Title: Bulk and spatially resolved extracellular metabolomics of free-living nitrogen fixation

Author names and affiliations:

Darian N Smercina ^a, Young-Mo Kim ^a, Mary S Lipton ^b, Dusan Velickovic ^b, Kirsten S Hofmockel ^{a, c}

^a Biological Sciences Division, Earth and Biological Sciences Directorate, Pacific Northwest National Laboratory, Richland, WA, USA

^b Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, WA 99354, USA

^c Department of Agronomy, Iowa State University, Ames, IA 50010

Corresponding author: Darian Smercina, Battelle for the US DOE, 3335 Innovation Blvd., Richland, WA 99354, (509) 375-2950, darian.smercina@pnnl.gov

ORCID IDs:

Darian Smercina: 0000-0002-8484-3827

Young-Mo Kim: 0000-0002-8972-7593

Mary Lipton: 0000-0001-7749-0077

Dusan Velickovic: 0000-0001-7945-9620

Kirsten Hofmockel: 0000-0003-1586-2167

Keywords: Soil microbial ecology, multi-scale, spatially resolved, nitrogen fixation, microbial metabolomics, GC-MS, MALDI MSI

Competing interests: The authors declare no conflicts of interest.

Funding: D.N.S. is also grateful for the support of the Linus Pauling Distinguished Postdoctoral Fellowship program through Pacific Northwest National Laboratory. Pacific Northwest National Laboratory is a multiprogram national laboratory operated by Battelle for the U.S. Department of Energy under contract DE-AC05-76RL01830. This research was performed using resources at the Environmental Molecular Sciences Laboratory (EMSL; grid.436923.9), a DOE Office of Science User Facility sponsored by the Biological and Environmental Research program.

Abstract

Soil microorganisms drive ecosystem function, but challenges of scale between microbe and ecosystem hinder our ability to accurately quantify and predictively model the soil microbe-ecosystem function relationship. Quantifying this relationship necessitates studies that systematically characterize multi-omics of soil microorganisms and their activity across sampling scales from spatially resolved to bulk measures, and structural complexity, from liquid pure culture to *in situ*. To address this need, we cultured two diazotrophic bacteria in liquid and solid media, with and without nitrogen (N) to quantify differences in extracellular metabolites associated with nitrogen fixation under increasing environmental structural complexity. We also quantified extracellular metabolites across sampling scales including bulk sampling via GC-MS analysis and spatially resolved analysis via MALDI mass spectrometry imaging. We found extracellular production of inorganic and organic N during free-living nitrogen fixation activity, highlighting a key mechanism of terrestrial N contributions from this process. Additionally, our results emphasize the need to consider the structural complexity of the environment and spatial scale when quantifying microbial activity. We found differences in metabolite profiles between culture conditions, supporting previous work indicating environmental structure influences microbial function, and across scales, underscoring the need to quantify microbial scale conditions to accurately interpret microbial function.

Importance: Studying soil microorganisms, both who is present and what they are doing, is a challenge because of vast differences in scale between microorganism and ecosystem and because of inherent complexities of the soil system (e.g., opacity, chemical complexity). This makes measuring and predicting important ecosystem processes driven by soil microorganisms,

like free-living nitrogen fixation, difficult. Free-living nitrogen fixing bacteria play a key role in terrestrial nitrogen contributions and may represent a significant, yet overlooked, nitrogen source in agricultural systems like bioenergy crops. However, we still know very little about how free-living nitrogen fixation contributes nitrogen to terrestrial systems. Our work provides key insight by hierarchically increasing structural complexity (liquid vs. solid culture) and scale (spatially resolved vs. bulk) to address the impact of environmental structure and sampling scale on detection of free-living nitrogen fixation and to identify the forms of nitrogen contributed to terrestrial systems by free-living nitrogen bacteria.

Introduction

Soil microorganisms are a key link between above and belowground ecosystem function, driving energy and nutrient transfer between the atmosphere, biosphere, and pedosphere (1-3). Multi-omic analysis aimed at understanding the structure and function of these microorganisms has become a routine tool for environmental samples. Despite generating large amounts of data, even quantitatively linking ‘omics of a specific function (e.g. functional genes and proteins) to measures of that function is often unsuccessful (2-6) because of inherent challenges of studying soils (5-7) and soil microorganisms. Given these challenges, the importance of linking microbial community structure to function has been questioned (8, 9), however without clearer understanding of how microorganisms relate to observed functions we cannot yet rule out the importance of individual microbial community members. Our limited ability to quantitatively link soil microbial communities to ecosystem function hinders our understanding of ecosystem processes, leaving us vulnerable to losing vital ecosystem services provided by our soils, particularly in the face of climate change (10).

Quantifying the link between soil microorganisms and ecosystem function requires systematic studies that characterize multi-omics of soil microorganisms and their functions hierarchically across scales of space and complexity (Fig. 1; (4, 11)). *In vitro* studies using pure cultures or limited species are an appealing option and have the potential to provide fundamental microbial and ecological knowledge (10, 12, 13). However, culturing conditions are often quite different from those experienced by microorganisms in soil and attachment to surfaces has been shown to impact microbial growth and function (14, 15). Thus, physical structure influences microbial function, and it is therefore essential for studies to systematically characterize multi-omics and function *in vitro* under growth conditions of increasing complexity (e.g. environmental structure) in order to determine how culture work may better inform *in situ* processes (Fig. 1).

In this study, we explored two questions surrounding the microbe-ecosystem function relationship: (1) Are there microbial metabolomic signatures associated with target ecosystem processes? (2) How do growth conditions and sampling scale influence metabolomic signatures? We used free-living nitrogen fixation (FLNF), biological nitrogen fixation (BNF) carried out by heterotrophic bacteria (diazotrophs), as a model microbial process and ecosystem function to address these questions. FLNF is carried out by a wide diversity of soil bacteria and occurs in all terrestrial biomes, contributing significantly to terrestrial N (16, 17). These contributions are thought to occur predominately through release of ammonium and organic N sources, like amino acids, as occurs during symbiotic BNF (18-21). Thus, extracellular metabolites may be valuable to understanding FLNF and identifying biosignatures.

