related to respiration and decomposition, changes in the oxygen requirement of microbes contributing to expressed proteins, and shifts in the taxonomic composition of microbes contributing to protein expression. Specifically, we predicted an abundance of contributing anaerobic bacteria in enriched pitchers relative to controls. We identified and found detectable differences in taxa, proteins, and pathways common to a wide range of aquatic ecosystems. Our results suggest that environmental proteomics can be a useful tool for detecting alternative enriched and unenriched states in aquatic ecosystems and may serve as a means to identify protein biomarkers of impending shifts between such states.

Methods

Enrichment experiment

The field experiment was conducted in Tom Swamp, a nutrient-poor fen located at the northern end of Harvard Pond (42.51 N, −72.21 W) at Harvard Forest, Worcester County, Massachusetts. Newly opened pitchers were identified and randomly assigned to an ambient control or detritus-enriched treatment (Appendix S1). Previous work by Peterson et al. (2008) using culture-independent methods revealed that newly opened pitchers are sterile and impermeable to bacteria, so we are reasonably sure that our experimental pitchers did not harbor diverse bacterial communities prior to the start of the experiment. Detritus-enriched pitchers received 1 $mg{\cdot}mL^{-1}{\cdot}d^{-1}$ of oven-dried, finely ground wasps (*Dolichovespula maculate*; Appendix S1), which have elemental ratios (C:N, 5.99:1, N:P:K, 10.7:1.75:1.01) similar to those of *Sarracenia*'s natural ant prey (C:N, 5.9:1; N:P:K, 12.1:1.52:0.93; Farnsworth and Ellison 2008). Proteomic analysis of the ground wasp (not reported here) failed to identify microbial proteins, so we are confident that microbial contribution to enriched pitchers from the wasps was minimal. Enrichment treatments were applied for 14 consecutive days; all pitchers were otherwise unmanipulated. Pitcher fluid was sampled on the first and last days of the experiment, filtered to remove microbes >30 μm, pelleted, and stored at −80°C until processed (Appendix S1).

Protein extraction, SDS-PAGE, and mass spectrometry

Six of ten replicate microbial pellets from each treatment yielded enough protein for analysis via tandem mass spectrometry. All replicates were analyzed separately using sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and Coomassie staining (Fig. 1; Appendix S1: Fig. S1a, b). All six of the enriched pitchers and five of the six control pitchers had visible protein staining levels and were chosen for mass spectrometry. Proteins were subjected to a tryptic digest (Appendix S1) and to LC-MS/MS as previously described (Cheerathodi and Ballif 2011) using a linear ion trap mass spectrometer (Thermo Electron, Waltham, Massachusetts, USA). MS/MS spectra were matched to peptides in a custom protein database using SEQUEST (Thermo Fisher Scientific, Waltham, Massachusetts, USA) software as described below.

Custom metagenomic databases

We generated a custom protein database from a six-frame forward and reverse translation of a metagenomic database constructed from microbial communities of three previously collected pitchers that had captured diverse amounts of prey (Appendix S1: Fig. S2). Pitchers were collected from Molly Bog, an ombrotrophic bog located in Morristown, Vermont (44.50 N, −72.64 W), on 18 August 2008 and transported in a cooler directly from the field to the University of Vermont. Microbial pellets were obtained immediately as described above. DNA was extracted, prepared, and sent for library construction, sequencing, and assembly to Genome Quebéc (Montréal, Quebéc, Canada) with the 454 GS-FLX Titanium Sequencing System (Roche, Basel, Switzerland; Appendix S1). Contigs were assembled de novo with Roche's Newbler assembler v2.3 (release 091027_1459) using default parameters (minimum read length = 20; overlap seed step = 12; overlap seed length = 16; overlap min seed count = 1; overlap seed hit limit = 70; overlap min match length = 40; overlap min match identity = 90; overlap match ident score = 2; overlap match diff score = −3; overlap match unique thresh = 12; map min contig depth = 1; all contig thresh = 100), with the exception of minimum read length (20 bp) and overlap hit position limit (1,000,000). The assembled contigs were imported into MG-RAST 4.0.2 (Argonne National Laboratory, Argonne, Illinois, USA) to assess functional and taxonomic potential (Meyer et al. 2008). Taxonomic assignments were visualized using the Krona plugin and the following cutoffs were applied to both taxonomic and subsystem functional category assignments: minimum identity = 60%, *e*-value of 1×10^{-5} or less, and a minimum alignment length of 15 bp (Appendix S1: Fig. S3). We calculated Hurlbert's probability of an interspecific encounter (PIE) to estimate the evenness of bacterial classes in the metagenome (Hurlbert 1971; Appendix S1). Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways (level 2 and level 3) were assigned to contigs using the KEGG database via MG-RAST (we report only the top 73 level 3 pathways here; Appendix S1: Fig. S4).

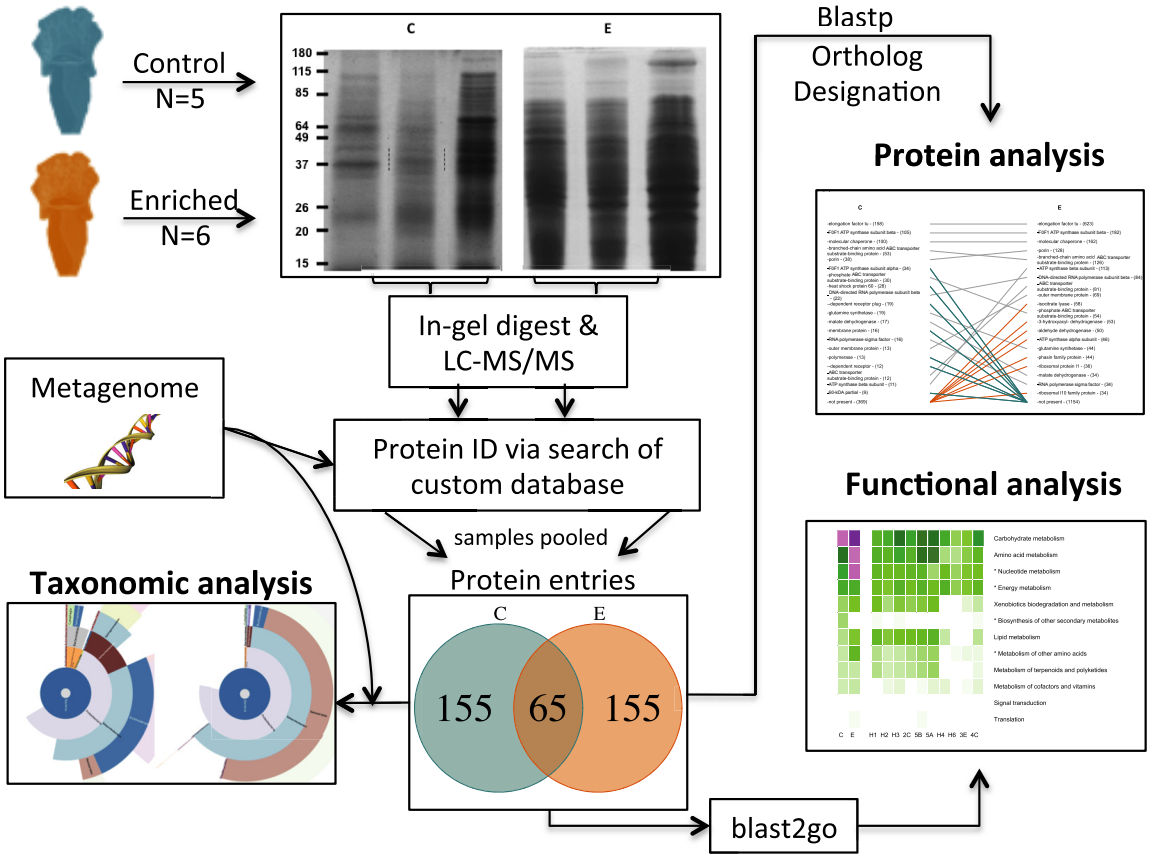

Fig. 1. Pipeline for data collection and analysis. Proteins from the microbial communities in experimentally enriched and ambient control pitcher fluid were processed using SDS-PAGE, tryptic digest, LC-MS/MS, and a SEQUEST search of a custom metagenomic database. The composition of microbial communities was determined using a BLAST homology search of metagenomic data associated with identified proteins. Protein identity and annotation was determined via a blastp search to identify orthologs and blast2go.

