bioRxiv preprint doi: https://doi.org/10.1101/2025.07.22.666210; this version posted July 26, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

Mono-mix strategy enables comparative proteomics of a cross-kingdom microbial symbiosis

Sunnyjoy Dupuis[1,2], Usha F. Lingappa[2], Samuel O. Purvine[3], Lauren Chiang[1], Sean D. Gallaher[2], Carrie D. Nicora[3], Mary S. Lipton[3], Sabeeha S. Merchant[1,2,4,5]*

[1]Department of Plant and Microbial Biology, University of California, Berkeley, CA 94720, USA

[2]California Institute for Quantitative Biosciences, University of California, Berkeley, CA 94720, USA

[3]Earth and Biological Sciences Division, Pacific Northwest National Laboratory, Richland, WA 99352, USA

[4]Environmental Genomics and Systems Biology Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

[5]Department of Molecular and Cell Biology, University of California, Berkeley, CA 94720, USA

*Corresponding author
Email: sabeeha@berkeley.edu (SSM)

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.22.666210; this version posted July 26, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

Abstract

Cross-kingdom microbial symbioses, such as those between algae and bacteria, are key players in biogeochemical cycles. The molecular changes during initiation and establishment of symbiosis are of great interest, but quantitatively monitoring such changes can be challenging, particularly when the microorganisms differ greatly in size or are intimately associated. Here, we analyze output from data-dependent acquisition (DDA) LC-MS/MS proteomics experiments investigating the well-studied interaction between the alga *Chlamydomonas reinhardtii* and the heterotrophic bacterium *Mesorhizobium japonicum*. We found that detection of bacterial proteins decreased in coculture by 50% proteome-wide due to the abundance of algal proteins. As a result, standard differential expression analysis led to numerous false-positive reports of significantly downregulated proteins, where it was not possible to distinguish meaningful biological responses to symbiosis from artifacts of the reduced protein detection in coculture relative to monoculture. We show that data normalization alone does not eliminate the impact of altered detection on differential expression analysis of the cross-kingdom symbiosis. We assessed two additional strategies to overcome this methodological artifact inherent to DDA proteomics. In the first, we combined algal and bacterial monocultures at a relative abundance that mimicked the coculture, creating a "mono-mix" control to which the coculture could be compared. This approach enabled comparable detection of bacterial proteins in the coculture and the monoculture control. In the second strategy, we enhanced detection of lowly abundant bacterial proteins by using sample fractionation upstream of LC-MS/MS analysis. When these simple approaches were combined, they allowed for meaningful comparisons of nearly 10,000 algal proteins and over 4,000 bacterial proteins in response to symbiosis by DDA. They successfully recovered expected changes in the bacterial proteome in response to algal coculture, including upregulation of sugar-binding proteins and transporters. They also revealed novel proteomic responses to coculture that guide hypotheses about algal-bacterial interactions.

Keywords

Dual-proteomics; shotgun proteomics; label-free quantitation; mixed cultures; chlorophyte; rhizobia

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.22.666210; this version posted July 26, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

Introduction

Cross-kingdom microbial symbioses abound in nature. Interactions between bacteria, archaea, fungi, and protists provide important ecosystem services, shape evolution, and drive biogeochemical cycling. For instance, syntrophic bacteria and archaea of the rumen degrade recalcitrant plant biomass in the ruminant digestive tract and constitute a major source of methane emissions globally [1,2]. Mycorrhizal fungi and symbiotic bacteria in their "hyphosphere" provide nitrogen, phosphorus, and other minerals to plants in exchange for carbon, increasing plant productivity [3–5]. Lichenized photobionts release glucose to symbiotic fungi and are protected by fungal structures and metabolites [6]. Phytoplankton are responsible for much of the Earth's primary productivity, serving as a major source of organic carbon for marine food webs and providing a habitat for heterotrophic bacteria in the "phycosphere" [7,8]. In each instance, tight association of the symbionts, their divergent structural characteristics, and large differences in their abundances and metabolic rates can pose challenges to studying the microorganisms' molecular responses to one another.

Cross-kingdom algal-bacterial interactions are an important component of global carbon cycling [9–11]. Among the most well studied algae is the unicellular green alga *Chlamydomonas reinhardtii*. This alga is experimentally tractable, can perform both photosynthesis and heterotrophy, and exhibits rhythms in growth and metabolism over diurnal cycles [12,13]. In recent decades, studies have begun to elucidate how *C. reinhardtii* interacts with various bacteria [14–23]. For example, the interaction between the alga and the vitamin B_{12}-producing α-proteobacterium *Mesorhizobium japonicum* has been well documented [15,22,24–26]. This bacterium receives a small but significant amount of photosynthetically fixed carbon in coculture, primarily via algal cell lysis [26]. Yet, it is unknown what specific alga-derived substrates *M. japonicum* and other bacteria may receive from *C. reinhardtii*, and how such partners sense and respond to one another.

Quantitative investigations of algal-bacterial interactions can be challenging due to technical barriers for studying eukaryotic and prokaryotic microbes simultaneously without disrupting the interaction, particularly when the partners differ in abundance and size by orders of magnitude [27]. Algal cells are typically much larger than prokaryotic cells. For example, *C. reinhardtii* cell volume ranges from roughly 70–3000 μm^3, whereas *M. japonicum*'s cell volume is around 1–5 μm^3 [26]. Although this bacterium can reach up to 10^9 cells/ml in coculture, 100-fold higher than the maximum algal cell density, *M. japonicum* makes up only 10% of the total biomass in mixed cultures, at best. Accurate quantification of bacterial cell density itself is challenging in algal cocultures: optical density cannot be used, as algal pigments influence absorbance, so cells must be counted or estimated from colony forming units (CFU). In addition, lowly abundant bacterial molecules (e.g., proteins) are difficult to detect quantitatively amidst an abundance of algal molecules.

Liquid chromatography and tandem mass spectrometry (LC-MS/MS) proteomics is a useful tool for studying microbial interactions [28–31]. The most widely used methodology is label-free, data-dependent acquisition (DDA). In this "shotgun" approach, peptide precursor ions are chosen stochastically for fragmentation from an MS scan. Precursor ion selection is influenced by the complexity of the analyte, and ions are more likely to be chosen by the mass spectrometer if they are highly abundant at the expense

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.22.666210; this version posted July 26, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

of less abundant ions [32,33]. Therefore, although eukaryotic and prokaryotic peptides present in a sample are generally both detectable by MS with respect to their chemical properties, lowly abundant bacterial peptides may be missed in coculture samples that are dominated by algal peptides. Systematic differences in detection related to differences in analyte complexity complicate differential expression analysis, particularly between coculture and monoculture samples [34–36]. Such comparisons are of great interest, as they have the potential to reveal proteomic responses to initiation and establishment of symbiosis.

Several studies have reported attempts to capture proteomic changes during algal-bacterial interactions using DDA approaches. Proteomic analysis of cocultures of the alga *Lobomonas rostrata* with *M. japonicum* using the iTRAQ labelling approach coupled with DDA yielded identification of only 4 bacterial proteins (in addition to 588 algal proteins) [37]. In a study of the interaction between *C. reinhardtii* and the growth promoting bacterium *Arthrobacter* strain P2b, many fewer bacterial proteins were detected in samples collected from coculture than from bacterial monocultures during label-free DDA proteomic analysis [21]. In another label-free DDA study investigating symbiosis between *C. reinhardtii* and the algicidal bacterium *Paenibacillus polymyxa*, hundreds of bacterial proteins were identified, many of which were significantly altered in abundance in the presence of the alga, but successful detection may have only been possible because of the minimal algal biomass resulting from algicidal activity [38].

Here, we quantitatively assess label-free DDA proteomics of *C. reinhardtii* and *M. japonicum* from mixed cultures. We show that the presence of algal proteins leads to a 50% reduction in bacterial protein detection in a typical shotgun proteomics experiment of cocultures and monocultures. These global differences in detection resulted in false positives during typical differential expression analyses, complicating identification of meaningful changes to the bacterial proteome in response to symbiosis. We explored three accessible strategies to improve DDA proteomics analysis of lowly abundant bacterial proteins from mixed *C. reinhardtii-M. japonicum* cultures: data normalization, the use of a “mono-mix” control, and sample fractionation. Together, these accessible strategies allowed us to confidently determine significant changes in bacterial protein abundance in coculture when it is relying on an alga for its growth.

Materials and Methods

Strains and culture conditions

The *C. reinhardtii* strain closely related to strain S24-, which is often studied with *M. japonicum*, was used in the first experiment (Sample Set 1) under continuous light [26]. The *C. reinhardtii* cell-wall reduced strain CC-5390 was used in the experiments comparing continuous and diurnal light (Sample Set 2). This strain is amenable to synchronization under diurnal cycles, providing a system with exceptional signal-to-noise for resolving diurnal patterns [39]. *M. japonicum* strain MAFF303099 was used in all experiments.

Cultures were grown in amended High Salt minimal medium (HSM) [40] with a modified trace element solution [41], 2 mM $MgSO_4 \cdot 7H_2O$, 10 µg/l biotin, and 10 µM $CoCl_2$ unless otherwise specified. Cultures were grown in 250 ml Erlenmeyer flasks containing

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.22.666210; this version posted July 26, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

100 ml medium and bubbled with filter-sterilized air (provided by an aquarium pump). Algal cultures and cocultures were incubated at 28 °C, agitated at 110–130 rpm, and illuminated with 200 µmol photons/m^2/s cool white light (Sylvania Dulux L55WFT55DL/841) continuously or under diurnal 12-h-light/12-h-dark cycles.

C. reinhardtii cells were precultivated axenically in photoautotrophic conditions. For the experiment comparing continuous and diurnal light, algal precultivation occurred in flat-panel photobioreactors (Photon System Instruments, Drásov, Czechia) as previously described under the respective light regime to achieve adequate synchrony of the diurnal populations [39]. Bacterial cells for all experiments were precultivated in 14 ml polystyrene round-bottom tubes with 3 ml amended HSM containing 0.2% sucrose, agitated at 200 rpm, and incubated at 28 °C. Bacterial cells were collected by centrifugation at 8000 x*g* and washed three times in 0.85% NaCl to remove sucrose before use as inoculum. Optical density at 600 nm (OD_{600}) was used to estimate the cell density of washed cell suspensions used for inoculation. Biological replicates refer to replicate flasks inoculated and incubated together in an experiment.

Estimation of cell density

C. reinhardtii cell density was determined using a Z2 Coulter Particle Count and Size Analyzer (Beckman Coulter, CA, USA). *M. japonicum* cell density was estimated as CFU determined from serial dilutions of culture samples on Tryptone Yeast (TY) agar plates after 4 d of incubation at 30 °C in the dark.

In silico assessment of protein physicochemical characteristics

The physicochemical properties of proteins encoded by *C. reinhardtii* and *M. japonicum* were assessed according to Shi *et al.* 2024 [36]. The *C. reinhardtii* reference genome assembly and annotations v6.1 protein FASTA file from was downloaded from https://phytozome-next.jgi.doe.gov [42]. The *M. japonicum* MAFF303099 RefSeq protein FASTA file was downloaded from https://www.ncbi.nlm.nih.gov/data-hub/taxonomy/266835/. FASTA files were read into R using the “seqinr” package (v4.2-36). The isoelectric point (pI) of each protein was estimated using the pI function in the R package “Peptides” (v2.4.6) using pKscale = “EMBOSS”. The grand average of hydropathy (GRAVY) hydrophobicity score of each protein was estimated using the hydrophobicity function from the R package “Peptides” (v2.4.6) using scale = “KyteDoolittle”. The molecular weight of each protein was determined using the mw function from the R package “Peptides” (v2.4.6) using monoisotopic = FALSE, avgScale = "expasy", label = "none", and aaShift = NULL.