We examined extracellular metabolites from two diazotrophic bacteria cultured under conditions that promote (N-free) or inhibit (N-rich) FLNF. Additionally, cultures were grown in

liquid or on solid media to examine the impact of physical structure on detected metabolites. Lastly, we examined extracellular metabolites across sampling scale from spatially resolved measures using matrix-assisted laser desorption/ionization mass spectrometry imaging (MALDI-MSI) to bulk sampling via gas chromatography-mass spectrometry (GC-MS) analysis. We hypothesized: (1) bulk metabolite profiles differ between N-rich and N-free conditions and also between liquid and solid cultures, (2) there are extracellular metabolites associated with FLNF activity and only observed under N-free conditions, (3) these extracellular metabolites are detectable with bulk and spatially resolved sampling, and (4) N-containing compounds are produced during FLNF and are readily detectable in bulk and spatially resolved extracellular metabolites under N-free conditions.

Results

Microbial biomass – total biomass, biomass C, and biomass N

Total microbial biomass, including cells and associated debris such as EPS, was collected from all treatments except for the AV N-rich solid treatment. In this case, microbial colonies had grown into and below the agar surface and it was not possible to collect biomass. Total biomass was highly variable across all treatments (CV across treatments ranged from 4.6% to 100.4%) and there were no significant differences observed with N treatment, culture type, or organism (Fig. 2D). Both biomass C and N differed significantly across treatments (Fig. 2A and 2B, respectively). In general, C and N content were greater in N-rich treatments where N was readily available compared to N-free treatments. There was also a trend towards greater biomass C and N from solid media, but this was mostly observed in the N-rich treatments. These C and N values translated to C:N ratios that predominately differed only between N treatments with the N-free

treatment resulting in greater biomass C:N ratios than N-rich conditions, regardless of culture type or organism (Fig. 2C).

Extracellular ammonium availability

Extracellular ammonium availability was measured supernatant and rinsate samples and was detected in all treatments regardless of culture type, N treatment, or organism. On a per unit biomass basis, ammonium concentrations differed significantly by organism ($F = 16.390$, $p = 0.0012$) and by the interaction between culture type and N treatment ($F = 35.411$, $p < 0.0001$). Extracellular ammonium availability per unit biomass was over 8x greater in PP than AV cultures (Fig. 3A). Under N-free conditions, ammonium availability was greater in liquid than in solid culture while in N-rich conditions the opposite was observed (Fig. 3B). Extracellular ammonium availability is of particular interest in N-free treatments as ammonium is hypothesized to be released from cells actively fixing N and thus represents a major form of N contributed by FLNF to terrestrial systems (18, 33). Therefore, we also calculated the percent of fixed N available as extracellular ammonium (Fig. 4). Total fixed N was estimated as total biomass N measured in N-free treatment samples. We find only upwards of 7.5% (± 2.3) of fixed N is readily available as extracellular ammonium. This did not differ significantly between culture types ($F = 2.171$, $p = 0.184$), though ammonium concentrations in solid culture tended to be lower than those in liquid culture.

Bulk extracellular metabolites

Across all treatments, 307 metabolites were detected with bulk sampling and of these 93 were successfully annotated (>80% confidence). The total number of detected metabolites

differed between treatment groups (Supp. Fig. 2) with generally more metabolites detected in N-free treatments, the majority of which were within the unannotated portion of detected metabolites. Distinct metabolite profiles, represented by Bray-Curtis and Jaccard distance based on all detected metabolites, were observed between N treatment, culture type, and their interaction (Fig. 5; Table 1). N treatment and culture type, together with their interaction, explain over 50% of the variance in metabolite profiles based on peak intensity (Fig. 5A) and based on presence-absence (Fig. 5B). Metabolite profiles based on abundance separated predominantly by culture type and then by N treatment (Fig. 5A). Metabolite profiles based on presence-absence show clear separation between culture types for N-rich treatments but have little separation under N-free conditions (Fig. 5B).

Because FLNF activity is hypothesized to result in the release of N-containing metabolites (33), we focused on N-containing extracellular metabolites. Of the 93 annotated metabolites detected through bulk sampling, 35 were N-containing. We found significant differences in N-containing compounds across treatments, after correcting for background metabolites, with significant interactions between N treatment, culture type, and organism ($F = 10.695$, $p = 0.0048$). N-free treatments were richer in N-containing metabolites (Supp. Fig. 3) but had similar or significantly lower total abundances of N-containing metabolites compared to N-rich treatments (Fig. 6). Examining the specific composition of these N-containing compounds, we found a variety of amino acids in N-free samples not well represented in N-rich samples (Supp. Fig. 3), but only a few N-containing metabolites were unique to N-free conditions including pantothenic acid, L-pyroglutamic acid, L-glutamic acid, and 4-pyridoxic acid.

Spatially resolved extracellular metabolites

Across all treatments, METASPACE analysis identified 69 metabolites in spatially resolved samples of which 41 were N-containing. However, only a few potential amino acids were detected at this resolved microbial scale including L-leucine and L-valine. These were only at detectable concentrations within the N-rich treatment (Fig. 7) unlike the diversity of amino acids detected in bulk samples predominately in association with N-free treatments. Observationally, N-free treatments seemed to be characterized by unique presence of organic acids rather than N-containing compounds. However, we did identify a few N-containing compounds unique to N-free treatments at the microbial scale including inosine and 4-pyridoxic acid (Fig. 7). Inosine was detected in N-free treatments of both AV and PP and was not at detectable levels in N-rich treatments. Also, much like bulk sampling scale detection, 4-pyridoxic acid was exclusively detected in AV N-free treatment samples.

Discussion

We explored the impact of N availability and physical structure on the extracellular metabolomics of diazotrophic bacteria across sampling scales from the spatially resolved to bulk. We find evidence of extracellular organic and inorganic N contributions from FLNF underscoring a key mechanism of terrestrial N contributions from FLNF. In general, we find physical structure and microbial function (e.g. FLNF) alter extracellular metabolite profiles and influence the detection of metabolites at bulk and spatially resolved scales.

