MPMI Vol. 35, No. 11, 2022, pp. 977–988, https://doi.org/10.1094/MPMI-05-22-0116-TA

TECHNICAL ADVANCE

Evaluation of Protein Extraction Methods for Metaproteomic Analyses of Root-Associated Microbes

Fernanda Salvato,[1,†] Simina Vintila,[1] Omri M. Finkel,[2] Jeffery L. Dangl,[2] and Manuel Kleiner[1,†]

[1] Department of Plant and Microbial Biology, North Carolina State University, Raleigh, NC 27607, U.S.A.
[2] Department of Biology and Howard Hughes Medical Institute, University of North Carolina, Chapel Hill, NC 27599, U.S.A.

Accepted for publication 21 July 2022.

Metaproteomics is a powerful tool for the characterization of metabolism, physiology, and functional interactions in microbial communities, including plant-associated microbiota. However, the metaproteomic methods that have been used to study plant-associated microbiota are very laborious and require large amounts of plant tissue, hindering wider application of these methods. We optimized and evaluated different protein extraction methods for metaproteomics of plant-associated microbiota in two different plant species (*Arabidopsis* and maize). Our main goal was to identify a method that would work with low amounts of input material (40 to 70 mg) and that would maximize the number of identified microbial proteins. We tested eight protocols, each comprising a different combination of physical lysis method, extraction buffer, and cell-enrichment method on roots from plants grown with synthetic microbial communities. We assessed the performance of the extraction protocols by liquid chromatography-tandem mass spectrometry–based metaproteomics and found that the optimal extraction method differed between the two species. For *Arabidopsis* roots, protein extraction by beating whole roots with small beads provided the greatest number of identified microbial proteins and improved the identification of proteins from gram-positive bacteria. For maize, vortexing root pieces in the presence of large glass beads yielded the greatest number of microbial proteins identified. Based on these data, we recommend the use of these two methods for metaproteomics with *Arabidopsis* and maize. Furthermore, detailed descriptions of the eight tested protocols will enable future optimization of protein extraction for metaproteomics in other dicot and monocot plants.

Keywords: endosphere, metaproteomics, microbiome, microbiota, phytobiome, plant-microbe interactions, rhizoplane

[†]Corresponding authors: F. Salvato; fersalvato@gmail.com, and M. Kleiner; manuel_kleiner@ncsu.edu

Current address for Omri M. Finkel: Department of Plant and Environmental Sciences, Institute of Life Science, The Hebrew University of Jerusalem, Jerusalem, Israel.

Funding: This work was supported by the North Carolina State Plant Soil Microbial Community Consortium, the United States National Science Foundation under award number IOS-2033621 (M. Kleiner), the U.S. Department of Agriculture National Institute of Food and Agriculture under award number 2021-67013-34537 (M. Kleiner) and the Novo Nordisk Foundation INTERACT project under award number NNF19SA0059360 (M. Kleiner). This work was also supported by a National Science Foundation grant IOS-1917270 and by the Office of Science (BER), U.S. Department of Energy, grant DE-SC0014395 to J. L. Dangl. J. L. Dangl is an Investigator of the Howard Hughes Medical Institute (HHMI) and is supported by HHMI.

***e*-Xtra:** Supplementary material is available online.

The author(s) declare no conflict of interest.

Plant-associated microbiota play important roles in plant growth, health, stress resilience, and evolution (Hassani et al. 2018; Reinhold-Hurek et al. 2015; Vorholt 2012). To understand plant-microbe interactions, most studies have used amplicon sequencing or shotgun sequencing approaches to define the core microbiota of many plant-species and to characterize how the microbiota composition varies across plant species, genotypes, geographic distance, soil types, and abiotic stresses (Barret et al. 2015; Castrillo et al. 2017; Finkel et al. 2020; Hamonts et al. 2018; Jochum et al. 2019; Lundberg et al. 2012; Niu et al. 2017; Simonin et al. 2020; Xu et al. 2018; Wagner et al. 2020).

The microbial functions that drive microbiota assembly and the various beneficial and detrimental impacts on plant phenotype are much less understood. Some initial comparative genomics, metagenomics, metabolomics, metatranscriptomics, and metaproteomics studies have shed light on microbial genes, pathways, and functions potentially critical for plant-microbe interactions (Broberg et al. 2018; Levy et al. 2018; Ofek-Lalzar et al. 2014; Zhalnina et al. 2018). However, the scope of these meta-omics studies has been limited and more studies are needed to characterize microbial genes expressed during plant-microbe interactions (López-Mondéjar et al. 2017).

Here, we focus on metaproteomics, which allows for the large-scale identification and quantification of microbial and host proteins. Proteins do most of the work inside the cell, and thus their characterization provides insights into the microorganisms' functional roles and interactions with the host (Bossche et al. 2021; Kleiner 2019; Salvato et al. 2021; Wilmes and Bond 2004). Metaproteomics has been used to study microbial communities in animals, including humans (Blakeley-Ruiz et al. 2019; Zhang et al. 2018a), and in plants (Bao et al. 2014; Delmotte et al. 2009; Knief et al. 2012), soil (Zampieri et al. 2016), and the ocean (Cohen et al. 2021; Mikan et al. 2020; Morris et al. 2010). Although metaproteomics has proven its value for the study of functional host-microbe interactions, its use for plant-associated microbiota has so far been limited.

As with other meta-omics methods, there are several steps in a metaproteomics workflow that need to be optimized for each studied system. These steps include sample preservation,

protein extraction, peptide preparation, choice of liquid chromatography-tandem mass spectrometry (LC-MS/MS) approaches, and the development of a suitable protein sequence database for peptide and protein identification (Kunath et al. 2019; Salvato et al. 2021). For metaproteomics of plant-associated microbiota, protein extraction is one of the most challenging steps for several reasons. The specific hurdles that need to be overcome in plant metaproteomics include i) the high ratio of plant to microbial proteins in the sample, which can make detecting microbial proteins difficult, ii) the massive amount of plant tissue needed if methods require separating microbes from the plant tissue prior to protein extraction (Bao et al. 2014; Knief et al. 2012), iii) interfering compounds (phenols, carbohydrates, pigments) present in the sample (Keiblinger et al. 2012), iv) the large diversity of microbial species with properties (e.g., cell wall thickness) that differentially impact protein extraction, and, finally, v) the large range of relative abundances of microbial species.

While similar challenges have been addressed for other systems, development of extraction methods specifically for plant metaproteomics has been limited. For example, to improve protein extraction from gram-positive bacteria for metaproteomics of intestinal microbiota, researchers suggested the combination of sodium dodecyl sulfate (SDS) buffer, bead-beating, and freeze-thaw cycles (Tanca et al. 2014), while others suggested a combination of SDS buffer and ultrasonication (Zhang et al. 2018b). In plant metaproteomics, the plant cell wall must be disrupted to release endophytic microorganisms, but this step is hindered by the lignin, cellulose, and cellulose derivatives that constitute the cell wall (Houston et al. 2016). In addition, plant tissues contain a wide variety of compounds, including polysaccharides, pigments, phenolics, and lipids, that can form complexes with proteins (Wu et al. 2014) that hamper protein extraction. Therefore, most plant-metaproteomics studies thus far have employed laborious procedures for the physical separation of microorganisms from plant tissue prior to protein extraction (Bao et al. 2014; Delmotte et al. 2009; Knief et al. 2012). These separation methods included density-gradient centrifugations and filtrations. For example, one study of the rice root metaproteome (Bao et al. 2014) enriched for microbes by starting with large amounts of plant tissue (100 g) that then required homogenization, filtration, and density-gradient centrifugation (Ikeda et al. 2009). Similarly, several tens to hundreds of grams of plant tissues were necessary in other studies using Percoll density gradients to enrich for bacterial cells (Delmotte et al. 2009; Knief et al. 2012).

Our goal for this study was to develop sample preparation and extraction methods for metaproteomics of root-associated microbiota that allow for increased throughput, high numbers of identified and quantified microbial proteins, efficient extraction from diverse microbial species, and the use of small amounts of roots (i.e., from single plants). For this, we evaluated different protein extraction methods in roots inoculated with synthetic communities (SynComs) previously isolated from *Arabidopsis* (Levy et al 2018) and maize (Niu et al. 2017). The *Arabidopsis* SynCom is derived from a culture collection from surface-sterilized *Arabidopsis* roots (Levy et al 2018) and contains four taxa known to influence plant growth (Finkel et al 2020). The maize SynCom is a simplified and representative bacterial community isolated from maize roots (Niu et al. 2017).

RESULTS

In order to develop and evaluate protein extraction methods for metaproteomic studies of root-associated microbiota, we used two different plant species—the dicot *Arabidopsis thaliana* and the monocot maize (*Zea mays* cv. Sugar Buns)—grown in different substrates with different synthetic microbiota. Sterile *Arabidopsis* seedlings were inoculated and were grown for 12 days on agar plates with a SynCom consisting of four bacterial strains (Levy et al. 2018): *Variovorax paradoxus* (CL014), *Arthrobacter* sp. (CL028), *Agrobacterium* sp. (MF389), and *Pseudomonas* sp. (MF397). For maize, sterile seeds were grown in sterile calcinated clay for 13 days after inoculation with a seven-strain SynCom (Niu et al. 2017) comprising *Stenotrophomonas maltophilia*, *Brucella pituitosa* (previously *Ochrobacterium pituitosum* [Hördt et al. 2020]), *Curtobacterium pusillum*, *Enterobacter ludwigii* (previously *Enterobacter cloacae*), *Chryseobacterium indologenes*, *Herbaspirillum robiniae* (previously *H. frisingense*), and *Pseudomonas putida*. In a first round of experiments, we tested the impact of six different protein extraction methods on the metaproteomes. These methods employed different mechanical cell-disruption strategies, such as bead-beating, using different matrices (BB_matrixZ and BB_matrixE), homogenization in a DUALL glass homogenizer (GH_PBS and GH_BCE), liquid nitrogen grinding (N2_PBS), and freeze-thaw cycles (FT_PBS). We also evaluated the combination of SDS and Triton X-100–based lysis buffers and the utilization of differential centrifugation to enrich for microbial cells (Table 1). To evaluate samples prepared with these different extraction methods, we prepared peptides from the extracts, using the filter-aided sample preparation (FASP) protocol (Wiśniewski et al. 2009), and injected equal amounts of peptides into the LC-MS/MS for metaproteome analyses.