Proteomic sample collection and digestion

40 ml of culture was collected either 24, 36, or 48 h after inoculation. Cell pellets were collected by centrifugation at 8000 x*g* for 2 min at 4 °C and were resuspended in 10 mM Na-PO_4 buffer (pH 7.0). Cell pellets were again collected by centrifugation at 8000 x*g* for 1 min at 4 °C, snap frozen in liquid N_2, and stored at –80 °C until further processing.

Cell pellets were diluted in 300 µl of 8 M urea and cells were lysed by bead beating in Micro-Organism Lysing Mix glass bead tubes for 45 s at 5.5 m/s using a Bead Ruptor

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.22.666210; this version posted July 26, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

Elite bead mill homogenizer (OMNI International, Kennesaw, GA). Lysate was collected by centrifugation at 1,000 x*g* for 10 min at 4 °C. Protein concentration was determined by bicinchoninic acid (BCA) assay (Thermo Scientific, Waltham, MA USA). Then, 10 mM dithiothreitol (DTT) was added and samples were incubated at 60 °C for 30 min with constant shaking at 800 rpm. 40 mM iodoacetamide (IAA) was then added and samples were incubated at ambient temperature in the dark for 30 min.

Protein samples were then digested with sequencing grade trypsin (United States Biological, Salem, MA) at a 1:50 (w/w) trypsin-to-protein ratio for 3 h at 37 °C in the presence of 100 mM NH_4HCO_3 and 1 mM $CaCl_2$. Digested samples were desalted using a 4-probe positive pressure Gilson GX-274 ASPEC™ system (Gilson Inc., Middleton, WI) with Discovery C18 50 mg/1 ml solid phase extraction tubes (Supelco, St. Louis, MO), using the following protocol: 3 ml of methanol was added for conditioning followed by 3 ml of 0.1% trifluoroacetic acid (TFA) in H_2O. The samples were then loaded onto the column followed by 4 ml of 95:5 H_2O:acetonitrile (ACN), 0.1% TFA. Samples were eluted with 1 ml 80:20 ACN:H_2O, 0.1% TFA. The solvent was completely removed in a vacuum concentrator (Labconco, Kansas City, MO) and the peptides subsequently reconstituted in 100 µl pure water. Peptide concentration of the final mixture was determined using a BCA assay, and samples were diluted to 0.10 µg/µl with pure water for global proteomics LC-MS/MS analysis.

Reversed phase fractionation of peptides

Peptide mixtures were separated into 6, 12, or 24 fractions by high-resolution reversed phase UPLC using a nanoACQUITY UPLC® system (Waters Corporation, Milford, MA) equipped with an autosampler. Capillary columns, 200 µm i.d. x 65 cm long, were packed with Jupiter 3.0 µm C18 300 Å bonded particles (Phenomenex, Torrance, CA). Separations were performed at a flow rate of 2.2 µl/min on binary pump systems, using 10 mM ammonium formate (pH 7.5) as mobile phase A and 100% ACN as mobile phase B. 48 µl of the peptide mixtures (0.25 µg/µl) were loaded onto the column and separated using a binary gradient of 1% B for 35 min, 1–10% B in 2 min, 10–15% B in 5 min, and 15–25% B in 35 min, 25–32% B in 25min, 32–40% B in 13 min, 40–90% B in 43 min, held at 90% B for 2 min (column washed and equilibrated from 90–50% B in 2 min, 50–95% B in 2 min, held at 95% B for 2 min and 95–1% in 4 min). The capillary column eluent was automatically deposited every minute into 12 mm x 32 mm polypropylene vials (Waters Corporation, Milford, MA) starting at minute 60 and ending at minute 170, over the course of the 180 min LC run. Fractions were concatenated into 6, 12, or 24 vials by collecting fractions from vial 1 to vial 6–24 and then returning to vial 1 (and so on). Prior to peptide fraction collection, 100 µl of water was added to each vial to facilitate the recovery of small droplets added into the vial. Each vial was completely dried in a vacuum concentrator (Labconco, Kansas City, MO) and reconstituted in 20 µl 2% ACN, 0.1% formic acid.

LC-MS/MS proteomics acquisition

A nanoACQUITY dual pumping UPLC system (Waters Corporation, Milford, MA) was configured for on-line trapping of a 5 µl injection at 5 µl/min for 10 min followed by reverse-flow elution onto the analytical column at 200 nl/min. The trapping column was slurry packed in-house using 360 µm o.d. x 150 µm i.d. fused silica (Polymicro

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.22.666210; this version posted July 26, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

Technologies Inc., Phoenix, AZ) Jupiter 5µm C18 media (Phenomenex, Torrence, CA) with 2mm sol-gel frits on either end. The analytical column was slurry packed in-house using Waters BEH 1.7 µm particles packed into a 35cm long, 360 µm o.d. x 75 µm i.d. column with an integrated emitter (New Objective, Inc., Littleton, MA). Mobile phases consisted of (A) 0.1% formic acid in water and (B) 0.1% formic acid in ACN with the following gradient profile (min, %B): 0, 1; 10, 8; 105, 25; 115, 35; 120, 75; 123, 95; 129, 95; 130, 50; 132, 95; 138, 95; 140, 1.

MS analysis was performed using a Thermo Eclipse mass spectrometer (Thermo Scientific, San Jose, CA). The ion transfer tube temperature and spray voltage were 300 ºC and 2.4 kV, respectively. Data were collected for 120 min following a 27 min delay from when the trapping column was switched in-line with the analytical column. FAIMS was used at CVs –40 V, –60 V, and –80 V. FT-MS spectra were acquired from 300–1800 m/z at a resolution of 120 k (AGC target 4e5) and while the top 12 FT-HCD-MS/MS spectra were acquired in data-dependent mode with an isolation window of 0.7 m/z and at an orbitrap resolution of 50 k (AGC target 5e4) using a fixed collision energy (HCD) of 32 and a 30 s exclusion time.

LC-MS/MS proteomics data processing

All LC-MS/MS datafiles were converted to mzML using MSConvert with all peaks kept during centroiding [43]. MZRefinery was used to calibrate search mass values from the MS2 data, which were then input to MS-GF+ [44]. Data were searched against a single concatenated FASTA file containing amino acid sequences from the *C. reinhardtii* reference genome assembly and annotations v6.1 (https://phytozome-next.jgi.doe.gov) [42], the *M. japonicum* MAFF303099 genome assembly and annotations (https://www.ncbi.nlm.nih.gov/data-hub/taxonomy/266835/), and commonly observed contaminant proteins. We employed a target/decoy approach with +/–20 ppm parent mass tolerance, dynamic modification of oxidized methionine, and allowance of partially tryptic peptide candidates [45]. MS abundances were extracted using StatsMomentsArea via the MASIC software to provide area-under-the-elution-curve values (i.e., extracted ion chromatograms) [46]. Data were collated using the MAGE software suite (https://github.com/PNNL-Comp-Mass-Spec/Mage), imported into SQL Server, and cross-checked for completeness. FDR was controlled to 1% (q-value ≤ 0.010297 gave 16,354 decoy identifications out of 1,635,330 total). Peptide-to-protein relationships were subjected to parsimony such that a protein was assigned if a) the peptide was uniquely mapped to that protein, b) was non-uniquely mapped to only one protein with a different uniquely mapping peptide, c) was non-uniquely mapped to a protein more other non-uniquely mapping peptides than any others, or d) was mapped to the first occurrence of the peptide in the search FASTA file when more than one protein had similar maximal peptide counts.

Data from samples that had been fractionated were then processed separately from data from samples that had not been fractionated. Peptide abundances were $\log_2$ transformed, mean central tendency normalized, then grouped by their gene ID and summed (simple rollup) per sample, resulting in protein abundances as $\log_2$-transformed MASIC values. Only proteins with at least two unique peptides for quantitation that were detected in at least two of the three biological replicates in any condition were analyzed. Log_2-transformed MASIC values were quantile normalized using the normalize.quantiles

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.22.666210; this version posted July 26, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

function from the R package *preprocessCore* (v1.66.0). Quantile normalization was performed independently for the distribution of *C. reinhardtii* proteins and the distribution of *M. japonicum* proteins across the relevant samples. In the unfractionated pilot experiment (Sample Set 1), this included 9871 proteins (6113 *C. reinhardtii* proteins, 3758 *M. japonicum* proteins). Normalized and unnormalized abundances for these proteins are provided in S2 Table. In the unfractionated mono-mix experiment (Sample Set 2), this included 9784 proteins (7039 *C. reinhardtii* proteins, 2745 *M. japonicum* proteins). Normalized and unnormalized abundances for these proteins are provided in S3 Table. In the mono-mix experiment comparing protein detection before and after peptide sample fractionation (Sample Set 2, fractionated), this included 15,464 proteins (10,858 *C. reinhardtii* proteins, 4606 *M. japonicum* proteins). Normalized and unnormalized abundances for these proteins are provided in S5 Table.

The average protein abundances were calculated after excluding non-detected values and the $\log_2$-fold change in protein abundance was calculated as (mean abundance in coculture) – (mean abundance in monoculture). To calculate the Z-score, non-detected values were excluded, then the mean and standard deviation of quantile-normalized MASIC values across the six conditions was calculated, and the Z-score was calculated as (mean norm. MASIC – (mean of all mean norm. MASIC))/(sd of all mean norm. MASIC).

Clustering analyses

k-means clustering was performed on the Z-score of quantile-normalized mean protein abundances in the unfractionated mono-mix experiment (Sample Set 2), separately for *C. reinhardtii* proteins and *M. japonicum* proteins that were detected in at least two biological replicates in all conditions (4737 *C. reinhardtii* proteins, 828 *M. japonicum* proteins). The fviz_nbclust function from the R package *factoextract* (v1.0.7) was used to generate elbow plots to determine the optimal *k* for clustering. The kmeans function from the R package *stats* (v4.4.0) was then applied using iter.max = 1000, centers = 6 for *C. reinhardtii* proteins or centers = 3 for *M. japonicum* proteins. Cluster assignment is provided in S3B Table. Heatmaps were generated using the pheatmap function from the R package *pheatmap* (v1.0.12).

Principal component analysis (PCA) was performed on quantile-normalized protein abundances of individual replicate samples in the unfractionated mono-mix experiment. This was done separately for *C. reinhardtii* proteins and *M. japonicum* proteins that were detected above the limit of quantitation (two standard deviations below the mean MASIC value in the experiment) in all samples (all replicates in all conditions: 3817 *C. reinhardtii* proteins, 569 *M. japonicum* proteins). The pca function was applied from the R package *PCAtools* (v2.16.0) and the results were visualized using the biplot function.

Statistical tests

Significant differences in culture densities and the number of proteins detected in different sample groups were assessed by two-tailed Student's t-tests. Significant differences in the distribution of protein abundances were assessed by Wilcoxon rank-sum tests (Mann-Whitney tests) using the wilcox.test function in the R package *stats* (v4.4.1).

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.22.666210; this version posted July 26, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

Significant differences in the abundance of a given protein in coculture relative to monoculture were assessed using the t.test function in the R package *stats* (v4.4.0). Only proteins that were detected in at least two of the three biological replicates in both the monoculture control condition and the respective coculture condition were analyzed. In the unfractionated pilot experiment (Sample Set 1), this included 5368 *C. reinhardtii* proteins and 1737 *M. japonicum* proteins. In the unfractionated mono-mix experiment (Sample Set 2), it included 6246 *C. reinhardtii* proteins and 2128 *M. japonicum* proteins. In the fractionated mono-mix experiment (Sample Set 2, fractionated), it included 9934 *C. reinhardtii* proteins and 4179 *M. japonicum* proteins). *p* values were adjusted for the number of t-tests conducted per species in each experiment using a Benjamini-Hochberg correction by applying the p.adjust function with method = “fdr”. Significant differences in protein abundance were defined as those where $|\log_2\text{-fold change}| > 1$, *p-adj.* < 0.05, and the mean $\log_2$-transformed MASIC value was greater than the limit of quantitation in both the coculture and the monoculture control.