Nitrogen contributions from FLNF

Products of BNF by symbiotic diazotrophs are well-studied and typically observed as ammonia and ammonium with contested evidence for production of amino acids (18-21). This

knowledge of symbiotic BNF is thought to translate directly to FLNF leading to the assumption that free-living diazotrophs also excrete ammonia/ammonium into the surrounding environment during BNF. However, ammonia produced during FLNF is rapidly assimilated through conversion to glutamine or glutamate via the glutamine synthetase (GS) and glutamate synthetase (GOGAT) pathways (34). Thus, excreted ammonium would necessarily be in excess of these assimilation pathways (34). Ammonium excretion has been observed in wild-type *Azotobacter vinelandii* DJ, at concentrations between ~2 and ~25 μM (35, 36), values within range of those measured in this study (Supp. Fig. 4). However, in many cases measurable ammonium excretion was only observed from *Azotobacter vinelandii* cultures genetically altered to disrupt the GS-GOGAT pathways or facilitate constitutive nitrogenase synthesis (37-40).

An alternative hypothesis to ammonium excretion is that N contributions occur as organic N, either through direct release of N-rich compounds like amino acids (18, 21) or through turnover of dead biomass (41). Our bulk metabolomics data support this hypothesis with many N-containing organic compounds, including amino acids detected in N-free treatments. In fact, N-free treatments were richer in N-containing compounds than N-rich conditions, particularly when comparing against N-rich solid media which had few N-containing metabolites (Supp. Fig. 3). The structure of our study did not allow us to determine whether these organic molecules were directly excreted by active, N-fixing cells or released during cell turnover. However, other metabolites detected in the system suggest cell turnover contributed at least partially to this N release. For example, we detected inosine in both bulk and spatially resolved analysis, and it was unique to N-free treatments in spatially resolved samples. Inosine, a metabolic product of adenine degradation likely indicates salvage activities by the bacterial populations (42, 43) and could indicate freely available nucleotides from cell lysis and turnover. FLNF may therefore

contribute available N through increasing microbial biomass and turnover, but this needs to be verified in future studies. Regardless of whether these N-containing compounds are actively excreted or released after cell death, this metabolic exchange with the surrounding environment highlights a key mechanisms of terrestrial N contributions from FLNF.

In vitro vs in situ analysis – identification of biosignatures

Through bulk and spatially resolved analysis, we found few N-containing metabolites exclusive to N-free treatments. At bulk scale, these include pantothenic acid, L-pyroglutamic acid, L-glutamic acid, and 4-pyridoxic acid. We similarly find 4-pyridoxic acid at the spatially resolved microbial scale as well as nine other metabolites including inosine. 4-pyridoxic acid was unique to AV N-free treatments at the microbial scale. However, despite being uniquely associated with N-free treatments and therefore microbial populations actively fixing N, it may be difficult to assign these as a signature of FLNF function. Of these compounds, only L-glutamic acid has a direct association with the FLNF pathway. Other metabolites seem more indicative of microbial nutrient needs and function. For example, pantothenic acid, vitamin B₅, is involved in the synthesis of coenzyme A and is a coenzyme for many reactions involved in protein and lipid metabolism (44-46). This is particularly important for the processing of organic acids like malate, the main C source provided in this study. Thus, the detection of vitamin B₅ is likely indicative of malate metabolism via the TCA cycle and its unique detection in the N-free treatment suggests a higher respiration rate in these N-fixing populations than in the N-rich populations. Increased respiration is a common response among diazotrophs in oxygenated environments as a protection mechanism to prevent or reduce denaturation of nitrogenase via oxygen (17, 47, 48). We also identified 4-pyridoxic acid, a derivative of pyridoxine (vitamin B₆).

Pyridoxine is a key cofactor in amino acid, fatty acid, and carbohydrate metabolisms, but can also act as an oxygen protectant (46). During this redox reaction, pyridoxine degrades and can result in 4-pyridoxic acid. *A. vinelandii* has been observed to produce B vitamins while under diazotrophic conditions and this seems to be a hallmark of FLNF for this organism (46, 49, 50). Though not directly associated with the N-fixation pathway, these vitamins may tangentially indicate bacterial functions surrounding FLNF such as oxygen regulation and highlight the need to analyze bacterial function holistically rather than focusing on single reactions or pathways.

Additionally, the limited number of unique extracellular metabolites detected in N-free treatments suggests some microbial functions may not have detectable or unique biosignatures, in the form of extracellular metabolites. This is an important consideration when applying metabolomics to the study of complex soil systems. Soil metabolomics are increasingly being used to study soil microbial ecology and biogeochemical function and have been successfully applied to soil carbon cycling (51-54). However, metabolites are by definition the by-products of and substrates for metabolic function, and turnover rapidly in soils (55, 56). Therefore, typical soil extractions to collect extracellular components (e.g. K₂SO₄ extracts, leachate; (57, 58) only capture what is not consumed by the microbial community. This includes metabolites available in dissolved organic matter pools at the time of sampling and metabolites readily exchangeable from mineral surfaces (59). In both cases, metabolites could be temporally separated from their originating processes making it difficult to trace back the associated metabolic pathway. It could be even more challenging to capture metabolic biosignatures from nutrient-limited communities, such as those in bulk soil. Under nutrient-limited conditions, resulting metabolic products are likely to be rapidly assimilated or, in the case of processes like FLNF, not released to the surrounding environment. The potential signature compounds of FLNF found here (e.g. amino

acids and B-vitamins) are also not uniquely produced by FLNF processes and would be difficult to directly link to FLNF *in situ*. We acknowledge our study system may provide a biased view on this issue, being a closed incubation system unlike soils where metabolites may diffuse away from microbes and persist in the environment. However, these findings highlight a key need to understand the soil microhabitat (11).

Lastly, our work highlights the importance of considering culture conditions and their association to *in situ* conditions as we observed clear differences in metabolite profiles between liquid and solid culture. Interestingly, though culture type strongly influenced metabolite profiles, it played a secondary role to N treatment in influencing the number of detected metabolites. This was particularly notable when N was readily available, where presence-absence based profiles were distinct between liquid and solid culture, but only under N-rich conditions. These differences in metabolite profiles were likely not driven by differences in biomass production as culture type had small and non-significant impacts on microbial metrics, like total biomass, and biomass C and N content. Thus, these responses seem specifically associated with the presence or absence of physical structure in the environment. Additionally, these findings suggest nutrient limitation, as experienced in the N-free treatments, may be a stronger driver of microbial activity than physical structure and simplified liquid culture may be somewhat informative to nutrient-limited *in situ* conditions.