Methods that maximize protein identification differ for *Arabidopsis* and maize.

We found large differences in the performance of specific extraction methods between the two plant species (Fig. 1). These differences were already visible in how well extraction methods homogenized the root tissue from either species. For example, maize roots remained almost intact when bead-beaten with matrix E, while *Arabidopsis* roots were fully disrupted by the same exact method. The differences were also large on the level of peptide and protein identification. For *Arabidopsis*, we were able to identify on average 1,259 to 3,166 bacterial proteins, 3,989 to 11,570 bacterial peptides, and 13,659 to 30,422 peptide-spectrum matches (PSMs), depending on the method. For maize,

Table 1. Summary of protein extraction methods tested[a]

Method name	Mechanical disruption method	Extraction buffers	Differential centrifugation
BB_matrixZ	Bead-beating with matrix Z, two cycles of 45 s at 8.0 m/s	SDS	
BB_matrixE	Bead-beating with matrix E, one cycle of 30 s at 4.0 m/s	SDS	
GH_PBS	DUALL glass homogenizer in PBS	PBS and SDS	X
GH_BCE	DUALL glass homogenizer in BCE	Triton X100- (BCE) and SDS	X
N_2_PBS	Grind in liquid nitrogen in PBS	PBS and SDS	X
FT_PBS	Freeze-thaw cycles + bead-beating	PBS and SDS	
Vortexing	Vortexing in PBS	PBS and SDS	
Vortexing + ultrasonication	Vortexing in PBS + ultrasonication in SDS	PBS and SDS	

[a] SDS = sodium dodecyl sulfate, PBS = phosphate buffered saline, and BCE = Triton-based buffer. X indicates that differential centrifugation was used.

these numbers were at least sixfold lower, demonstrating the need for further optimization (Fig. 1B, D, and F). In terms of reproducibility, we observed that BB_matrixE, GH_PBS, and GH_BCE showed the highest R^2 correlations between replicates in *Arabidopsis* samples (Supplementary Fig. S1). For maize, the reproducibility of these methods was far beyond optimal (Supplementary Fig. S2). The complete list of proteins identified in roots of *Arabidopsis* and maize are reported in Supplementary Tables S1 and S2, respectively.

We obtained the highest peptide yield for *Arabidopsis* with the bead-beating methods (BB_matrixZ and BB_matrixE), whereas, for maize, we obtained the highest yield using the DUALL glass homogenizer (GH_PBS) or liquid nitrogen grinding (N2_PBS) followed by differential centrifugation (Fig. 1A and B). The peptide yield did not necessarily correlate with the number of identified bacterial peptides, most likely because the prepared samples also contained plant peptides likely in different concentrations. For example, BB_matrixZ yielded the highest peptide amounts for *Arabidopsis,* however, it performed much worse in terms of peptide and protein identifications as compared with most other tested methods. Supplementary Table S1 contains plant proteins, peptides, and PSM numbers for all extraction methods. When

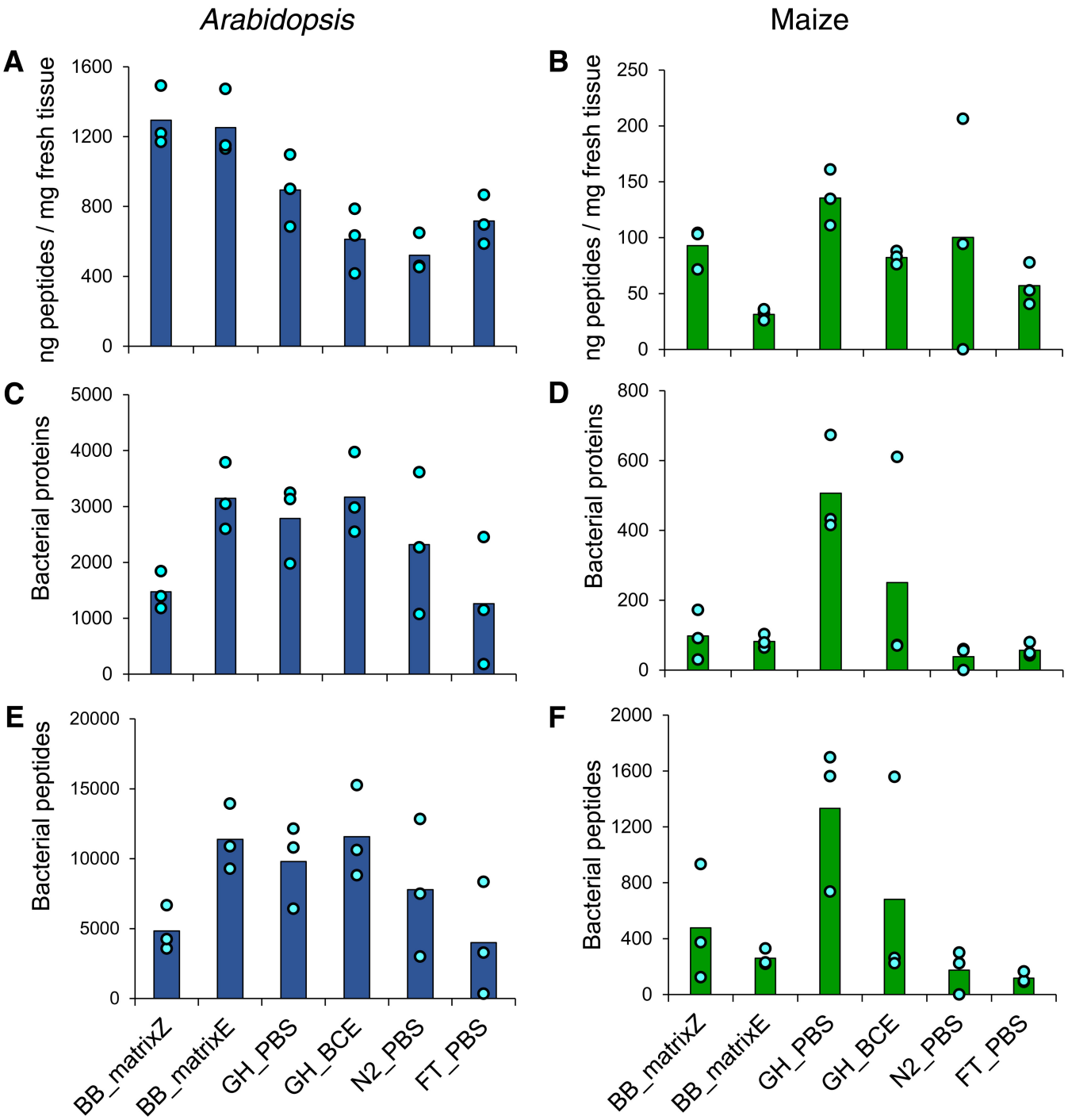

Fig. 1. Protein extraction methods differently impact the root metaproteomes of *Arabidopsis* (blue bars) and maize (green bars). **A and B,** Peptide yield per root fresh weight (nanograms/milligrams) determined by Micro BCA after sample preparation, **C and D,** number of bacterial proteins, and **E and F,** number of bacterial peptides. The data are expressed as means with error bars representing standard deviations ($n = 3$).

comparing the BB_matrixZ and BB_matrixE chromatograms for *Arabidopsis*, we observed that retention times for peptides in the BB_matrixZ samples were shifted by up to 120 min (25% of total run time). Since all samples were run together in one run with a blocked-randomized design, we could exclude LC issues as an explanation. This indicates that the difference in extraction method strongly impacted LC performance. Currently, we do not have an explanation for what caused this surprisingly large impact of extraction method on LC retention times.

To test which methods favored the extraction of bacterial proteins over plant proteins, we evaluated the total relative abundance of bacterial versus plant proteins for each method. For *Arabidopsis*, the use of the Triton-based buffer (bacterial cell extraction [BCE]) buffer during homogenization in combination with differential centrifugation (GH_BCE) resulted in the highest bacteria to plant ratio (0.37) (Fig. 2), which can be helpful for the identification of microbial proteins. However, when we compared the protein IDs between BB_matrixE and GH_BCE methods, we observed that 81% of the proteins were identified using both methods (Supplementary Fig. S3) and the majority of proteins exclusively identified using GH_BCE (73%) were lowly abundant (<2 PSMs; Supplementary Fig. S3). For maize, the GH_BCE method also performed best, but the bacteria to plant ratio was, on average, only 5% (Fig. 2), which indicates that these methods for maize needed further optimization. Supplementary Table S3 contains the number of proteins, peptides, and PSMs derived from the plant (*Arabidopsis* and maize).

Optimization of extraction methods for maize root–associated microbiota.

While, in our first round of tests, we found good methods for *Arabidopsis* root metaproteome extraction, metaproteomes from maize roots were of much lower quality (Figs. 1 and 2), indicating the protein extraction methods needed to be further optimized for maize. Since the best method for maize root metaproteomics involved homogenization using the DUALL glass homogenizer (GH_PBS) followed by differential centrifugation, we decided to optimize this method by increasing the speed and reducing the duration of the first centrifugation step (discussed below). In addition, we adapted a method that was previously employed to obtain viable cells from maize roots (Niu et al. 2017). This method employs vortexing of maize root fragments with glass beads in order to mildly disrupt root surface tissues and dislodge bacterial cells from root surfaces (Fig. 3A). To potentially enhance bacterial cell lysis in the vortexing method, we also tested the addition of an ultrasonication step after the vortexing step in a separate experiment (Fig. 3B). In this experiment, we also extracted the roots again with the BB_matrixE method for comparison. The complete list of proteins identified in this experiment can be found in Supplementary Table S4. Reproducibility of these methods are shown in Supplementary Table S5.