Gene ontology (GO) enrichment analysis

GO term assignments for the *C. reinhardtii* v6.1 proteome were downloaded from Phytozome (https://phytozome-next.jgi.doe.gov/). GO terms were assigned *de novo* to the *M. japonicum* MAFF303099 proteome (available from the NCBI with accession GCF_000009625.1) using InterProScan (v5.59-91.0) with --iprlookup –goterms [47]. Then, the lists of proteins with significant changes in their abundance in coculture relative to monoculture were tested for GO term enrichment using the enricher function in the R package *clusterProfiler* (v3.0.4). *p* values were corrected using a Bonferroni-Holm adjustment, and enriched GO terms where *p-adj.* < 0.05 were considered significant.

Accession numbers

Raw proteomics data have been deposited at http://massive-ftp.ucsd.edu/ under the accession MSV000098360 and are available as of the date of publication.

Results

Bacterial protein detection by DDA LC-MS/MS is impaired in the presence of an algal partner

To determine whether label-free DDA LC-MS/MS could be used to monitor changes to a bacterium’s proteome during coculture with an alga, we grew the model alga *C. reinhardtii* and the bacterium *M. japonicum* under four conditions and performed proteomics: algal monoculture, bacterial monoculture supplemented with sucrose, coculture, and coculture supplemented with sucrose (S1 and S2 Tables). We found that the presence of *C. reinhardtii* reduced detection of bacterial proteins substantially (Fig 1). In the absence of the alga, nearly 4000 unique bacterial proteins were detected, but in coculture, the number of unique bacterial proteins detected decreased as the relative abundance of the alga increased, such that <500 bacterial proteins could be detected in the coculture in the absence of sucrose (Figs 1A and 1B). For bacterial proteins detected in all three sample types, the measured abundance was also significantly decreased proteome-wide in coculture samples relative to monoculture samples (Fig 1C, Wilcoxon

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.22.666210; this version posted July 26, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

rank-sum tests, $p < 0.05$). Whereas the histogram of algal protein abundances in the monoculture samples overlapped with the histograms of algal protein abundances in coculture samples (muddied color from overlapping purple, turquoise, and orange histograms), the histograms of bacterial protein abundances were shifted to lower abundances as the relative abundance of the alga increased (Fig 1C, Wilcoxon rank-sum tests, $p < 0.05$). These results suggest a systematic decrease in the selection of bacterial peptide precursor ions in the presence of relatively abundant algal peptides in coculture samples.

Similar observations were apparent upon reanalysis of proteomics data from a published study on the interaction between *C. reinhardtii* and the actinomycete *Arthrobacter* strain P2b [21]. Although bacterial density was not impacted by the presence of the alga, the number of bacterial proteins detected by label-free DDA LC-MS/MS decreased in coculture relative to monoculture (S1A Fig), and the measured abundance of bacterial proteins that were detected in both conditions was decreased globally in samples containing *C. reinhardtii* proteins (S1B Fig, Wilcoxon rank-sum tests, $p < 0.05$) [21]. We wondered if decreased detection of bacterial proteins in coculture samples was unique to those containing eukaryotic partners with larger cells, or rather generalizable to any sample with another microbe. Therefore, we also reanalyzed label-free LC-MS/MS proteomics data from syntrophic bacterial cultures of *Dehalococcoides ethenogenes* strain 195 and *Desulfovibrio vulgaris* Hildenborough [48]. In this case, the two organisms are closer in size, coculture increases growth of *D. ethenogenes*, and *D. ethenogenes* cells make up roughly 80% of the cocultures. Nevertheless, detection of *D. ethenogenes* proteins was markedly decreased in cocultures relative to monocultures (S1C and S1D Figs). Furthermore, a similar observation was reported from proteomic analysis of the archaeon *Nanoarchaeum equitans* in the presence and absence of another archaeon, the epibiont *Ignicoccus hospitalis* [34]. These data on diverse symbiotic interactions suggest that increased complexity of coculture samples decreases microbial protein detection regardless of the identity or relative abundance of species present.

We sought to assess how differences in detection influence differential expression analysis of algal and bacterial proteins. We used standard criteria for defining significant changes in protein abundance: $|\log_2\text{-fold change}| > 1$ and *p-adj.* < 0.05 from a Student's t-test of the protein's abundance (MASIC values) in replicate coculture and monoculture samples. We found that no *C. reinhardtii* proteins showed significant changes in abundance in coculture relative to monoculture (S2A Fig). This was not surprising, as we previously showed that wild-type *C. reinhardtii* growth and physiology are not significantly impacted by *M. japonicum* [26]. When we compared the abundance of the 1737 *M. japonicum* proteins detected in bacterial monocultures and cocultures with sucrose, we found that 1350 proteins (78%) were significantly decreased in abundance in the coculture samples, whereas only 12 (0.7%) were increased (S2B Fig). Downregulation of such a large proportion of the bacterium's proteome was not expected, particularly since *M. japonicum's* preferred carbon source, sucrose, was provided in both conditions. Given the systematic depression of bacterial protein abundance observed in Fig 1C, it was not possible to distinguish whether these differences in protein abundance were biological or methodological.

Data normalization has been shown to be important for differential expression analysis using label-free DDA proteomics [49], and it has been employed in order to

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.22.666210; this version posted July 26, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

perform differential expression analysis on data collected from microbial cocultures and monocultures [34,36,50]. Therefore, we tested whether data normalization could allow for appropriate comparisons of the abundance of bacterial proteins in monoculture and algal coculture. Since the distributions of algal protein abundances were markedly different from the distributions of bacterial protein abundances (Fig 1C), proteomics data from each organism were normalized independently. The abundances of each organism's proteins in each replicate sample were quantile normalized relative to the abundances of that organism's proteins in all relevant samples. For example, the abundances of *M. japonicum* proteins in Replicate A of the monoculture condition were quantile normalized together with the abundances of *M. japonicum* proteins in the other monoculture replicate samples, in the replicate samples from cocultures, and in the replicate samples from cocultures with sucrose, but not with the abundances of proteins in the *C. reinhardtii* monoculture samples. We found that this quantile normalization strategy led to greater overlap in the distributions of protein abundances for both species (Fig 1D, a greater proportion of the histograms are overlapping and their color is muddied; S2B Table). However, significant differences in the distributions remained (Fig 1D, Wilcoxon rank-sum tests, $p < 0.05$). Furthermore, differential expression analysis using the quantile-normalized values still suggested that most bacterial proteins were significantly decreased in abundance in coculture relative to monoculture in the presence of sucrose (Fig 1E). In this case, 949 bacterial proteins (55%) decreased in abundance in coculture samples, whereas only 28 bacterial proteins (1.6%) increased. Thus, comparison of protein abundance determined by shotgun LC-MS/MS analysis of monoculture and algal coculture samples is not an appropriate way to discern symbiosis-induced changes in the bacterial proteome, even after data normalization.

The "mono-mix" strategy allows comparable detection of bacterial proteins from a coculture and a monoculture control by DDA proteomics

In order to enable differential expression analysis of the bacterial proteome in cocultures relative to monocultures from label-free DDA, we sought to develop a more appropriate monoculture control that could emulate the analyte complexity of the coculture condition. In a previous study investigating the interaction between the archaeon *Methanococcus maripaludis* and the bacterium *D. vulgaris*, protein abundances in coculture were compared to abundances in a "synthetic blend," where equal amounts of total protein from each organism had been mixed together prior to analysis [51]. Then, the fold change in each protein's abundance in cocultures relative to the synthetic blend was multiplied by a constant correction factor according to the estimated relative abundance of the organisms in coculture. This correction factor approach assumes that differences in organism relative abundance would impact detection equally for all peptides regardless of their physicochemical properties or their abundances. Proteins encoded by *C. reinhardtii* and *M. japonicum* exhibit a wide range of isoelectric points, hydropathies, and molecular weights, which can influence their detection (S3 Fig).

We conceived of a modified experimental design that we term the "mono-mix" strategy, in which DDA peptide detection in control samples may more closely emulate detection in coculture samples. We grew *C. reinhardtii* monocultures and *M. japonicum* monocultures supplemented with sucrose in parallel to cocultures in minimal medium.

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.22.666210; this version posted July 26, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

Then, immediately prior to collecting cell pellets for proteomics, we combined the two monocultures to approximate the bacteria-to-algae ratio expected in the cocultures (Fig 2A). We call this combined monoculture the "mono-mix" control. We applied this strategy to cultures grown under continuous light and diurnal light. Like other algae, *C. reinhardtii* coordinates its metabolism and gene expression with the time of day under diurnal light. The alga performs photosynthesis, cell growth, and starch synthesis during the day period, initiates replication at dusk, then degrades starch and induces fermentative metabolism during the night in the G0 phase of the cell cycle [52,39]. To assess whether these diurnal rhythms influence *M. japonicum* in coculture, we sampled cultures grown in diurnal light at both the end of the night (36 h after inoculation) and the end of the day (48 h after inoculation), and we compared them to cultures grown in continuous light sampled 36 h after inoculation (Fig 2A, S1 and S3 Tables).

Preparing a mono-mix control with the same bacteria-to-algae ratio as the parallel coculture was challenging. We determine the number of bacterial cells in a culture as CFU on TY agar plates, which is assessed only after 4 d of incubation. Thus, only a rough estimate of the bacterial cell density is available at the time samples are collected for proteomics. Therefore, there was variation in how comparable the cell densities were in the mono-mix controls and their parallel cocultures (Fig 2B). We found that although the number of bacterial proteins detected in the cocultures (roughly 800–2200) was more comparable to that of the mono-mix controls (roughly 1600–2300), it was still significantly lower in the samples collected from the continuous light cultures and the diurnal light cultures at the end of the day (Fig 2C, Student's t-test, $p < 0.05$). The mono-mix control was most comparable to the coculture in the diurnal night samples, both in terms of organism relative abundance and of the number of bacterial proteins observed (Figs 2B and 2C). In this condition, there was no significant difference in the total number of bacterial proteins detected (Fig 2C, Student's t-test, $p = 0.48$) nor in the distribution of bacterial protein abundances in the coculture relative to the mono-mix control (Fig 2E, Wilcoxon rank-sum test, $p = 0.30$). This was true whether data were quantile normalized or not (Fig 2E and S4B Fig, respectively).

Data normalization and comparison to mono-mix controls reveal that light regime may influence the *M. japonicum* proteome during coculture with *C. reinhardtii*

Having achieved more comparable detection of the bacterial proteome in cocultures and mono-mix controls, we next sought to determine whether these data could reveal the impact of light regime and time of day on the organisms during their interaction. We performed *k*-means clustering and principal component analysis (PCA) on the quantile-normalized abundances of proteins that were detected in all conditions (the cocultures and mono-mix controls in continuous light, diurnal night, and diurnal day). The *C. reinhardtii* proteome could be separated into 6 clusters that exhibited distinct patterns with respect to light regime and the presence of the bacterium (Fig 3A, S3B Table). For example, Cluster Cr1 proteins were more abundant in coculture under all light regimes, and Cluster Cr5 proteins were more abundant in samples collected during the light phase (either continuous or diurnal) regardless of the presence of the bacterium during growth.

On the other hand, the *M. japonicum* proteome could only be separated into 2-to-3 distinct clusters (Fig 3B, S3B Table). There was only a subtle difference in the pattern

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.22.666210; this version posted July 26, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

of expression between bacterial proteins in Clusters Mj1 and Mj2 across the different light and culture conditions. Most bacterial proteins appeared to have lower abundance in the continuous light coculture samples (for which the distribution of protein abundances was significantly lower, Fig 2E, Wilcoxon rank-sum test, $p < 2x10^{-16}$) and high abundance in the continuous light mono-mix controls and the diurnal light cultures collected at the end of the night. Bacterial proteins in Cluster Mj3 appeared higher in cocultures than in monocultures, particularly in the diurnal night samples. These results suggest that complexity-related differences in detection of bacterial peptides in the various samples still influence the data.