A metaproteomic database was created with a six-frame forward and reverse translation of the assembled metagenome using open-source Ruby software (http://www.ruby-lang.org). Sequences with >100 amino acids (n = 184,128) in length were retained. A decoy database was constructed by reversing the retained sequences and concatenating them to the forward database to allow for an estimation of the false discovery rate as has been described (Elias and Gygi 2007).

Protein orthologue identification

Peptide and protein identifications were made via a SEQUEST search of the tandem mass spectral data against the custom pitcher plant microbial community protein database described above (Appendix S1). The number of protein hits varied substantially among replicates, so to have enough proteins for treatment comparisons, peptides and proteins from the five control samples and six enriched samples were pooled after LC-MS/MS and the SEQUEST search into a single control and a single enriched sample dataset. The doubly and triply charged peptide ions were further considered and each dataset was filtered by first adjusting the cutoffs for X_{corr} and ΔC_n until the false discovery rate was <10%. The final filters were as follows: X_{corr} ≥3.0 for doubly charged ions, X_{corr} ≥3.3 for triply charged ions, and unique Δcorr ≥0.15. The resulting list of protein hits for each treatment was then ranked by unique number of peptides and the top 220 proteins from each treatment were selected so that the false discovery rate for control and enriched

treatments was 6.6% and 0%, respectively. These top 220 proteins and their associated peptides are found in Data S1.

In the list of control peptides, a protein hit from the decoy database was represented by 25 total peptides; therefore, we suspected that this hit was a true positive not represented in our target database. However, a BLAST search of the full amino acid sequence did not yield an identical match, so we cannot definitively claim it is a true positive; therefore, we removed this peptide from our top 220 list of control peptides. With this peptide removed, the false discovery rate for the control treatment was 4.3%.

All peptide hits were pooled within treatments and mapped back to their source sequences in the custom protein database. Those source sequences were imported in fasta file format into blast2go v.2.8.0 (Conesa et al. 2005) for identification and annotation using the following configuration settings: blastp program, blast expect value of 1.0E-3, 10 blast hits, annotation cutoff >55, GO (gene ontology) weight >5.

Analysis of the top proteins shared between treatments

A randomization test was done using RStudio (v. 0.98.1059, RStudio, Boston, Massachusetts, USA) to test the hypothesis that there was a single common protein pool for both the control and enriched treatments and that the number of observed shared proteins between treatments reflects chance effects resulting from random draws from this single protein pool (Appendix S1). We conducted an additional simulation in R to determine the likelihood of a type I error in our randomization test (Appendix S1).

Comparison of the top 20 proteins from each treatment

We downloaded the sequence by annotation file from the blast2go search for each treatment to get the protein names associated with each protein hit (sequence description in blast2go). Each of the top 220 identified proteins in each treatment, ordered by the number of total peptides associated with the protein hit, was matched to a protein name using R software. If multiple protein hits within a treatment matched a single protein name, the protein names were merged in silico and the total peptides representing them were summed. Protein names were ranked in order of the abundance of total peptides for each treatment.

Taxonomic analysis

To determine the taxonomic composition of the microbes contributing to identified proteins in our treatments, we conducted a BLAST homology search of the metagenomic sequence data for protein hits. All peptides from the top 220 identified proteins in each treatment were mapped back to their contigs of origin to obtain nucleotide sequences. Because contigs were at least 500 base pairs in length, we felt confident that a BLAST search of the nucleotide sequences would yield correct taxonomic identifications at course taxonomic levels and acknowledge that ambiguity can remain in the taxonomic identification from a metacommunity at genus and species levels. The top BLAST hit was retained for each nucleotide sequence associated with an identified protein and linked to a bacterial class (Appendix S1). For each bacterial class identified, a 2×2 contingency table was created with treatments as columns and the number of peptides associated and not associated with the taxon as rows. A chi-square test was then used to determine whether the abundance of the bacterial class was significantly different between treatments. All P values were adjusted using the Benjamini–Hochberg method (Benjamini and Hochberg 1995; Table 1). Species composition was visualized using Krona (Ondov et al. 2011;

Table 1. Results of chi-square analysis of bacterial classes in control and enriched pitchers.

Class	Control peptides	Enriched peptides	Adjusted chi-square
Acidobacteria	6	0	0.000
Actinobacteria	32	3	**0.000**
Alphaproteobacteria	276	196	**0.000**
Bacteroidia	12	16	0.059
Betaproteobacteria	469	2448	**0.000**
Chloroflexi	0	3	0.816
Clostridia	3	7	0.959
Cytophagia	14	17	**0.021**
Deltaproteobacteria	2	0	0.132
Flavobacteria	0	3	0.816
Gammaproteobacteria	50	146	0.816
Gloeobacteria	8	0	0.000
Sphingobacteria	48	53	**0.000**
Spirochaetia	11	9	**0.006**

Note: Values in boldface are those in which the adjusted P value is <0.05.

Appendix S1: Fig. S5). In addition to the BLAST homology search, we used Unipept (Mesuere et al. 2016) to map tryptic peptides to the UniprotKB database and retrieve the least common taxonomic ancestor (most derived shared taxonomic node) associated with each peptide for pooled replicates (Appendix S1: Fig. S6). We calculated Hurlbert's PIE to estimate the evenness of bacterial classes contributing to expressed proteins in control and enriched pitchers (Hurlbert 1971; Appendix S1).

Functional analysis

Functional pathways (two levels) associated with each identified protein from each treatment were retrieved using the KEGG (Kanehisa et al. 2014) mapping function of blast2go v.2.8.0. Each pathway was weighted by the total number of peptides associated with protein hits, or the number of spectral counts, mapping to that pathway (Appendix S1: Fig. S7). For each pathway identified, a 2 × 2 contingency table was created with treatments as columns and the number of peptides associated and not associated with the pathway as rows. A chi-square test was used to determine whether each pathway was significantly over- or under-represented in enriched pitchers relative to controls. All *P* values were adjusted using the Benjamini–Hochberg method (Benjamini and Hochberg 1995; Appendix S1: Table S1).

To determine whether bacteria contributing to expressed proteins in control and enriched ecosystems differed in their O_2 requirements, we mapped each bacterial species identified in our BLAST search to its O_2 requirement using data from the Integrated Microbial Genomes database (IMG; Timinskas et al. 2014, Reddy et al. 2015; Appendix S1). The IMG database contains six classes of O_2 requirements: aerobe, anaerobe, facultative, microaerophilic, obligate aerobe, and obligate anaerobe. The latter three categories make up <7% of the database. We merged any species classified as obligate aerobes or obligate anaerobes into the aerobe and anaerobe classes, respectively.

Analysis of unpooled data

In addition to analyzing pooled data, we used ordination and permutation analyses to determine the effect of enrichment on microbial community protein expression, taxonomic contribution to expressed proteins at the class and family levels, and KEGG pathways. We tested the similarity within and among replicates of control and enriched microbial communities using ADONIS, a nonparametric permutation test in the "vegan" package (v. 2.4.1) in R (Oksanen et al. 2016). We used a multivariate homogeneity of group dispersions test (betadisper function in the "vegan" package) to determine whether the composition of contributing microbial taxa was more divergent in control replicates than in enriched replicates. The permutation tests used 999 permutations and were done using total peptide counts associated with protein identifications, microbial classes, microbial families, and KEGG pathways (Table 2). To visualize the similarities among replicate ecosystems, we used the "vegan" package function metaMDS to perform non-metric multidimensional scaling ordination using Bray–Curtis distances. Data were square-root-transformed and standardized using Wisconsin double standardization. To determine which taxa contributed the most to Bray–Curtis dissimilarity of taxa contributing to protein expression between the treatments, we did a similarity percentages test using the simper function in the "vegan" package.