As shown in Figure 3A, the vortexing method improved bacterial protein identification fourfold and the bacteria to plant protein abundance ratio 13-fold, as compared with BB_matrixE and the GH_PBS methods. The addition of ultrasonication to the vortexing method did not significantly enhance the overall number of bacterial protein identifications, although a small trend towards higher identification numbers was observed (Fig. 3B). The difference in the number of proteins associated with the BB_matrixE method between the first experiment (Fig. 1D) and the second (Fig. 3A) is due to experimental batch effect. These two sets of experiments correspond to different plants, extractions, and mass spectrometry analyses performed months apart. The complete list of proteins from this experiment is described in Supplementary Table S4.

Differences between extraction methods in the physicochemical properties of identified bacterial proteins.

To evaluate if different extraction methods were biased against proteins with specific properties, we analyzed the impact of four protein properties, including molecular mass (MM), isoelectric point (pI), GRAVY score (negative score = protein is overall hydrophilic, positive score = hydrophobic), and number of transmembrane helices (TMHMM) on the likelihood of a protein being identified (Supplementary Fig. S4). All *Arabidopsis* methods were compared, except the N2_PBS and FT_PBS methods, because they showed lower reproducibility (Supplementary Fig. S1). For maize, only the methods from the second set of optimization experiments were compared. Overall, none of the

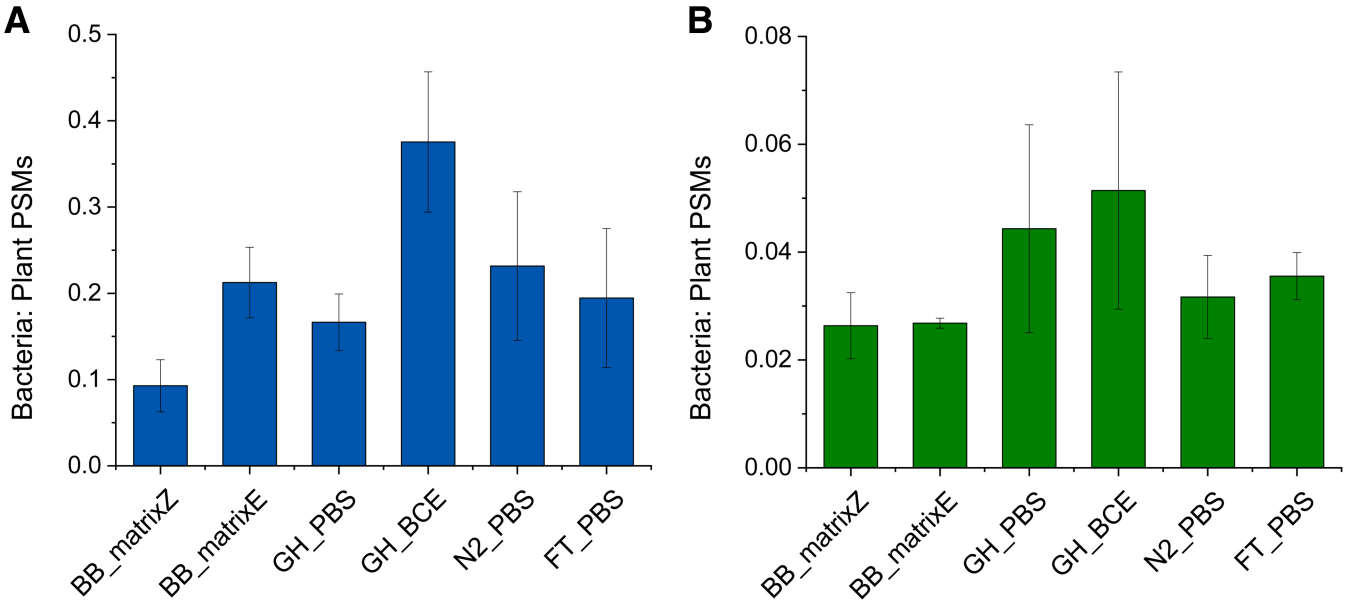

Fig. 2. Some extraction methods strongly favor extraction of bacterial proteins over plant proteins. **A,** Bacteria to plant protein ratios were calculated for *Arabidopsis* and **B,** maize datasets. Peptide-spectrum matches (PSMs) were summed separately for the plant and the bacteria, as a measure of total protein abundance for each organism type, and were used to calculate ratios. The data are expressed as means with error bars representing standard deviations.

methods were strongly biased for or against proteins with specific properties and the distribution profiles of both plant species were very similar. However, we observed some small but significant differences between methods (analysis of variance [ANOVA] followed by Tukey's honestly significant difference (HSD) test, false discovery rate [FDR] < 0.05). For example, the vortexing method used in maize showed a small enrichment of proteins in the 0- to 40-kDa range as compared with the BB_matrixE and GH_PBS methods (Supplementary Fig. S4).

Effect of extraction methods on specific bacterial species.

Efficient lysis of heterogeneous types of microbial cells is a major challenge in meta-"omics" studies, for example, due to species differences in cell-wall composition. To evaluate whether the tested methods differentially affected protein extraction from different species, we compared the metaproteomics-based relative abundances of species across extraction methods (Fig. 4). We paid particular attention to differences between gram-positive and gram-negative bacteria due to differences in their cell-wall structures.

For the *Arabidopsis*-associated species, we observed that the bead-beating methods (BB_matrixZ and BB_matrixE) provided significantly higher protein abundances for the gram-positive bacterium *Arthrobacter* sp. (CL028), which has a thick wall (Fig. 4A). The relative abundance of *Arthrobacter* sp. (CL028) was five- and twofold higher for the BB_matrixE method as compared with the glass homogenizer methods (GH_PBS and GH_BCE) and to the BB_matrixZ method, indicating that the matrix E–based protocol significantly enhanced protein extraction from gram-positive bacteria.

For maize-associated species, we compared both vortexing methods, which were the most promising ones in terms of number of identified proteins (Fig. 3). We found that the addition of ultrasonication after the vortexing step significantly increased the total protein abundance for *Herbaspirillum robiniae* (Fig. 4B), while the protein abundance for *Enterobacter ludwigii* was reduced (Fig. 4B). Unexpectedly, the abundance of the only gram-positive bacterium in the maize SynCom—*Curtobacterium pusillum*—did not differ between the two methods (Fig. 4B).

Different protein extraction protocols affect relative abundances of bacterial proteins.

We next evaluated whether the abundances of individual proteins in the *Arabidopsis* samples clustered within or significantly differed between the BB_matrixZ, BB_matrixE, and GH_BCE extraction methods. Comparing protein abundances between methods showed that around 44% of all identified proteins significantly differed (ANOVA, FDR < 0.05) in abundance across the three methods, with -about 77% of these differences involving the BB_matrix Z method (Fig. 5B). We observed the highest fold difference for the oligopeptide transport system substrate-binding protein from the *Arthrobacter* sp. (CL028_83599; Supplementary Table S6). This protein was 169-fold more abundant in BB_matrixE samples compared with GH_BCE samples.

In a two-way hierarchical clustering analysis based on relative abundances of all differentially abundant proteins (Supplementary Table S6), samples clustered according to their extraction method (Fig. 5A), indicating that the extraction methods impacted protein abundance patterns in the metaproteomes. A similar pattern was observed in the hierarchical clustering of all proteins identified by the three methods (Supplementary Fig. S5). The proteins formed three distinct clusters in the hierarchical clustering (Fig. 5A, clusters I to III), using a distance threshold cutoff of 3.69. Cluster I included 745 proteins that had higher relative abundances in both BB_matrixE and GH_BCE methods compared with BB_matrixZ (Fig. 5A). Cluster II contained 493 proteins that represented the proteins that were more abundant when using the BB_matrixE method. A large proportion (72%) (Fig. 5E) of these proteins were from the *Arthrobacter* sp. (gram-positive, CL028) and exclusively detected when using matrix E. And, finally, cluster III included 147 proteins more abundant in BB_matrixZ method (Fig. 5A).