Implications scaling from microbial scale to bulk sampling

The combination of techniques in this study allowed us to explore detection of metabolites across scales from spatially resolved, relevant to microorganisms, to bulk, relevant for soil microbial ecology analysis. MALDI MSI allowed us to resolve abundances of

extracellular metabolites on solid media at a microbial scale. Using GC-MS, we were able to evaluate detection of extracellular metabolites at the bulk scale. While the detection ranges of these two techniques do not fully overlap (50 – 500 m/z for GC-MS and 92 – 700 m/z for MALDI), many metabolites of interest to this study are measurable with both techniques providing valuable information about metabolite detection and sampling scale. It is also important to note that a lack of detection is not equivalent to metabolite absence but only indicates metabolite concentrations were below detection.

Through bulk sampling, we found a wide variety of N-containing compounds in N-free samples, but generally lower abundances of N-containing compounds than in the N-rich treatment. While N-containing compounds are characteristic of N-free samples at a bulk scale, these treatments had fewer N-containing metabolites when spatially resolved at the microbial scale. Interestingly, there was a shift in amino acid detection between spatially resolved and bulk scales where amino acids were commonly detected in N-free samples at bulk scale, but in N-rich samples when spatially resolved. This somewhat counterintuitive result highlights differences in N competition at the microbial scale and its influence on bulk measurements.

First, detection of a diverse array of amino acids in the N-free treatment in bulk sample, but not in spatially resolved samples suggests N competition at the microbial scale resulted in rapid uptake of amino acids, while extraction of the bulk metabolite pool likely captured the cumulative low abundance signal of the entire system. Amino acids are shown to have short residence times in soils and experience rapid uptake and turnover (60, 61). In the case of microbial vs. bulk scale, it is likely the spatially resolved pool of extracellular amino acids collected from microbial colonies (~200 μm spatial resolution) was small and often below detection. However, in bulk sampling of millions of cells, a larger pool of amino acids coupled to

our sampling method could have allowed amino acids to diffuse away and accumulate to detectable levels.

Second, biofilm formation is likely to influence diffusion of metabolites into the surrounding environment (62, 63). Bacteria tend to live in biofilms in their natural environments rather than as individually dispersed cells (64). However, the impact of surrounding environmental conditions, including nutrient availability, on biofilm production is unclear. For example, some studies suggest nutrient limiting conditions may promote greater biofilm formation (65), while others suggest biofilm formation is greater under more favorable growth conditions (66, 67). This is particularly notable for diazotrophs as biofilms can play a role in oxygen protection (47), thus investment in biofilm could be beneficial to FLNF activity. Yet, under severe N limitation imposed by an N-free environment the high energy demands of FLNF may limit investment in biofilm. While not directly measured in this study, we noted solid agar plates of *Azotobacter vinelandii* and *Paenibacillus polymyxa* cultures had visually greater biofilm production under N-rich than N-free conditions. Thus, diffusion of amino acids away from populations would have been more easily achieved in the N-free treatment. This is evidenced by the similarity between metabolite chemistry in liquid and solid culture under N-free treatments (Fig. 5B). Similarly, a small number of amino acids were detected in spatially resolved samples from N-rich treatments, but not in bulk samples for the similar N-rich solid media treatment. This could have resulted from greater biofilm formation under N-rich conditions and limited diffusion of small molecules away from cell populations.

The detection of small molecules across scales has important implications for the influence of soil microbial communities on their surrounding environment. In general, our results indicate microbial scale processes drive bulk metabolite availability. The N-rich treatment in this

experiment is an optimal environment and most representative of carbon and nutrient rich soil environments like the rhizosphere or detritosphere. Our findings suggest these conditions would result in production of valuable small molecules, like N-rich amino acids, potentially exchangeable with the immediate environment, but biofilm formation may limit diffusion far into the soil environment. Under limiting conditions of the N-free treatment, similar those of bulk soil, microbial activity produces valuable metabolites, like amino acids, but competition between microbes reduces the exchange of these molecules. Understanding how these differences in microbial scale conditions influence microbial activity and detectability of function is crucial to accurately linking microbe and ecosystem.

Conclusions

We demonstrate extracellular production of inorganic and organic N during FLNF and reveal the importance of habitat conditions and sampling scale when quantifying microbial activity. Across bulk and spatially resolved sampling scales, we identified metabolites uniquely associated with FLNF activity including several B-vitamins, which may play roles in mitigating oxygen damage to nitrogenase. Despite finding unique metabolites and potential biosignatures, many detected metabolites are not exclusively produced through FLNF related pathways, thus would be difficult to assign to FLNF for *in situ* soil samples. This would likely hold true for other processes under nutrient limited conditions where metabolic products are rapidly assimilated and not captured during sampling. Our findings highlight the need to carefully consider both structural complexity and sampling scale when quantifying microbial function. We found culture conditions to be a key driver of metabolite chemistry under N-rich and N-free conditions, indicating physical structure influences microbial processes. Across scales, our

results indicate high N competition at the microbial scale under N-free conditions, while at the bulk scale N appeared readily available within the microbial environment. These differences in environmental conditions across scales could lead to incorrect interpretations of microbial function as immediate conditions surrounding microorganisms will drive their activity and may not necessarily match what is measured through bulk or composite sampling.

Materials and Methods

Culture conditions

Two diazotrophic bacteria with distinct growth strategies (e.g. gram-negative vs gram-positive) and fully sequenced genomes (22, 23) were chosen for this study, *Azotobacter vinelandii* DJ (ATCC BAA 1303; hereafter AV) and *Paenibacillus polymyxa* (ATCC 842; hereafter PP). Bacteria were cultured under N-free and N-rich conditions, respectively promoting or inhibiting FLNF. Nfb media, commonly used to isolate diazotrophs (24), was used for N-free treatments, and was supplemented with tryptone for N-rich treatments. Both treatments contained 1.79 g C L⁻¹ as malic acid and N-rich media contained ~1.33 g N L⁻¹ as tryptone. Cultures were grown in liquid or solid agar media and all media was autoclave sterilized prior to inoculation.