Results

From 243 Mb of DNA sequence information, roughly 54% of 567,549 filtered reads (median

Table 2. Effect of treatment on microbial proteins, contributing taxa (class and family), and pathways.

	Proteins				Taxa (Class)				Taxa (Family)				KEGG pathways			
Variable	df	*F*	R^2	*P*	df	*F*	R^2	*P*	df	*F*	R^2	*P*	df	*F*	R^2	*P*
Treatment	1	4.217	0.319	**0.004**	1	3.766	0.295	**0.022**	1	4.218	0.319	**0.003**	1	4.753	0.373	**0.024**
Residuals	9		0.681		9		0.705		9		0.681		9		0.627	
Total	10		1.000		10		1.000		10		1.000		10		1.000	

Notes: KEGG, Kyoto Encyclopedia of Genes and Genomes; df, degrees of freedom. Values in boldface are those in which the adjusted *P* value is <0.05.

read length = 482 bp) were assembled into 26,713 contigs ranging from 500 to 43,200 bp (N50 = 1135; Appendix S1: Fig. S2b, c). All the contigs passed MG-RAST quality control. The metagenome was dominated by bacteria (99.11%) at the domain level. The top five bacterial classes were Betaproteobacteria (31.99%), Alphaproteobacteria (19.42%), Sphingobacteria (13.32%), Gammaproteobacteria (10.10%), and Acidobacteria (7.04%). The top five genera comprising the genome were Burkholderia (8.87%), Variovorax (6.50%), Pedobacter (5.24%), Mucilaginibacter (4.04%), and Lutiella (3.91%). Within the metagenome, 23% of aligned contigs were mapped to the order Burkholderiales, while only 7% mapped to Neisseriales (Appendix S1: Fig. S3). Taxonomic evenness of the metagenome, calculated using Hurlbert's PIE, was equal to 0.79.

Representation of the contigs mapping to functional pathways was dominated by amino acid metabolism (20.6%), followed by membrane transport (12.9%), carbohydrate metabolism (11.9%), translation (7.2%), and metabolism of cofactors and vitamins (6.4%). Within amino acid metabolism, pathways were represented primarily by glycine, serine, and threonine metabolism (17.1%), alanine, aspartate, and glutamate metabolism (13.8%), and valine, leucine, and isoleucine degradation (12.7%). Membrane transport was represented by ABC transporters (78.2%), bacterial secretion system (19.4%), and phosphotransferase system (2.4%). Carbohydrate metabolism was dominated by pyruvate metabolism (13.9%), glycolysis/glucogenesis (12.6%), and pentose phosphate pathway (11.6%). Overall, the top five level 3 KEGG categories included ABC transporters (10.1%), two-component system (4.8%), aminoacyl-tRNA biosynthesis (3.8%), glycine, serine, and threonine metabolism (3.5%), and ribosome (3.3%; Appendix S1: Fig. S4).

We identified a total of 986 proteins in the enriched treatment and 616 proteins in the control treatment. Of the 220 most abundant protein identifications for each treatment, 65 were shared between treatments leaving 155 unique to each treatment (Fig. 2a). The randomization test revealed significantly fewer protein hits shared between the treatments than expected by chance (Fig. 2b). In both treatments, the top three of the 20 most abundant proteins, as measured by the total number of matched peptides (spectral counts), were the same in the control and enriched treatments. However, the relative abundance of the remaining 17 proteins in this top list differed strongly between treatments, with only seven of the 20 proteins unique to each treatment (Fig. 2c).

The majority of identified proteins were associated with bacteria. The most common microbial class contributing to identified proteins in both treatments was Betaproteobacteria, but the contribution was higher in enriched (84.4%) vs. control (50.3%) treatments (Table 1, Fig. 3a; Appendix S1: Figs. S5, S6). This difference was driven by a higher abundance of Alphaproteobacteria in multiple families, including Sphingobacteriaceae, Phyllobacteriaceae, Xanthomonadaceae, and Rhizobiaceae, in control ecosystems relative to the enriched ecosystems. The similarity percentages test identified Betaproteobacteria (38.8%) and Alphaproteobacteria (9.9%) as the main contributors to dissimilarity of active microbial class composition between treatments and Neisseriaceae (23.8%) and Comamonadaceae (9.7%) as the main contributors to active microbial family dissimilarity between treatments. Although both treatments yielded similar numbers of identified microbial classes (control = 12, enriched = 11), taxonomic evenness of microbial classes contributing to identified proteins was substantially higher in the controls (PIE = 0.71) than in the enriched pitchers (PIE = 0.31). Similar taxonomic profiles were obtained using Unipept's search for the least common taxonomic ancestors of the pooled data (Appendix S1: Fig. S6). For the unpooled data, taxonomic and functional variability among treatments was greater than variability among replicate ecosystems within treatments (Figs. 3, 4). Multivariate analysis of group dispersion revealed that composition of microbes contributing to protein expression was significantly more variable in control replicates than in enriched replicates at both the family ($P = 0.003$) and class ($P = 0.023$) levels.

The BLAST search yielded taxonomic assignments for 191 and 173 of the 220 sequences in enriched and control treatments, respectively, and all *E*-values were $<10^{-5}$. Of top species hits identified in the BLAST search, *Variovorax paradoxus* and *Chromobacterium violaceum* were the only two of the most six abundant "species" contributing to identified proteins common to both