PCA revealed that 27% of the variation in *C. reinhardtii's* proteome could be explained by the presence or absence of light, and 13% of the variation was driven by the interaction with the bacterium (Fig 3C). PCA of *M. japonicum's* proteome showed that the continuous light cocultures (white triangles) were most distinct from their respective mono-mix controls (white circles) (Fig 3D). This was also evident in the *k*-means clustered heatmap, where the majority of bacterial proteins detected appeared less abundant in the cocultures grown in continuous light than in the mono-mix controls grown in continuous light (Fig 3B). As discussed, bacterial protein detection was significantly lower in the continuous light cocultures relative to their mono-mix controls (Figs 2C and 2E). In addition, differential expression analysis suggested that in continuous light, many more proteins were significantly decreased in abundance in coculture than were increased (S5 Fig, S4 Table). Therefore, the patterns observed by clustering analyses likely reflect differences in organism relative abundance and consequent differences in peptide detection in the coculture and mono-mix samples collected from continuous light.

Interestingly, however, PCA of *M. japonicum's* proteome also revealed that while the mono-mix controls (circles) clustered relatively close to one another on both the first and second principal components, the cocultures (triangles) were quite separated from one another by the first two principal components (Fig 3D). This suggested that light regime had a greater influence on the bacterial proteome in coculture than it did in monoculture.

Comparative analysis of the significant changes in bacterial protein abundance revealed that the majority of changes occurred uniquely in either continuous light, diurnal night, or diurnal day (Fig 3E, S5 Fig, and S4 Table). Two bacterial proteins were significantly increased regardless of light regime: WP_010914712.1, annotated as an extracellular solute-binding protein, and WP_010912842.1, annotated as a sugar-binding protein. Light-insensitive changes in proteins with such functions could be expected in coculture, where the bacterium relies on *C. reinhardtii* for reduced carbon substrates, compared to monocultures grown with sucrose, *M. japonicum's* preferred carbon source.

Taken together, these results suggest that although complexity-related differences in peptide detection still influence proteomic comparisons between cocultures and monocultures, the mono-mix strategy together with data normalization may allow one to capture more meaningful changes in the proteome resulting from an algal-bacterial interaction by label-free DDA.

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.22.666210; this version posted July 26, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

Fractionation improves detection and differential expression analysis of sparse proteins in coculture

Sample fractionation is a strategy that physically separates analytes in complex samples (e.g., by liquid chromatography (LC)) to reduce the diversity of ions entering the mass spectrometer at any given moment [53–57]. This simplification enables the instrument to perform additional fragmentation scans to identify and quantify detectable compounds, decreasing the impact of complexity, particularly for low-abundance analytes.

To test whether fractionation could improve detection of bacterial proteins from coculture by label-free DDA, we analyzed the proteome of a coculture sample after separation into 1, 6, 12, or 24 fractions. Each level of fractionation substantially increased the number of proteins that could be detected in coculture, both for *M. japonicum* as well as for *C. reinhardtii* (Fig 4A). We next sought to combine fractionation with the mono-mix strategy to determine changes in the *M. japonicum* proteome upon interaction with the alga. Since the diurnal night samples had the most comparable cocultures and mono-mix controls (Fig 2), we chose these samples as our test-case and analyzed them each as a single fraction or as 12 fractions by label-free DDA (S1 and S5 Tables). This level of fractionation resulted in a 1.8-fold increase in the number of algal proteins detected and a 1.9-fold increase in bacterial proteins detected (Fig 4B). In addition, the fractionated samples showed no significant difference in the distribution of protein abundances between the coculture and the respective mono-mix controls for either organism (Fig 4C bottom, Wilcoxon rank-sum tests, $p > 0.05$).

We next determined how combining fractionation with the mono-mix strategy influenced differential expression analysis of cocultures relative to monocultures. We found that running samples as 12 fractions nearly doubled the number of proteins that met the detection criteria for differential expression analysis. Over 4000 more algal proteins and 2000 more bacterial proteins could be assessed for changes in coculture relative to mono-mix controls when samples were fractionated (Figs 5A and 5B). Most of these newly captured proteins were of lower abundance on average.

While sample fractionation did not reveal any additional significant changes in the algal proteome (Fig 5A), nearly 4X the number of significant changes in bacterial protein abundance could be observed if samples were fractionated (Fig 5B and S6 Fig, S6 Table). Of the 55 *M. japonicum* proteins that exhibited a significant change in abundance in coculture relative to the mono-mix controls, 15 did not meet our detection criteria without sample fractionation. The remaining 40 proteins were detected in samples regardless of fractionation, but a significant change ($|\log_2\text{-fold change}| > 1$, *p-adj.* < 0.05) was only observable if the samples were fractionated. Thus, sample fractionation improves differential expression analysis of bacterial proteins in coculture not only by increasing the detection of lowly abundant proteins, but also by increasing the signal-to-noise, such that changes in abundance can be observed.

Of the 36 significant increases in protein abundance in coculture observed if samples were analyzed as 12 fractions, 34 of them were missed if samples were analyzed as 1 fraction (Fig 5C). Gene ontology (GO) enrichment analysis revealed that in addition to “outer membrane-bounded periplasmic space” proteins, *M. japonicum* increased the abundance of groups of proteins significantly enriched for transport, thiamine-binding, and carbohydrate-binding activity (Fig 5D). These additional changes detectable upon sample

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.22.666210; this version posted July 26, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

fractionation may provide clues about the kinds of metabolites that are exchanged between *C. reinhardtii* and the bacterium.

Discussion

Here, we have documented a concerning artifact in label-free DDA LC-MS/MS data from mixed algal-bacterial cocultures, developed several strategies to minimize and even overcome this artifact, and enabled more meaningful differential expression analysis of algal and bacterial proteins in cocultures relative to monocultures from such data. We found that detection of bacterial proteins was decreased in the presence of an alga, both in terms of the number of bacterial proteins detected and in terms of their measured abundances (Fig 1). Standard differential expression analysis on these data comparing monocultures and cocultures with sucrose suggests that *M. japonicum* downregulates 78% of its proteome in the presence of *C. reinhardtii* – an artifact that reflects increased suppression of bacterium-derived peptide ions by abundant alga-derived peptide ions in mixed samples (S2 Fig). Reanalysis of publicly available label-free DDA proteomics data from other alga-bacterium and bacterium-bacterium interactions suggests that increased complexity in coculture samples decreases microbial protein detection regardless of the identity or relative abundance of species present (S1 Fig) [21,48].

We first assessed an *in silico* data normalization strategy to improve DDA proteomic analysis of the algal-bacterial interaction, as has been suggested in the literature [36]. Although data normalization led to greater overlap in the distributions of protein abundances for both species, differential expression analysis using quantile-normalized values still suggested that most bacterial proteins were significantly decreased in abundance in coculture relative to monoculture (Fig 1). Thus, we found that data normalization alone does not eliminate the influence of altered detection on differential expression analysis of sample types with different relative abundances of two organisms.

To enable differential expression analysis of DDA data from mixed algal-bacterial samples, we conceived of a modified experimental design that we term the “mono-mix” strategy, where *C. reinhardtii* monocultures and *M. japonicum* monocultures are combined into a single mono-mix control that emulates the bacteria-to-algae ratio of the coculture samples. Using this strategy, detection of bacterial proteins was more comparable in cocultures and monocultures (Fig 2). However, we found that even small differences in relative abundance still appear to shape peptide detection and differential expression analysis.

It was challenging to achieve comparable bacteria-to-algae ratios in the mono-mix samples, as bacterial density is estimated *post hoc* from CFU counts of viable cells. Furthermore, previous work has shown that these counts are not an accurate reflection of bacterial biomass in this system, as *M. japonicum* undergoes reductive division when in coculture with *C. reinhardtii* [26]. For future work, we recommend preparing several mono-mix samples of various target relative abundances and selecting the samples whose actual relative abundance of biomass best matches the parallel coculture for comparison. Alternatively, monoculture samples could be frozen and then combined *post hoc* once measurements of relative abundance in the coculture samples have been completed. In addition, fluorescence in situ hybridization (FISH), flow cytometry, or total

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.22.666210; this version posted July 26, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

organic carbon measurements may be used to more accurately estimate bacterial biomass [58].

Even though complexity-related differences in peptide detection still influenced the data, the mono-mix experiment revealed that light regime shapes the *M. japonicum* proteome during coculture with *C. reinhardtii*, but not during monoculture (Fig 3). *C. reinhardtii* exhibits massive changes in its proteome and metabolome over diurnal cycles [39,59,60]. Previous work showed that the exchange of reduced carbon between *C. reinhardtii* and *M. japonicum*, as well as the resulting amount of *M. japonicum* growth, is reduced in diurnal light compared to continuous light [26]. Unsurprisingly, we found that the bacterial proteome was distinct in continuous light and diurnal light, likely reflecting more drastic carbon limitation in the latter situation (Fig 3D). Interestingly though, we found that the bacterial proteome is also quite distinct in the day and night phases of the same diurnal period, showing that *M. japonicum* is shaped by *C. reinhardtii's* diurnal program.

Finally, we augmented the mono-mix approach by using sample fractionation to improve detection of lowly abundant bacterial proteins. By analyzing peptide samples divided into 12 fractions, we achieved over a 1.8-fold increase in the number of proteins that could be detected from either organism (Fig 4B). Even analyzing samples as 6 fractions allows a substantial increase in protein detection (Fig 4A), and thus users may weigh the costs and benefits of fractionation level depending on the number of experimental conditions. Not only did sample fractionation increase the number of proteins detected, it also increased the signal-to-noise of measurements, capturing nearly 4X the number of significant changes in bacterial protein abundances (Fig 5). This included an increase in proteins with peptide transport, carbohydrate binding, and thiamine binding activities. Thus, our proteomic analysis has yielded novel clues about an algal-bacterial interaction.

Our analysis recognizes and emphasizes the potential for misinterpreting proteomic comparisons of systems that differ in their ecological complexity. The problem is not unique to this study or these organisms, since the phenomenon is evident also in independent work in the literature [21,34,36,48,51]. Furthermore, we found that even subtle differences in the relative abundance of an alga and a bacterium across different samples could lead to differences in peptide detection (Fig 2). Thus, we predict that in addition to monocultures and cocultures, any two treatments that exhibit substantial differences in the relative abundance of interacting organisms would suffer from altered detection by DDA that could result in false positives during differential expression analysis.

Here, we have shown that when used in tandem, data normalization, sample fractionation, and the mono-mix strategy enable discrimination of meaningful changes that arise during a microbial symbiosis from label-free DDA proteomics. These simple strategies may also strengthen other approaches that provide deeper coverage of the proteome. For example, an accurate mass and time (AMT) tag approach could be employed [61,62], in which an initial shotgun LC-MS/MS analysis with the high mass accuracy of Fourier transform ion cyclotron resonance (FTICR) MS could be performed on monoculture samples to generate a database of peptide markers for each species of interest. AMT databases are used to identify peptides by their characteristic accurate mass and elution times during LC-MS alone, allowing detection of many more peptides

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.22.666210; this version posted July 26, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

675 [61,62]. Alternatively, data-independent acquisition (DIA) LC-MS/MS proteomics could be
676 used, in which powerful instrumentation enables fragmentation and monitoring of all
677 precursor ions within a given mass scan [63,64]. DIA has recently been used to improve
678 proteomic coverage of complex microbial community samples [63–67]. We posit that such
679 approaches applied in conjunction with the mono-mix strategy would not only allow
680 unbiased, high coverage of the proteomes of multiple microbes from mixed culture
681 samples, but also high-fidelity determination of significant changes in protein abundance
682 resulting from symbiosis.

683 Acknowledgments

684 We thank Dr. Rhona Stuart for providing data on a different symbiotic system and her
685 encouragement of our work. We are grateful to Dr. Joseph Loo, Dr. Rachel Loo, and Dr.
686 Anne G. Glaesener for critical feedback on the manuscript.