Thirty samples were cultured (2 organisms x 2 N treatments x 2 media types x 3 replicates, plus controls) for bulk analysis with an additional set of 14 samples (2 organisms x 2 N treatments x 3 replicates, plus controls) on solid media for spatially resolved analysis. Cultures were grown in a temperature-controlled incubator at 25°C to 10⁷ CFU mL⁻¹, based on liquid cultures, and then harvested for analysis of extracellular metabolites at two scales – bulk sampling via MPLEx extraction and GC-MS (25) and spatially-resolved sampling via colony

analysis with MALDI MSI. Extracellular ammonium availability and microbial biomass, including total biomass, biomass carbon (C) and biomass N, were also measured. Because FLNF activity is necessary for microbial growth under N-free conditions, measures of total biomass and biomass N are used as estimates of FLNF (24).

Sample collection

Extracellular metabolites were collected from liquid culture by centrifuging culture tubes to pellet cells and collecting the resulting supernatant for bulk analysis as described below. Cell pellets were resuspended in autoclave sterilized nanopure water, immediately flash frozen on liquid nitrogen, and stored at -80°C until further analysis. Extracellular metabolites were collected from solid media for bulk and spatially resolved analysis. Bulk samples were collected by washing culture plate surfaces with autoclave sterilized nanopure water and collecting the resulting rinsate. Samples were collected for spatially resolved analysis as described below. Lastly, microbial colonies from rinsate plates were collected from the surface by gentle scraping, transferred to autoclave sterilized nanopure water, flash frozen on liquid nitrogen, and stored at -80°C until further processing.

Microbial biomass – total biomass, biomass C, and biomass N

Frozen cell pellets and colonies were lyophilized until completely dry and weighed to obtain total biomass, including cells and associated debris such as exopolysaccharides (EPS). Dried biomass was ground using sterile steel beads and then analyzed for C and N content on a VarioTOC Cube (Elementar, Langenselbold, Germany).

Extracellular ammonium availability

We measured extracellular ammonium concentrations in supernatant and rinsate samples using a high-throughput colorimetric ammonium assay (26). Briefly, samples were pipetted in

triplicate into clear 96-well plates and incubated with ammonium salicylate and ammonium cyanurate reagents to facilitate color change via the Berthelot reaction. Plates were read for absorbance at 610 nm on a Synergy H1 plate reader (BioTek Instruments, Inc., Winooski, VT, USA).

Bulk metabolomics – GC-MS

Bulk metabolomics analysis was conducted on 1 ml subsamples of undiluted, supernatant and rinsate samples. Supernatant and rinsate samples were prepared for metabolite analysis via GC-MS following the MPLEx protocol for simultaneous metabolite, protein, and lipid extraction (25). Additionally, 1 ml of supernatant and rinsate from sterile liquid culture and solid culture plates were also extracted via MPLEx as background controls. This extraction method allows simultaneously collection of metabolites, lipids, and proteins; however, lipid fractions were not analyzed in this study. Additionally, protein yields were too low for downstream analysis. Metabolite samples were completely dried under speed-vacuum concentrator and chemically derivatized prior to analysis by GC-MS as reported previously (27). The m/z range of derivatized metabolites scanned was 50 – 550 m/z which can detect organic acids, amino acids, and mono to tri-saccharides. Raw GC-MS data were processed using the PNNL in-house metabolomics database, which can identify metabolites using two dimensional matching factors (fragmented spectrum + retention index (28), and with cross-checking against commercially available NIST 20/Wiley 11th GC-MS spectral databases (25, 29).

Spatially resolved metabolomics – MALDI MSI

Samples were prepared for spatially resolved analysis via MALDI-MSI using a previously described workflow (30). Briefly, areas of agar were excised from Petri dishes and placed onto double-sided adhesive copper tape adhered to indium tin oxide (ITO)-coated glass

slides (Bruker Daltonics; Supp. Fig 1.). This approach enhanced our sensitivity for analysis in negative ionization mode and improved adhesion of agar onto the MALDI target. Samples were dried at room temperature overnight, then treated with MALDI matrix using a HTX TM-Sprayer (HTX Technologies). For analysis in negative-ion mode, 7 mg mL⁻¹ of N-(1-naphthyl) ethylenediamine dihydrochloride (NEDC) in 70% MeOH was sprayed with eight passes at 1,200 µL min⁻¹, 75°C, a spray spacing of 3 mm, and a spray velocity of 1,200 mm min⁻¹. MALDI-MSI was performed on a 15-Tesla Fourier transform ion cyclotron resonance (FTICR)-MS (Bruker Daltonics, Billerica, MA, USA) equipped with SmartBeam II laser source (355 nm) using 200 shots pixel⁻¹ with a frequency of 2 kHz and a step size of 200 µm. FTICR-MS was operated to collect m/z 92–700, using a 209-ms transient, which translated to a mass resolution of R ~ 70,000 at 400 m/z. Metabolites in this range can typically be detected to fmol concentrations.

Data Analysis

A factorial ANOVA with N treatment, culture type, organism and their interactions as main effects followed by a Tukey's post hoc test was used to determine treatment differences for measured variables. Prior to statistical analysis, bulk metabolite values were blank corrected by subtracting peak intensities identified in background controls of the associated treatment. Differences in bulk metabolite chemistry were evaluated using distance matrices based on peak intensities (Bray-Curtis) presence-absence (Jaccard) generated from all detected metabolites using R *vegan* (31). Differences between culture type, N treatment, and organism were determined via PERMANOVA using *adonis* in R *vegan*. Spatially resolved metabolite data was acquired using FlexImaging (v 4.1, Bruker Daltonics), and image processing, segmentation, co-localization analysis and visualization were performed using SCLS (Bruker Daltonics). The list of m/z values that colocalized with the colonies were uploaded to the METLIN

(<https://metlin.scripps.edu>) for putative molecular annotations based only on accurate m/z, secured by using a 3-ppm window during the search. imzML files (created by SCiLS) of our analyses were also uploaded to METASPACE (32) for metabolite annotation based on both accurate m/z and a comprehensive bioinformatics framework that considers the relative intensities and spatial colocalization of isotopic peaks as well as quantifies spatial information with a measure of spatial chaos followed by the estimation of the False Discovery Rate. For this purpose, we used KEGG-v1 and NPA-2019-08 (Natural Product Atlas) databases that are available in METASPACE. METASPACE uses by default 3 ppm window in its annotation engine.