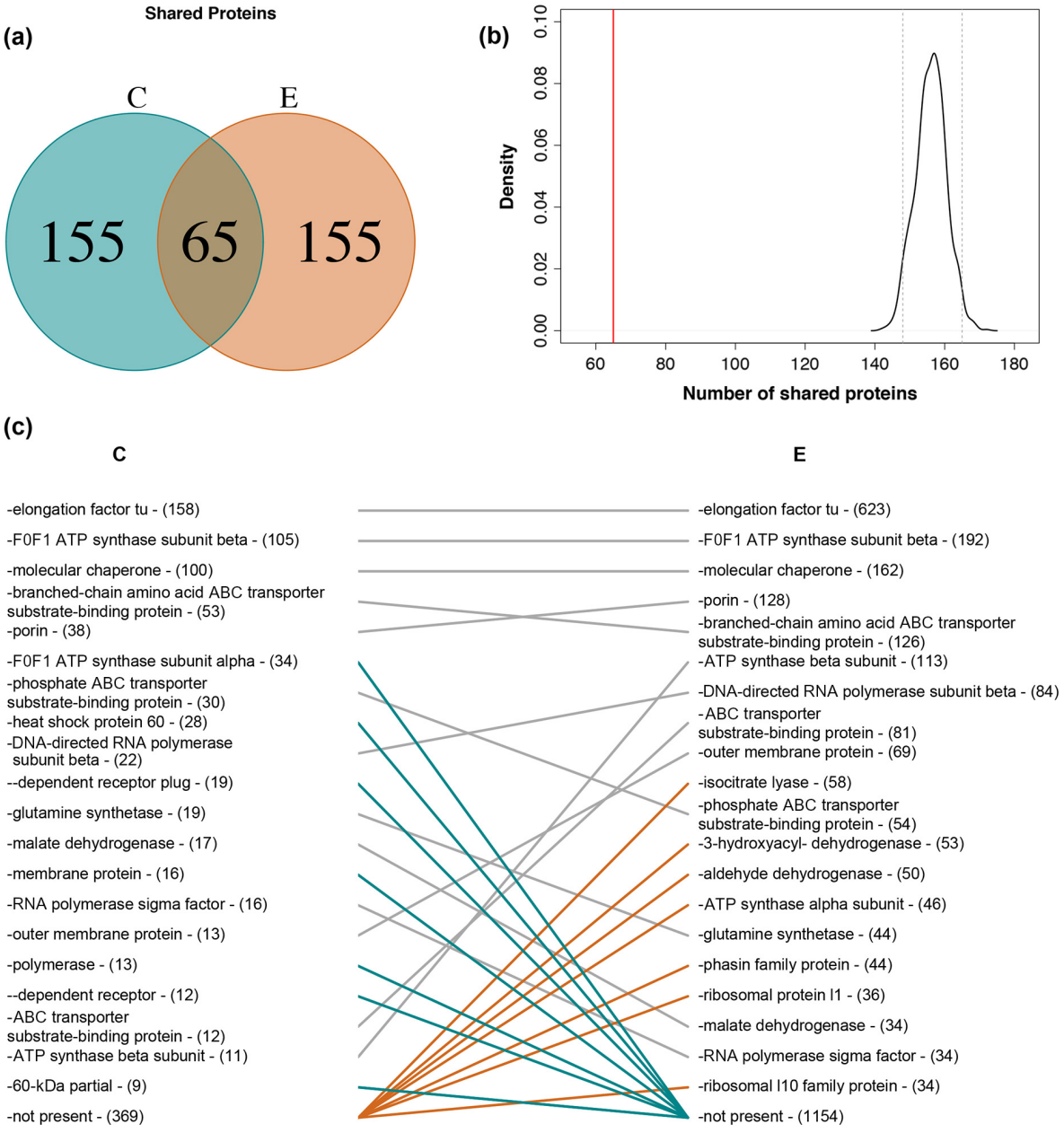

Fig. 2. Protein identifications differed between control and enriched pitchers. (a) Protein hits shared between control and enriched treatments. (b) Results of a randomization test in which 220 protein hits were randomly assigned to each treatment and the number of shared protein hits was calculated. Red line indicates the actual shared number of proteins. Gray probability density function indicates the 95% confidence interval for the simulated shared protein hit values. (c) Top 20 proteins in rank order for each treatment. Proteins are ranked by the number of total peptides associated with them (in parentheses). Identical proteins in both treatments are connected by lines. Blue lines indicate proteins unique to the top 20 in control pitchers (C), and brown lines indicate proteins unique to the top 20 in enriched pitchers (E).

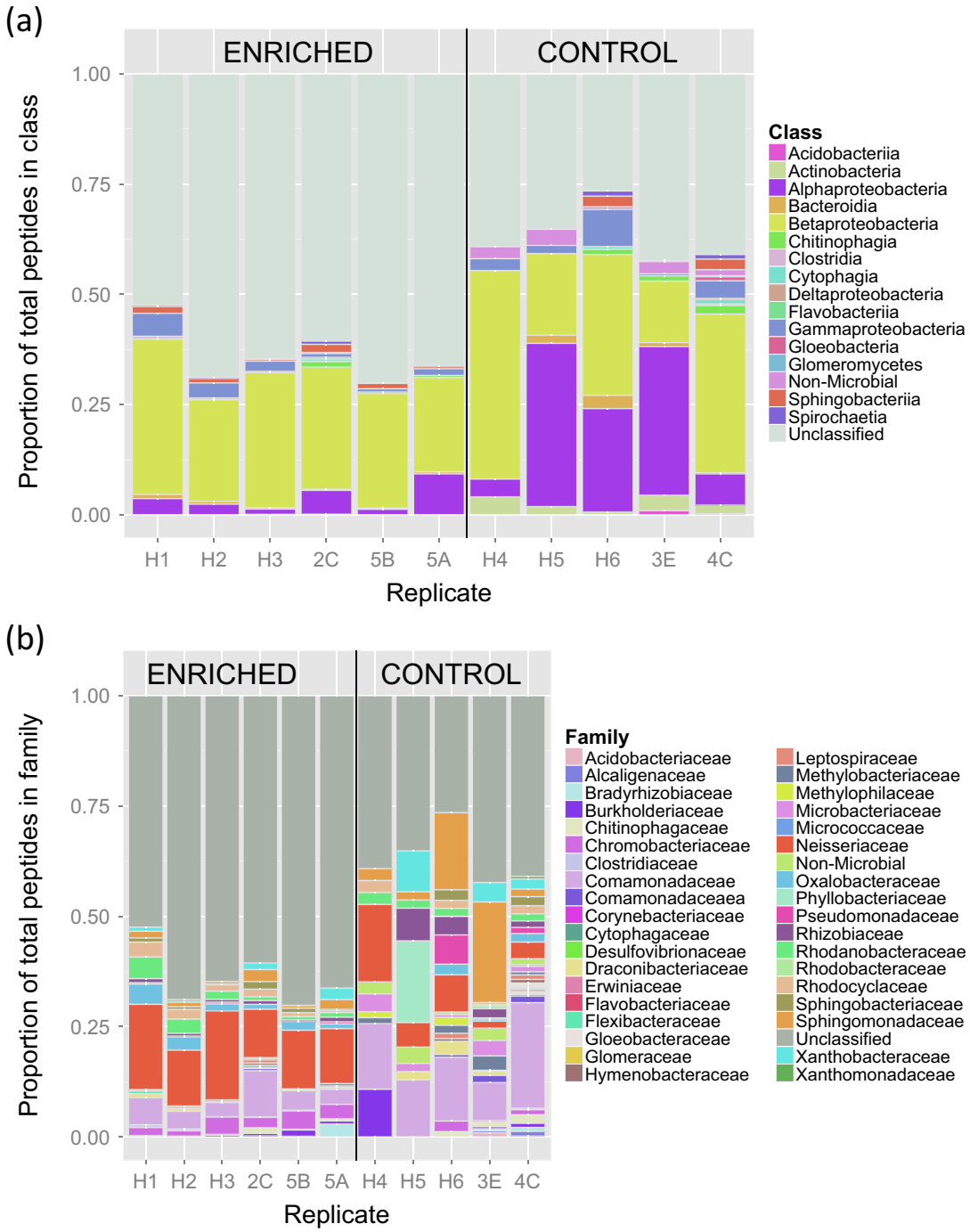

Fig. 3. Distinctly different microbial communities contributed to protein expression in control and enriched pitchers. The proportion of total peptides from the top 220 proteins associated with particular microbial classes (a) and families (b) in all enriched and control replicate pitchers.

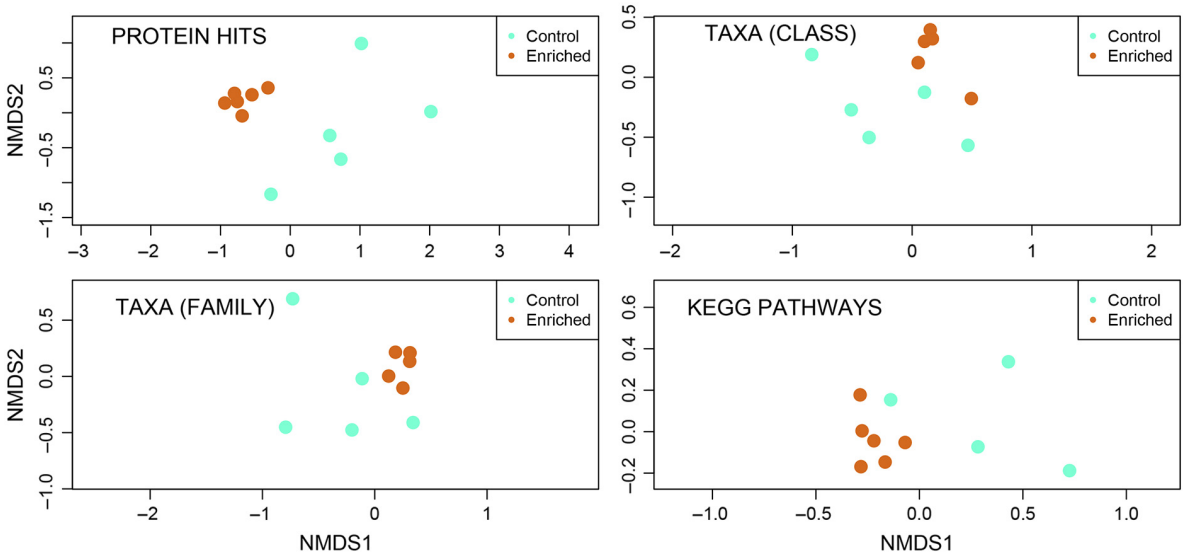

Fig. 4. Microbial communities in control and enriched pitchers differ in the proteins they produce, taxa that contribute to protein expression, and function. Ordination of Bray–Curtis dissimilarities of total peptides shows clustering of pitcher microbial communities by treatment for protein hits (adonis $P = 0.004$), microbial classes (adonis $P = 0.022$), microbial families (adonis $P = 0.003$), and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways (adonis $P = 0.003$) as a function of treatment (control or enriched).

treatments. *Novosphingobium aromaticivorans*, *Starkeya novella*, *Sphingomonas wittichii*, and *Sphingomonas* sp. were among the six most abundant contributors in control pitchers. *Pseudogulbenkiania* sp., *Rhodanobacter denitrificans*, *Janthinobacterium* sp., and *Dechlorosoma suillum* were among the six most abundant contributors in enriched pitchers (Appendix S1: Table S2). Obligate aerobic bacteria contributed the most to identified proteins in the control pitchers, while facultative anaerobic bacteria contributed the most in enriched pitchers (Fig. 5b).