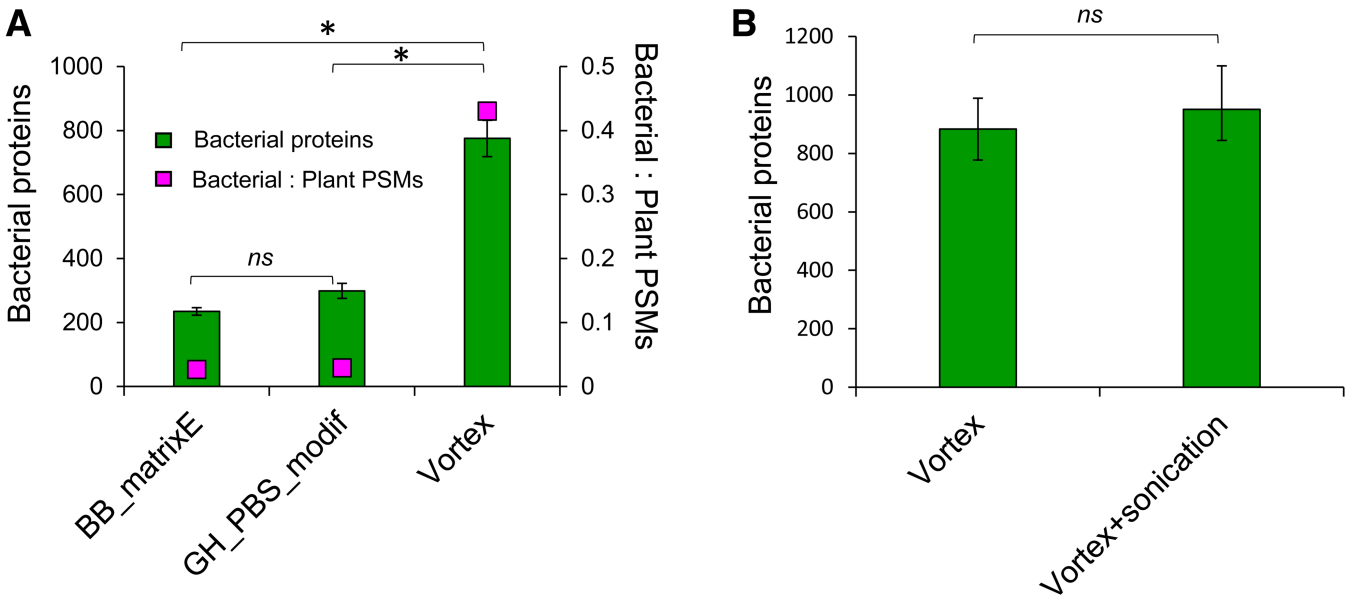

Fig. 3. Optimization of protein extraction protocols for metaproteomics of maize roots. **A,** For each of the three protein extraction methods, the bacterial protein count (green bars) and the ratio of total bacterial versus plant protein abundance are shown, calculated by summing peptide-spectrum matches (PSMs) for each organism type (pink squares). BB_matrixE = bead beating using matrix E; GH_PBS_modif = DUALL glass homogenization in phosphate buffered saline (PBS) protocol, modified for maize roots and vortexing protocol. **B,** Results from a separate experiment comparing the number of bacterial proteins obtained using the vortexing method alone or with the addition of an ultrasonication step. Error bars represent standard deviations, and comparisons between methods were assessed by one-way analysis of variance followed by Tukey's honestly significant difference test (false discovery rate < 0.05). Asterisks represent significant differences between means of pairwise comparisons and *ns* indicates the means were not significantly different.

We analyzed the GRAVY score, MM, and species of origin for all proteins in the three clusters, to identify the main drivers of the differentiation (Fig. 5C to E). While both GRAVY score and MM showed some small differences between clusters, the main difference between clusters was the number of identified *Arthrobacter* sp. proteins. This indicates that the main driver of significant differences between extraction methods was related to the extraction of this gram-positive SynCom species.

DISCUSSION

Different plant types require distinct extraction methods for root metaproteomics.

Our data show that root samples from different plant species require different protein extraction methods for best metaproteomics results (Fig. 1). We found that the BB_matrixE method, which worked well for the *Arabidopsis* samples, performed poorly for the maize samples, as compared with the other methods in the initial experiment (Fig. 1). Ultimately, for maize, the vortexing method, which required an extra development effort, improved protein and peptide identification numbers. We can think of three reasons that might explain these differences in extraction efficiency: i) root anatomy, ii) cell-wall composition, and iii) microbial load and localization. First, the root tissue anatomy of *Arabidopsis* and maize are very distinct, with the maize tissue being much denser, as it is composed of 10 to 15 cortical cell layers compared with just one layer in *Arabidopsis* (Hochholdinger and Zimmermann 2008). Denser tissues can potentially yield more plant proteins compared with lower density tissues. This would lead to higher relative abundances of plant to microbial cells in maize root samples, which would increase the mass spectrometer measurement time spent on plant peptides and make it more difficult to identify microbial proteins. This explanation is in line with the lower ratio of bacterial to plant PSMs for maize roots (highest bacteria/plant PSM ratio of 0.05) (Fig. 2B), as compared with *Arabidopsis* roots (highest bacteria/plant PSM ratio of 0.37) (Fig. 2A). Second, plant cell-wall composition might affect the efficiency of mechanical disruption, thereby hampering the release of bacteria, particularly if they were endophytes. For example, suberin and lignin are cell-wall polymers that provide a rigid protection to inner tissues. The suberin amount in *Arabidopsis* roots is approximately three to 15 times lower than the amount found in monocot roots (Kreszies et al. 2018). This difference in tissue rigidity affected, for example, the ability of matrix E to disrupt the tissue. Maize roots remained almost intact when bead-beaten with matrix E, while *Arabidopsis* roots were fully disrupted, possibly releasing more endophytes. The caveat for this explanation is that, currently, the localization of the members of both the *Arabidopsis* and maize SynComs within or on the root has not been characterized and we rely on limited observations to speculate on root localization of the bacteria. Third, microbial load and localization within the roots might be another reason for the differences found between species. We, for example, observed that maize roots stayed intact with the vortexing method, indicating that the higher bacterial protein identification rate of this method was likely due to bacteria being primarily localized on the surface or in the outer layers of the endodermis of the root. Additionally, by keeping most of the root intact, less plant protein was likely released, leading to an increased bacteria to plant protein ratio in the extracts. The vortexing method enhanced the microbial protein identification rate in maize by at least fourfold compared with other methods tested (Fig. 3A). On the other hand, we observed that the matrix E method, which performed best in terms of identified bacterial proteins from *Arabidopsis* roots, completely disintegrated the root material, suggesting that the bacteria associated with *Arabidopsis* roots might be more widely distributed within the root tissues (endosphere and rhizoplane). While this last explanation of differences in extraction method performance for different plant species is quite compelling,

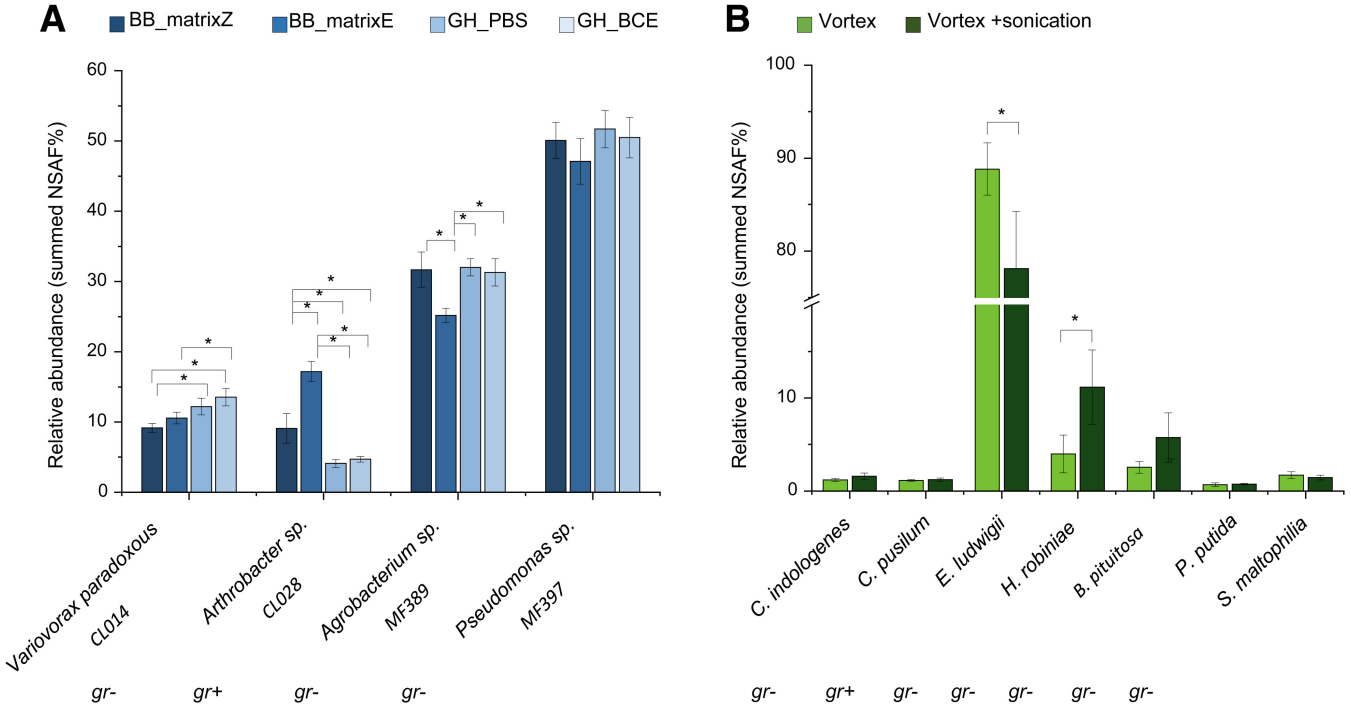

Fig. 4. A, Extraction of proteins from different bacterial species in *Arabidopsis* and **B,** in maize. The data are expressed as summed protein abundance (normalized spectral abundance factors multiplied by 100 [NSAF%]) means of all proteins for each strain in the SynCom. Asterisks and brackets show pairwise comparisons that were statistically significantly different (analysis of variance followed by Tukey's honestly significant difference test, false discovery rate < 0.05).

further research is needed to characterize microbial cell densities and distributions in *Arabidopsis* and maize roots.