Acknowledgements: We would like to thank Heather Olson, Carrie Nicora, and Jesse Trejo for sampling processing work and for taking time to carefully train D.N.S. on the methods. D.N.S. is also grateful for the support of the Linus Pauling Distinguished Postdoctoral Fellowship program through Pacific Northwest National Laboratory. Pacific Northwest National Laboratory is a multiprogram national laboratory operated by Battelle for the U.S. Department of Energy under contract DE-AC05-76RL01830. This research was performed using resources at the Environmental Molecular Sciences Laboratory (EMSL; grid.436923.9), a DOE Office of Science User Facility sponsored by the Biological and Environmental Research program.

References

1. Falkowski PG, Fenchel T, Delong EF. The microbial engines that drive Earth's biogeochemical cycles. *science*. 2008;320(5879):1034-9.
2. Hall EK, Bernhardt ES, Bier RL, Bradford MA, Boot CM, Cotner JB, et al. Understanding how microbiomes influence the systems they inhabit. *Nature microbiology*. 2018;3(9):977-82.
3. Jansson JK, Hofmockel KS. The soil microbiome—from metagenomics to metaphenomics. *Current opinion in microbiology*. 2018;43:162-8.
4. Graham EB, Knelman JE, Schindlbacher A, Siciliano S, Breulmann M, Yannarell A, et al. Microbes as engines of ecosystem function: when does community structure enhance predictions of ecosystem processes? *Frontiers in microbiology*. 2016;7:214.
5. Martiny JB, Jones SE, Lennon JT, Martiny AC. Microbiomes in light of traits: a phylogenetic perspective. *Science*. 2015;350(6261).
6. Carini P, Marsden PJ, Leff JW, Morgan EE, Strickland MS, Fierer N. Relic DNA is abundant in soil and obscures estimates of soil microbial diversity. *Nature microbiology*. 2016;2(3):1-6.
7. Kleiner M. Metaproteomics: much more than measuring gene expression in microbial communities. *Msystems*. 2019;4(3):e00115-19.
8. Prosser JJ. Ecosystem processes and interactions in a morass of diversity. *FEMS Microbiology Ecology*. 2012;81(3):507-19.
9. Schimel J, Schaeffer SM. Microbial control over carbon cycling in soil. *Frontiers in microbiology*. 2012;3:348.

503 10. Jansson JK, Hofmockel KS. Soil microbiomes and climate change. *Nature Reviews*
504 *Microbiology*. 2020;18(1):35-46.

505 11. Baveye PC, Otten W, Kravchenko A. elucidating microbial processes in soils and
506 sediments: microscale measurements and modeling. *Frontiers in Environmental Science*.
507 2019;7:78.

508 12. Jessup CM, Kassen R, Forde SE, Kerr B, Buckling A, Rainey PB, et al. Big questions,
509 small worlds: microbial model systems in ecology. *Trends in ecology & evolution*.
510 2004;19(4):189-97.

511 13. Prosser JI, Bohannon BJ, Curtis TP, Ellis RJ, Firestone MK, Freckleton RP, et al. The
512 role of ecological theory in microbial ecology. *Nature Reviews Microbiology*. 2007;5(5):384-92.

513 14. Vandevivere P, Kirchman DL. Attachment stimulates exopolysaccharide synthesis by a
514 bacterium. *Applied and environmental microbiology*. 1993;59(10):3280-6.

515 15. Prigent-Combaret C, Vidal O, Dorel C, Lejeune P. Abiotic surface sensing and biofilm-
516 dependent regulation of gene expression in *Escherichia coli*. *Journal of bacteriology*.
517 1999;181(19):5993-6002.

518 16. Reed SC, Cleveland CC, Townsend AR. Functional ecology of free-living nitrogen
519 fixation: a contemporary perspective. *Annual review of ecology, evolution, and systematics*.
520 2011;42:489-512.

521 17. Smercina DN, Evans SE, Friesen ML, Tiemann LK. To fix or not to fix: controls on free-
522 living nitrogen fixation in the rhizosphere. *Applied and Environmental Microbiology*.
523 2019;85(6).

- 524 18. Allaway D, Lodwig E, Crompton L, Wood M, Parsons R, Wheeler T, et al. Identification
525 of alanine dehydrogenase and its role in mixed secretion of ammonium and alanine by pea
526 bacteroids. *Molecular microbiology*. 2000;36(2):508-15.
- 527 19. Li Y, Parsons R, Day DA, Bergersen FJ. Reassessment of major products of N₂ fixation
528 by bacteroids from soybean root nodules. *Microbiology*. 2002;148(6):1959-66.
- 529 20. Udvardi M, Poole PS. Transport and metabolism in legume-rhizobia symbioses. *Annual*
530 *review of plant biology*. 2013;64:781-805.
- 531 21. Waters JK, Mawhinney TP, Emerich DW. Nitrogen assimilation and transport by ex
532 planta nitrogen-fixing *Bradyrhizobium diazoefficiens* bacteroids is modulated by oxygen,
533 bacteroid density and L-malate. *International Journal of Molecular Sciences*. 2020;21(20):7542.
- 534 22. Setubal JC, Dos Santos P, Goldman BS, Ertesvåg H, Espin G, Rubio LM, et al. Genome
535 sequence of *Azotobacter vinelandii*, an obligate aerobe specialized to support diverse anaerobic
536 metabolic processes. *Journal of bacteriology*. 2009;191(14):4534-45.
- 537 23. Jeong H, Park S-Y, Chung W-H, Kim SH, Kim N, Park S-H, et al. Draft genome
538 sequence of the *Paenibacillus polymyxa* type strain (ATCC 842T), a plant growth-promoting
539 bacterium. *Am Soc Microbiol*; 2011.
- 540 24. Baldani JI, Reis VM, Videira SS, Boddey LH, Baldani VLD. The art of isolating
541 nitrogen-fixing bacteria from non-leguminous plants using N-free semi-solid media: a practical
542 guide for microbiologists. *Plant and soil*. 2014;384(1):413-31.
- 543 25. Nakayasu ES, Nicora CD, Sims AC, Burnum-Johnson KE, Kim Y-M, Kyle JE, et al.
544 MPLEEx: a robust and universal protocol for single-sample integrative proteomic, metabolomic,
545 and lipidomic analyses. *MSystems*. 2016;1(3).