Functional pathways represented by the top 220 expressed microbial proteins also differed between control and enriched pitchers. We detected significant differences in metabolic pathways, including those involved in the metabolism of amino acids, carbohydrates, lipids, secondary metabolites, cofactors and vitamins, and terpenoids and polyketides (Appendix S1: Table S1, Figs. S7, S8a) and, at courser pathway levels, energy metabolism, nucleotide metabolism, and amino acid metabolism (Fig. 5a). In the control treatment, 161 of the top 220 protein hits were not assigned to a KEGG pathway (represented by 906 total peptides). Of the 220 top protein hits in the enriched treatment, 129 were not assigned to a pathway (represented by 2375 total peptides).

Discussion

We hypothesized that there would be detectable differences in the taxonomic composition of microbes contributing to expressed proteins. Indeed, we observed striking differences between unenriched and enriched ecosystems in the taxonomic composition of the microbes contributing to identified proteins (Fig. 3). The taxonomic composition of bacteria contributing to protein expression in our study, and in our metagenome, is consistent with findings of previous studies of bacterial communities in *Sarracenia* species. *Sarracenia purpurea* pitchers contain more than 1000 species of bacteria and a negligible amount of archaea (Paisie et al., 2014). One genomic study of *Sarracenia alata* pitcher bacterial communities revealed an abundance of Proteobacteria (primarily Gammaproteobacteria). Taxonomic groups within the Betaproteobacteria had relative abundance similar to our metagenome and to control pitcher communities in our experiment, with a

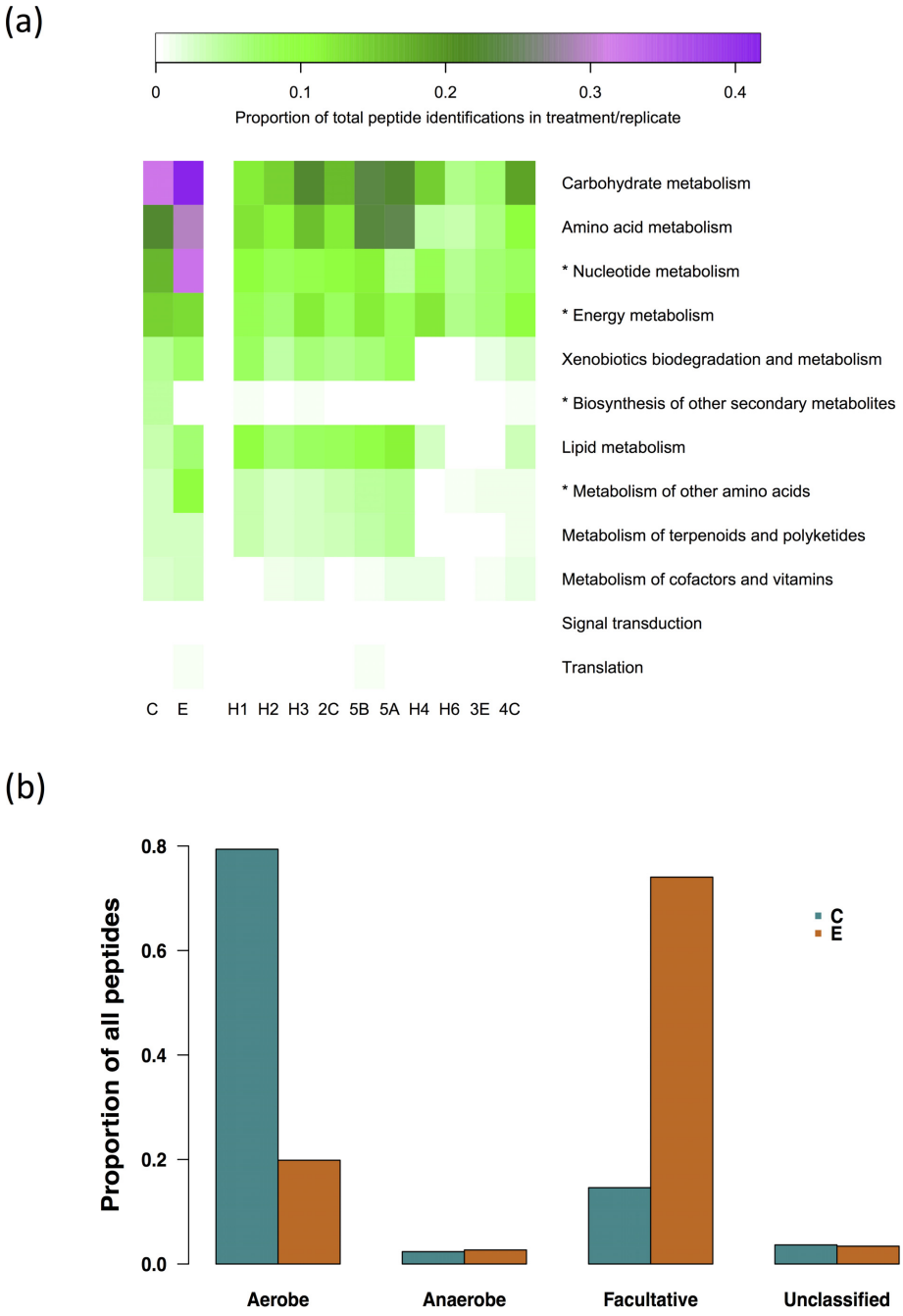

Fig. 5. Microbial function differed between control and enriched pitchers. (a) Heat map of the proportional representation of course-level Kyoto Encyclopedia of Genes and Genomes pathways between control pitchers (C) and enriched pitchers (E) and individual control (H4, H6, E3, C4) and enriched (H1, H2, H3, 2C, 5B, 5A) replicates. Significantly different pathways between pooled control and enriched samples are indicated with an asterisk. (b) Oxygen requirement of microbial classes contributing to protein expression as a proportion of all peptides in control (C) and enriched (E) pitchers.

high percentage of sequences derived from Burkholderiales and a lower proportion from the Neisseriales (Koopman et al. 2010). A study of sub-habitats in *S. purpurea* revealed an abundance of Betaproteobacteria (primarily Burkholderiales) on the pitcher walls and in the sediment, co-dominance in pitcher liquid by Beta- and Alphaproteobacteria, and the presence of Bacteroidetes and Firmicutes, though in a low proportion, in the sediment, fluid, and pitcher walls (Krieger and Kourtev 2012). This finding is fairly consistent with the taxonomic potential revealed by our metagenome, in which 35%, 23%, 14%, and 1% of identified contigs were mapped to Betaproteobacteria, Alphaproteobacteria, Bacteroidetes, and Firmicutes, respectively. Gray et al. (2012) found that *S. purpurea* pitchers were composed primarily of Proteobacteria and Bacteroidetes, with Gammaproteobacteria, Alphaproteobacteria, or Betaproteobacteria dominating within the Proteobacteria, but that taxonomic composition varied from pitcher to pitcher within and across geographic regions.