Since the vortexing method resulted in the best results in maize, we added an ultrasonication step to the vortexed suspension to further optimize the extraction of proteins from gram-positive bacteria. Ultrasonication has been shown to improve protein extraction from gram-positive strains (Zhang et al. 2018b). The only gram-positive strain in the maize SynCom–*Curtobacterium pusillum*–did not, however, show any abundance changes with the addition of ultrasonication (Fig. 4B). The reason for this might be low relative abundance of this strain compared with other microbial members of the SynCom in maize roots, as observed in a previous study (Niu et al. 2017). On the other hand, the relative abundances of two gram-negative strains (*Enterobacter ludwigii* and *Herbaspirillum robiniae*) significantly changed when the ultrasonication step was added to the

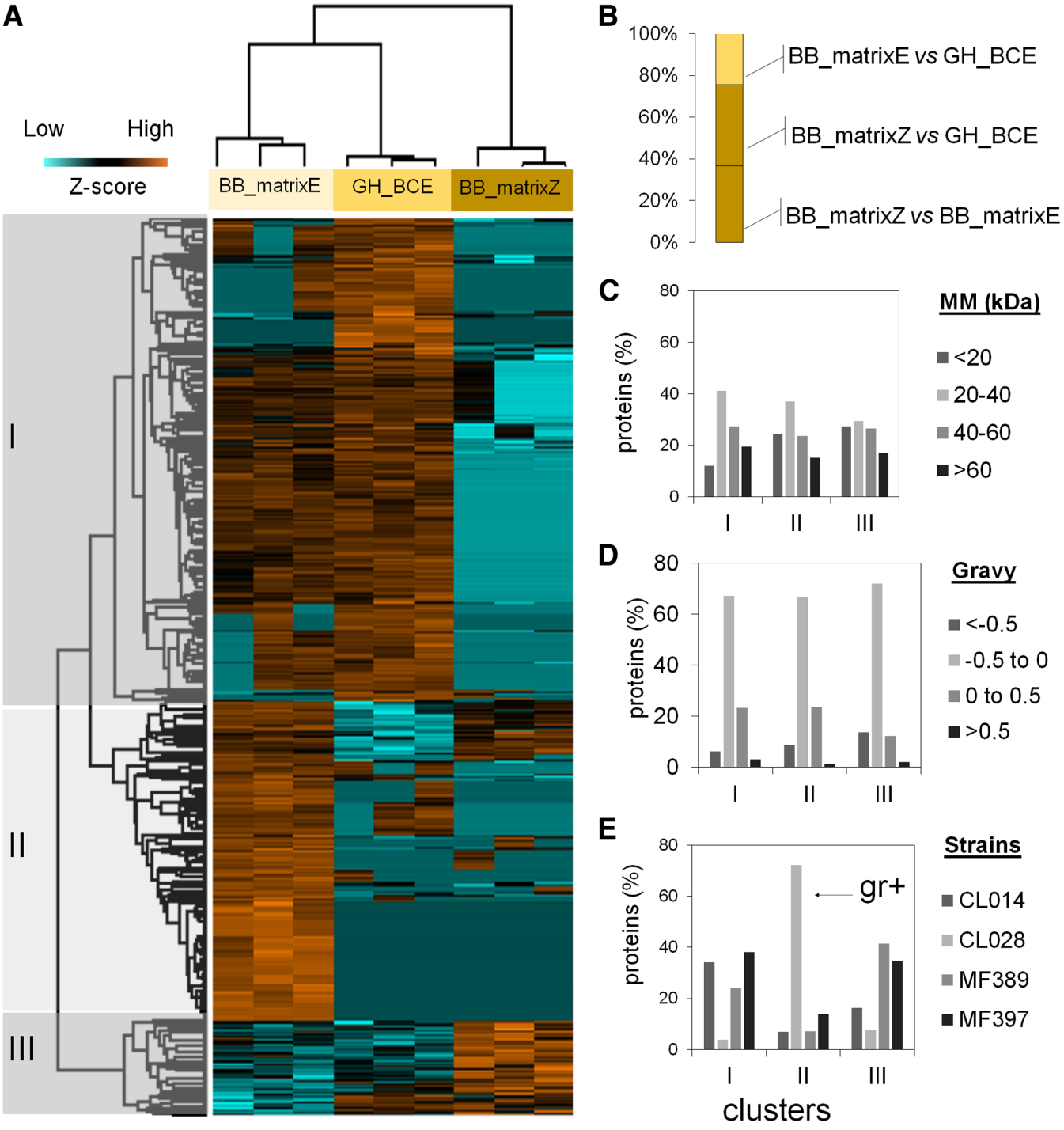

Fig. 5. Extraction of a gram-positive bacterium is the main factor driving differentiation of extraction methods for *Arabidopsis*. **A,** Two-way hierarchical clustering (Euclidean, average linkage) of log-transformed and *z*-score normalized spectral abundance factors of differentially abundant proteins (rows; analysis of variance, Tukey's honestly significant difference test, false discovery rate < 0.05) between protein extraction methods (columns). **B,** Percentage of differentially abundant proteins per pairwise comparison of protein extraction methods. **C,** Molecular mass (MM) distribution of proteins from each cluster identified in A. **D,** GRAVY score distribution of proteins from each cluster. **E,** Protein distribution of strains (CL014 = *Variovorax paradoxus*, CL028 = *Arthrobacter* sp., MF389 = *Agrobacterium* sp., and MF397 = *Pseudomonas* sp.) in each cluster.

vortexing protocol. We currently have no good explanation for these results, but the trend, although not statistically significant, indicated that ultrasonication may improve overall numbers.

MATERIALS AND METHODS

SynCom culture for *Arabidopsis* inoculation.

The four-member SynCom used to inoculate *Arabidopsis thaliana* Col-0 seedlings contained genome-sequenced isolates of *Variovorax paradoxus* CL14, *Arthrobacter* sp. UNCCL28, *Agrobacterium* sp. 719_389, and *Pseudomonas* sp. 397MF (Levy et al. 2018). Each strain was grown from freezer stocks, separately, overnight in King's B medium. Then, media was washed off by centrifugation at 8,000 × *g* and culture pellets were resuspended in 10 μM $MgCl_2$. Optical density at 600 nm (OD_{600}) for each culture was measured and adjusted to an OD_{600} of 0.01. From these suspensions, equal volumes of each culture were mixed to form the SynCom. This SynCom (100 μl) was spread onto 12 × 12 cm agar plates with ½ strength Murashige-Skoog media (without sucrose).

SynCom culture for maize inoculation.

The seven-member SynCom used to inoculate maize (*Zea mays* cv. Sugar Buns [Johnny's Selected Seeds]) plants contained genome-sequenced isolates of *Stenotrophomonas maltophilia*, *Brucella pituitosa* (previously *Ochrobacterium pituitosum*), *Curtobacterium pusillum*, *Enterobacter ludwigii* (previously *Enterobacter cloacae* [Hördt et al. 2020]), *Chryseobacterium indologenes*, *Herbaspirillum robiniae* (previously *Herbaspirillum frisingense*), and *Pseudomonas putida* (Niu et al. 2017). For cultivation of these strains, we followed the previously published protocol by Niu and Kolter (2018). Briefly, we streaked each bacterial species from frozen stocks separately onto 0.1× TSA (tryptic soy agar) plates and incubated the plates at 30°C for 48 h. Then, we inoculated one colony of each species separately into 5 ml of tryptic soy broth (TSB) and let them grow under 180 rpm agitation at 30°C for 14 h. After that, we transferred 100 μl of each culture to 250 ml of fresh TSB media and let them grow overnight in the same conditions. We recovered the bacterial cells by centrifuging the cultures at 8,000 × *g*, at 4°C for 10 min. Bacterial pellets were resuspended in 20 ml of phosphate buffered saline (PBS), pH 7.4, and the OD_{600} of resuspended pellets was measured, using a spectrophotometer (Ultraspec 10-cell density meter; Biochrom). Based on the spectrophotometer readings, we calculated the cells per milliliter for each strain. Then, we diluted each resuspended culture in phosphate-buffered saline (PBS) to obtain 10^8 cells per milliliter according to the specific OD_{600} (Supplementary Table S7). After that, the suspensions from the seven species were mixed in equal parts to produce the inoculum. Finally, the inoculum containing the seven strains was mixed into ½ strength Murashige-Skoog basal solution to obtain a 10^7–cells per milliliter solution.

Plant growth conditions and inoculation.

For *Arabidopsis*, we surface-sterilized seeds with a mixture of 70% (vol/vol) bleach and 0.2% (vol/vol) Tween-20 for 8 mins to eliminate any seed-borne microorganisms on the seed surface. We rinsed the seeds three times with sterile distilled water. Seeds were stratified at 4°C in the dark for 2 days. Plant germination and inoculation procedures were performed as previously described (Finkel et al. 2020). Sterilized seeds were germinated and grown on vertical square 12 × 12 cm agar plates containing 1/2 strength Murashige-Skoog media supplemented with 0.5% sucrose for 7 days. Then, we transferred 10 plants to new plates containing ½ strength Murashige-Skoog media (without sucrose) inoculated with the SynCom. Five replicates (plates) were used. Plates were randomized in a growth chamber under a 16-h-light and 8-h-dark regime at 21 and 18ºC, respectively, for 12 days.

For maize, we surface-sterilized seeds with 70% (vol/vol) ethanol for 3 min, then with 2% sodium hypochlorite for 3 min, and then, rinsed the seeds five times with sterile water. We deposited seeds on the bottom of sterile bags (Nasco) and added 150 ml of autoclaved, calcinated clay (Pro's Choice Rapid Dry, Oil-Dri Corporation). After that, we added 120 ml of ½ strength Murashige-Skoog media containing 10^7 cells per milliliter to the bag. Plants in the bags were randomized in a growth chamber under a 16-h light and 8-h-dark regime at 27 and23°C, respectively, for 13 days.

Sample collection.

We collected the *Arabidopsis* roots from agar plates using a sterile razor blade and tweezers. Before freezing the samples, we pooled together *Arabidopsis* roots from different plants, resulting in a master mix to avoid the introduction of biological variability in the method tests. The master mix was flash-frozen in liquid nitrogen and was stored at –80°C.

For maize, we first gently washed the roots with distilled water, to remove clay particles, and then cut them into 1-cm pieces. We also produced a master mix of maize roots to be used with all protein extraction methods and stored at –80°C. The only exception was for the vortexing method, for which we used fresh tissues (described below).

Protein extraction protocols.

We performed all extractions in triplicate, using the same pool of root samples. Around 20 mg of root tissue were used for *Arabidopsis* and 70 mg for maize protein extractions.

Bead-beating using matrix Z in SDS buffer (BB_matrixZ).

We added 300 μl of SDT lysis buffer (4% SDS, 100 mM Tris-HCl, 0.1 M dithiothreitol) to the frozen roots. Then, we performed cell disruption by bead-beating (Bead Ruptor Elite; OMNI) the samples with lysing matrix Z (MP Biomedicals), using two cycles of 45 s at 8.0 m/s. Matrix Z is composed of 2-mm diameter zirconium oxide beads and is recommended for tougher tissues as roots. After that, we incubated the samples at 95°C for 10 min and centrifuged them at 14,000 × *g* for 5 min. Supernatant was loaded four times (4 × 60 μl) on centrifugal filters (10 kDa molecular weight cut-off [MWCO]) (VWR International) to perform the FASP protocol.