687 Funding

688 This work was supported by The Gordon and Betty Moore Foundation Symbiosis in
689 Aquatic Systems Initiative Investigator Award GBMF9203 to S.S.M.
690 (https://doi.org/10.37807/GBMF9203). Proteomic analysis was performed on a Facilities
691 Integrating Collaborations for User Science (FICUS) project award
692 (10.46936/fics.proj.2021.60010/60000390) from the Environmental Molecular Sciences
693 Laboratory, a Department of Energy Office of Science User Facility sponsored by the
694 Biological and Environmental Research program under Contract No. DE-AC05-
695 76RL01830. S.D. acknowledges support from the National Institutes of Health T32
696 Genetic Dissection of Cells and Organisms training grant 1T32GM132022-01. U.F.L.
697 acknowledges support from the μBiospheres Scientific Focus Area grant SCW1039
698 supported by the Genome Sciences Program of the U.S. Department of Energy's Office
699 of Biological and Environmental Research. The funders had no role in study design, data
700 collection and analysis, decision to publish, or preparation of the manuscript.

701 Author Contributions

702 S.D. and S.S.M. conceived of the project and procured funding for the study. S.D. and
703 U.F.L. designed and carried out experiments. M.S.L. directed proteomics analyses.
704 C.D.N. performed sample processing and LC-MS/MS data collection. S.O.P. processed
705 LC-MS/MS data. S.D.G. performed genome analysis. S.D. and L.C. analyzed proteomics
706 data. S.D. wrote the manuscript with input from all authors.

707 Competing Interests

708 The authors have declared that no competing financial interests exist.

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.22.666210; this version posted July 26, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

References

1. Robert E. Hungate. The Rumen and its Microbes. Elsevier; 1966. doi:10.1016/C2013-0-12555-X

2. Mizrahi I, Wallace RJ, Moraïs S. The rumen microbiome: balancing food security and environmental impacts. Nat Rev Microbiol. 2021;19: 553–566. doi:10.1038/s41579-021-00543-6

3. Zhang L, Feng G, Declerck S. Signal beyond nutrient, fructose, exuded by an arbuscular mycorrhizal fungus triggers phytate mineralization by a phosphate solubilizing bacterium. The ISME Journal. 2018;12: 2339–2351. doi:10.1038/s41396-018-0171-4

4. Rozmoš M, Bukovská P, Hršelová H, Kotianová M, Dudáš M, Gančarčíková K, et al. Organic nitrogen utilisation by an arbuscular mycorrhizal fungus is mediated by specific soil bacteria and a protist. The ISME Journal. 2022;16: 676–685. doi:10.1038/s41396-021-01112-8

5. Duan S, Feng G, Limpens E, Bonfante P, Xie X, Zhang L. Cross-kingdom nutrient exchange in the plant–arbuscular mycorrhizal fungus–bacterium continuum. Nat Rev Microbiol. 2024;22: 773–790. doi:10.1038/s41579-024-01073-7

6. Spribille T, Resl P, Stanton DE, Tagirdzhanova G. Evolutionary biology of lichen symbioses. New Phytologist. 2022;234: 1566–1582. doi:10.1111/nph.18048

7. Field CB, Behrenfeld MJ, Randerson JT, Falkowski P. Primary Production of the Biosphere: Integrating Terrestrial and Oceanic Components. Science. 1998;281: 237–240. doi:10.1126/science.281.5374.237

8. Seymour JR, Amin SA, Raina J-B, Stocker R. Zooming in on the phycosphere: the ecological interface for phytoplankton–bacteria relationships. Nature Microbiology. 2017;2: 17065. doi:10.1038/nmicrobiol.2017.65

9. Pomeroy LR. The ocean's food web, a changing paradigm. BioScience. 1974;24: 499–504. doi:10.2307/1296885

10. Azam F, Fenchel T, Field J, Gray J, Meyer-Reil L, Thingstad F. The ecological role of water-column microbes in the sea. Marine Ecology Progress Series. 1983;10: 257–263. doi:10.3354/meps010257

11. Azam F, Malfatti F. Microbial structuring of marine ecosystems. Nature Reviews Microbiology. 2007;5: 782–791. doi:10.1038/nrmicro1747

12. Salomé PA, Merchant SS. A Series of Fortunate Events: Introducing Chlamydomonas as a Reference Organism. Plant Cell. 2019;31: 1682–1707. doi:10.1105/tpc.18.00952

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.22.666210; this version posted July 26, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

13. Dupuis S, Merchant SS. *Chlamydomonas reinhardtii*: a model for photosynthesis and so much more. Nature Methods. 2023;20: 1441–1442. doi:10.1038/s41592-023-02023-6

14. Xie B, Bishop S, Stessman D, Wright D, Spalding MH, Halverson LJ. *Chlamydomonas reinhardtii* thermal tolerance enhancement mediated by a mutualistic interaction with vitamin B_{12}-producing bacteria. The ISME Journal. 2013;7: 1544–1555. doi:10.1038/ismej.2013.43

15. Helliwell KE, Collins S, Kazamia E, Purton S, Wheeler GL, Smith AG. Fundamental shift in vitamin B_{12} eco-physiology of a model alga demonstrated by experimental evolution. ISME Journal. 2015;9: 1446–1455. doi:10.1038/ismej.2014.230

16. Aiyar P, Schaeme D, García-Altares M, Carrasco Flores D, Dathe H, Hertweck C, et al. Antagonistic bacteria disrupt calcium homeostasis and immobilize algal cells. Nature Communications. 2017;8: 1756. doi:10.1038/s41467-017-01547-8

17. Fakhimi N, Gonzalez-Ballester D, Fernández E, Galván A, Dubini A. Algae-bacteria consortia as a strategy to enhance H_2 production. Cells. 2020;9: 1353. doi:10.3390/cells9061353

18. Krespach MKC, García-Altares M, Flak M, Schoeler H, Scherlach K, Netzker T, et al. Lichen-like association of *Chlamydomonas reinhardtii* and *Aspergillus nidulans* protects algal cells from bacteria. The ISME Journal. 2020;14: 2794–2805. doi:10.1038/s41396-020-0731-2

19. Krespach MKC, Stroe MC, Flak M, Komor AJ, Nietzsche S, Sasso S, et al. Bacterial marginolactones trigger formation of algal gloeocapsoids, protective aggregates on the verge of multicellularity. Proceedings of the National Academy of Sciences. 2021;118. doi:10.1073/pnas.2100892118

20. Durán P, Flores-Uribe J, Wippel K, Zhang P, Guan R, Melkonian B, et al. Shared features and reciprocal complementation of the Chlamydomonas and Arabidopsis microbiota. Nat Commun. 2022;13: 406. doi:10.1038/s41467-022-28055-8

21. Windler M, Stuart R, Deutzmann JS, Mayali X, Navid A, D'haeseleer P, et al. Bacterial exometabolites influence *Chlamydomonas* cell cycle and double algal productivity. FEMS Microbiology Ecology. 2022;98. doi:10.1093/femsec/fiac091

22. Bunbury F, Deery E, Sayer AP, Bhardwaj V, Harrison EL, Warren MJ, et al. Exploring the onset of B_{12}-based mutualisms using a recently evolved *Chlamydomonas* auxotroph and B_{12}-producing bacteria. Environmental Microbiology. 2022;24: 3134–3147. doi:10.1111/1462-2920.16035

23. Carrasco Flores D, Hotter V, Vuong T, Hou Y, Bando Y, Scherlach K, et al. A mutualistic bacterium rescues a green alga from an antagonist. Proc Natl Acad Sci USA. 2024;121: e2401632121. doi:10.1073/pnas.2401632121

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.22.666210; this version posted July 26, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

24. Bunbury F, Helliwell KE, Mehrshahi P, Davey MP, Salmon DL, Holzer A, et al. Responses of a newly evolved auxotroph of Chlamydomonas to B_{12} deprivation. Plant Physiology. 2020;183: 167–178. doi:10.1104/pp.19.01375

25. Kazamia E, Czesnick H, Nguyen TTV, Croft MT, Sherwood E, Sasso S, et al. Mutualistic interactions between vitamin B_{12}-dependent algae and heterotrophic bacteria exhibit regulation. Environmental Microbiology. 2012;14: 1466–1476. doi:10.1111/j.1462-2920.2012.02733.x

26. Dupuis S, Lingappa UF, Mayali X, Sindermann ES, Chastain JL, Weber PK, et al. Scarcity of fixed carbon transfer in a model microbial phototroph–heterotroph interaction. The ISME Journal. 2024;18: wrae140. doi:10.1093/ismejo/wrae140

27. Kleiner M, Thorson E, Sharp CE, Dong X, Liu D, Li C, et al. Assessing species biomass contributions in microbial communities via metaproteomics. Nat Commun. 2017;8: 1558. doi:10.1038/s41467-017-01544-x

28. Banfield JF, Verberkmoes NC, Hettich RL, Thelen MP. Proteogenomic Approaches for the Molecular Characterization of Natural Microbial Communities. OMICS: A Journal of Integrative Biology. 2005;9: 301–333. doi:10.1089/omi.2005.9.301

29. Hettich RL, Sharma R, Chourey K, Giannone RJ. Microbial metaproteomics: identifying the repertoire of proteins that microorganisms use to compete and cooperate in complex environmental communities. Current Opinion in Microbiology. 2012;15: 373–380. doi:10.1016/j.mib.2012.04.008

30. Hettich RL, Pan C, Chourey K, Giannone RJ. *Metaproteomics*: Harnessing the Power of High Performance Mass Spectrometry to Identify the Suite of Proteins That Control Metabolic Activities in Microbial Communities. Anal Chem. 2013;85: 4203–4214. doi:10.1021/ac303053e

31. Otto A, Becher D, Schmidt F. Quantitative proteomics in the field of microbiology. Proteomics. 2014;14: 547–565. doi:10.1002/pmic.201300403

32. Liu H, Sadygov RG, Yates JR. A Model for Random Sampling and Estimation of Relative Protein Abundance in Shotgun Proteomics. Anal Chem. 2004;76: 4193–4201. doi:10.1021/ac0498563

33. Guo J, Huan T. Comparison of Full-Scan, Data-Dependent, and Data-Independent Acquisition Modes in Liquid Chromatography–Mass Spectrometry Based Untargeted Metabolomics. Anal Chem. 2020;92: 8072–8080. doi:10.1021/acs.analchem.9b05135

34. Giannone RJ, Huber H, Karpinets T, Heimerl T, Küper U, Rachel R, et al. Proteomic Characterization of Cellular and Molecular Processes that Enable the *Nanoarchaeum equitans-Ignicoccus hospitalis* Relationship. Randau L, editor. PLoS ONE. 2011;6: e22942. doi:10.1371/journal.pone.0022942

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.22.666210; this version posted July 26, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

35. Chignell JF, Park S, Lacerda CMR, De Long SK, Reardon KF. Label-Free Proteomics of a Defined, Binary Co-culture Reveals Diversity of Competitive Responses Between Members of a Model Soil Microbial System. Microb Ecol. 2018;75: 701–719. doi:10.1007/s00248-017-1072-1

36. Shi M, Evans CA, McQuillan JL, Noirel J, Pandhal J. LFQRatio: A Normalization Method to Decipher Quantitative Proteome Changes in Microbial Coculture Systems. J Proteome Res. 2024;23: 999–1013. doi:10.1021/acs.jproteome.3c00714

37. Helliwell KE, Pandhal J, Cooper MB, Longworth J, Kudahl UJ, Russo DA, et al. Quantitative proteomics of a B_{12}-dependent alga grown in coculture with bacteria reveals metabolic tradeoffs required for mutualism. New Phytologist. 2018;217: 599–612. doi:10.1111/nph.14832