The composition of bacteria contributing to protein expression in our experiment varied between control replicates, much more so than between enriched pitcher communities. This pattern is likely the result of a combination of factors. First, pitchers contain distinct sub-habitats that vary in light availability and concentration of dissolved oxygen and organic matter and therefore provide multiple habitats for a diverse set of microbes (Krieger and Kourtev 2012). As organic matter enrichment increases biological oxygen demand, the subsequent decline in dissolved oxygen may create a more homogenous oxygen environment such that microbes sensitive to oxygen conditions can no longer compete against low-oxygen-tolerant bacteria, decreasing bacterial diversity.

Low bacterial diversity in enriched pitchers echoes findings in larger enriched aquatic ecosystems. Analysis of the 16S rRNA gene product of bacterial communities in nutrient-enriched salt marsh sediments revealed that the bacterial diversity of active bacteria decreased relative to that of communities in unenriched sediments (Kearns et al. 2016). Similarly, enrichment of heterotrophic stream biofilm communities yielded lowered diversity; however, in contrast to our enriched pitcher communities, the stream biofilm communities diverged in composition (Van Horn et al. 2011).

The composition of microbes contributing to protein expression in *S. purpurea* pitchers was similar to the composition of larger freshwater aquatic ecosystems. Betaproteobacteria dominated microbes contributing to protein expression in both enriched and control pitchers though in higher abundance in enriched pitchers relative to control pitchers. Betaproteobacteria are generally the most abundant class of bacteria in freshwater lakes (Percent et al. 2008, Newton et al. 2011) and dominate contaminated sediments (Haller et al. 2011) and organic aggregates in eutrophic lakes (Tang et al. 2009). Betaproteobacteria populations associated with the beta II clade have been shown to increase rapidly with the addition of organic carbon in humic lakes (Burkert et al. 2003, Kent et al. 2006). Furthermore, experimental dissolved organic matter additions to microcosms containing alpine lake bacteria cultures led to a near-dominance of Betaproteobacteria, suggesting that these bacteria are good competitors in enriched aquatic ecosystems (Perez and Sommaruga 2006). These results suggest that bacterial communities in *S. purpurea* pitchers are structured and behave like bacterial communities in larger lakes and ponds in response to enrichment. It is important to note that most existing literature on freshwater bacteria and *S. purpurea* bacterial communities relies primarily on genomic methods for identification and therefore is likely capturing functionally active and inactive bacteria, whereas our methods are capturing only the functionally active bacteria. As a result, we use caution when directly comparing the results of our study to those in larger aquatic ecosystems. However, the Unipept search of our identified tryptic peptides and NCBI BLAST search of their contigs of origin yielded remarkably similar results (Fig. 3a; Appendix S1: Fig. S6), suggesting that tryptic peptides could be used to correctly identify microbes contributing to identified proteins, though at coarser taxonomic levels than can be achieved by nucleic acid analysis.

We hypothesized that there would be detectable differences in the function of microbial communities in control and enriched pitchers. We measured function in two ways: First, we mapped identified bacterial classes associated

with proteins to their oxygen requirements and second, we mapped peptides to functional KEGG pathways. Oxygen requirements differed significantly between taxa contributing to protein expression in control and enriched microbial communities. Bacteria contributing to protein expression in control pitchers were predominately aerobic, whereas bacteria contributing to protein expression in enriched pitchers were primarily facultatively anaerobic. The difference in oxygen requirement of contributing bacteria between the two treatments was driven largely by two taxa: the obligate aerobe *Variovorax paradoxus* (28.4% of total peptides in the control treatment and 7.2% in the enriched treatment) and the facultative anaerobe *Chromobacterium violaceum* (53.3% of total peptides in the enriched treatment and 6.6% in the control treatment; Appendix S1: Table S2). Peptides that mapped to *C. violaceum* in the BLAST search mapped in the Unipept search to *Aquitalea magnusonii*, a Betaproteobacteria most closely related to *C. violaceum*, isolated from a humic lake in Wisconsin, USA (Lau et al. 2006). Although we did not measure dissolved oxygen during the field experiment, enriched pitchers in a subsequent experiment enriched with the same concentration of organic matter became hypoxic within 48 h, suggesting that pitchers in the field were likely hypoxic (Sirota et al. 2013). Dissolved oxygen concentration is one of three primary drivers of bacterial community composition in eutrophic, dimictic lakes (Shade et al. 2007) and appears to also drive the composition of functionally active bacteria in enriched *S. purpurea* pitchers.

We expected to see a high proportion of obligate anaerobic bacteria in enriched pitchers. Bacteroidetes and Firmicutes, to a lesser degree, have been found to inhabit *S. purpurea* pitchers (Krieger and Kourtev 2012); however, we identified very few proteins associated with these taxa. Of the 3008 and 969 peptides associated with the top 220 proteins in enriched and control treatments, respectively, we found only 17 peptides associated with obligate anaerobes in the enriched pitchers (seven of which were associated with Firmicutes) and 13 associated with obligate anaerobes in the control pitchers (three of which were associated with Firmicutes). Though we did find a higher number of peptides associated with Bacteroidetes (74 peptides in control pitchers and 89 in enriched pitchers), they were facultative anaerobes and not strict anaerobes. It is likely that the low numbers of identified peptides associated with these taxa in experimental pitchers are the result of a skewed protein database. Our database was built using metagenomic data from pitchers in the field, the majority of which are oxygen rich (Adlassnig et al. 2011), and likely contained nucleotide sequences primarily from aerobic and facultative anaerobic bacteria. Additionally, pitchers are generally oxygen rich due to photosynthetic activity of the plant and therefore primarily harbor aerobic inquilines (Adlassnig et al. 2011). Even when dissolved oxygen is low, there is a constant flux of oxygen into the pitcher fluid and so the pitchers are rarely ever truly anoxic. It is not surprising, therefore, that peptides associated with anaerobic bacteria were rare. In the absence of a fully representative database, we feel that the higher number of proteins represented by facultative bacteria in enriched pitchers relative to control pitchers is a good indicator of changing oxygen conditions. These results are consistent with the shift to a hypoxic state when *S. purpurea* is enriched with additional prey (Sirota et al. 2013).

We assigned KEGG pathways to contigs in the metagenome and to protein identifications in the metaproteomes to compare microbial community function between control and enriched pitchers, and between the metaproteomes and functional potential in the metagenome. Not surprisingly, the functional potential revealed by the metagenome differed from function revealed by the metaproteomes. Amino acid metabolism and carbohydrate metabolism were represented in the top five rank-ordered pathways in both the metaproteomes and the metagenome; however, carbohydrate metabolism was ranked first in the metaproteomes (~34–40% of total peptides) and third in the metagenome (~12% of mapped contigs). Nucleotide metabolism and energy metabolism were represented in the top five in the metaproteomes (~18% of total peptides in controls and ~34% of total peptides in enriched pitchers), but were ranked ninth (~4%) and seventh (~5%) in the metagenome, respectively. Such differences could be a result of not all nucleotide sequences being transcribed and translated to proteins, but may also be an artifact of only including 220 proteins from each treatment in the metaproteome analysis.

We hesitate to hypothesize broader relevance of our functional pathway results for two reasons. First, we are most interested in the identification of proteins that can serve as biomarkers of aquatic ecosystem state changes. Although we expect that functional information will be useful for determining the utility and generality of such biomarkers, it is not necessary for finding useful biomarkers. Second, it seems impossible, with our limited data, to identify a complete set of functions. With that caveat, we found that coarse KEGG pathway assignments differed between control and enriched microbial communities. Enriched pitchers contained significantly more microbial biomass, as evidenced by the size of the microbial pellets post-centrifugation. When samples were pooled and total peptide counts were normalized, chi-square analysis revealed an enrichment of peptides associated with energy metabolism in enriched pitchers.