Bead-beating using matrix E in SDS buffer (BB_matrixE).

We added 300 μl of SDT lysis buffer to the frozen roots. We performed cell disruption by bead-beating (Bead Ruptor Elite; OMNI) the samples, using lysing matrix E (MP Biomedicals), using one cycle of 30 s at 4.0 m/s. Matrix E is composed of 1.4-mm ceramic spheres, 0.1-mm silica spheres, and 4-mm glass beads and is recommended for use with microbe infected tissues and environmental samples. After that, we incubated the samples at 95°C for 10 min and centrifuged the samples at 5,000 × *g* for 5 min and supernatant was loaded twice (2 × 60 μl) on centrifugal filters (10 kDa MWCO, VWR International) to perform the FASP protocol.

DUALL glass homogenizer disruption and differential centrifugation (GH_PBS).

We added 1 ml of PBS (pH 7.4) with proteinase inhibitors (cOmplete, Roche) to the frozen roots and used a DUALL glass homogenizer to disrupt the root tissue. We transferred the suspension to 1.7-ml tubes and centrifuged them at 500 × *g* for 5 min at 4°C. We collected the supernatant and centrifuged at 21,000 × *g* for 5 min at 4°C. The resulting pellet was resuspended in SDT lysis buffer (1:10 ratio). We incubated samples

at 95°C for 10 min. Finally, we loaded the sample (50 µl) onto centrifugal filters (10 kDa MWCO, VWR International) to perform the FASP protocol.

DUALL glass homogenizer disruption and differential centrifugation (GH_BCE).

We added 1 ml of BCE containing 50 mM Tris HCl, pH 7.5, 0.3% Triton X-100, and 2 mM 2-β-mercaptoethanol (Ikeda et al. 2009) to the frozen roots and used a DUALL glass homogenizer to disrupt the root tissue. We transferred the suspension to 1.7-ml tubes and centrifuged them at 500 × *g* for 5 min at 4°C. We collected the supernatant and centrifuged at 21,000 × *g* for 5 min at 4°C. The resulting pellet was resuspended in SDT lysis buffer (1:10 ratio). We incubated samples at 95°C for 10 min. Finally, we loaded the sample (50 µl) on centrifugal filters (10 kDa MWCO, VWR International) to perform the FASP protocol.

Nitrogen grinding protocol (N2_PBS).

We ground the frozen roots in liquid nitrogen using a mortar and pestle. Then, we transferred the pulverized tissue to 1.7 ml tubes filled with 1 ml of PBS with proteinase inhibitors (cOmplete). After that, we performed the same steps as those for protocol GH_PBS, starting at the first centrifugation.

Freeze-thaw cycles protocol (FT_PBS).

We added 1 ml of PBS with proteinase inhibitors (cOmplete) to the frozen roots and disrupted the tissue using a DUALL glass homogenizer. We transferred the suspension to 1.7-ml tubes and centrifuged them at 500 × *g* for 5 min at 4°C. We collected the supernatant and centrifuged it at 21,000 × *g* for 5 min at 4°C. The resulting pellet was resuspended in SDT lysis buffer (1:10 ratio). We transferred the supernatant to 2-ml tubes containing lysing matrix Z (MP Biomedicals) and incubated them at 95°C for 10 min. After that, we disrupted cells by bead-beating (Bead Ruptor Elite; OMNI) the samples for two cycles of 45 s at 6 m/s. Then, we performed two cycles of freeze-thawing: incubation at –80°C for 10 min, proceeded by incubation at 95°C for 10 min, then bead-beating (same as previous) and repeating the steps of freeze and heat incubation. Finally, we centrifuged samples at 14,000 × *g* for 5 min at 4°C and supernatants were collected to load (50 µl) onto centrifugation filters (10 kDa MWCO, VWR International) for the FASP protocol.

Modified DUALL glass homogenizer + differential centrifugation (only for maize samples, GH_PBS_modif).

This method was the same as the GH_PBS method described above but with a modified centrifugation step. The first centrifugation was performed at 1,000 × *g* for 2 min. All the other steps were the same.

Vortexing method (only for maize roots).

Six 1-in pieces of fresh maize primary root were suspended in 1 ml of PBS and were vortexed by hand (on a Genie 2 mixer at speed 8) with six 3-mm sterile glass beads. The roots were vortexed for three cycles of 1 min with a 10-s dwell between cycles. Samples were kept on ice between cycles. The root slurry was collected and centrifuged at 15,000 × *g* and 4°C for 7 min. Supernatant was discarded and the pellet was resuspended in 120 µl of SDT lysis buffer and was heated to 95°C for 10 min. The cell lysis was mixed with an 800-µl UA solution (8 M urea in 0.1 M Tris HCl, pH 8.5), was loaded twice on 10-kDa centrifugation filters (VWR International), and was subjected to standard FASP protocol.

Vortexing + ultrasonication (only for maize roots).

Maize roots were processed as described above for the vortexing method, except we added an ultrasonication in SDT lysis buffer step before boiling the samples. We ultrasonicated samples using a single microtip, in three pulses of 30 s with 1 min off between pulses, at 10% amplitude (Qsonica Q700). Samples were kept on ice during sonication. After that, we boiled samples at 95°C and proceeded with the FASP protocol.

FASP.

We prepared the peptide samples following the FASP protocol described previously (Wiśniewski et al. 2009) with some modifications (Kleiner et al. 2017). We mixed 60 µl of protein suspension into 400 µl of UA solution on centrifugal filters (10 kDa MWCO, VWR International). Then, we centrifuged samples at 14,000 × *g* for 30 min. After loading, we added 200 µl of UA solution to the filter and centrifuged again at 14,000 × *g* for 20 min. After that, we added 100 µl of indole acetic acid (IAA) solution (0.05 M iodoacetamide in UA solution) to the filter, followed by incubation at room temperature for 20 min. We removed the IAA solution by centrifugation and we washed the filter three times, by adding 100 µl of UA solution followed by centrifugation. After washing, we changed the buffer to ABC solution (50 mM ammonium bicarbonate) by washing the filter three times with 100 µl of ABC and centrifuging. For protein digestion, we added 1 µg of MS-grade trypsin (Thermo Scientific Pierce) in 40 µl of ABC solution in the filters. Samples were placed inside a sealed plastic box with wet paper towels, to maintain humidity, and the box was placed in an incubator at 37°C for overnight incubation. The next day, we eluted peptides by centrifuging the samples at 14,000 × *g* for 20 min and adding 50 µl of 0.5 M NaCl and further centrifugating at 14,000 × *g* for 30 min. Finally, we measured peptide concentrations using the Pierce Micro BCA assay (Thermo Scientific Pierce), following manufacturer instructions.

LC-MS/MS.

We analyzed the peptide samples generated in the first experiment with the six extraction methods (BB_matrixZ, BB_matrixE, GH_PBS, GH_BCE, N2_PBS, and FT_PBS) using the same LC-MS/MS method and in a single run per plant species. Samples were block-randomized in the LC-MS/MS run sequence to minimize LC-MS/MS batch effects, as recommended by Oberg and Vitek (2009), and one wash run with 100% acetonitrile was added between samples. The LC-MS/MS method was similar to the one described previously (Kleiner et al. 2017). For each run 1.5 µg of peptides were loaded onto a 5-mm, 300-µm ID C18 Acclaim PepMap 100 pre-column (Thermo Fisher Scientific) using an UltiMate 3000 RSLCnano liquid chromatography (Thermo Fisher Scientific), using 100% loading solvent A (98% water, 2% acetonitrile, 0.05% triflouroacetic acid). After loading, peptides were eluted onto a 75 cm × 75 µm analytical EASY-Spray column packed with PepMap RSLC C18 2-µm material (Thermo Fisher Scientific) and were heated to 60°C. Peptides were separated on the analytical column at a flow rate of 300 nl/min, using a 460-min gradient of eluent A (0.1% formic acid in water) and eluent B (80% acetonitrile, 0.1% formic acid). The gradient performed was as follows: from 5% B to 31% B in 364 min, in 76 min up to 50% B, 20 min at 99% B. Eluting peptides were ionized with electrospray ionization using the Easy-Spray source (Thermo Fisher Scientific). Mass spectra were acquired in a Q Exactive HF or HF-X hybrid quadrupole-Orbitrap mass spectrometer (Thermo Fisher Scientific). Full scans within the range of 380 to 1600 *m*/*z* were acquired in the Orbitrap at 60,000 with the maximum injection time set to 200 ms. MS/MS scans of the 15 most abundant precursor ions were acquired at 7,500 (HF-X) or 15,000 (HF) resolution with maximum injection time of 150 ms. The mass (*m*/*z*) 445.12003 was used as lock mass. Normalized collision

energy was set to 24. Ions with charge state +1 were excluded from MS/MS analysis. Dynamic exclusion was set to 20 s.

The second batch of tests that included only maize samples extracted with the BB_matrixE, GH_PBS_modif, and vortexing methods were analyzed using a 140-min method gradient on the same LC-MS/MS instrumentation and columns. For this method, we loaded 600 ng of peptides for each run as described above, and peptides were separated on the analytical column at a flow rate of 300 nl/min, using a 140-min gradient as follows: from 5% B to 31% B in 117 min, in 4 min up to 50% B, 20 min at 99% B. Ionized peptides were analyzed in the Q Exactive HF hybrid quadrupole-Orbitrap mass spectrometer (Thermo Fisher Scientific). Full scans within the range of 380 to 1,600 m/z were acquired in the Orbitrap at 60,000, and maximum injection time was set to 200 ms. MS/MS scans of the 15 most abundant precursor ions were acquired at 15,000 resolution with maximum injection time equals to 100 ms. The mass (m/z) 445.12003 was used as lock mass. Normalized collision energy was set to 24. Ions with charge state +1 were excluded from MS/MS analysis. Dynamic exclusion was set to 25 s.

Protein identification.