- 546 26. Rhine E, Mulvaney R, Pratt E, Sims G. Improving the Berthelot reaction for determining
547 ammonium in soil extracts and water. *Soil Science Society of America Journal*. 1998;62(2):473-
548 80.
- 549 27. Kim Y-M, Nowack S, Olsen MT, Becraft ED, Wood JM, Thiel V, et al. Diel
550 metabolomics analysis of a hot spring chlorophototrophic microbial mat leads to new hypotheses
551 of community member metabolisms. *Frontiers in microbiology*. 2015;6:209.
- 552 28. Snijders AM, Langley SA, Kim Y-M, Brislawn CJ, Noecker C, Zink EM, et al. Influence
553 of early life exposure, host genetics and diet on the mouse gut microbiome and metabolome.
554 *Nature microbiology*. 2016;2(2):1-8.
- 555 29. Kim Y-M, Schmidt BJ, Kidwai AS, Jones MB, Kaiser BLD, Brewer HM, et al.
556 *Salmonella* modulates metabolism during growth under conditions that induce expression of
557 virulence genes. *Molecular BioSystems*. 2013;9(6):1522-34.
- 558 30. Nagy G, Veličković D, Chu RK, Carrell AA, Weston DJ, Ibrahim YM, et al. Towards
559 resolving the spatial metabolome with unambiguous molecular annotations in complex biological
560 systems by coupling mass spectrometry imaging with structures for lossless ion manipulations.
561 *Chemical Communications*. 2019;55(3):306-9.
- 562 31. Dixon P. VEGAN, a package of R functions for community ecology. *Journal of*
563 *Vegetation Science*. 2003;14(6):927-30.
- 564 32. Palmer A, Phapale P, Chernyavsky I, Lavigne R, Fay D, Tarasov A, et al. FDR-
565 controlled metabolite annotation for high-resolution imaging mass spectrometry. *Nature*
566 *methods*. 2017;14(1):57-60.
- 567 33. Sprent JI. *The biology of nitrogen-fixing organisms*: McGraw-Hill Book Co.(UK) Ltd.;
568 1979.

34. Colnaghi R, Green A, He L, Rudnick P, Kennedy C. Strategies for increased ammonium production in free-living or plant associated nitrogen fixing bacteria. *Plant and Soil*. 1997;194(1):145-54.
35. Ortiz-Marquez JCF, Do Nascimento M, Dublan MdlA, Curatti L. Association with an ammonium-excreting bacterium allows diazotrophic culture of oil-rich eukaryotic microalgae. *Applied and environmental microbiology*. 2012;78(7):2345-52.
36. Barney BM, Eberhart LJ, Ohlert JM, Knutson CM, Plunkett MH. Gene deletions resulting in increased nitrogen release by *Azotobacter vinelandii*: application of a novel nitrogen biosensor. *Applied and environmental microbiology*. 2015;81(13):4316-28.
37. Ambrosio R, Ortiz-Marquez JCF, Curatti L. Metabolic engineering of a diazotrophic bacterium improves ammonium release and biofertilization of plants and microalgae. *Metabolic Engineering*. 2017;40:59-68.
38. Smith MJ, Francis MB. A designed *A. vinelandii*–*S. elongatus* coculture for chemical photoproduction from air, water, phosphate, and trace metals. *ACS synthetic biology*. 2016;5(9):955-61.
39. Ortiz-Marquez JCF, Do Nascimento M, Curatti L. Metabolic engineering of ammonium release for nitrogen-fixing multispecies microbial cell-factories. *Metabolic Engineering*. 2014;23:154-64.
40. Wu C, Herold RA, Knoshaug EP, Wang B, Xiong W, Laurens LM. Fluxomic analysis reveals central carbon metabolism adaptation for diazotroph *azotobacter vinelandii* ammonium excretion. *Scientific reports*. 2019;9(1):1-11.

- 590 41. Vitousek PM, Cassman K, Cleveland C, Crews T, Field CB, Grimm NB, et al. Towards
591 an ecological understanding of biological nitrogen fixation. The nitrogen cycle at regional to
592 global scales: Springer; 2002. p. 1-45.
- 593 42. Kilstrup M, Hammer K, Ruhdal Jensen P, Martinussen J. Nucleotide metabolism and its
594 control in lactic acid bacteria. FEMS microbiology reviews. 2005;29(3):555-90.
- 595 43. Beaman TC, Hitchins A, Ochi K, Vasantha N, Endo T, Freese E. Specificity and control
596 of uptake of purines and other compounds in *Bacillus subtilis*. Journal of bacteriology.
597 1983;156(3):1107-17.
- 598 44. Yao C, Chou J, Wang T, Zhao H, Zhang B. Pantothenic acid, vitamin C, and biotin play
599 important roles in the growth of *Lactobacillus helveticus*. Frontiers in microbiology.
600 2018;9:1194.
- 601 45. Peterson W, Peterson MS. Relation of bacteria to vitamins and other growth factors.
602 Bacteriological reviews. 1945;9(2):49-109.
- 603 46. Palacios OA, Bashan Y, de-Bashan LE. Proven and potential involvement of vitamins in
604 interactions of plants with plant growth-promoting bacteria—an overview. Biology and fertility
605 of soils. 2014;50(3):415-32.
- 606 47. Inomura K, Bragg J, Follows MJ. A quantitative analysis of the direct and indirect costs
607 of nitrogen fixation: a model based on *Azotobacter vinelandii*. The ISME journal.
608 2017;11(1):166-75.
- 609 48. Dixon R, Kahn D. Genetic regulation of biological nitrogen fixation. Nature Reviews
610 Microbiology. 2004;2(8):621-31.