These results are consistent with patterns seen in larger aquatic ecosystems: Mineralization of organic matter, an effect of microbial energy metabolism, has been shown to increase along trophic gradients, with bacteria contributing most to mineralization in eutrophic freshwater lakes (Simcic 2005). Not surprisingly, peptides associated with processes requiring oxygen including oxidative phosphorylation and the citric acid cycle were enriched in oxygen-rich control pitcher microbial communities. One protein associated with the citric acid cycle, isocitrate lyase, was present in the top 20 rank-ordered protein identifications in the enriched treatment, but not in the control treatment. This protein, which has been found to be upregulated during periods of oxygen depletion in *Mycobacterium tuberculosis* (Wayne and Lin 1982), could be a candidate biomarker for an impeding tipping point in the *S. purpurea* microecosystem. Though we did not find a significant difference in lipid metabolism pathways between control and enriched pitcher proteins, there was a trend for increased pathway representation of unsaturated fatty acid biosynthesis and fatty acid elongation in enriched pitchers. Such an increase has been found in bacteria in low-oxygen or anaerobic conditions, primarily resulting from an increase in membrane lipids (Lemmer et al. 2015). While these differences do not immediately reveal a functional explanation, it is promising that there were signatures of detectable differences in the protein profiles between treatments. Such differences imply that there are changes in the expression of the most abundant proteins in the most abundant taxa related to organic matter loading.

In larger aquatic systems, traditional water quality indicators may not provide enough lead-time to forecast a tipping point (Contamin and Ellison 2009), especially if they lag behind changes in the microbial community. We hypothesize that microbial proteins may be more sensitive and timely indicators of impending tipping points than traditional chemical markers of water quality. We argue that even though metatranscriptomic and metagenomic methods have superior throughput, metaproteomic methods can inexpensively and rapidly simultaneously characterize the function and (indirectly) composition of the active microbial community members responsible for processes related to aquatic ecosystem state changes. Our study includes a semi-quantitative small initial sampling at only a single time point and therefore does not yet enable a comprehensive enough proteomic analysis to determine the identity of biomarkers or place them in an ecological context. Future studies using more sensitive instrumentation will allow for the identification of a larger number of proteins. Time series of environmental proteomics data and quantitative analysis of changes in protein abundance prior to state changes will allow for the identification and ecological characterization of tipping point biomarkers.

Acknowledgments

This work was funded by the National Science Foundation (Grant Numbers 1144055 and 1144056). Research reported in this publication was supported by an Institutional Development Award (IDeA) from the National Institute of General Medical Sciences of the National Institutes of Health under grant number P20GM103449. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of NIGMS or NIH. The authors thank Hailee Tenander for assisting with preparation of samples for mass spectrometry analysis. The authors declare no conflict of interest. This Whole Genome Shotgun Project has been deposited at DDB/ENA/GenBank under the accession NMRC01000000. The version described in this paper is version NMRC01000000. The

protein database and all codes used to analyze the data are freely available on the Harvard Forest Data Archive under ID Number HF295.

Literature Cited

Adlassnig, W., M. Peroutka, and T. Lendi. 2011. Traps of carnivorous pitcher plants as a habitat: composition of the fluid, biodiversity and mutualistic activities. Annals of Botany 107:181–194.

Benjamini, Y., and Y. Hochberg. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society Series B 57:289–300.

Benndorf, D., G. U. Balcke, H. Harms, and M. von Bergen. 2007. Functional metaproteome analysis of protein extracts from contaminated soil and groundwater. ISME Journal 1:224–234.

Bestelmeyer, B. T., et al. 2011. Analysis of abrupt transitions in ecological systems. Ecosphere 2:129. https://doi.org/10.1890/ES11-00216.1

Biggs, R., S. R. Carpenter, and W. A. Brock. 2009. Turning back from the brink: detecting an impending regime shift in time to avert it. Proceedings of the National Academy of Sciences USA 106:826–831.

Burkert, U., F. Warnecke, D. Babenzien, E. Zwirnmann, and J. Pernthaler. 2003. Members of a readily enriched beta-proteobacterial clade are common in surface waters of a humic lake. Applied and Environmental Microbiology 69:6550–6559.

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

Campos, A., S. Tedesco, V. Vasconcelos, and S. Cristobal. 2012. Proteomic research in bivalves: towards the identification of molecular markers of aquatic pollution. Journal of Proteomics 75:4346–4359.

Carpenter, S. R., W. A. Brock, J. J. Cole, J. F. Kitchell, and M. L. Pace. 2008. Leading indicators of trophic cascades. Ecology Letters 11:128–138.

Cheerathodi, M., and B. A. Ballif. 2011. Identification of CrkL-SH3 binding proteins from embryonic murine brain: implications for reelin signaling during brain development. Journal of Proteome Research 10:4453–4462.

Chrost, R. J., and W. Siuda. 2006. Microbial production, utilization, and enzymatic degradation of organic matter in the upper trophogenic layer in the pelagial zone of lakes along a eutrophication gradient. Limnology and Oceanography 51:749–762.

Colatriano, D., A. Ramachandran, E. Yergeau, R. Maranger, Y. Gelinas, and D. A. Walsh. 2015. Metaproteomics of aquatic microbial communities in a deep and stratified estuary. Proteomics 15:3566–3579.

Conesa, A., S. Gotz, J. M. Garcia-Gomez, J. Terol, M. Talon, and M. Robles. 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21: 3674–3676.

Contamin, R., and A. M. Ellison. 2009. Indicators of regime shifts in ecological systems: What do we need to know and when do we need to know it? Ecological Applications 19:799–816.

Correll, D. L. 1998. The role of phosphorus in the eutrophication of receiving waters: a review. Journal of Environmental Quality 27:261–266.

Dakos, V., S. R. Carpenter, E. H. van Nes, and M. Scheffer. 2015. Resilience indicators: prospects and limitations for early warnings of regime shifts. Philosophical Transactions of the Royal Society of London B: Biological Sciences 370:20130263.

Dennison, W. C., R. J. Orth, K. A. Moore, J. C. Stevenson, V. Carter, S. Kollar, P. W. Bergstrom, and R. A. Batiuk. 1993. Assessing water-quality with submersed aquatic vegetation. BioScience 43:86–94.

Elias, J. E., and S. P. Gygi. 2007. Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. Nature Methods 4:207–214.

Ellison, A. M., N. J. Gotelli, J. S. Brewer, D. L. Cochran-Stafira, J. M. Kneitel, T. E. Miller, A. C. Worley, and R. Zamora. 2003. The evolutionary ecology of carnivorous plants. Advances in Ecological Research 33:1–74.

Farnsworth, E. J., and A. M. Ellison. 2008. Prey availability directly affects physiology, growth, nutrient allocation and scaling relationships among leaf traits in 10 carnivorous plant species. Journal of Ecology 96:213–221.

Gray, S. M., D. M. Akob, S. J. Green, and J. E. Kostka. 2012. The bacterial composition within the *Sarracenia purpurea* model system: local scale differences and the relationship with the other members of the food web. PLoS ONE 7:e50969.

Haller, L., M. Tonolla, J. Zopfi, R. Peduzzi, W. Wildi, and J. Pote. 2011. Composition of bacterial and archaeal communities in freshwater sediments with different contamination levels (Lake Geneva, Switzerland). Water Research 45:1213–1228.

Hargeby, A., I. Blindow, and G. Andersson. 2007. Long-term patterns of shifts between clear and turbid states in Lake Krankesjon and Lake Takern. Ecosystems 10:29–36.

Hurlbert, S. H. 1971. The nonconcept of species diversity: a critique and alternative parameters. Ecology 52:577–586.

Jayapal, K. P., R. J. Philp, Y. J. Kok, M. G. S. Yap, D. H. Sherman, T. J. Griffin, and W. S. Hu. 2008.

Uncovering genes with divergent mRNA-protein dynamics in *Streptomyces coelicolor*. PLoS ONE 3: e2097.

Kanehisa, M., S. Goto, Y. Sato, M. Kawashima, M. Furumichi, and M. Tanabe. 2014. Data, information, knowledge and principle: back to metabolism in KEGG. Nucleic Acids Research 42:D199–D205.