38. Zhao N, Yi L, Ren S, Yin Q, Xiang W, Zhang X, et al. Algicidal interaction between *Paenibacillus polymyxa* MEZ6 and microalgae. Journal of Applied Microbiology. 2022;133: 646–655. doi:10.1111/jam.15592

39. Strenkert D, Schmollinger S, Gallaher SD, Salomé PA, Purvine SO, Nicora CD, et al. Multiomics resolution of molecular events during a day in the life of Chlamydomonas. Proceedings of the National Academy of Sciences. 2019;116: 2374–2383. doi:10.1073/pnas.1815238116

40. Sueoka N. Mitotic replication of deoxynucleic acid in *Chlamydomonas reinhardi*. Proceedings of the National Academy of Sciences. 1960;46: 83–91. doi:10.1073/pnas.46.1.83

41. Kropat J, Hong-Hermesdorf A, Casero D, Ent P, Castruita M, Pellegrini M, et al. A revised mineral nutrient supplement increases biomass and growth rate in *Chlamydomonas reinhardtii*. The Plant Journal. 2011;66: 770–780. doi:10.1111/j.1365-313X.2011.04537.x

42. Craig RJ, Gallaher SD, Shu S, Salomé PA, Jenkins JW, Blaby-Haas CE, et al. The Chlamydomonas Genome Project, version 6: Reference assemblies for mating-type *plus* and *minus* strains reveal extensive structural mutation in the laboratory. The Plant Cell. 2023;35: 644–672. doi:10.1093/plcell/koac347

43. Chambers MC, Maclean B, Burke R, Amodei D, Ruderman DL, Neumann S, et al. A cross-platform toolkit for mass spectrometry and proteomics. Nat Biotechnol. 2012;30: 918–920. doi:10.1038/nbt.2377

44. Kim S, Pevzner PA. MS-GF+ makes progress towards a universal database search tool for proteomics. Nat Commun. 2014;5. doi:10.1038/ncomms6277

45. Elias JE, Gygi SP. Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. Nat Methods. 2007;4: 207–214. doi:10.1038/nmeth1019

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.22.666210; this version posted July 26, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

855 46. Monroe ME, Shaw JL, Daly DS, Adkins JN, Smith RD. MASIC: A software program
856 for fast quantitation and flexible visualization of chromatographic profiles from
857 detected LC–MS(/MS) features. Computational Biology and Chemistry. 2008;32:
858 215–217. doi:10.1016/j.compbiolchem.2008.02.006

859 47. Jones P, Binns D, Chang H-Y, Fraser M, Li W, McAnulla C, et al. InterProScan 5:
860 genome-scale protein function classification. Bioinformatics. 2014;30: 1236–1240.
861 doi:10.1093/bioinformatics/btu031

862 48. Men Y, Feil H, VerBerkmoes NC, Shah MB, Johnson DR, Lee PKH, et al. Sustainable
863 syntrophic growth of *Dehalococcoides ethenogenes* strain 195 with *Desulfovibrio*
864 *vulgaris* Hildenborough and *Methanobacterium congolense*: global transcriptomic
865 and proteomic analyses. The ISME Journal. 2012;6: 410–421.
866 doi:10.1038/ismej.2011.111

867 49. Peng H, Wang H, Kong W, Li J, Goh WWB. Optimizing differential expression
868 analysis for proteomics data via high-performing rules and ensemble inference. Nat
869 Commun. 2024;15: 3922. doi:10.1038/s41467-024-47899-w

870 50. Giannone RJ, Wurch LL, Heimerl T, Martin S, Yang Z, Huber H, et al. Life on the
871 edge: functional genomic response of *Ignicoccus hospitalis* to the presence of
872 *Nanoarchaeum equitans*. The ISME Journal. 2015;9: 101–114.
873 doi:10.1038/ismej.2014.112

874 51. Walker CB, Redding-Johanson AM, Baidoo EE, Rajeev L, He Z, Hendrickson EL, et
875 al. Functional responses of methanogenic archaea to syntrophic growth. The ISME
876 Journal. 2012;6: 2045–2055. doi:10.1038/ismej.2012.60

877 52. Zones JM, Blaby IK, Merchant SS, Umen JG. High-resolution profiling of a
878 synchronized diurnal transcriptome from *Chlamydomonas reinhardtii* reveals
879 continuous cell and metabolic differentiation. Plant Cell. 2015;27: 2743–2769.
880 doi:10.1105/tpc.15.00498

881 53. Wang Y, Yang F, Gritsenko MA, Wang Y, Clauss T, Liu T, et al. Reversed-phase
882 chromatography with multiple fraction concatenation strategy for proteome profiling
883 of human MCF10A cells. Proteomics. 2011;11: 2019–2026.
884 doi:10.1002/pmic.201000722

885 54. Chiu C, Chang C, Chen S. Evaluation of peptide fractionation strategies used in
886 proteome analysis. J of Separation Science. 2012;35: 3293–3301.
887 doi:10.1002/jssc.201200631

888 55. Yang F, Shen Y, Camp DG, Smith RD. High-pH reversed-phase chromatography with
889 fraction concatenation for 2D proteomic analysis. Expert Review of Proteomics.
890 2012;9: 129–134. doi:10.1586/epr.12.15

891 56. Law HCH, Kong RPW, Li M, Szeto SSW, Chu IK. Implementation of a multiple-
892 fraction concatenation strategy in an online two-dimensional high-/low-pH reversed-

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.22.666210; this version posted July 26, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

phase/reversed-phase liquid chromatography platform for qualitative and quantitative shotgun proteomic analyses. J Mass Spectrom. 2021;56: e4591. doi:10.1002/jms.4591

57. Lenčo J, Jadeja S, Naplekov DK, Krokhin OV, Khalikova MA, Chocholouš P, et al. Reversed-Phase Liquid Chromatography of Peptides for Bottom-Up Proteomics: A Tutorial. J Proteome Res. 2022;21: 2846–2892. doi:10.1021/acs.jproteome.2c00407

58. Daims H. Use of Fluorescence In Situ Hybridization and the *daime* Image Analysis Program for the Cultivation-Independent Quantification of Microorganisms in Environmental and Medical Samples. Cold Spring Harb Protoc. 2009;2009: pdb.prot5253. doi:10.1101/pdb.prot5253

59. Jüppner J, Mubeen U, Leisse A, Caldana C, Brust H, Steup M, et al. Dynamics of lipids and metabolites during the cell cycle of *Chlamydomonas reinhardtii*. The Plant Journal. 2017;92: 331–343. doi:10.1111/tpj.13642

60. Dupuis & Ojeda, Gallaher SD, Purvine SO, Glaesener AG, Ponce R, Nicora CD, et al. Too dim, too bright, and just right: Systems analysis of the *Chlamydomonas* diurnal program under limiting and excess light. The Plant Cell. 2025; koaf086. doi:10.1093/plcell/koaf086

61. Paša-Tolić L, Masselon C, Barry RC, Shen Y, Smith RD. Proteomic Analyses using an Accurate Mass and time Tag Strategy. BioTechniques. 2004;37: 621–639. doi:10.2144/04374RV01

62. Zimmer JSD, Monroe ME, Qian W, Smith RD. Advances in proteomics data analysis and display using an accurate mass and time tag approach. Mass Spectrometry Reviews. 2006;25: 450–482. doi:10.1002/mas.20071

63. Kitata RB, Yang J, Chen Y. Advances in data-independent acquisition mass spectrometry towards comprehensive digital proteome landscape. Mass Spectrometry Reviews. 2023;42: 2324–2348. doi:10.1002/mas.21781

64. Guzman UH, Martinez-Val A, Ye Z, Damoc E, Arrey TN, Pashkova A, et al. Ultra-fast label-free quantification and comprehensive proteome coverage with narrow-window data-independent acquisition. Nat Biotechnol. 2024;42: 1855–1866. doi:10.1038/s41587-023-02099-7

65. Aakko J, Pietilä S, Suomi T, Mahmoudian M, Toivonen R, Kouvonen P, et al. Data-Independent Acquisition Mass Spectrometry in Metaproteomics of Gut Microbiota—Implementation and Computational Analysis. J Proteome Res. 2020;19: 432–436. doi:10.1021/acs.jproteome.9b00606

66. Pietilä S, Suomi T, Elo LL. Introducing untargeted data-independent acquisition for metaproteomics of complex microbial samples. ISME Communications. 2022;2: 51. doi:10.1038/s43705-022-00137-0

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.22.666210; this version posted July 26, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

67. Zhao J, Yang Y, Xu H, Zheng J, Shen C, Chen T, et al. Data-independent acquisition boosts quantitative metaproteomics for deep characterization of gut microbiota. npj Biofilms Microbiomes. 2023;9: 4. doi:10.1038/s41522-023-00373-9

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.22.666210; this version posted July 26, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

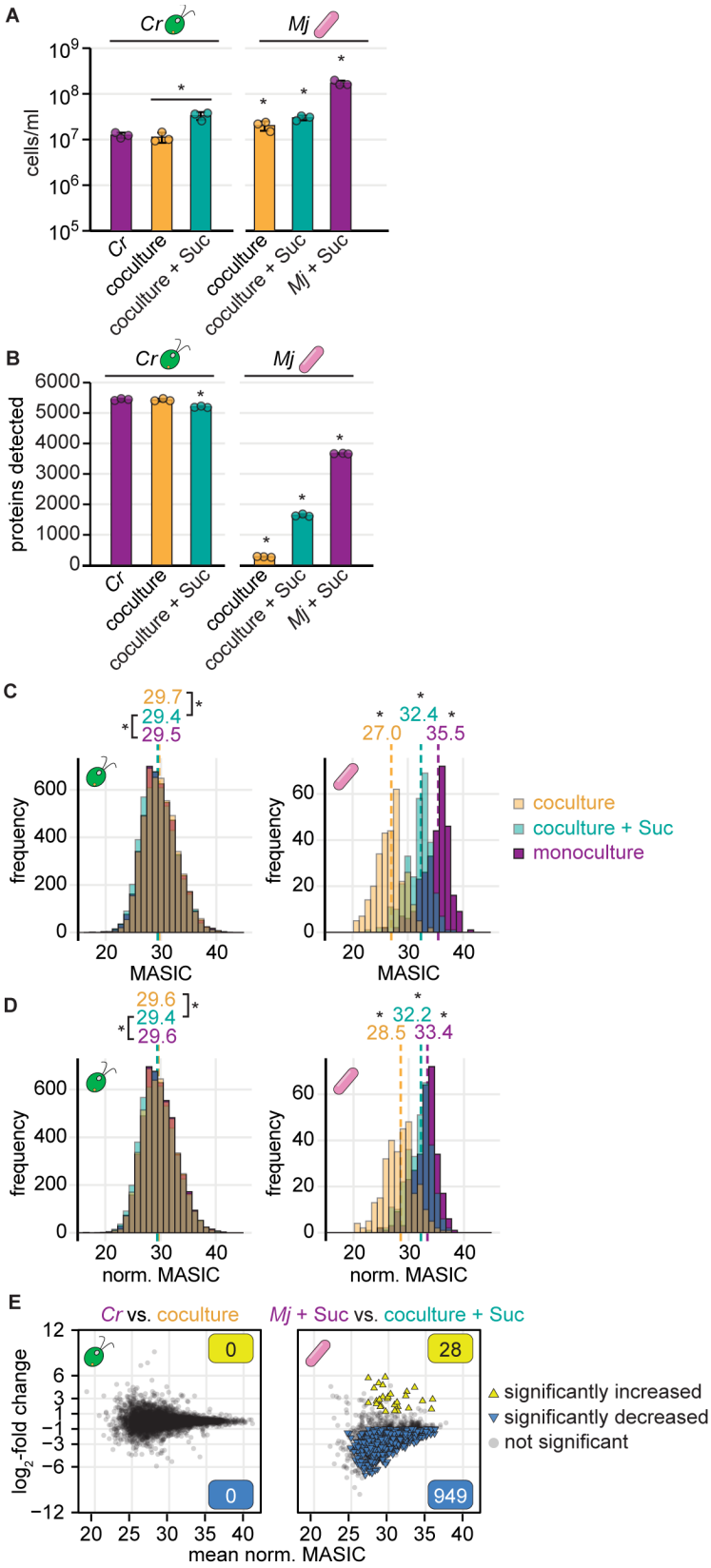


Fig 1. Detection of bacterial proteins decreases in coculture samples.