49. Gonzalez-Lopez J, Salmeron V, Moreno J, Ramos-Cormenzana A. Amino acids and vitamins produced by *Azotobacter vinelandii* ATCC 12837 in chemically-defined media and dialysed soil media. *Soil Biology and Biochemistry*. 1983;15(6):711-3.
50. Martinez-Toledo MV, Rodelas B, Salmeron V, Pozo C, Gonzalez-Lopez J. Production of pantothenic acid and thiamine by *Azotobacter vinelandii* in a chemically defined medium and a dialysed soil medium. *Biology and fertility of soils*. 1996;22(1):131-5.
51. Danczak RE, Chu RK, Fansler SJ, Goldman AE, Graham EB, Tfaily MM, et al. Using metacommunity ecology to understand environmental metabolomes. *Nature communications*. 2020;11(1):1-16.
52. Danczak RE, Goldman AE, Chu RK, Toyoda JG, Garayburu-Caruso VA, Tolić N, et al. Ecological theory applied to environmental metabolomes reveals compositional divergence despite conserved molecular properties. *Science of The Total Environment*. 2021;788:147409.
53. Wilson RM, Tfaily MM, Kolton M, Johnston ER, Petro C, Zalman CA, et al. Soil metabolome response to whole-ecosystem warming at the Spruce and Peatland Responses under Changing Environments experiment. *Proceedings of the National Academy of Sciences*. 2021;118(25).
54. Song Y, Li X, Yao S, Yang X, Jiang X. Correlations between soil metabolomics and bacterial community structures in the pepper rhizosphere under plastic greenhouse cultivation. *Science of The Total Environment*. 2020;728:138439.
55. Boddy E, Hill PW, Farrar J, Jones DL. Fast turnover of low molecular weight components of the dissolved organic carbon pool of temperate grassland field soils. *Soil Biology and Biochemistry*. 2007;39(4):827-35.

56. Hill PW, Farrar JF, Jones DL. Decoupling of microbial glucose uptake and mineralization in soil. *Soil Biology and Biochemistry*. 2008;40(3):616-24.
57. Pétriacq P, Williams A, Cotton A, McFarlane AE, Rolfe SA, Ton J. Metabolite profiling of non-sterile rhizosphere soil. *The Plant Journal*. 2017;92(1):147-62.
58. Swenson TL, Jenkins S, Bowen BP, Northen TR. Untargeted soil metabolomics methods for analysis of extractable organic matter. *Soil Biology and Biochemistry*. 2015;80:189-98.
59. Dungait JA, Hopkins DW, Gregory AS, Whitmore AP. Soil organic matter turnover is governed by accessibility not recalcitrance. *Global Change Biology*. 2012;18(6):1781-96.
60. Wilkinson A, Hill PW, Farrar JF, Jones DL, Bardgett RD. Rapid microbial uptake and mineralization of amino acids and peptides along a grassland productivity gradient. *Soil Biology and Biochemistry*. 2014;72:75-83.
61. Hobbie JE, Hobbie EA. Amino acid cycling in plankton and soil microbes studied with radioisotopes: measured amino acids in soil do not reflect bioavailability. *Biogeochemistry*. 2012;107(1):339-60.
62. Drescher K, Nadell CD, Stone HA, Wingreen NS, Bassler BL. Solutions to the public goods dilemma in bacterial biofilms. *Current Biology*. 2014;24(1):50-5.
63. Stewart PS. Diffusion in biofilms. *Journal of bacteriology*. 2003;185(5):1485-91.
64. Flemming H-C, Wingender J. The biofilm matrix. *Nature reviews microbiology*. 2010;8(9):623-33.
65. Wang D, Xu A, Elmerich C, Ma LZ. Biofilm formation enables free-living nitrogen-fixing rhizobacteria to fix nitrogen under aerobic conditions. *The ISME journal*. 2017;11(7):1602-13.

66. Rinaudi L, Fujishige NA, Hirsch AM, Banchio E, Zorreguieta A, Giordano W. Effects of nutritional and environmental conditions on *Sinorhizobium meliloti* biofilm formation. *Research in microbiology*. 2006;157(9):867-75.
67. Jung J-H, Choi N-Y, Lee S-Y. Biofilm formation and exopolysaccharide (EPS) production by *Cronobacter sakazakii* depending on environmental conditions. *Food microbiology*. 2013;34(1):70-80.

Figure legends

Fig. 1: Depiction of the systematic scaling of system complexity from liquid and solid culture to synthetic soils (represented here as a microfluidic chip) to *in situ* soil conditions. In our study we focus on the first two steps, relating metabolomics in liquid and solid media culture.

Fig. 2: Microbial (A) C content, (B) N content, (C) C:N ratio, and (D) total biomass. Bars represent average values $\pm$ standard error and are colored by nitrogen treatment. Figures are faceted by culture type. Lowercase letters represent significant difference at $p < 0.05$.

Fig. 3: Extracellular ammonium availability per unit biomass shown by (A) diazotrophic organism and by (B) N treatment. Bars represent average values $\pm$ standard error. Bars in (B) are colored by culture type. Lowercase letters indicate significant difference at $p < 0.05$.

Fig. 4: Percent of fixed nitrogen available as extracellular ammonium in N-free treatments. Bars represent average $\pm$ standard error. No significant difference was observed between culture types.

Fig. 5: Principal coordinates analysis (PCoA) of metabolite chemistry based on (A) Bray-Curtis of peak intensity and (B) Jaccard of presence-absence including all detected metabolites. Each point represents a single sample and are colored by treatment group (organism, N treatment, culture type). 95% confidence ellipses are shown for culture type, represented by color, and N treatment, represented by line type.

Fig. 6: Peak intensities of N-containing metabolites detected across treatment groups. Bars represent average peak intensity $\pm$ standard error. Lowercase letters indicate significant difference at $p < 0.05$.

Fig. 7: Examples of the N-containing metabolites detected at the microscale using MALDI MSI. All ions are annotated as $[M-H]^-$ adducts. Ion images of individual m/z values were generated on the same color bar scale for visual comparison in terms of relative ion abundance.

Table 1: PERMANOVA results for Bray-Curtis and Jaccard distance of macroscale metabolite abundance.

		Df	Sums of Squares	Mean Squares	F model	R squared	p-value
Bray-Curtis	N treatment	1	1.422	1.422	7.899	0.192	0.0001
	Culture type	1	1.548	1.548	8.599	0.209	0.0001
	Interaction	1	0.841	0.842	4.676	0.114	0.0001
	Residual	20	3.601	0.180		0.486	
	Total	23	7.413			1	
Jaccard	N treatment	1	1.350	1.350	11.547	0.279	0.0001
	Culture type	1	0.639	0.639	5.460	0.132	0.0003
	Interaction	1	0.515	0.515	4.406	0.106	0.0016
	Residual	20	2.339	0.117		0.483	
	Total	23	4.843			1	