The database for protein identification from *Arabidopsis* roots consisted of *Arabidopsis* protein sequences downloaded from UNIPROT (UP000006548) and the bacterial protein sequences of the four species used in the SynCom downloaded from the Joint Genome Institute IMG database under the IMG genome identification numbers 2643221508 (*Variovorax paradoxus*), 2593339130 (*Arthrobacter* sp.), 2521172559 (*Agrobacterium* sp.), and 2643221503 (*Pseudomonas* sp.). We added the cRAP protein sequence database containing protein sequences of common laboratory contaminants. Similarly, the maize database consisted of *Zea mays* protein sequences downloaded from UNIPROT (UP000007305), the cRAP database, and the bacterial protein sequences of the seven species used in the SynCom downloaded from GenBank under accession numbers CP001918 (*Enterobacter ludwigii*, previously *Enterobacter cloacae*), CP018756 (*Stenotrophomonas maltophilia*), CP018779 (*Brucella pituitosa*, previously *Ochrobactrum pituitosum*), CP018783 (*Curtobacterium pusillum*), CP018786 (*Chryseobacterium indologenes*), CP018845 (*Herbaspirillum robiniae*, previously *H. frisingense*) and CP018846 (*Pseudomonas putida*). The final *Arabidopsis* and maize databases contained 61,313 and 133,246 protein sequences, respectively. These databases are available at the PRIDE repository (PXD026330 and PXD026369, respectively).

For protein identification, we searched the MS raw files against the databases using the Sequest HT search engine embedded in the Proteome Discoverer version 2.2 software (Thermo Fisher Scientific), with the following parameters: trypsin (Full), maximum two missed cleavages, 10 ppm precursor mass tolerance, 0.1 Da fragment mass tolerance, and maximum three equal dynamic modifications per peptide. We considered the following dynamic modifications: oxidation on M (+15.995 Da), carbamidomethyl on C (+57.021 Da), and acetyl on the protein N terminus (+42.011 Da). For peptide FDR calculation, we used the Percolator node in Proteome Discoverer with the following parameters: maximum Delta Cn 0.05, a strict target FDR of 0.01, a relaxed target FDR of 0.05, and validation based on q-value. Only proteins identified with medium or high confidence were retained. Multiconsensus tables were exported and used for subsequent analysis. For protein quantification, we calculated the normalized spectral abundance factors (NSAFs) (Florens et al. 2006). The NSAF multiplied by 100 (NSAF%) gives the relative abundance of a protein within a sample as a percentage.

Protein chemical properties.

We retrieved the GRAVY score of proteins using the GRAVY calculator and transmembrane helices were predicted using the TMHMM server version 2.0. We obtained the protein MM and pI from Proteome Discoverer (version 2.2) output.

Statistical analyses and hierarchical clustering.

We conducted all statistical analyses in Perseus software (version 1.6.14.0) (Tyanova et al. 2016). Missing values were imputed using a constant that was lower than the lowest NSAF value within each treatment. Statistical significance was assessed by ANOVA, using permutation-based FDR correction (FDR < 5%) followed by posthoc Tukey's HSD test (FDR < 5%) to compare multiple groups.

For hierarchical clustering analysis, we included only the differentially abundant proteins detected in at least all replicates of one group. We first log-transformed the NSAF% values followed by z-score normalization. After that, we proceeded with the two-dimensional hierarchical clustering on the z-scored normalized data using the Euclidean distance and average linkage.

Conclusion.

Our study showed that sample preparation methods for metaproteomics of root-associated microbes cannot be generalized for all plant species and that protein extraction protocols must be optimized for each species to achieve the best results. Here, we have identified methods that enable metaproteomic experiments with low input material (40 to 70 mg) for *Arabidopsis* (BB_matrixE method) and maize (vortexing method) roots. These methods yielded the highest microbial protein numbers for these species and were much less labor-intensive compared with other methods tested. In addition, the BB_matrixE method clearly performed better for the gram-positive bacterium included in the SynCom. Based on our observations, we speculate that plant tissue characteristics and microbial localization were the main drivers of protein extraction success and that finding the right mechanical disruption method was essential to optimize microbial protein identification. Thus, we recommend the BB_matrixE method for soft root tissues and the vortexing method for rigid root tissues with bacteria localized superficially. Our results do not allow us to determine which root characteristic (dicot versus monocot architecture, bacterial localization, or something else) was responsible for the success of the two identified methods in the different plant species. However, since these two methods were very different from each other, they represent ideal starting points for optimizing metaproteomics for root-associated microorganisms in other plant species. We thus recommend including both methods in any testing design for other plant species.

Data availability.

The mass spectrometry proteomics data and protein sequence database have been deposited to the ProteomeXchange Consortium via the PRIDE (Vizcaíno et al. 2010) partner repository with the dataset identifier PXD026330 for the *Arabidopsis* data and PXD026369 for the maize data.

AUTHOR-RECOMMENDED INTERNET RESOURCES

cRAP protein sequence database: http://www.thegpm.org/crap
GRAVY calculator: http://www.GRAVY-calculator.de
Joint Genome Institute IMG database:
https://img.jgi.doe.gov/cgi-bin/m/main.cgi
TMHMM server version 2.0:
http://www.cbs.dtu.dk/services/TMHMM-2.0

ACKNOWLEDGMENTS

We thank B. Niu and R. Kolter for providing the maize SynCom strains. We thank three anonymous reviewers and B. Day for thoughtful feedback that improved the quality of this manuscript. The LC-MS/MS measurements were made using equipment in the Molecular Education, Technology, and Research Innovation Center (METRIC) at North Carolina State University (NCSU). Thanks to the NCSU Phytotron installations team for allowing the cultivation of maize plants.

LITERATURE CITED

Bao, Z., Okubo, T., Kubota, K., Kasahara, Y., Tsurumaru, H., Anda, M., Ikeda, S., and Minamisawa, K. 2014. Metaproteomic identification of diazotrophic methanotrophs and their localization in root tissues of field-grown rice plants. Appl. Environ. Microbiol. 80:5043-5052.

Barret, M., Briand, M., Bonneau, S., Préveaux, A., Valière, S., Bouchez, O., Hunault, G., Simoneau, P., and Jacques, M.-A. 2015. Emergence shapes the structure of the seed microbiota. Appl. Environ. Microbiol. 81: 1257-1266.

Blakeley-Ruiz, J. A., Erickson, A. R., Cantarel, B. L., Xiong, W., Adams, R., Jansson, J. K., Fraser, C. M., and Hettich, R. L. 2019. Metaproteomics reveals persistent and phylum-redundant metabolic functional stability in adult human gut microbiomes of Crohn's remission patients despite temporal variations in microbial taxa, genomes, and proteomes. Microbiome 7:18.

Bossche, T. V. D., Kunath, B. J., Schallert, K., Schäpe, S. S., Abraham, P. E., Armengaud, J., Arntzen, M. Ø., Bassignani, A., Benndorf, D., Fuchs, S., Giannone, R. J., Griffin, T. J., Hagen, L. H., Halder, R., Henry, C., Hettich, R. L., Heyer, R., Jagtap, P., Jehmlich, N., Jensen, M., Juste, C., Kleiner, M., Langella, O., Lehmann, T., Leith, E., May, P., Mesuere, B., Miotello, G., Peters, S. L., Pible, O., Reichl, U., Renard, B. Y., Schiebenhoefer, H., Scryba, A., Tanca, A., Trappe, K., Trezzi, J.-P., Uzzau, S., Verschaffelt, P., von Bergen, M., Wilmes, P., Wolf, M., Martens, L., and Muth, T. 2021. Critical assessment of metaproteome investigation (CAMPI): A multi-lab comparison of established workflows. bioRxiv: 2021.03.05.433915.

Broberg, M., Doonan, J., Mundt, F., Denman, S., and McDonald, J. E. 2018. Integrated multi-omic analysis of host-microbiota interactions in acute oak decline. Microbiome 6:21.

Castrillo, G., Teixeira, P. J. P. L., Paredes, S. H., Law, T. F., de Lorenzo, L., Feltcher, M. E., Finkel, O. M., Breakfield, N. W., Mieczkowski, P., Jones, C. D., Paz-Ares, J., and Dangl, J. L. 2017. Root microbiota drive direct integration of phosphate stress and immunity. Nature 543: 513-518.

Cohen, N. R., McIlvin, M. R., Moran, D. M., Held, N. A., Saunders, J. K., Hawco, N. J., Brosnahan, M., DiTullio, G. R., Lamborg, C., McCrow, J. P., Dupont, C. L., Allen, A. E., and Saito, M. A. 2021. Dinoflagellates alter their carbon and nutrient metabolic strategies across environmental gradients in the central Pacific Ocean. Nat. Microbiol. 6:173-186.

Delmotte, N., Knief, C., Chaffron, S., Innerebner, G., Roschitzki, B., Schlapbach, R., von Mering, C., and Vorholt, J. A. 2009. Community proteogenomics reveals insights into the physiology of phyllosphere bacteria. Proc. Natl. Acad. Sci. U.S.A. 106:16428-16433.

Finkel, O. M., Salas-González, I., Castrillo, G., Conway, J. M., Law, T. F., Teixeira, P. J. P. L., Wilson, E. D., Fitzpatrick, C. R., Jones, C. D., and Dangl, J. L. 2020. A single bacterial genus maintains root growth in a complex microbiome. Nature 587:103-108.

Florens, L., Carozza, M. J., Swanson, S. K., Fournier, M., Coleman, M. K., Workman, J. L., and Washburn, M. P. 2006. Analyzing chromatin remodeling complexes using shotgun proteomics and normalized spectral abundance factors. Methods 40:303-311.

Hamonts, K., Trivedi, P., Garg, A., Janitz, C., Grinyer, J., Holford, P., Botha, F. C., Anderson, I. C., and Singh, B. K. 2018. Field study reveals core plant microbiota and relative importance of their drivers. Environ. Microbiol. 20:124-140.

Hassani, M. A., Durán, P., and Hacquard, S. 2018. Microbial interactions within the plant holobiont. Microbiome 6:58.