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.22.666210; this version posted July 26, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

Triplicate continuous-light-grown monocultures and cocultures with and without sucrose were collected for LC-MS/MS proteomics. See also S1 and S2 Figs.

(A) Density of *C. reinhardtii* cells (left) and *M. japonicum* CFU (right) in triplicate monocultures (purple), cocultures (orange), and cocultures with sucrose (turquoise) upon collection for proteomics. The bars represent the mean of three replicate cultures, and error bars represent the standard deviation from the mean. Asterisks indicate significant differences by two-tailed Student's t-tests ($p < 0.05$).

(B) Number of unique *C. reinhardtii* proteins (left) and *M. japonicum* proteins (right) detected in samples from (**A**) by LC-MS/MS proteomics. Asterisks indicate significant differences as in (**A**).

(C) Distribution of algal (left) and bacterial protein abundances (right) in cocultures (transparent orange), cocultures with sucrose (transparent turquoise), and monocultures (purple) for proteins detected in all three conditions. When distributions overlap, the color is muddied. The unnormalized MASIC values were averaged across the three biological replicates. Median values are shown above the dashed lines. Asterisks next to the medians indicate significant differences between the indicated distributions by a Wilcoxon rank-sum test ($p < 0.05$).

(D) Distribution of algal (left) and bacterial protein abundances (right) in cocultures (transparent orange), cocultures with sucrose (transparent turquoise), and monocultures (purple) after quantile normalization, presented as in (**C**).

(E) Changes in quantile-normalized protein abundances in coculture relative to monoculture for *C. reinhardtii* (left) and *M. japonicum* (right). Significant differences (triangles) were defined as those where $|\log_2\text{-fold change}| > 1$, *p-adj.* < 0.05 from a Student's t-test of the quantile-normalized MASIC values, and the mean MASIC value was greater than the limit of quantitation in both the coculture and the monoculture.

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.22.666210; this version posted July 26, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

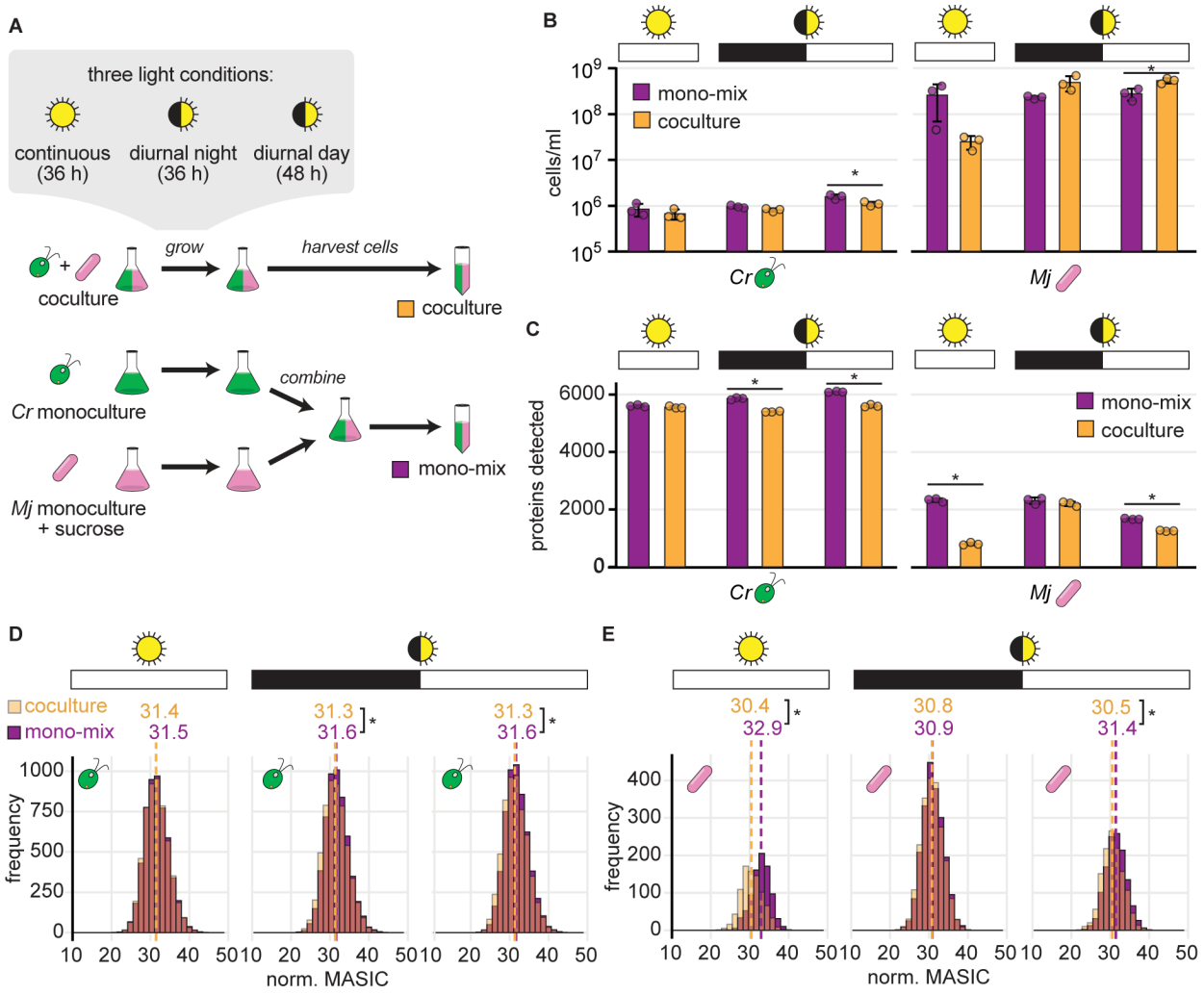


Fig 2. The mono-mix strategy enables comparable DDA detection of bacterial proteins in the coculture and the monoculture control when a similar bacteria-to-algae ratio is achieved. See also S3 and S4 Figs.

(A) Schematic of experiment. Triplicate cocultures of *C. reinhardtii* (green) and *M. japonicum* (pink) were grown in parallel with triplicate monocultures of *C. reinhardtii* and *M. japonicum* with 150 μg/ml sucrose. Then, the *M. japonicum* monocultures were added to the *C. reinhardtii* monocultures to achieve a "mono-mix" control (purple) with a similar bacteria-to-algae ratio as the coculture (orange). Cultures were grown in either continuous (sun icon) or diurnal light (eclipsed sun icon). Continuous light cultures were collected 36 h after inoculation, and the diurnal light cultures were collected at the end of the night (36 h after inoculation) and the end of the day (48 h after inoculation).

(B) Density of *C. reinhardtii* cells and *M. japonicum* CFU in triplicate cocultures (orange) and mono-mix controls (purple) grown in continuous light (sun icon) or diurnal light (eclipsed sun icon) upon collection for proteomics at the end of the dark or light phases (black or white bars, respectively). The bars represent the mean of three replicate cultures, and error bars represent the standard deviation from the mean. Asterisks indicate significant differences by two-tailed Student's t-tests ($p < 0.05$).

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.22.666210; this version posted July 26, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

(C) Number of unique algal (left) and bacterial proteins (right) detected in the samples from (**B**). Asterisks indicate significant differences as in (**B**).

(D) Distribution of normalized abundances of algal proteins in cocultures (transparent orange) and mono-mix controls (purple) grown in continuous light (sun icon) or diurnal light (eclipsed sun icon) collected at the end of the dark or light phases (black or white bars, respectively). When distributions overlap, the color is muddied. The quantile-normalized MASIC values were averaged across the three biological replicates. Median values are shown above the dashed lines. Asterisks next to the medians indicate significant differences between the indicated distributions by a Wilcoxon rank-sum test ($p < 0.05$).

(E) Distribution of normalized abundances of bacterial proteins presented as in (**D**).

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.22.666210; this version posted July 26, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

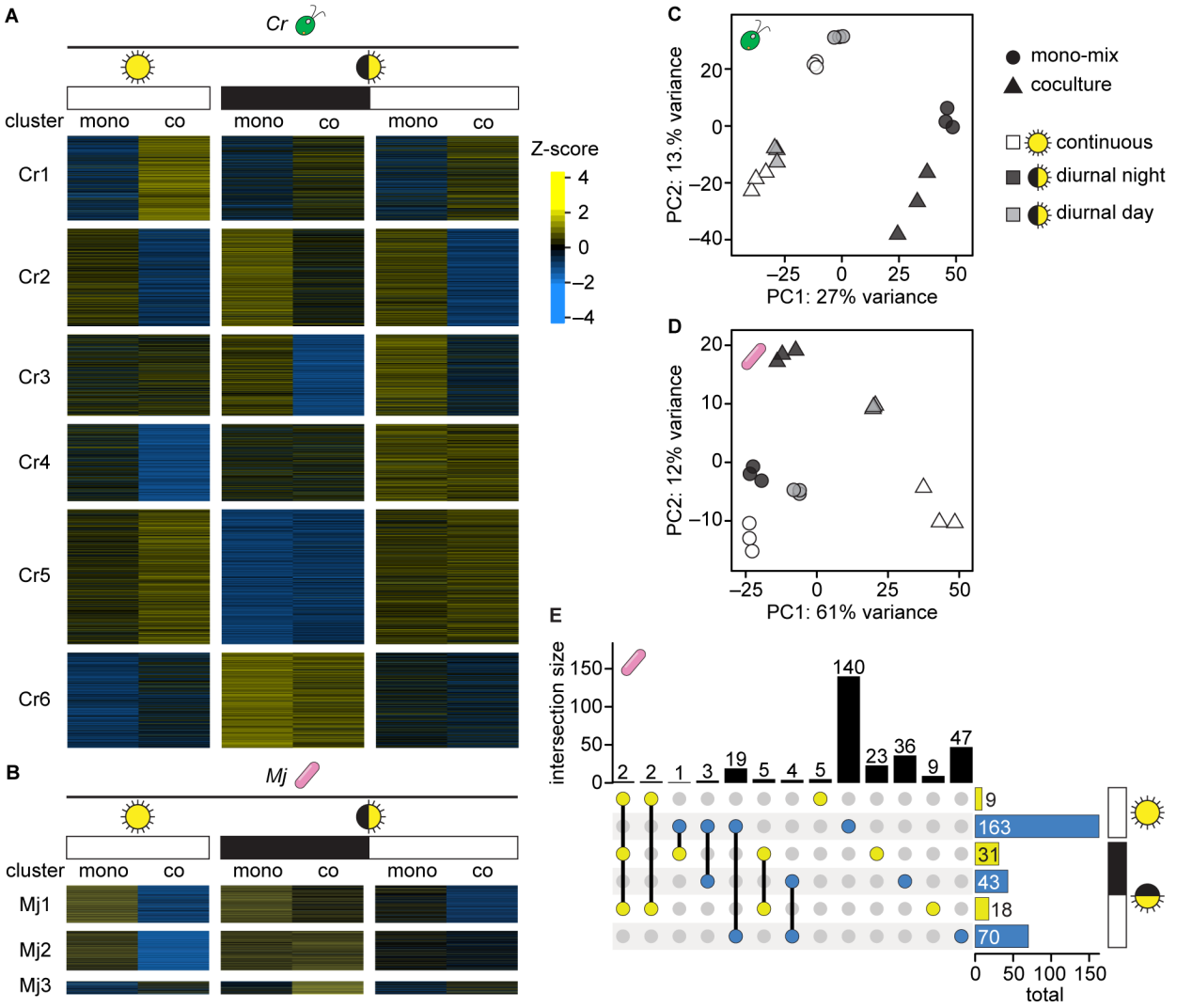


Fig 3. Data normalization and comparison to mono-mix controls reveals that light regime may influence the *M. japonicum* proteome during coculture with *C. reinhardtii*.