Kearns, P. J., J. H. Angell, E. M. Howard, L. A. Deegan, R. H. R. Stanley, and J. L. Bowen. 2016. Nutrient enrichment induces dormancy and decreases diversity of active bacteria in salt marsh sediments. Nature Communications 7:12881.

Kent, A. D., S. E. Jones, G. H. Lauster, J. M. Graham, R. J. Newton, and K. D. McMahon. 2006. Experimental manipulations of microbial food web interactions in a humic lake: shifting biological drivers of bacterial community structure. Environmental Microbiology 8:1448–1459.

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

Koopman, M. M., D. M. Fuselier, S. Hird, and B. C. Carstens. 2010. The carnivorous pale pitcher plant harbors diverse, distinct, and time-dependent bacterial communities. Applied and Environmental Microbiology 76:1851–1860.

Krieger, J. R., and P. S. Kourtev. 2012. Bacterial diversity in three distinct sub-habitats within the pitchers of the northern pitcher plant, *Sarracenia purpurea*. FEMS Microbiology Ecology 79:555–567.

Lau, H. T., J. Faryna, and E. W. Triplett. 2006. *Aquitalea magnusonii* gen. nov., sp nov., a novel gram-negative bacterium isolated from a humic lake. International Journal of Systematic and Evolutionary Microbiology 56:867–871.

Lauro, F. M., et al. 2011. An integrative study of a meromictic lake ecosystem in Antarctica. ISME Journal 5:879–895.

Lemmer, K. C., A. C. Dohnalkova, D. R. Noguera, and T. J. Donohue. 2015. Oxygen-dependent regulation of bacterial lipid production. Journal of Bacteriology 197:1649–1658.

Levin, P. S., and C. Mollmann. 2015. Marine ecosystem regime shifts: challenges and opportunities for ecosystem-based management. Philosophical Transactions of the Royal Society of London B: Biological Sciences 370:20130275.

Mesuere, B., T. Williams, F. Van der Jeugt, B. Devreese, P. Vandamme, and P. Dawyndt. 2016. Unipept web services for metaproteomics analysis. Bioinformatics 32:1746–1748.

Meyer, F., et al. 2008. The metagenomics RAST server: a public resource for the automatic phylogenetic and functional analysis of metagenomes. BMC Bioinformatics 9:386.

Newton, R. J., S. E. Jones, A. Eiler, K. D. McMahon, and S. Bertilsson. 2011. A guide to the natural history of freshwater lake bacteria. Microbiology and Molecular Biology Reviews 75:14–49.

Oksanen, J., et al. 2016. Vegan: community ecology package. R Package Version 2.4-1. https://cran.r-project.org/package=vegan

Ondov, B. D., N. H. Bergman, and A. M. Phillippy. 2011. Interactive metagenomic visualization in a Web browser. BMC Bioinformatics 12:385.

Paisie, T. K., T. E. Miller, and O. U. Mason. 2014. Effects of a ciliate protozoa predator on microbial communities in pitcher plant (Sarracenia purpurea) leaves. PLoS ONE 9:e113384.

Pan, Y. D., R. J. Stevenson, B. H. Hill, A. T. Herlihy, and G. B. Collins. 1996. Using diatoms as indicators of ecological conditions in lotic systems: a regional assessment. Journal of the North American Benthological Society 15:481–495.

Percent, S. F., M. E. Frischer, P. A. Vescio, E. B. Duffy, V. Milano, M. McLellan, B. M. Stevens, C. W. Boylen, and S. A. Nierzwicki-Bauer. 2008. Bacterial community structure of acid-impacted lakes: What controls diversity? Applied and Environmental Microbiology 74:1856–1868.

Perez, M. T., and R. Sommaruga. 2006. Differential effect of algal- and soil-derived dissolved organic matter on alpine lake bacterial community composition and activity. Limnology and Oceanography 51:2527–2537.

Peterson, C. N., S. Day, B. E. Wolfe, A. M. Ellison, R. Kolter, and A. Pringle. 2008. A keystone predator controls bacterial diversity in the pitcher-plant (*Sarracenia purpurea*) microecosystem. Environmental Microbiology 10:2257–2266.

Rabalais, N. N., R. E. Turner, R. J. Diaz, and D. Justic. 2009. Global change and eutrophication of coastal waters. Ices Journal of Marine Science 66:1528–1537.

Reddy, T. B. K., A. D. Thomas, D. Stamatis, J. Bertsch, M. Isbandi, J. Jansson, J. Mallajosyula, I. Pagani, E. A. Lobos, and N. C. Kyrpides. 2015. The Genomes OnLine Database (GOLD) v. 5: a metadata management system based on a four level (meta) genome project classification. Nucleic Acids Research 43:D1099–D1106.

Scheffer, M. 2009. Critical transitions in nature and society. Princeton University Press, Princeton, New Jersey, USA.

Scheffer, M., J. Bascompte, W. A. Brock, V. Brovkin, S. R. Carpenter, V. Dakos, H. Held, E. H. van Nes, M. Rietkerk, and G. Sugihara. 2009. Early-warning signals for critical transitions. Nature 461:53–59.

Shade, A., A. D. Kent, S. E. Jones, R. J. Newton, E. W. Triplett, and K. D. McMahon. 2007. Interannual dynamics and phenology of bacterial communities

in a eutrophic lake. Limnology and Oceanography 52:487–494.

Simcic, T. 2005. The role of plankton, zoobenthos, and sediment in organic matter degradation in oligotrophic and eutrophic mountain lakes. Hydrobiologia 532:69–79.

Sirota, J., B. Baiser, N. J. Gotelli, and A. M. Ellison. 2013. Organic-matter loading determines regime shifts and alternative states in an aquatic ecosystem. Proceedings of the National Academy of Sciences USA 110:7742–7747.

Sondergaard, M., L. S. Johansson, T. L. Lauridsen, T. B. Jorgensen, L. Liboriussen, and E. Jeppesen. 2010. Submerged macrophytes as indicators of the ecological quality of lakes. Freshwater Biology 55:893–908.

Sowell, S. M., P. E. Abraham, M. Shah, N. C. Verberkmoes, D. P. Smith, D. F. Barofsky, and S. J. Giovannoni. 2011. Environmental proteomics of microbial plankton in a highly productive coastal upwelling system. ISME Journal 5:856–865.

Srivastava, D. S., J. Kolasa, J. Bengtsson, A. Gonzalez, S. P. Lawler, T. E. Miller, P. Munguia, T. Romanuk, D. C. Schneider, and M. K. Trzcinski. 2004. Are natural microcosms useful model systems for ecology? Trends in Ecology & Evolution 19:379–384.

Tang, X. M., G. Gao, B. Q. Qin, L. P. Zhu, J. Y. Chao, J. J. Wang, and G. J. Yang. 2009. Characterization of bacterial communities associated with organic aggregates in a large, shallow, eutrophic freshwater lake (Lake Taihu, China). Microbial Ecology 58:307–322.

Timinskas, K., M. Balvociute, A. Timinskas, and C. Venclovas. 2014. Comprehensive analysis of DNA polymerase III alpha subunits and their homologs in bacterial genomes. Nucleic Acids Research 42:1393–1413.

Ullrich, N., P. Casper, A. Otto, and M. O. Gessner. 2016. Proteomic evidence of methanotrophy in methane-enriched hypolimnetic lake water. Limnology and Oceanography 61:S91–S100.

Van Horn, D. J., R. L. Sinsabaugh, C. D. Takacs-Vesbach, K. R. Mitchell, and C. N. Dahm. 2011. Response of heterotrophic stream biofilm communities to a gradient of resources. Aquatic Microbial Ecology 64:149–161.

Vogel, C., and E. M. Marcotte. 2012. Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. Nature Reviews Genetics 13:227–232.

Wayne, L. G., and K. Lin. 1982. Glyoxylate metabolism and adaptation of *Mycobacterium tuberculosis* to survival under aerobic conditions. Infection and Immunity 37:1042–1049.

Supporting Information

Additional Supporting Information may be found online at: http://onlinelibrary.wiley.com/doi/10.1002/ecs2.1954/full