Hochholdinger, F., and Zimmermann, R. 2008. Conserved and diverse mechanisms in root development. Curr. Opin. Plant Biol. 11:70-74.

Hördt, A., López, M. G., Meier-Kolthoff, J. P., Schleuning, M., Weinhold, L.-M., Tindall, B. J., Gronow, S., Kyrpides, N. C., Woyke, T., and Göker, M. 2020. Analysis of 1,000+ type-strain genomes substantially improves taxonomic classification of alphaproteobacteria. Front. Microbiol. 11:468.

Houston, K., Tucker, M. R., Chowdhury, J., Shirley, N., and Little, A. 2016. The plant cell wall: a complex and dynamic structure as revealed by the responses of genes under stress conditions. Front. Plant Sci. 7:984.

Ikeda, S., Kaneko, T., Okubo, T., Rallos, L. E. E., Eda, S., Mitsui, H., Sato, S., Nakamura, Y., Tabata, S., and Minamisawa, K. 2009. Development of a bacterial cell enrichment method and its application to the community analysis in soybean stems. Microb. Ecol. 58:703-714.

Jochum, M. D., McWilliams, K. L., Pierson, E. A., and Jo, Y.-K. 2019. Host-mediated microbiome engineering (HMME) of drought tolerance in the wheat rhizosphere. PLoS One 14:e0225933.

Keiblinger, K. M., Wilhartitz, I. C., Schneider, T., Roschitzki, B., Schmid, E., Eberl, L., Riedel, K., and Zechmeister-Boltenstern, S. 2012. Soil metaproteomics—Comparative evaluation of protein extraction protocols. Soil Biol. Biochem. 54:14-24.

Kleiner, M. 2019. Metaproteomics: Much more than measuring gene expression in microbial communities. mSystems 4:e00115-19.

Kleiner, M., Thorson, E., Sharp, C. E., Dong, X., Liu, D., Li, C., and Strous, M. 2017. Assessing species biomass contributions in microbial communities via metaproteomics. Nat. Commun. 8:1-38.

Knief, C., Delmotte, N., Chaffron, S., Stark, M., Innerebner, G., Wassmann, R., von Mering, C., and Vorholt, J. A. 2012. Metaproteogenomic analysis of microbial communities in the phyllosphere and rhizosphere of rice. ISME J. 6:1378-1390.

Kreszies, T., Schreiber, L., and Ranathunge, K. 2018. Suberized transport barriers in *Arabidopsis*, barley and rice roots: From the model plant to crop species. J. Plant Physiol. 227:75-83.

Kunath, B. J., Minniti, G., Skaugen, M., Hagen, L. H., Vaaje-Kolstad, G., Eijsink, V. G. H., Pope, P. B., and Arntzen, M. Ø. 2019. Metaproteomics: Sample preparation and methodological considerations. Pages 187-215 in: Emerging Sample Treatments in Proteomics, Advances in Experimental Medicine and Biology. J.-L. Capelo-Martínez, ed. Springer International Publishing, Cham, Switzerland.

Levy, A., Salas Gonzalez, I., Mittelviefhaus, M., Clingenpeel, S., Herrera Paredes, S., Miao, J., Wang, K., Devescovi, G., Stillman, K., Monteiro, F., Rangel Alvarez, B., Lundberg, D. S., Lu, T.-Y., Lebeis, S., Jin, Z., McDonald, M., Klein, A. P., Feltcher, M. E., Rio, T. G., Grant, S. R., Doty, S. L., Ley, R. E., Zhao, B., Venturi, V., Pelletier, D. A., Vorholt, J. A., Tringe, S. G., Woyke, T., and Dangl, J. L. 2018. Genomic features of bacterial adaptation to plants. Nat. Genet. 50:138-150.

López-Mondéjar, R., Kostovčík, M., Lladó, S., Carro, L., and García-Fraile, P. 2017. Exploring the plant microbiome through multi-omics approaches. Pages 233-268 in: Probiotics in Agroecosystem. V. Kumar, M. Kumar, S. Sharma, and R. Prasad, eds. Springer, Singapore.

Lundberg, D. S., Lebeis, S. L., Paredes, S. H., Yourstone, S., Gehring, J., Malfatti, S., Tremblay, J., Engelbrektson, A., Kunin, V., del Rio, T. G., Edgar, R. C., Eickhorst, T., Ley, R. E., Hugenholtz, P., Tringe, S. G., and Dangl, J. L. 2012. Defining the core *Arabidopsis thaliana* root microbiome. Nature 488:86-90.

Mikan, M. P., Harvey, H. R., Timmins-Schiffman, E., Riffle, M., May, D. H., Salter, I., Noble, W. S., and Nunn, B. L. 2020. Metaproteomics reveal that rapid perturbations in organic matter prioritize functional restructuring over taxonomy in western Arctic Ocean microbiomes. ISME J. 14: 39-52.

Morris, R. M., Nunn, B. L., Frazar, C., Goodlett, D. R., Ting, Y. S., and Rocap, G. 2010. Comparative metaproteomics reveals ocean-scale shifts in microbial nutrient utilization and energy transduction. ISME J. 4: 673-685.

Niu, B., and Kolter, R. 2018. Quantification of the composition dynamics of a maize root-associated simplified bacterial community and evaluation of its biological control effect. Bio-Protoc. 8:e2885-e2885.

Niu, B., Paulson, J. N., Zheng, X., and Kolter, R. 2017. Simplified and representative bacterial community of maize roots. Proc. Natl. Acad. Sci. U.S.A. 114:E2450-E2459.

Oberg, A., and Vitek, O. 2009.Statistical design of quantitative mass spectrometry-based proteomic experiments. J. Proteome Res. 8:2144-2156.

Ofek-Lalzar, M., Sela, N., Goldman-Voronov, M., Green, S. J., Hadar, Y., and Minz, D. 2014. Niche and host-associated functional signatures of the root surface microbiome. Nat. Commun. 5:4950.

Reinhold-Hurek, B., Bünger, W., Burbano, C. S., Sabale, M., and Hurek, T. 2015. Roots shaping their microbiome: Global hotspots for microbial activity. Annu. Rev. Phytopathol. 53:403-424.

Salvato, F., Hettich, R. L., and Kleiner, M. 2021. Five key aspects of metaproteomics as a tool to understand functional interactions in host-associated microbiomes. PLOS Pathog. 17:e1009245.

Simonin, M., Dasilva, C., Terzi, V., Ngonkeu, E. L. M., Diouf, D., Kane, A., Béna, G., and Moulin, L. 2020. Influence of plant genotype and soil on the wheat rhizosphere microbiome: Evidences for a core microbiome across eight African and European soils. FEMS Microbiol. Ecol. 96:fiaa067.

Tanca, A., Palomba, A., Pisanu, S., Deligios, M., Fraumene, C., Manghina, V., Pagnozzi, D., Addis, M. F., and Uzzau, S. 2014. A straightforward

and efficient analytical pipeline for metaproteome characterization. Microbiome 2:49.

Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M. Y., Geiger, T., Mann, M., and Cox, J. 2016. The Perseus computational platform for comprehensive analysis of (prote)omics data. Nat. Methods 13: 731-740.

Vizcaíno, J. A., Reisinger, F., Côté, R., and Martens, L. 2010. PRIDE: Data submission and analysis. Curr. Protoc. Protein Sci. 60: 25.4.1-25.4.11.

Vorholt, J. A. 2012. Microbial life in the phyllosphere. Nat. Rev. Microbiol. 10:828-840.

Wagner, M. R., Roberts, J. H., Balint-Kurti, P., and Holland, J. B. 2020. Microbiome composition differs in hybrid and inbred maize. bioRxiv: 2020.01.13.904979.

Wilmes, P., and Bond, P. L. 2004. The application of two-dimensional polyacrylamide gel electrophoresis and downstream analyses to a mixed community of prokaryotic microorganisms. Environ. Microbiol. 6:911-920.

Wiśniewski, J. R., Zougman, A., Nagaraj, N., and Mann, M. 2009. Universal sample preparation method for proteome analysis. Nat. Methods 6: 359-362.

Wu, X., Gong, F., and Wang, W. 2014. Protein extraction from plant tissues for 2DE and its application in proteomic analysis. Proteomics 14: 645-658.

Xu, J., Zhang, Y., Zhang, P., Trivedi, P., Riera, N., Wang, Y., Liu, X., Fan, G., Tang, J., Coletta-Filho, H. D., Cubero, J., Deng, X., Ancona, V., Lu, Z., Zhong, B., Roper, M. C., Capote, N., Catara, V., Pietersen, G., Vernière, C., Al-Sadi, A. M., Li, L., Yang, F., Xu, X., Wang, J., Yang, H., Jin, T., and Wang, N. 2018. The structure and function of the global citrus rhizosphere microbiome. Nat. Commun. 9:4894.

Zampieri, E., Chiapello, M., Daghino, S., Bonfante, P., and Mello, A. 2016. Soil metaproteomics reveals an inter-kingdom stress response to the presence of black truffles. Sci. Rep. 6:25773.

Zhalnina, K., Louie, K. B., Hao, Z., Mansoori, N., da Rocha, U. N., Shi, S., Cho, H., Karaoz, U., Loqué, D., Bowen, B. P., Firestone, M. K., Northen, T. R., and Brodie, E. L. 2018. Dynamic root exudate chemistry and microbial substrate preferences drive patterns in rhizosphere microbial community assembly. Nat. Microbiol. 3:470-480.

Zhang, X., Deeke, S. A., Ning, Z., Starr, A. E., Butcher, J., Li, J., Mayne, J., Cheng, K., Liao, B., Li, L., Singleton, R., Mack, D., Stintzi, A., and Figeys, D. 2018a. Metaproteomics reveals associations between microbiome and intestinal extracellular vesicle proteins in pediatric inflammatory bowel disease. Nat. Commun. 9:2873.

Zhang, X., Li, L., Mayne, J., Ning, Z., Stintzi, A., and Figeys, D. 2018b. Assessing the impact of protein extraction methods for human gut metaproteomics. J. Proteomics. 180:120-127.