(A) *k*-means clustering (*k* = 6) of the Z-score of quantile-normalized mean algal protein abundances in the mono-mix experiment. 4737 *C. reinhardtii* proteins that were detected in at least two biological replicates in all conditions were included in the analysis.

(B) *k*-means clustering (*k* = 3) of the Z-score of quantile-normalized mean bacterial protein abundances in the mono-mix experiment. 828 *M. japonicum* proteins that were detected in at least two biological replicates in all conditions were included in the analysis.

(C) PCA of quantile-normalized protein abundances for the 3817 *C. reinhardtii* proteins that were detected in all samples: three biological replicates of the mono-mix (circles) and coculture samples (triangles) under continuous light (white), diurnal night (dark grey), and diurnal day (light grey).

(D) PCA of quantile-normalized protein abundances for the 569 *M. japonicum* proteins that were detected in all samples, presented as in (**C**).

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.22.666210; this version posted July 26, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

(E) Intersections and differences in the bacterial proteins that were significantly increased (yellow) and significantly decreased (blue) in samples collected from continuous light, diurnal night, and diurnal day. Intersection sizes (indicated by lines connecting colored dots) and differences (indicated by single colored dots) are shown as the vertical bars, while the total number of bacterial proteins that are significantly increased and decreased are shown as the horizontal bars. Significant differences were defined as those where $|\log_2\text{-fold change}| > 1$, *p-adj.* < 0.05 from a Student's t-test of the quantile-normalized MASIC values, and the mean MASIC value was greater than the limit of quantitation in both the coculture and the monoculture. See also S5 Fig.

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.22.666210; this version posted July 26, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

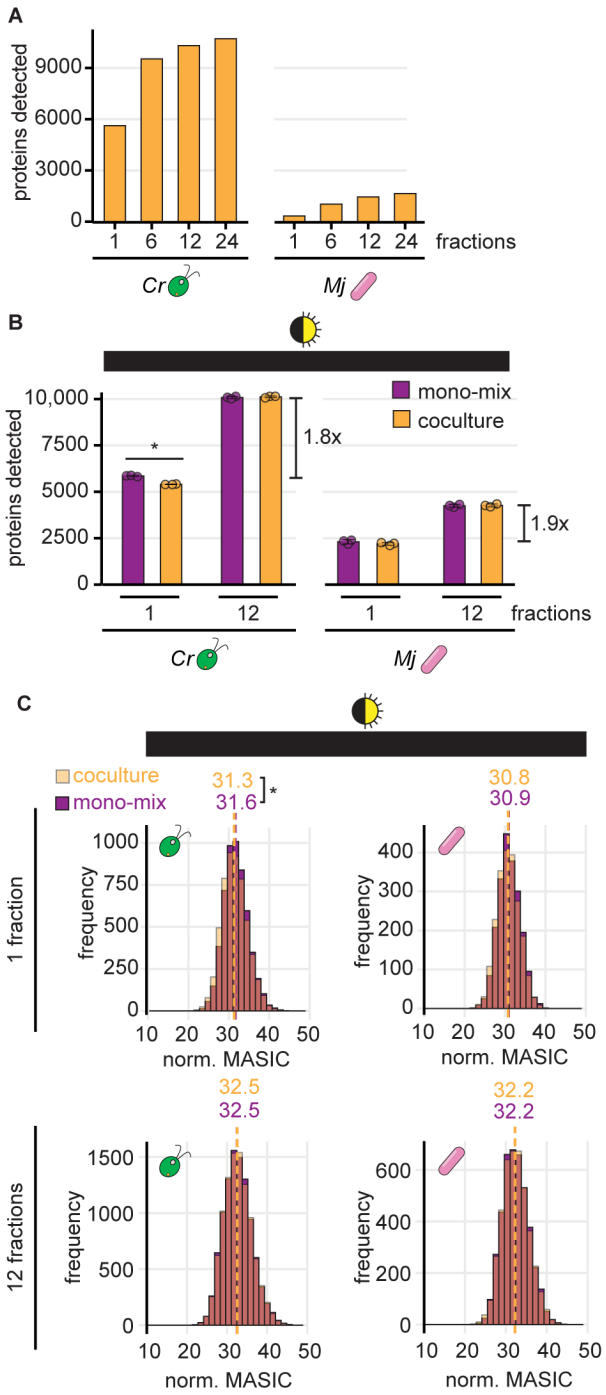


Fig 4. Fractionation improves detection of sparse bacterial proteins in coculture.

(A) The number of unique algal (left) and bacterial proteins (right) detected when a coculture sample was analyzed as 1, 6, 12, or 24 fractions.

(B) The number of unique algal (left) and bacterial proteins (right) detected from cocultures (orange) and mono-mix controls (purple) collected at the end of the night from diurnal light-grown cultures when analyzed as 1 or 12 fractions. Bars represent the mean of three replicate cultures, and error bars represent the standard deviation from the mean. Asterisks indicate significant differences by two-tailed Student's t-tests ($p < 0.05$).

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.22.666210; this version posted July 26, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

(C) Distribution of normalized abundances of algal (left) and bacterial proteins (right) in cocultures (transparent orange) and mono-mix controls (purple) collected at the end of the night when analyzed as 1 (top) or 12 fractions (bottom). When distributions overlap, the color is muddied. The quantile-normalized MASIC values were averaged across the three biological replicates. Median values are shown above the dashed lines. Asterisks next to the medians indicate significant differences between the indicated distributions by a Wilcoxon rank-sum test ($p < 0.05$).

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.22.666210; this version posted July 26, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

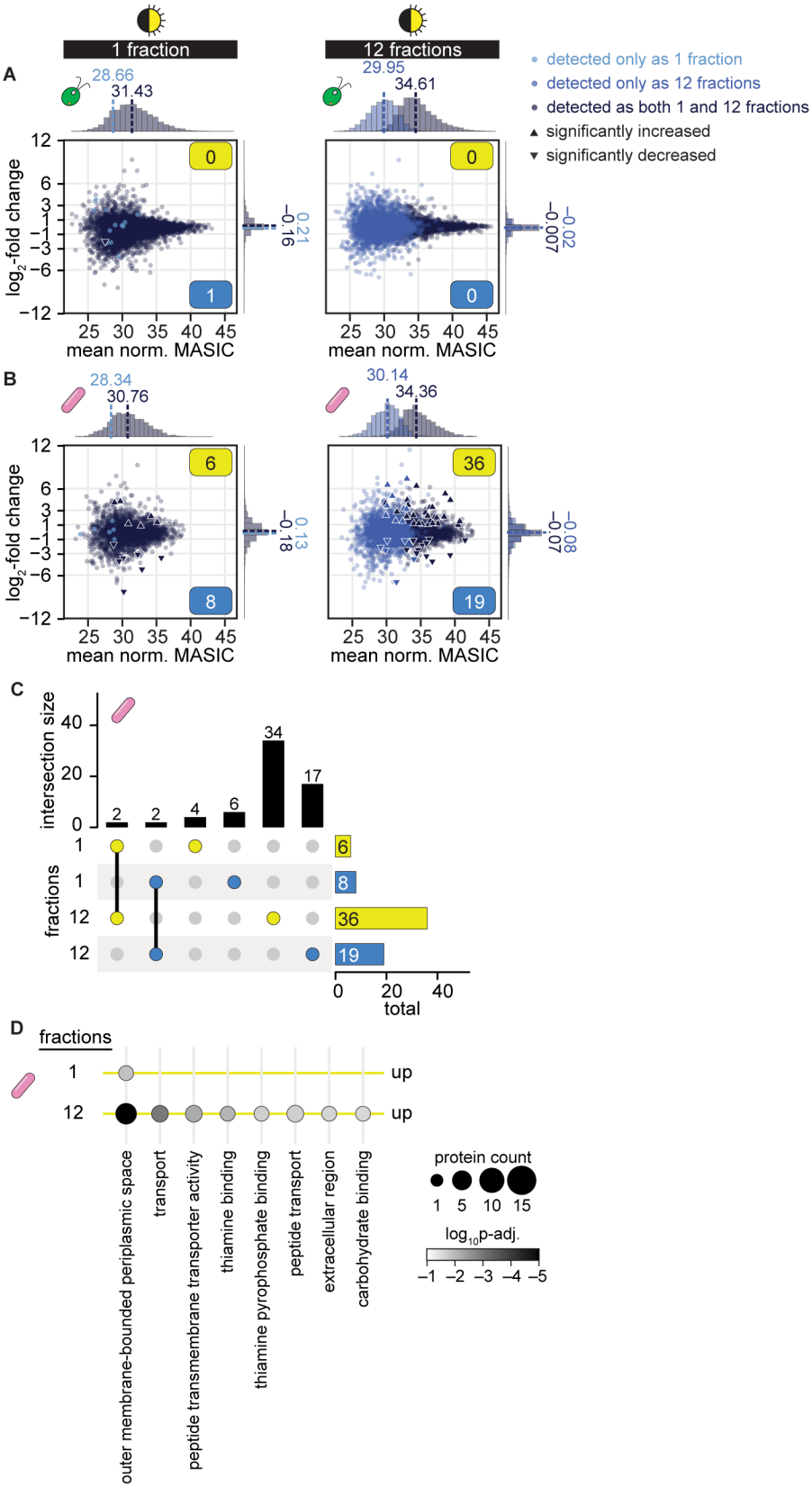


Fig 5. Sample fractionation improves differential expression analysis of proteins in cocultures relative to mono-mix controls. See also S6 Fig.

bioRxiv preprint doi: https://doi.org/10.1101/2025.07.22.666210; this version posted July 26, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

(A) Changes in quantile-normalized *C. reinhardtii* protein abundance in coculture relative to mono-mix controls at the end of the night when samples were analyzed as 1 (left) or 12 fractions (right). Points are colored based on their detection criteria: lighter blue indicates that the protein only met the criteria when the sample was analyzed as either 1 or 12 fractions, navy blue indicates that the protein met the criteria when the sample was analyzed both as 1 and as 12 fractions. Significant differences (triangles) were defined as those where $|\log_2\text{-fold change}| > 1$, *p-adj.* < 0.05 from a Student's t-test of the quantile-normalized MASIC values, and the mean MASIC value was greater than the limit of quantitation in both the coculture and the mono-mix. Yellow and blue boxes indicate the number of significantly increased and decreased proteins, respectively. Histograms show the distribution of mean protein abundances (*x* axes) and of $\log_2$-fold changes (*y* axes) for proteins that only met the detection criteria when the sample was analyzed as either 1 or 12 fractions (lighter blue) and for proteins that met the detection criteria when the sample was analyzed both as 1 and as 12 fractions (navy blue); median values are shown above the dashed lines.

(B) Changes in quantile-normalized *M. japonicum* protein abundance in coculture relative to mono-mix controls at the end of the night when samples were analyzed as 1 (left) or 12 fractions (right) presented as in (**A**).

(C) Intersections and differences in the bacterial proteins that were significantly increased (yellow) and significantly decreased (blue) in coculture when samples were analyzed as 1 or 12 fractions. Intersection sizes (indicated by lines connecting colored dots) and differences (indicated by single colored dots) are shown as the vertical bars, while the total number of bacterial proteins that are significantly increased and decreased are shown as the horizontal bars.

(D) GO term enrichment in bacterial proteins that were significantly increased in quantile-normalized abundance in coculture when samples were analyzed as 1 or 12 fractions. The number of proteins with the corresponding GO term in each comparison is indicated by dot size and the *p-adj.* by the shading. No GO terms were significantly enriched in the bacterial proteins that were significantly decreased in quantile-normalized abundance in coculture, regardless of sample fractionation.